

Die Last des Wissens

Eine Analyse von Publikations- und Autoredaten

Dissertation
zur Erlangung des Grades eines Doktors der Wirtschaftswissenschaft
der Rechts- und Wirtschaftswissenschaftlichen Fakultät
der Universität Bayreuth

vorgelegt von
Jan Brendel
aus
Danzig

Dekan:	Prof. Dr. Martin Leschke
Erstberichterstatter:	Prof. Dr. Stefan Seifert
Zweitberichterstatter:	Prof. Dr. David Stadelmann
Tag der mündlichen Prüfung:	24.07.2018

Danksagung

An dieser Stelle möchte ich mich bei allen Personen bedanken, die bewusst oder unbewusst zur Erstellung dieser Dissertation beigetragen haben. Zuvorderst gilt der Dank meinem Doktorvater Stefan Seifert, der mir die wissenschaftliche Arbeit an der Universität ermöglicht und mich bereits während der Diplomarbeit an das wissenschaftliche Arbeiten herangeführt hat. Er lehrte mich u. a. das kritische Denken und die kritische Auseinandersetzung mit Wissenschaft, die wissenschaftliche Kommunikation sowie die Präzision von Formulierungen.

Ein großer Dank gilt auch Sascha Schweitzer, der mir stets mit Rat und Tat zur Seite stand. Von ihm habe ich im Rahmen vieler Diskussionen, gemeinsamer Lehrveranstaltungen und Paper-Projekte ebenfalls vieles über das wissenschaftliche Arbeiten und Denken gelernt. Er half mir ferner bei der Vertiefung technischer Fertigkeiten, was letztlich u. a. die Erhebung der Publikationsdaten für diese Arbeit ermöglichte. Er war mir in den letzten Jahren ein guter Kollege und Freund.

Bedanken möchte ich mich auch bei Tim Kessler und Sven Pietrzyk. Mit Tim habe ich meinen ersten wissenschaftlichen Artikel veröffentlicht und bin dabei zum ersten Mal mit den Einreichungs-, Review- und Publikationsprozessen in Berührung gekommen. Mit Sven habe ich mehrere Seminare gemeinsam betreut und dabei vieles zum wissenschaftlichen Denken und Argumentieren, sowie zur Übertragung theoretischer Inhalte in die Praxis mitgenommen.

Für ihre hilfreichen Kommentare und Ideen zum vorliegenden Text bedanke ich mich außerdem bei Stefan Seifert, David Stadelmann, Sascha Schweitzer und Marek, Lukas sowie Leonie Brendel.

Abschließend gilt ein herzlicher Dank auch meinen lieben Eltern und Brüdern sowie Freunden. Sie haben mich stets unterstützt und mich zu dem Menschen erzogen und geformt, der ich heute bin.

Inhaltsverzeichnis

Abbildungsverzeichnis	ix
Tabellenverzeichnis	xv
1 Einleitung	1
1.1 Wissen, Fortschritt und Wachstum	1
1.2 Die Last des Wissens	4
1.3 Motivation, Forschungsfrage und Methodik	6
1.4 Ergebnisübersicht und Aufbau der Arbeit	9
2 Grundlagen und Ergebnisse verwandter Arbeiten	13
2.1 Detailliertere Darstellung ausgewählter Arbeiten	15
2.2 Wachstum, Fortschritt und Produktivität	21
2.3 Koautorenschaften	26
2.4 Quellenangaben und Zitationen	30
2.5 Altersdaten und Produktivität von Individuen	32
2.6 Einordnung und Beitrag der vorliegenden Arbeit	34
3 Daten, Maße und Kontrollvariablen	39
3.1 Der EconLit-Datensatz	40
3.1.1 Herkunft und Umfang	40

3.1.2	JEL-Felder	43
3.2	Der JSTOR-Datensatz	47
3.2.1	Erhebung	47
3.2.2	Umfang	55
3.3	Der <i>Science</i> -Datensatz	62
3.4	Identifikation von Autoren	66
3.4.1	Problematik	66
3.4.2	Methodischer Ansatz	67
3.4.3	Güte des verwendeten Ansatzes	69
3.5	Ergänzung der Publikationsdaten	71
3.5.1	Persönliche und institutionelle Informationen	71
3.5.2	Altersdaten	73
3.6	Untersuchte Maße	76
3.7	Kontrollvariablen bei den Regressionen	80
4	Ergebnisse der Datenauswertung	83
4.1	Teamgröße	84
4.1.1	Teamgröße für alle Artikel	84
4.1.2	Teamgröße bei erstem Artikel	96
4.2	Anzahl der Quellenangaben	102
4.2.1	Anzahl der Quellenangaben für alle Artikel	102
4.2.2	Anzahl der Quellenangaben bei erstem Artikel	115
4.3	Länge von Titeln und Abstracts	122
4.4	Alter bei erstem Artikel	129
4.5	Spezialisierung	135
4.6	Zusammenhang zwischen den betrachteten Maßen und dem Wissensinput	139

4.7	Vergleich der Ergebnisse für JEL-Felder und Disziplinen	144
4.7.1	Vergleich von JEL-Feldern im Econlit-Datensatz	144
4.7.2	Vergleich der untersuchten Disziplinen	147
5	Umgang mit der Last des Wissens	153
5.1	Prinzipien zum Umgang mit der Last des Wissens	155
5.1.1	Das Prinzip der Erhöhung eines Inputfaktors	157
5.1.2	Das Prinzip des parallelen Einsatzes von Inputfaktoren	158
5.1.3	Das Prinzip der Problemreduktion	160
5.1.4	Implikationen	161
5.2	Beispiele für die Umsetzung der drei Prinzipien	161
5.2.1	Erhöhung eines Inputfaktors	161
5.2.2	Paralleler Einsatz von Inputfaktoren	164
5.2.3	Problemreduktion	166
6	Zusammenfassung, Schlussfolgerungen und Ausblick	169
6.1	Zusammenfassung und Diskussion der Ergebnisse	169
6.2	Limitationen der vorliegenden Arbeit	174
6.3	Ausblick und zukünftige Forschung	177
	Literaturverzeichnis	183
	Anhang A Ergänzende Informationen zu den Datensätzen	199
A.1	Ergänzungen zum EconLit-Datensatz	199
A.1.1	Anzahl der Zeitschriften pro Jahr im EconLit-Datensatz	199
A.1.2	Liste der 50 Top-Institutionen	200
A.1.3	Anzahl erfasster Artikel pro Jahr und Zeitschrift	201

A.2	Ergänzungen zum JSTOR-Datensatz	202
A.2.1	Aufbau einer Webseite mit Artikel-Metadaten bei JSTOR . . .	202
A.2.2	Bestimmung von XPath-Ausdrücken	203
A.2.3	Zeitschriften im JSTOR-Datensatz (1878–2014)	204
A.2.4	Anzahl erfasster Artikel	207
Anhang B Ergänzende Auswertungen		209
B.1	Teamgröße	209
B.2	Anzahl der Quellenangaben	211
B.3	Anzahl der Quellenangaben bei erstem Artikel	214
B.4	Länge von Titeln und Abstracts	218
B.5	Alter bei erstem Artikel	219

Abbildungsverzeichnis

1.1	Auswahl zentraler Ergebnisse für <i>Betriebswirtschaftslehre und Finance</i>	9
1.2	Auswahl zentraler Ergebnisse für die Volkswirtschaftslehre	11
2.1	Produktivität von Chemikern in Abhängigkeit von ihrem Alter	16
2.2	Produktivität von Thomas E. Edison in Abhängigkeit von seinem Alter	17
3.1	Anzahl der Artikel pro Jahr im EconLit-Datensatz	41
3.2	Definition von Top-Artikeln und Top-Autoren im EconLit-Datensatz .	42
3.3	Anzahl der Top-Artikel pro Jahr im EconLit-Datensatz	44
3.4	Themenspezifische Suche nach Publikationen bei JSTOR	48
3.5	Themenbereich <i>Business & Economics</i> bei JSTOR	48
3.6	Information zum Umfang der Zeitschrift <i>Econometrica</i> bei JSTOR . .	49
3.7	Anordnung von Ausgaben der Zeitschrift <i>Econometrica</i> bei JSTOR .	49
3.8	Erste Beiträge in der ersten Ausgabe des 83. Bandes von <i>Econometrica</i>	51
3.9	Erste Metadaten zu einem Beitrag bei JSTOR	51
3.10	HTML-Code zu den Metadaten aus Abbildung 3.9	52
3.11	Menü mit weiteren Informationen zu einem Artikel bei JSTOR	53
3.12	Information zur Anzahl der Verweise und Ausschnitt der Liste von Fußnoten mit Quellenangaben für den Artikel aus Abbildung 3.9 . . .	53
3.13	Anfang der Liste mit Quellen aus dem Literaturverzeichnis für den Artikel aus Abbildung 3.9	54

3.14	Anzahl der Artikel pro Jahr im JSTOR-Datensatz (1960–2009)	58
3.15	Beispiel eines XML-Dokuments zum <i>Science</i> -Datensatz	64
3.16	Anzahl der Artikel pro Jahr für die Zeitschrift <i>Science</i>	65
3.17	Anzahl der Artikel pro Jahr ab 1960 für die Zeitschrift <i>Science</i>	65
3.18	Minimaler Graph der Beziehungen zwischen verschiedenen Vornamen	68
3.19	Beispiel eines Namensclusters	70
4.1	Mittlere Teamgröße im EconLit-Datensatz	84
4.2	Mittlere Teamgröße im JSTOR-Datensatz	86
4.3	Mittlere Teamgröße in <i>Science</i>	86
4.4	Anzahl der Top-Artikel im EconLit-Datensatz nach Teamgrößen . . .	87
4.5	Anteile der Teamgrößen für Top-Artikel im EconLit-Datensatz	88
4.6	Anteile der Teamgrößen im EconLit-Datensatz	88
4.7	Anteile der Teamgrößen im JSTOR-Datensatz	90
4.8	Anteile der Teamgrößen in der Zeitschrift <i>Science</i>	90
4.9	Anteile der Artikel mit mindestens 4 Autoren in <i>Science</i>	91
4.10	Mittlere Teamgröße bei erstem Top-Artikel im EconLit-Datensatz . .	98
4.11	Mittlere Teamgröße bei erstem Artikel im EconLit-Datensatz	98
4.12	Mittlere Teamgröße bei erstem Artikel im JSTOR-Datensatz	100
4.13	Mittlere Teamgröße bei erstem Artikel in <i>Science</i>	100
4.14	Mittlere Anzahl Quellenangaben pro Artikel im EconLit-Datensatz . .	103
4.15	Mittlere Anzahl Quellenangaben pro Artikel im JSTOR-Datensatz . .	105
4.16	Mittlere Anzahl Quellenangaben pro Artikel in <i>Science</i>	105
4.17	Mittlere Anzahl Quellenangaben pro Artikel im EconLit-Datensatz für verschiedene Teamgrößen	107

4.18	Mittlere Anzahl Quellenangaben pro Artikel in <i>Science</i> für verschiedene Teamgrößen	108
4.19	Mittlere Anzahl Quellenangaben pro Artikel im EconLit-Datensatz in Abhängigkeit von der Teamgröße	109
4.20	Mittlere Anzahl Quellenangaben pro Artikel in der Zeitschrift <i>Science</i> in Abhängigkeit von der Teamgröße	109
4.21	Mittlere Anzahl Quellenangaben pro Artikel und Autor für Top-Artikel im EconLit-Datensatz für verschiedene Teamgrößen	114
4.22	Mittlere Anzahl Quellenangaben pro Artikel und Autor in der Zeitschrift <i>Science</i> für verschiedene Teamgrößen	114
4.23	Mittlere Anzahl Quellenangaben bei erstem Top-Artikel im EconLit-Datensatz	116
4.24	Mittlere Anzahl Quellenangaben bei erstem Artikel im EconLit-Datensatz (Solo-Artikel)	120
4.25	Mittlere Anzahl Quellenangaben bei erstem Artikel im JSTOR-Datensatz (Solo-Artikel)	121
4.26	Mittlere Anzahl Quellenangaben bei erstem Artikel in <i>Science</i> (Solo-Artikel)	121
4.27	Mittlere Anzahl Wörter im Titel von Artikeln im EconLit-Datensatz	123
4.28	Mittlere Anzahl Wörter im Abstract von Top-Artikeln im EconLit-Datensatz	124
4.29	Mittlere Anzahl Wörter im Titel von Artikeln in <i>Science</i>	125
4.30	Mittlere Anzahl Wörter im Abstract von Artikeln in <i>Science</i>	126
4.31	Mittlere Anzahl Wörter im Titel von Artikeln im JSTOR-Datensatz	127
4.32	Alter bei erstem Artikel in der Volkswirtschaftslehre	130
4.33	Alter bei erstem Top-Artikel in der Volkswirtschaftslehre	131
4.34	Alter bei erstem Artikel in der Mathematik	134
4.35	Relative Häufigkeit eines Feldsprungs im EconLit-Datensatz	136

4.36	Relative Häufigkeit eines Feldsprungs für Top-Autoren im EconLit-Datensatz	137
5.1	Die drei Prinzipien zum Umgang mit komplexen Problemen	156
5.2	Beispiele für die konkrete Umsetzung der drei Prinzipien zum Umgang mit komplexen Problemen	163
A.1	Anzahl der Zeitschriften pro Jahr im EconLit-Datensatz	199
A.2	Mittlere Anzahl Artikel pro Zeitschrift und Jahr im EconLit-Datensatz	201
A.3	Aufbau einer Webseite mit Metainformationen zu einem Beitrag bei JSTOR (Teil 1)	202
A.4	Aufbau einer Webseite mit Metainformationen zu einem Beitrag bei JSTOR (Teil 2)	203
A.5	Bestimmung eines XPath-Ausdrucks	204
A.6	Anzahl erfasster Artikel im JSTOR-Datensatz pro zwei Jahre	207
A.7	Mittlere Anzahl Artikel pro Zeitschrift und Jahr im JSTOR-Datensatz	208
B.1	Anzahl der Artikel im JSTOR-Datensatz nach Teamgrößen	209
B.2	Anzahl der Artikel im EconLit-Datensatz nach Teamgrößen	210
B.3	Anzahl der Artikel in <i>Science</i> nach Teamgrößen	210
B.4	Mittlere Teamgröße in der Zeitschrift <i>Journal of Geology</i> und dem Rest der Disziplin <i>Geographie und Geologie</i>	210
B.5	Mittlere Anzahl Quellenangaben pro Artikel im JSTOR-Datensatz für verschiedene Teamgrößen	211
B.6	Mittlere Anzahl Quellenangaben pro Artikel im JSTOR-Datensatz in Abhängigkeit von der Teamgröße	212
B.7	Mittlere Anzahl Quellenangaben pro Artikel und Autor im EconLit-Datensatz für verschiedene Teamgrößen	212
B.8	Mittlere Anzahl Quellenangaben pro Artikel und Autor im JSTOR-Datensatz für verschiedene Teamgrößen	213

B.9	Mittlere Anzahl Quellenangaben bei erstem Artikel im EconLit-Datensatz (Solo-Artikel)	215
B.10	Mittlere Anzahl Quellenangaben bei erstem Top-Artikel im EconLit-Datensatz (Solo-Artikel)	216
B.11	Mittlere Anzahl Quellenangaben bei erstem Artikel im EconLit-Datensatz (Solo-Artikel)	216
B.12	Mittlere Anzahl Quellenangaben bei erstem Artikel im JSTOR-Datensatz	217
B.13	Mittlere Anzahl Quellenangaben bei erstem Artikel in <i>Science</i>	217
B.14	Anzahl der Wörter im Abstract von Artikeln im JSTOR-Datensatz . . .	218
B.15	Mittleres Alter bei erstem Top-Artikel in der Volkswirtschaftslehre (Altersspanne 25–35)	219
B.16	Mittleres Alter bei erstem Artikel in der Volkswirtschaftslehre (Altersspanne 25–35)	219
B.17	Mittleres Alter bei erstem Artikel in der Mathematik (Altersspanne 25–35)	220

Tabellenverzeichnis

3.1	Top-Zeitschriften im EconLit-Datensatz	43
3.2	JEL-Felder und zugehörige JEL-Codes	45
3.3	Wissenschaftliche Disziplinen im JSTOR-Datensatz (1878–2014)	55
3.4	Wissenschaftliche Disziplinen im JSTOR-Datensatz (1960–2009)	59
3.5	Zeitschriften im JSTOR-Datensatz (1960–2009)	59
3.6	Anzahl identifizierter Personen in den Datensätzen	70
3.7	Anteile ermittelter Geschlechter für die Datensätze	72
3.8	Verteilung von Top-Autoren bei ihren ersten Artikeln auf Top-Zeitschriften im EconLit-Datensatz	74
3.9	Verteilung von Autoren bei ihren ersten Artikeln auf Zeitschriften der <i>Volkswirtschaftslehre</i>	75
4.1	Regressionen: Teamgröße (EconLit-Datensatz)	93
4.2	Regressionen: Teamgröße (JSTOR-Datensatz)	95
4.3	Regressionen: Teamgröße (<i>Science</i>)	95
4.4	Regressionen: Teamgröße bei erstem Artikel (EconLit-Datensatz)	101
4.5	Regressionen: Teamgröße bei erstem Artikel (JSTOR-Datensatz)	102
4.6	Regressionen: Quellenangaben pro Artikel (EconLit-Datensatz)	110
4.7	Regressionen: Quellenangaben pro Artikel (JSTOR-Datensatz)	112
4.8	Regressionen: Quellenangaben pro Artikel (<i>Science</i>)	113

4.9	Regressionen: Quellenangaben bei erstem Artikel (EconLit-Datensatz)	118
4.10	Regressionen: Quellenangaben bei erstem Artikel (JSTOR-Datensatz)	119
4.11	Regressionen: Quellenangaben bei erstem Artikel (JSTOR-Datensatz & <i>Science</i> , Solo-Artikel)	122
4.12	Regressionen: Länge des Titels in Wörtern	128
4.13	Regressionen: Alter bei erstem Artikel in der <i>Volkswirtschaftslehre</i> . .	132
4.14	Regressionen: Alter bei erstem Artikel in der Mathematik	135
4.15	Regressionen: Relative Häufigkeit von Feldsprüngen im EconLit-Datensatz	138
4.16	Regressionen: Teamgröße vs. Wissensinput (EconLit-Datensatz) . . .	141
4.17	Regressionen: Teamgröße vs. Wissensinput (JSTOR-Datensatz & <i>Science</i>)	142
4.18	Regressionen: Titellänge vs. Wissensinput (EconLit-Datensatz)	142
4.19	Regressionen: Titellänge vs. Wissensinput (JSTOR-Datensatz & <i>Science</i>)	143
4.20	Vergleich der JEL-Felder bzgl. der betrachteten Maße	146
4.21	Vergleich der Disziplinen bzgl. der mittleren Teamgrößen	148
4.22	Vergleich der Disziplinen bzgl. der mittleren Anzahl Quellenangaben .	149
4.23	Vergleich der Disziplinen bzgl. der mittleren Anzahl der Wörter im Titel	151
A.1	Zeitschriften im JSTOR-Datensatz (1878–2014)	204
B.1	Regressionen: Quellenangaben bei erstem Artikel im EconLit-Datensatz (Artikel-Ebene)	214
B.2	Regressionen: Quellenangaben bei erstem Artikel im JSTOR-Datensatz (Artikel-Ebene)	215
B.3	Regressionen: Alter bei erstem Artikel in der Mathematik (25–35) . .	220

Kapitel 1

Einleitung

1.1 Wissen, Fortschritt und Wachstum

„If I have seen further it is by standing on ye shoulders of Giants.“
(Isaac Newton, 1676)

Mit dieser Aussage bezieht sich Isaac Newton, einer der wohl bekanntesten und bedeutendsten Wissenschaftler, bereits im 17. Jahrhundert auf die Tatsache, dass seine wissenschaftlichen Erkenntnisse stets auf den Ergebnissen vorhergehender Forschungsarbeiten aufbauten.¹ Für Newton war es somit ein Segen, dass er bestehende grundlegende Erkenntnisse innerhalb seiner Forschungsgebiete (z. B. Mathematik oder Physik) nicht von Neuem entwickeln musste, sondern sich auf die Ergebnisse vorangehender Wissenschaftler berufen konnte.² Auch heutzutage berufen sich Wissenschaftler auf bestehende Forschungsergebnisse und zitieren die entsprechenden Werke in ihren Arbeiten. Die Menge des bestehenden Wissens wächst mit jedem neuen Beitrag bzw. jeder neuen Erkenntnis in einem Fachbereich und mit jedem neuen Artikel oder Buch steigt die Anzahl der Werke, in denen Wissen niedergeschrieben ist.

Die ökonomische Bedeutung von Wissenschaft und ihr Beitrag als Ressource des ökonomischen Wachstums sind unstrittig (Solow 1962; Ayres 1996; Stephan 1996).

¹Das Gleichnis wird 1159 zum ersten Mal mit einem Verweis auf den Gelehrten Bernhard von Chartres von Johannes von Salisbury zitiert (Leuker 1997).

²Aus Gründen der leichteren Lesbarkeit wird in der vorliegenden Arbeit die männliche Sprachform bei personenbezogenen Substantiven und Pronomen verwendet (z. B. Autor oder Wissenschaftler). Sie bezieht sich auf Personen beiderlei Geschlechts und soll im Sinne der sprachlichen Vereinfachung als geschlechtsneutral zu verstehen sein. Sind im Text explizit Frauen gemeint, so wird entsprechend die weibliche Form verwendet.

Diverse endogene Wachstumsmodelle erklären wirtschaftliches Wachstum durch die Akkumulation von Wissen (z. B. Romer 1986; Lucas 1988; Romer 1990). Wissen und Bildung, im Sinne von verarbeitetem und angewandtem Wissen, wirken sich positiv auf technologischen Fortschritt und damit auf die Produktivität sowie das Wirtschaftswachstum aus (Nelson und Phelps 1966; Galor und Tsiddon 1997). Davenport und Prusak (1998) betonen, dass Wissen ein beständiger Vermögensgegenstand eines Unternehmens ist, der dazu eingesetzt werden kann, langfristig Einkommen zu generieren. Daher ist die Schaffung von Wissen ein wichtiges politisches Ziel und seit jeher Gegenstand wissenschaftlicher Untersuchungen. Allerdings sind wissenschaftlicher Fortschritt in Form von neuem Wissen und technologischer Fortschritt zwar notwendige, aber keine hinreichenden Voraussetzungen für wirtschaftliches Wachstum, wie beispielsweise Solow (1962), Mokyr (2002) oder Audretsch und Keilbach (2008) argumentieren.^{3,4}

Blickt man auf die vergangenen Raten technologischen und wissenschaftlichen Fortschritts zurück, so herrscht gemeinhin die Überzeugung, dass die Geschwindigkeit, mit der Fortschritt stattfindet, weiterhin zunehmen wird. Kurzweil zitiert in seinem Buch „The Age of Intelligent Machines“ Péguy (1913) wie folgt: „The world has changed less since Jesus Christ than it has in the last thirty years.“⁵ Tatsächlich scheint dieses Argument heute ebenso valide wie vor 100 Jahren. Blickt man beispielsweise in das Jahr 2000 zurück, so sieht man eine Welt ohne Internet (World Wide Web), Google, Facebook oder Youtube sowie ohne Smartphones und damit ohne technologische Erfindungen, die das heutige gesellschaftliche Leben und soziale Normen stark prägen. Diese historischen Entwicklungen begründeten Vorhersagen von exponentiellem technologischen Wachstum von Unternehmern und Wissenschaftlern wie Gordon E. Moore (1975), der eine Verdopplung der Rechenleistung von Computern alle zwei Jahre vorhersagt (auch als *Moore'sches Gesetz* bekannt) oder Ray Kurzweil (1990), der enorme Fortschritte im Bereich der künstlichen Intelligenz vorhersagt.⁶

Kurzweil (1990) ist von einer exponentiellen technologischen Entwicklung überzeugt. Er prognostiziert eine Beschleunigung des Fortschritts bis zu einem Punkt, an dem Wissen und Fortschritt vorwiegend von künstlicher Intelligenz generiert werden.

³Es wird zwischen technologischem und wissenschaftlichem Fortschritt unterschieden, denn „technology is knowledge, even if not all knowledge is technological“ wie es Mokyr (2002, S. 2) ausdrückt.

⁴Kuznets (1965) verwendet beispielsweise die Ausdrücke *useful knowledge* bzw. *tested knowledge*, um sich auf Wissen zu beziehen, dass in der Produktion bzw. Wirtschaft Anwendung findet und zu deren Wachstum beiträgt.

⁵Das Zitat findet sich in Kurzweils Buch auf Seite 465. Im Original heißt es in Péguy (1913, S. 12 f.): „Le monde a moins changé depuis Jésus-Christ qu'il n'a changé depuis trente ans.“

⁶Rechenleistung wird in diesem Kontext als die Dichte verstanden, mit der Transistoren in Computer-Chips verbaut werden.

Kurzweil illustriert exponentielles Wachstum anhand einer altertümlichen Geschichte über den Erfinder des Schachspiels:⁷ Dieser soll für die Erfindung des Spiels belohnt werden und der König des Landes entscheidet, dass er seine Belohnung selbst wählen darf. Der Mann bittet darum, in Reiskörnern entlohnt zu werden. Ein Reiskorn auf dem ersten Feld eines Schachbretts, zwei auf dem zweiten, vier auf dem dritten und so weiter. Der König gibt sich mit dem Vorschlag einverstanden, weil er glaubt, dass diese Belohnung gering sei. Kurzweil vergleicht die wenigen Reiskörner auf den ersten Feldern mit einzelnen technologischen Innovationen, die technischen Fortschritt nicht enorm beeinflussen, wie auch die wenigen ersten Reiskörner den König nicht störten. Bewegt man sich allerdings weiter entlang des Schachbretts, steigt die Anzahl der Reiskörner derart an, dass sie die kumulierte Menge, jemals von Menschen produzierten Reises übersteigt (Gerbert, Gauger und Steinhäuser 2015).

Brynjolfsson und McAfee (2011) folgen der Argumentation von Kurzweil, indem sie die obige Geschichte und das Mooresche Gesetz bemühen. Sie argumentieren, dass Verbesserungen in einem bestimmten Feld häufig mit wenigen kleineren Verbesserungen bzw. Innovationen beginnen. Wenn allerdings die Anzahl der Innovationen zunehmend schnell steigt, wie der Reisberg aus obiger Geschichte, sind die resultierenden Folgen für den technologischen Fortschritt enorm. Von ähnlich positiven Wachstumsraten, wie denen im technologischen Bereich, wird auch bezüglich des wissenschaftlichen Fortschritts berichtet. Beispielsweise berichtet Price (1963, S. 20 f.) von einer Verdopplung wissenschaftlichen Wissens alle 10–15 Jahre. Seit seiner grundlegenden Arbeit bestätigen diverse bibliometrische Studien den Fortgang dieses Wachstums in den vergangenen Jahrzehnten (Tague, Beheshti und Rees-Potter 1981; Archibald und Line 1991; Tabah 1999; Mabe und Amin 2001; Fernández-Cano, Torralbo und Vallejo 2004; Bornmann und Mutz 2015).

Der optimistischen Sichtweise von Moore (1975), Kurzweil (1990) sowie Brynjolfsson und McAfee (2011) steht allerdings eine zweite, pessimistischere Schule gegenüber. Im Jahr 2013 schrieb *The Economist*, dass verschiedene Indikatoren wie wirtschaftliches Wachstum, Reallöhne in der Produktion und Lebenserwartungen in den letzten Jahren stagnierten.⁸ Brynjolfsson (1993) diskutiert das sogenannte *productivity paradoxon*, das sich darin äußert, dass sich die wachsende Verbreitung von Informationstechnologien nicht in den Produktivitätszahlen widerspiegelt. Robert Solow beschreibt dieses Phänomen mit den Worten: „we see computers everywhere but in the productivity statistics“

⁷Laut Wikipedia hieß der vermeintliche Erfinder des Schachspiels Sissa ibn Dahir und lebte im dritten oder vierten Jahrhundert n. Chr. in Indien.

⁸Der Autor des Artikels ist nicht bekannt.

(Brynjolfsson 1993, S. 67). Gordon (1999, 2012) berichtet von abnehmenden Wachstumsraten der US-Ökonomie und Bloom u. a. (2017) stellen verschiedene Fallstudien aus betrieblichen, technologischen sowie wissenschaftlichen Kontexten vor und zeigen, dass die Produktivität trotz vermehrt konstanter Wachstumsraten abnimmt.

Hinsichtlich des Wachstums wissenschaftlicher Erkenntnisse und Publikationen merkt Price (1963) an, dass dieses wesentlich durch eine wachsende Menge an Wissenschaftlern getrieben ist. Somit lenkt das absolute wissenschaftliche Wachstum leicht von seinem zentralen Treiber, nämlich der Anzahl der Wissenschaftler ab. Studien, die sich auf den Output pro Wissenschaftler fokussieren, kommen zu etwas anderen Ergebnissen als die oben genannten Studien. Beispielsweise berichtet Evenson (1984, 1993) von abnehmendem Output (in Form von Patenten) pro Erfinder bezogen auf die Ausgaben für bzw. Investitionen in Forschung und Entwicklung. Schankerman und Pakes (1986, S. 1072, Tabelle 6) berichten von abnehmenden Patentanwendungen und Werten von Patenten pro Erfinder. Conley, Crucini u. a. (2013) untersuchen die Produktivitätsrate (im Sinne wissenschaftlicher Veröffentlichungen) junger Wissenschaftler, innerhalb eines Zeitraums von sechs Jahren nach ihrer Promotion. Aus ihren Ergebnissen geht hervor, dass die durchschnittliche Produktivitätsrate dieser jungen Wissenschaftler im Zeitraum zwischen 1987 und 1999 abgenommen hat.

1.2 Die Last des Wissens

Wie bereits angesprochen hat sich das allgemein zugängliche Wissen im Laufe der Geschichte vermehrt. Diese zunehmende Wissensmenge kann jedoch auch als Last empfunden werden, da Wissensarbeiter (z. B. Wissenschaftler oder Erfinder) einen erheblichen Teil ihres Lebens in Schulen, an Universitäten, in Doktorandenprogrammen und in Forschungspositionen damit verbringen, zunehmend viel bestehendes Wissen zu verarbeiten und sich anzueignen. Dieses Wissen ermöglicht ihnen gegebenenfalls selbst zum Wissen in ihrem eigenen Feld beizutragen. Paradoxe Weise trägt aber jeder dieser neuen Beiträge wiederum zur wachsenden Menge an Wissen bei, die kommende Generationen von Wissenschaftlern verarbeiten müssen, bevor sie selbst neue Erkenntnisse gewinnen können.

Mit diesem Phänomen beschäftigt sich Jones (2009a) in seinem Artikel „The Burden of Knowledge and the ‘Death of the Renaissance Man’: Is Innovation Getting Harder?“ und bezeichnet den Mechanismus als eine *Last des Wissens*. Diese manifestiert sich

dadurch, dass neues Wissen stets auf vorhandenem Wissen aufbaut. Je mehr Wissen eine Generation von Wissenschaftlern erzeugt, desto mehr Wissen müssen folgende Generationen verarbeiten, um an die bestehenden Grenzen des Wissens in ihrem Fachbereich vorzustoßen und selbst zu diesem Wissenspool beizutragen. Einmal an diesem Punkt angekommen, müssen Wissenschaftler ferner zunehmend viele Arbeiten lesen, um in ihrem Fachbereich auf dem neuesten Stand zu bleiben. Jones (2009a) spekuliert, dass sich wissenschaftlicher Fortschritt in Zukunft verlangsamen könnte und dass Wissensarbeiter verschiedene Strategien anwenden könnten, um der steigenden Menge an Wissen zu begegnen. Hierzu gehören beispielsweise längere Ausbildungszeiten, eine zunehmende Spezialisierung auf kleinere Teilgebiete ihres Fachbereichs oder die Arbeit in Teams, das heißt die Kooperation mit anderen Wissenschaftlern.

Tatsächlich gibt es diverse Anzeichen, die für eine solche Last des Wissens in der Wissenschaft sprechen. Beispielsweise berichtet Ellison (2002), dass die Anzahl der Publikationen in den renommiertesten Zeitschriften der Volkswirtschaftslehre langsamer wächst und dass die Begutachtungsprozesse für Einreichungen zunehmend aufwändig sind. Für einzelne naturwissenschaftliche Forschungsbereiche berichtet Arbesman (2011), dass es zunehmend schwerer ist, neue Erkenntnisse zu erbringen. Für Physik und Geowissenschaften kommen Levin und Stephan (1991) zu dem Ergebnis, dass die Produktivität von Wissenschaftlern seit den 1960er Jahren abgenommen hat. Gonzalez-Brambila und Veloso (2007) replizieren diese Ergebnisse für Biologen, Ingenieure und Sozialwissenschaftler in Mexiko. Für die Volkswirtschaftslehre berichtet Hamermesh (2013) von einem Anstieg des mittleren Alters publizierender Wissenschaftler und Conley, Crucini u. a. (2013) berichten von abnehmender pro-Kopf-Produktivität junger Wissenschaftler.

Jones (2009a) untersucht verschiedene Indikatoren für eine Last des Wissens anhand von US-Patentdaten. Zu diesen Indikatoren gehören das Alter eines Erfinders bei seinem ersten Patent, die Menge des zugrunde liegenden Wissens (bemessen anhand der Anzahl zitierter Patente), die Teamgröße (d. h. die Anzahl der Patentanmelder für ein Patent) und die Spezialisierung. Die Spezialisierung bemisst Jones anhand der Wahrscheinlichkeit, dass ein Patentanmelder das technologische Feld zwischen zwei aufeinander folgenden Patenten wechselt. Er zeigt, dass das Alter der Patentanmelder bei ihrer ersten Patentanmeldung, die Anzahl zitierter Patente und die Teamgröße steigen und dass die Wahrscheinlichkeit, dass Erfinder ihr technologisches Feld zwischen zwei aufeinander folgenden Patentanmeldungen wechseln abnimmt.

1.3 Motivation, Forschungsfrage und Methodik

Der Ansatz von Jones (2009a), die Last des Wissens in der industriellen Forschung und Entwicklung zu quantifizieren, kann prinzipiell auf die wissenschaftliche Forschung übertragen werden. Sowohl Patente, als auch wissenschaftliche Artikel weisen Meta-informationen bzgl. der Autoren und der Quellenangaben auf. Ferner weisen beide ähnliche Erfolgs-Entlohnungen auf (Stephan 1996) und lassen sich als ein „winner-takes-it-all“ Wettbewerb interpretieren. Während ein Patent die Nutzungsrechte an einer Erfindung schützt, die zum ersten Mal von jemandem angemeldet wird, sichert das Publizieren eines wissenschaftlichen Artikels dem Autor bzw. den Autoren den Vorrang in der Reihenfolge für diese Entdeckung zu. Merton (1957) fasst das Publizieren einer wissenschaftlichen Erkenntnis in einem Artikel als eine Art Eigentumsrecht auf, dass die Reputation und den Erfolg einer Karriere erhöht. Stephan (1996) zeigt, dass die Bereiche der industriellen und akademischen Forschung nicht disjunkt sein müssen und Wissenschaftler bzw. Erfinder teilweise gleichzeitig in beiden tätig sind.

Primäres Ziel der vorliegenden Arbeit ist es, für die Wissenschaft empirisch zu überprüfen, ob es Indikatoren für eine Last des Wissens, wie sie Jones (2009a) beschreibt, gibt. Ferner sollen die allgemeinen Prinzipien hinter möglichen Ansätzen zum Umgang mit einer solchen Last des Wissens herausgearbeitet werden, die ggf. auch die beobachteten empirischen Trends erklären.

Um den empirischen Teil der Forschungsfrage zu beantworten, werden Publikationsdaten aus den Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* sowie *Psychologie* über einen Zeitraum von jeweils 50 Jahren (1960–2009) und *Volkswirtschaftslehre* (1970–2014) sowie Artikel der Zeitschrift *Science* (1960–2011), in der überwiegend naturwissenschaftliche Forschungsergebnisse publiziert werden, ausgewertet. Zweck der Betrachtung von *Science* ist es ein Schlaglicht auf eine besonders renommierte Zeitschrift zu werfen. Als einzelne Zeitschrift ist sie jedoch nicht repräsentativ für die gesamte Wissenschaft. Eine Methode der Datensammlung, die im Kontext dieser Studie relevant ist, ist die Erhebung zusätzlicher Informationen aus öffentlich verfügbaren Quellen. Zum Beispiel recherchierte Jones (2009a) Geburtsdaten von Erfindern, Card und DellaVigna (2013) ermittelten Zitationen wissenschaftlicher Artikel bei *Google Scholar* und Hayne, Smith und Vijayasaraty (2003) sammelten Auktionsdaten von eBay. In der vorliegenden Arbeit betrifft dies zusätzliche institutionelle (z. B. Zugehörigkeit zu einer akademischen Institution) und persönliche Informationen (z. B. Geburtsjahr) der Autoren. Für die Auswertung werden

bibliometrische Methoden nach den Beispielen von Yitzhaki und Ben-Tamar (1991), Ucar u. a. (2013) sowie Bornmann und Mutz (2015) für die Anzahl an Quellenangaben in Artikeln oder McDowell und Melvin (1983), Hudson (1996) sowie Card und DellaVigna (2013) für die Teamgröße verwendet. Die Daten werden einerseits mittels deskriptiver Statistik (z. B. Abbildungen und Tabellen) veranschaulicht. Weitere statistische Auswertungen umfassen Regressionsanalysen, in denen für diverse Faktoren kontrolliert wird. Im konzeptionellen Teil der Arbeit wird die Perspektive der *Design Science* eingenommen (vgl. Vaishnavi und Kuechler 2008; Hevner und Chatterjee 2010). Den Schritten von Peffers u. a. (2008) zur Entwicklung eines Rahmenwerks folgend, werden dabei zunächst Ansätze zum Umgang mit komplexen Problemen identifiziert und systematisiert. Anschließend wird ein Rahmenwerk herausgearbeitet, das die grundlegenden Prinzipien hinter den unterschiedlichen Ansätzen umfasst. Schließlich wird das Rahmenwerk anhand von Beispielen veranschaulicht und diskutiert.

Ein zentraler Teil der Untersuchung beschäftigt sich mit Publikationen aus der Volkswirtschaftslehre. Dieser Fachbereich diente bereits vielfach als Beispiel, um wissenschaftliche Produktivität zu untersuchen (z. B. McDowell und Melvin 1983; Oster und Hamermesh 1998; Conley und Önder 2014). Im Rahmen dieser Arbeit hat der Fachbereich der Volkswirtschaftslehre zwei gewichtige Vorteile. Einerseits erlaubt die Publikationskultur in der Volkswirtschaftslehre, wissenschaftliche Erkenntnisse individuellen Forschern zuzurechnen. Forschung in den Naturwissenschaften wird beispielsweise zunehmend von einer Vielzahl beteiligter Personen sowie großen Forschungseinrichtungen bzw. Laboren durchgeführt (Cronin 2001; Simonton 2013; Franzoni und Sauerermann 2014; Castelvechi 2015), wobei auch nicht-wissenschaftliches Personal als Autoren bei den Publikationen gelistet wird oder gar das gesamte Forschungsinstitut ohne die Nennung einzelner Personen. In der vorliegenden Arbeit liegt der Fokus allerdings auf der Untersuchung von Effekten der Last des Wissens auf individuelle Forscher und nicht auf ganze Forschungseinrichtungen. Andererseits hat sich für die Volkswirtschaftslehre ein Klassifizierungssystem etabliert, das von der Zeitschrift *Journal of Economic Literature* (JEL) vorgeschlagen wurde (vgl. hierzu auch Abschnitt 3.1.2, S. 43). Dieses ermöglicht es, spezifischer zu untersuchen, in welchen Teilfeldern ihrer Disziplin Autoren tätig waren und beigetragen haben.

Die folgenden Indikatoren dienen als Proxys für die Last des Wissens und die Strategien, wie Wissenschaftler mit dieser umgehen: Die Teamgröße, also die Anzahl der Autoren, die an einem Artikel gearbeitet haben, dient als Maß dafür, wie viele (spezialisierte) Wissenschaftler dazu beigetragen haben, die notwendige Breite an Wissen abzudecken, um einen wissenschaftlichen Artikel zu verfassen. Die Anzahl

der angegebenen Quellen in Artikeln bemisst die Menge an bestehendem Wissen, das in dem Artikel verarbeitet wurde. Das Alter von Autoren zum Zeitpunkt ihrer ersten Veröffentlichung deutet darauf hin, wie lange die Wissenschaftler brauchen, um das Grundwissen auf ihrem Gebiet zu erlangen, zu dem sie gegebenenfalls beitragen werden. Zuletzt beschreibt die Spezialisierung von Autoren, wie viele verschiedene Felder ihres Fachbereichs Wissenschaftler abdecken. Sie wird anhand der relativen Häufigkeit bemessen, dass Autoren ihre ersten beiden Artikel, bei denen sie alleiniger Autor sind, zu unterschiedlichen Teilgebieten ihres Fachbereichs (Felder) verfassen (auch als *Feldsprung* bezeichnet).

Vor dem Hintergrund einer Last des Wissens ist die Hypothese bzgl. des Alters von Autoren zum Zeitpunkt ihres ersten Artikels, dass dieses im Laufe der Zeit steigt, da Wissenschaftler mehr Zeit aufwenden müssen, um die zunehmende Menge des bestehenden Wissens zu verarbeiten und auf diese Weise das nötige Wissen sowie die nötigen Fertigkeiten zu entwickeln, um ihren ersten größeren wissenschaftlichen Beitrag zu publizieren. Die Anzahl der angegebenen Quellen, die Wissenschaftler in ihrer ersten Publikation ausweisen, ist eine mögliche Manifestation dieser Prozesse der Wissensverarbeitung. Je mehr externe Informationen ein Autor sammeln und verarbeiten musste, um zu seinen Ergebnissen zu gelangen, desto mehr Quellen wird er für seine Arbeit angeben. Entsprechend sollte die Anzahl der Quellenangaben, die in Artikeln (speziell ersten Artikeln von Autoren) angegeben werden, steigen. Ferner wird angenommen, dass sich Wissenschaftler zunehmend spezialisieren und die ersten beiden Artikel eines Autors daher thematisch zunehmend nahe beieinander liegen. Wenn die Menge des zu verarbeitenden Wissens steigt und Wissenschaftler sich zunehmend spezialisieren, dann sollte auch die Anzahl der Autoren pro Artikel steigen, da das notwendige Wissen dann von mehreren Personen abgedeckt werden kann.

Die Ergebnisse der vorliegenden Arbeit gehen teilweise auf gemeinsame Arbeiten mit Sascha Schweitzer zurück. So werden Ergebnisse für die Volkswirtschaftslehre zum Alter, der Teamgröße sowie der Anzahl Quellenangaben beim ersten Artikel eines Autors und zur Spezialisierung von Autoren in einem gemeinsamen Arbeitspapier berichtet.⁹ Entsprechend finden sich in der vorliegenden Arbeit in verschiedenen Kapiteln auch einzelne Textpassagen, die aus dieser Arbeit übernommen sind und lediglich aus dem Englischen übersetzt wurden (z. B. in Kapitel 2 bei der Beschreibung von Ergebnissen verwandter Arbeiten oder in Kapitel 4 bei der Vorstellung von Ergebnissen zur Spezialisierung von Autoren sowie dem Vergleich der Ergebnisse für verschiedene

⁹Eine aktuelle Version des Arbeitspapiers findet sich unter https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2895685 (Stand: 09.06.2018).

Felder der Volkswirtschaftslehre). Auch das konzeptionelle Rahmenwerk zur Einordnung von Ansätzen zum Umgang mit der Last des Wissens (vgl. Kapitel 5) entstammt gemeinsamen Ideen und Diskussionen mit Sascha Schweitzer und ersten, (noch) nicht veröffentlichten Aufzeichnungen.

1.4 Ergebnisübersicht und Aufbau der Arbeit

Die empirischen Ergebnisse der vorliegenden Arbeit zur mittleren Teamgröße und der mittleren Anzahl der Quellenangaben pro Artikel (vgl. Abschnitt 4.1.1 bzw. Abschnitt 4.2.1) zeigen für alle untersuchten Disziplinen in den jeweiligen Beobachtungszeiträumen deutlich steigende Trends und bestätigen damit die Ergebnisse bestehender Literatur hinsichtlich dieser Maße. In Abbildung 1.1 sind diese Trends beispielhaft für Artikel der *Betriebswirtschaftslehre und Finance* abgebildet. Für diese steigt die mittlere Teamgröße zwischen 1960 und 2009 von 1,2 auf 2,3 Autoren (vgl. Abbildung 1.1a). Im gleichen Zeitraum vervierfacht sich die mittlere Anzahl der Quellenangaben von 10,9 auf 42,5 (vgl. Abbildung 1.1b). Die Arbeit erweitert die Ergebnisse bestehender Literatur und zeigt, dass die beschriebenen Trends für die Teamgröße und die Anzahl der Quellenangaben auch für die jeweils ersten Artikel der Autoren gelten (vgl. Abschnitt 4.1.2 bzw. Abschnitt 4.2.2). Die mittlere Teamgröße bei diesen Artikeln ist ferner größer als für alle Artikel in den jeweiligen Disziplinen.

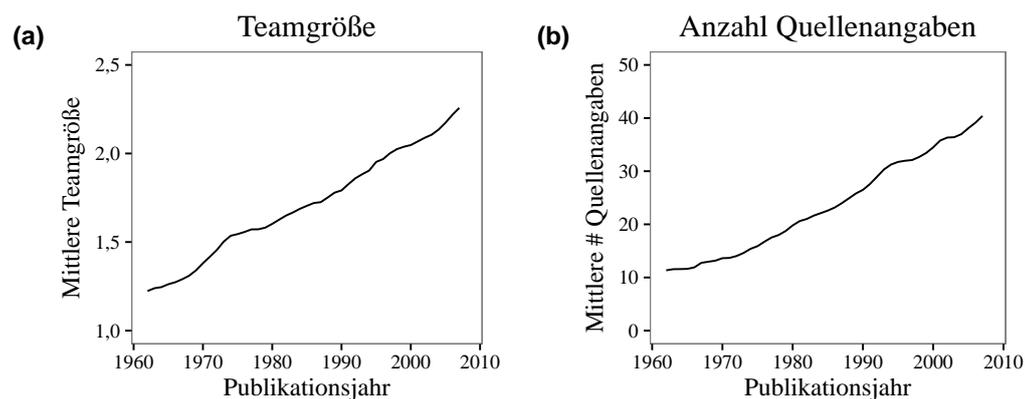


Abbildung 1.1: Mittlere Teamgröße und mittlere Anzahl Quellenangaben für *Betriebswirtschaftslehre und Finance*. Abgebildet sind gleitende 5-Jahres-Mittelwerte.

Die größten Teams und die größte absolute und relative Zunahme der Teamgröße sind für Artikel der Zeitschrift *Science* zu beobachten. Hier ist die mittlere Teamgröße

von 1,9 im Jahr 1960 auf 5,2 im Jahr 2011 angestiegen. Ferner werden Artikel in der Zeitschrift *Science* in den letzten Beobachtungsjahren häufig von sehr großen Teams verfasst, sodass 13,7 % der Artikel im Jahr 2011 von zehn oder mehr Autoren verfasst wurden. Die *Geographie und Geologie* weist die geringsten Teamgrößen sowie den geringsten absoluten und relativen Anstieg der Teamgröße auf. Die mittlere Teamgröße steigt von 1,1 im Jahr 1960 auf 1,4 im Jahr 2009. Für alle Disziplinen ist der Anteil von Artikeln, die lediglich von einem Autor verfasst werden, zurückgegangen. Für *Betriebswirtschaftslehre und Finance* ist dieser Rückgang am stärksten und der Anteil dieser Artikel fällt von 86,3 % im Jahr 1960 auf 16,5 % im Jahr 2009. Für *Geographie und Geologie* fällt dieser Rückgang zwar unter den untersuchten Disziplinen am geringsten aus, dennoch ist auch hier ein Rückgang im entsprechenden Zeitraum von 85,1 % auf 74,3 % zu beobachten.

Wie oben erwähnt, sind auch für die mittlere Anzahl der Quellenangaben für alle untersuchten Disziplinen deutlich steigende Trends zu beobachten. Den stärksten absoluten und relativen Anstieg weist die *Psychologie* auf, in der die mittlere Anzahl der Quellenangaben zwischen 1960 und 2009 um den Faktor 5 von 12,1 auf 60,1 gestiegen ist. Auch für dieses Maß weisen Artikel der *Geographie und Geologie* den geringsten Zuwachs auf. Dennoch hat sich der Mittelwert etwa verdoppelt und ist von 19,8 im Jahr 1960 auf 42,8 im Jahr 2009 gestiegen. Für alle Disziplinen erweist sich der geschätzte Effekt der Teamgröße auf die Anzahl der Quellenangaben als positiv. Je mehr Autoren also an der Verfassung eines Artikels beteiligt sind, desto mehr Quellenangaben weist der Artikel auf. Dieser Effekt ist für alle Disziplinen statistisch hoch signifikant.

Ein zentraler Beitrag dieser Arbeit, der sich in analoger Form in der Literatur lediglich für Patentanmelder findet, bezieht sich auf das Alter von Autoren bei ihrem ersten Artikel (vgl. Abschnitt 4.4). Für die Disziplinen *Volkswirtschaftslehre* und *Mathematik* ist über die Zeit ein Anstieg des mittleren Alters bei Erstpublikation zu beobachten. Abbildung 1.2 (a) veranschaulicht diesen Trend für die Volkswirtschaftslehre. Das mittlere Alter von Autoren bei ihrer ersten Publikation in einer volkswirtschaftlichen Zeitschrift ist von 32,5 Jahren im Jahr 1970 auf 37,2 Jahre im Jahr 2014 gestiegen. Ein weiterer Beitrag der vorliegenden Arbeit, der sich in der Literatur bislang ebenfalls nur für Patentanmelder findet, ist die Untersuchung der Spezialisierung von Autoren der Volkswirtschaftslehre (vgl. Abschnitt 4.5). Diese wird mittels der relativen Häufigkeit bemessen, dass ein Wissenschaftler seine ersten beiden Publikationen, bei denen er alleiniger Autor ist, zu unterschiedlichen Teilbereichen (Feldern) der Volkswirtschaftslehre verfasst. Die relative Häufigkeit solcher Feldsprünge ist von 33,2 % im Jahr 1973 auf 14,4 % im Jahr 2014 gefallen (Abbildung 1.2b). Dies impliziert, dass Wissenschaftler zu-

mindest zu Beginn ihrer Karriere seltener zwischen den Teilgebieten ihres Fachbereichs wechseln und sich in diesem Sinne im Laufe der Zeit zunehmend spezialisieren.

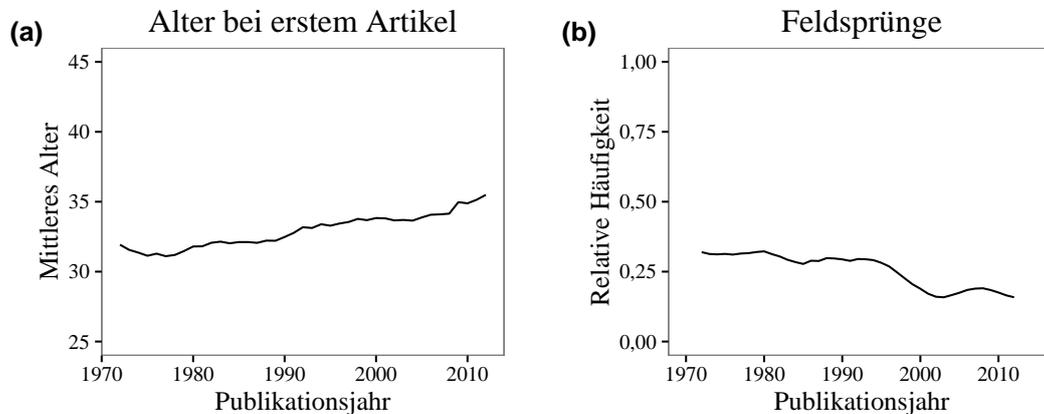


Abbildung 1.2: Mittleres Alter beim ersten Artikel eines Autors und relative Häufigkeit von Feldsprüngen zwischen den ersten beiden Publikationen eines Autors in der Volkswirtschaftslehre. Abgebildet sind gleitende 5-Jahres-Mittelwerte.

Neben den empirischen Ergebnissen wird ein konzeptionelles Rahmenwerk herausgearbeitet, das eine Einordnung von Methoden und Techniken zur Lösung großer und komplexer Probleme ermöglicht (vgl. Kapitel 5). Dieses Rahmenwerk umfasst drei grundsätzliche Prinzipien. Das Prinzip der *Erhöhung eines einzelnen Inputfaktors* umfasst Ansätze, die darauf abzielen beispielsweise die Arbeitszeit eines einzelnen Wissensarbeiters zu erhöhen oder dessen Effizienz durch gezieltes Training zu steigern (z. B. Speed Reading). Das Prinzip des *parallelen Einsatzes von Inputfaktoren* zielt darauf ab ein gegebenes Problem in möglichst disjunkte Teilprobleme aufzuteilen und diese gleichzeitig unter dem Einsatz mehrerer Inputfaktoren zu lösen. Ein Beispiel für dieses Prinzip in der Wissenschaft ist Kooperation bzw. Teamarbeit, bei der beispielsweise die Arbeit am Verfassen eines Artikels unter mehreren Autoren aufgeteilt wird. Schlussendlich gehört das Prinzip der *Problemreduktion* zum Rahmenwerk. Ansätze, die dieses Prinzip verfolgen, versuchen das gegebene Problem zu transformieren und ein ggf. einfacheres und/oder abstraktes, verwandtes Problem zu lösen, dessen Lösung die des ursprünglichen Problems ausreichend gut annähert. Ein Beispiel aus dem wissenschaftlichen Kontext ist die Modellierung von (realen) Phänomenen unter festen, unter Umständen vereinfachenden Annahmen. Im Kontext der Wissensverarbeitung sowie -schaffung und damit im Umgang mit einer möglichen Last des Wissens, erklären diese Prinzipien einige empirische Trends, wie die Zunahme von Teamarbeit oder den gestiegenen Ressourceneinsatz in Form von Geld und Zeit.

Der Rest der Arbeit ist wie folgt aufgebaut: In Kapitel 2 wird zunächst auf verwandte Literatur eingegangen und die vorliegende Arbeit sowie deren Ergebnisse entsprechend eingeordnet. Kapitel 3 umfasst eine Vorstellung der dieser Arbeit zugrunde liegenden Daten und ihres Umfangs sowie Erläuterungen dazu, wie diese Daten erhoben wurden. Ferner werden in diesem Kapitel die untersuchten Maße vorgestellt und diskutiert. In Kapitel 4 finden sich die empirischen Ergebnisse. Das Kapitel schließt in Abschnitt 4.7.1 mit einer Zusammenfassung und einem Vergleich der Ergebnisse sowohl für die einzelnen Felder in der Volkswirtschaftslehre als auch für die in dieser Arbeit untersuchten Disziplinen. Im darauffolgenden Kapitel 5 wird eine konzeptionelle Einordnung von Methoden zum Umgang mit der Last des Wissens vorgestellt, bevor die Arbeit mit einer Zusammenfassung und Diskussion der Ergebnisse sowie einem Fazit und Ausblick in Kapitel 6 schließt.

Kapitel 2

Grundlagen und Ergebnisse verwandter Arbeiten

In diesem Kapitel werden die Grundlagen der vorliegenden Arbeit vorgestellt, die Arbeit in die bestehende Literatur eingeordnet und entsprechend verwandte Arbeiten und deren Ergebnisse vorgestellt. Dabei werden zunächst zentrale Begriffe wie „Wissen“ oder „Wissenschaft“ erläutert, bevor in Abschnitt 2.1 die besonders eng verwandten und relevanten Arbeiten auch methodisch etwas ausführlicher dargestellt werden, um die Verwandtschaft noch besser herauszustellen und den engeren Rahmen der vorliegenden Arbeit aufzuzeigen. Die Arbeit ist nah verwandt mit zwei Strängen wissenschaftlicher Literatur. Zum einen beschäftigt sich seit Lehmans (1953) grundlegender Arbeit „Age and Achievement“ ein beachtlicher Wissenschaftszweig mit dem Alter und anderen Eigenschaften von Erfindern und Wissenschaftlern. Im Zentrum der vorliegenden Arbeit stehen ebenfalls die individuellen Wissenschaftler. Untersucht werden Wissen und Innovation, die zu den zentralen Treibern allgemeinen Fortschritts und damit auch wirtschaftlichen Wachstums gehören (z. B. Nelson und Phelps 1966; Galor und Tsiddon 1997). Damit ist das Thema zum anderen mit theoretischer und praktischer Forschung verwandt, die sich mit den Determinanten volkswirtschaftlichen Wachstums beschäftigt (z. B. Romer 1986; Lucas 1988; Aghion und Howitt 1992; Segerstrom 1998).

Gemeinhin wird zwischen den Begriffen „Daten“, „Informationen“ und „Wissen“ nicht ausreichend unterschieden. Die Begriffe stehen allerdings in einer Begriffshierarchie zueinander, die auch als *Wissenspyramide* bekannt ist (Ackoff 1989; Miller 1996; Bates 2005; Bodendorf 2006). Unter einem *Zeichen* versteht man beispielsweise einzelne Buchstaben oder Ziffern. Einzelne Zeichen, wie „a“, „b“, „c“, „=“, „+“ oder „2“

die in einer bestimmten Reihenfolge und Anordnung (z. B. Hochstellung von Zeichen), der sogenannten *Syntax*, zusammengesetzt werden, ergeben ein *Datum*. Beispielsweise ergeben sich aus obigen Zeichen die Daten „ a^2 “ oder „ b^2 “. Daten wiederum werden durch einen *Kontext* geordnet und so zu einer für eine Person interpretierbaren *Information*. Aus den obigen Zeichen und Daten lässt sich beispielsweise die Formel „ $a^2 + b^2 = c^2$ “ aufstellen. Durch die Vernetzung von Informationen und die „Kenntnisse darüber, in welchem Zusammenhang die Informationen zueinander stehen“ (Bodendorf 2006, S. 1) entsteht *Wissen*, als höchste Stufe der Wissenspyramide. Wissen kann somit auch als gelernte Information aufgefasst werden (Köhler und Oswald 2009, S. 15). Die Information „ $a^2 + b^2 = c^2$ “ beschreibt, in den richtigen Zusammenhang gestellt, entsprechend dem Satz von Pythagoras, Wissen über das Verhältnis der Seitenlängen in einem rechtwinkligen Dreieck. Auch für Informationen bedarf es also eines bestimmten Kontextes und Erfahrung, um diese miteinander zu brauchbarem Wissen zu verbinden (Bodendorf 2006; Köhler und Oswald 2009; North, Brandner und Steininger 2016).

Wissen wird seit Polanyi (1969) in implizites und explizites Wissen unterteilt. Das Aufkommen dieser Unterscheidung des Wissensbegriffs führte zu einer Kritik vieler ökonomischer Wachstumsmodelle, die Wissen als einfachen Inputfaktor in einer Produktionsfunktion betrachten und ein Mehr an Wissen einem Zuwachs an technischem Fortschritt gleichsetzen. Nelson und Winter (1982, S. 73) unterstreichen die Bedeutung des bis dato kaum beachteten impliziten Wissens für den wirtschaftswissenschaftlichen Entwicklungsprozess: „... the knowledge that underlies a skillfull performance is in large measure tacit knowledge ...“. Sie gehen dabei ebenso wie Polanyi (1969) davon aus, dass dieses nicht artikulierbar ist. Dies und die mangelnde Formalisierbarkeit führen ferner dazu, dass es nicht zwischen Individuen transferierbar ist (Grant 1996). Osterloh und Frey (2000, S. 539) beschreiben die Transferierbarkeit der Wissensarten folgendermaßen: „Tacit knowledge is acquired by and stored within individuals and cannot be transferred or traded as a separate entity. Explicit knowledge has the character of a public good [...]“. Die vorliegende Arbeit beschäftigt sich vorwiegend mit dem expliziten Wissen, das in Form von Patenten, Büchern oder wissenschaftlichen Artikeln festgehalten ist (Riviera 2013).

„Wissenschaft“ bezeichnet grob gesagt eine Tätigkeit, die einer Gesellschaft bislang unbekanntes Wissen bereitstellt (Bak 2016, S. 1). Aufgrund der Vielzahl an Forschungsdisziplinen ist es kaum möglich Wissenschaft einheitlich zu beschreiben. Daher wird zumeist versucht ihren Kerngehalt einzugrenzen (Rost 2012). Die Brockhaus Enzyklopädie definiert Wissen dabei wie folgt: „Wissenschaft, das System des durch Forschung, Lehre und überlieferter Literatur gebildeten, geordneten und begründeten,

für gesichert erachteten Wissens einer Zeit“ (Brockhaus 1999). Dieser Definition nach, bezeichnet Wissenschaft also die Gesamtheit des Wissens der Menschheit, das durch nachvollziehbare Forschung entstanden und durch Dokumentation (z. B. in Büchern oder wissenschaftlichen Artikeln) festgehalten ist sowie durch Lehre vermittelt wird.

Die Publikation von Forschungsergebnissen stellt eine Form der Kommunikation dar (Riviera 2013). Sie wird auch als Wissenschaftskommunikation bzw. im Englischen als *scholarly communication* und *science communication* bezeichnet. Die beiden englischen Begriffe unterscheiden sich allerdings nach dem Adressatenkreis. Dabei bezieht sich die *scholarly communication* auf die interne Kommunikation zwischen Wissenschaftlern, während *science communication* die Kommunikation wissenschaftlicher Ergebnisse an die breite Öffentlichkeit meint. Die Kommunikation nach außen hat beispielsweise zum Ziel Mittel für zukünftige Projekte einzuwerben, während die interne Kommunikation auf das Problemlösen, den Erkenntnisgewinn und die Wissensgenerierung abzielt (Ball 2009; Seidenfaden, Ortelbach und Hagenhoff 2005).

2.1 Detailliertere Darstellung ausgewählter Arbeiten

Bereits Mitte des 20. Jahrhunderts betrachtet Lehman (1953) individuelle Leistungen von Individuen in verschiedenen Disziplinen, wie den Naturwissenschaften, der Philosophie, Musik, Kunst und Literatur. Für die einzelnen Disziplinen leitet er daraus unter anderem die „kreativen Lebensjahre“ eines Menschen her, das heißt die Jahre, in denen er besonders produktiv ist. Lehman betrachtet ferner die Unterschiede zwischen allen Beiträgen der Individuen in seinem Datensatz und den besonders wichtigen Beiträgen. Die Wichtigkeit der Beiträge lässt er dabei von ausgewählten Experten der einzelnen Disziplinen bewerten. Für alle Personen einer Disziplin aus seiner Stichprobe und deren Beiträge zählt er, wie viele der Beiträge im Alter zwischen 20 und 24 Jahren, 25 und 29 Jahren, 30 und 34 Jahren usw. geleistet wurden. Aus der Anzahl der Beiträge in jedem der 5-Jahres-Intervalle und der Anzahl der Personen errechnet er die mittlere Anzahl der Beiträge pro Kopf und Jahr in diesem Altersintervall. Diese relativen Beiträge setzt er anschließend in Relation zum jeweils höchsten ermittelten Wert und trägt diese Anteile gegen die 5-Jahres-Intervalle ab, wie dies am Beispiel der Chemie in Abbildung 2.1 zu sehen ist. Der Abbildung entnimmt man, dass Lehmans Ergebnissen zufolge Chemiker beispielsweise zu Beginn ihrer Dreißiger (30–34 Jahre) am produktivsten sind.

Dies ist an dem Hochpunkt der unterbrochenen Linie in Abbildung 2.1 auszumachen.¹ Die wichtigsten Erkenntnisse werden in der Chemie allerdings bereits in einem etwas jüngeren Alter, nämlich zwischen 25 und 30 Jahren getätigt. In Abbildung 2.1 ist dies daran zu erkennen, dass der Hochpunkt der durchgezogenen Linie, die die wichtigsten Beiträge repräsentiert, ein 5-Jahres-Intervall vor dem Hochpunkt der unterbrochenen Linie liegt.

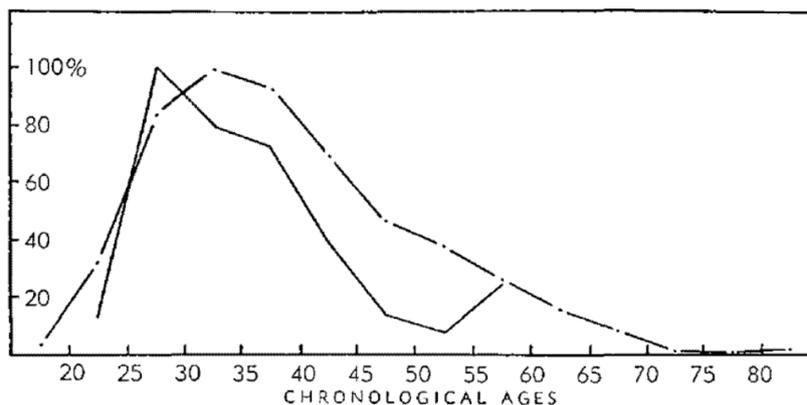


Abbildung 2.1: Produktivität von Chemikern in Abhängigkeit von ihrem Alter (Quelle: Lehman 1953, S. 7, Abbildung 2). Abgetragen ist das Verhältnis wissenschaftlicher Beiträge in jedem Jahr in Relation zum Lebensjahr mit den meisten Beiträgen. Durchgezogene Linie: wichtigste Beiträge, unterbrochene Linie: alle Beiträge in Lehmans Datensatz.

Zu ähnlichen Ergebnissen kommt Lehman auch für die anderen Disziplinen, wobei die Höchstleistung bei Physik oder Schach eher zu Beginn der Dreißiger liegt, für Medizin und Mathematik tendenziell Mitte der Dreißiger und die Astronomie mit Mitte 40 einen Ausreißer nach oben darstellt. Auch für den Bereich der Erfindungen bzw. Patente liegt das Hoch der Schaffenskraft laut Lehman Mitte Dreißig. Für Patente betrachtet Lehman unter anderem detailliert die Karriere von Thomas E. Edison, der im Laufe seines Lebens sehr produktiv war und in den USA insgesamt 1.086 Patente anmeldete. Ähnlich wie oben am Beispiel der Chemie beschrieben, zählt Lehman, wie viele Patente Edison in welchem Alter anmeldete. Für jedes Lebensjahr setzt er die Anzahl der angemeldeten Patente in Relation zur Anzahl der Anmeldungen im Lebensjahr mit den meisten Patentanmeldungen. Anschließend trägt er, wie dies in Abbildung 2.2 zu sehen ist, diese relativen Werte gegen die Lebensjahre von Edison ab.

¹Im Altersintervall zwischen 30 und 34 Jahren zählt Lehman für die Chemie 201 Beiträge in seinem Datensatz. Diese entsprechen dem höchsten jährlichen pro-Kopf-Beitrag aller 5-Jahres-Intervalle von 0,165. Diesem Wert ordnet Lehman den Wert 100 % zu. Alle weiteren Prozentwerte ergeben sich aus dem Quotienten des jeweiligen jährlichen pro-Kopf-Beitrags und dem Wert 0,165 (z. B. 0,115/0,165 für die Beiträge im Alter zwischen 40 und 44 Jahren).

In Abbildung 2.2 ist deutlich zu erkennen, dass Edison die meisten seiner Patente im Alter von 35 Jahren anmeldete und in diesem Sinne zu dem Zeitpunkt am produktivsten war.

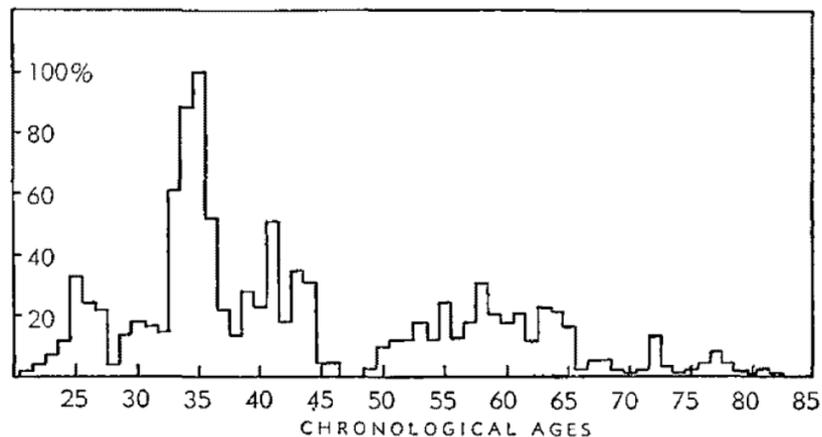


Abbildung 2.2: Produktivität von Thomas E. Edison hinsichtlich seiner in den USA patentierten Erfindungen in Abhängigkeit von seinem Alter (Quelle: Lehman 1953, S. 11, Abbildung 7). Abgetragen ist das Verhältnis angemeldeter Patente in jedem Jahr in Relation zum Lebensjahr mit den meisten Patentanmeldungen.

Während Lehman die Produktivität von Individuen im Verlauf ihres Lebenszyklus untersucht, fokussiert sich Jones (2009a) auf den ersten Beitrag und somit auf den Beginn der Karriere von Erfindern. Wie in der Einleitung bereits erwähnt, beschreibt Jones das Phänomen der Last des Wissens, welche langfristig zu einer Verlangsamung wissenschaftlichen Fortschritts führen könnte. Um seine Hypothesen zu stützen, verwendet Jones (2009a) einen US-amerikanischen Patentdatensatz von Hall, Jaffe und Trajtenberg (2001), der alle Patente beinhaltet, die die US-amerikanische Patentbehörde USPTO (*U.S. Patent and Trademark Office*) zwischen 1963 und 1999 veröffentlicht hat. Seine Untersuchung basiert auf 2,9 Millionen Patenten im Zeitraum zwischen 1975 und 1999. Die Ergebnisse zeigen, dass das mittlere Alter von Erfindern bei ihrer ersten Patentanmeldung und die mittlere Anzahl gelisteter Erfinder für ein Patent im Beobachtungszeitraum jeweils gestiegen sind. Dabei sind die Erfinder bei ihrem ersten Patent pro Dekade im Mittel um etwa 0,66 Jahre älter geworden und die Anzahl der pro Patent genannten Erfinder stieg von 1,73 auf 2,33. Jones zeigt ferner, dass die Anzahl der Erfinder positiv mit der Menge des einem Patent zugrunde liegenden Wissens korreliert. Dabei handelt es sich um jenes Vorwissen, das benötigt wird, um eine Erfindung zu machen. Er bemisst dieses anhand des Zitationsbaums eines gegebenen Patents. Dieser entsteht dadurch, dass bei Anmeldung eines Patents andere Patente und

wissenschaftliche Arbeiten zitiert werden, die wiederum weitere Quellen zitieren usw.² Jones wählt als einfaches Maß für die Größe dieses Zitationsbaums die Anzahl seiner Knoten.

Ferner zeigt Jones, dass Erfinder zunehmend spezialisiert sind. Jedes Patent in seinem Datensatz ist bestimmten technologischen Feldern zugeordnet. Um den Grad der Spezialisierung von Erfindern zu messen, betrachtet Jones aufeinander folgende Patentanmeldungen desselben Erfinders und vergleicht die technologischen Felder dieser Patentanmeldungen. Er berechnet daraus für Erfinder in einem gegebenen Jahr die mittlere Wahrscheinlichkeit, dass diese das technologische Feld von einem zum nächsten Patent wechseln. Jones' Ergebnisse zeigen, dass die Wahrscheinlichkeit das technologische Feld zu wechseln im beobachteten Zeitraum von ca. 57 % auf ca. 51 % gefallen ist, was er als Zunahme der Spezialisierung auffasst.

Wie der Titel „The Burden of Knowledge and the 'Death of the Renaissance Man': Is Innovation Getting Harder?“ bereits ausdrückt, interpretiert Jones (2009a) seine Ergebnisse so, dass es den klassischen Universalgelehrten aufgrund der Menge des bestehenden und zu verarbeitenden Wissens heutzutage kaum noch gibt und Wissensarbeiter sich, wie seine empirischen Ergebnisse zeigen, stärker spezialisieren. Melero und Palomeras (2015) greifen diese Interpretation auf und untersuchen, ebenfalls im Kontext von Patentanmeldungen, die Rolle von Generalisten innerhalb von Teams von Erfindern, die gemeinsam ein Patent anmelden. Generalisten identifizieren sie dabei über die Anzahl der technologischen Felder, denen Patente eines Erfinders im Laufe seiner Karriere zugeordnet wurden. Wie Jones (2009a) verwenden sie den Datensatz von Hall, Jaffe und Trajtenberg (2001) und ergänzen diesen um detailliertere Informationen zu den technologischen Feldern, denen die einzelnen Patente zugeordnet werden. Melero und Palomeras (2015) beschränken sich auf genehmigte Patente privater Firmen aus der Elektronikindustrie. Sie begründen dies damit, dass Firmen in dieser Industrie laut Hall (2005) besonders häufig jede ihrer Erfindungen und Verbesserungen patentieren lassen. Somit sollten bei Betrachtung dieser Industrie weniger verzerrende Effekte dadurch entstehen, dass Melero und Palomeras lediglich patentierte Innovationen betrachten. Ihre Ergebnisse zeigen, dass Generalisten unter gewissen Umständen tatsächlich zum Erfolg eines Innovationsteams beitragen. Sie interpretieren dies so, dass breiter aufgestellte Erfinder die Koordination und Kommunikation innerhalb eines Teams von Spezialisten übernehmen und damit die Leistungsfähigkeit eines solchen Teams steigern können. Damit widerlegen Sie Jones (2009a) zwar nicht, der etwas überspitzt vom Aussterben

²In der Sprache der Graphentheorie stellt jede der zitierten Quellen einen Knoten des Zitationsbaums dar.

des Universalgelehrten spricht, aber sie zeigen, dass Generalisten durchaus weiterhin ihre Berechtigung haben.

Card und DellaVigna (2013) untersuchen Publikationen der renommiertesten Zeitschriften („top-five“) der Volkswirtschaftslehre im Zeitraum von 1970–2012. Zu diesen Zeitschriften gehören *American Economic Review*, *Econometrica*, *Journal of Political Economy*, *Quarterly Journal of Economics* und *Review of Economic Studies*. Dabei stützt sich ihre Klassifizierung dieser Zeitschriften auf kein bestehendes Ranking und wird auch nicht anderweitig begründet (z. B. über den *Impact Factor*). Die Autoren verwenden bei ihrer Untersuchung Publikationsdaten der Publikationsdatenbank *Econ-Lit*, Statistiken bzgl. Einreichungen bei den genannten Zeitschriften (aus Berichten der Redakteure und persönlicher Kommunikation mit diesen) sowie *Google Scholar* für Informationen zu Zitationshäufigkeiten.

Wie der Titel ihres Artikels „Nine Facts about Top Journals in Economics“ bereits verspricht, fassen Card und DellaVigna ihre Untersuchungen in neun zentralen Ergebnissen zusammen. Sie zeigen unter anderem, dass die Annahmequote bei diesen Zeitschriften von 15 auf 6 % gefallen ist, da sich die Anzahl der Einreichungen zwischen 1990 und 2012 beinahe verdoppelt hat, während die Anzahl jährlich veröffentlichter Artikel sogar zurückgegangen ist. Sie berichten ferner, dass Artikel in den späteren Jahren des untersuchten Zeitraums etwa dreimal so lang sind, wie Artikel zu Beginn des Beobachtungszeitraums und sich die mittlere Anzahl der Autoren eines Artikels von 1,3 auf 2,3 erhöht. Somit gibt es in diesen Zeitschriften zwar weniger „freie Plätze“ für zu publizierende Artikel, aber dafür mehr „Plätze“ für Autoren, da die Anzahl der Autoren pro Artikel schneller gestiegen ist, als die Anzahl publizierter Artikel abgenommen hat.

Card und DellaVigna zeigen, dass die Reihenfolge der untersuchten fünf Zeitschriften bzgl. der erhaltenen Zitationen im Zeitverlauf weitestgehend konstant geblieben ist, mit der Ausnahme des *Quarterly Journal of Economics*, das sich von Platz vier auf Platz eins verbessert hat. Ferner werden längere Artikel und solche, die von mehr Autoren geschrieben wurden, häufiger zitiert. Letzteres Ergebnis lässt auf stärkere Netzwerkeffekte bei einer größeren Zahl an Autoren schließen. Die thematische Ausrichtung der Zeitschriften ist weitestgehend gleich geblieben.³ Allerdings gibt es für einzelne Forschungsbereiche sichtliche Veränderungen bei der Zitationshäufigkeit der entsprechenden Artikel. So werden in den letzten Jahren des Beobachtungszeitraums

³Die thematische Ausrichtung der Zeitschriften ermitteln Card und DellaVigna (2013) dabei anhand der auf den entsprechenden Artikeln angegebenen JEL-Codes (vgl. Abschnitt 3.1.2 ab S. 43 für eine Erläuterung zu JEL-Codes).

verstärkt Artikel aus dem Themenfeld *Entwicklungsökonomik und Außenwirtschaft* zitiert, während ökonometrische und theoretische Papiere immer seltener zitiert werden.

Hamermesh (2013) betrachtet Publikationen aus drei der renommiertesten Zeitschriften der Volkswirtschaftslehre, nämlich *American Economic Review*, *Journal of Political Economy* und *Quarterly Journal of Economics*, aus den Jahren 1963, 1973, 1983, 1993, 2003 und 2011 (insgesamt 748 Artikel). Die Publikationsdaten wurden um Informationen zum Alter der Autoren und der in den Artikeln verwendeten Methodik ergänzt. Hamermeshs Ergebnisse zeigen, dass der Anteil Autoren, die zum Zeitpunkt der Publikation 50 Jahre oder älter sind, im betrachteten Zeitraum deutlich angestiegen ist. So lag der Anteil der Autoren, die bei ihrer Publikation älter waren als 50, im Jahr 1963 noch bei ca. 4,2 % und stieg bis auf 18,9 % im Jahr 2011. Hamermesh diskutiert außerdem unterschiedliche Gründe für diese Entwicklung. Diese reichen von der verbesserten Gesundheit im Alter, über den zunehmend langen Publikationsprozess in der Volkswirtschaftslehre (vgl. auch Ellison 2002), bis hin zu längeren Ausbildungszeiten der publizierenden Wissenschaftler. Diesem letzten, scheinbar offensichtlichen Grund stehen allerdings Statistiken zum Durchschnittsalter von Promovierenden in den Vereinigten Staaten entgegen. So ist dieses seit Beginn der 1960er Jahre bis in die Mitte der 2000er Jahre im Mittel sogar um etwa ein Jahr zurückgegangen (Scott und Siegfried 2008, Tabelle A auf S. 633). Hamermeshs Ergebnisse bzgl. der Anzahl der Autoren decken sich mit denen von Card und DellaVigna (2013). Auch er zeigt, dass der Anteil der Artikel mit mehreren Autoren zugenommen hat. Hamermesh geht allerdings noch einen Schritt weiter und betrachtet die Altersdifferenzen innerhalb dieser Koautorenschaften. Er findet dabei keine eindeutigen Hinweise darauf, dass die meisten Koautorenschaften zwischen Personen mit hohem Altersunterschied stattfinden, wie dies beispielsweise bei der Zusammenarbeit eines Doktoranden mit seinem Doktorvater der Fall sein könnte.

Hamermesh (2013) untersucht ferner, welche Methoden in den Artikeln in seinem Datensatz angewandt werden und ob sich die Anteile der Methoden im Verlauf der Jahre ändern. Dabei stützt er sich im Gegensatz zu Card und DellaVigna (2013) nicht auf die Identifikation mittels JEL-Codes, sondern auf eine manuelle Zuordnung der verwendeten Methodik.⁴ Dennoch decken sich seine Ergebnisse weitestgehend mit denen von Card und DellaVigna. So zählen in den ersten drei Jahrzehnten mehr als 50 % der veröffentlichten Artikel zur Kategorie der theoretischen Papiere. Ab den 1990er Jahren sinkt dieser Anteil bis auf 19,1 % im Jahr 2011, wobei gleichzeitig der Anteil

⁴Vgl. Abschnitt 3.1.2 ab S. 43 für eine Erläuterung zu JEL-Codes.

empirischer Arbeiten, mit eigens erhobenen Daten, auf 34 % steigt (Hamermesh 2013, Tabelle 4 auf S. 168).

2.2 Wachstum, Fortschritt und Produktivität

Es gibt unterschiedliche Ansätze, das Wachstum der Wissenschaft und wissenschaftlichen Fortschritt zu messen. Smith (2005) gibt einen Überblick zur Messung kommerzieller Innovationen, während Ansätze zur Bemessung wissenschaftlicher Beiträge in Verbeek u. a. (2002) sowie Raan (2005) nachzulesen sind. Ein Ansatz kommerzielle Innovation zu messen sind Ausgaben für Forschung und Entwicklung sowie Maße, die den Output eines Unternehmens (z. B. Umsatz, Gewinn, neue Produkte) in Relation zu diesen setzen. Ein anderer Ansatz betrachtet die Anzahl angemeldeter und genehmigter Patente oder die Zitation von Patenten (Smith 2005). Die meisten Ansätze zur Messung wissenschaftlichen Wachstums und Fortschritts verwenden bibliometrische Maße (Verbeek u. a. 2002; Raan 2005). So wird in vielen Studien, von denen im Folgenden einige erwähnt werden, die Anzahl veröffentlichter wissenschaftlicher Arbeiten herangezogen, die als Proxy für die festgehaltene Wissensmenge und damit das Wachstum der Wissenschaft stehen. Hinsichtlich des Fortschritts wird versucht, die Relevanz wissenschaftlicher Arbeiten zu bemessen, wozu beispielsweise Zitationszahlen herangezogen werden. Bornmann und Mutz (2015) geben zu bedenken, dass man das Wachstum der Wissenschaft grundsätzlich auch anhand der Anzahl Wissenschaftler bemessen könnte. Allerdings habe dieses Maß ähnliche Limitationen wie die Anzahl der Publikationen und historische Informationen zur Anzahl aktiver Wissenschaftler seien nicht systematisch (z. B. anhand einer Datenbank) verfügbar.

In der Einleitung (Kapitel 1) wurde schon erwähnt, dass die absolute Menge wissenschaftlicher Publikationen in den vergangenen Jahrzehnten zugenommen hat. Eine Zunahme der Publikationen wird bereits von Price (1963, S. 20 f.) berichtet, der eine jährliche Wachstumsrate von 4,7 % ermittelt und somit von einer Verdopplung wissenschaftlichen Wissens alle 10–15 Jahre ausgeht. Auch in den Jahrzehnten nach der grundlegenden Arbeit von Price (1963) zählen diverse Studien die Anzahl der Publikationen in verschiedenen Disziplinen und bestätigen den Fortgang dieses Wachstums. Price (1963) zufolge sagte Barnaby Rich bereits 1613 aus: „One of the diseases of this age is the multiplicity of books; they doth so overcharge the world that it is not able to digest the abundance of idle matter that is every day hatched and brought forth into the world.“ Tague, Beheshti und Rees-Potter (1981) betrachten Veröffentlichungen

unterschiedlicher Disziplinen (u. a. Chemie, Physik, Biologie und Geschichte) und zeigen, dass deren Zahl wächst. Für die Chemie errechnen sie, unter der Annahme exponentiellen Wachstums, im Zeitraum zwischen 1907 und 1979 eine Verdopplung alle 15 Jahre (4,7 % p. a.). Sie betrachten das Wachstum zwischen 1960 und 1979 genauer und kommen zu dem Ergebnis, dass dieses für die meisten der von ihnen untersuchten Disziplinen abnimmt und dass es adäquater durch einen linearen als einen exponentiellen Wachstumsprozess beschrieben wird. Stuhlhofer (1983) kritisiert die Messung des Wachstums von Wissen anhand der Zählung von Publikationen, da ein Teil der Informationen in neuen wissenschaftlichen Arbeiten bereits bekannt sei und man diese daher nicht in vollem Umfang einem Wissenszuwachs zurechnen könne. Er zählt daher unter anderem die Anzahl bedeutender Entdeckungen oder bekannter Wissenschaftler im Zeitverlauf und kommt zu etwas niedrigeren Wachstumsraten, die zu einer Verdopplung des Wissens alle 100 Jahre führen.

Archibald und Line (1991) kommen zu ähnlichen Ergebnissen wie Tague, Beheshti und Rees-Potter (1981). Auch sie untersuchen Veröffentlichungen verschiedener Disziplinen wie der Mathematik, Philosophie, Medizin oder Geographie und zeigen, dass die Anzahl der Artikel pro Zeitschrift und Jahr zwischen 1950 und 1987 zugenommen hat. Allerdings finden sie auch, dass das Wachstum zwischen 1950 und 1970 am größten war und danach für die meisten Disziplinen abflacht. Mabe und Amin (2001) berichten ebenfalls von zunächst exponentiell steigenden Zahlen veröffentlichter Artikel in einzelnen Zeitschriften und einem Abflachen dieses Wachstums ab den 1980er Jahren. Das Abflachen der Anzahl Artikel pro Zeitschrift schreiben sie der Verteilung wissenschaftlicher Publikationen auf zunehmend viele Zeitschriften zu. Das Wachstum der Anzahl der Zeitschriften errechnen sie auf 3,3 % pro Jahr. Mabe und Amin zufolge ist das langsamere Wachstum der Anzahl Artikel pro Zeitschrift somit Ausdruck der gestiegenen Anzahl an Zeitschriften, in der sich wiederum das Wachstum der Wissenschaft und die Spezialisierung von Fachbereichen zeigen. Sie geben auch zu bedenken, dass die jährliche Wachstumsrate in der Wissenschaft tätiger Personen etwa 3,2 % beträgt und damit fast den Wachstumsraten wissenschaftlicher Publikationen entspricht. Larsen und Ins' (2010) Ergebnisse deuten darauf hin, dass es langsam und schnell wachsende Disziplinen gibt. Zu den langsam wachsenden gehören traditionelle Disziplinen wie Chemie, Mathematik und Physik, während Ingenieur- und Computerwissenschaften zu den schnell wachsenden gehören. Insgesamt schätzen Larsen und Ins' das Wachstum der Wissenschaft, gemessen an der Anzahl veröffentlichter Artikel, auf ca. 4,7 %, was in etwa der Schätzung von Price (1963) oder Tague, Beheshti und Rees-Potter (1981) entspricht.

In einer neueren Studie bemessen Bornmann und Mutz (2015) das Wachstum wissenschaftlicher Publikationen auf zwei Arten. Einerseits betrachten sie von der Literaturdatenbank *Web of Science* erfasste Artikel für den Zeitraum 1980 bis 2012 und ermitteln eine jährliche Wachstumsrate von 3 % für die Anzahl wissenschaftlicher Veröffentlichungen. Dies entspricht einer Verdopplung alle 24 Jahre. Andererseits betrachten sie die Quellenangaben aus obigen Artikeln genauer und zählen, wie viele davon aus welchem Jahr stammen. Auf Basis dieser Verteilung von erwähnten Arbeiten, kommen sie auf deutlich höhere jährliche Wachstumsraten von 9 bis 10 % und schließen daraus, dass das *Web of Science* nur einen Bruchteil der tatsächlich veröffentlichten Werke abdeckt und dass das Wachstum der Wissenschaft tatsächlich stärker ist als bislang angenommen.

Eine Zunahme an Publikationen ist dabei nicht zwingend einer entsprechenden Zunahme an Wissen gleichzusetzen. Tague, Beheshti und Rees-Potter (1981, S. 126), Stuhlhofer (1983) sowie Martin (1996) geben zu bedenken, dass man zwischen einem Anstieg der Anzahl von Publikationen und einer Zunahme von Wissen unterscheiden muss. Da die Anzahl der Veröffentlichungen eines Wissenschaftlers, im Allgemeinen, für die Beurteilung seiner wissenschaftlichen Leistung herangezogen wird (z. B. Hornbostel, Klingsporn und Ins 2009), gelten „[...] Publikationen als Erfolgskriterium im Wettbewerb der Wissenschaftlerinnen und Wissenschaftler um Karrierechancen, Forschungsmittel etc. [...]“ (Deutsche Forschungsgesellschaft 2013, S. 43). Aus dieser Motivation heraus versuchen Wissenschaftler möglichst viele Publikationen zu erzielen. Ash u. a. (2015, S. 32) bezeichnen dies als „[...] ein ‚inhaltsleeres‘ Wachstum, bei dem die Menge der Publikationen ansteigt, ohne dass die Menge an Forschungsergebnissen im gleichen Umfang zunimmt.“

Die starken Anreize zu publizieren führen ferner zu Praktiken, die den wissenschaftlichen Output steigern sollen und die Anzahl der Veröffentlichungen unnötig in die Höhe treiben. Eine dieser Strategien ist das sogenannte „salami sliced publishing“ bzw. „salami science“ (Broad 1981; Huth 1986; Bornmann und Daniel 2007). Dabei werden wissenschaftliche Ergebnisse in möglichst vielen einzelnen Artikeln veröffentlicht, anstatt diese in einer größeren Publikation niederzuschreiben. Es gibt außerdem Hinweise darauf, dass eine zunehmende Anzahl an aktiven Wissenschaftlern und ein damit einhergehender, verstärkter Wettbewerb, Anreize dafür setzen, Werke niedriger Qualität zu publizieren. Bobtcheff, Bolte und Mariotti (2017) argumentieren, dass sich Wissenschaftler einem Kompromiss gegenüber sehen zwischen dem Hervorbringen von Arbeiten hoher Qualität und dem Druck, mit Ergebnissen als erstes an die Öffentlichkeit zu gehen, um diese für sich beanspruchen zu können. Ball (2009, S. 48) weist darauf

hin, dass es einen zunehmenden Anteil wissenschaftlicher Artikel gibt, die nie zitiert werden und in diesem Sinne wertlos sind. Einzelne Wissenschaftler fordern sogar eine Reform in Bezug auf die Karriereanreize, um den quantitativen Anstieg an Publikationen einzudämmen und stattdessen auf weniger Publikationen von höherer Qualität zu setzen (Marx 2011). Die Zunahme an wissenschaftlichen Publikationen stellt Wissenschaftler zunehmend vor Herausforderungen, da sie die relevanten Informationen aus einer wachsenden Menge an redundanten Publikationen und solchen von minderer Qualität extrahieren müssen (Ash u. a. 2015, S. 33).

Laut Price (1963), Stuhlhofer (1983) oder Mabe und Amin (2001) ist der wissenschaftliche Fortschritt in Form wissenschaftlicher Erkenntnisse und Publikationen wesentlich durch eine wachsende Menge an Wissenschaftlern getrieben. Studien, die sich auf den Output pro Wissenschaftler fokussieren, stützen diese Sichtweise und kommen zu anderen Ergebnissen als die oben genannten Studien, die lediglich die absolute Anzahl an Publikationen betrachten. Mabe und Amin (2002) berichten, dass nicht nur die Anzahl der Publikationen gestiegen ist, sondern auch die Anzahl der Autoren pro Artikel. Dies ist in dem Maße geschehen, dass der Output pro Autor zurückgegangen ist, wenn man den Autoren ihre Artikel nur anteilig zurechnet. In diesem Sinne ist die Produktivität von Wissenschaftlern zurückgegangen. Larsen und Ins (2010) zufolge, ist eine mögliche Erklärung für diesen Rückgang der gemessenen Produktivität, dass das Verfassen und Publizieren von Artikeln zunehmend viel Arbeit erfordert. Eine andere Erklärung könnte sein, dass heutzutage vermehrt Arbeiten in Tagungsbänden oder in Open-Access-Portalen veröffentlicht werden, die gängige Literaturdatenbanken nicht alle erfassen. Entsprechend decken letztere einen zunehmend kleinen Teil tatsächlich veröffentlichter Werke ab, weshalb die verwendeten Produktivitätsmaße zu verzerrten Ergebnissen kommen. Auch die Ergebnisse der Studie von Bornmann und Mutz (2015) lassen auf die letzte Ursache schließen.

Jones (2009a) erläutert, dass sich Humankapital insofern von anderen Ressourcen unterscheidet, dass es nicht einfach übertragen werden kann. Er schreibt „the vessel of human capital—the individual—is born with little knowledge and absorbs information at a limited rate” (S. 283). Wie oben erläutert, ist die Menge an festgehaltenen Informationen im Laufe der Geschichte und speziell in den vergangenen Jahrzehnten stark gestiegen (vgl. speziell Lyman und Varian (2003) für einen Anstieg der Menge an festgehaltenen Daten und Informationen). Typischerweise wird diese Menge an Informationen erst nutzbar bzw. nützlich, wenn sie von Individuen (z. B. Wissenschaftlern oder Entwicklern) verarbeitet und verstanden wird. Die Verarbeitung von Informationen ist daher eine notwendige Bedingung, um die Wissensbasis eines Individuums aufzubauen und

seine Arbeit zu befruchten. Der Mensch ist allerdings natürlichen Grenzen ausgesetzt, die beispielsweise seine kognitive Leistung beeinflussen (Miller 1955; Kotovsky, Hayes und Simon 1985; Sweller 1994; Marois und Ivanoff 2005). Ferner argumentiert Marx (2011), dass das beobachtete starke Wachstum der Wissenschaft auch aufgrund des „nicht beliebig vermehrbaren Anteil[s] der Wissenschaftler an der Gesamtbevölkerung“ anhalten kann. In diesem Sinne erscheint das Potenzial des Menschen als Input für wissenschaftliches und volkswirtschaftliches Wachstum limitiert. Gegebenenfalls könnten somit diese Limitationen zu einer Verminderung der Produktivität und zu einer Verlangsamung des wissenschaftlichen Fortschritts führen.

In einer aktuellen Studie beschäftigen sich Bloom u. a. (2017) näher mit Forschungsproduktivität. Sie greifen die grundlegende Annahme des Modells von Solow (1962) und weiterer Wachstumsmodelle auf, dass ökonomisches Wachstum durch die Ideen und Erfindungen von Individuen getrieben wird. Die Wachstumsraten wiederum ergeben sich aus dem Produkt der Anzahl an Forschern und ihrer Produktivität. Bloom u. a. diskutieren diverse Beispielen dafür, dass das beobachtete, konstante Wachstum v. a. durch eine erhebliche Erhöhung des Forschungsaufwands, wie der Anzahl an Forschern oder der Ausgaben für Forschung und Entwicklung, erzielt wird, während die Forschungsproduktivität stark abnimmt. Ein Beispiel hierfür ist Moore's Law, das besagt, dass sich die Rechenleistung von Computern alle zwei Jahre verdoppelt (Moore 1975).⁵ Diese Wachstumsrate ist über fast ein halbes Jahrhundert erstaunlich konstant geblieben (Bloom u. a. 2017, Abbildung 3, S. 17). Bloom u. a. (2017, S. 18 f.) geben jedoch zu bedenken, dass heutzutage mehr als 18 mal so viele Forscher daran beteiligt sind, diese Wachstumsrate aufrechtzuerhalten, als es noch in den 1970er Jahren der Fall war. Sie zeigen ferner, dass die Forschungsproduktivität in den USA seit den 1930er Jahren um den Faktor 41 zurückgegangen ist, was einer jährlichen Abnahme von 5 % entspricht (Bloom u. a. 2017, S. 8).

Wenn individuelle Errungenschaften zunehmend schwieriger werden, mehr Einsatz erfordern, beispielsweise in Form von Kollaborationen oder Zeit, sollte sich dies auch in technologischem und wirtschaftlichem Output widerspiegeln. Tatsächlich berichten mehrere Studien von einer Verlangsamung des technologischen Fortschritts. Kortum (1993, 1997) zeigt ein abnehmendes Verhältnis von angemeldeten Patenten zu Ausgaben für Forschung und Entwicklung in den USA, während Evenson (1984, 1993) von abnehmendem Output (in Form von Patenten) pro Erfinder bezogen auf die Ausgaben für bzw. Investitionen in Forschung und Entwicklung berichtet. Schankerman und

⁵Rechenleistung wird in diesem Kontext als die Dichte verstanden, mit der Transistoren in Computer-Chips verbaut werden.

Pakes (1986, S. 1072, Tabelle 6) berichten von abnehmenden Patentanwendungen und Werten von Patenten pro Erfinder. In der Literatur finden sich drei verschiedene Erklärungen für diese Beobachtungen. Erstens mindert, wie Evenson ausführt, eine Ausschöpfung technologischer Möglichkeiten die Forschungsproduktivität. Segerstrom (1998) zeigt, dass Wachstumsraten unter der Annahme stetig wachsender Forschungs- und Entwicklungsausgaben nicht steigend sind. Zweitens, suggeriert Kortum (1997), dass eine Zunahme der Patentwerte den Wettbewerb in der Forschung verstärkt und zu erhöhten Ausgaben für Forschung und Entwicklung geführt hat. Drittens, weist Griliches (1989, 1990) auf zunehmend hohe Kosten des Patentierungsprozesses und damit eine erhöhte Hürde zur Anmeldung eines Patentes hin.

Trotz der verschiedenen Lasten, die wissenschaftlichen Fortschritt verlangsamen könnten, gibt es drei zentrale Stellschrauben für Verbesserungen, um mit der steigenden Wissensmenge adäquat umzugehen. Dies sind die schiere Erhöhung der Arbeitskraft, die Entwicklung neuer Technologien zur Informationsverarbeitung und die bereits erwähnte Zunahme in der Zusammenarbeit zwischen Wissenschaftlern. Price (1963) spricht von der Entwicklung zu immer mehr Zusammenarbeit als „one of the most violent transitions that can be measured in recent trends of scientific manpower and literature“ (S. 79). Aus heutiger Sicht erscheinen Informationstechnologien als vielversprechendes Instrument, um eine Last des Wissens wie sie Jones (2009a) beschreibt zu mindern. Verbesserte technologische Möglichkeiten, Daten zu speichern, zu verarbeiten und auszutauschen erleichtern das weitere Wachstum von Daten und Informationen sowie deren Verarbeitung und Nutzung (Hilbert und López 2011). Auf der anderen Seite tragen Informationstechnologien paradoxerweise selbst zur Informationsüberlastung bei (Hitt und Brynjolfsson 1996; Depken und Ward 2009).

2.3 Koautorenschaften

Shapin (1989) berichtet, dass Wissenschaft im 17. Jahrhundert mit Hilfe technischer Assistenten betrieben wurde. In diesem Sinne ist die Wissenschaft schon seit Langem keine Disziplin einsamer Wissenschaftler. Allerdings blieben die Assistenten eines Forschers zur damaligen Zeit weitestgehend unsichtbar und wurden nur in seltenen Fällen in den Schriften ihrer Arbeitgeber erwähnt. Heutzutage sind Online-Kommunikation und andere digitale Technologien zentrale Treiber von Kollaboration bzw. Zusammenarbeit. Mehrere Studien berichten von positiven Effekten von Informationstechnologien auf Kollaborationen über größere geografische Entfernungen, sowohl über Stadtgrenzen

im gleichen Land, als auch über Länder hinweg (z. B. Adams u. a. 2005; Jones 2009b; Freeman, Ganguli und Murciano-Goroff 2014). Freeman, Ganguli und Murciano-Goroff (2014) betonen allerdings die Wichtigkeit persönlicher Treffen für den Erfolg dieser Kooperationen. Manche Studien versuchen, direkt die Vorteile digitaler Medien für Kollaborationen abzuleiten. Ding u. a. (2010) berichten beispielsweise über positive Auswirkungen von Informationstechnologien auf die Produktivität von Wissenschaftlern und ihrer kollaborativen Netzwerke. Hamermesh und Oster (2002) beobachten einen Anstieg bei Koautorenschaften zwischen Autoren, die nicht an der gleichen Universität tätig sind, den sie auf die Verwendung von Informationstechnologien zurückführen. Überraschenderweise zeigen sie, dass die Produktivität dieser entfernten Koautorenschaften geringer ist.

Diverse Studien belegen seit den 1950ern für verschiedene wissenschaftliche Disziplinen einen langfristigen Trend hin zur Zusammenarbeit in größeren Teams. Hierzu zählen beispielsweise McDowell und Melvin (1983), Hudson (1996), Laband und Tollison (2000), Adams u. a. (2005) sowie Hamermesh (2015) für die Volkswirtschaftslehre, Zuckerman (1977) für die Medizin, Laband und Tollison (2000) für die Biologie, Henriksen (2016) für verschiedene Fachbereiche der Sozialwissenschaften oder Jones (2009a) sowie Singh und Fleming (2010) für Patente. Für die Human-, Natur- und Sozialwissenschaften kommen Larivière, Gingras und Archambault (2006) ferner zu dem Ergebnis, dass internationale Koautorenschaften zugenommen haben. Greene (2007) berichtet von einer beinahe verschwindenden Menge an Publikationen mit lediglich einem Autor.

Cronin, Shaw und La Barre (2003) zeigen, dass der Anteil der Artikel mit mehr als einem Autor in der Zeitschrift *Psychological Review* zwischen 1900 und 1999 gestiegen ist. Wuchty, Jones und Uzzi (2007) analysieren nahezu 20 Millionen Artikel im Zeitraum von 1955 bis 2000 aus den Natur- und Ingenieurwissenschaften, den Sozial- und Geisteswissenschaften, sowie der Kunst. Bis auf die Fachbereiche der Kunst und Geisteswissenschaften, in denen weiterhin 90 % der Arbeiten von einem einzigen Autor verfasst werden, ist ein Trend hin zur Koautorenschaft bei wissenschaftlichen Artikeln zu beobachten. Beispielsweise steigen die Anteile von Artikeln mit mehr als einem Autor für die Mathematik von 19 auf 57 % und die Sozialwissenschaften von 17,5 auf 51,5 %. Am stärksten fällt dieser Anstieg laut Wuchty, Jones und Uzzi (2007) in den Natur- und Ingenieurwissenschaften aus, in denen sich die Anzahl der Autoren innerhalb des Beobachtungszeitraums von 45 Jahren von 1,9 auf 3,5 Autoren nahezu verdoppelt. In den Naturwissenschaften ist die Anzahl der Autoren eines Artikels besonders stark angewachsen, sodass beispielsweise Cronin (2001) von „hyperauthorship“ spricht.

Beispielsweise berichtet Castelvechi (2015) von Artikeln der Physik und der Biologie mit bis zu mehreren tausend Autoren. Simonton (2013) argumentiert, dass ein Großteil der wissenschaftlichen Erkenntnisse heutzutage in großen Forschungsgruppen erarbeitet wird, da die Fachbereiche in den Naturwissenschaften gewachsen seien, die nötige Wissensbasis komplex ist und Wissenschaftler sich zunehmend spezialisieren.

Rath und Wohlrabe (2016) sowie Nowell und Grijalva (2011) finden, dass die Teamgröße positiv mit dem Ranking bzw. dem *Impact Factor* der Zeitschrift, in der der entsprechende Artikel publiziert wird, korreliert. Boschini und Sjögren (2007) zeigen, dass die Wahrscheinlichkeit für Koautorenschaften mit Autoren unterschiedlicher Geschlechter geringer ist, als die für Teams von Autoren des gleichen Geschlechts. Ferner finden sie heraus, dass Frauen Artikel häufiger alleine verfassen als ihre männlichen Kollegen und dass Frauen grundsätzlich weniger zur Zusammenarbeit mit anderen Autoren neigen.

Manche Arbeiten untersuchen auch weniger explizite Formen der Zusammenarbeit. Beispielsweise zählen Laband und Tollison (2000) die Anzahl der erwähnten Kollegen in den Danksagungen und anderen Textteilen wissenschaftlicher Arbeiten und finden für den Zeitraum von 1950 bis 1995 einen steigenden Trend. Cronin, Shaw und La Barre (2003) untersuchen Artikel der Zeitschrift *Psychological Review* und finden heraus, dass die Anzahl der Artikel mit Danksagungen im Zeitraum zwischen 1900 und 1999 zugenommen hat. Franzoni und Sauerermann (2014) berichten vom Phänomen der *Crowd Science*, das vor allem in den Naturwissenschaften zu beobachten ist. Dabei handelt es sich um wissenschaftliche Projekte, bei denen hunderttausende Personen an der Lösung eines gegebenen Problems beteiligt sind, allerdings nicht namentlich auf den entsprechenden Artikeln erwähnt werden. Solche Projekte ermöglichen eine schnelle und kostengünstige Lösung von Problemen und entsprechende Ergebnisse werden häufig in renommierten Zeitschriften wie *Nature* veröffentlicht. Die genannten Studien lassen auch auf weniger formelle Formen der intellektuellen Zusammenarbeit und der Anerkennung des Beitrags von anderen schließen, die nicht in Form von Koautorenschaften sichtbar sind.

Für die Entwicklung hin zu mehr Kooperation zwischen Wissenschaftlern finden sich in der Literatur verschiedene Erklärungen. Melin (2000) berichtet von den Ergebnissen einer Umfrage unter Wissenschaftlern zu den Motiven für Kollaboration. Zu den am häufigsten genannten gehören, dass der Koautor spezifische Fertigkeiten oder spezifisches Wissen hat, dass dieser im Besitz spezieller Daten oder Forschungsausstattung ist und dass man sich eine Erhöhung der Qualität der Forschung durch die Zusammenarbeit

erhofft. McDowell und Melvin (1983), Barnett, Ault und Kaserman (1988) sowie Jones (2009a) argumentieren, dass wenn Individuen Wissen heute nicht mehr in der gleichen Breite wie früher abdecken können, es eine mögliche Strategie sei, diesem Problem mit größeren Teams zu begegnen, in denen sich jedes Mitglied nur mit einem bestimmten Teil des Problems beschäftigt. Für die Volkswirtschaftslehre erklärt Hudson (1996) die Zunahme in der Teamgröße (d. h. der Anzahl der Autoren pro Artikel) durch das Wachstum der einzelnen Fachbereiche, das sich unter anderem durch eine Zunahme in der Quantität wissenschaftlicher Arbeiten äußert. Er bemerkt ferner, dass sich das Wachstum einer Disziplin auch in der Anzahl der Wissenschaftler widerspiegelt, die in dieser tätig sind. Mit einer zunehmenden Menge an Wissenschaftlern im eigenen Fachgebiet eines Wissenschaftlers, gibt es ferner mehr Möglichkeiten zur Kooperation, was ebenfalls eine Erklärung für die Zunahme an Kooperationen sowie die gestiegenen Teamgrößen sein kann.

Neben der Erklärung der Arbeitsteilung finden Barnett, Ault und Kaserman (1988), auf Basis einer Untersuchung von Artikeln im *American Economic Review*, Evidenz für zwei weitere Hypothesen, die die Zunahme der Koautorenschaften erklären könnten. Eine davon ist, dass die Opportunitätskosten von Wissenschaftlern gestiegen sind, da sie bei der steigenden Menge an Publikationen mehr Zeit aufwenden müssen, um auf dem neuesten Stand in ihrem Forschungsbereich zu bleiben und zusätzlich die zeitintensive Einwerbung von Drittmitteln an Bedeutung gewonnen hat. Entsprechend würden heute Kollegen als Koautoren gelistet, die eher die Rolle eines Gutachters inne hätten, an der eigentlichen Verfassung des Artikels kaum beteiligt waren und früher beispielsweise lediglich in den Danksagungen erwähnt worden wären. Die zweite Hypothese betrifft eine Diversifikation hinsichtlich des Begutachtungsprozesses von Artikeln. Durch die Zusammenarbeit mit anderen Autoren lässt sich die Anzahl verfasster Artikel erhöhen und somit die Chance, dass zumindest eine der Arbeiten zeitnah zur Publikation angenommen wird. Laut Hamermesh (2015) sind unterschiedliche Forschungsgewohnheiten von Wissenschaftlern unterschiedlicher Generationen eine weitere Erklärung. Wissenschaftler der jüngsten Generation würden häufiger kollaborieren, als ihre erfahreneren Kollegen.

Größere Teams ermöglichen ferner eine heterogene Zusammenstellung von Wissenschaftlern, die wiederum heterogenes Wissen zum Team beitragen, welches laut Aral, Brynjolfsson und Van Alstyne (2012) das Leistungspotenzial steigern kann. Zu den Vorteilen größerer Teams gehört außerdem, dass Artikel, die von größeren Teams verfasst werden, häufiger zitiert werden (Wuchty, Jones und Uzzi 2007; Franceschet und Costantini 2010; Larivière, Gingras, Sugimoto u. a. 2015; Mukherjee u. a. 2017). Singh

und Flemings (2010) Patentanalyse deutet darauf hin, dass Teams eher dazu neigen bahnbrechende Erfindungen hervorzubringen, während die Qualität bzw. die Relevanz der Patente mit nur einem Patentanmelder geringer ausfällt.⁶ Melero und Palomeras (2015) untersuchen Patentierungsprozesse und finden heraus, dass in 12–15 % der Fälle, in denen „the innovation process involves an especially high degree of uncertainty“, Teams von Generalisten, das heißt weniger spezialisierten Erfindern, profitieren können. In der Mehrzahl der Fälle hatten Generalisten allerdings keinen oder sogar negativen Einfluss auf die Forschungsergebnisse eines Teams. Die Ergebnisse von Lee und Bozeman (2005) zeigen, dass die Arbeit in Teams den Zugang zu öffentlichen Geldern erleichtert und in einer höheren Produktivität resultiert.

Vereinzelt werden auch nachteilige Effekte der Arbeit in Teams berichtet. Hierzu zählen zum Beispiel die Kosten der Koordination (Organisation und Kommunikation) innerhalb des Teams (Hudson 1996; Cummings und Kiesler 2007; Bercovitz und Feldman 2011). Dabei kann es zu Differenzen in den Vorstellungen der einzelnen Teammitglieder bzw. deren Institutionen kommen, beispielsweise bezüglich der Zeitschrift, in der die Ergebnisse veröffentlicht werden sollen oder bezüglich der Verantwortlichkeiten für einzelne Aufgaben (Cummings und Kiesler 2007). Daher ist das Eingehen von Kompromissen bei der Zusammenarbeit in Teams unerlässlich (Hudson 1996). Ferner erhöht sich der Aufwand der Überwachung für den Teamleiter, da die Komplexität eines Teams, das gegebenenfalls über verschiedene Institutionen verteilt ist, einen großen Teil seiner eingeschränkten Aufmerksamkeit in Anspruch nimmt (z. B. Masten, Meehan und Snyder 1991). Bercovitz und Feldman (2011) zeigen, dass die höheren Kosten der Koordination eines Teams, dessen Mitglieder in unterschiedlichen Disziplinen tätig sind oder von unterschiedlichen Institutionen stammen, die Wahrscheinlichkeit der erfolgreichen Kommerzialisierung einer Erfindung senken. Die Kosten der Koordination und der Konflikte steigen ferner mit der Größe des Teams (Holmstrom 1982; Becker und Murphy 1992; Hudson 1996).

2.4 Quellenangaben und Zitationen

Die Anzahl der Quellenangaben in und der Zitationen von Artikeln sind gängige und vielfach studierte bibliometrische Maße. Ersteres ist dabei rückwärts gerichtet und gibt an, auf welche Quellen sich der Autor bzw. die Autoren bei der Verfassung eines

⁶Singh und Fleming (2010) bemessen die Qualität bzw. Relevanz von Patenten daran, wie häufig diese zitiert werden.

Artikels berufen und damit die Arbeiten anderer Wissenschaftler anerkennen. Letzteres ist vorwärts gerichtet und gibt an, wie häufig ein Artikel nach seiner Veröffentlichung zitiert wurde. Trotz einiger Limitationen, die die Aussagekraft dieser Maße betreffen, werden in diversen wissenschaftlichen Arbeiten entweder die Maße selbst und deren Entwicklung über die Zeit untersucht oder sie werden zur Ableitung weiterer Indikatoren verwendet.⁷ Beispielsweise wird die Anzahl der Zitationen, das heißt wie häufig ein Artikel oder Patent zitiert wurde, als Proxy für die Qualität und Relevanz von Artikeln oder Patenten (Gross und Gross E. M. 1927; Garfield 1972; Nicolaisen 2002; Bornmann und Daniel 2008; Mukherjee u. a. 2017), Autoren (z. B. Hirsch 2005) und Zeitschriften (Gross und Gross E. M. 1927; Garfield 1972; Engemann und Wall 2009) herangezogen. Die Anzahl der Quellenangaben in einem Artikel oder Patent dient zum Beispiel als Proxy für die in dem Artikel oder Patent verarbeitete Wissensmenge (z. B. Jones 2009a) sowie für das Wachstum der Wissenschaft (Bornmann und Mutz 2015). Bornmann und Mutz (2015) analysieren mehr als 38,5 Millionen Publikationen im Zeitraum 1980 bis 2012 aus verschiedenen wissenschaftlichen Disziplinen. Aus den Quellenangaben in obigen Arbeiten leiten sie eine historische Verteilung der entsprechenden zitierten Arbeiten (insgesamt über 755 Millionen) über die Jahre 1650 bis 2012 ab. Aus einer segmentierten linearen Regressionsanalyse erhalten sie drei größere Phasen des historischen Wachstum der Wissenschaft, die unterschiedliche Wachstumsraten aufweisen. Sie unterscheiden den Zeitraum zwischen Mitte des 17. und Mitte des 18. Jahrhunderts, die darauffolgende Zeit bis zur Zeit zwischen den beiden Weltkriegen und die Zeit danach mit Wachstumsraten von weniger als 1 %, 2–3 % bzw. 8–9 %.

Arbeiten, die sich mit der Anzahl der Quellenangaben in Artikeln beschäftigen, kommen für verschiedene Disziplinen zu dem Ergebnis, dass diese in den vergangenen Jahrzehnten gestiegen ist. Für die Biochemie berichten (Yitzhaki und Ben-Tamar 1991) von einer Zunahme der Anzahl der Quellenangaben. Sie betrachten Artikel, die in den Jahren zwischen 1910 und 1985 in der Zeitschrift *Journal of Biological Chemistry* erschienen sind und ermitteln einen Anstieg der mittleren Anzahl Quellenangaben von 10,0 im Jahr 1910 auf 37,0 im Jahr 1985. Ähnliche Trends werden von Krampen (2010) für die Psychologie berichtet, der eine Verdopplung der Anzahl der Quellenangaben zwischen 1985 und 2005 für 90 psychologische Artikel beobachtet. Auch für die Biologie (Cordero u. a. 2016), die Ingenieur- (Ucar u. a. 2013) sowie Finanzwissenschaften (Esen, Tunahan und Takil 2016) zeigt sich dieser positive Trend. Cordero u. a. (2016) untersuchen dabei 1.464 Artikel im Zeitraum zwischen 1993 und 2013 aus zehn verschiedenen Zeitschriften mit dem Fokus auf Biologie. Ucar u. a. (2013) untersuchen über 70.000

⁷Die Limitationen dieser Maße werden in Kapitel 6, S. 169, diskutiert.

Artikel aus ingenieurwissenschaftlichen Zeitschriften im Zeitraum zwischen 1972 und 2013. Im Beobachtungszeitraum wächst die mittlere Anzahl der Quellenangaben von 8 auf 25. Esen, Tunahan und Takil (2016) untersuchen 999 Artikel aus türkischen finanzwissenschaftlichen Zeitschriften zwischen 2010 und 2014. Sie beobachten in diesem Zeitraum einen leichten Anstieg der mittleren Anzahl Quellenangaben für alle vier untersuchten Zeitschriften. Eine umfangreiche Untersuchung von Kuld und O'Hagan (2018) für die Volkswirtschaftslehre in den Jahren 1996–2014 wurde kürzlich veröffentlicht. Kuld und O'Hagan untersuchen 175.000 Artikel aus 255 Zeitschriften und kommen zu dem Ergebnis, dass die Anzahl der Quellenangaben von ca. 24 auf ca. 38 gestiegen ist.

2.5 Altersdaten und Produktivität von Individuen

Diverse Arbeiten haben in der Vergangenheit das Alter von Individuen im Zusammenhang mit deren Produktivität bzw. Errungenschaften (wissenschaftliche und nicht-wissenschaftliche) untersucht. Wie bereits ausgeführt, betrachtet Jones (2009a) vor dem Hintergrund einer möglichen Last des Wissens das Alter von Patentanmeldern und findet, dass dieses im Zeitraum zwischen 1975 und 1999 gestiegen ist. Jones und Weinberg (2011) betrachten insgesamt 525 Nobelpreisträger für Physik, Chemie und Medizin. Sie untersuchen, wie alt diese zum Zeitpunkt der Veröffentlichung jener Erfindung waren, für die sie später einen Nobelpreis erhielten. Aus den Ergebnissen geht hervor, dass die relevanten Arbeiten vor 1905 für alle drei Bereiche in über 60 % der Fälle vor dem 40. Lebensjahr abgeschlossen wurden. Während der Anteil der unter 30-jährigen vor 1905 noch bei etwa 20 % lag, sinkt er bis zum Jahr 2000 auf fast 0 %. Der Anteil der unter 40-jährigen bleibt für Medizin fast unverändert, sinkt für Physik allerdings auf 19 % und für Chemie gar auf fast 0 %.

Weiter oben wurden ferner auch Lehmans (1953) Ergebnisse zum Alter höchster Produktivität in den Naturwissenschaften, der Mathematik, Medizin, Musik, Philosophie und weiteren Disziplinen ausgeführt. Einige Jahre später untersucht Lehman (1962) Lebensläufe von Menschen in Führungspositionen, wie Staatsmännern, Richtern oder Universitätspräsidenten. Auch für diese Personengruppen findet er heraus, dass die Rate kreativer Arbeit von höchster Qualität eines Individuums während dessen Dreißigern am höchsten ist. Mehrere Studien berichten einen Anstieg beim Alter in dem große Leistungen erbracht werden bzw. damit im Einklang, einen späteren Beginn der kreativen Schaffenszeit. Beispielsweise zeigt Jones (2010), dass wissenschaftliche

Leistungen zunehmend in einem höheren Alter erbracht werden und argumentiert, dass dies in längeren Phasen der Ausbildung im frühen Lebenszyklus eines Individuums begründet sein könnte. Zu ähnlichen Ergebnissen kommen Jones und Weinberg (2011) für den Höhepunkt der wissenschaftlichen Kreativität und Conley, Crucini u. a. (2013) für die wissenschaftliche Produktivität junger Wissenschaftler in den ersten Jahren nach Erlangung ihres Dokortitels. Hamermesh (2013) kommt zu dem Ergebnis, dass Autoren von Artikeln, die in den Zeitschriften *American Economic Review*, *Journal of Political Economy* und *Quarterly Journal of Economics* veröffentlicht werden, zunehmend älter sind.

Becker (1962) und Schultz (1963) leisten grundlegende Arbeit bzgl. des theoretischen Verständnisses möglicher Ursachen für die Produktivität von Individuen im Laufe ihres Lebenszyklus. Ihre Modelle besagen, dass eine inverse U-Form der Produktivitätskurve das Ergebnis von Entscheidungen bzgl. des Humankapitals sein kann. Levin und Stephan (1991) untersuchen die Produktivität von Ph.D.-Studenten in sechs Teilbereichen der Physik und der Geowissenschaften. Sie zeigen, dass die Produktivität für alle bis auf einen der sechs Bereiche der Form eines inversen U folgt, ihren Hochpunkt in den frühen Vierzigern hat und danach stark abfällt.⁸ Ähnliche Verläufe der Produktivität über den Lebenszyklus hinweg werden auch von Jones (2010) für Erfinder und von Conley, Crucini u. a. (2013) für Wissenschaftler aus der Volkswirtschaftslehre berichtet.

Andere empirische Studien untersuchen weitere individuelle und institutionelle Variablen, die die wissenschaftliche Produktivität beeinflussen könnten. Beispielsweise betrachten Buchmueller, Dominitz und Hansen (1999) die langfristigen Effekte von Hochschulbildung, der ersten Anstellung nach Abschluss des Studiums und der individuellen Charakteristiken auf die Produktivität von Volkswirten. Sie zeigen, dass die ersten Erfahrungen mit Forschung und dem Publikationsprozess während der Promotionszeit positiv mit dem Publikationserfolg in späteren Jahren korrelieren. Gonzalez-Brambila und Veloso (2007) kommen zu dem Schluss, dass Anreizsysteme eine entscheidende Rolle dabei spielen, ob Wissenschaftler auch im höheren Alter produktiv bleiben.

Franzoni, Scellato und Stephan (2011) betrachten verschiedene nationale Anreizsysteme und zeigen, dass institutionelle und finanzielle Anreize, sowie Karriereanreize positiv mit den Einreichungsquoten der angestellten Wissenschaftler korrelieren. In-

⁸Erstaunlicherweise scheint sich die landläufige Meinung zu diesem Thema zwischen 1953 und 1991 geändert zu haben. Während Lehman (1953, S. 12) bemerkt, dass „school texts usually picture the renowned scientist as a man of rather advanced years“, machen Levin und Stephan (1991, S. 114) die gegensätzliche Beobachtung, dass „a popular belief held by scientists and the lay public alike is that science is a young person’s game“.

teressanterweise, führen lediglich Karriereanreize auch zu höheren Annahmehquoten. Finanzielle Anreize sind hingegen negativ mit den Annahmehquoten korreliert. Die geringen Auswirkungen von monetären Anreizen werden auch von Jacob und Lefgren (2011) berichtet, die einen Anstieg von 7 % bei den Publikationen von Wissenschaftlern feststellen, die Gelder von nationalen Gesundheitsinstituten bekommen. Im Gegensatz dazu zeigen Önder und Schweitzer (2017) den substanziellen Einfluss von Karriereanreizen auf die Publikationsaktivität am Beispiel deutscher Ökonomen. Sie berichten, dass der Anteil publizierender Doktoranden, nach einer Reform des Betreuungssystems von Doktoranden, hin zu einem System mit individuellen Karriereanreizen zu publizieren, von 18 % im Jahr 1991 auf 46 % im Jahr 2008 angestiegen ist. Eine noch ungelöste Herausforderung für Anreizsysteme in der Forschung, ist der Unterschied zwischen den Geschlechtern bzgl. des Publizierens. Beispielsweise untersuchen Beaudry und Larivière (2016) mögliche Unterschiede zwischen den Geschlechtern bzgl. des wissenschaftlichen Impacts. Sie kontrollieren dabei für Faktoren wie Forschungsgelder oder die Anzahl der Koautoren und kommen zu dem Ergebnis, dass Wissenschaftler in den Naturwissenschaften und der Medizin seltener zitiert werden, die mit mehr weiblichen Koautoren zusammenarbeiten.

2.6 Einordnung und Beitrag der vorliegenden Arbeit

Die meisten der bislang erwähnten Studien, die sich mit Fortschritt in und dem Wachstum der Wissenschaft beschäftigen, messen in irgendeiner Weise wissenschaftlichen und technologischen Output. Hamermesh und Oster (2002), Ellison (2002), Conley, Crucini u. a. (2013), Hamermesh (2013) untersuchen Output in Form von Publikationen, Evenson (1984, 1993), Schankerman und Pakes (1986), Griliches (1989, 1990) sowie Jones (2009a) in Form von Patenten und Lehman (1953, 1962) untersucht zusätzlich weitere andere individuelle Leistungen. Im Sinne von Maßen des Fortschritts ist die vorliegende Studie am nächsten verwandt mit der ersten Gruppe, die wissenschaftliche Publikationen untersucht. Ein Aspekt, der diese Arbeit von den zuvor genannten Studien unterscheidet, ist, dass zusätzlich zum Output auch der Input von Wissen anhand der in einem Artikel angegebenen Quellen beschrieben bzw. gemessen werden soll. Hierbei orientiert sich die vorliegende Arbeit an den Beispielen von Yitzhaki und Ben-Tamar (1991), Ucar u. a. (2013) sowie Bornmann und Mutz (2015), die die Anzahl an Quellenangaben in Artikeln untersuchen.

Ein Beitrag der vorliegenden Arbeit ist ihr ganzheitlicher Ansatz, bei dem für verschiedene in einem Datensatz enthaltene Disziplinen, unterschiedliche Maße sowohl mittels deskriptiver als auch induktiver Statistik unter einer einheitlichen Fragestellung untersucht werden. Im Gegensatz dazu stellen die meisten der oben erwähnten Arbeiten punktuelle Untersuchungen dar, die beispielsweise lediglich eine Zeitschrift (z. B. Cronin, Shaw und La Barre 2003; Krampen 2010), eine Auswahl von Zeitschriften *einer* Disziplin (Card und DellaVigna 2013; Hamermesh 2015) oder lediglich ein Maß (Yitzhaki und Ben-Tamar 1991; Hudson 1996; Henriksen 2016) untersuchen. Einige Arbeiten beschränken sich ferner auf die deskriptive Untersuchung von Trends ohne belastbare Tests bzw. Ursache-Wirkungs-Analysen (z. B. Hudson 1996; Henriksen 2016). Henriksen (2016) untersucht zwar mehr psychologische Zeitschriften als die vorliegende Arbeit, beschränkt sich allerdings auf das Maß der Teamgröße sowie den Zeitraum von 1980–2013 und berichtet keine Regressionsergebnisse. Yitzhaki und Ben-Tamar (1991) beschränken sich auf die Untersuchung der Anzahl der Quellenangaben im *Journal of Biological Chemistry*. Sie betrachten allerdings einen längeren Zeitraum (1910–1985) als die vorliegende Arbeit. Card und DellaVigna (2013) wiederum wählen einen ähnlichen Untersuchungszeitraum (1970–2012) und untersuchen beispielsweise mit der Anzahl Zitationen sowie Einreichungs- und Annahmequoten Maße, die in der vorliegenden Arbeit nicht betrachtet werden. Ihre Studie beschränkt sich allerdings auf die fünf renommiertesten Zeitschriften der Volkswirtschaftslehre. Grundsätzlich lässt sich somit nicht allgemein sagen, dass die Arbeit längere Zeiträume oder mehr Daten, Disziplinen und/oder Maße untersucht als alle bestehenden Arbeiten. Allerdings findet sich die Breite der vorliegenden Untersuchung hinsichtlich dieser Aspekte bislang nicht in der Literatur.

Wie in den vorhergehenden Abschnitten dieses Kapitels beschrieben, wurden etablierte Maße wie die Anzahl der Autoren eines Artikels oder die Anzahl der Quellenangaben in einem Artikel bereits in vergangenen Arbeiten untersucht, wobei die bestehende Literatur zur Teamarbeit wesentlich breiter ist als die zu den Quellenangaben. Insofern werden bestehende Ergebnisse hinsichtlich dieser Maße auch für die in dieser Arbeit untersuchten Datensätze, Disziplinen und zugehörigen Zeitschriften sowie Untersuchungszeiträume bestätigt. Eine Bestätigung der Ergebnisse anhand verschiedener Datensätze ist wichtig, da einschlägige Datenbanken Publikationen und die zugehörigen Metainformationen unterschiedlich vollständig abdecken (z. B. Mabe und Amin 2001; Bornmann und Mutz 2015).

Die erwähnten Studien, die sich mit der Anzahl der Quellenangaben oder der Teamgröße von Artikeln beschäftigen, betrachten entweder eine Auswahl der renom-

miertesten Zeitschriften einer Disziplin (z. B. McDowell und Melvin 1983; Hamermesh 2013; Card und DellaVigna 2013) oder auf aggregierter Ebene eine Vielzahl von Zeitschriften ohne zwischen diesen zu unterscheiden (Bornmann und Mutz 2015; Henriksen 2016). Für die Volkswirtschaftslehre unterscheidet die vorliegende Arbeit zwischen Artikeln und Autoren der neun renommiertesten Zeitschriften (Top-Zeitschriften) sowie aller volkswirtschaftlichen Zeitschriften und ermöglicht so Unterschiede hinsichtlich der untersuchten Maße zwischen diesen beiden Gruppen aufzuzeigen. Ein ähnlicher Vergleich findet sich in einer neueren Studie von Kuld und O'Hagan (2018), die Artikel der renommiertesten 20 bzw. 255 Zeitschriften der Volkswirtschaftslehre zwischen 1996 und 2014 gegenüberstellen. Hinsichtlich derselben Maße, decken sich die Ergebnisse mit denen aus der vorliegenden Arbeit.

Die vorliegende Arbeit erweitert die bestehenden Ergebnisse hinsichtlich der Teamgröße und der Anzahl der Quellenangaben, indem diese auch für den jeweils *ersten* Artikel eines Autors untersucht werden. Dieser zeitliche Aspekt wurde bislang lediglich von Jones (2009a) für das Alter von Erfindern bei ihrer ersten Patentanmeldung untersucht. Im Hinblick auf die Forschungsfrage dieser Arbeit, ermöglicht diese Betrachtungsweise eine Abschätzung des Wissens, das junge Wissenschaftler bis zu ihrer ersten Publikation verarbeiten mussten und der Dauer dieses Verarbeitungsprozesses. Die Einschränkung auf erste Artikel von Autoren stellt ferner sicher, dass die beobachteten Trends nicht durch Veränderungen im Karriere- bzw. Lebenszyklus von Wissenschaftlern getrieben sind und ermöglicht so die zeitlichen Trends für die einzelnen Disziplinen besser herauszuarbeiten. Beispielsweise könnten Wissenschaftler im Laufe ihrer Karriere in zunehmend großen Teams arbeiten oder zunehmend viele Quellen in ihren Arbeiten angeben. In diesem Fall würden diese individuellen Effekte der Karriere- und Lebenszyklen die zeitlichen Effekte für die Disziplin überlagern.

Die vorliegende Arbeit trägt ferner mit der Untersuchung des Alters von Autoren bei ihrem ersten Artikel zur bestehenden Literatur bei. Entsprechende Informationen sind für Autoren der Volkswirtschaftslehre sowie Autoren der Mathematik verfügbar. In der Literatur finden sich nur wenige Untersuchungen zum Alter von Autoren, vermutlich, weil es aufwändig ist diese Information zu ermitteln. Wie erwähnt betrachtet Lehman (1953, 1962) das Alter von Individuen hinsichtlich ihrer Produktivität, Jones (2010) untersucht das Alter von Nobelpreisträgern bei der Erfindung, für die sie den Nobelpreis erhielten und Hamermesh (2013) betrachtet das mittlere Alter bei Publikation von Autoren in drei der renommiertesten Zeitschriften der Volkswirtschaftslehre. Der Ansatz das mittlere Alter beim ersten wissenschaftlichen Artikel eines Autors zu bemessen ist neu und wurde in der Form bislang lediglich von Jones (2009a) für Patentanmelder

durchgeführt. Auch die Ergebnisse zur Spezialisierung von Autoren liefern einen neuen Beitrag zur bestehenden Literatur. Verwandte Untersuchungen finden sich auch für dieses Maß bislang im Kontext von Patentanmeldungen (Jones 2009a; Melero und Palomeras 2015).

Ein Beitrag der sich in dieser Form nicht in der bestehenden Literatur findet, ist ferner die Entwicklung eines konzeptionellen Rahmenwerks (vgl. Kapitel 5), das die Einordnung von Ansätzen zum Umgang mit und zur Lösung von komplexen Problemen im Allgemeinen bzw. zum Umgang mit der Last des Wissens im Speziellen ermöglicht. Es basiert auf drei Prinzipien, die helfen, die grundlegenden Mechanismen hinter den verwendeten Techniken und/oder Methoden zu verstehen, und die beobachteten empirischen Ergebnisse erklären können.

Kapitel 3

Daten, Maße und Kontrollvariablen

Die dieser Arbeit zugrunde liegenden Daten stammen aus verschiedenen Quellen und wurden auf unterschiedliche Weise erhoben. Dabei handelt es sich einerseits um Publikationsdaten verschiedener wissenschaftlicher Disziplinen sowie andererseits um individuelle Informationen, wie das Alter oder das Geschlecht der Autoren der erhobenen Artikel. Im Folgenden wird näher auf die Herkunft der Daten bzw. die Methoden der Datenerhebung, die Anreicherung dieser Daten mit weiteren Informationen und den Datenumfang eingegangen. Ferner werden die in dieser Arbeit untersuchten Maße sowie die Kontrollvariablen bei den Regressionen erläutert. Die Publikationsdaten stammen aus den Literaturdatenbanken JSTOR und EconLit.¹ Im Folgenden wird auf die Publikationsdaten von JSTOR ohne die Publikationen aus der Zeitschrift *Science* auch mit „JSTOR-Datensatz“ und auf die Daten von EconLit mit „EconLit-Datensatz“ Bezug genommen. In beiden wurden auf semi-automatisiertem Weg Metadaten wissenschaftlicher Artikel zu ausgewählten Wissenschaftsdisziplinen und zugehörigen Zeitschriften gesammelt (vgl. Abschnitt 3.2.1).

Daten, die von JSTOR stammen, umfassen 20 der renommiertesten Zeitschriften der *Volkswirtschaftslehre*, renommierte Zeitschriften aus den Fachbereichen *Betriebswirtschaftslehre und Finance* (16 Zeitschriften), *Geographie und Geologie* (4), *Mathematik* (12), *Psychologie* (8), sowie Publikationen der Zeitschrift *Science*.² In den Abschnitten 3.1 bis 3.3 finden sich nähere Informationen zu den Datensätzen (z. B. bzgl. des Umfangs) und in Kapitel 4 die Auswertung der Daten im Hinblick auf die Forschungsfragen dieser Arbeit. Die 20 volkswirtschaftlichen Zeitschriften aus dem JSTOR-Datensatz,

¹Siehe <http://www.jstor.org/> und <https://www.aeaweb.org/econlit/>.

²Genaugenommen stammen die Daten zu Artikeln der Zeitschrift *Science* von *JSTOR for science*, wie später in Abschnitt 3.3 (S. 62) erläutert wird.

ergänzen den EconLit-Datensatz punktuell, der viele weitere Zeitschriften der Volkswirtschaftslehre umfasst.

Bei der Datenerhebung wurden Webscraper auf Python-Basis eingesetzt. Diese Skripte besuchten die entsprechenden Seiten der obigen Literaturdatenbanken und extrahierten aus den HTML-Dokumenten die notwendigen Informationen. In Abschnitt 3.2.1 wird beispielhaft für den JSTOR-Datensatz dargestellt, wie die Informationen anhand eines solchen Skripts erhoben wurden. Der EconLit-Datensatz wurde von Sascha Schweitzer im Rahmen eines gemeinsamen Projekts mit dem Autor der vorliegenden Arbeit aufbereitet und zur Verfügung gestellt. Nähere Informationen zu diesem Datensatz finden sich im folgenden Abschnitt 3.1.

3.1 Der EconLit-Datensatz

3.1.1 Herkunft und Umfang

Zentrale Datenquelle für die volkswirtschaftlichen Artikel ist die Publikationsdatenbank EconLit. In dieser Datenbank finden sich Metadaten zu wissenschaftlichen Artikeln aus volkswirtschaftlichen Zeitschriften seit 1886. Der EconLit-Datensatz umfasst 762.826 Artikel mit Informationen zum Publikationsjahr, dem Namen der Zeitschrift, dem Titel des Artikels, den zugehörigen JEL-Codes und den Namen der Autoren.³ Diese Daten wurden ferner mit Informationen des *Web of Science* und *Business Source Premier* ergänzt,⁴ beispielsweise hinsichtlich der Institutszugehörigkeiten der einzelnen Autoren und der Anzahl der Quellenangaben eines Artikels.⁵ Der EconLit-Datensatz ist ferner auf Artikel von Autoren limitiert, die ihren ersten Artikel frühestens im Jahr 1970 in einer volkswirtschaftlichen Zeitschrift veröffentlicht haben. Hierbei wurden alle Autoren und deren Artikel ab 1970 ausgeschlossen, die bereits mit einem Artikel zwischen 1950 und 1969 in EconLit vertreten sind. Eine Verteilung der Artikel aus dem EconLit-Datensatz auf die einzelnen Jahre des Untersuchungszeitraums findet sich in Abbildung 3.1. Die gestrichelte Linie markiert die Anzahl der Artikel, die ursprünglich erhoben wurden, d. h. vor dem Ausschluss der Artikel von Autoren, die bereits zwischen

³Vgl. Abschnitt 3.1.2 ab S. 43 für eine Erläuterung zu JEL-Codes. Informationen zu den JEL-Codes sind nicht für alle Artikel verfügbar, wie im selben Abschnitt ausgeführt wird.

⁴Siehe www.webofknowledge.com und <https://www.ebsco.com/products/research-databases/business-source-premier>.

⁵Institutszugehörigkeiten sind für 85,4 % der Autoren im gesamten EconLit-Datensatz verfügbar und für 68,8 % der Top-Autoren (vgl. unten), während die Anzahl der Quellenangaben für 27,8 % aller Artikel und für 86,6 % der Top-Artikel (vgl. unten) bekannt ist.

1950 und 1969 mit einem Artikel vertreten sind. Ähnliche Untersuchungszeiträume finden sich auch in verwandten Studien zu volkswirtschaftlichen Veröffentlichungen wie Card und DellaVigna (2013), Hamermesh (2013) oder Kosnik (2015).

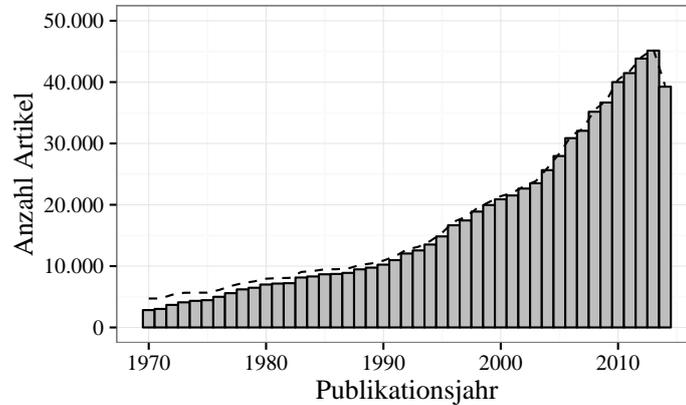


Abbildung 3.1: Anzahl der Artikel pro Jahr im EconLit-Datensatz. Die gestrichelte Linie gibt die Anzahl ursprünglich erfasster Artikel wieder (vgl. Erläuterung im Text).

Die Tradition volkswirtschaftlicher Zeitschriften reicht ins 19. Jahrhundert zurück und deren Landschaft unterlag seit ihren Anfängen unterschiedlichen Veränderungen. Beispielsweise hat die Anzahl volkswirtschaftlicher Zeitschriften in den Jahren nach 1970 stark zugenommen (vgl. z. B. Abbildung A.1, S. 199, für die Anzahl der Zeitschriften pro Jahr im EconLit-Datensatz). Ein Grund für diese Zunahme ist, dass Online-Zeitschriften die Kosten für die Gründung und Unterhaltung einer Zeitschrift gesenkt (Odlyzko 1997; Van Noorden 2013) haben. Ferner sind Zeitschriften hinzugekommen, die spezielle Teilgebiete der Volkswirtschaftslehre abdecken und ein kleineres Publikum ansprechen. In diesem Sinne sind volkswirtschaftliche Zeitschriften heterogener geworden. Wie Larsen und Ins (2010) anmerken, erschwert die sich ändernde Zeitschriftenlandschaft bibliometrische Trendanalysen, da es keine perfekte Lösung gibt, für alle Effekte die damit einhergehen zu kontrollieren. Beispielsweise nimmt der Wettbewerb um Veröffentlichungen in den renommiertesten Zeitschriften zu, da diese einen zunehmend kleinen Teil eines wachsenden Marktes ausmachen (Mabe 2003). Aber auch das Einbeziehen im Zeitverlauf hinzukommender Zeitschriften ist mit Problemen behaftet, da diese in den späteren Jahren des Beobachtungszeitraums gegebenenfalls zu Effekten führen, die in den ersten Jahren nicht wirkten.

Betrachtet man zur vorliegenden Arbeit verwandte Studien, die sich mit wissenschaftlicher Literatur beschäftigen, so gehen diese in unterschiedlicher Weise mit der unterschiedlichen Qualität von Zeitschriften bzw. Artikeln um. Beispielsweise schränken Laband und Tollison (2000), Ellison (2002), Laband, Tollison und Karahan (2002),

Ellison (2011), Card und DellaVigna (2013), sowie Hamermesh (2013) ihre Untersuchungen auf besonders renommierte Zeitschriften in ihren Bereichen ein. Andere Studien, wie Conley, Crucini u. a. (2013), gewichten Publikationen mit einem Qualitätsindex oder nehmen eine aggregierte Sichtweise ein, in der nicht für einzelne Zeitschriften bzw. deren Qualität kontrolliert wird (Bornmann und Mutz 2015). Ein Vorteil der Beschränkung auf renommierte Zeitschriften ist, dass die Stichprobe leicht zu handhaben und die Auswahl der Zeitschriften über einen längeren Zeitraum konstant bleibt, während die Kontrolle mittels eines Qualitätsindex ein vollständigeres Bild des gesamten Publikationsspektrums ermöglicht.

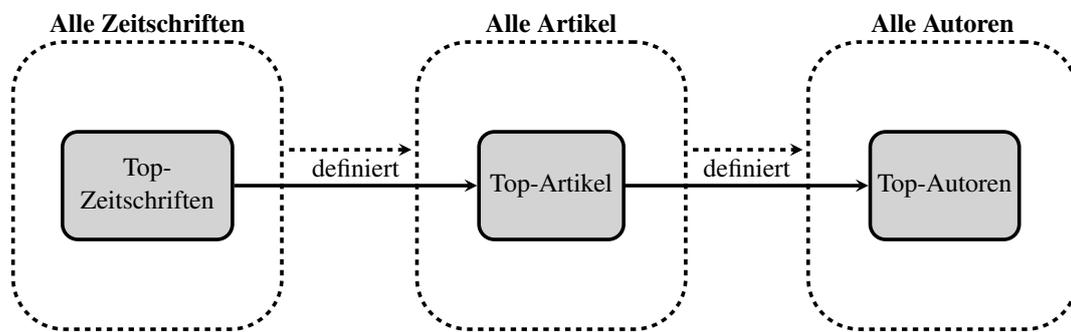


Abbildung 3.2: Definition von Top-Artikeln und Top-Autoren im EconLit-Datensatz.

Auch wenn der Fokus dieser Arbeit nicht auf dem Aspekt der sich ändernden Landschaft volkswirtschaftlicher Zeitschriften liegt, sollen diese Effekte zumindest ansatzweise berücksichtigt werden. Hierzu wird einerseits in den Regressionen für die einzelnen Zeitschriften kontrolliert. Andererseits, ist die Untersuchung der volkswirtschaftlichen Artikel zweigeteilt. Im einen Teil der Untersuchung werden alle Zeitschriften, Artikel und Autoren aus dem EconLit-Datensatz berücksichtigt. Der andere Teil beschränkt sich auf Artikel und Autoren aus neun der renommiertesten, begutachteten Zeitschriften der Volkswirtschaftslehre (i. F. auch als *Top-Zeitschriften* bezeichnet).⁶ Die Top-Zeitschriften sind *American Economic Review*, *Econometrica*, *International Economic Review*, *Journal of Political Economy*, *Economic Journal*, *Journal of Finance*, *Quarterly Journal of Economics*, *Review of Economic Studies* und *Review of Economics and Statistics* (vgl. auch Tabelle 3.1). Bei der Auswahl dieser neun Zeitschriften wurden verschiedene Rankings wie Ritzberger (2008), Engemann und Wall (2009), Kodrzycki und Yu (2006) sowie www.journal-ranking.com zu Rate gezogen. Die gewählten Zeitschriften werden ferner seit mindestens 50 Jahren herausgegeben und erlauben somit die kontinuierliche Untersuchung von Forschungsartikeln hoher Qualität über

⁶Vgl. Kapitel 6 und speziell Abschnitt 6.2 (S. 174) für eine kritische Diskussion der Probleme und Verzerrungen einer fixen Auswahl von besonders renommierten Zeitschriften.

einen längeren Zeitraum. Auch andere Studien, wie Kim, Morse und Zingales (2006), Hamermesh (2013) oder Card und DellaVigna (2013) beziehen sich auf eine ähnliche Auswahl der renommiertesten Zeitschriften.

Tabelle 3.1: Top-Zeitschriften im EconLit-Datensatz.

Zeitschrift	Publikations- historie	Anzahl Artikel	Anzahl Autoren
American Economic Review	1911–2018	7.503	6.566
Econometrica	1933–2018	2.973	2.569
Economic Journal	1960–2018	2.560	2.941
International Economic Review	1892–2018	2.256	2.420
Journal of Finance	1891–2018	3.865	3.421
Journal of Political Economy	1946–2018	2.313	2.383
Quarterly Journal of Economics	1886–2018	1.958	2.097
Review of Economic Studies	1933–2018	1.938	2.046
Review of Economics and Statistics	1946–2018	3.493	4.153

Bem.: Die Anzahl der Artikel und die Anzahl der Autoren beziehen sich nicht auf den Zeitraum der Publikationshistorie, sondern auf den erhobenen und letztlich untersuchten Datensatz von Artikeln aus den Jahren 1970–2014.

Artikel, die in einer der neun Top-Zeitschriften erschienen sind, werden im Folgenden auch als *Top-Artikel* bezeichnet. Diese wiederum definieren *Top-Autoren*. Hierzu gehören Autoren, die mindestens einen Top-Artikel veröffentlicht haben. Die beschriebenen Beziehungen zwischen den Zeitschriften, zugehörigen Artikeln und Autoren sind in Abbildung 3.2 dargestellt. Der EconLit-Datensatz der Top-Artikel umfasst 28.859 Artikel im Zeitraum von 1970–2014. Die Top-Zeitschriften und Informationen zum Zeitraum, in dem sie herausgegeben wurden (Publikationshistorie), sowie zur Anzahl Artikel und Top-Autoren für jede dieser Zeitschriften im EconLit-Datensatz finden sich in Tabelle 3.1. Die Verteilung der Top-Artikel auf die einzelnen Jahre des Untersuchungszeitraums ist in Abbildung 3.3 dargestellt. Es ist deutlich zu sehen, dass die Verteilung der Artikel hier gleichmäßiger ist, als für den gesamten EconLit-Datensatz (vgl. Abbildung 3.1), da über den gesamten Zeitraum eine konstante Anzahl von neun Zeitschriften betrachtet wird.

3.1.2 JEL-Felder

Bereits zu Beginn des 20. Jahrhunderts, gab es in der Volkswirtschaftslehre Bestrebungen die eigene Disziplin, speziell Publikationen und Wissenschaftler, zu klassifizieren

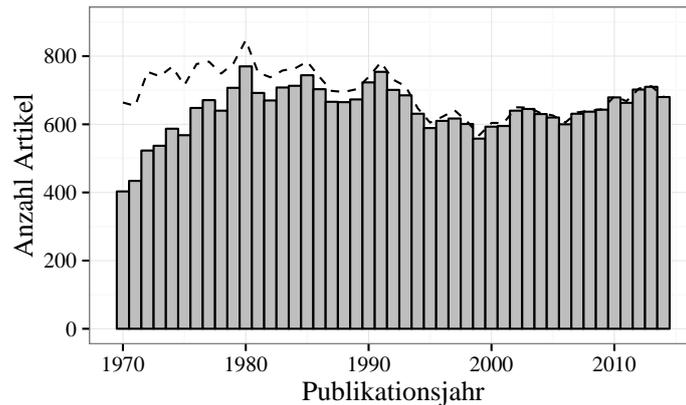


Abbildung 3.3: Anzahl der Top-Artikel pro Jahr im EconLit-Datensatz. Die gestrichelte Linie gibt die Anzahl ursprünglich erfasster Artikel wieder (vgl. Erläuterung im Text).

und damit konkreten Fachbereichen zuzuordnen (Cherrier 2017, S. 548). Das *JEL-Klassifikationssystem* dient der Zuordnung von Artikeln zu spezifischen Fachbereichen der Volkswirtschaftslehre. Das bestehende System wurde im Jahr 1991 vom *Journal of Economic Literature* (JEL) eingeführt (Cherrier 2017); daher auch der Name. Es basiert auf Codes aus Zahlen und Buchstaben, den sogenannten *JEL-Codes*, die jeweils für einen bestimmten Teilbereich der Volkswirtschaftslehre stehen.⁷ Die Codes des bestehenden Systems sind dreistellig und beginnen mit einem Buchstaben, dem eine zweistellige Zahl folgt (z. B. D47 für *Market Design* oder H41 für *Public Goods*).

Card und DellaVigna (2013) ordnen die einzelnen JEL-Codes dreizehn größeren volkswirtschaftlichen Teilbereichen zu.⁸ Diese Zuordnung wird auch in der vorliegenden Arbeit verwendet und Artikel im EconLit-Datensatz entsprechend ihren JEL-Codes den gelisteten *JEL-Feldern* zugeordnet (vgl. Tabelle 3.2). Der Untersuchungszeitraum für Artikel im EconLit-Datensatz reicht von 1970–2014. Wie oben erwähnt, wurde das bestehende JEL-Klassifikationssystem 1991 eingeführt. Entsprechend gibt es im Datensatz Artikel, die anhand von JEL-Codes nach dem „neuen“ System ab 1991 klassifiziert wurden, aber auch solche, die noch nach dem „alten“ System, das seit 1967 bestand hatte (Cherrier 2017), den JEL-Feldern zugeordnet wurden. Daher wurden für die Auswertung „alte“ JEL-Codes, nach dem Vorbild von Card und DellaVigna (2012, S. 23), in neue JEL-Codes übersetzt und den entsprechenden JEL-Feldern zugeordnet. Der Übergang vom alten in das neue System im Jahr 1991 hat allerdings keinen systematischen Einfluss auf die Ergebnisse in Abschnitt 4.5 (S. 135 ff.), wie dort

⁷Eine Liste der JEL-Codes findet sich unter <https://www.aeaweb.org/econlit/jelCodes.php>.

⁸Die konkrete Zuordnung findet sich im zugehörigen elektronischen Anhang (Card und DellaVigna 2012, S. 22–23).

auch nochmals ausgeführt wird. Ist in Tabelle 3.2 für die neuen JEL-Codes lediglich ein Buchstabe oder ein Buchstabe und eine Ziffer angegeben, so schließt das JEL-Feld alle weiteren, spezifischeren JEL-Codes aus dieser Kategorie ein. Beispielsweise schließt die Angabe „F“ somit auch die JEL-Codes „F00“, „F01“, „F43“, etc. ein. Die Felder selbst sind dabei überschneidungsfrei, ein Artikel kann allerdings, abhängig von den für ihn angegebenen JEL-Codes, mehreren JEL-Feldern zugeordnet sein. Die Anzahl der JEL-Felder, denen ein Artikel zugeordnet ist, kann jedoch maximal der Anzahl JEL-Codes entsprechen.

Tabelle 3.2: JEL-Felder und zugehörige JEL-Codes nach altem und neuem System (vgl. Card und DellaVigna 2012).

JEL-Feld (engl. Entsprechung)	neue JEL-Codes (ab 1991)	alte JEL-Codes (1967–1990)
Arbeitsmarktökonomik (<i>labor</i>)	J, I2	811-813, 821-826, 831-833, 841, 851, 912, 917, 918
Außenwirtschaft (<i>international</i>)	F	111, 400, 411, 421-423, 431-433, 441-443
Entwicklungsökonomik (<i>development</i>)	O	621
Experimentalökonomik (<i>lab-based experiments</i>)	C9	215
Finance (<i>finance</i>)	G	310, 312-315, 521
Geschichte (<i>history</i>)	N	041-048
Gesundheits- und Stadtökonomik (<i>health and urban economics</i>)	I0, I1, R, K	731, 913, 916, 931-933, 941
Industrieökonomik (<i>industrial organization</i>)	L	514, 611-616, 619, 631-636
Makroökonomie (<i>macroeconomics</i>)	E, O11, O4, O5	023, 112, 120-124, 131-134, 221, 223, 226, 311

Tabelle 3.2: (Fortsetzung)

JEL-Feld (engl. Entsprechung)	neue JEL-Codes (ab 1991)	alte JEL-Codes (1967–1990)
Mikroökonomie (<i>microeconomics</i>)	D außer D11, D5, D21, D85 und D86	022, 024, 025, 114, 224, 511-513, 522, 921
Öffentliche Wirtschaft (<i>public economics</i>)	H	320-325, 641, 915
Ökonometrie (<i>econometrics</i>)	C0-C5, C6, C8	211-214, 220, 222, 229
Theorie (<i>theory</i>)	C7, D11, D5, D21, D85, D86	021, 026
Andere (<i>other</i>)	A, B, I3, M, P, Q, Y, Z	011, 012, 027, 031, 036, 050-053, 113, 531, 541, 710, 711, 713-718, 721- 723, 911, 914

Artikeln (Top-Artikeln) im EconLit-Datensatz sind zu 4,8 % (5,7 %) keine JEL-Codes zugeordnet, zu 24,3 % (35,5 %) einer, zu 27,7 % (32,6 %) zwei, zu 19,7 % (15,7 %) drei und zu 23,5 % (10,5 %) zwischen vier und sieben.⁹ Durchschnittlich sind für alle Artikel im EconLit-Datensatz 2,5 und für Top-Artikel 2,0 JEL-Codes angegeben. Wie oben erläutert, werden die JEL-Codes einzelnen JEL-Feldern zugeordnet. Entsprechend sind allen Artikeln (Top-Artikeln) im EconLit-Datensatz zu 40,1 % (52,8 %) ein JEL-Feld, zu 35,2 % (31,5 %) zwei, zu 15,3 % (8,2 %) drei und zu 9,4 % (7,5 %) vier bis sechs JEL-Felder zugeordnet. Im Mittel ist ein Artikel 1,8 JEL-Feldern und ein Top-Artikel 1,6 JEL-Feldern zugeordnet.

⁹Gibt es für Artikel keine Information bzgl. der JEL-Codes, so fehlt diese Information typischerweise in EconLit. Alle Top-Artikel ohne Angabe eines JEL-Codes stammen aus den Jahren vor 1991, d. h. aus der Zeit vor der letzten Reform des JEL-Klassifikationssystems.

3.2 Der JSTOR-Datensatz

Im Zeitraum von März bis August 2015 wurden über die Publikationsdatenbank JSTOR Informationen zu Artikeln aus den wissenschaftlichen Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik*, *Psychologie* und *Volkswirtschaftslehre* erhoben. Hierzu gehören die folgenden Metadaten:

- Name der Zeitschrift,
- Band und Ausgabe,
- Publikationsjahr,
- Titel des Beitrags,
- Abstract (wenn vorhanden),
- Vor-, Mittel- und Nachnamen der Autoren,
- Seiten (z. B. „43–85“) und
- Anzahl der Quellenangaben.

Im Folgenden Abschnitt 3.2.1 wird die Funktionsweise des verwendeten Python-basierten Webscrapers und entsprechend das Vorgehen bei der Datenerhebung erläutert. Alle Ausführungen, z. B. Angaben konkreter URLs oder Beschreibungen der Navigation auf den Webseiten von JSTOR, beziehen sich auf den Zeitpunkt der Datenerhebung bzw. der Erstellung dieser Arbeit. In der Zwischenzeit können sich die URLs oder Strukturen der HTML-Dokumente bei JSTOR teilweise geändert haben, sodass der verwendete und beschriebene Webscraper gegebenenfalls nicht mehr ohne Anpassungen lauffähig ist. Bei den Abbildungen in diesem Abschnitt handelt es sich zum größten Teil um Screenshots relevanter Ausschnitte der entsprechenden Webseiten.

3.2.1 Erhebung

In der Publikationsdatenbank JSTOR lassen sich wissenschaftliche Publikationen themenspezifisch durchsuchen. Auf die entsprechende Übersicht gelangt man entweder direkt über die URL www.jstor.org/subjects oder indem man auf der Startseite www.jstor.org im Reiter „Browse“ auf „by Subject“ klickt (vgl. Abbildung 3.4).¹⁰ In der dann folgenden Übersicht finden sich Themenbereiche wie *Kunst* (engl. *Arts*), *Geisteswissenschaften* (engl. *Humanities*), *Naturwissenschaften und Mathematik* (engl. *Science & Mathematics*) oder *Sozialwissenschaften* (engl. *Social Sciences*), die wiederum weiter in spezifischere Themengebiete aufgegliedert sind, wie in Abbildung 3.5 für

¹⁰Die Ellipse in Abbildung 3.4 wurde vom Autor nachträglich zur Hervorhebung eingefügt.

den Themenbereich *Betriebs- und Volkswirtschaftslehre* (engl. *Business & Economics*) zu sehen ist.

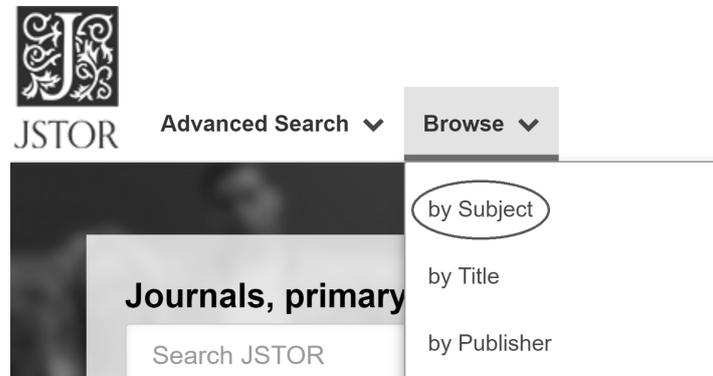


Abbildung 3.4: Themenspezifische Suche nach Publikationen bei JSTOR.



Abbildung 3.5: Themengebiete im Themenbereich *Business & Economics* bei JSTOR.

Hinter jedem der spezifischeren Themengebiete findet sich eine Liste mit Zeitschriften, die diesem Gebiet bei JSTOR zugeordnet sind. Beispielsweise umfasst das Themengebiet „Business“ zum Zeitpunkt der Erstellung dieser Arbeit 208 und das Themengebiet „Economics“ 180 Zeitschriften, wobei diese Listen nicht immer überschneidungsfrei sind. So wird zum Beispiel die Zeitschrift *American Economic Review* sowohl dem Themengebiet „Business“, als auch dem Themengebiet „Economics“ und die Zeitschrift *Econometrica* sowohl „Economics“ als auch „Mathematics“ zugeordnet. Hinter jedem Namen der hier gelisteten Zeitschriften verbirgt sich wiederum ein weiterführender Link auf eine Übersichtsseite mit einer kurzen Information zum Umfang der jeweiligen Zeitschrift bei JSTOR und weiterführenden Links (vgl. Abbildung 3.6), sowie allen bei JSTOR verfügbaren Ausgaben der Zeitschrift, wobei diese nach Bänden bzw. Dekaden aufgeführt sind (vgl. Abbildung 3.7). Im Folgenden dient die Zeitschrift *Econometrica* als illustratives Beispiel. Nach dem Laden der entsprechenden Übersichtsseite ist zunächst lediglich die aktuelle Dekade „aufgeklappt“, so wie es in Abbildung 3.7 auf der linken Seite zu sehen ist, während die übrigen Dekaden noch zugeklappt sind, wie man in derselben Abbildung auf der rechten Seite sieht.



Abbildung 3.6: Information zum Umfang der Zeitschrift *Econometrica* bei JSTOR.

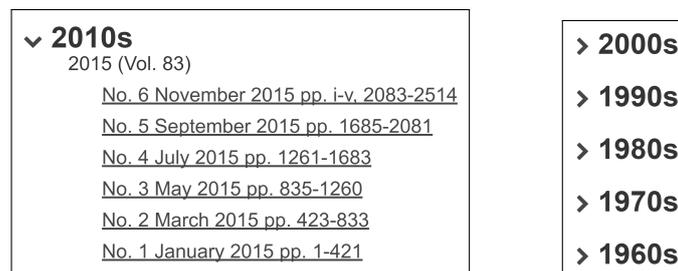


Abbildung 3.7: Ausschnitt der Anordnung von Ausgaben der Zeitschrift *Econometrica* nach Bänden und Dekaden bei JSTOR.

Diese Übersichtsseite ist der Ausgangspunkt für den Webscraper.¹¹ Die entsprechende URL (im Beispiel www.jstor.org/journal/econometrica) und die zugehörige älteste bei JSTOR gelistete Dekade (d. h. im Beispiel *Econometrica* die Dekade der 1930er Jahre) werden im Skript manuell eingetragen.¹² Nach dem Start öffnet das Skript die URL in einem Browser (in der konkreten Implementierung ist dies *Google Chrome*).¹³ Anschließend durchsucht das Skript alle Ausgaben und die entsprechenden Beiträge der jeweiligen Zeitschrift, angefangen bei der ersten Ausgabe, das heißt der ersten Ausgabe im ersten Band der ältesten Dekade. Im Beispiel wäre dies also die Ausgabe „No. 1“ aus dem Jahr 1933. Der Webscraper navigiert dabei auf den Seiten unter Nutzung von XPath (XML Path Language), einer Abfragesprache, um Elemente (z. B. einen Hyperlink) oder Teile eines XML-Dokumentes (z. B. Liste in einem HTML-Dokument) zu adres-

¹¹Die Steuerung des Webscrapers bzw. die Befehle zu dessen Navigation im Webbrowser (z. B. Anklicken eines Links), entstammen dem Python-Paket *Selenium*.

¹²An dieser Stelle muss nicht zwingend die älteste Dekade angegeben werden. Möchte man beispielsweise lediglich Artikel ab 1970 erheben, kann man als „Startdekade“ für den Webscraper auch diese angeben.

¹³Genau genommen öffnet das Skript ein Browserfenster, trägt die URL in die Adresszeile ein und öffnet die zugehörige Seite anschließend.

sieren.¹⁴ In Abschnitt A.2.2 (S. 203 f. im Anhang) findet sich eine Erläuterung dazu, wie die XPath-Ausdrücke einzelner Seiten-Elemente bestimmt wurden. Im Zeitraum der Datenerhebung wurde der Ausdruck `//*[@id="1960s"]/a` verwendet, um die Dekade der 1960er Jahre anzusprechen.¹⁵ Ersetzt man in diesem Ausdruck 1960 durch 1970, so spricht man die Dekade der 1970er Jahre an. Die Steuerung des Webscrapers ermöglicht es ferner, Elemente einer Webseite nicht nur zu lokalisieren, sondern beispielsweise auch, einem Hyperlink zu folgen oder Text in Eingabefelder einzugeben.

Hinter jeder der in einer Dekade gelisteten Ausgaben verbirgt sich wiederum ein weiterführender Link, der auf eine Auflistung aller Beiträge in dieser Ausgabe führt. Die Links zu den einzelnen Ausgaben innerhalb einer Dekade wurden beispielsweise für die 1940er Jahre über den XPath-Ausdruck `//*[@id="C1940s"]/div[j]/ul/li[i]/div/span/a` adressiert, wobei das `j` in `[j]` bzw. das `i` in `[i]` für den `j`-ten Band in dieser Dekade bzw. die `i`-te Ausgabe stehen. In Abbildung 3.8 ist beispielsweise der Beginn der Liste aller Beiträge für die erste Ausgabe des 83. Bandes von *Econometrica* aus dem Jahr 2015 abgebildet. Wie man sieht, gehören zu den gelisteten Beiträgen nicht nur wissenschaftliche Artikel, sondern auch andere Beiträge, wie zum Beispiel das Deckblatt (engl. *Front Matter*). Das Skript durchläuft auch in dieser Liste wiederum alle Einträge, ungeachtet der Art des Beitrags, das heißt zunächst werden auch Kommentare der Redaktion, Konferenzankündigungen, Listen neu erschienener Bücher oder Nachrufe erhoben. Die Daten wurden nach der Erhebung um diese nicht-wissenschaftlichen Beiträge bereinigt und der Datensatz so auf wissenschaftliche Artikel reduziert.

Die Titel der einzelnen Beiträge führen wiederum auf eine weitere Seite, auf der detailliertere Informationen zu jedem Beitrag ausgewiesen werden.¹⁶ Diese Seite ist die Stelle, an der die Metadaten für jeden Beitrag erhoben werden. Hier sind die interessie-

¹⁴XML steht für *Extensible Markup Language* und ist eine Auszeichnungssprache, mit der sich hierarchisch strukturierte Daten als Textdatei darstellen lassen. Eine Auszeichnungssprache wiederum ist eine maschinenlesbare Sprache, um Texte oder andere Daten zu gliedern und zu formatieren. Ein bekanntes Beispiel für eine solche Sprache ist HTML (Hypertext Markup Language) zur Gestaltung von Webseiten.

¹⁵Wenn Webseiten überarbeitet werden, ist es möglich, dass sich auch die Strukturen der HTML-Dokumente ändern und Elemente der Seite mit anderen XPath-Ausdrücken angesprochen werden müssen. Beispielsweise müsste zum Zeitpunkt des Schreibens dieser Arbeit der Ausdruck `//*[@id="content"]/div[1]/div[2]/div[3]/div/div/dl/dl[6]/dt/a` verwendet werden, um die Dekaden anzusprechen. In diesem Ausdruck spricht der Teil `dl[6]` die Dekade der 1960er Jahre an, während `dl[5]` bzw. `dl[7]` die Dekaden der 1970er bzw. 1950er Jahre ansprechen würden. Für den Start des bei der Datenerhebung verwendeten Skripts würde daher nicht mehr die älteste Dekade, sondern die Anzahl Dekaden angegeben werden.

¹⁶Die einzelnen Titel lassen sich über den XPath-Ausdruck `//*[@id="results"]/ol/li[n]/div/div[2]/div[2]/span/a` adressieren, wobei `li[n]` für den `n`-ten Beitrag (von oben gezählt) in dieser Auflistung steht.

<input type="checkbox"/>	Front Matter Stable URL: http://www.jstor.org/stable/43616920 Read Online Download PDF Add to My Lists Cite this Item
<input type="checkbox"/>	ESTIMATION OF NONPARAMETRIC MODELS WITH SIMULTANEITY (pp. 1-66) Rosa L. Matzkin Stable URL: http://www.jstor.org/stable/43616921 Read Online Download PDF Add to My Lists Cite this Item
<input type="checkbox"/>	NECESSITY IS THE MOTHER OF INVENTION: INPUT SUPPLIES AND DIRECTED TECHNICAL CHANGE (pp. 67-100) W. Walker Hanlon Stable URL: http://www.jstor.org/stable/43616922 Read Online Download PDF Add to My Lists Cite this Item

Abbildung 3.8: Erste Beiträge in der ersten Ausgabe des 83. Bandes von *Econometrica*.

renden Informationen an unterschiedlichen Stellen zu finden. Im oberen Bereich der Seite gibt es eine Übersicht mit grundlegenden Informationen zum Titel des Beitrags, zu den Autoren, den Seitenangaben und zu der Anzahl Seiten, wie beispielhaft in Abbildung 3.9 zu sehen ist. Wurde ein Artikel von mehreren Autoren geschrieben, so sind alle Autoren gelistet und deren Namen durch Kommata bzw. das Wort „and“ voneinander getrennt. Abbildungen A.3 und A.4 (S. 202 f.) im Anhang zeigen ein Beispiel einer ganzen Seite.

<p>CONSUMER HETEROGENEITY AND PAID SEARCH EFFECTIVENESS: A LARGE-SCALE FIELD EXPERIMENT</p> <p>Thomas Blake, Chris Nosko and Steven Tadelis <i>Econometrica</i> Vol. 83, No. 1 (January 2015), pp. 155-174</p> <p>Published by: The Econometric Society Stable URL: http://www.jstor.org/stable/43616924 Page Count: 20</p>

Abbildung 3.9: Erste Metadaten zu einem Beitrag bei JSTOR.

Abbildung 3.10 zeigt den HTML-Code des in Abbildung 3.9 abgebildeten Ausschnitts. Wie man sieht, sind die einzelnen Informationen im HTML-Dokument gut anhand entsprechender Tags strukturiert.¹⁷ So sind der Titel des Beitrags als Überschrift deklariert und beispielsweise die Autoren, der Name der Zeitschrift oder die Seitenzahl in entsprechend benannten Tags ("class=contrib", "class=journal", "class=count") ausgewiesen. Diese Elemente lassen sich wiederum durch entsprechende XPath-Ausdrücke adressieren und die Texte dieser Elemente extrahieren. Teil-

¹⁷HTML-Tags sind Strukturelemente eines HTML-Dokumentes. Beispielsweise markieren die Tags <h1> bzw. </h1> den Anfang bzw. das Ende einer Überschrift, während bzw. den Beginn und das Ende einer nicht-nummerierten Auflistung markieren.

weise benötigen die extrahierten Texte im Python-Skript weiterer Bearbeitung. So muss beispielsweise bei der Seitenzahl die Zeichenkette "Page Count: " entfernt werden und die Zeichenkette, in der die Autoren stehen, muss mit Hilfe von String-Operatoren an den Stellen mit einem Komma bzw. dem Wort "and" getrennt werden, sodass man letztlich die Namen der einzelnen Autoren erhält. Am Ende der URL eines jeden Beitrags findet sich eine mehrstellige Zahl. Diese ist für jeden Beitrag eindeutig und wurde daher als Identifikationsnummer für die einzelnen Beiträge erhoben. Hierzu wurde die aktuelle URL abgefragt und der Teil "http://www.jstor.org/stable/" von der Zeichenkette entfernt.

```

Elements Console Sources Network Performance >> 1 3
▼ <h1 class="title medium-heading">
  "
  CONSUMER HETEROGENEITY AND PAID SEARCH
  EFFECTIVENESS: A LARGE-SCALE FIELD EXPERIMENT
  "
</h1>
<div class="contrib">
  Thomas Blake, Chris Nosko and Steven Tadelis
</div>
▶ <div class="journal">...</div>
<div class="src mbl">
  Vol. 83, No. 1 (January 2015), pp. 155-174
</div>
▶ <div class="publisher">...</div>
<div class="stable">
  Stable URL: http://www.jstor.org/stable/43616924
</div>
<div class="count">
  Page Count: 20
</div>
▶ <div class="topics-list mtl">...</div>
▶ <div class="mtl mbl left">...</div>

```

Abbildung 3.10: HTML-Code zu den Metadaten aus Abbildung 3.9.

Auf jeder Seite, die zu einem spezifischen Beitrag gehört, findet sich etwas unterhalb der Informationen, die in Abbildung 3.9 beispielhaft dargestellt sind, ein Menü mit drei Reitern (vgl. Abbildung 3.11), von denen zu Beginn der erste („Article“) aktiv ist. Bei der Auswahl dieses Reiters lassen sich die einzelnen Seiten des Beitrags betrachten, wobei man von der aktuell betrachteten Seite jeweils nur zur vorherigen oder nächsten Seite wechseln kann. Hinter dem Reiter „Thumbnails“ versteckt sich eine Übersicht mit Miniaturbildern der einzelnen Seiten des Beitrags, über die man schneller zu spezifischen Seiten gelangt, wenn man das entsprechende Miniaturbild anklickt.

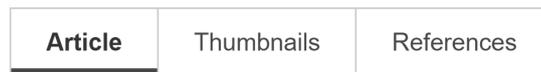


Abbildung 3.11: Menü mit weiteren Informationen zu einem Artikel bei JSTOR.

Hinter dem Reiter „References“ wiederum verbergen sich weitere Metainformationen zum Beitrag.¹⁸ Hierzu gehören, wenn vorhanden, der Abstract, eine Angabe zur Anzahl der Verweise, eine Liste mit Fußnoten und eine Liste der Quellen aus dem Literaturverzeichnis. Die Anzahl der Verweise wird dabei in dem Satz „This item contains ... references.“ angegeben; im Beispiel also 42 (vgl. Abbildung 3.12). Sie entspricht der Summe aus der Anzahl der Quellenangaben im Literaturverzeichnis und der Anzahl an dieser Stelle gelisteter Fußnoten. Die an dieser Stelle gelisteten Fußnoten entsprechen nicht allen Fußnoten des Beitrags, sondern nur solchen mit Quellenverweisen. Ausgenommen sind beispielsweise Fußnoten mit Danksagungen oder Begriffserläuterungen. Dabei enthalten die Fußnoten teilweise auch Verweise auf Quellen aus dem Literaturverzeichnis (vgl. Fußnote „3“ in Abbildung 3.12 und den dritten Eintrag im Literaturverzeichnis in Abbildung 3.13). Somit besteht zwischen den hier gelisteten Fußnoten und den Einträgen im Literaturverzeichnis eine gewisse Schnittmenge, weshalb die Anzahl Verweise tatsächlich geringer sein kann. Somit kann diese Angabe zur Anzahl Verweise nicht als Wert für die Anzahl der Quellenangaben im Artikel verwendet werden. Gleichzeitig bedeutet dies auch, dass nicht alle Quellen, die in einem Artikel angegeben werden, auch im Literaturverzeichnis gelistet sind. Zum Beispiel enthält der Artikel aus Abbildung 3.9 Fußnoten mit Verweisen auf Webseiten, die nicht im Literaturverzeichnis am Ende des Artikels gelistet sind. So wird beispielsweise die Quelle aus Fußnote „2“ in Abbildung 3.12 nicht im Literaturverzeichnis gelistet.



Abbildung 3.12: Information zur Anzahl der Verweise und Ausschnitt der Liste von Fußnoten mit Quellenangaben für den Artikel aus Abbildung 3.9.

¹⁸Auch dieser Reiter wurde vom Skript mit Hilfe des entsprechenden XPath ausgwählt bzw. aktiviert.

REFERENCES

- Akerberg, d. A. (2001): "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, 32 (2), 316-333. [159]
- Bagwell, K. (2007): "The Economic Analysis of Advertising," in *Handbook of Industrial Organization*, Vol. 3. Amsterdam: Elsevier, 1701-1844. [159]
- Berndt, E. R. (1991): *The Practice of Econometrics: Classic and Contemporary*. Reading, MA: Addison-Wesley. [156]

Abbildung 3.13: Anfang der Liste mit Quellen aus dem Literaturverzeichnis für den Artikel aus Abbildung 3.9.

Für die Anzahl der Quellenangaben wurde daher die Anzahl gelisteter Elemente im Literaturverzeichnis erhoben und zur späteren Auswertung (vgl. Kapitel 4, S. 83) herangezogen.¹⁹ Dabei wurde im HTML-Code der zum Literaturverzeichnis gehörende Block identifiziert und in der zugehörigen nicht-nummerierten Liste die Anzahl der Listenelemente bestimmt, indem die Anzahl schließender HTML-Tags (``) für ein Listenelement gezählt wurde. Nachdem die Metadaten für einen Artikel wie beschrieben identifiziert waren, wurden sie vom Python-Skript in eine MySQL-Datenbank geschrieben.

Die in diesem Abschnitt beschriebene Erhebung des JSTOR-Datensatzes soll lediglich als Anschauung dienen, woher die Informationen stammen und eine Idee davon geben, wie der Webscraper funktioniert, mit dem diese erhoben wurden. Die Darstellung ist keinesfalls vollständig, da bewusst Ausnahmefälle und Probleme beim Umgang mit den Daten ausgespart wurden, die hier nicht im Detail erläutert werden sollen. So war die Liste der Literaturangaben beispielsweise nicht immer wie im Beispiel mit „REFERENCES“ benannt, sondern, teilweise abhängig von der Zeitschrift, teilweise abhängig vom Publikationsjahr, auch mit „Bibliography“, „References cited“ u. v. m. Dies erschwerte entsprechend die Identifikation der Stelle, an der sich das Literaturverzeichnis auf der Seite befand. Fehlte eines der Elemente (z. B. der Abstract), dann wirkte sich dies auf die weiteren HTML-Elemente auf dieser Seite aus, da sich die XPath-Ausdrücke änderten, mit denen diese adressiert wurden. Bei den Autorennamen mussten Ausnahmen abgefangen werden, falls beispielsweise einer der Autoren ein Suffix „Jr.“ oder „III.“ führte. Im Verlauf der Datenerhebung wurden hin und wieder auch Kleinigkeiten auf den Seiten von JSTOR angepasst, die im Skript beachtet werden

¹⁹Da die Nutzung und damit auch die Angabe von Quellen aus dem World Wide Web, seien es Webseiten oder dort auffindbare Dokumente, erst seit Bestehen des World Wide Web möglich ist, dürfte die Anzahl tatsächlich verwendeter Quellen, v. a. in späteren Jahren des Untersuchungszeitraums, höher ausfallen und ein Einbeziehen dieser Quellen die in Abschnitt 4.2 (S. 102 ff.) gefundenen positiven Trends sogar noch verstärken.

mussten. Auch weitestgehend ausgespart wurde in den obigen Ausführungen, wie genau die Verarbeitung von Zeichenketten in Python geschah oder wie genau die Befehle des Python-Pakets *Selenium* lauten, mit denen der Webscraper gesteuert wurde.

3.2.2 Umfang

Wie zu Beginn dieses Kapitels bereits erwähnt, umfasst der JSTOR-Datensatz Zeitschriften verschiedener Wissenschaftsdisziplinen. Hierzu zählen *Betriebswirtschaftslehre und Finance* (i. F. auch kurz *BWL und Finance*), *Geographie und Geologie* (i. F. auch kurz *Geo*), *Mathematik, Psychologie*, sowie *Volkswirtschaftslehre* (VWL). Insgesamt wurden aus 55 renommierten Zeitschriften der obigen Disziplinen 178.050 Artikel erfasst. Die frühesten Artikel stammen aus der Zeitschrift *American Journal of Mathematics* und dem Jahr 1878, die jüngsten Artikel stammen aus dem Jahr 2014. Tabelle 3.3 gibt eine Übersicht über die einzelnen Felder, mit Informationen zur Anzahl erfasster Zeitschriften, der Anzahl Artikel sowie der mittleren Anzahl Quellenangaben und der mittleren Anzahl Autoren (Teamgröße) pro Artikel. Bereits in dieser aggregierten Darstellung sind Unterschiede zwischen den einzelnen Disziplinen erkennbar. Im Mittel weisen die Artikel der *Psychologie* beispielsweise die meisten Quellenangaben pro Artikel auf (33,2). Dieser Wert ist dabei mehr als doppelt so hoch wie der von Artikeln in der *Mathematik*. In der *Psychologie* sind ferner auch die größten Teams zu beobachten (im Mittel 2,2 Autoren pro Artikel), das heißt ein Artikel wird hier durchschnittlich von mehr Autoren verfasst, als in den anderen Disziplinen. Mit einem Mittelwert von 1,2 Autoren pro Artikel, sind in der *Geographie und Geologie* die kleinsten Teams zu beobachten.

Tabelle 3.3: Wissenschaftliche Disziplinen im JSTOR-Datensatz (1878–2014).

Disziplin	Anzahl Zeit- schriften	Anzahl Artikel	∅ Anzahl Quellen- angaben	∅ Team- größe
BWL und Finance	16	54.392	28,6	1,8
Geographie und Geologie	4	14.065	30,5	1,2
Mathematik	12	60.331	15,6	1,5
Psychologie	8	12.503	33,2	2,2
VWL	20	66.171	24,7	1,5

Bem.: Für die Berechnung der mittleren Anzahl der Quellenangaben wurden nur Artikel mit mindestens einer Quellenangabe berücksichtigt.

Eine Übersicht aller Zeitschriften im JSTOR-Datensatz mit Informationen zum erhobenen Zeitraum und Umfang findet sich in Tabelle A.1 (S. 204 f.) im Anhang. Wie man dieser Tabelle entnehmen kann, unterscheiden sich die Zeiträume, die für die einzelnen Zeitschriften im Datensatz verfügbar sind. Beispielsweise umfasst der Datensatz für die Zeitschrift *American Economic Review* die Jahre 1911–2014, während er für *Review of Financial Studies* die Jahre 1988–2013 abdeckt. Diese unterschiedlichen Zeiträume ergeben sich einerseits aus dem Zeitraum, in dem die Zeitschrift herausgegeben wurde und andererseits aus den Ausgaben der Zeitschriften, die bei JSTOR verfügbar sind. Ferner bestehen zwischen den Zeitschriften teilweise auch Unterschiede in der Anzahl der Artikel pro Jahr, die innerhalb der Zeitschriften veröffentlicht wurden. In der Disziplin *BWL und Finance* sind beispielsweise mit dem *American Economic Review* oder *Management Science* Zeitschriften mit mehr als 100 Artikeln pro Jahr vertreten, aber auch solche, wie *Administrative Science Quarterly* oder *Journal of Business*, die im Mittel mit weniger als 30 Artikeln pro Jahr vertreten sind.

Die Anzahl der Artikel im JSTOR-Datensatz pro zwei Jahre nimmt für alle Disziplinen bis auf *Geographie und Geologie* im Laufe der Zeit zu, wobei insbesondere in den Jahren nach 1950 ein deutlicher Zuwachs zu beobachten ist (vgl. Abbildung A.6, S. 207, im Anhang). Dieser Anstieg ist einerseits darauf zurückzuführen, dass nicht alle Zeitschriften über den gesamten Zeitraum herausgegeben wurden (die frühesten Artikel im JSTOR-Datensatz stammen wie bereits erwähnt aus dem Jahr 1878) und der Datensatz in späteren Jahren mehr Zeitschriften umfasst. Andererseits hat die Anzahl der erfassten Artikel pro Jahr für die meisten Zeitschriften zugenommen. Abbildung A.7 (S. 208) im Anhang stellt die mittlere Anzahl erfasster Artikel pro Jahr und Zeitschrift für die einzelnen Disziplinen dar. Auch wenn es sich hierbei um ein grobes Maß auf Disziplin-Ebene handelt und entsprechend „kleinere“ Zeitschriften (i. S. der Anzahl veröffentlichter Artikel pro Jahr) von größeren Zeitschriften überlagert werden, sieht man, dass die mittlere Anzahl erfasster Artikel pro Zeitschrift und Jahr für alle Disziplinen außer *Geographie und Geologie* gestiegen ist.

Artikel ab dem Jahr 1960 machen mit 80,5 % den größten Teil des Gesamtdatensatzes aus. Für die *Volkswirtschaftslehre* liegt dieser Anteil bei 76,2 %, für *Mathematik* bei 85,2 %, für *Psychologie* bei 87,7 % und für *Betriebswirtschaftslehre und Finance* gar bei 89,7 %. Für *Geographie und Geologie* ist die Anzahl veröffentlichter Artikel pro zwei Jahre bis auf einige Schwankungen etwa gleich geblieben und hat nicht, wie bei den

anderen Disziplinen, in späteren Jahren stark zugenommen.²⁰ Um eine Vergleichbarkeit zu anderen Studien, wie Hamermesh (2013) und Kosnik (2015), die jeweils Artikel ab 1963 bzw. 1960 betrachten, oder Card und DellaVigna (2013), sowie dem EconLit-Datensatz (vgl. Abschnitt 3.1) zu gewährleisten (jeweils Artikel ab 1970), wird der Untersuchungszeitraum für den JSTOR-Datensatz auf die Jahre ab 1960 eingeschränkt.

Die Anzahl erfasster Artikel pro Jahr geht für alle Disziplinen in den letzten Jahren mitunter stark zurück (vgl. Abbildung A.6, S. 207, im Anhang). Dies liegt vor allem daran, dass zum Zeitpunkt der Datenerhebung für alle Disziplinen in den späteren Jahren nicht alle erhobenen Zeitschriften bei JSTOR verfügbar waren oder nicht mehr herausgegeben wurden (vgl. Tabelle A.1, S. 204, im Anhang). Manche Zeitschriften waren in den letzten Jahren auch nur teilweise bei JSTOR verfügbar (z. B. nur einzelne Ausgaben). Speziell für *Geographie und Geologie*, *Psychologie*, sowie *Mathematik* ist ein besonders starker Rückgang der erfassten Artikel ab dem Jahr 2010 zu beobachten. Um mögliche verzerrende Effekte hierdurch bei der Auswertung der Daten in den letzten Jahren auszuschließen und um die Vergleichbarkeit sowie Einheitlichkeit innerhalb des JSTOR-Datensatzes zu gewährleisten, werden die Publikationsdaten auf den 50-Jahres-Zeitraum von 1960 bis 2009 eingeschränkt.²¹ Insgesamt umfasst der JSTOR-Datensatz nach dieser Einschränkung 133.814 Artikel. Abbildung 3.14 zeigt die Anzahl Artikel pro Jahr für den eingeschränkten Zeitraum von 1960–2009.²² Zwar bestehen zwischen den Disziplinen weiterhin Unterschiede in der Verteilung der Artikel und für alle außer *Geographie und Geologie* werden es im Zeitverlauf mehr Artikel pro Jahr, allerdings ist deutlich zu sehen, dass die Anzahl der Artikel pro Jahr über diesen eingeschränkten Zeitraum gleichmäßiger über die Jahre verteilt ist als für den gesamten Zeitraum (vgl. Abbildung A.6, S. 207, im Anhang).

Analog zu Tabelle 3.3 findet sich in Tabelle 3.4 eine Übersicht der Disziplinen mit Informationen zur Anzahl der Artikel, der mittleren Anzahl Quellenangaben und der mittleren Teamgröße im JSTOR-Datensatz für die Jahre 1960–2009. Vergleicht man die Mittelwerte in den Tabellen 3.3 und 3.4, so lassen sich bereits erste Unterschiede und Trends bzgl. der mittleren Anzahl Quellenangaben und der mittleren Teamgröße ausmachen. Bis auf den Wert für die mittlere Anzahl der Quellenangaben für *Mathematik*, sind

²⁰Wie man Abbildung A.7 (S. 208) im Anhang entnimmt, ist die mittlere Anzahl Artikel pro Jahr und Zeitschrift für *Geographie und Geologie* leicht zurückgegangen. Der Anteil Artikel ab 1960 macht hier 45,3 % aus.

²¹Alle Untersuchungen für den JSTOR-Datensatz aus Kapitel 4 wurden zusätzlich für alle erfassten Artikel ab 1960 durchgeführt (d. h. auch solche aus den Jahren 2010–2014) und liefern, bezogen auf grafischen Trends, sowie die Effektrichtungen und Signifikanzniveaus in den Regressionen, grundsätzlich die gleichen Ergebnisse.

²²Man beachte, dass die Dimensionen der Ordinate in dieser Darstellung nicht übereinstimmen.

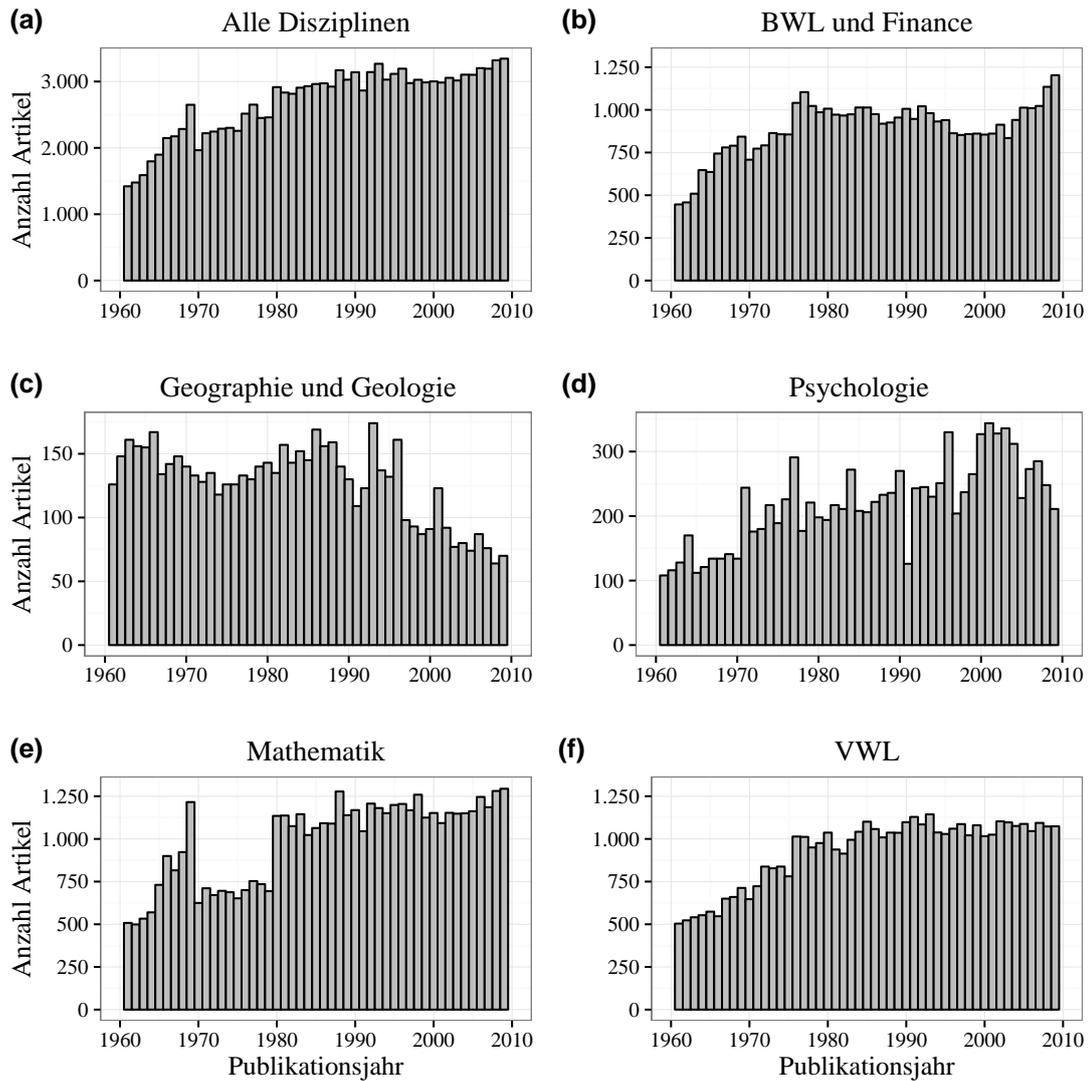


Abbildung 3.14: Anzahl der Artikel pro Jahr im JSTOR-Datensatz (1960–2009).

die jeweiligen Mittelwerte in Tabelle 3.4 mindestens so groß wie die in Tabelle 3.3.²³ Wie in Abschnitt 4.1.1 (S. 84) und Abschnitt 4.2.1 (S. 102) gezeigt wird, sind für die einzelnen Disziplinen tatsächlich steigende Trends bzgl. dieser Kennzahlen zu beobachten. Entsprechend müssen die Werte steigen, wenn man den Datensatz, wie in Tabelle 3.4, auf einen späteren Zeitraum einschränkt. Eine Übersicht aller Zeitschriften aus dem JSTOR-Datensatz eingeschränkt auf den für die spätere Untersuchung (Kapitel 4) relevanten Zeitraum von 1960–2009, mit Informationen zum erfassten Zeitraum und zur Anzahl erfasster Artikel, findet sich in Tabelle 3.5. Aus der Tabelle geht auch hervor, dass selbst nach der Einschränkung auf den Zeitraum ab 1960–2009 noch Unterschiede in den erhobenen Zeiträumen und der mittleren Anzahl Artikel pro Jahr und Zeitschrift bestehen. Daher wird in den späteren Regressionsanalysen (vgl. Kapitel 4) auch für die einzelnen Zeitschriften kontrolliert, um mögliche verzerrende Effekte einzelner Zeitschriften auf den Gesamttrend abzufangen.

Tabelle 3.4: Wissenschaftliche Disziplinen im JSTOR-Datensatz (1960–2009).

Disziplin	Anzahl Zeit- schriften	Anzahl Artikel	Ø Anzahl Quellen- angaben	Ø Team- größe
BWL und Finance	16	44.082	26,8	1,8
Geographie und Geologie	4	6.337	32,7	1,3
Mathematik	12	48.890	15,1	1,6
Psychologie	8	10.806	34,3	2,3
VWL	20	46.082	23,4	1,5

Bem.: Für die Berechnung der mittleren Anzahl der Quellenangaben wurden nur Artikel mit mindestens einer Quellenangabe berücksichtigt.

Tabelle 3.5: Zeitschriften im JSTOR-Datensatz (1960–2009).

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
BWL und Finance				
Academy of Management Journal	1963–2009	2.631	86,8 %	56,0
Administrative Science Quarterly	1960–2009	1.384	78,5 %	27,7
American Economic Review	1960–2009	7.397	88,4 %	147,9
Financial Analysts Journal	1960–2009	3.486	15,3 %	69,7

²³Tatsächlich sind bis auf die mittlere Anzahl der Quellenangaben in der *Mathematik* alle Mittelwerte in Tabelle 3.4 größer, als die entsprechenden Werte in Tabelle 3.3. Dies wird allerdings aufgrund der Rundung auf eine Nachkommastelle nicht immer ersichtlich.

Tabelle 3.5: (Fortsetzung)

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
Financial Management	1972–2009	1.333	92,3 %	35,1
International Economic Review	1960–2009	2.399	96,6 %	48,0
Journal of Business	1960–1997	1.129	46,9 %	29,7
Journal of Consumer Research	1974–2009	1.692	94,3 %	47,0
Journal of Finance	1960–2009	4.574	77,2 %	91,5
Journal of Marketing	1960–2009	2.253	50,4 %	45,1
Journal of Marketing Research	1964–2009	2.488	94,2 %	54,1
Management Science	1960–2009	5.986	93,9 %	119,7
Marketing Science	1982–2009	964	59,5 %	34,4
MIS Quarterly	1977–2009	956	28,6 %	29,0
Operations Research	1960–2009	4.469	96,3 %	89,4
Review of Financial Studies	1988–2009	941	99,0 %	42,8
Geographie und Geologie				
Geographical Review	1960–2009	1.525	19,5 %	30,5
Geographische Zeitschrift	1963–2009	783	6,4 %	16,7
Geography	1960–2009	2.178	49,8 %	43,6
Journal of Geology	1960–1997	1.851	81,1 %	48,7
Mathematik				
American Journal of Mathematics	1960–2009	2.460	85,3 %	49,2
Annals of Mathematics	1960–2009	2.299	22,5 %	46,0
Bulletin of Symbolic Logic	1995–2009	251	72,9 %	16,7
Econometrica	1960–2009	3.544	94,5 %	70,9
Journal of the Am. Mathematical Society	1988–2009	677	53,5 %	30,8
Mathematics of Computation	1960–2009	4.952	84,3 %	99,0
Mathematics of Operations Research	1976–2009	1.699	96,4 %	50,0
Operations Research	1960–2009	4.469	96,3 %	89,4
Proc. of the Am. Mathematical Society	1960–2009	17.212	96,7 %	344,2
SIAM Journal on Applied Mathematics	1966–2009	4.576	39,3 %	104,0
SIAM Journal on Numerical Analysis	1966–2009	4.141	51,6 %	94,1
SIAM Review	1960–2009	2.610	28,9 %	52,2
Psychologie				
Child Development	1960–2009	6.414	98,7 %	128,3

Tabelle 3.5: (Fortsetzung)

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
Journal of Business and Psychology	1986–2009	802	97,1 %	33,4
Journal of Consumer Psychology	1992–2004	337	97,9 %	25,9
Journal of Occupational Behaviour	1980–1987	188	100,0 %	23,5
Journal of Organizational Behavior	1988–2009	1.137	96,5 %	51,7
Psychological Inquiry	2000–2007	384	98,2 %	48,0
Social Psychology Quarterly	1979–2009	884	62,6 %	28,5
Sociometry	1960–1977	660	0,0 %	36,7
VWL				
Am. Economic Review	1960–2009	7.397	88,4 %	147,9
American Economist	1960–2009	1.175	69,2 %	23,5
Am. Journal of Economics and Sociology	1970–2009	2.232	34,5 %	55,8
Econometric Theory	1985–2009	1.601	80,5 %	64,0
Econometrica	1960–2009	3.544	94,5 %	70,9
Economic Geography	1960–1989	668	61,5 %	22,3
Economic Theory	1991–2009	1.568	98,1 %	82,5
International Economic Review	1960–2009	2.399	96,6 %	48,0
Journal of Economic Growth	1996–2009	198	100,0 %	14,1
Journal of Economic Literature	1969–2009	628	81,5 %	15,3
Journal of Labor Economics	1983–2009	793	57,4 %	29,4
Journal of Political Economy	1960–2009	3.328	66,3 %	66,6
Economic History Review	1960–2009	1.625	3,4 %	32,5
Economic Journal	1960–2009	3.381	77,0 %	67,6
Journal of Economic Perspectives	1987–2009	1.192	92,2 %	51,8
Journal of Finance	1960–2009	4.574	77,2 %	91,5
Quarterly Journal of Economics	1960–2009	2.475	55,9 %	49,5
RAND Journal of Economics	1984–2009	1.091	99,9 %	42,0
Review of Economic Studies	1960–2009	2.221	94,2 %	44,4
Review of Economics and Statistics	1960–2009	3.992	79,6 %	79,8

Bem.: (i) Die Abkürzungen Am. bzw. Proc. stehen für American bzw. Proceedings.
(ii) Die Spalte „Info zu Quellen“ gibt an, zu welchem Anteil der Artikel eine Information bzgl. der Anzahl der Quellenangaben vorliegt.

In Tabelle 3.5 sind die Zeitschriften *Econometrica* und *Operations Research* jeweils zwei Disziplinen zugeordnet.²⁴ Erstere den Disziplinen *Mathematik* und *Volkswirtschaftslehre*, letztere den Disziplinen *Betriebswirtschaftslehre und Finance* sowie *Mathematik*.²⁵ Diese Zuordnung entspricht der Zuordnung von Zeitschriften bei JSTOR und erscheint inhaltlich sinnvoll. Daher wurde die Zuordnung dieser Zeitschriften auch in dieser Arbeit übernommen.

3.3 Der Science-Datensatz

Die Fachzeitschrift *Science*, auch bekannt unter dem Namen *Science Magazine*, gehört zu den wichtigsten wissenschaftlichen Zeitschriften weltweit und erreicht nach eigenen Angaben über 400.000 Leser jede Woche AAAS (2017). Herausgegeben wird die Zeitschrift von der gemeinnützigen Organisation *American Association for the Advancement of Science* (AAAS) mit Sitz in Washington, D.C. Die wöchentlich erscheinenden Ausgaben sind nicht auf ein Fachgebiet beschränkt, sondern decken verschiedene Wissenschaftsfelder ab, wobei der Fokus auf den Naturwissenschaften liegt. Die Zeitschrift weist auch bzgl. ihrer Leserschaft ein breites Spektrum auf. So wird *Science* nach eigenen Angaben vor allem an Universitäten gelesen. Zur Leserschaft gehören aber auch Industrieunternehmen, die Politik, das Gesundheitswesen und gemeinnützige Organisationen.

Angetrieben von dem Erfolg der Fachzeitschrift *Nature* und dem damaligen Mangel an wissenschaftlichen Fachzeitschriften wurde die Zeitschrift 1880 vom New Yorker Journalisten John Michels gegründet. Der erste Geldgeber war kein Geringerer als der amerikanische Erfinder und Geschäftsmann Thomas Alva Edison. Einige Jahre später wurde das Magazin von Graham Bell übernommen.²⁶ Heutzutage ist *Science* die führende wissenschaftliche Zeitschrift weltweit. Mit 129.559 Abonnenten liegt sie nach eigenen Angaben deutlich vor den Zeitschriften *BioTechniques* (80.008), *Genetic Engineering & Biotechnology News* (64.776) und *Nature* (52.613).

Die Publikationsdaten für die Zeitschrift *Science* stammen wie auch die Publikationsdaten für die Disziplinen *Betriebswirtschaftslehre und Finance*, *Geologie und*

²⁴Dies gilt auch für Tabelle A.1 (S. 204) im Anhang, die für alle erfassten Artikel die gleichen Informationen wie Tabelle 3.5 ausweist, allerdings ohne die Einschränkung auf den Zeitraum 1960–2009.

²⁵Entsprechend addieren sich die Artikelzahlen für die einzelnen Disziplinen in den Tabellen 3.5 und A.1 (S. 204 im Anhang) nicht zur Gesamtzahl der Artikel von 178.050 bzw. 133.814.

²⁶Diese und weitere interessante Details zu den mitunter schwierigen Anfängen und zur weiteren Geschichte der Zeitschrift *Science* sind in Kohlstedt (1980) nachzulesen.

Geographie, Mathematik, Psychologie sowie *Volkswirtschaftslehre* aus der Publikationsdatenbank von JSTOR. Diese wurden jedoch nicht selbst erhoben, sondern über JSTOR *Data for Research* von JSTOR bezogen. Diese Dienstleistung ermöglicht es Wissenschaftlern, individuelle Anfragen zu Publikationsdaten zu stellen.²⁷ Diese Anfragen werden anschließend geprüft und der Datensatz dem Antragsteller bei erfolgreicher Prüfung (es werden keine Kriterien bzgl. dieser Entscheidung genannt) im *XML-Format* zur Verfügung gestellt (vgl. Seite 50, Fußnote 14, zu XML).

Die zur Verfügung gestellten Daten kommen somit nicht in Form einer einzelnen Datei, die man direkt mit einem der gängigen Statistikprogramme (z. B. R, SPSS oder STATA) auswerten kann, wie beispielsweise eine CSV-Datei, sondern als ein gepackter ZIP-Ordner, innerhalb dessen wiederum die einzelnen Bände und Ausgaben der Zeitschrift *Science* in einer Ordnerstruktur vorliegen, in der es für jeden Artikel einen separaten Ordner und darin ein XML-Dokument mit den entsprechenden Metadaten gibt. Die Metadaten umfassen den Publikationstyp („*book-review*“, „*editorial*“, „*misc*“, „*news*“ und „*research-article*“), den Titel des Artikels, falls vorhanden den Abstract, die Vor- und Zunamen der einzelnen Autoren, das Publikationsjahr, den Band, die Ausgabe, die Seitenangabe und die Anzahl der Quellenangaben.²⁸

Für die vorliegende Arbeit wurden die einzelnen XML-Dokumente mit Hilfe eines Python-Skripts gelesen und für die weitere Verarbeitung zeilenweise in eine Excel-Datei geschrieben. In Abbildung 3.15 ist beispielhaft der Teil mit den Metadaten zum Artikel für ein solches XML-Dokument abgebildet (geöffnet mit Syntax-Highlighting im Editor *Notepad++*). Wie man sieht, ist das Dokument anhand entsprechender Tags strukturiert.²⁹ So lässt sich beispielsweise gezielt der Titel des Beitrags auslesen, da er innerhalb des öffnenden und schließenden `article-title`-Tags steht. Aus den Metadaten der XML-Dokumente lassen sich ferner weitere Angaben ableiten, nämlich die Anzahl der Autoren, die Länge des Artikels in Seiten, sowie das Geschlecht der Autoren.³⁰

Bei der Auswertung des Datensatzes (vgl. Kapitel 4, S. 83 ff.) wurden aus obigem Datensatz lediglich Publikationen mit dem Publikationstyp „*research-article*“ berücksich-

²⁷Vgl. www.jstor.org/dfr für nähere Informationen zu JSTOR *Data for Research*.

²⁸Ein Abstract liegt für 78,3 % der Artikel im Datensatz ab 1960, d. h. dem Untersuchungszeitraum für die Zeitschrift *Science* in der vorliegenden Arbeit (vgl. unten), vor. Die Anzahl der Quellenangaben ist für 81,8 % der Artikel ab 1960 bekannt.

²⁹Tags sind Strukturelemente eines XML-Dokumentes. Beispielsweise markieren die Tags `<article-title>` bzw. `</article-title>` den Anfang bzw. das Ende des Titels (vgl. Abbildung 3.15). Vgl. auch Fußnote 17 (S. 51) bzgl. HTML-Tags.

³⁰Vgl. Abschnitt 3.5.1 (S. 71) für eine Erläuterung zur Bestimmung des Geschlechts.

sichtigt, da es sich bei den restlichen Typen nicht um wissenschaftliche Beiträge im engeren Sinne handelt. Der Datensatz reicht bis ins 19. Jahrhundert und umfasst 101.307 Artikel im Zeitraum von 1880–2011. Abbildung 3.16 zeigt die Entwicklung der Anzahl der Artikel pro zwei Jahre über diesen Zeitraum. Wie man sieht, ist die Anzahl veröffentlichter Artikel pro zwei Jahre bis auf wenige, kleinere Schwankungen kontinuierlich gestiegen. Analog zum JSTOR-Datensatz wird auch der *Science*-Datensatz auf Artikel ab dem Jahr 1960 eingeschränkt.³¹ Der finale Datensatz umfasst 70.028 Artikel, deren Anteil am Gesamtdatensatz 69,1 % beträgt. Die Anzahl der Artikel pro Jahr im Zeitraum 1960–2011 ist in Abbildung 3.17 abgebildet.

```

<article xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns:xlink="http://www.w3.org/1999/xlink"
  xmlns:mml="http://www.w3.org/1998/Math/MathML"
  dtd-version="1.0"
  article-type="research-article">
  <front>
    <journal-meta>
    <article-meta>
      <article-id pub-id-type="jstor">1697758</article-id>
      <article-categories>
        <subj-group>
          <subject>Research News</subject>
        </subj-group>
      </article-categories>
      <title-group>
        <article-title>Egg Laying is for the Birds</article-title>
      </title-group>
      <contrib-group>
        <contrib contrib-type="author">
          <string-name>
            <given-names>Roger</given-names>
            <surname>Lewin</surname>
          </string-name>
        </contrib>
      </contrib-group>
      <pub-date>
        <day>17</day>
        <month>10</month>
        <year>1986</year>
      </pub-date>
      <volume>234</volume>
      <issue>4774</issue>
      <issue-id>i298828</issue-id>
      <fpage>285</fpage>
      <lpage>285</lpage>
      <page-range>285</page-range>
    
```

Abbildung 3.15: Beispiel eines XML-Dokuments zum *Science*-Datensatz.

³¹Die in Kapitel 4 beschriebenen Ergebnisse bleiben jedoch weitestgehend vergleichbar, wenn man längere Beobachtungszeiträume betrachtet und erweisen sich in diesem Sinne als robust.

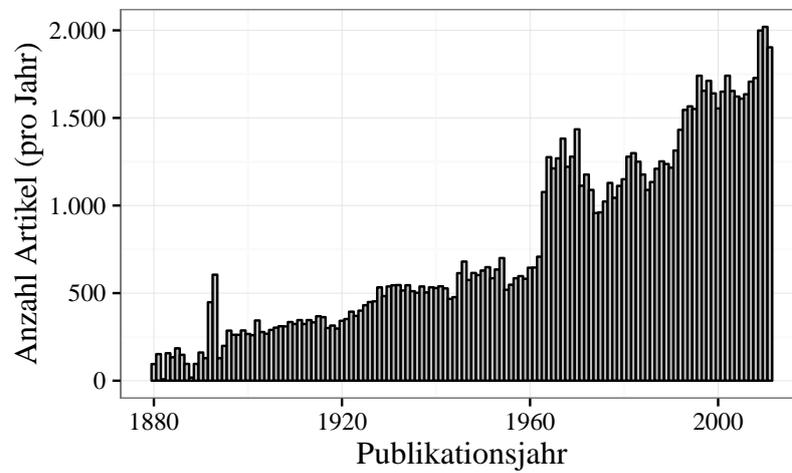


Abbildung 3.16: Anzahl der Artikel pro Jahr ab 1880 für die Zeitschrift *Science*.

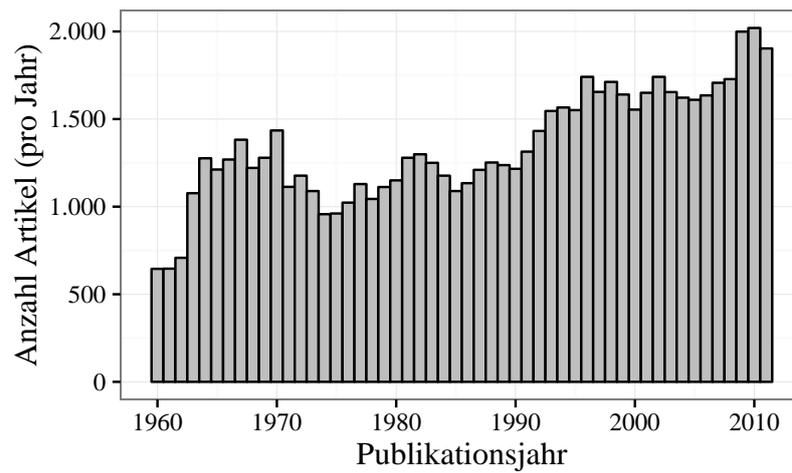


Abbildung 3.17: Anzahl der Artikel pro Jahr ab 1960 für die Zeitschrift *Science*.

3.4 Identifikation von Autoren

3.4.1 Problematik

In den Publikationsdaten dieser Arbeit sind zu jedem Artikel die Namen des zugehörigen Autors bzw. der zugehörigen Autoren aufgeführt. Allerdings lässt sich aus einem Namen nicht immer zweifelsfrei auf eine eindeutige Person schließen. Eine solche Situation tritt zum Beispiel dann ein, wenn ein und derselbe Autor mit mehreren Publikationen in der Datenbank unter verschiedenen Namen auftritt. Beispielsweise könnte ein „Peter Michael Mustermann“ auch mit den folgenden Namensvarianten als Autor von Artikeln angegeben sein:³²

- Peter M. Mustermann,
- P. M. Mustermann,
- Peter Mustermann und
- P. Mustermann.

Möglich ist auch, dass ein Autor im Laufe seiner Karriere seinen Namen ändert. Dies kann sowohl eine Änderung des Vornamens, als auch des Nachnamens sein. Letzteres kann beispielsweise bei Eheschließungen auftreten, bei denen einer der Ehegatten den Nachnamen des Partners oder einen Doppelnamen annimmt. Im Falle, dass ein Autor nach Eheschließung unter einem gänzlich anderen Nachnamen auftritt, ist eine automatisierte Zuordnung dieses Namens zu anderen Namen unter denen er auftritt aus den vorliegenden Daten quasi unmöglich und es bedarf zusätzlicher manueller oder automatisierter Recherche aus teilweise kostenpflichtigen Datenquellen wie ORCID oder Scopus.³³ Nimmt ein Autor einen Doppelnamen an, ist nicht zwingend eine manuelle Recherche notwendig. Dieser Fall dürfte gerade bei bereits publizierten und weiterhin aktiven Wissenschaftlern häufiger vorkommen, die weiterhin als Autoren ihrer vergangenen Publikationen erkannt werden möchten.

Neben der Schwierigkeit, dass der gleiche Autor in einem Datensatz unter verschiedenen Namen auftritt, sind auch Fälle denkbar, in denen die Namen verschiedener Autoren übereinstimmen. Dies umfasst einerseits die Fälle, in denen Vor- und Nachnamen mindestens zweier Autoren exakt übereinstimmen, aber auch solche, bei denen beispielsweise nur die Initialen und der Nachname übereinstimmen, sofern die Namen

³²Die Art der Angabe muss dabei nicht zwingend auf den Autor selbst zurückzuführen sein, sondern kann auch durch unterschiedliche Zitierweisen bzw. Formen der Angabe in Zeitschriften, von Herausgebern und in Datenbanken begründet sein.

³³Siehe www.orcid.org und www.scopus.com für die Webseiten der genannten Datenquellen.

entsprechend gekürzt angegeben werden. Zum Beispiel könnten zwei unterschiedliche Autoren „Michael Mustermann“ und „Markus Mustermann“ bei ihren Publikationen als „M. Mustermann“ angegeben sein. In diesem Fall ist es schwierig Personen ohne manuelle Recherche auseinanderzuhalten.

Leider sind zum Zeitpunkt dieser Arbeit weder bei *Web of Science* noch bei JSTOR die Namen der Autoren mit den zugehörigen Personen verknüpft. Denkbar wäre zum Beispiel, dass den Autorennamen bereits durch den Verlag bzw. die Verlage oder Datenbanken wie JSTOR und *Web of Science* eindeutige personenbezogene IDs zugeordnet sind. Unter Umständen könnten auch diese zuvor von den Betreibern der Literaturdatenbanken heuristisch bestimmt worden sein. Die Schwierigkeit, der man sich somit gegenüber sieht, besteht also darin, zu einer gegebenen Liste an Namen möglichst gut die zugehörigen Personen, das heißt in diesem Fall die einzelnen Autoren, zu bestimmen. Solange alle Personen in der Datenbank unterschiedliche Namen haben und jede einzelne Person stets unter demselben Namen auftritt, ist dies nicht weiter problematisch. Leider ist dies in den vorliegenden Datensätzen nicht der Fall, weshalb es hier einer Nachbearbeitung der Daten bedarf, um das Problem zu lösen.

3.4.2 Methodischer Ansatz

Für die Identifikation von Personen (engl. *author disambiguation*) wurde in dieser Arbeit ein graphentheoretischer Ansatz gewählt. Die verwendete heuristische Methode, um Personen aus einer Liste von Namen zu identifizieren, wurde im Laufe der Erstellung dieser Arbeit von Sascha Schweitzer in ein Python-Paket verpackt und auf der Seite *github.com* veröffentlicht und beschrieben (Schweitzer 2017). Dabei handelt es sich um einen graphentheoretischen Ansatz, der einem hierarchischen Prozess folgt. Zunächst werden alle Namen in Mengen zusammengefasst, die den gleichen Nachnamen tragen. Innerhalb einer jeden solchen Menge werden jeweils paarweise Beziehungen zwischen den unterschiedlichen Vornamen untersucht.³⁴ Die Vornamen werden entweder als *gleich*, *verschieden*, *Teilmengen voneinander* oder als *partiell kompatibel* kategorisiert. Beispielsweise ist der Vorname „Peter“ identisch mit bzw. gleich wie „Peter“ und verschieden von „Wolfgang“. „Peter M.“ ist in dieser Terminologie Teilmenge von „Peter“, auch wenn es sich für die beiden Zeichenketten entgegengesetzt verhält. Das liegt daran, dass „Peter M.“ mehr spezifische Informationen beinhaltet als „Peter“

³⁴In der Terminologie der Graphentheorie werden die Beziehungen zwischen den einzelnen Namen in einer Adjazenzmatrix (deren Zeilen und Spalten den einzelnen Namen entsprechen) abgespeichert, aus der sich die grafische Repräsentationsform des entsprechenden Graphen ableiten lässt.

und mit weniger anderen Vornamenskombinationen kompatibel ist, als „Peter“. Falls keiner der drei genannten Fälle auf zwei Einträge zutrifft, so werden diese als partiell kompatibel definiert. Beispielsweise stimmt „P. Michael“ teilweise mit „Peter M.“ überein bzw. es ist nicht auszuschließen, dass die Namensangaben sich beide auf einen „Peter Michael“ beziehen.

Nachdem alle binären Beziehungen bestimmt sind, entsteht aus den Namen als Knoten und den Beziehungen zwischen ihnen als Pfeile für jeden Nachnamen ein Graph wie in Abbildung 3.18.³⁵ Die Pfeile zeigen dabei stets von einem Knoten bzw. Namen in Richtung der Knoten bzw. Namen, die Teilmengen dieses Namens darstellen, das heißt mehr spezifische Informationen enthalten. Zuletzt werden „Abkürzungen“ von einem Knoten zum anderen eliminiert, um einen minimalen Graphen zu erhalten. Abkürzungen sind dabei alle Wege von einem Knoten A zu einem anderen Knoten B , die „kürzer“ sind (gemessen an der Anzahl Pfeile bzw. Knoten, die entlang dieses Weges abgelaufen bzw. besucht werden), als der „längste“ Weg zwischen diesen beiden Knoten im Graphen.³⁶ In Abbildung 3.18 (b) sind zwei solche Abkürzungen als gestrichelte Pfeile dargestellt.

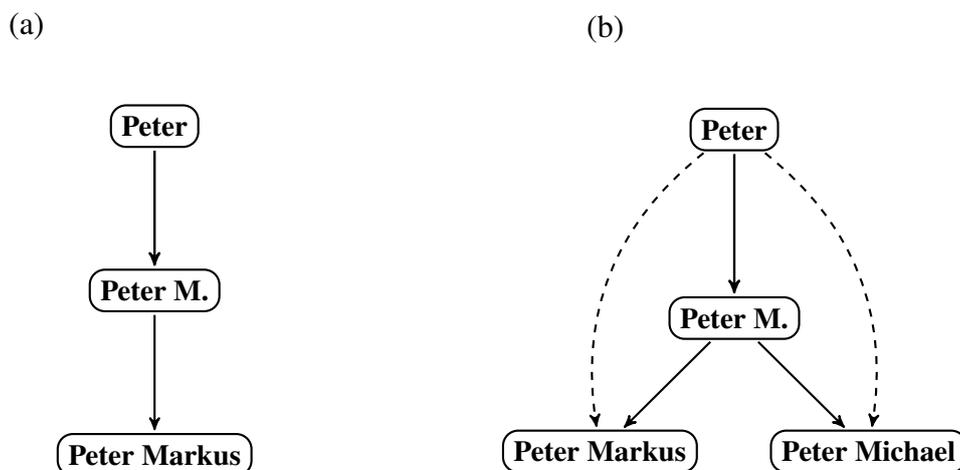


Abbildung 3.18: Beispiel eines minimalen Graphen für die Beziehungen zwischen verschiedenen Vornamen.

Alle Namen, die einer Person im Datensatz zugeordnet werden, liegen im minimalen Graphen, ausgehend von den Blättern des Graphen, das heißt solchen Knoten, die zwar Vorgängerknoten aber keine Nachfolger haben, entlang der nicht-verzweigenden Wege

³⁵Daraus folgt auch, dass die Anzahl identifizierter Personen aus einer Namensliste mindestens der Anzahl unterschiedlicher Nachnamen innerhalb der Liste entspricht.

³⁶In der Graphentheorie bezeichnet ein *Weg* eine Pfeilfolge, deren Knoten sämtlich verschieden sind. Der Fakt, dass von Pfeilen gesprochen wird, impliziert, dass es sich um einen gerichteten Graphen handelt.

aufwärts. Im Beispiel in Abbildung 3.18 (a) würden also drei Autorennennungen mit identischen Nachnamen und den Vornamen „Peter“, „Peter M.“, sowie „Peter Michael“ der gleichen Person zugeschrieben, falls sie die einzigen drei Vornamen für diesen Nachnamen sind. Sie würden allerdings nicht der gleichen Person zugeschrieben, falls es einen „Peter“, „Peter M.“, „Peter Markus“ und „Peter Michael“ gäbe, da es in diesem Fall eine Verzweigung unterhalb von „Peter M.“ gäbe (vgl. Beispiel (b) in Abbildung 3.18). Im Beispiel (b) würden somit drei Personen unterschieden. „Peter Markus“, „Peter Michael“ und eine Person, die unter den Namen „Peter“ und „Peter M.“ auftritt. Dies gilt natürlich nur, solange es zum gleichen Nachnamen beispielsweise keinen Autor mit der Vornamensvariante „Peter B.“ im Datensatz gibt.

Am Beispiel der Artikel und Autoren aus den psychologischen Zeitschriften des JSTOR-Datensatzes wird an dieser Stelle ein Eindruck von den Verhältnissen zwischen den Autorennennungen, der Anzahl unterschiedlicher Nachnamen im Datensatz und der Anzahl identifizierter Personen gegeben. Für jeden Artikel in den Datensätzen sind die Namen von mindestens einem Autor angegeben. Somit werden mindestens so viele Namen genannt, wie es Artikel gibt. Für die insgesamt 12.503 erhobenen Artikel der Disziplin *Psychologie* gibt es insgesamt 27.161 Autorennennungen, das heißt im Mittel 2,2 Autorennennungen pro Artikel (vgl. hierzu auch die mittlere Teamgröße in Tabelle 3.3, S. 59). Innerhalb der genannten Namen von Autoren finden sich 10.296 verschiedene Nachnamen. Diese Zahl stellt eine untere Schranke für die Anzahl identifizierter Personen dar, da Autoren mit unterschiedlichen Nachnamen auch als unterschiedliche Personen identifiziert werden. Unter den Namensnennungen finden sich ferner 16.482 verschiedene Vorname-Nachname-Kombinationen. Deren Anzahl stellt eine obere Schranke für die Anzahl identifizierter Personen dar, da wie oben beschrieben ggf. Autoren mit demselben Nachnamen, aber unterschiedlichen Vornamensvarianten einer Person zugeordnet werden. Letztlich identifiziert die verwendete Methode 15.641 Personen innerhalb des Teildatensatzes der *Psychologie*. Informationen zu allen Datensätzen für den jeweiligen Untersuchungszeitraum (d. h. für den EconLit-Datensatz 1970–2014, den JSTOR-Datensatz 1960–2009 und die Zeitschrift *Science* 1960–2011) finden sich in Tabelle 3.6.

3.4.3 Güte des verwendeten Ansatzes

Der verwendete Algorithmus wurde im Rahmen einer Abschlussarbeit anhand von Publikationsdaten für elf renommierte Zeitschriften der Volkswirtschaftslehre getestet. Dieser Datensatz umfasste insgesamt 95.503 Einträge mit Informationen zur Zeitschrift,

Tabelle 3.6: Anzahl identifizierter Personen in den einzelnen Datensätzen.

Datensatz	# Artikel	# Autoren- nennungen	# Nachnamen	# Personen
EconLit	762.826	1.285.534	169.782	381.964
EconLit (Top)	28.859	47.383	11.543	17.390
JSTOR				
BWL und Finance	44.082	77.219	18.148	31.227
Geo	6.337	9.218	4.290	5.432
Mathematik	48.890	76.983	22.512	35.463
Psychologie	10.806	24.667	9.591	15.686
<i>Science</i>	70.028	271.698	64.182	135.747

dem Publikationsjahr, Vor- und Nachnamen der Autoren, Titel des Artikels, etc. Dabei entspricht jeder Eintrag einer Autorennennung. Das heißt, dass ein Artikel, der von mehreren Autoren geschrieben wurde, mehrfach im Datensatz vertreten ist, nämlich jeweils einmal für jeden seiner Autoren. Der Test-Datensatz unterscheidet sich in zweierlei Hinsicht von den Top-Artikeln aus dem EconLit-Datensatz. Einerseits umfasst er neben den neun Top-Zeitschriften (vgl. Abschnitt 3.1, S. 42), noch die Zeitschriften *Review of Financial Studies* und *Journal of Economic Growth*. Dieser Unterschied ist dadurch zu erklären, dass der obige Algorithmus bereits implementiert und getestet wurde, bevor die letztlichen Datensätze für die vorliegende Untersuchung definiert wurden. Ferner umfasst der Test-Datensatz nicht nur den Zeitraum 1970–2014, sondern alle Jahre ab 1886.

id	frm_normalized	snm_normalized	source	source_id	frm	snm	year	maximum_time_gap	matching
211	dennis j	aigner	citedata	59976	Dennis J.	Aigner	1966	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	39677	Dennis J.	Aigner	1968	11	{'vertical', 'equal'}
211	d j	aigner	citedata	39830	D. J.	Aigner	1968	11	{'vertical', 'equal'}
211	d j	aigner	citedata	77598	D. J.	Aigner	1968	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	5016	Dennis J.	Aigner	1971	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	17320	Dennis J.	Aigner	1971	11	{'vertical', 'equal'}
211	d j	aigner	citedata	13614	D. J.	Aigner	1973	11	{'vertical', 'equal'}
211	d j	aigner	citedata	49166	D. J.	Aigner	1973	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	816	Dennis J.	Aigner	1974	11	{'vertical', 'equal'}
211	d j	aigner	citedata	873	D. J.	Aigner	1974	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	402	Dennis J.	Aigner	1977	11	{'vertical', 'equal'}
211	dennis j	aigner	citedata	1634	Dennis J.	Aigner	1988	11	{'vertical', 'equal'}

Abbildung 3.19: Beispiel eines Namensclusters für den Nachnamen „Aigner“ (Quelle: Wang 2015, S. 57, Abbildung 16).

Insgesamt gelangt der Algorithmus für den Test-Datensatz zu 27.407 Personen bzw. den zugehörigen Namensclustern. Die Namenscluster enthalten den Nachnamen und alle

Vornamen, die einer Person vom Algorithmus zugeordnet werden. In Abbildung 3.19 ist ein solches Namenscluster für „Dennis J. Aigner“ abgebildet, der dem Algorithmus zufolge auch unter der Vornamensvariante „D. J.“ auftritt. Jeder Person und all ihren Namensvarianten ist im Datensatz eine eindeutige Personen-ID zugeordnet (im Beispiel ist dies die ID 211). 24.232 (88,4 %) der ermittelten Namenscluster sind in dem Sinne eindeutig, dass sie lediglich eine Vornamensvariante enthalten, die zugehörige Person im Datensatz also stets unter dem gleichen eindeutigen Namen auftritt.³⁷ Entsprechend gibt es im Test-Datensatz 3.175 (d. h. 11,6 %) nicht eindeutige Namenscluster, die somit Namensteilmengen enthalten, wie dies auch in Abbildung 3.19 der Fall ist.

Aus den nicht eindeutigen Namensclustern wurden zufällig zehn zur näheren Überprüfung herangezogen. Dabei wurde nach den Autoren und ihren Publikationslisten recherchiert und diese mit den zugeordneten Einträgen abgeglichen. Für 8 der 10 untersuchten Cluster stimmte die Publikationsliste mit den Artikeln der ermittelten Autoren überein. Für zwei identifizierte Autoren gehörten die zugehörigen Artikel tatsächlich jeweils zu unterschiedlichen Personen. Damit ordnet der Algorithmus in dieser Stichprobe in einem von fünf Fällen zwei Namensvarianten fälschlicherweise der gleichen Person bzw. dem gleichen Autor zu. Allerdings entstammt die Stichprobe der zehn getesteten Namenscluster den nicht eindeutigen Zuordnungen, die 11,6 % bzw. etwa ein Neuntel aller Namenscluster im Test-Datensatz ausmachen. Somit liegt die Fehlerquote, das heißt der Anteil fälschlicher Zuordnungen von Namen zur gleichen Person, für den gesamten Test-Datensatz bei 2,3 %.

3.5 Ergänzung der Publikationsdaten

3.5.1 Persönliche und institutionelle Informationen

Die oben vorgestellten Publikationsdaten wurden ferner um das Geschlecht der Autoren ergänzt. Für die Bestimmung des Geschlechts wurde das R-Paket „R_gender“ verwendet. Dieses bestimmt heuristisch auf Basis des Vornamens einer Person ihr Geschlecht. Es basiert auf Daten der *U.S. Social Security Administration* zu den beliebtesten männlichen und weiblichen Vornamen in den USA. Konkret ist in diesen Daten für jedes der Jahre ab 1880 erfasst, welcher Name wie häufig in diesem Jahr an Neugeborene vergeben

³⁷Theoretisch ist es denkbar, dass in einem Datensatz zwei unterschiedliche Personen mit identischen Vor- und Nachnamen auftreten. Solche Fälle lassen sich mittels des Algorithmus nicht erkennen bzw. die entsprechenden Personen nicht auseinanderhalten.

wurde.³⁸ Damit lässt sich für alle Namen, die in diesen Daten auftreten, entscheiden, ob sie eher zu einer männlichen oder einer weiblichen Person gehören. Dabei kann nicht allen Autoren aus den obigen Publikationsdaten ein Geschlecht zugewiesen werden, da das verwendete Paket in den USA verbreitete Vornamen umfasst. Ferner sind für manche Autoren lediglich die Initialen bekannt, sodass hier eine Schätzung des Geschlechts nicht möglich ist. Tabelle 3.7 veranschaulicht, wie viele Autoren in den einzelnen Datensätzen als männlich oder weiblich identifiziert wurden und bei wie vielen eine Identifizierung nicht möglich war.

Tabelle 3.7: Anteile ermittelter Geschlechter für die einzelnen Datensätze.

Datensatz / Disziplin	# Autoren	männlich	weiblich	unbekannt
EconLit	381.964	54,8 %	24,5 %	20,7 %
EconLit (Top)	17.390	72,4 %	14,6 %	13,0 %
JSTOR				
BWL und Finance	31.227	69,5 %	13,0 %	17,5 %
Geographie und Geologie	5.432	57,7 %	12,1 %	30,2 %
Mathematik	35.463	54,2 %	8,3 %	37,5 %
Psychologie	15.686	48,4 %	41,9 %	9,7 %
<i>Science</i>	135.747	55,1 %	18,4 %	26,5 %

Bem.: Die Zahlen für JSTOR und *Science* beziehen sich jeweils auf den Zeitraum ab 1960.

Der EconLit-Datensatz und die Disziplin *Mathematik* aus dem JSTOR-Datensatz wurden ferner, mittels manueller Recherche, um weitere individuelle Informationen zu den einzelnen Autoren ergänzt. Für den EconLit-Datensatz zählen zu den zusätzlichen Informationen:

- Zugehörigkeit zu einer nicht-akademischen Institution,
- Zugehörigkeit zu einer Top-Institution,
- Migrationshintergrund,
- US-Staatsbürgerschaft,
- Geburtsjahr.

Die Publikationen der Disziplin *Mathematik* im JSTOR-Datensatz wurden lediglich um die Geburtsjahre der Autoren ergänzt (vgl. folgenden Abschnitt 3.5.2). Der EconLit-Datensatz enthält Informationen zur auf einem Artikel angegebenen Instituts-

³⁸Weitere Informationen und Downloadlinks zu den Daten finden sich unter <https://www.ssa.gov/OACT/babynames/limits.html>.

zugehörigkeit eines Autors. Auf Basis dieser Information wurden die einzelnen Autoren *nicht-akademischen* Institutionen bzw. *Top-Institutionen* zugeordnet. Zu den nicht-akademischen Institutionen gehören solche, deren primärer Fokus nicht auf Forschung liegt. Hierzu zählen Regierungen, kommerzielle und nicht-kommerzielle Organisationen, wie Ausschüsse, Räte, Banken, Unternehmen, Fonds und Stiftungen. Im Gegensatz dazu beschäftigen sich akademische Institutionen vornehmlich mit Forschung und Lehre bzw. Ausbildung, so wie Universitäten, Forschungsinstitute, Schulen und Akademien.

Die Zugehörigkeit zu einer Top-Institution wurde anhand einer Liste der 50 renommiertesten Institutionen weltweit bestimmt (vgl. S. 200). Diese Liste wurde auf Basis der Ausgaben von 2015 und 2016 des *QS World University Rankings for Social Sciences and Management* und des *RePEc/IDEAS Ranking of Economics Institutions* abgeleitet.³⁹ Das zweite Ranking wurde herangezogen, da es auch nicht-akademische Institutionen wie den internationalen Währungsfonds und die Weltbank umfasst. Die Informationen zum Migrationshintergrund eines Autors und dazu, ob er die US-Staatsbürgerschaft besitzt, sind nur für Autoren verfügbar, zu denen auch das Geburtsjahr bekannt ist. Sie wurden beispielsweise den Lebensläufen der Autoren oder entsprechenden Wikipedia-Artikeln entnommen.

3.5.2 Altersdaten

Wie in der Einleitung bereits beschrieben, ist eines der Ziele dieser Arbeit herauszufinden, ob es Indikatoren dafür gibt, dass die zunehmende Wissensmenge und die zunehmende Komplexität von Forschungsprojekten sich in irgendeiner Weise auf die Karrieren von Wissenschaftlern auswirken. Eine Information, die in diesem Zusammenhang betrachtet wird, ist das Alter der Autoren bei ihrem ersten Artikel. Um dieses zu bestimmen, bedarf es des Geburtsjahres der Autoren und des Zeitpunkts ihres ersten Artikels. Da das Alter der Autoren nicht in den oben genannten Datenbanken verfügbar ist, wurde dieses aus öffentlich verfügbaren, externen Quellen ergänzt. Zu diesen Quellen gehören Lebensläufe der Autoren, ihre Online-Profile auf ihren privaten oder universitären Webseiten, Wikipedia-Artikel und einschlägige Online-Suchmaschinen wie prabook.com, intelius.com und birthdatabase.com. Um zu validieren, dass es sich um die richtige Person handelt, wurden Informationen wie Titel von Artikeln, Arbeitsplatz und Wohnort eines gefundenen Autors, mit den Informationen im Datensatz abgeglichen.

³⁹Die Rankings finden sich unter <https://www.topuniversities.com/university-rankings/faculty-rankings/social-sciences-and-management/2015/> bzw. <http://ideas.repec.org/top/>.

An der manuellen Recherche von Geburtsjahren waren unterschiedliche Personen beteiligt. Neben dem Verfasser dieser Arbeit und Sascha Schweitzer, haben Studierende der Universität Bayreuth, im Rahmen ihrer Abschlussarbeiten oder als studentische Hilfskräfte, bei der Suche nach Geburtsdaten mitgewirkt.⁴⁰ Die Suche im Bereich der Volkswirtschaftslehre war dabei zweigeteilt, da in einem ersten Schritt zunächst nach den Geburtsdaten von Top-Autoren (vgl. Abschnitt 3.1, S. 40) gesucht wurde. Hierzu wurde nach Autoren einer zufälligen Stichprobe von 5.000 der 17.390 Top-Autoren gesucht. In einem zweiten Schritt wurden zwei weitere zufällige Stichproben mit 3.000 Top-Autoren und 12.000 Autoren, die nicht in einer der renommiertesten Zeitschriften publiziert haben, betrachtet. Ferner wurde eine zufällige Stichprobe von 8.000 Autoren der *Mathematik* recherchiert.⁴¹ Dabei wurden für die *Mathematik* Autoren von Artikeln, die ab 1950 veröffentlicht wurden, bei der Suche berücksichtigt (für den EconLit-Datensatz waren es Autoren von Artikeln ab 1970). Auf diese Weise konnten Geburtsjahre von 2.656 Autoren der Volkswirtschaftslehre (1.635 Top-Autoren und 1.021 nicht Top-Autoren) und 840 Autoren der Mathematik ermittelt werden.

Tabelle 3.8: Verteilung von Top-Autoren bei ihren ersten Artikeln und von Top-Autoren mit Geburtsdaten auf die Top-Zeitschriften im EconLit-Datensatz.

Zeitschrift	Autoren mit erstem Artikel		Autoren mit Geburtsjahr	
	Anzahl	Anteil	Anzahl	Anteil
American Economic Review	4.106	23,6 %	371	22,7 %
Econometrica	1.487	8,6 %	178	10,9 %
Economic Journal	1.887	10,9 %	132	8,1 %
International Economic Review	1.283	7,4 %	99	6,1 %
Journal of Finance	2.805	16,1 %	202	12,4 %
Journal of Political Economy	1.164	6,7 %	173	10,6 %
Quarterly Journal of Economics	996	5,7 %	144	8,8 %
Review of Economic Studies	944	5,4 %	115	7,0 %
Review of Economics and Statistics	2.718	15,6 %	221	13,5 %

Tabelle 3.8 zeigt, wie viele der 1.635 Top-Autoren ihren ersten Artikel in welcher der Top-Zeitschriften veröffentlicht haben (Spalte 2), den Anteil dieser Autoren an allen

⁴⁰Die studentischen Hilfskräfte waren am Lehrstuhl für Technologie- und Innovationsmanagement von Professor Stefan Seifert angestellt, an dem der Verfasser dieser Arbeit zu dem Zeitpunkt beschäftigt war.

⁴¹Auch wenn die Zeitschrift *Econometrica* bei JSTOR bzw. im JSTOR-Datensatz der *Mathematik* zugeordnet ist, wurden bei der Altersdatensuche für die Disziplin *Mathematik* keine Autoren berücksichtigt, die ihren ersten Artikel in dieser Zeitschrift veröffentlicht haben. Eine Auswertung für *Econometrica* geschieht bereits im Rahmen des EconLit-Datensatzes.

Top-Autoren (Spalte 3), für jede Top-Zeitschrift die Anzahl der Autoren zu denen das Geburtsjahr bekannt ist (Spalte 4) und deren Anteil an allen ermittelten Geburtsjahren (Spalte 5). Die ersten Artikel der 2.656 Autoren, zu denen das Geburtsjahr bekannt ist, verteilen sich auf 592 verschiedene Zeitschriften im gesamten EconLit-Datensatz. Auf eine Auflistung nach Zeitschriften wie in Tabelle 3.8 wird daher verzichtet. Der Anteil weiblicher Autoren im gesamten EconLit-Datensatz liegt bei 24,5 % und für Top-Autoren bei 14,6 %. Von allen Autoren bzw. Top-Autoren, zu denen das Geburtsjahr ermittelt werden konnte, sind 12,9 % bzw. 11,9 % weiblich. In dieser Hinsicht erscheint die Stichprobe der Top-Autoren, zu denen das Alter bekannt ist repräsentativ für die Gesamtheit der Top-Autoren im EconLit-Datensatz zu sein. Analoge Informationen wie für die Top-Autoren im EconLit-Datensatz in Tabelle 3.8 finden sich für Autoren der *Mathematik* in Tabelle 3.9. Der Anteil weiblicher Autoren mit ersten Artikeln ab 1950 liegt bei 8,5 %, während der Anteil weiblicher Autoren, zu denen das Alter ermittelt werden konnte, bei 6,0 % liegt. Die größeren Diskrepanzen für den gesamten EconLit-Datensatz und die Mathematiker rühren vermutlich aus den verhältnismäßig geringeren Stichprobenumfängen von Autoren, zu denen das Alter ermittelt werden konnte, im Vergleich zu den Top-Autoren.

Tabelle 3.9: Verteilung von Autoren bei ihren ersten Artikeln und von Autoren mit Geburtsdaten auf Zeitschriften der *Mathematik*.

Zeitschrift	Autoren mit erstem Artikel		Autoren mit Geburtsjahr	
	Anzahl	Anteil	Anzahl	Anteil
American Journal of Mathematics	1.667	4,8 %	131	15,5 %
Annals of Mathematics	1.473	4,2 %	203	24,1 %
Journal of the Am. Mathematical Society	444	1,3 %	42	5,0 %
Mathematics of Computation	3.488	10,0 %	64	7,6 %
Mathematics of Operations Research	1.233	3,6 %	17	2,0 %
Operations Research	4.626	13,3 %	18	2,1 %
Proc. of the Am. Mathematical Society	12.617	36,3 %	252	29,9 %
SIAM Journal on Applied Mathematics	4.257	12,3 %	57	6,8 %
SIAM Journal on Numerical Analysis	3.004	8,6 %	40	4,7 %
SIAM Review	1.667	4,8 %	16	1,9 %
The Bulletin of Symbolic Logic	256	0,7 %	3	0,4 %

Bem.: Die Abkürzungen Am. bzw. Proc. stehen für American bzw. Proceedings.

3.6 Untersuchte Maße

Bei der Betrachtung der in dieser Arbeit untersuchten Maße können zwei Ebenen unterschieden werden, die sich aus der Form der Aggregation der Daten ergeben. Einerseits ist das die *Artikel-Ebene*, die sich gewissermaßen aus den „Eigenschaften“ eines Artikels ergibt. Hierzu zählen beispielsweise etablierte bibliometrische Maße, wie die *Anzahl der Quellenangaben* in einem Artikel (z. B. Bornmann und Mutz 2015) oder die *Anzahl der Autoren* eines Artikels (z. B. Card und DellaVigna 2013; Hamermesh 2013). Diese Ebene dient unter anderem dazu, mögliche Trends in der Welt des wissenschaftlichen Publizierens aufzuzeigen, wie beispielsweise Veränderungen in der Zusammenarbeit von Wissenschaftlern.

Andererseits werden Maße auf *Autoren-Ebene* betrachtet. Hier werden die Daten somit nicht auf Artikel-Ebene aggregiert, sondern auf einzelne Autoren heruntergebrochen. Dieser Teil der Untersuchung zielt darauf ab, die Hypothesen der Last des Wissens von Jones (2009a) für Autoren wissenschaftlicher Artikel zu prüfen. Zu den betrachteten Maßen gehören beispielsweise das Alter eines Autors bei seinem ersten Artikel oder die Anzahl der Quellen, die er bei seinem ersten Artikel angibt. Insofern werden manche Größen sowohl auf Autoren-Ebene, als auch auf Artikel-Ebene betrachtet (z. B. Anzahl der Quellenangaben). Im weiteren Verlauf der Arbeit werden die folgenden Maße betrachtet:

Maß	Beschreibung
Teamgröße	Anzahl der Autoren, die als Autoren eines Artikels gelistet sind.
Anzahl Quellenangaben	Anzahl der Quellen, die im Literaturverzeichnis eines Artikels gelistet sind.
Länge des Titels	Anzahl der Wörter im Titel eines Artikels.
Länge des Abstracts	Anzahl der Wörter im Abstract eines Artikels.
Alter bei erstem Artikel	Alter eines Autors bei seinem ersten Artikel.
Spezialisierung	Relative Häufigkeit von JEL-Feld-Wechseln zwischen den beiden ersten Artikeln von Autoren, bei denen sie alleinige Autoren sind.

Die Teamgröße, also die Anzahl der Autoren, die an einem Artikel gearbeitet haben und als Autoren dieses Artikels gelistet sind, dient als Maß dafür, wie viele (spezialisierte) Wissenschaftler es braucht, um die Breite an Wissen und den entsprechenden Aufwand bzw. die entsprechende Arbeit abzudecken, um einen wissenschaftlichen Artikel zu schreiben. Beispielsweise argumentieren McDowell und Melvin (1983), Barnett, Ault und Kaserman (1988) sowie Jones (2009a), dass ein Grund für Kooperation unter Wissenschaftlern die Möglichkeit zur Aufteilung eines Forschungsproblems unter den Teammitgliedern ist.

Die Anzahl der Quellenangaben dient als grobe Schätzung für das Wissen, das verarbeitet wurde, um den betrachteten Artikel zu schreiben bzw. zu den Erkenntnissen zu gelangen, die in diesem Artikel aufgeschrieben sind. Wie in Kapitel 3 beschrieben, ist die Information zur Anzahl der Quellenangaben nicht für alle Artikel verfügbar. Bei den Auswertungen zur Anzahl der Quellenangaben wurden daher lediglich Artikel mit mindestens einer Quellenangabe berücksichtigt.⁴² Auf Basis dieses Maßes lässt sich nicht zwischen Breite und Tiefe des Wissens unterscheiden, das sich hinter den angegebenen Quellen verbirgt, da aus den vorhandenen Informationen nicht hervorgeht, ob mehrere Quellen dazu dienen ein möglichst breites Wissen abzudecken oder ein spezielles Thema zu vertiefen. Im weiteren Verlauf der Arbeit wird teilweise auch im Text kurz „Anzahl Quellenangaben“ statt „Anzahl der Quellenangaben“ geschrieben, wenn der Charakter des Maßes stärker betont werden soll.

Titel und Abstracts von Artikeln sind wichtige Anhaltspunkte für die Inhalte eines Artikels. Der Titel ist für gewöhnlich das erste Element eines Artikels, das der Leser zu sehen bekommt und stellt daher den ersten Anhaltspunkt für die Einordnung des Artikels und seiner Inhalte dar. Während die Informationen bzgl. der Inhalte eines wissenschaftlichen Artikels im Titel stark komprimiert sind, finden sich hierzu, in Form einer Zusammenfassung, im Abstract detailliertere Informationen. Diener (1984), der Titel von wissenschaftlichen Artikeln hinsichtlich ihres Informationsgehalts untersuchte, schreibt über diese: „They constitute the most concise statement of the content of a document.“ (S. 222). Bereits in den 1970er Jahren zeigten Tocatlian (1970) sowie Buxton und Meadows (1977) für Artikel der Chemie sowie Buxton und Meadows (1977) für Artikel der Botanik, dass der Informationsgehalt der zugehörigen Titel, gemessen

⁴²In Artikeln, die laut Datensatz null Quellenangaben aufweisen, wurden im Artikel zwar typischerweise auch Quellen genannt, allerdings nicht in systematischer Weise in Form eines Literaturverzeichnisses zusammengefasst. Entsprechend ist die Anzahl der Quellenangaben für diese Artikel nicht verfügbar.

an der Anzahl der Schlüsselwörter, zugenommen hat.⁴³ White und Hernandez (1991) argumentieren, dass die Länge des Titels mit der Komplexität des zugehörigen Artikels korreliert. Für verschiedene Zeitschriften aus den Bereichen Pädagogik, Psychologie, Soziologie, sowie Volkswirtschaftslehre zeigen sie auf Basis der Anzahl Schlüsselwörter im Titel, dass die Komplexität wissenschaftlicher Artikel zugenommen hat.⁴⁴ Vor dem Hintergrund einer möglichen Last des Wissens, die sich u. a. durch zunehmende Komplexität bei der Wissensverarbeitung und -generierung äußert, werden daher in Abschnitt 4.3 die Titel der in dieser Arbeit untersuchten Artikel im Hinblick auf ihre Länge näher untersucht.⁴⁵ Die Länge der Titel entspricht dabei der Anzahl der Wörter (vgl. z. B. Diener 1984).⁴⁶ Damit stellt das verwendete Maß ein etwas gröberes Maß dar, als es beispielsweise von White und Hernandez (1991) oder Yitzhaki (1994) verwendet wird, die die Anzahl der Schlüsselwörter zählen.

Ein Abstract fasst die Inhalte (z. B. Thematik, Methode, Ergebnisse) des zugehörigen Artikels zusammen. Entsprechend sollten sich die Informationen eines Titels auch im Abstract widerspiegeln. Tatsächlich kommt Diodato (1982) zu dem Ergebnis, dass Abstracts den Großteil der Informationen eines Titels enthalten. Konkret untersucht Diodato (1982), welche der Substantive aus dem Titel auch im Abstract des Artikels vorkommen. Somit kann die Länge des Abstracts als ein zur Länge des Titels verwandtes Maß angesehen werden und ggf. tiefere Erkenntnisse liefern, als eine isolierte Betrachtung des Titels. Daher wird in Abschnitt 4.3 zusätzlich zur Länge der Titel auch die Länge der Abstracts betrachtet.⁴⁷ Für die Länge von Abstracts gibt es seitens der Zeitschriften meist explizite Vorgaben. Abstracts im *American Economic Review* sollten nicht länger sein als 100 Wörter, während die Abstracts in *Econometrica* bis zu 150 Wörter lang sein dürfen.⁴⁸ Gängige Empfehlungen zum Verfassen von Abstracts empfehlen diese kurz und einfach zu gestalten (vgl. Weinberger, Evans und Allesina 2015). Letchford, Preis und Moat (2016) zeigen für die laut *Web of Science* am häufigsten zitierten Artikel, dass ein längerer Abstract tendenziell zu weniger Zitationen führt. Entsprechend sollten Autoren darauf bedacht sein die Zusammenfassung des Inhalts

⁴³Schlüsselwörter, auch substantielle Wörter genannt, sind dabei als Wörter definiert, die einen hohen Informationsgehalt aufweisen. Sie werden ermittelt, indem man nicht-substantielle Wörter eines Textes, wie Artikel, Präpositionen, Konjunktionen oder Hilfsverben, streicht (Tocatlian 1970, S. 346).

⁴⁴Ihre Stichprobe umfasst die volkswirtschaftlichen Zeitschriften *Economic Journal* und *Journal of Political Economy*.

⁴⁵In die Untersuchung gehen Titel mit mindestens drei Wörtern ein.

⁴⁶Die Länge der Titel wurde zusätzlich auch mittels der Anzahl Zeichen bemessen. Die Ergebnisse für dieses Maß decken sich weitestgehend mit denen für die Anzahl der Wörter.

⁴⁷In die Untersuchung gehen Abstracts mit mindestens zehn Wörtern ein.

⁴⁸Diese Informationen stammen aus den Leitfäden zu Einreichungen der Zeitschriften auf den entsprechenden Webseiten (Stand: 10.06.2018).

und der Ergebnisse ihres Artikels im Abstract kurz und prägnant zu halten. Zunehmend lange Abstracts könnten daher ein Anzeichen für immer längere und komplexere Artikel bzw. Inhalte sein, die Autoren nicht mehr ausreichend im bisherigen Umfang von Abstracts zusammenfassen können.⁴⁹ Es sind allerdings auch weitere Erklärungen denkbar: Weinberger, Evans und Allesina (2015), die Artikel verschiedener Disziplinen aus Zeitschriften unterschiedlicher Qualität untersuchen, kommen beispielsweise zu dem Ergebnis, dass längere und kompliziertere Abstracts zu mehr Zitationen führen.⁵⁰ Die Autoren mutmaßen, dass Abstracts auch aus strategischen Gründen länger gefasst werden könnten. Diese zielen darauf ab, dass die eigene Arbeit, aufgrund von mehr Stichworten und Inhalten im Abstract, häufiger von anderen Autoren gefunden wird und auf deren Interesse stößt.

Das Alter eines Autors bei seinem ersten Artikel dient der Quantifizierung des Effekts der Last des Wissens auf die Karrieren von Wissenschaftlern. Es markiert den Übergang aus der Phase der Wissensverarbeitung in die aktive Phase der Wissensgenerierung. Unter der Annahme, dass sich Wissenschaftler eine zunehmend große Menge an Wissen aneignen müssen, bevor sie zum bestehenden Wissen beitragen können, könnte sich diese Last darin äußern, dass Autoren zum Zeitpunkt ihres ersten Artikels zunehmend älter werden (vgl. Jones 2009a, für eine analoge Überlegung bzgl. des Alters bei erster Patentanmeldung).

Je spezialisierter Autoren sind, desto schwieriger fällt es ihnen, ihr Feld und damit die JEL-Felder, die sich aus den auf den Artikeln angegebenen JEL-Codes ergeben (vgl. Abschnitt 3.1.2, S. 43), zu wechseln. Um die binäre Natur der Feldsprünge in Jones (2009a) nachzuahmen, wird ein *Feldsprung* als ein Wechsel aller angegebenen JEL-Felder definiert. Somit wird von einem Feldsprung gesprochen, wenn die Schnittmenge der JEL-Felder zwischen zwei aufeinander folgenden Publikationen leer ist. Da das Maß der Spezialisierung auf dem von Jones (2009a) basiert, weist es ähnliche Schwächen auf. Es ist nicht einfach, die Wahrscheinlichkeit für einen Feldsprung für einen Autor zu bestimmen, der mit anderen Autoren gemeinsam publiziert. Beispielsweise ist es denkbar, dass solche Autoren zwar immer spezialisierter werden, aber in zunehmend großen Teams arbeiten, in denen jeder Autor sein Wissen aus einem anderen Bereich beiträgt und die Spezialisierung in den entsprechenden JEL-Codes der Artikel nicht sichtbar wird. Solche Entwicklungen bezüglich der in einem Artikel angegebenen JEL-Codes,

⁴⁹Tatsächlich berichten beispielsweise Card und DellaVigna (2013) von zunehmend langen Artikeln.

⁵⁰Die Unterschiede zwischen den Ergebnissen von Letchford, Preis und Moat (2016) und Weinberger, Evans und Allesina (2015) könnten ggf. aus den qualitativen Unterschieden der jeweils betrachteten Artikel sowie den Vorgaben zur Länge von Abstracts seitens der untersuchten Zeitschriften rühren.

könnten das Bild eines zunehmend spezialisierten Autors (und damit eines Autors, der immer seltener seine JEL-Felder wechselt) verzerren und er könnte fälschlicherweise eine konstante Wahrscheinlichkeit eines Feldsprungs aufweisen. Daher konzentriert sich die Untersuchung der Spezialisierung in Abschnitt 4.5 (ab S. 135) auf Autoren, die alleine publizieren bzw. Artikel mit nur einem Autor (i. F. auch als *Solo-Artikel* bezeichnet). Die Regressionen für Autoren, die in Teams arbeiten, zeigen allerdings grundsätzlich die gleichen Effekte und sprechen somit für die Robustheit der Ergebnisse. Das Maß der Spezialisierung wird lediglich für Autoren aus der Volkswirtschaftslehre betrachtet, da für die entsprechenden Artikel JEL-Codes angegeben sind, aus denen sich wiederum die JEL-Felder ergeben (vgl. Abschnitt 3.1.2, S. 43). Für die anderen Disziplinen, die in dieser Arbeit betrachtet werden, gibt es entweder kein Äquivalent zu den JEL-Codes oder die Publikationsdaten umfassen diese Information nicht. Entsprechend lassen sich die Artikel nicht Teilbereichen (analog zu den JEL-Feldern) der betrachteten Disziplinen zuordnen, weshalb auch keine Wechsel zwischen Teilbereichen identifiziert werden können bzw. keine relativen Häufigkeiten für diese.

3.7 Kontrollvariablen bei den Regressionen

In den Regressionen bzgl. der einzelnen Maße wird für die folgenden Variablen kontrolliert:

Kontrollvariable	Typ	Beschreibung
Publikationsjahr	skalar	Jahr, in dem der jeweils betrachtete Artikel veröffentlicht wurde.
Teamgröße	skalar	Anzahl der Autoren eines Artikels.
Zeitschrift	kategorial, Dummies	Kontrollen für fixe Effekte (engl. <i>fixed effects</i>) der einzelnen Zeitschriften.
JEL-Feld	kategorial, Dummies	Kontrollen für fixe Effekte (engl. <i>fixed effects</i>) der einzelnen JEL-Felder.
Autorin	binär, Dummy	Wert 1, wenn der betrachtete Autor des Artikels eine Frau ist und 0 für Männer.
Migration	binär, Dummy	Wert 1, wenn der betrachtete Autor aus seinem Geburtsland ausgewandert ist und 0 sonst.
Nicht-US	binär, Dummy	Wert 1, wenn der betrachtete Autor nicht die US-Staatsbürgerschaft besitzt und 0 für US-Bürger.

Kontrollvariable	Typ	Beschreibung
Nicht-Aka	binär, Dummy	Wert 1 für Autoren von nicht akademischen Institutionen und 0 für Autoren akademischer Institutionen.
Top-Inst	binär, Dummy	Wert 1 für die Zugehörigkeit zu einer Top-Institution und 0 sonst.

Die Teamgröße bzw. die Anzahl der Autoren eines Artikels wird in Abschnitt 3.6 diskutiert. Erläuterungen zu JEL-Feldern und dazu, wie die Informationen zu den binären Kontrollvariablen erhoben wurden, finden sich in Abschnitt 3.1.2 (S. 43) bzw. Abschnitt 3.5 (S. 71). In den Regressionen wird ferner dafür kontrolliert, ob zum Beispiel alle Autoren innerhalb eines Teams weiblich sind oder alle einer Top-Institution angehören. Die obigen Kontrollvariablen übertragen sich entsprechend und finden sich in den Regressionstabellen unter den Begriffen *Autorinnen*, *Alle Nicht-US*, *Alle Nicht-Aka*, sowie *Alle Top-Inst*.

Grundsätzlich wurde bei der Darstellung der Ergebnisse auf Konsistenz geachtet. Koeffizienten, Standardabweichungen und R^2 -Werte in den Regressionstabellen sind auf drei Nachkommastellen gerundet. Findet sich in einer der Tabellen ein Eintrag mit dem Wert „0,000“, so bedeutet dies, dass der entsprechende Wert kleiner als 0,0005 ist. Informationen zu Signifikanzniveaus bei den Regressionen finden sich unterhalb der Regressionstabellen. Die Regressionsergebnisse beziehen sich, wenn nicht anders angegeben, auf die Beobachtungszeiträume 1970–2014 für den EconLit-Datensatz, 1960–2009 für den JSTOR-Datensatz und 1960–2011 für die Zeitschrift *Science*.

Kapitel 4

Ergebnisse der Datenauswertung

In diesem Kapitel werden die Ergebnisse aus der Analyse der in Kapitel 3 beschriebenen Publikations- und Autoren Daten vorgestellt. Die Untersuchungen bzgl. der einzelnen in Abschnitt 3.6 (S. 76) vorgestellten Maße sind zweigeteilt. So werden beobachtete Trends zunächst grafisch dargestellt und anschließend mittels Regressionsanalyse näher untersucht.¹ Dabei wurden für den Großteil der Analysen lineare Regressionsmodelle betrachtet.^{2,3} Lediglich für binäre abhängige Variablen (z. B. für das Maß der Spezialisierung) wurde eine Probit-Regression herangezogen. Ferner ist die Darstellung der grafischen Trends und der Regressionsergebnisse bzgl. eines jeden Maßes, entsprechend der Disziplinen *Volkswirtschaftslehre* (EconLit-Datensatz), *Business und Finance*, *Geographie und Geologie*, *Mathematik*, *Psychologie* (JSTOR-Datensatz) und der Artikel aus der Zeitschrift *Science* separat dargestellt.⁴ Für jeden der Datensätze werden die einzelnen Maße getrennt untersucht, wobei sich die Ergebnisse der einzelnen Datensätze und Disziplinen bzgl. des gleichen Maßes jeweils auch im gleichen Abschnitt finden.

¹Die Trendlinien in den Abbildungen dieser Arbeit entstehen durch Verbindungsgeraden zwischen den einzelnen Punkten für die Jahresmittelwerte.

²Die Regressionen wurden mit Hilfe der Statistiksoftware R berechnet.

³Da es sich bei den untersuchten Variablen vorwiegend um Zählvariablen handelt, die keine negativen Werte annehmen, wurden der Vollständigkeit halber jeweils auch negativ binomiale Regressionsmodelle gerechnet. Die qualitativen Ergebnisse bzgl. der Effektrichtungen und Signifikanzniveaus decken sich mit denen der linearen Regressionen.

⁴Der JSTOR-Datensatz umfasst zwar auch Artikel der Volkswirtschaftslehre, allerdings wird diese Disziplin bereits im Rahmen des EconLit-Datensatzes in größerem Umfang untersucht. Die Daten zur Volkswirtschaftslehre aus dem JSTOR-Datensatz wurden daher u. a. dazu verwendet den EconLit-Datensatz zu ergänzen, falls dort einzelne Informationen zu Artikeln fehlten. Beispielsweise umfasst der EconLit-Datensatz keine Abstracts, sodass bei der Untersuchung der Länge von Abstracts in der Volkswirtschaftslehre in Abschnitt 4.3 (S. 122) auf Abstracts aus dem JSTOR-Datensatz zurückgegriffen wurde.

4.1 Teamgröße

4.1.1 Teamgröße für alle Artikel

Zunächst wird die Teamgröße eines Artikels betrachtet, das heißt die Anzahl der Autoren pro Artikel. Abbildungen 4.1 bis 4.3 veranschaulichen die zeitlichen Trends. Für alle Datensätze und Disziplinen ist die Teamgröße im Beobachtungszeitraum deutlich gestiegen. Grob gesagt steigt die mittlere Teamgröße im beobachteten Zeitraum von etwas mehr als einem Autor auf etwa zweieinhalb Autoren. Betrachtet man alle volkswirtschaftlichen Artikel des EconLit-Datensatzes, so steigt die mittlere Teamgröße eines Artikels im Zeitraum zwischen 1970 und 2014 von 1,3 auf 2,1, während sie für Top-Artikel insgesamt etwas höher liegt und im gleichen Zeitraum von 1,3 auf 2,3 steigt (vgl. Abbildung 4.1).⁵ Dies entspricht relativen Anstiegen von 63,8 % bzw. 76,9 %.⁶

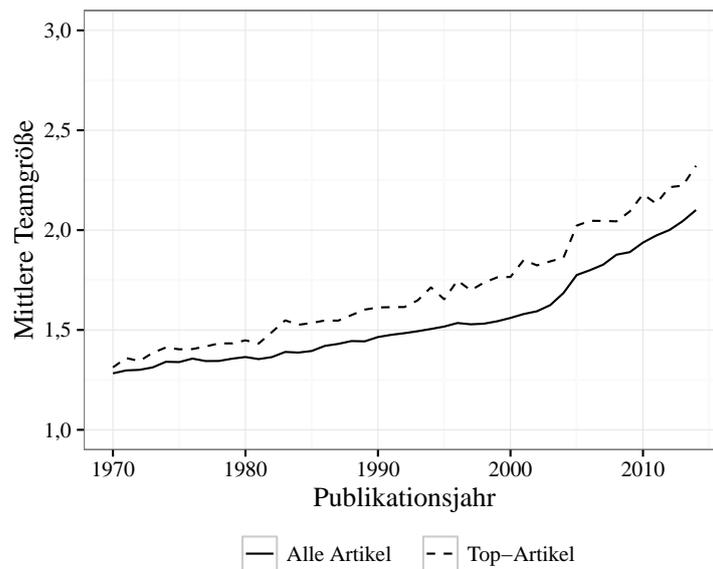


Abbildung 4.1: Mittlere Teamgröße im EconLit-Datensatz.

In *Betriebswirtschaftslehre und Finance* verdoppelt sich die mittlere Anzahl der Autoren pro Artikel und steigt im betrachteten 50-Jahres-Zeitraum von 1,2 im Jahr 1960 auf 2,3 im Jahr 2009. Der relative Anstieg der mittleren Teamgröße von 101,4 % für diese Disziplin ist damit für die Disziplinen im JSTOR-Datensatz am größten.

⁵In der vorliegenden Arbeit wird die Bezeichnung „alle Artikel“ für die Beschreibung zweier unterschiedlicher Mengen von Artikeln verwendet. Einerseits wird für den EconLit-Datensatz zwischen *allen Artikeln* und *Top-Artikeln* unterschieden. Andererseits wird für alle Datensätze zwischen *allen* und *ersten Artikeln* der Autoren unterschieden. Aus dem jeweiligen Kontext sollte hervorgehen, welche Artikel gemeint sind.

⁶Hier und im Folgenden wurden Werte, wie der prozentuale Anstieg, die sich aus mehreren anderen Werten ergeben, nicht aus den gerundeten Angaben im Text berechnet, sondern mit den exakten Werten.

Auch Artikel in der *Psychologie* erfahren einen starken Anstieg in der Teamgröße. Die mittlere Anzahl der Autoren pro Artikel steigt hier im Beobachtungszeitraum von 1,7 auf 3,0 (relativer Anstieg von 74,2 %). Die *Psychologie* weist damit innerhalb des JSTOR-Datensatzes den größten absoluten Anstieg in der mittleren Teamgröße auf. Ähnlich fällt auch die relative Zunahme in der Teamgröße für die *Mathematik* aus, für die ein Zuwachs von 1,2 im Jahr 1960 auf 2,0 im Jahr 2009 zu beobachten ist (67,9 %). Für Artikel aus der *Geographie und Geologie* fällt sowohl der absolute als auch der relative Anstieg am geringsten aus. Die mittlere Teamgröße steigt im Beobachtungszeitraum lediglich von 1,1 auf 1,4 (20,6 %). Die geringeren mittleren Teamgrößen in der *Geographie und Geologie* im Vergleich zu den anderen Disziplinen im JSTOR-Datensatz sind in Abbildung 4.2 an den niedrigeren Durchschnittswerten deutlich erkennbar. Der strukturelle Bruch für *Geographie und Geologie*, mit einer deutlich erhöhten mittleren Teamgröße zwischen 1980 und Mitte der 1990er Jahre ist auf die deutlich höheren Teamgrößen für die Zeitschrift *The Journal of Geology* in diesem Zeitraum zurückzuführen. Während die mittlere Teamgröße für alle Zeitschriften in *Geographie und Geologie* über den Beobachtungszeitraum steigt, ist der Anstieg für das *Journal of Geology* in diesen Jahren besonders hoch und beeinflusst entsprechend die mittlere Teamgröße der ganzen Disziplin (vgl. Abbildung B.4, S. 210, im Anhang). Ab 1998 sind keine Artikel aus dieser Zeitschrift im Datensatz verfügbar, weshalb die mittlere Teamgröße wieder auf ein niedrigeres Niveau sinkt. Auch für Artikel in der Zeitschrift *Science* ist im Laufe der Jahre ein deutlicher Anstieg der Teamgröße zu beobachten (Abbildung 4.3). Während es im Jahr 1960 noch durchschnittlich 1,9 Autoren pro Artikel sind, verdoppelt sich dieser Wert bis zum Jahr 2000 fast auf 3,6 (relativer Anstieg von 93,3 %) und steigt bis zum Jahr 2011 auf 5,2 (179,1 %).

Betrachtet man die absoluten Anzahlen von Artikeln nach Teamgrößen, so ist leicht ersichtlich, warum die mittlere Teamgröße in den Beobachtungszeiträumen gestiegen ist. In Abbildung 4.4 sind beispielhaft die Top-Artikel aus dem EconLit-Datensatz in einem Histogramm dargestellt, wobei die Balken unterschiedliche Teamgrößen repräsentieren und der Untersuchungszeitraum in drei konsekutive 15-Jahres-Zeiträume aufgeteilt ist. Wie man sieht, finden sich in den ersten 30 Jahren (1970–1999), also den ersten beiden 15-Jahres-Zeiträumen, keine Artikel mit vier oder mehr Autoren, in den letzten 15 Jahren allerdings schon. Ferner ist zu erkennen, dass die Anzahl der Artikel mit zwei und drei Autoren zunimmt und die Anzahl der Artikel mit nur einem Autor zurückgeht. Analoge Histogramme für alle Artikel im EconLit-Datensatz, für die Disziplinen im JSTOR-Datensatz und die Zeitschrift *Science* finden sich in den Abbildungen B.1 bis B.3 im Anhang (S. 210 ff.).

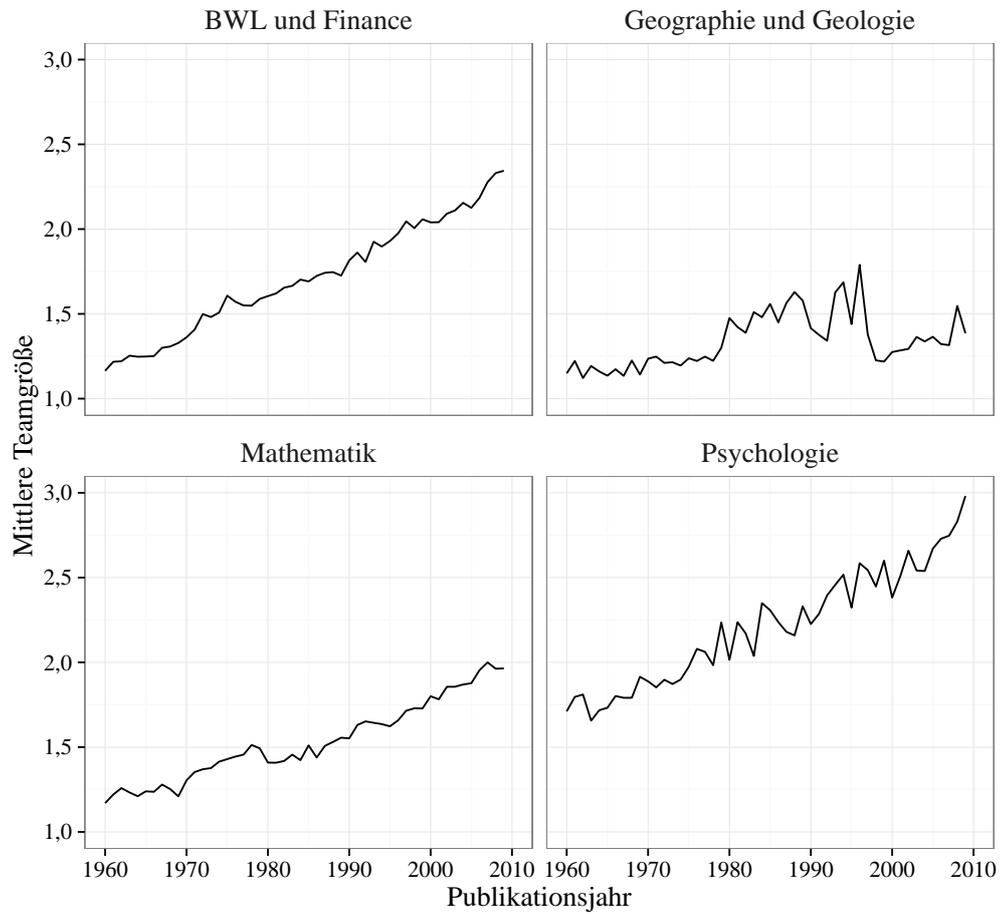


Abbildung 4.2: Mittlere Teamgröße im JSTOR-Datensatz.

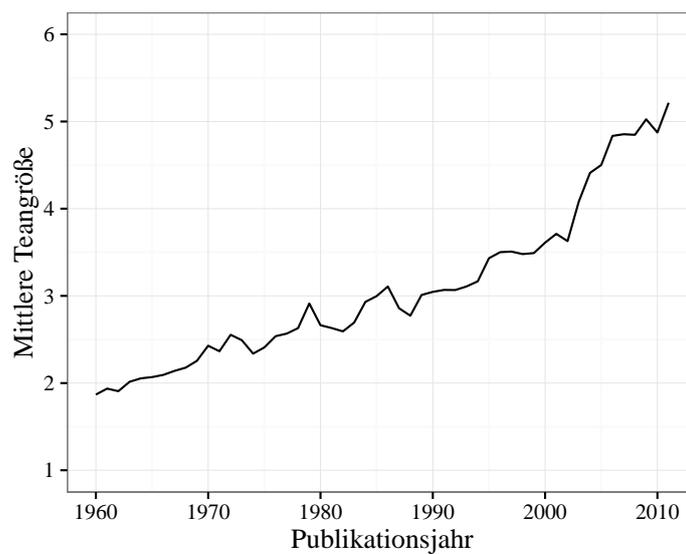


Abbildung 4.3: Mittlere Teamgröße in der Zeitschrift *Science*.

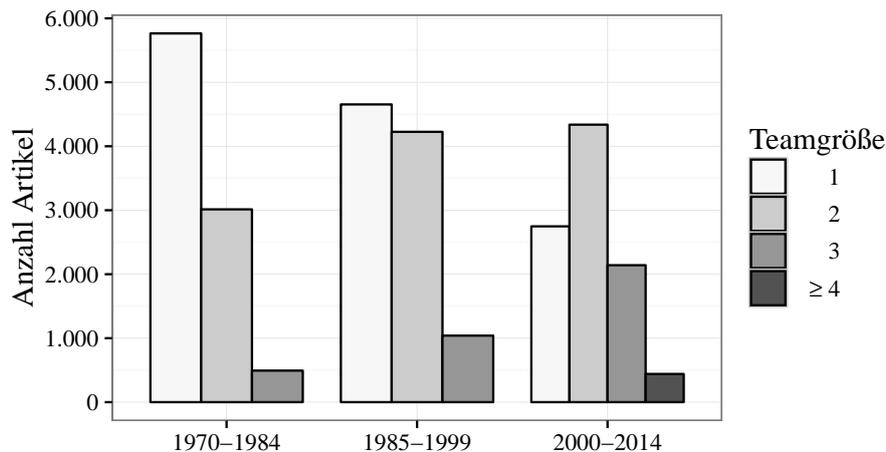


Abbildung 4.4: Anzahl der Top-Artikel im EconLit-Datensatz nach Teamgrößen.

Die absoluten Zahlen aus Abbildung 4.4 geben einen ersten Eindruck über die Verhältnisse der einzelnen Teamgrößen, allerdings unterscheidet sich die Gesamtzahl der Artikel für die drei Teilzeiträume. Daher sind in Abbildung 4.5 die relativen Anteile der einzelnen Teamgrößen-Gruppen an der Gesamtzahl der Top-Artikel pro Jahr abgetragen. Man sieht deutlich die fallende durchgezogene Linie, die für Top-Artikel mit lediglich einem Autor steht. Der Anteil dieser Artikel fällt von 71,7 % im Jahr 1970 auf 20,4 % im Jahr 2014 und sinkt im Beobachtungszeitraum somit jährlich um mehr als einen Prozentpunkt. Im Jahr 1994 liegt ihr Anteil erstmals unter dem der Artikel mit zwei Autoren. Ferner ist der Anteil Artikel mit nur einem Autor gegen Ende des Beobachtungszeitraums (ab dem Jahr 2012) sogar geringer, als der Anteil der Artikel mit drei Autoren. Gleichzeitig ist zu erkennen, wie die Anteile der anderen drei Gruppen steigen. Ab der Jahrtausendwende sinkt nicht nur der Anteil von Artikeln mit lediglich einem Autor, sondern auch der Anteil der Top-Artikel mit zwei Autoren ist rückläufig und stattdessen nehmen Artikel mit mindestens drei Autoren zu. Während der Anteil der Top-Artikel mit zwei Autoren im Jahr 1970 noch 25,3 % beträgt, erfährt diese Gruppe im Jahr 1999 ein Hoch und macht mit 50,9 % mehr als die Hälfte der Top-Artikel aus. Danach fällt ihr Anteil bis 2014 auf 39,9 %.

Für alle volkswirtschaftlichen Artikel im EconLit-Datensatz, das heißt ohne die Einschränkung auf Artikel aus den renommiertesten Zeitschriften, ergibt sich ein analoges Bild, auch wenn einzelne kleinere Unterschiede erkennbar sind (Abbildung 4.6). Die Anteile verlaufen ähnlich wie für Top-Artikel, allerdings fällt beispielsweise der Anteil der Artikel mit nur einem Autor von 74,8 % im Jahr 1970 auf 33,4 % im Jahr 2014 nicht ganz so stark wie für Top-Artikel. Ferner ist optisch kaum ein Rückgang des Anteils von Artikeln mit zwei Autoren (gepunktete Linie) auszumachen, der für

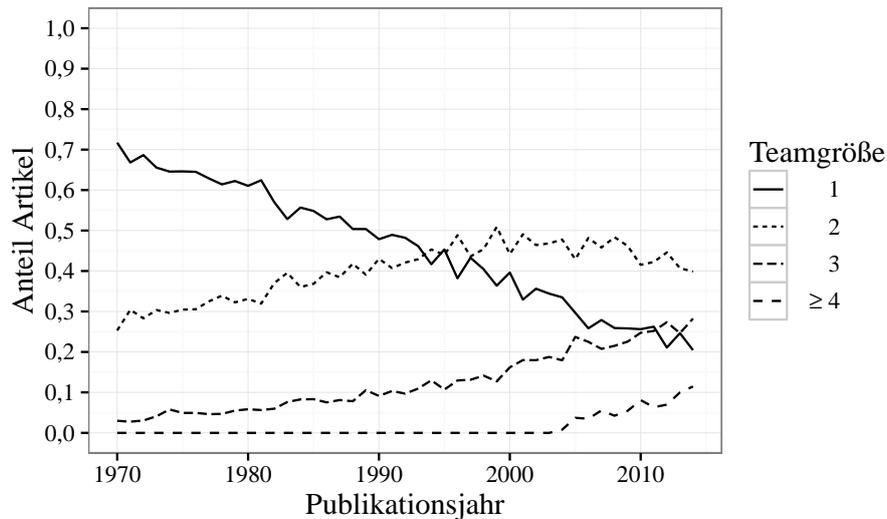


Abbildung 4.5: Anteile der Teamgrößen für Top-Artikel im EconLit-Datensatz.

Top-Artikel ab dem Jahr 1999 klar zu erkennen ist. Tatsächlich flacht die Kurve der Anteile von Artikeln mit zwei Autoren in Abbildung 4.6 leicht ab, erreicht im Jahr 2010 ihren Hochpunkt mit einem Anteil von 36,0 % und fällt bis zum Jahr 2014 leicht auf einen Anteil von 35,2 %. Im Jahr 2014 ist der Anteil von Artikeln mit zwei Autoren erstmals größer als der Anteil der Artikel mit nur einem Autor.

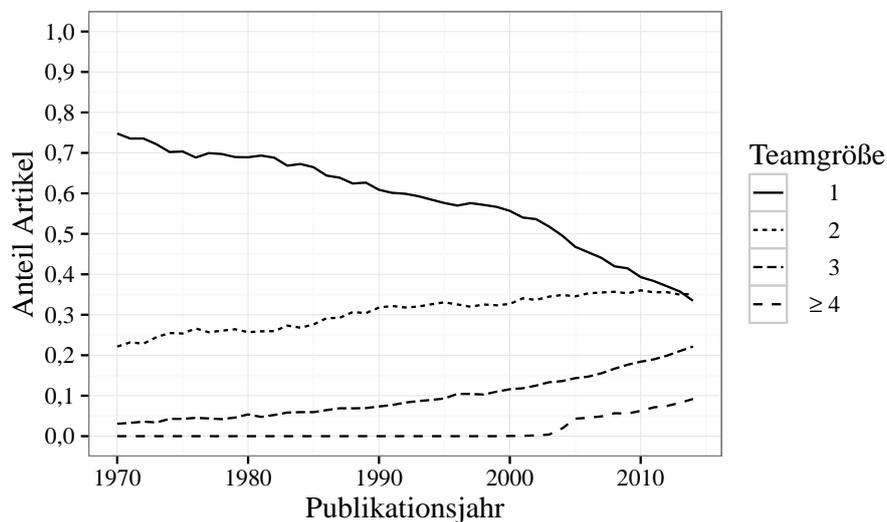


Abbildung 4.6: Anteile der Teamgrößen im EconLit-Datensatz.

Auch für die unterschiedlichen wissenschaftlichen Disziplinen im JSTOR-Datensatz zeigen sich weitestgehend die gleichen Trends (Abbildung 4.7). Auch hier fallen die Anteile der Artikel mit nur einem Autor im Beobachtungszeitraum teilweise sehr stark. Für *BWL* und *Finance* ist dieser Rückgang besonders deutlich. Der Anteil von Solo-

Artikeln fällt von 86,3 % im Jahr 1960 auf 16,5 % im Jahr 2009. Für Artikel der *Mathematik* sinkt der Anteil von Solo-Artikeln im gleichen Zeitraum von 84,1 % auf 34,4 %. In der *Psychologie* ist der Rückgang geringer, da der Anteil bereits im Jahr 1960 nur 52,6 % beträgt und bis zum Jahr 2009 auf 15,6 % fällt. Wie weiter oben bereits beschrieben und in Abbildung 4.2 (S. 86) zu sehen, liegt die mittlere Teamgröße für Artikel aus der *Geographie und Geologie* deutlich unter der der anderen Disziplinen. Ein Grund hierfür, ist in Abbildung 4.7 ersichtlich. Solo-Artikel machen über den gesamten Beobachtungszeitraum den größten Anteil aus und sorgen für einen insgesamt niedrigeren Mittelwert bzgl. dieses Maßes. In dieser Disziplin ist auch der geringste Rückgang von Artikeln mit nur einem Autor zu verzeichnen. Zwischen 1960 und 2009 fällt deren Anteil von 85,1 % auf 74,3 %.

Ähnliche Trends wie für den EconLit-Datensatz bzgl. des Anteils von Artikeln mit zwei Autoren sind auch für die Disziplinen im JSTOR-Datensatz zu beobachten. Deren Anteil steigt zu Beginn des Beobachtungszeitraums. Allerdings flacht dieser Anstieg wie im Falle der *Mathematik* ab oder ihr Anteil geht sogar zurück. Für *BWL und Finance* beginnt dieser Rückgang, ähnlich wie für Top-Artikel im EconLit-Datensatz, um die Jahrtausendwende und in der *Geographie und Geologie* sowie *Psychologie* bereits ab Anfang der 1980er Jahre. Interessant sind darüber hinaus die Unterschiede bzgl. der Artikel mit vier oder mehr Autoren, verglichen mit den Resultaten für den EconLit-Datensatz. In letzterem findet sich der erste Top-Artikel mit mindestens vier Autoren erst im Jahr 2003. Der erste Artikel mit so vielen Autoren im gesamten EconLit-Datensatz ist zwar bereits im Jahr 1989 zu beobachten, allerdings ist dies auch nur ein einziger Artikel und die Zahl solcher Artikel bzw. deren Anteil bleibt bis zum Jahr 2004 verschwindend gering (z. B. 2 Artikel im Jahr 1999 und 9 Artikel im Jahr 2000). Wie man Abbildung 4.7 entnimmt, liegen die Anteile von Artikeln mit vier oder mehr Autoren für die Disziplinen im JSTOR-Datensatz, im Gegensatz zum EconLit-Datensatz (Abbildungen 4.5 und 4.6), über den größten Teil des Beobachtungszeitraums nicht bei 0 %. Sowohl für *BWL und Finance* als auch *Mathematik* sind solche Artikel bereits in den 1970er Jahren häufiger zu beobachten und ihr Anteil steigt bereits ab den 1980er Jahren. In der *Psychologie* finden sich schon im Jahr 1960 Artikel mit vier oder mehr Autoren und ihr Anteil nimmt bereits ab Anfang der 1970er Jahre zu.

Für Artikel der Zeitschrift *Science* verhalten sich die Teamgrößen bzw. deren Anteile im Zeitverlauf etwas anders als beispielsweise die in der *Volkswirtschaftslehre* oder *Mathematik*. Hier ist der Anteil der Solo-Artikel nur unwesentlich gefallen, wie man der durchgezogenen Linie in Abbildung 4.8 entnehmen kann (von 44,0 % im Jahr 1960 auf 39,1 % im Jahr 2011). Im Gegensatz zu den bislang betrachteten Disziplinen

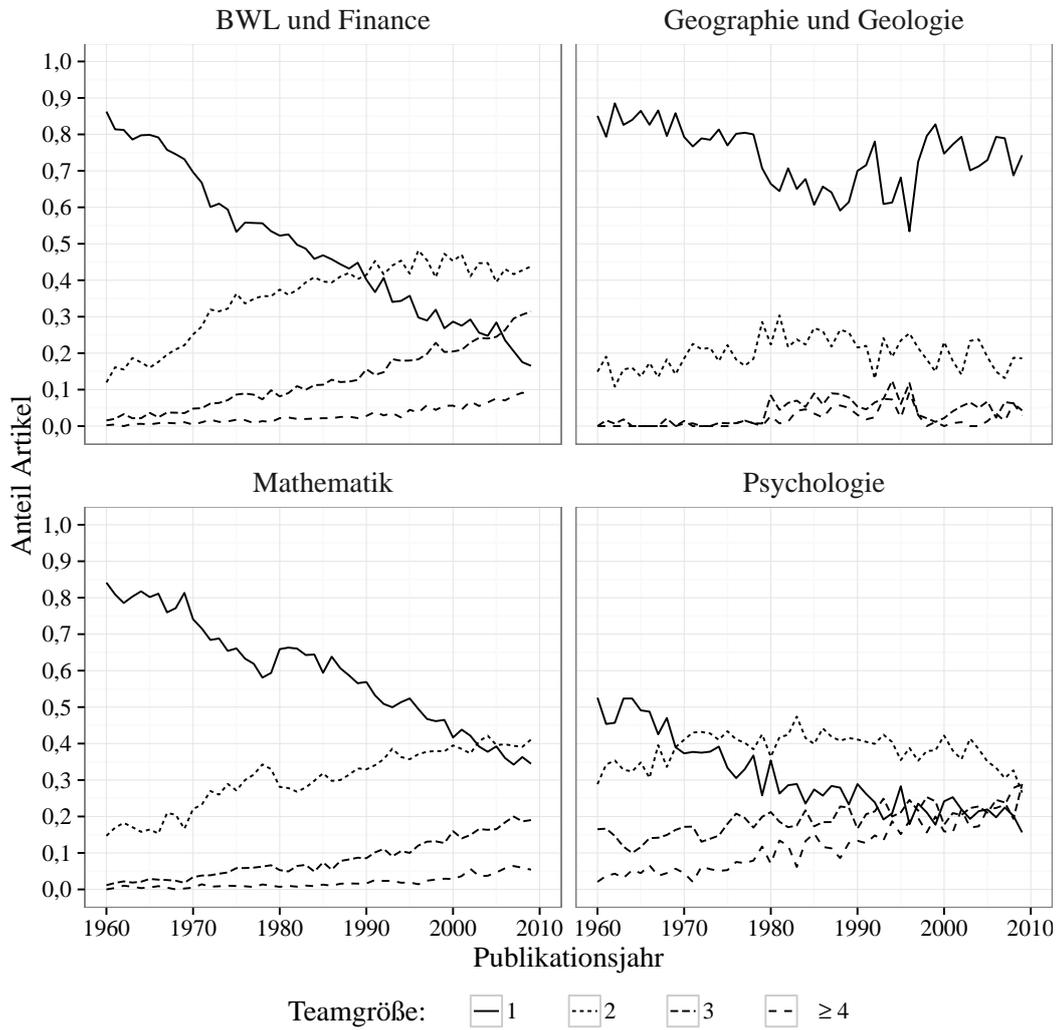


Abbildung 4.7: Anteile der Teamgrößen im JSTOR-Datensatz.

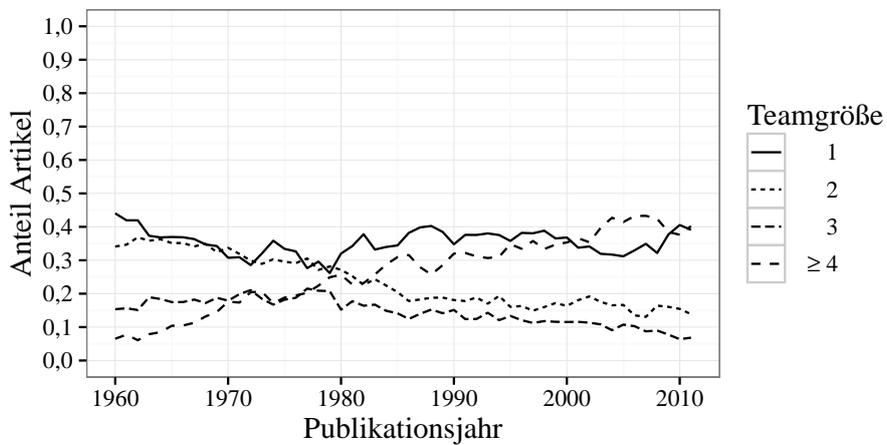


Abbildung 4.8: Anteile der Teamgrößen in der Zeitschrift *Science*.

sind die Anteile von Artikeln mit zwei oder drei Autoren im Beobachtungszeitraum zurückgegangen. Waren es 1960 noch 34,1 % Artikel mit zwei Autoren und 15,3 % Artikel mit drei Autoren, sind es im Jahr 2011 nur noch 13,9 % bzw. 6,8 %. Dafür fällt der Anstieg der Artikel mit vier oder mehr Autoren sehr stark aus. Im Jahr 1960 lag der Anteil dieser Beiträge bei 6,5 % und im Jahr 2011 ist deren Anteil mit 40,2 % etwa gleichauf mit dem Anteil der Artikel, die nur von einer Person geschrieben wurden.

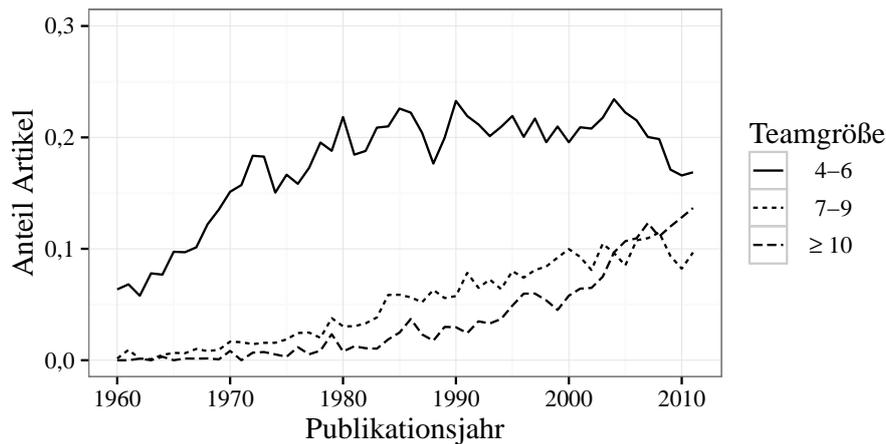


Abbildung 4.9: Anteile der Artikel mit mindestens 4 Autoren in der Zeitschrift *Science*.

Interessant ist hier ein genauerer Blick auf Artikel mit mindestens vier Autoren. Die Anteile für diese Teamgrößen sind nochmals detaillierter in Abbildung 4.9 abgebildet. Sie beziehen sich nicht auf die Anteile innerhalb der Kategorie „ ≥ 4 “ aus Abbildung 4.8, sondern auf die Anteile innerhalb des gesamten Datensatzes. Wie man sieht, machen in dieser Kategorie Artikel mit 4–6 Autoren über den gesamten Beobachtungszeitraum den größten Teil am Gesamtdatensatz aus. Allerdings haben die Anteile für Teamgrößen von 7–9 Autoren und speziell die der Teams mit mindestens 10 Autoren stark zugenommen. Für Teamgrößen von 7–9 Autoren liegt der Anteil 1960 bei 0,16 %, während es 1960 noch keine Artikel mit zehn oder mehr Autoren gibt. Im Jahr 2011 machen Artikel mit zehn oder mehr Autoren bereits einen Anteil von 13,7 % am Gesamtdatensatz aus. Dies ist ein erheblicher Unterschied zu den Ergebnissen für die anderen Disziplinen. Beispielsweise erreicht in der *Volkswirtschaftslehre* der Anteil Artikel mit mindestens vier Autoren im Jahr 2014 weniger als 10 % (vgl. Abbildung 4.6, S. 88).

Die grafischen Zeittrends werden im Folgenden mittels linearer Regression näher untersucht. In dieser Analyse lässt sich für verschiedene Faktoren kontrollieren, die die Teamgröße in einem Artikel beeinflussen könnten. Dadurch kann beispielsweise ausgeschlossen werden, dass die steigenden Trends in den Abbildungen 4.1 bis 4.3

durch bestimmte Zeitschriften getrieben sind, deren Anteile im Datensatz im Zeitverlauf ggf. zunehmen und in denen überdurchschnittlich große Teams zu beobachten sind.

Notation in Regressionstabellen

Da im Folgenden zum ersten Mal in dieser Arbeit Ergebnisse anhand einer Regressionstabelle diskutiert werden, folgt an dieser Stelle zunächst eine kurze Erläuterung der Notation der Regressionstabellen, die in dieser Arbeit grundsätzlich alle ähnlich aufgebaut sind. In der obersten Zeile einer jeden Tabelle findet sich ein Hinweis auf die betrachtete abhängige Variable. Die Zeile darunter ist nicht in allen Tabellen gleich. Für einzelne Regressionstabellen folgt hier direkt eine Nummerierung der betrachteten Regressionsmodelle, zum Beispiel (1), (2) oder (3). Dies ist beispielsweise für Tabelle 4.14 (S. 135) zum Alter beim ersten Artikel eines Autors in der *Mathematik* der Fall. Diese Modelle unterscheiden sich primär dadurch, welche Kontrollvariablen berücksichtigt werden. Teilweise sind die unterschiedlichen Modelle nicht wie eben beschrieben nummeriert, sondern anhand von Kurzbeschreibungen für die betrachteten Artikel benannt. So finden sich beispielsweise in den Regressionstabellen zum JSTOR-Datensatz die Modelle „BWL“, „Geo“, „Mathe“ sowie „Psycho“ für die vier untersuchten Disziplinen und das Modell „Gesamt“ für ihre Gesamtheit (z. B. Tabelle 4.2, S. 95). Diese Modelle unterscheiden sich nicht nach den berücksichtigten Kontrollvariablen, sondern nach der zugrunde liegenden Stichprobe. Die Kontrollvariablen (z. B. *Publikationsjahr* oder *Teamgröße*) finden sich jeweils in der ersten Spalte der Tabellen.

In den meisten Tabellen sind Regressionsergebnisse abgebildet, die sich nach einem weiteren Kriterium unterscheiden und entsprechend gruppiert sind. So wird beispielsweise in Tabelle 4.1 für den EconLit-Datensatz zwischen Modellen unterschieden, die alle Artikel betrachten und solchen die Top-Artikel betrachten. In anderen Tabellen (z. B. Tabelle 4.3, S. 95) wird zwischen Ergebnissen für alle Artikel und erste Artikel von Autoren im Datensatz unterschieden. Für diese Tabellen wird eine solche Unterscheidung bzw. Gruppierung in der zweiten Zeile angegeben. Die Nummerierung der Regressionsmodelle folgt eine Zeile später.

Für die meisten Kontrollvariablen sind die geschätzten Regressionskoeffizienten, mit einer Kennung bzgl. des Signifikanzniveaus, in den entsprechenden Zeilen ausgewiesen. Eine Information bzgl. der Signifikanzniveaus findet sich unterhalb einer jeden Tabelle. Unterhalb der Regressionskoeffizienten findet sich in Klammern die jeweilige Standardabweichung. Eine Ausnahme bilden

Kontrollen für die Zeitschrift, das JEL-Feld oder die Disziplin innerhalb eines Datensatzes. Hier ist jeweils nur mittels eines „Ja“ vermerkt, ob in dem jeweiligen Modell dafür kontrolliert wurde oder nicht. Findet sich für eine Kontrollvariable in einem der Modelle in der entsprechenden Zeile kein Eintrag, so wurde in diesem Modell nicht für diese Variable kontrolliert. Am Ende der Tabellen finden sich ferner Informationen zur Anzahl der Beobachtungen und, wo angebracht, zum Bestimmtheitsmaß R^2 . Zusätzliche Informationen oder solche, die in den beschriebenen Dimensionen nicht abbildbar sind, finden sich in einer Bemerkung unterhalb der Tabelle.

Tabelle 4.1: Lineare Regressionen für die Teamgröße im EconLit-Datensatz.

	Abhängige Variable: Teamgröße					
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,021*** (0,000)	0,019*** (0,000)	0,022*** (0,000)	0,021*** (0,000)	0,019*** (0,000)	0,025*** (0,000)
Autorinnen			–0,638*** (0,022)			–0,532*** (0,003)
Alle Top-Inst			–0,345*** (0,012)			–0,319*** (0,004)
Alle Nicht-Aka			–0,349*** (0,028)			–0,248*** (0,004)
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Beobachtungen	28.859	28.859	17.973	762.826	762.826	623.704
R^2	0,123	0,138	0,199	0,069	0,176	0,211

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Die Regressionen für volkswirtschaftliche Artikel aus dem EconLit-Datensatz finden sich in Tabelle 4.1. Die Modelle (1) bis (3) betrachten dabei lediglich Top-Artikel, während die Regressionen in den Modellen (4) bis (6) für alle Artikel im EconLit-Datensatz durchgeführt wurden. Wie man der Tabelle entnimmt, wird der steigende Trend bzgl. der Teamgröße pro Artikel in der Volkswirtschaftslehre unter diversen Kontrollen bestätigt und der zeitliche Trend erweist sich jeweils als statistisch hoch signifikant. In

den Modellen (3) und (6) mit allen Kontrollen fällt der zeitliche Effekt am größten aus und ist für alle Artikel mit 0,025 etwas größer als der für Top-Artikel mit 0,022. Der Regressionskoeffizient bzgl. des Publikationsjahres entspricht einer Erhöhung der Teamgröße um einen Autor alle 40 bis 45 Jahre, was mit den grafischen Beobachtungen in Abbildung 4.1 (S. 84) übereinstimmt. Teams, die vollständig aus Frauen bestehen, sind insgesamt kleiner als rein männliche oder gemischt-geschlechtliche Teams, wie aus dem Koeffizienten der Kontrollvariable „Autorinnen“ hervorgeht. Diese Beobachtung ist nicht verwunderlich, da mit zunehmender Zahl an Autoren in einem Team natürlicherweise die Wahrscheinlichkeit sinkt, dass alle Autoren das gleiche Geschlecht haben. Die Wahrscheinlichkeit für größere Autorinnen-Teams ist ferner zusätzlich geringer als die für reine Autoren-Teams, da die Zahl der Frauen im Datensatz geringer ist als die der Männer. Im Jahr 1970 lag der Frauenanteil im gesamten EconLit-Datensatz bei lediglich 6,1 % und stieg bis zum Jahr 2014 auf 29,6 %.

Auch Teams, bei denen alle Autoren einer Top-Institution angehören, sind etwas kleiner als andere Teams. Dies könnte einerseits darauf zurückzuführen sein, dass diese Autoren leistungsfähiger sind und die Arbeit an einem Artikel daher unter weniger Autoren aufteilen. Ferner haben diese Autoren an ihren Institutionen ggf. günstigere Bedingungen hinsichtlich der Forschung und Publikation als Autoren anderer Universitäten und Organisationen. Für die höhere Leistungsfähigkeit spricht, dass der Anteil der Autoren von Top-Artikeln, die zum Zeitpunkt des Schreibens einer Top-Institution zugehörig waren, beispielsweise im Jahr 2014 bei 44,8 % lag.⁷ Gegeben, dass die 50 als „Top“ klassifizierten Institutionen lediglich einen Bruchteil aller Institutionen im Datensatz ausmachen, ist dies ein beachtlicher Anteil.⁸ Zuletzt gelten ähnliche Überlegungen wie für die kleineren Autorinnen-Teams: Erstens ist es in größeren Teams rein rechnerisch unwahrscheinlicher, dass alle Autoren einer Top-Institution angehören. Ferner machen Autoren von Top-Institutionen im Datensatz schlichtweg den kleineren Anteil aller Autoren aus, was die Wahrscheinlichkeit von Teams, deren Mitglieder alle von Top-Institutionen stammen, weiter senkt. Den Regressionen in Tabelle 4.1 ist ferner zu entnehmen, dass die Teamgröße auch für Teams kleiner ausfällt, deren Autoren alle einer nicht-akademischen Institution angehören. Wieder greifen die Erklärungen, die bereits für Autorinnen-Teams und die „Top-Institutionen-Teams“ erläutert wurden. Der Anteil der Autoren, die einer nicht-akademischen Institution angehören, liegt im

⁷Für den gesamten EconLit-Datensatz liegt der Anteil Autoren von Top-Institutionen im Jahr 2014 bei 9,5 %.

⁸Für den gesamten EconLit-Datensatz finden sich 170.040 verschiedene Institutszugehörigkeiten. Für die Top-Autoren sind es 6.745. Tatsächlich dürften es etwas weniger Institutionen sein, da bei der Ermittlung dieser Zahlen nicht berücksichtigt wurde, dass dieselbe Institution im Datensatz mehrmals unter leicht unterschiedlichen Schreibweisen auftreten kann.

Jahr 2014 bei 8,1 % für den gesamten Datensatz und bei 8,8 % für den Datensatz der Top-Artikel.

Tabelle 4.2: Lineare Regressionen für die Teamgröße im JSTOR-Datensatz.

	Abhängige Variable: Teamgröße				
	Gesamt	BWL	Geo	Mathe	Psycho
Publikationsjahr	0,021*** (0,000)	0,023*** (0,000)	0,012*** (0,001)	0,018*** (0,000)	0,037*** (0,001)
Autorinnen	-0,606*** (0,011)	-0,616*** (0,018)	-0,272*** (0,030)	-0,510*** (0,017)	-0,807*** (0,030)
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	Ja
Beobachtungen	105.646	44.082	6.337	48.890	10.806
R^2	0,218	0,192	0,161	0,155	0,171

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Tabelle 4.3: Lineare Regressionen für die Teamgröße in der Zeitschrift *Science*.

	Abhängige Variable: Anzahl Autoren			
	Alle Artikel: (1)–(2)		Erste Artikel: (3)–(4)	
	(1)	(2)	(3)	(4)
Publikationsjahr	0,060*** (0,001)	0,064*** (0,001)	0,094*** (0,001)	0,095*** (0,001)
Autorinnen		-2,623*** (0,067)		-3,088*** (0,131)
Beobachtungen	70.028	70.028	49.202	49.202
R^2	0,038	0,059	0,074	0,085

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Bem.: Die Modelle (1) und (2) betrachten alle Artikel. Die Modelle (3) und (4) betrachten Artikel, die für mindestens einen der Autoren dessen erster Artikel sind.

Der Koeffizient bzgl. des Zeittrends entspricht für alle Artikel im JSTOR-Datensatz (Modell „Gesamt“), *Betriebswirtschaftslehre und Finance* (Modell „BWL“) sowie *Mathematik* (Modell „Mathe“) (Tabelle 4.2) etwa dem im EconLit-Datensatz (vgl. Tabelle 4.1) und ist statistisch ebenfalls hoch signifikant. Der geschätzte Effekt für

Geographie und Geologie (Modell „Geo“) ist mit einem Koeffizienten von 0,012 etwas geringer, während er für *Psychologie* (Modell „Psycho“) mit 0,037 größer ist. Wie oben bereits diskutiert, sind Teams von zehn oder mehr Autoren in *Science* in den letzten Jahren immer häufiger geworden (vgl. Abbildung 4.9, S. 91). Dies spiegelt sich auch beim Anstieg in der mittleren Teamgröße wider (vgl. Abbildung 4.3 (S. 86) und Regressionsmodelle (1) und (2) in Tabelle 4.3). Der zeitliche Effekt ist hier mit 0,060 bzw. 0,064 fast dreimal so hoch, wie für Artikel aus dem EconLit- und dem JSTOR-Datensatz.

4.1.2 Teamgröße bei erstem Artikel

In Abschnitt 4.1.1 wurde die mittlere Teamgröße für alle Artikel in den jeweiligen Datensätzen betrachtet. Von speziellem Interesse ist in dieser Arbeit die erste Publikation bzw. der erste Artikel eines Autors, da dieser vermeintlich seinen ersten bedeutenden Beitrag zur Wissenschaft markiert. Daher wird im Folgenden die Teamgröße speziell für Artikel betrachtet, die für mindestens einen der Autoren dessen erster Artikel sind. Autoren werden zum Zeitpunkt ihres ersten Artikels im Folgenden auch als Debütanten oder debütierende Autoren bezeichnet.⁹ Diese Definition von ersten Artikeln, impliziert für diese zumindest theoretisch bzw. statistisch eine Verzerrung hin zu größeren Teams. Geht man davon aus, dass sich Autoren mit gleicher Wahrscheinlichkeit auf die einzelnen Artikel verteilen, so ist die Wahrscheinlichkeit mindestens einen debütierenden Autor in einem Team anzutreffen für größere Teams höher. Dennoch erscheint es wichtig und interessant zu prüfen, ob dies empirisch auch tatsächlich der Fall ist.

Betrachtet man die Daten für die einzelnen Disziplinen, so machen Artikel mit mindestens einem Debütanten bei größeren Teams tatsächlich den größeren Teil aus, bleiben dabei allerdings unterhalb der theoretischen Verteilung. Beispielsweise sind für die *Mathematik* im Jahr 2009 46,4 % der Autoren Debütanten. Daher sollten bei Gleichverteilung 46,4 % der Artikel mit genau einem Autor von Debütanten verfasst worden sein. Entsprechend sollten 71,3 % der Artikel mit zwei Autoren, 84,6 % der Artikel mit drei Autoren sowie 91,8 % der Artikel mit vier Autoren mindestens einen debütierenden Autor aufweisen. Im Datensatz wurde für die *Mathematik* allerdings mehr als die Hälfte der Artikel mit nur einem Autor (51,3 %) von einem debütierenden

⁹Für die einzelnen Disziplinen bleibt der Anteil debütierender Autoren über den jeweiligen Beobachtungszeitraum etwa konstant oder ist leicht fallend. Gleiches gilt mehrheitlich auch für die Anteile von Artikeln mit mindestens einem debütierenden Autor, wobei diese für *Mathematik* und *Psychologie* leicht steigen.

Autor verfasst und damit ein größerer Anteil als unter der Annahme einer Gleichverteilung. Hingegen sind nur 62,0 %, 70,0 % bzw. 84,5 % der Artikel mit zwei, drei bzw. vier Autoren für mindestens ein Teammitglied der erste Artikel. Ähnlich verhält es sich auch für die anderen Disziplinen im JSTOR-Datensatz und die volkswirtschaftlichen Artikel aus dem EconLit-Datensatz. Eine mögliche Erklärung für diese Unterschiede zwischen Theorie und Empirie könnte sein, dass sich Autoren in der Praxis vermutlich nicht gleich auf Artikel verteilen. Debütanten könnten sogar versuchen ihren ersten Artikel alleine oder mit möglichst wenigen Autoren zu verfassen, um so ein starkes *Job-Market-Paper* für ihre weitere Karriere zu erzielen. Gegebenenfalls verfassen sie ihren ersten Artikel aber auch in kleineren Teams, da ihnen zu Beginn der Karriere noch das nötige Netzwerk bzw. die Kontakte zu Koautoren fehlen. Eine Ausnahme hinsichtlich der beschriebenen Unterschiede zwischen den theoretischen Wahrscheinlichkeiten und beobachteten Anteilen für Artikel debütierender Autoren stellt die Zeitschrift *Science* in den späteren Beobachtungsjahren dar. So sind im Jahr 2009 61,4 % der Autoren Debütanten. Entsprechend sollten beispielsweise 61,4 % der Artikel mit nur einem Autor und 97,7 % der Artikel mit vier Autoren von mindestens einem Debütanten verfasst worden sein. Tatsächlich wurden nur 21,4 % der Artikel mit nur einem Autor von Debütanten verfasst, aber 94,0 % der Artikel mit vier Autoren von mindestens einem Debütanten. Somit wird der Großteil von Artikeln mit kleinen Teams nicht von Debütanten verfasst, während der Anteil von Artikeln mit größeren Teams und mindestens einem debütierenden Autor fast der theoretischen Wahrscheinlichkeit entspricht.

In den Abbildungen 4.10 bis 4.13 ist jeweils die mittlere Teamgröße für Artikel mit debütierenden Autoren (durchgezogene Linie) und alle Artikel im jeweiligen Datensatz (gepunktete Linie) abgebildet, wobei letztere bereits in Abschnitt 4.1.1 zu sehen waren. Für den EconLit-Datensatz ist in den Abbildungen 4.10 und 4.11 deutlich zu sehen, dass die mittlere Teamgröße beim ersten Artikel einen steigenden Trend aufweist. Außerdem ist zu erkennen, dass die mittlere Teamgröße bei den ersten Artikeln, bis auf das Jahr 1970 für Top-Artikel, größer ist, als für alle Publikationen der entsprechenden Autoren. Das heißt, dass Artikel mit debütierenden Autoren im Mittel mehr Autoren aufweisen. Es ist ferner zu beobachten, dass sich die beiden Linien im Laufe der Zeit weiter voneinander entfernen. Dies würde bedeuten, dass der Unterschied zwischen Artikeln mit debütierenden Autoren und allen Artikeln zunehmend größer wird. Dieses Phänomen ist allerdings auf den Datensatz zurückzuführen. Wie in Abschnitt 3.1 (S. 40) erläutert, umfasst der EconLit-Datensatz lediglich Autoren und deren Artikel, die ihren ersten Artikel in oder nach 1970 veröffentlicht haben. Entsprechend finden sich im Jahr 1970 beinahe ausschließlich erste Artikel, außer ein Autor hat im selben Jahr

mehrere Artikel veröffentlicht. Daher zeigt sich der Unterschied zwischen ersten und allen Artikeln der Autoren für den EconLit-Datensatz im Laufe der Zeit zunehmend.

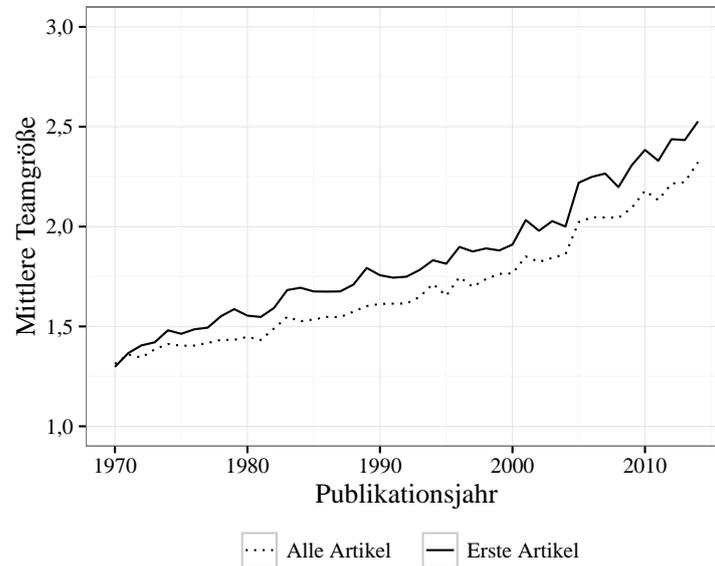


Abbildung 4.10: Mittlere Teamgröße beim ersten Top-Artikel eines Autors im EconLit-Datensatz.

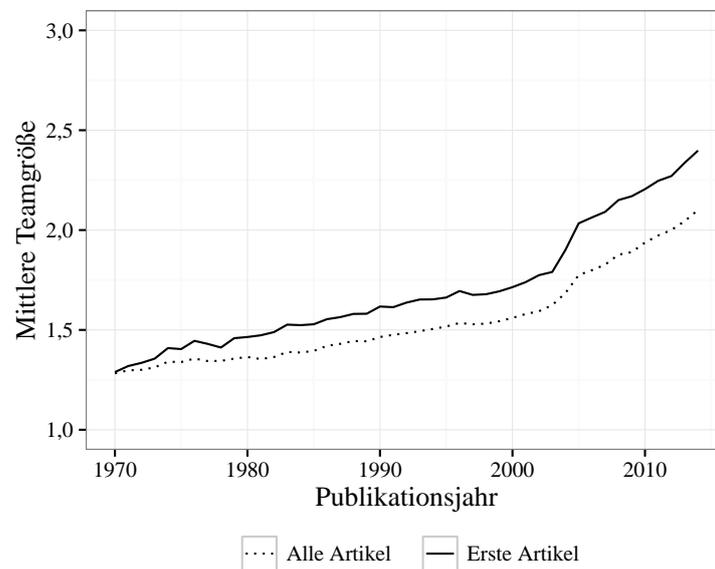


Abbildung 4.11: Mittlere Teamgröße beim ersten Artikel eines Autors im EconLit-Datensatz.

Auch für die weiteren in dieser Arbeit untersuchten Datensätze bzw. Disziplinen zeigt sich, dass die mittlere Teamgröße beim ersten Artikel im jeweiligen Beobachtungszeitraum deutlich gestiegen ist. Abbildung 4.12 zeigt die mittleren Teamgrößen für die Disziplinen aus dem JSTOR-Datensatz. Wie für volkswirtschaftliche Artikel zeigt

sich auch für Artikel der Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* sowie *Psychologie*, dass die mittlere Teamgröße bei den ersten Artikeln eines Autors größer ist als für alle Artikel. Beispielsweise weisen alle Artikel der *Mathematik* im Jahr 1960 eine mittlere Teamgröße von 1,2 auf, während erste Artikel im selben Jahr im Mittel von 1,3 Autoren geschrieben wurden. Im Jahr 2009 betragen die entsprechenden mittleren Teamgrößen 2,0 bzw. 2,1. Abgesehen von den höheren Mittelwerten für erste Artikel, verlaufen die beiden Kurven sehr ähnlich. Abbildung 4.13 veranschaulicht den Trend in der Teamgröße beim ersten Artikel in der Zeitschrift *Science*. Auch hier zeigt sich, dass die mittlere Teamgröße bei ersten Artikeln deutlich gestiegen ist (durchgezogene Linie) und dass Teams mit mindestens einem debütierenden Autor größer sind als bei allen Artikeln. Ferner ist hier zu erkennen, dass die mittlere Teamgröße für erste Artikel stärker steigt als für alle Artikel. Als besonders renommierte Zeitschrift, nimmt *Science* im Gegensatz zu den untersuchten Disziplinen mit mehreren Zeitschriften eine Sonderstellung ein und weist einige Besonderheiten auf, wie im weiteren Verlauf der Arbeit noch zu sehen sein wird. So sind beispielsweise Teams in *Science*, wie bereits gesehen, in den letzten Beobachtungsjahren besonders stark gewachsen. Ferner wurde weiter oben beschrieben (S. 97), dass sich debütierende Autoren in *Science* in den späteren Jahren des Beobachtungszeitraums vorwiegend in größeren Teams finden und Artikel mit ein oder zwei Autoren zum größten Teil nicht von Debütanten geschrieben bzw. mitverfasst werden. Entsprechend sind die unterschiedlich steilen Verläufe der beiden Kurven in Abbildung 4.13 zu erklären.

Auf Basis der grafischen Trends kann somit insgesamt festgehalten werden, dass die mittlere Teamgröße beim ersten Artikel eines Autors sowie die mittlere Teamgröße bei allen Artikeln (vgl. Abschnitt 4.1.1) in den vergangenen Jahrzehnten deutlich gestiegen sind. Dabei ist die mittlere Teamgröße beim ersten Artikel eines Autors größer als bei allen Artikeln. Die grafischen Zeittrends werden auch in den entsprechenden Regressionen bestätigt und erweisen sich als hoch signifikant.

Tabelle 4.4 zeigt die Regressionsergebnisse für erste Artikel im EconLit-Datensatz. Die Ergebnisse stimmen hinsichtlich der Effektrichtungen mit denen für alle Artikel überein und weisen positive Regressionskoeffizienten hinsichtlich des Publikationsjahres aus (vgl. Tabelle 4.1, S. 93). Das heißt, dass die Teamgröße, wie bereits grafisch gesehen, mit der Zeit auch für erste Artikel angestiegen ist, sodass Autoren ihr Debüt in zunehmend großen Teams geben. Die geschätzten Koeffizienten sind ähnlich groß wie für alle Artikel und hoch signifikant. Beispielsweise beträgt der geschätzte Koeffizient bzgl. des Publikationsjahres für alle Top-Artikel im Modell (3) 0,022 (vgl. Tabelle 4.1, S. 93) und der für Artikel mit debütierenden Autoren 0,026 im gleichen

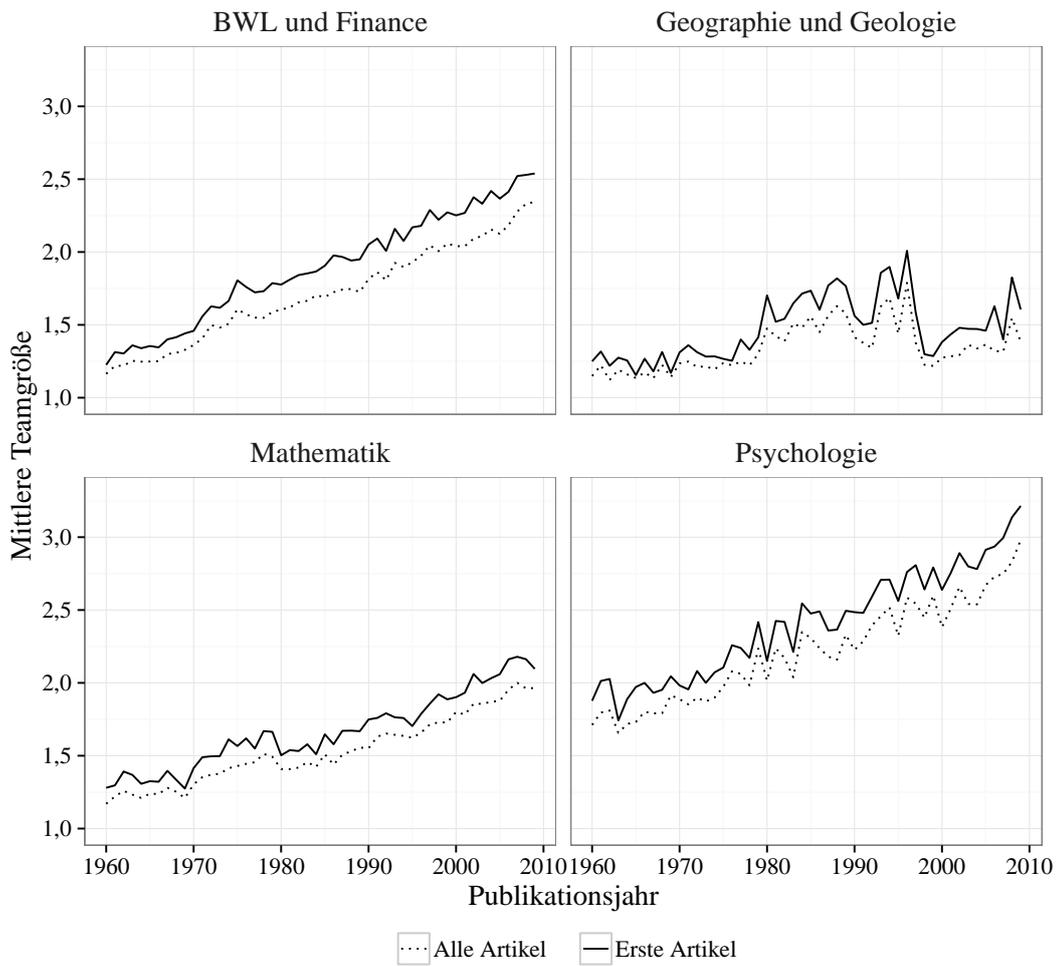


Abbildung 4.12: Mittlere Teamgrößen beim ersten Artikel eines Autors im JSTOR-Datensatz.

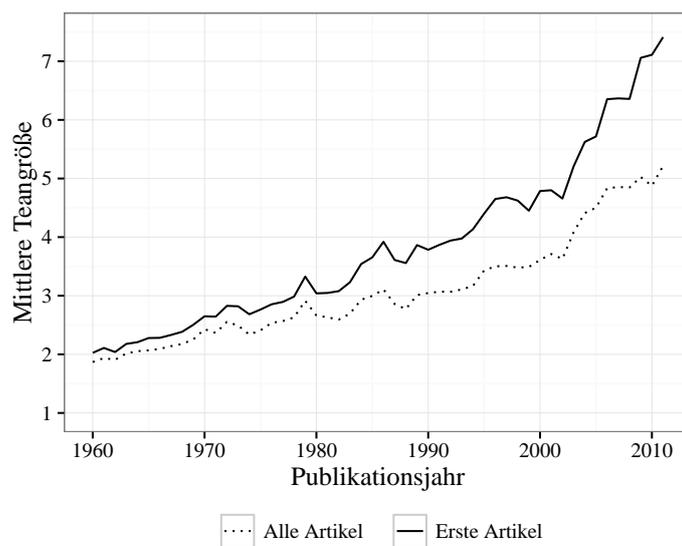


Abbildung 4.13: Mittlere Teamgröße beim ersten Artikel eines Autors in der Zeitschrift *Science*.

Modell in Tabelle 4.4.¹⁰Die Regressionen für die Teamgröße beim ersten Artikel für den JSTOR-Datensatz finden sich in Tabelle 4.5. Analog zu den Regressionen für den EconLit-Datensatz bestätigen auch diese einen signifikant steigenden Zeittrend, dessen geschätzte Koeffizienten ähnlich groß sind, wie die für alle Artikel. Der Koeffizient ist dabei für *Psychologie* am größten und für *Geographie und Geologie* am kleinsten. Die Regressionen für den jeweils ersten Artikel eines Autors in der Zeitschrift *Science* finden sich in Tabelle 4.3 (S. 95) in den Modellen (3) und (4). Auch hier spiegeln sich die in Abbildung 4.13 beobachteten Trends wider und weisen einen signifikant positiven Zeittrend aus.

Tabelle 4.4: Lineare Regressionen für die Teamgröße beim ersten Artikel eines Autors im EconLit-Datensatz.

	Abh. Variable: Teamgröße					
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,024*** (0,000)	0,024*** (0,001)	0,026*** (0,001)	0,026*** (0,000)	0,026*** (0,000)	0,033*** (0,000)
Autorinnen			–0,776*** (0,032)			–0,659*** (0,006)
Alle Top-Inst			–0,310*** (0,019)			–0,293*** (0,008)
Alle Nicht-Aka			–0,331*** (0,039)			–0,267*** (0,007)
Zeitschrift (Dummies)		Ja	Ja		Ja	Ja
JEL-Feld (Dummies)		Ja	Ja		Ja	Ja
Beobachtungen	14.190	14.190	8.901	300.915	300.915	241.464
R ²	0,161	0,175	0,228	0,099	0,244	0,268

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

¹⁰Die Koeffizienten für die Kontrollen „Autorinnen“, „Alle Top-Inst“ sowie „Alle Nicht-Aka“ wurden hier nicht weiter diskutiert, da die Wahrscheinlichkeit für größere Teams, in denen alle Autoren eine dieser Eigenschaften teilen, wie bereits in Abschnitt 4.1.1 diskutiert, geringer ist. Daher gibt es konstruktionsbedingt einen negativen Einfluss dieser Kontrollen auf die Teamgröße.

Tabelle 4.5: Lineare Regressionen für die Teamgröße beim ersten Artikel eines Autors für den JSTOR-Datensatz.

	Abhängige Variable: Teamgröße				
	Gesamt	BWL	Geo	Mathe	Psycho
Publikationsjahr	0,024*** (0,000)	0,026*** (0,000)	0,017*** (0,001)	0,020*** (0,000)	0,040*** (0,001)
Autorinnen	-0,690*** (0,015)	-0,716*** (0,025)	-0,437*** (0,042)	-0,606*** (0,024)	-0,822*** (0,037)
Zeitschrift (Dummies)	Ja	Ja	Ja	Ja	Ja
Beobachtungen	60.274	23.408	3.925	27.110	8.199
R^2	0,247	0,221	0,187	0,179	0,183

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

4.2 Anzahl der Quellenangaben

4.2.1 Anzahl der Quellenangaben für alle Artikel

In diesem Abschnitt wird die Anzahl der Quellenangaben pro Artikel betrachtet. In den Abbildungen 4.14 bis 4.16 ist für alle betrachteten wissenschaftlichen Disziplinen eine deutliche Zunahme der mittleren Anzahl Quellenangaben pro Artikel zu sehen. Für alle volkswirtschaftlichen Artikel aus dem EconLit-Datensatz stieg deren Zahl um 157,6 % von 14,7 im Jahr 1970 auf fast 37,9 im Jahr 2014 (vgl. Abbildung 4.14), was einem Anstieg um den Faktor 2,5 entspricht. Dieser Trend zeigt sich sowohl für den gesamten Datensatz, als auch für die Top-Artikel.¹¹ In Abbildung 4.14 ist ferner zu erkennen, dass die mittlere Anzahl der Quellenangaben bei Top-Artikeln zwischen 1970 und Mitte der 1990er Jahre geringer ist als für alle Artikel im EconLit-Datensatz, während sie gegen Ende des Beobachtungszeitraums, also in den 2010er Jahren etwas größer ist. Entsprechend fällt der relative Anstieg in der mittleren Anzahl Quellenangaben pro Artikel für Top-Zeitschriften sogar höher aus. Sind es 1970 noch durchschnittlich 13,2 Quellenangaben pro Artikel, steigt dieser Wert bis ins Jahr 2014 auf 38,6. Dies

¹¹Die Trends zeigen sich auch, wenn man diese getrennt nach der Teamgröße eines Artikels betrachtet. In Abbildung 4.17 (S. 107) ist dies beispielhaft für den gesamten EconLit-Datensatz abgebildet. Erstaunlicherweise liegen die Mittelwerte für die einzelnen Teamgrößen nahe beieinander und für einzelne Jahre weisen Artikel mit nur einem Autor im Mittel sogar mehr Quellenangaben auf als Artikel mit mehreren Autoren.

entspricht einem Mehr an Quellenangaben von 193,0 % und damit beinahe einem Anstieg um den Faktor 3.

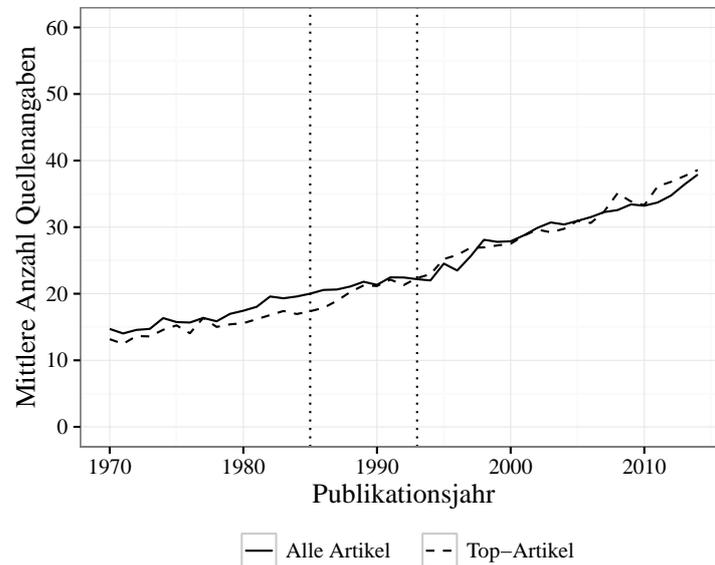


Abbildung 4.14: Mittlere Anzahl Quellenangaben pro Artikel im EconLit-Datensatz. Die gepunkteten Linien markieren die Jahre 1985 und 1993.

Für den Trend für alle Artikel im EconLit-Datensatz ist ferner zu erkennen, dass es einen leichten Knick Mitte der 1990er Jahre gibt, sodass die mittlere Anzahl der Quellenangaben ab diesem Zeitpunkt etwas schneller wächst als in den Jahren davor. Tatsächlich ergibt eine Analyse auf das Vorliegen eines segmentierten linearen Zusammenhangs zwischen der mittleren Anzahl Quellenangaben und dem Publikationsjahr das Jahr 1993 als Pivotpunkt.¹² Vergleicht man die Koeffizienten der linearen Modelle für Daten bis zu diesem Punkt und für Daten danach, so ergeben sich signifikant unterschiedliche Steigungen für die beiden Regressionsgeraden.¹³ Diese strukturelle Veränderung im Trend könnte möglicherweise auf das Aufkommen des Internets sowie die verstärkte Nutzung von Computern und Literaturverwaltungssoftware zurückzuführen sein.¹⁴ Eine analoge Analyse des Trends für Top-Artikel ergibt allerdings eine solche strukturelle Änderung für das Jahr 1985.¹⁵ Allerdings ist ein schnellerer Anstieg

¹²Die Analyse geschah mittels des R-Pakets „segmented“ und der gleichnamigen Funktion.

¹³Der Chow-Test bestätigt einen signifikanten Unterschied zwischen den beiden Steigungen.

¹⁴Denkbar ist grundsätzlich auch, dass der Trend in der mittleren Anzahl Quellenangaben keinem linearen sondern beispielsweise einem exponentiellen Wachstum folgt. Im Fokus dieser Arbeit steht allerdings nicht so sehr die Art des Wachstums, sondern v. a. die Richtung des Trends. Optisch erscheint ferner eine lineare Approximation der beobachteten Trends im Beobachtungszeitraum weitestgehend passend.

¹⁵Analog zur strukturellen Veränderung für alle Artikel im EconLit-Datensatz ist diese Änderung, wenn auch schwach, bei genauem Hinsehen für die gestrichelte Linie der Top-Artikel in Abbildung 4.14 erkennbar.

für die mittlere Anzahl der Quellenangaben in Top-Artikeln ab diesem Jahr wohl nicht auf das Internet zurückzuführen, da dieses erst 1991 für die Öffentlichkeit zugänglich war. Es ist davon auszugehen, dass sich Einflüsse des Internets auf die Anzahl der Quellenangaben sowohl für den gesamten EconLit-Datensatz als auch für Top-Artikel zeigen würden. Da für letztere allerdings keine solchen Effekte zu beobachten sind, ist entsprechend die strukturelle Änderung für alle Artikel im EconLit-Datensatz ab dem Jahr 1993 vermutlich nicht auf das Internet zurückzuführen.¹⁶

Auch für die vier weiteren wissenschaftlichen Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik*, sowie *Psychologie* aus dem JSTOR-Datensatz sind ähnliche Trends zu beobachten (vgl. Abbildung 4.15). So steigt die mittlere Anzahl der Quellenangaben pro Artikel in *BWL und Finance* beinahe um den Faktor 4 von 10,9 Quellenangaben im Jahr 1960 auf 42,5 im Jahr 2009. Einen ähnlichen relativen Anstieg weist die Disziplin *Mathematik* auf. Hier ist zwischen 1960 und 2009 ebenfalls ein Anstieg um den Faktor 4 zu beobachten. Dabei wurden im Jahr 1960 noch durchschnittlich 6,0 Quellenangaben gemacht, während es 23,7 im Jahr 2009 sind. In *Geographie und Geologie* fällt dieser Trend mit einem Faktor von etwas mehr als 2 geringer aus. Hier steigt die mittlere Anzahl der Quellenangaben von 19,8 im Jahr 1960 auf 42,8 im Jahr 2009. Allerdings geht sie in den letzten Beobachtungsjahren zurück. Bezogen auf das Jahr 2005 (\emptyset 54,4 Quellenangaben) fällt der Anstieg beispielsweise höher aus. In Abbildung 4.15 ist ferner zu sehen, dass die Mittelwerte für *Geographie und Geologie* zwischen den Jahren stärker schwanken als für die anderen Disziplinen. Eine mögliche Erklärung hierfür ist die deutlich geringere Anzahl an Artikeln pro Jahr im Vergleich zu den anderen Disziplinen (vgl. Tabelle 3.5, S. 59). Den größten absoluten und relativen Anstieg in der mittleren Anzahl Quellenangaben weist die *Psychologie* auf. Es ist ein deutlicher Anstieg um den Faktor 5 zu beobachten, wobei die mittlere Anzahl der Quellenangaben ab dem Jahr 1980 stärker steigt als in den Jahren zuvor.¹⁷ Für die psychologischen Zeitschriften im JSTOR-Datensatz steigt die mittlere Anzahl der Quellenangaben von 12,1 im Jahr 1960 auf 60,1 im Jahr 2009.

¹⁶Aufgrund des historischen Zeitpunktes könnte man auch mutmaßen, dass die strukturellen Veränderungen auf geopolitische Veränderungen (Auflösung der Sowjetunion im Jahr 1991) zurückzuführen sind. Die Argumentation hinsichtlich des Internets im Text schließt wohl aber auch diese Einflüsse aus.

¹⁷Bis zum Jahr 1978 sind in der *Psychologie* lediglich für die Zeitschrift *Child Development* Artikel mit Informationen zur Anzahl der Quellenangaben im Datensatz verfügbar. Entsprechend beschreibt die Trendlinie in Abbildung 4.15 bis 1978 lediglich eine Zeitschrift. Allerdings ähnelt der Anstieg in der mittleren Anzahl der Quellenangaben für diese Zeitschrift ab 1980 dem für die gesamte *Psychologie*. Ferner überlagert der Trend dieser Zeitschrift auch nicht die anderen Zeitschriften, für die der Trend ab 1980 ähnlich stark steigend ist.

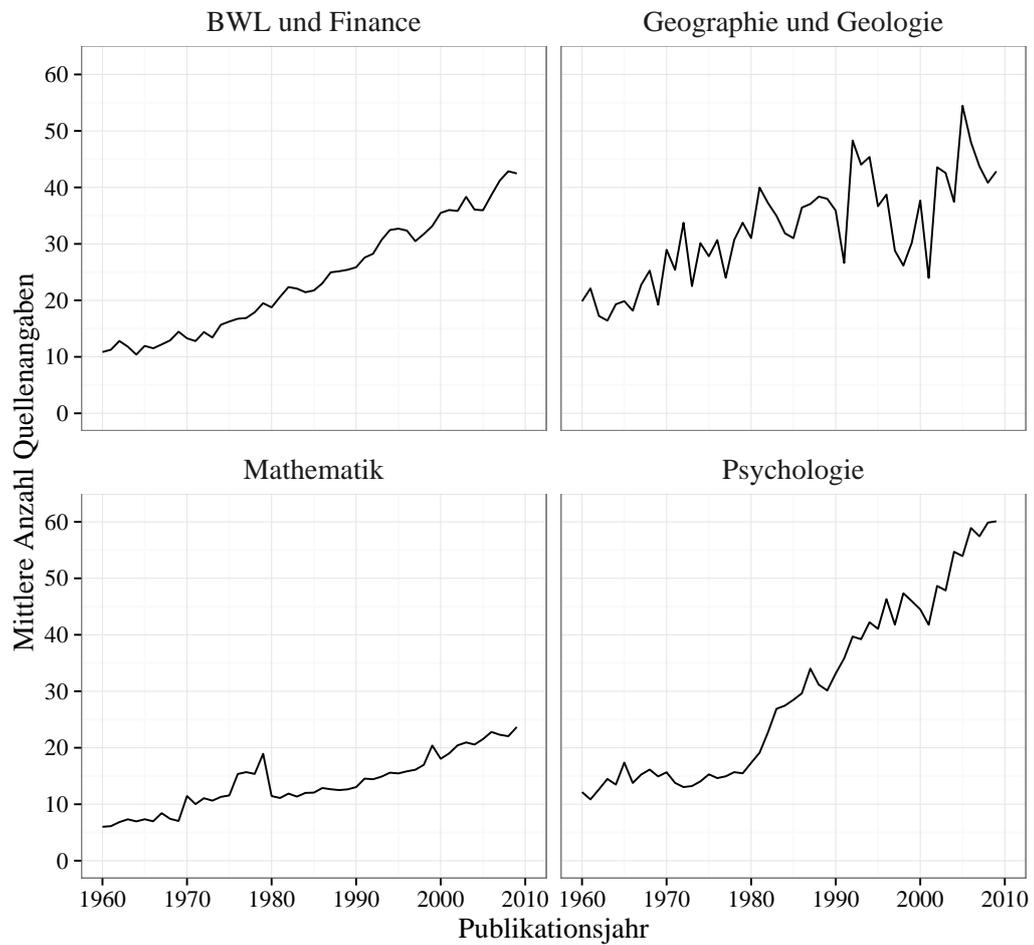


Abbildung 4.15: Mittlere Anzahl Quellenangaben pro Artikel im JSTOR-Datensatz.



Abbildung 4.16: Mittlere Anzahl Quellenangaben pro Artikel in der Zeitschrift *Science*.

Ähnliche Trends finden sich auch für Artikel aus der Zeitschrift *Science* wieder (vgl. Abbildung 4.16). Für diese steigt die Zahl der für einen Artikel angegebenen Quellen von 8,4 im Jahr 1960 auf 22,8 im Jahr 2011 mit Hochpunkten von über 25 Quellenangaben zu Beginn der 1990er Jahre und nach der Jahrtausendwende. Weiter oben wurde gemutmaß, dass der stärkere Anstieg in der Anzahl Quellenangaben pro Artikel ab Mitte der 1990er Jahre im EconLit-Datensatz womöglich mit dem Aufkommen des Internets und der stärkeren Nutzung von Literaturverwaltungssoftware zusammenhängt. Allerdings sind entsprechende strukturelle Veränderungen weder in den Trends für die einzelnen Disziplinen im JSTOR-Datensatz zu beobachten, noch zeigen sich diese für die Zeitschrift *Science*.¹⁸ Im Gegenteil sieht es so aus, als stagniere in *Science* die mittlere Anzahl der Quellenangaben pro Artikel seit 1990 bzw. als sei sie sogar leicht zurückgegangen (vgl. Abbildung 4.16). Es ist nicht klar, welche Ursache diese plötzliche Veränderung hat. Eine mögliche Erklärung sind die Vorgaben für Einreichungen seitens der Zeitschrift. Auf den entsprechenden Informationsseiten für Autoren auf der Homepage von *Science* findet sich eine Beschränkung von wissenschaftlichen Artikeln auf 4.500 Wörter.¹⁹ Da die Vorgabe von maximal 4.500 Wörtern auch Quellenangaben einschließt, sehen sich Autoren einem Kompromiss zwischen der Beschreibung eigener Inhalte und der Angabe von Quellen gegenüber. Möglicherweise wurde diese Beschränkung erst Anfang der 1990er Jahre eingeführt, wobei in dieser Zeit kein Rückgang der mittleren Länge von Artikeln (gemessen an der Anzahl der Seiten) im Datensatz zu beobachten ist. Denkbar ist ferner, dass ab Anfang der 1990er Jahre ein Verhältnis zwischen inhaltlichem Text und Quellenangaben erreicht wurde, das den Autoren noch eine ausreichende Darstellung der eigenen Ergebnisse erlaubt, bei dem sie sich allerdings beim Bezug auf vorhergehende Arbeiten einschränken müssen.²⁰

Wie in Abschnitt 4.1 (ab S. 84) zu sehen war, ist für die einzelnen Disziplinen im Beobachtungszeitraum auch für die mittlere Teamgröße ein deutlicher Anstieg zu beobachten. Somit könnte man vermuten, dass der Anstieg in der mittleren Anzahl der Quellenangaben auf zunehmend große Teams zurückzuführen ist, da man erwarten würde, dass die mittlere Anzahl der Quellenangaben mit der Teamgröße steigt. Wie in den späteren Regressionen zu sehen sein wird, korreliert die Anzahl der Quellenangaben tatsächlich positiv mit der Teamgröße. Allerdings ist dieser Effekt nicht sehr groß, wie

¹⁸Wie erwähnt, ist in der *Psychologie* auch ein struktureller Bruch zu beobachten (vgl. Abbildung 4.15), allerdings bereits im Jahr 1980, weshalb dieser nicht auf das Internet zurückgeführt werden kann.

¹⁹Siehe hierzu <http://www.sciencemag.org/authors/science-information-authors> (Stand: 08.05.2018).

²⁰Hinsichtlich des historischen Zeitpunktes, könnte der strukturelle Bruch auch geopolitische Hintergründe haben (z. B. Auflösung der Sowjetunion). Allerdings wären hiervon vermutlich auch andere Zeitschriften und Disziplinen betroffen. Außerdem ist nicht klar, warum dies eine derartige Auswirkung auf die Anzahl der Quellenangaben haben sollte und warum diese bis in das Jahr 2011 andauert.

aus den geschätzten Koeffizienten der Regressionen für die Teamgröße hervorgeht. Der Anstieg in der mittleren Anzahl Quellenangaben scheint allerdings nicht grundsätzlich auf zunehmend große Autorenteam zurückzuführen zu sein. Betrachtet man die Trends für die mittlere Anzahl der Quellenangaben separat nach Teamgrößen, so zeigen sich teilweise erstaunliche Verläufe, vor allem hinsichtlich der Mittelwerte und damit der Reihung der einzelnen Kurven. Die in den Abbildungen 4.14 bis 4.16 beobachteten Trends spiegeln sich weitestgehend auch bei dieser Betrachtungsebene wider. Für *Betriebswirtschaftslehre und Finance* sowie *Mathematik* ist größtenteils die erwartete Reihung der Kurven zu beobachten (vgl. Abbildung B.5, S. 211, im Anhang).

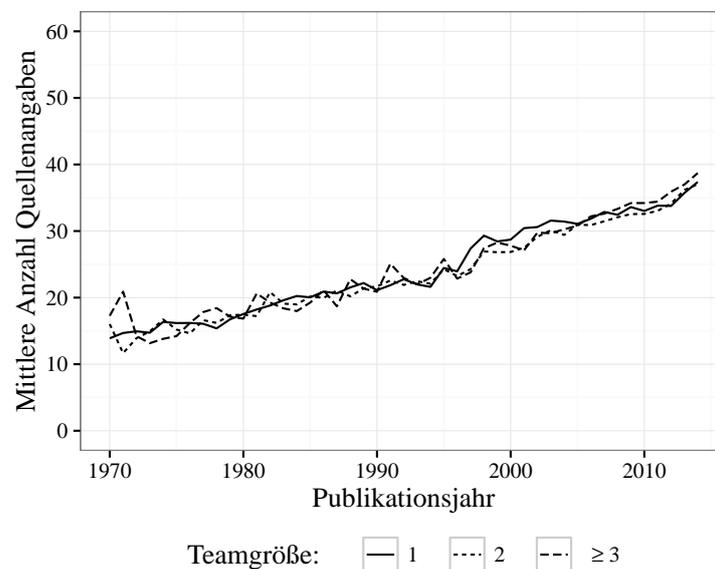


Abbildung 4.17: Mittlere Anzahl Quellenangaben pro Artikel im EconLit-Datensatz für verschiedene Teamgrößen.

Die Kurven für Teams von einem, zwei sowie mindestens drei Autoren für den gesamten EconLit-Datensatz sind in Abbildung 4.17 abgebildet. Wie man sieht, zeigen sich auch für die einzelnen Teamgrößen steigende Trends. Die drei Kurven liegen ferner nah beieinander, wobei die Kurve für Artikel mit einem Autor (durchgezogene Linie) teilweise sogar über den beiden anderen Kurven liegt. Das bedeutet, dass Artikel mit einem Autor in den entsprechenden Jahren im Mittel mehr Quellenangaben aufweisen als Artikel mit mehreren Autoren. Eine analoge Darstellung für Artikel der Zeitschrift *Science* findet sich in Abbildung 4.18.²¹ Wie man erkennt, liegt die Kurve der Artikel mit einem Autor bis Anfang der 1980er Jahre größtenteils deutlich über den Kurven für Teams von mindestens zwei Autoren. Dies ändert sich ab Mitte der 1980er Jahre,

²¹Die Abbildung für Top-Artikel ähnelt Abbildung 4.17 für den gesamten EconLit-Datensatz und ist daher nicht abgebildet.

wonach die erwartete Reihung der Mittelwerte bzw. Kurven zu beobachten ist. Auffällig ist ferner, dass die mittlere Anzahl der Quellenangaben für Artikel mit einem und zwei Autoren nach 1991 fallend ist. Dies erklärt zum Teil den Verlauf des ab Anfang der 1990er Jahre stagnierenden Trends aus Abbildung 4.16 für die mittlere Anzahl der Quellenangaben in der Zeitschrift *Science*.

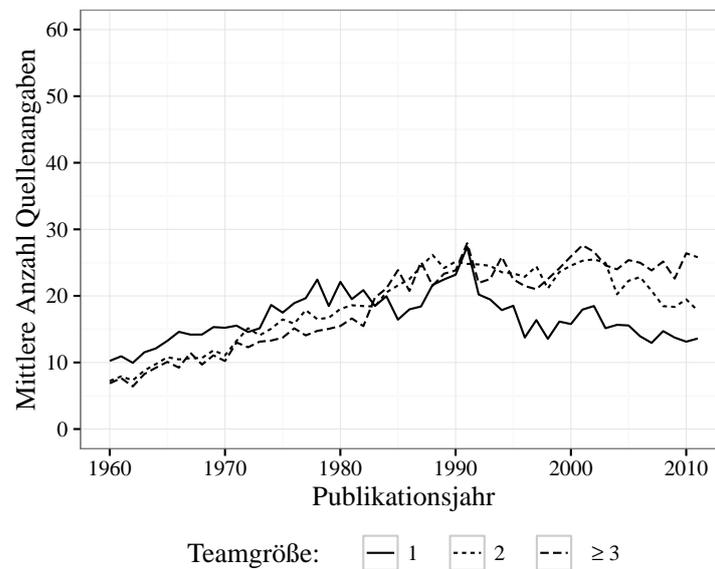


Abbildung 4.18: Mittlere Anzahl Quellenangaben pro Artikel in der Zeitschrift *Science* für verschiedene Teamgrößen.

Trägt man die mittlere Anzahl der Quellenangaben gegen die einzelnen Teamgrößen ab, so zeigt sich, dass die Anzahl der Autoren durchaus einen positiven Einfluss auf diese hat. Abbildung 4.19 zeigt dies für volkswirtschaftliche Artikel des EconLit-Datensatzes.²² Erstaunlicherweise ist die mittlere Anzahl der Quellenangaben für Artikel aus den Jahren 1992 sowie 2014 mit einem Autor sogar größer als für Artikel mit zwei Autoren. Auch für die übrigen Disziplinen ist der Effekt der Teamgröße auf die Anzahl der Quellenangaben insgesamt positiv und vor allem für die ersten zusätzlichen Autoren eines Artikels besonders ausgeprägt (vgl. Abbildung B.6, S. 212, im Anhang). Danach flacht die mittlere Anzahl der Quellenangaben für größere Teams deutlich ab oder ist sogar fallend, wie man Abbildung 4.20 für Artikel der Zeitschrift *Science* entnehmen kann. Für das Jahr 1960 (durchgezogene Linie in Abbildung 4.20) besteht sogar ein negativer Zusammenhang zwischen der Teamgröße und der Anzahl der Quellenangaben in *Science*.

Wie auch im vorangegangenen Abschnitt 4.1 für die Teamgröße werden die beobachteten Zeittrends im Folgenden mittels linearer Regression näher untersucht. Tabelle 4.6

²²Die Abbildung für Top-Artikel sieht Abbildung 4.19 sehr ähnlich und ist daher nicht abgebildet.

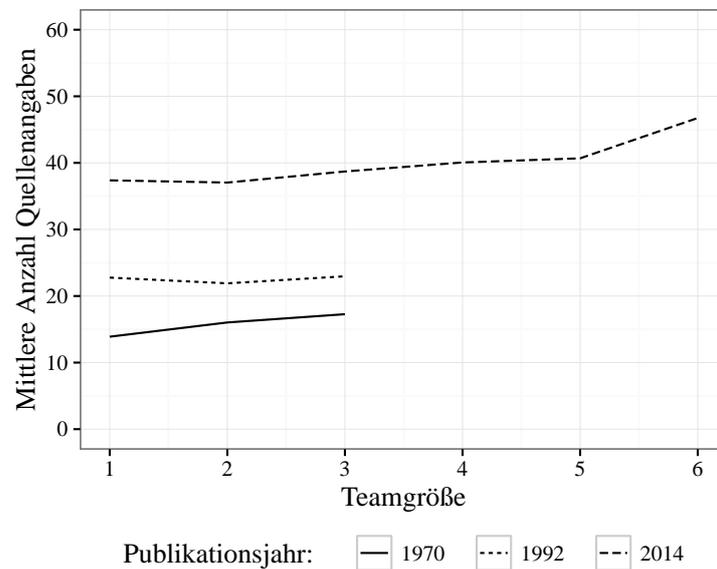


Abbildung 4.19: Mittlere Anzahl Quellenangaben pro Artikel im gesamten EconLit-Datensatz in Abhängigkeit von der Teamgröße.

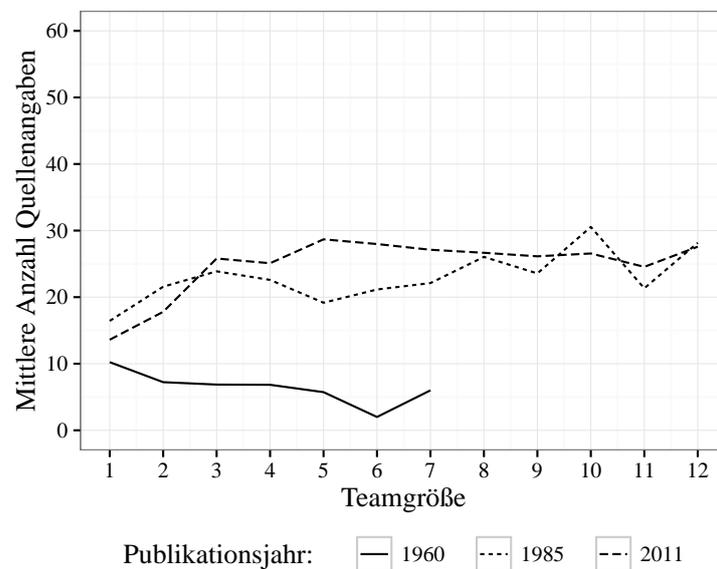


Abbildung 4.20: Mittlere Anzahl Quellenangaben pro Artikel in der Zeitschrift *Science* in Abhängigkeit von der Teamgröße.

ist zu entnehmen, dass der zeitliche Effekt bzgl. des Publikationsjahres auch unter diversen Kontrollen statistisch signifikant bleibt und im beobachteten Zeitraum einer Zunahme von etwas mehr als einer halben Quellenangabe pro Jahr entspricht. Dies gilt sowohl für Top-Artikel (Modelle (1) bis (3)) als auch für alle Artikel im EconLit-Datensatz (Modelle (4) bis (6)), wobei der rein zeitliche Trend für Top-Artikel mit einem Koeffizienten von 0,607 stärker ausfällt als der Trend für alle Artikel (0,557). Diese Beobachtung deckt sich mit dem grafischen Trend in Abbildung 4.14 (S. 103), für den die gestrichelte Linie der Top-Artikel zunächst unterhalb der Linie für alle Artikel liegt, jedoch zum Ende des Beobachtungszeitraums sogar leicht darüber liegt und damit stärker ansteigt.

Tabelle 4.6: Lineare Regressionen für die Anzahl der Quellenangaben pro Artikel im EconLit-Datensatz.

	Abhängige Variable: Anzahl Quellenangaben					
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,607*** (0,009)	0,574*** (0,010)	0,557*** (0,016)	0,557*** (0,005)	0,571*** (0,006)	0,587*** (0,007)
Teamgröße		0,131 (0,146)	0,170 (0,193)		0,549*** (0,051)	0,547*** (0,054)
Autorinnen			–1,818** (0,599)			0,552*** (0,166)
Alle Top-Inst			1,507*** (0,312)			0,534*** (0,148)
Alle Nicht-Aka			0,242 (0,723)			–2,506*** (0,216)
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Beobachtungen	24.997	24.997	16.577	212.366	212.366	188.299
R ²	0,164	0,224	0,189	0,066	0,290	0,284

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Die Regressionsergebnisse zeigen einen positiven Zusammenhang zwischen der Teamgröße und der Anzahl angegebener Quellen in einem Artikel. Das heißt, je mehr

Autoren am Verfassen des Artikels beteiligt sind, desto mehr Quellenangaben finden sich im Literaturverzeichnis. Dieser Trend ist für den gesamten EconLit-Datensatz signifikant. Für Top-Artikel fällt dieser Effekt geringer aus und ist statistisch nicht signifikant.²³ Hierin zeigen sich vermutlich die ähnlich hohen Mittelwerte und die nicht eindeutige Reihung der Kurven für verschiedene Teamgrößen aus Abbildung 4.17. Aus den Regressionen geht ferner hervor, dass Teams von Autorinnen etwas mehr zitieren, als andere Teams, wobei der Effekt für Top-Artikel gerade umgekehrt ist und damit nicht robust. Sind alle Autoren einer Top-Institution zugehörig („Alle Top-Inst“), so geben sie im Mittel etwas mehr Quellen an, wobei dieser Effekt bei Top-Artikeln stärker ausgeprägt ist als für alle Artikel. Stammen alle Autoren von einer nicht-akademischen Institution, so zitieren sie für den Gesamtdatensatz weniger. Allerdings geben sie bei Top-Artikeln sogar etwas mehr Quellen an als andere Teams. Die Effekte für die einzelnen Zeitschriften und JEL-Felder sind in Tabelle 4.6 nicht explizit ausgewiesen. Allerdings zeigen sich hier sowohl Unterschiede bezüglich der Zeitschriften, als auch bezüglich der JEL-Felder (vgl. hierzu auch Tabelle 4.20, S. 146). Beispielsweise werden in den Zeitschriften *Econometrica* oder *Economic Journal* mehr Quellenangaben gemacht, als im *American Economic Review*.

Auch für Artikel aus den wissenschaftlichen Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* sowie *Psychologie* im JSTOR-Datensatz und Artikel aus der Zeitschrift *Science* spiegeln sich die zeitlichen Trends aus den Abbildungen 4.14 und 4.16 in den Regressionen wider. Die jährliche Zunahme an Quellenangaben für Artikel aus dem gesamten JSTOR-Datensatz entspricht, ähnlich wie für den EconLit-Datensatz (vgl. Tabelle 4.6), etwas mehr als einer halben Quellenangabe (vgl. Tabelle 4.7). Die in Abbildung 4.15 beobachteten Unterschiede zwischen den Disziplinen zeigen sich auch, wenn man die Regressionen separat für Artikel der einzelnen Disziplinen durchführt (vgl. Tabelle 4.7). Für alle Disziplinen ist der Zeittrend weiterhin positiv und signifikant. An den einzelnen Koeffizienten bzgl. der Kontrolle für das Publikationsjahr erkennt man beispielsweise, dass der zeitliche Effekt für *Mathematik* mit einem Koeffizienten von 0,326 am geringsten und der für *Psychologie* am größten ist (Koeffizient von 1,219). Die Zunahme für Artikel aus der Zeitschrift *Science* fällt mit einer zusätzlichen Quelle alle vier Jahre nur etwa halb so groß aus wie für den gesamten JSTOR-Datensatz oder den EconLit-Datensatz (vgl. Modelle (1) bis (2) in Tabelle 4.8); beide Trends sind jeweils statistisch signifikant. Dieser Unterschied war bereits grafisch in den Abbildungen 4.14 bis 4.16 (S. 103 ff.) zu sehen, da der

²³Auch für Top-Artikel wird der Teamgrößen-Effekt für bestimmte Spezifikationen signifikant, wenn man beispielsweise Modelle betrachtet, in denen lediglich für das Publikationsjahr und die Teamgröße kontrolliert wird. Ein solches Modell ist in Tabelle 4.6 allerdings nicht abgebildet.

Anstieg in der Zeitschrift *Science* insgesamt geringer ausfällt und ab den 1990er Jahren sogar stagniert.

Tabelle 4.7: Lineare Regressionen für die Anzahl der Quellenangaben pro Artikel im JSTOR-Datensatz.

	Abhängige Variable: Anzahl Quellenangaben				
	Gesamt	BWL	Geo	Mathe	Psycho
Publikationsjahr	0,534*** (0,005)	0,652*** (0,010)	0,558*** (0,044)	0,326*** (0,005)	1,219*** (0,020)
Teamgröße	1,305*** (0,077)	0,425** (0,140)	2,371*** (0,641)	1,101*** (0,096)	1,158*** (0,171)
Autorinnen	0,808** (0,274)	0,876 (0,540)	-0,983 (1,901)	-0,146 (0,358)	0,304 (0,557)
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	Ja
Beobachtungen	81.066	34.828	2.933	37.954	9.654
R^2	0,357	0,303	0,170	0,251	0,368

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Für den JSTOR-Datensatz wirkt sich die Teamgröße ebenfalls positiv auf die Anzahl der Quellenangaben aus. Dieser Effekt ist ferner für alle Disziplinen statistisch signifikant, für *Geographie und Geologie* am größten (Koeffizient von 2,371) und für *Betriebswirtschaftslehre und Finance* am kleinsten (0,425). Der Effekt der Teamgröße auf die Anzahl der Quellenangaben fällt in der Zeitschrift *Science* mit einem Koeffizienten von 0,212 am geringsten aus (vgl. Modell (2) in Tabelle 4.8). Interpretiert man diesen Wert, so erscheint der Beitrag von 0,212 Quellenangaben eines zusätzlichen Autors sehr gering. Allerdings zeigt sich hierin die Arbeitsteilung der Koautoren im sinkenden Pro-Kopf-Beitrag hinsichtlich der Anzahl Quellenangaben. Oben war bereits zu sehen, dass die einzelnen Kurven für die Anzahl der Quellenangaben nach Teamgrößen nah beieinander liegen (vgl. z. B. Abbildung 4.17, S. 107). Betrachtet man die Pro-Kopf-Beiträge beispielsweise für das letzte Beobachtungsjahr 2011 für Artikel der Zeitschrift *Science*, so machen Autoren, die einen Artikel alleine schreiben im Mittel 13,6 Quellenangaben. Bei zwei bzw. drei Autoren sind es pro Artikel im Mittel 17,8 bzw. 25,8 Quellenangaben und pro Autor 8,9 bzw. 8,6. In dieser Weise sinkt die Anzahl der Quellenangaben pro Autor mit zunehmender Teamgröße noch

weiter. Diese Beobachtung ist in Abbildung 4.21 beispielhaft für Top-Artikel aus dem EconLit-Datensatz veranschaulicht.

Tabelle 4.8: Lineare Regressionen für die Anzahl der Quellenangaben pro Artikel in der Zeitschrift *Science*.

	Abhängige Variable: Anzahl Quellenangaben					
	Alle Artikel: (1)–(2)		Erste Artikel: (3)–(6)			
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,287*** (0,004)	0,271*** (0,004)	0,327*** (0,004)	0,311*** (0,004)	0,367*** (0,002)	0,335*** (0,003)
Teamgröße		0,212*** (0,013)		0,173*** (0,012)		0,082*** (0,002)
Autorinnen		−3,088*** (0,355)		−2,043*** (0,413)		
Autorin						0,646*** (0,091)
Beobachtungen	57.263	57.263	45.830	45.830	127.261	127.261
R^2	0,078	0,083	0,118	0,123	0,152	0,163

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Bem.: Modelle (1) bis (4) beobachten auf Artikel-Ebene, Modelle (5) und (6) auf Autoren-Ebene.

Wie man sieht weisen Artikel mit nur einem Autoren (durchgezogene Linie) die meisten Quellenangaben pro Autor und Artikel sowie die größte Steigung auf. Es ist ebenfalls zu erkennen, dass auch die Anzahl der Quellenangaben pro Kopf für die einzelnen Teamgrößen im Beobachtungszeitraum steigend ist. Auch für den gesamten EconLit-Datensatz sowie die Disziplinen im JSTOR-Datensatz zeigt sich ein ähnliches Bild, wie man den Abbildungen B.7 und B.8 im Anhang (S. 212) entnimmt. Ähnlich wie für die aggregierten Abbildungen 4.14 bis 4.16 (S. 103 ff.) zur mittleren Anzahl der Quellenangaben, fällt die Zeitschrift *Science* auch bei dieser Betrachtung etwas aus der Reihe. Zwar ist in Abbildung 4.22 ebenfalls die Reihung der Teamgrößen hinsichtlich der Anzahl Quellenangaben pro Autor zu erkennen, allerdings sind diese Trends nicht über den gesamten Beobachtungszeitraum steigend. Für Artikel mit nur einem Autor oder zwei Autoren ist die Anzahl der Quellenangaben nach 1990 sogar fallend. Dies spiegelt den aggregierten, ab 1990 stagnierenden Trend für die mittlere

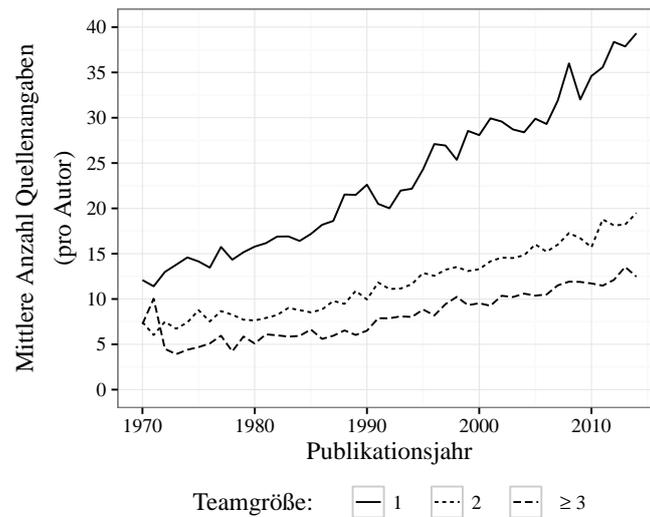


Abbildung 4.21: Mittlere Anzahl Quellenangaben pro Artikel und Autor für Top-Artikel im EconLit-Datensatz für verschiedene Teamgrößen.

Anzahl der Quellenangaben in der Zeitschrift *Science* aus Abbildung 4.16 (S. 105) wider. Wie in Abschnitt 4.1 gezeigt und diskutiert, hat in der Zeitschrift *Science* in den letzten Jahren die Anzahl der Artikel mit zehn oder mehr Autoren stark zugenommen (vgl. Abbildung 4.9, S. 91). Wie oben argumentiert, steigt die Anzahl der Quellenangaben zwar größtenteils mit der Teamgröße, allerdings sind die pro-Kopf-Beiträge zusätzlicher Autoren hinsichtlich der Quellenangaben fallend. Entsprechend führen die besonders großen Teams in der Zeitschrift *Science*, im Vergleich zu den übrigen Disziplinen, zu kleineren pro-Kopf-Beiträgen und somit zum kleineren geschätzten Effekt der Teamgröße auf die Anzahl der Quellenangaben in Tabelle 4.8.

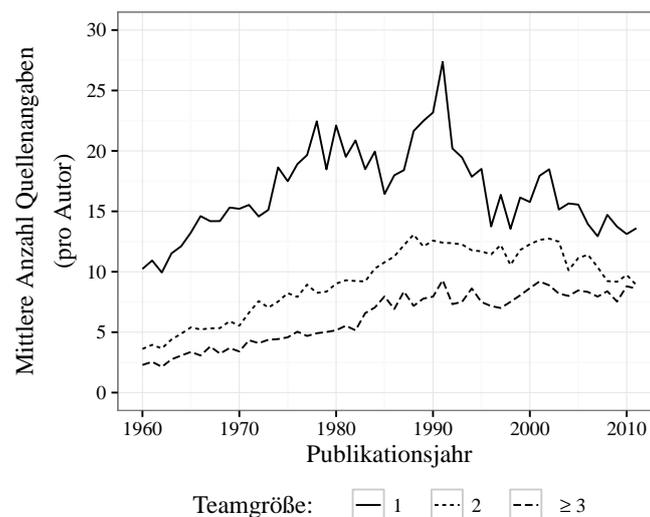


Abbildung 4.22: Mittlere Anzahl Quellenangaben pro Artikel und Autor in der Zeitschrift *Science* für verschiedene Teamgrößen.

Betrachtet man die Gesamtzahl der Artikel im JSTOR-Datensatz, so zeigt sich wie auch für alle Artikel im EconLit-Datensatz, dass Teams von Autorinnen mehr Quellen angeben als gemischt-geschlechtliche oder rein männliche Teams (vgl. Tabelle 4.7). Allerdings sind die Effekte dieser Kontrollvariablen für die einzelnen Disziplinen nicht signifikant und die Effektrichtung nicht für alle positiv. Der Effekt von Autorinnen-Teams auf die Anzahl der Quellenangaben für Artikel in *Science* ist negativ (Tabelle 4.8). Diese Effektrichtung deckt sich mit der für volkswirtschaftliche Top-Artikel (Tabelle 4.6).

4.2.2 Anzahl der Quellenangaben bei erstem Artikel

In Abschnitt 4.2.1 wurde die Anzahl der Quellenangaben in einem Artikel bereits auf Artikel-Ebene untersucht. Für alle Disziplinen sind deutlich steigende Zeittrends in der Anzahl Quellenangaben auszumachen. Diese werden ferner unter diversen Kontrollen (u. a. für die Teamgröße) in den Regressionen als statistisch signifikant bestätigt. Zieht man dieses einfache Maß heran, um die Menge des in einem Artikel verarbeiteten Wissens zu bemessen, so kommt man zu dem Schluss, dass im Laufe der Jahre zunehmend viel Wissen pro Artikel verarbeitet wurde.²⁴ In diesem Abschnitt wird die Anzahl der Quellenangaben sowohl auf Autoren- als auch auf Artikel-Ebene, wobei erstere im Vordergrund steht, für die jeweils ersten Artikel der Autoren betrachtet. Diese markieren den ersten bedeutenden Beitrag eines Wissenschaftlers zu seinem Fachgebiet und den Zeitpunkt, zu dem der Wissenschaftler ausreichend bestehendes Wissen verarbeitet und akkumuliert hat, um einen eigenen Beitrag zu diesem zu leisten. In Abschnitt 4.2.1 wurde also bislang gezeigt, dass in Artikeln der untersuchten Disziplinen zunehmend viel zitiert wird. Es wurde jedoch nicht geprüft, ob dies auch für debütierende Autoren gilt, die mit ihrem ersten Artikel den Übergang aus der Phase der Wissensaneignung und -verarbeitung in die Phase der Wissensschaffung vollziehen, auch wenn die Phase der Wissensaneignung und -verarbeitung vermutlich nie endet.

Im Folgenden werden die ersten Artikel von Autoren ferner danach unterschieden, ob diese lediglich von einem Autor verfasst wurden, ob ein Autor bei seinem ersten Artikel also alleiniger Autor des Artikels war, oder ob er seinen ersten Artikel in Zusammenarbeit mit anderen Autoren verfasst hat. Für beide Kategorien bestehen lediglich geringe Unterschiede zu den Ergebnissen für alle Artikel aus Abschnitt 4.2.1. Die grundsätzlichen qualitativen Trends zeigen sich somit auch für die mittlere Anzahl

²⁴Weitere mögliche Gründe für eine Zunahme der Anzahl Quellenangaben wurden bereits in Kapitel 2 angesprochen und werden in Abschnitt 6.1 nochmals diskutiert.

der Quellenangaben debütierender Autoren. Entsprechend weisen alle betrachteten Disziplinen zumeist deutlich steigende Trends hinsichtlich dieses Maßes auf.

In Abbildung 4.23 ist dies beispielhaft für Top-Artikel des EconLit-Datensatzes dargestellt.²⁵ Man erkennt, dass die beiden Kurven für alle Top-Artikel (gepunktete Linie) und erste Top-Artikel (durchgezogene Linie) nah beieinander liegen und dass die Kurve für alle Artikel etwas oberhalb der für erste Artikel liegt. Ein ähnliches Bild zeigt sich auch für die weiteren untersuchten Disziplinen. Entsprechende Darstellungen für Artikel der Zeitschrift *Science* und für die Disziplinen im JSTOR-Datensatz finden sich in den Abbildungen B.12 und B.13 im Anhang (S. 217). Während die Kurve für die mittlere Anzahl der Quellenangaben für alle Artikel in der Zeitschrift *Science* zunächst leicht über der Kurve für Artikel debütierender Autoren verläuft, liegen die beiden Kurven gegen Ende des Beobachtungszeitraums gleichauf. Für die Disziplinen aus dem JSTOR-Datensatz sind keine systematischen Unterschiede im Verlauf der beiden Kurven auszumachen. Lediglich für *Geographie und Geologie* ist auffällig, dass in den 1960ern und 1970ern für alle Artikel im Mittel deutlich mehr Quellen angegeben wurden als bei ersten Artikeln.

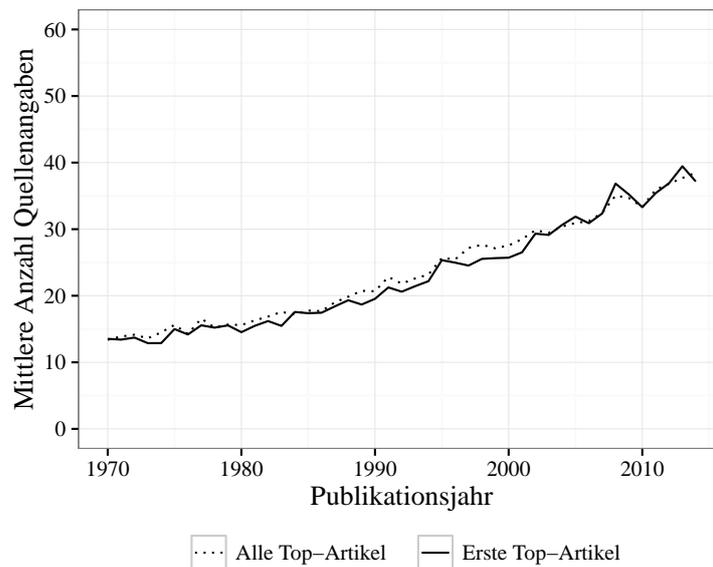


Abbildung 4.23: Mittlere Anzahl Quellenangaben beim ersten Top-Artikel eines Autors im EconLit-Datensatz.

Die steigenden Trends werden auch für erste Artikel in den Regressionen bestätigt und erweisen sich als statistisch signifikant. Die Regressionsergebnisse auf Autoren-Ebene für erste Artikel aus dem EconLit-Datensatz finden sich in Tabelle 4.9, wobei

²⁵Die Abbildungen B.13 und 4.23 beziehen sich auf Autoren-Ebene. Die entsprechenden Abbildungen auf Artikel-Ebene unterscheiden sich allerdings kaum von diesen.

sich die Modelle (1), (2), (4) sowie (5) auf erste Artikel mit nur einem Autor beziehen, während die Modelle (3) und (6) alle ersten Artikel betrachten.²⁶ Wie man sieht, sind die Regressionskoeffizienten bzgl. des Publikationsjahres sowohl für Top-Artikel als auch den gesamten EconLit-Datensatz positiv und hoch signifikant. Ähnlich wie für alle Artikel in Abschnitt 4.2.1 bzw. Tabelle 4.6 (S. 110), entspricht die jährliche Zunahme für die mittlere Anzahl der Quellenangaben bei ersten Artikeln etwas mehr als einer halben Quellenangabe. Wie bereits in Abschnitt 4.2.1 für alle Artikel gesehen, wirkt sich auch bei den ersten Artikeln die Teamgröße positiv auf die Anzahl der im Artikel angegebenen Quellen aus. Je mehr Autoren also an der Verfassung eines Artikels beteiligt sind, desto mehr Quellenangaben weist der Artikel auf. Dieser Effekt ist für alle Artikel im EconLit-Datensatz hoch signifikant und für Top-Artikel auf dem 5 %-Niveau signifikant. Autorinnen geben bei ihrem ersten Top-Artikel im EconLit-Datensatz signifikant weniger Quellen an als ihre männlichen Kollegen (vgl. Modell (3) in Tabelle 4.9). Dieser Effekt kehrt sich allerdings für alle Artikel im EconLit-Datensatz um (vgl. Modell (6) in Tabelle 4.9). Die Zugehörigkeit eines Autors zu einer der renommiertesten Institutionen oder einer akademischen Institution wirkt sich positiv auf die Anzahl der Quellenangaben in ersten Artikeln im EconLit-Datensatz aus (vgl. Zeilen „Top-Inst“ und „Nicht-Aka“).

Auch die Regressionen für erste Artikel der Disziplinen im JSTOR-Datensatz (vgl. Tabelle 4.10) sowie der Zeitschrift *Science* (vgl. Modelle (5) und (6) in Tabelle 4.8) weisen signifikant positive Zeittrends aus. Wie für den EconLit-Datensatz sind hier die Teamgrößeneffekte statistisch größtenteils positiv und statistisch signifikant. Eine Ausnahme hinsichtlich der Effektrichtung stellen erste Artikel der *Betriebswirtschaftslehre und Finance* dar. Für diese Disziplin weist die Regression einen negativen Zusammenhang zwischen der Teamgröße und der Anzahl der Quellenangaben bei ersten Artikeln auf (vgl. Modell „BWL“ in Tabelle 4.10).²⁷ Für die Disziplinen im JSTOR-Datensatz zeigen sich keine signifikanten Effekte einer Autorin auf die mittlere Anzahl angegebener Quellen. In der Zeitschrift *Science* hingegen ist der geschätzte Koeffizient für eine Autorin positiv und hoch signifikant. Das bedeutet, dass beim ersten Artikel einer

²⁶Die Tabellen 4.9 und 4.10 weisen Regressionsergebnisse für die Anzahl der Quellenangaben auf Autoren-Ebene aus. Die qualitativen Ergebnisse bzgl. der Effektrichtungen und Signifikanzniveaus bleiben allerdings erhalten, wenn man diese auf Artikel-Ebene durchführt, wie man den Tabellen B.1 und B.2 im Anhang (S. 214 und 215) für den EconLit- und JSTOR-Datensatz sowie den Modellen (3) und (4) in Tabelle 4.8 für die Zeitschrift *Science* entnehmen kann.

²⁷Berechnet man die Regressionen auf Artikel-Ebene, das heißt für Artikel mit mindestens einem Debütanten, so ist der geschätzte Teamgrößen-Koeffizient für *Betriebswirtschaftslehre und Finance* zwar nicht mehr negativ, allerdings beträgt er lediglich 0,065 und ist statistisch nicht signifikant (vgl. Modell „BWL“ in Tabelle B.2, S. 215, im Anhang).

Autorin in der Zeitschrift *Science* mehr Quellen angeben werden als bei ersten Artikeln von männlichen Autoren.

Tabelle 4.9: Lineare Regressionen für die Anzahl der Quellenangaben beim ersten Artikel eines Autors im EconLit-Datensatz.

	Abhängige Variable: Anzahl Quellenangaben					
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,595*** (0,016)	0,517*** (0,019)	0,573*** (0,012)	0,574*** (0,013)	0,525*** (0,020)	0,628*** (0,009)
Autorin		-1,677** (0,581)	-1,759*** (0,333)		-0,230 (0,385)	0,491** (0,155)
Top-Inst		3,276*** (0,530)	1,903*** (0,306)		2,889*** (0,500)	0,497* (0,228)
Nicht-Aka		-1,182 (0,958)	-0,933* (0,450)		-2,226*** (0,675)	-3,617*** (0,241)
Teamgröße			0,314* (0,147)			0,505*** (0,065)
Zeitschrift (Dummies)		Ja	Ja		Ja	Ja
JEL-Feld (Dummies)		Ja	Ja		Ja	Ja
Beobachtungen	4.406	4.406	15.176	19.214	19.214	87.319
R^2	0,238	0,297	0,308	0,089	0,359	0,325

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Bem.: Die Modelle (1), (2), (4) und (5) betrachten erste Artikel mit nur einem Autor. Die Modelle (3) und (6) betrachten alle ersten Artikel.

Zuletzt werden Artikel mit nur einem Autor betrachtet (wie oben erwähnt auch als Solo-Artikel bezeichnet). Bei diesen Artikeln lässt sich die Anzahl der Quellenangaben eindeutig auf einen Autor zurückführen. Wie gerade für alle ersten Artikel diskutiert, bleiben die qualitativen Trends auch für erste Artikel mit nur einem Autor bestehen. Das bedeutet, dass die mittlere Anzahl der Quellenangaben auch für diese Artikel klar steigend ist. Für den EconLit-Datensatz ist zu beobachten, dass die mittlere Anzahl der Quellenangaben beim ersten Artikel eines Autors, wenn er bei diesem alleiniger Autor ist, sogar etwas größer ist als die für alle Artikel (vgl. Abbildung 4.24), für alle Artikel mit nur einem Autor (vgl. Abbildung B.9, S. 215, im Anhang) sowie für alle

Tabelle 4.10: Lineare Regressionen für die Anzahl der Quellenangaben beim ersten Artikel eines Autors für den JSTOR-Datensatz.

	Abhängige Variable: Anzahl Quellenangaben				
	Gesamt	BWL	Geo	Mathe	Psycho
Publikationsjahr	0,651*** (0,006)	0,736*** (0,011)	0,510*** (0,058)	0,356*** (0,007)	1,327*** (0,016)
Teamgröße	0,960*** (0,057)	-0,386** (0,128)	2,566*** (0,485)	1,154*** (0,098)	0,573*** (0,083)
Autorin	0,558** (0,201)	0,230 (0,365)	-1,597 (1,487)	-0,251 (0,310)	-0,108 (0,361)
Beobachtungen	64.113	24.839	2.214	27.383	12.795
R^2	0,410	0,366	0,229	0,250	0,428

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

ersten Artikel (vgl. Abbildung B.11, S. 216, im Anhang). Im Gegensatz zum EconLit-Datensatz, gibt ein Autor in den Disziplinen im JSTOR-Datensatz bei seinem ersten Artikel, wenn er bei diesem alleiniger Autor ist, wie man erwarten würde, weniger Quellen an als Autoren bei allen ersten Artikeln (vgl. Abbildung 4.25). Gleiches gilt auch für die Zeitschrift *Science*. Dennoch weisen erste Artikel mit nur einem Autor in den einzelnen Disziplinen im JSTOR-Datensatz klar steigende Trends auf, die den grafischen Trends für alle Artikel in Abbildung 4.15 (S. 105) ähneln. Ähnlich wie dies bereits für alle Artikel der Zeitschrift *Science* mit nur einem oder zwei Autoren zu sehen war (vgl. Abbildung 4.18, S. 108), weisen auch erste Artikel mit nur einem Autor in der Zeitschrift *Science*, im Vergleich zu den anderen Disziplinen, eine Besonderheit auf. Abbildung 4.26 veranschaulicht die mittlere Anzahl der Quellenangaben bei allen Artikeln (gepunktete Linie) sowie ersten Artikeln mit nur einem Autor (durchgezogene Linie). Wie man sieht, unterscheiden sich die Linien in ihrem Verlauf kaum voneinander. Die durchgezogene Linie für erste Artikel mit nur einem Autor liegt für die meisten Jahre etwas unter der gepunkteten für alle Solo-Artikel. Beide steigen zunächst bis zum Jahr 1991 an und sind danach fallend.

In den Modellen (1) und (2) bzw. (4) und (5) in Tabelle 4.9 finden sich die Regressionen zur Anzahl der Quellenangaben beim ersten Artikel im EconLit-Datensatz, wenn der Autor bei diesem alleiniger Autor ist. Die grafischen Trends aus den Abbildungen

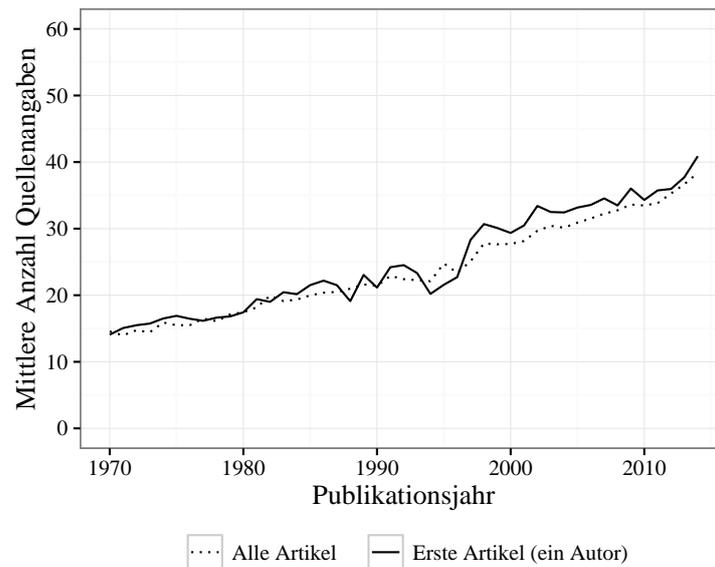


Abbildung 4.24: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors im EconLit-Datensatz, wenn er bei diesem alleiniger Autor war.

4.24 sowie B.9 bis B.11 (S. 215 ff. im Anhang) werden in den Regressionen bestätigt und erweisen sich als statistisch signifikant. In allen Modellen steigt die Anzahl der Quellenangaben mit jedem Jahr um etwas mehr als eine halbe Quellenangabe. Die Regressionsergebnisse zeigen ferner, dass Frauen bei ihrem ersten Artikel, wenn dieser ein Solo-Artikel ist, weniger Quellen angeben als ihre männlichen Kollegen. Dieser Effekt ist lediglich für Top-Zeitschriften signifikant. Autoren der renommiertesten Institutionen geben mehr Quellen an als ihre Kollegen von anderen Institutionen, während Autoren von nicht akademischen Institutionen weniger Quellenangaben machen als Autoren von Universitäten oder anderen Forschungseinrichtungen. Dieser Effekt ist für die Gesamtheit der ersten Artikel mit nur einem Autor des EconLit-Datensatzes signifikant, allerdings nicht für die Teilmenge der Top-Artikel.

Die Regressionsmodelle für erste Artikel mit nur einem Autor aus dem JSTOR-Datensatz und der Zeitschrift *Science*, finden sich in Tabelle 4.11. Wie für die volkswirtschaftlichen Artikel aus dem EconLit-Datensatz, bestätigen auch die Regressionen für Disziplinen im JSTOR-Datensatz die beobachteten positiven Zeittrends (vgl. Abbildung 4.25). Der Koeffizient bzgl. des Publikationsjahres ist signifikant und beträgt für alle vier Disziplinen des JSTOR-Datensatzes zusammengenommen 0,394 (Spalte „Gesamt“). Auch für die Zeitschrift *Science* (Spalte „*Science*“) ergibt die Regressionsanalyse einen positiven und statistisch signifikanten Zeittrend. Der Effekt eines weiblichen Autors ist in keinem der Modelle statistisch signifikant.²⁸ Im Modell „Ge-

²⁸Streng genommen ist der Effekt in der Spalte „Psycho“ auf dem 10 %-Niveau signifikant.

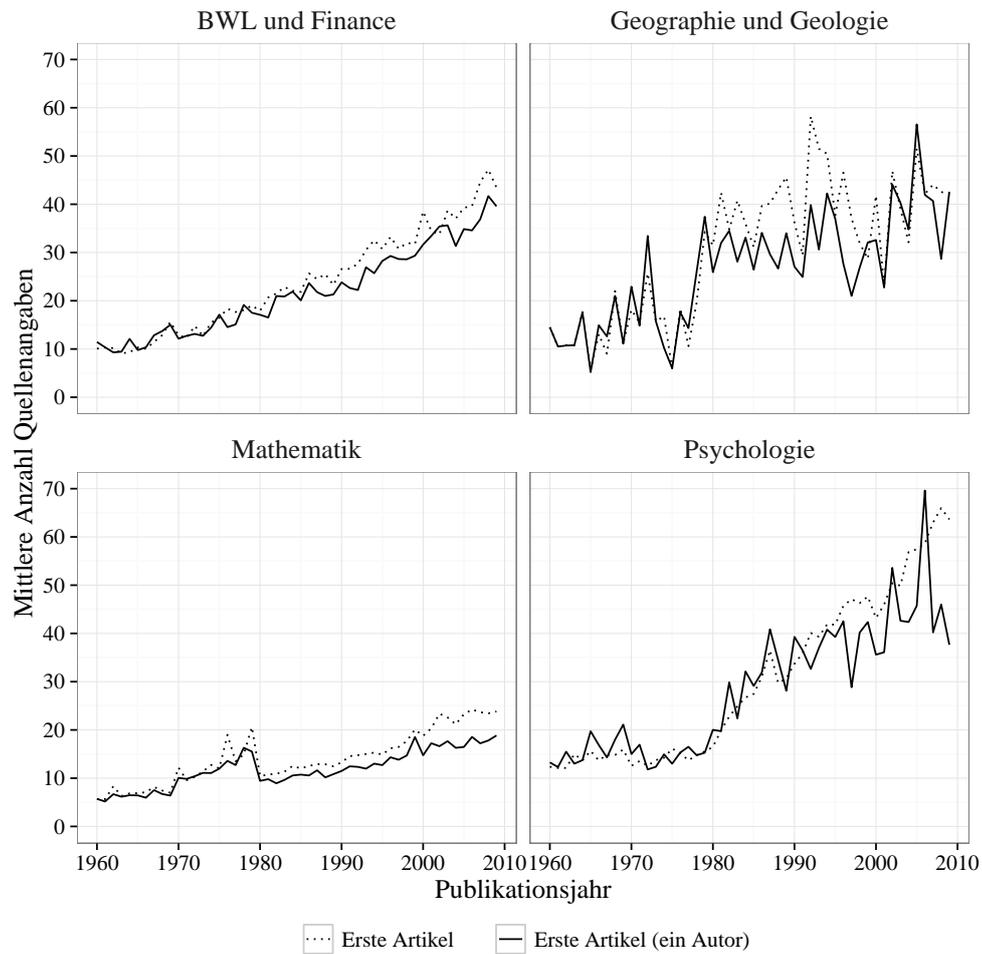


Abbildung 4.25: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors im JSTOR-Datensatz, wenn er bei diesem alleiniger Autor war.

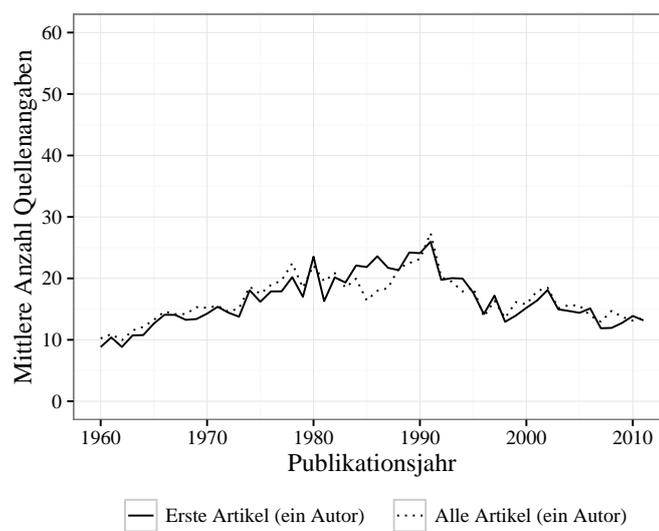


Abbildung 4.26: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors in der Zeitschrift *Science*, wenn er bei diesem alleiniger Autor war.

samt“ ist die Effektrichtung zwar positiv, das heißt die Anzahl der Quellenangaben ist für Frauen größer als für Männer, allerdings gibt es zwischen den Disziplinen Unterschiede. So ist die Effektrichtung für *BWL und Finance* sowie *Mathematik* negativ, während sie für *Geographie und Geologie* sowie *Psychologie* positiv ist. Der geschätzte Koeffizient für weibliche Autoren in der Zeitschrift *Science* ist nicht signifikant und fällt mit 0,073 deutlich geringer aus als für den EconLit- oder JSTOR-Datensatz. Wie erwähnt, ist der grafische Trend in Abbildung 4.26 für die Anzahl der Quellenangaben bei ersten Artikeln in der Zeitschrift *Science*, sofern diese von nur einem Autor verfasst wurden, nicht linear. Dies zeigt sich auch im geringen R^2 -Wert von 0,005 im Modell „*Science*“ in Tabelle 4.11.

Tabelle 4.11: Lineare Regressionen für die Anzahl der Quellenangaben beim ersten Artikel eines Autors im JSTOR-Datensatz und der Zeitschrift *Science*, wenn er bei diesem alleiniger Autor war.

	Abhängige Variable: Anzahl Quellenangaben					
	Gesamt	BWL	Geo	Mathe	Psycho	<i>Science</i>
Publikationsjahr	0,394*** (0,009)	0,597*** (0,020)	0,393*** (0,073)	0,275*** (0,008)	0,982*** (0,062)	0,071*** (0,013)
Autorin	0,480 (0,360)	-0,556 (0,734)	0,458 (2,279)	-0,586 (0,385)	1,992+ (1,200)	0,073 (0,640)
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	Ja	
Beobachtungen	17.578	5.787	787	10.290	1.359	6.996
R^2	0,344	0,295	0,225	0,255	0,280	0,005

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

4.3 Länge von Titeln und Abstracts

Wie bei der Vorstellung der Maße in Abschnitt 3.6 (S. 76) erläutert, wurde die Länge der Titel von Artikeln in diversen Studien herangezogen, um auf deren Informationsgehalt und den Informationsgehalt bzw. die Komplexität der zugehörigen Artikel zu schließen (z. B. Tocatlian 1970; Buxton und Meadows 1977; Diener 1984; White und Hernandez 1991). Diodato (1982) zeigt, dass die Länge von Titeln für alle von ihm untersuchten Disziplinen (Chemie, Ökonomie, Geschichte, Philosophie und Mathematik) zunimmt, außer für Geschichte und Philosophie. Buxton (1987) berichtet, dass sich das Wachstum

der Titellänge seit seiner Untersuchung in Buxton und Meadows (1977) verlangsamt hat. In diesem Abschnitt werden die Titel der dieser Arbeit vorliegenden Artikel untersucht.

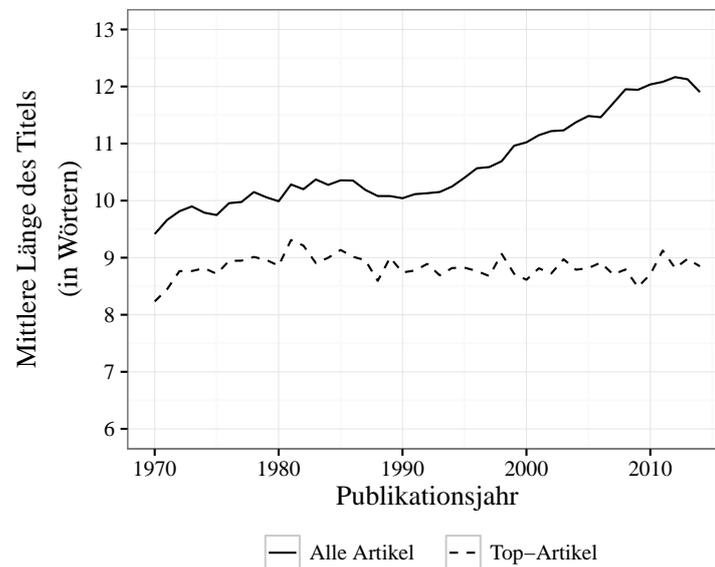


Abbildung 4.27: Mittlere Anzahl Wörter im Titel von Artikeln im EconLit-Datensatz.

Abbildung 4.27 zeigt den Trend für die mittlere Länge der Titel im EconLit-Datensatz. Für die Gesamtheit der Artikel im EconLit-Datensatz ist ein deutlicher Anstieg von 9,4 Wörtern im Jahr 1970 auf 11,9 im Jahr 2014 zu beobachten, was einem prozentualen Anstieg von 26,4 % entspricht. Interessanterweise ergibt sich für Top-Artikel im EconLit-Datensatz ein anderes Bild (gestrichelte Linie in Abbildung 4.27). Die Länge der Titel von Top-Artikeln liegt über den gesamten Zeitraum unter der für alle Artikel, wobei der Unterschied 1970 etwa ein Wort und 2014 fast drei Wörter beträgt. Zwischen 1970 und 2014 ist zwar auch für Top-Artikel ein Anstieg von 8,2 auf 8,8 Wörter zu verzeichnen, allerdings resultiert dieser v. a. aus einem starken Anstieg von 1970 bis 1981 zu Beginn des Beobachtungszeitraums. Daraufhin nimmt die mittlere Länge des Titels für Top-Artikel leicht ab.²⁹ Im Unterschied zu den Titeln von

²⁹In den Leitfäden zu Einreichungen in den einzelnen Top-Zeitschriften finden sich keine Restriktionen bzgl. der Länge des Titels. Lediglich das *Quarterly Journal of Economics* schreibt hierzu: „the title should be short, specific, and informative“ (vgl. https://academic.oup.com/qje/pages/Instructions_To_Authors, Stand: 31.05.2018).

Top-Artikeln, sind die entsprechenden Abstracts im Laufe der Zeit länger geworden (vgl. Abbildung 4.28).^{30,31,32}



Abbildung 4.28: Mittlere Anzahl Wörter im Abstract von Top-Artikeln im EconLit-Datensatz.

Wie ist zu erklären, dass die Länge der Titel für alle Artikel deutlich zugenommen hat, während die von Top-Artikeln stagniert? Eine Funktion der Titel von Artikeln ist diese von anderen Artikeln zu unterscheiden (Bird und Knight 1975, zitiert nach White und Hernandez, 1991). Beispielsweise differenziert Hirsch seinen Artikel aus dem Jahr 2010, in welchem er eine Weiterentwicklung des *h-Index* vorschlägt, von seinem ursprünglichen Artikel „An index to quantify an individual’s scientific research output“ aus dem Jahr 2005, indem er im Titel die Spezifizierung „that takes into account the effect of multiple coauthorship“ ergänzt. Ein anderes Beispiel ist Lehman, der zwei seiner Arbeiten zur Produktivität von Individuen mit „Age and Achievement“ (1953) bzw. „More about Age and Achievement“ (1962) betitelt. Ähnliche Differenzierungen, bei denen die Titel der Artikel mit der Zeit länger werden, sind aber auch für Artikel unterschiedlicher Autoren zu beobachten. Ein solches Beispiel sind die folgenden drei Artikel bzw. deren Titel: „The Demand for Leisure“ (Owen 1971), „The Demand for

³⁰Im Gegensatz zu den Titeln von Artikeln (vgl. Fußnote 29), finden sich in den Leitfäden zu Einreichungen bei Zeitschriften zumeist klare Vorgaben für die Länge des Abstracts. Für die Top-Zeitschriften reichen diese von 100 Wörtern (z. B. für das *American Economic Review*) bis 250 Wörtern im *Quarterly Journal of Economics* (Stand: Juni 2018). Inwiefern diese oberen Grenzen im Laufe der Zeit angepasst wurden ist nicht bekannt.

³¹Da der EconLit-Datensatz keine Abstracts umfasst, stammen diese für Top-Artikel, aus dem JSTOR-Datensatz, der die neun Top-Zeitschriften umfasst.

³²Der JSTOR-Datensatz umfasst insgesamt lediglich 20 renommierte Zeitschriften der Volkswirtschaftslehre, sodass der gesamte EconLit-Datensatz nicht adäquat repräsentiert wird. Dennoch sei erwähnt, dass sich die Kurve für die Anzahl der Wörter im Abstract aller volkswirtschaftlichen Artikel im JSTOR-Datensatz mit der für Top-Artikel in Abbildung 4.28 deckt.

Leisure and Money“ (Phlips 1978) und „The Demand for Leisure and Nonpecuniary Job Characteristics“ (Atrostic 1982).³³

Wie in Abschnitt 3.1 (S. 40) beschrieben, stammen die Top-Artikel aus neun der renommiertesten Zeitschriften der Volkswirtschaftslehre. Die Auswahl dieser Zeitschriften basiert unter anderem darauf, dass es sich um traditionsreiche Zeitschriften handelt, die über den gesamten Zeitraum von 1970–2014 herausgegeben wurden und die Artikel publizieren, die für ein breites Publikum von Interesse sind. Möglicherweise ist der beobachtete Unterschied in Abbildung 4.27 zwischen allen Artikeln im EconLit-Datensatz und den Top-Artikeln dadurch zu erklären, dass in Top-Artikeln größtenteils über Erkenntnisse berichtet wird, die eine ggf. grundlegende, allgemein relevante Fragestellung für ein breites Publikum beantworten, entsprechend nicht so stark von vorhandener Literatur differenziert werden müssen und daher kürzere Titel aufweisen. Diese Erklärung ist allerdings spekulativ und basiert nicht auf empirischer Evidenz.

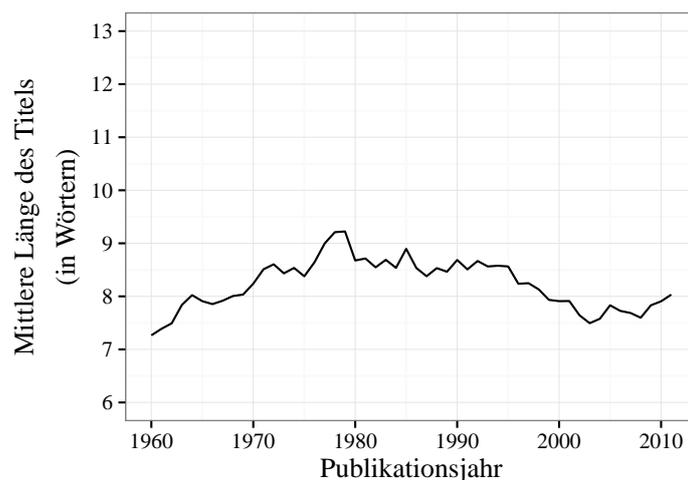


Abbildung 4.29: Mittlere Anzahl Wörter im Titel von Artikeln in *Science*.

Ähnlich wie für volkswirtschaftliche Top-Artikel, zeigt sich auch für Artikel der Zeitschrift *Science* kein klarer, steigender Trend für die Länge der Titel (Abbildung 4.29). Zwar sind Titel im Jahr 2011 mit durchschnittlich 8,0 Wörtern länger als zu Beginn des Beobachtungszeitraum im Jahr 1960 (\varnothing 7,3 Wörter), allerdings beschreibt die Kurve annähernd den Verlauf einer nach unten geöffneten Parabel mit einem Hoch Ende der 1970er Jahre. Ähnliche Argumente bzgl. der zunehmenden Anzahl spezialisierter

³³Es finden sich auch Beispiele, in denen später publizierte Artikel kürzere Titel aufweisen. Dies gilt beispielsweise für die beiden Artikel „A Bargaining Model with Incomplete Information about Time Preferences“ (Rubinstein 1985) und „A Bargaining Model with Incomplete Information“ (Bikhchandani 1992). Diese Fälle sind allerdings seltener zu beobachten als die im Text beschriebenen.

Zeitschriften, wie für die Länge von Titeln der Top-Artikel aus dem EconLit-Datensatz könnten mögliche Erklärungen für diesen Verlauf sein. Analog zu den Top-Artikeln im EconLit-Datensatz wurden auch für *Science* zusätzlich die Abstracts der Artikel betrachtet. Wie Abbildung 4.30 zu entnehmen ist, hat deren Länge im Laufe des Beobachtungszeitraums deutlich zugenommen, auch wenn der Verlauf ungleichmäßiger ist als für die Abstracts der volkswirtschaftlichen Top-Artikel (vgl. Abbildung 4.28).

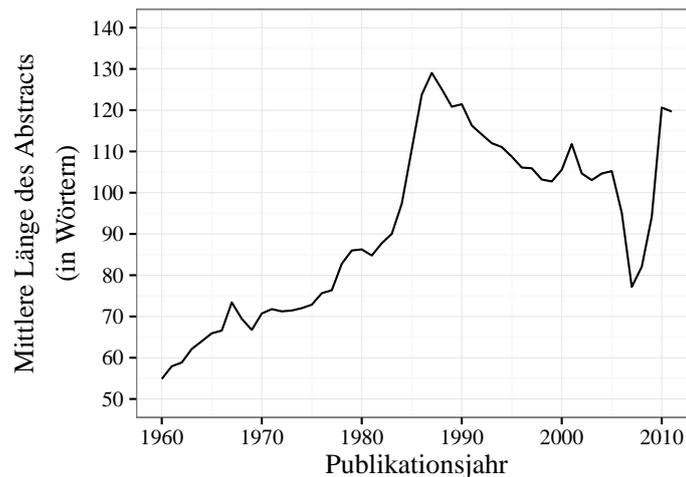


Abbildung 4.30: Mittlere Anzahl Wörter im Abstract von Artikeln in *Science*.

Wie in Abbildung 4.31 zu sehen ist, steigt die Länge des Titels, gemessen an der Anzahl der Wörter, auch für die Disziplinen im JSTOR-Datensatz an. Betrachtet man dabei die Mittelwerte über den gesamten Zeitraum von 1960–2009, so weist die *Psychologie* mit durchschnittlich 10,7 Wörtern die längsten Titel auf. Darauf folgen *Geographie und Geologie* (\varnothing 8,9 Wörter), *BWL und Finance* (\varnothing 8,9 Wörter) sowie *Mathematik* (\varnothing 8,6 Wörter). Vergleicht man die Länge der Titel zu Beginn und am Ende des Beobachtungszeitraums, erfahren Artikel aus *BWL und Finance* den größten prozentualen Zuwachs von 33,6 %. Hier liegt die mittlere Titellänge 1960 noch bei 7,2 und 2009 bei 9,7 Wörtern. Dabei ist dieser Anstieg gewissermaßen stabiler bzw. stetiger als beispielsweise für *Geographie und Geologie*, da die mittlere Titellänge geringeren Schwankungen unterliegt (vgl. Abbildung 4.31). Der zweitgrößte prozentuale Anstieg ist für die *Psychologie* zu beobachten. Mit Titellängen von 12,1 Wörtern im Jahr 2009 und 9,5 im Jahr 1960 entspricht der relative Anstieg hier 26,9 %. Die Länge der Titel in der *Geographie und Geologie* bzw. *Mathematik* hat jeweils um 22,0 % bzw. 16,2 % zugenommen. Artikel bzw. deren Titel aus *BWL und Finance* erfahren, ähnlich wie die volkswirtschaftlichen Artikel (vgl. Abbildung 4.27), einen besonders großen Zuwachs zu Beginn des Beobachtungszeitraums. Danach ist weiterhin ein Anstieg zu beobachten, allerdings fällt dieser flacher aus. Abbildung B.14 für die mittlere Länge der Abstracts

im JSTOR-Datensatz findet sich im Anhang (S. 218). Dort ist zu erkennen, dass auch für *Betriebswirtschaftslehre und Finance* sowie *Psychologie* die Länge der Abstracts zugenommen hat. Nachdem die Abstracts in der *Mathematik* zunächst kürzer werden, nimmt ihre mittlere Länge ab Anfang der 1980er Jahre zu. Für die *Geographie und Geologie* ist die mittlere Länge der Abstracts hingegen über den Beobachtungszeitraum hinweg leicht fallend.

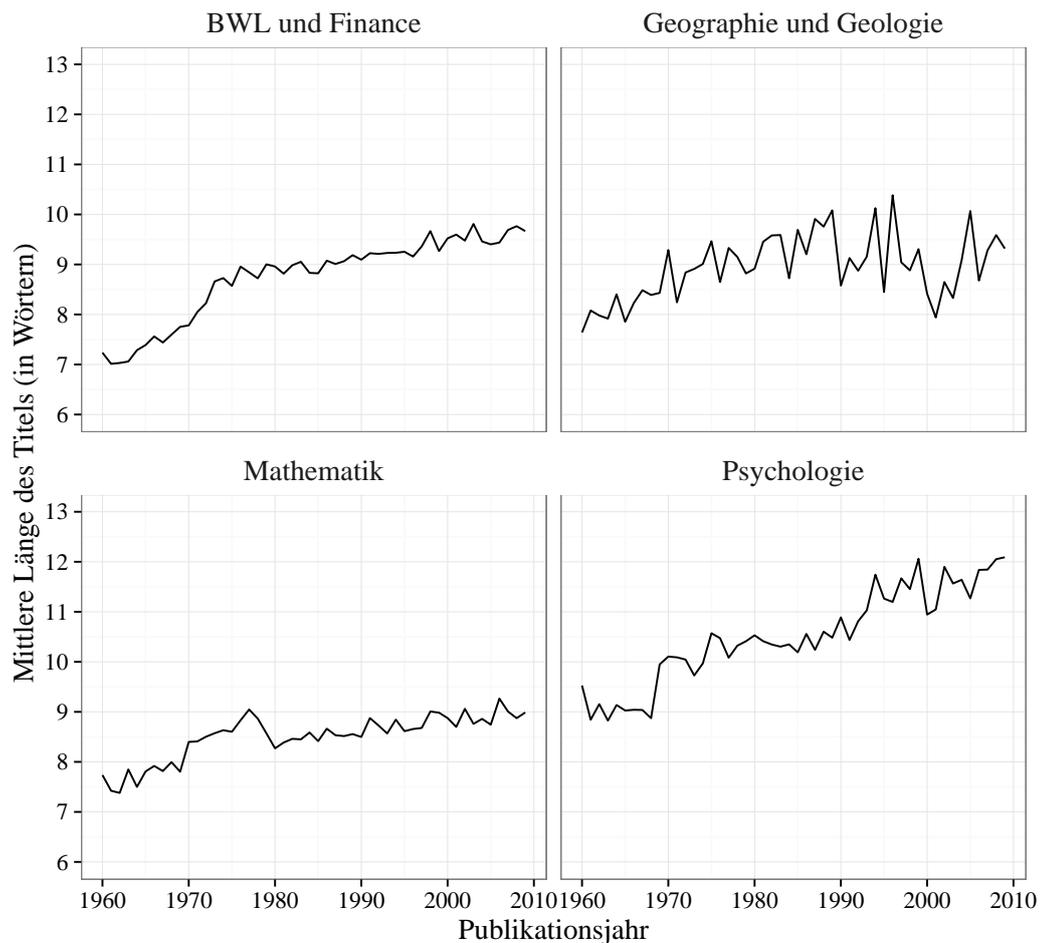


Abbildung 4.31: Mittlere Anzahl Wörter im Titel von Artikeln im JSTOR-Datensatz.

Tabelle 4.12 fasst die Regressionsergebnisse für die Länge des Titels zusammen. Die grafischen Beobachtungen werden weitestgehend bestätigt. So nimmt die Länge der Titel für alle Artikel im EconLit-Datensatz (Modell „EconLit“) und für die Disziplinen im JSTOR-Datensatz zu. Ferner ist der Tabelle zu entnehmen, dass die Länge der Titel für alle Datensätze positiv mit der Teamgröße korreliert und dass dieser Effekt bis auf das Modell für *Mathematik* hoch signifikant ist. Die Regression für Top-Artikel im EconLit-Datensatz ist in Tabelle 4.12 nicht abgebildet. Der Regressionskoeffizient bzgl. des Publikationsjahres ist für Top-Artikel negativ (-0,008). Dieser Effekt ist auf

dem %-Niveau signifikant. Der Koeffizient bzgl. der Teamgröße beträgt 0,111 und ist ebenfalls auf dem %-Niveau signifikant. Der Wert des R^2 ist 0,037. Analog zu den Ergebnissen für Top-Artikel aus dem EconLit-Datensatz, weist die Regression für *Science*-Artikel („*Science*“) einen fallenden Zeittrend aus, obwohl grafisch sogar ein leichter Anstieg der Länge der Titel zwischen 1960 und 2011 zu beobachten ist (Abbildung 4.29). Entsprechend ist der steigende grafische Trend vermutlich auf den Anstieg der Teamgröße zurückzuführen, die positiv mit der Anzahl der Wörter im Titel korreliert. Die geringen R^2 -Werte für die Top-Artikel und das Modell „*Science*“ bestätigen den optischen Eindruck, dass die Trends für volkswirtschaftliche Top-Artikel und *Science*-Artikel anhand der linearen Regression nicht adäquat abgebildet werden.

Tabelle 4.12: Lineare Regressionen für die Länge des Titels in Wörtern.

		Abhängige Variable: Titellänge					
		EconLit	BWL	Geo	Mathe	Psycho	<i>Science</i>
Publikationsjahr		0,017*** (0,001)	0,040*** (0,001)	0,044*** (0,004)	0,025*** (0,001)	0,065*** (0,003)	-0,018*** (0,001)
Teamgröße		0,243*** (0,007)	0,231*** (0,020)	0,596*** (0,073)	0,039+ (0,020)	0,294*** (0,028)	0,151*** (0,003)
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	Ja	Ja	
Beobachtungen		751.394	43.672	6.213	48.183	10.774	67.389
R^2		0,351	0,087	0,228	0,064	0,088	0,046

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Zusammenfassend lässt sich somit sagen, dass die Länge der Titel von Artikeln, gemessen an der Anzahl der Wörter, für alle untersuchten Disziplinen über die Zeit zunimmt und signifikant positiv mit der Teamgröße korreliert. Je mehr Autoren also an der Verfassung eines Artikels beteiligt sind, desto länger ist auch dessen Titel. Dieses Ergebnis bzgl. des Effekts der Teamgröße deckt sich mit den Ergebnissen von Yitzhaki (1994) sowie Rath und Wohlrabe (2016), die ebenfalls eine positive Korrelation zwischen Titellänge und Teamgröße für die Natur- und Sozialwissenschaften finden.³⁴ Der Anstieg in der Länge des Titels fällt für volkswirtschaftliche Top-Artikel und die Zeitschrift *Science* jeweils deutlich geringer aus als für die übrigen Disziplinen, wobei der Verlauf der Trendkurve für *Science* annähernd ein umgekehrtes U beschreibt. Auch

³⁴Der Fokus der Untersuchung von Rath und Wohlrabe (2016) liegt auf der Größe von Teams in der Volkswirtschaftslehre. Allerdings kontrollieren sie in den Regressionen u. a. für die Titellänge.

die mittlere Länge der Abstracts hat über den Beobachtungszeitraum größtenteils zugenommen. Allerdings zeigt sich auch, dass zwischen dem qualitativen Trend für die Länge des Titels und dem für die Länge des Abstracts teilweise Unterschiede bestehen, die darauf hindeuten, dass beide Elemente eines Artikels doch etwas unterschiedliche Informationen widerspiegeln. Obwohl beispielsweise die Titellänge für Top-Artikel seit Anfang der 1980er Jahre leicht rückläufig ist (vgl. gestrichelte Linie in Abbildung 4.27), steigt die mittlere Länge der entsprechenden Abstracts deutlich (vgl. Abbildung 4.28). Ferner steigt die Länge der Titel für *Geographie und Geologie*, während die mittlere Anzahl der Wörter in den zugehörigen Abstracts über den gesamten Beobachtungszeitraum fällt.

4.4 Alter bei erstem Artikel

Das Alter beim ersten Artikel eines Autors ist das zentrale Maß dieser Arbeit, um ggf. die Last des Wissens zu bemessen. Junge Wissenschaftler müssen vermeintlich eine zunehmend große Menge an Wissen verarbeiten, bevor sie selbst neue Erkenntnisse und neues Wissen zu ihrem Fachbereich beitragen können. Jones' (2009a) Hypothese folgend müssten diese Wissenschaftler (in seiner Studie Patentanmelder) bei ihrer ersten bedeutenden Entdeckung im Laufe der Zeit zunehmend älter werden. Jones (2009a) untersuchte Patentanmelder und betrachtete deren Alter bei ihrer ersten Patentanmeldung. Die Patentanmelder aus Jones' Studie entsprechen in der vorliegenden Untersuchung den Autoren von Artikeln und der Zeitpunkt ihrer ersten bedeutenderen Entdeckung (bei Jones der Zeitpunkt der ersten Patentanmeldung), dem Zeitpunkt der Veröffentlichung ihres ersten Artikels.

Um Ausreißer zu eliminieren, ist die Untersuchung des Alters auf das Zeitfenster eingeschränkt, in dem Menschen am produktivsten sind. Wie bereits in Kapitel 2 beschrieben, untersuchte Lehman (1953) den Zusammenhang zwischen dem Alter und der Produktivität von Individuen in verschiedenen Bereichen und fand heraus, dass Menschen vermehrt in ihren Dreißigern bzw. Anfang ihrer Vierziger am produktivsten sind. Diese Beobachtung deckt sich auch mit den in der vorliegenden Arbeit erhobenen Daten. Entsprechend wurde die Untersuchung auf Autoren eingeschränkt, die zum Zeitpunkt ihres ersten Artikels zwischen 25 und 45 Jahre alt sind. Diese Einschränkung nimmt auch Jones (2009a) für die Patentanmelder bei ihrer ersten Patentanmeldung vor. Um die Robustheit der Ergebnisse zu prüfen und um die Vergleichbarkeit mit Jones (2009a) zu gewährleisten, wurden ferner weitere Altersspannen untersucht. Die

Ergebnisse decken sich dabei auch für andere Altersspannen weitestgehend mit denen für die Altersspanne 25–45. Analog zu Jones (2009a) sind in den Regressionstabellen auch Ergebnisse zur Altersspanne 25–35 ausgewiesen. Die entsprechenden Abbildungen finden sich im Anhang in Abschnitt B.5 (S. 219).

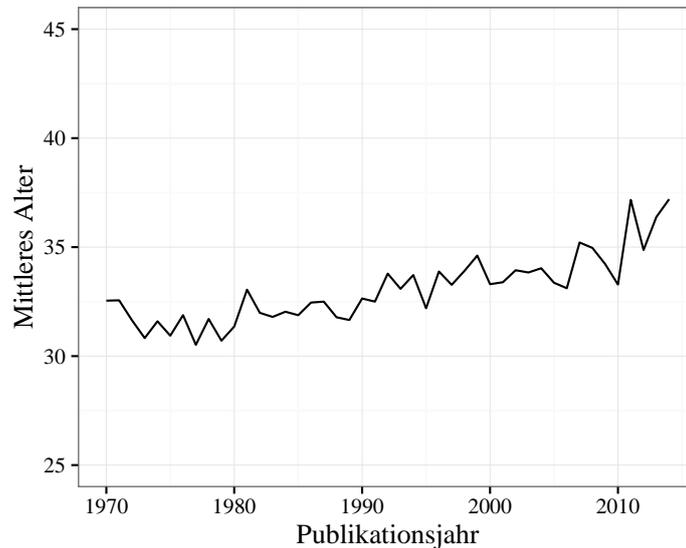


Abbildung 4.32: Mittleres Alter bei erstem Artikel in der Volkswirtschaftslehre.

Altersdaten in der Altersspanne 25–45 konnten, wie bereits in Abschnitt 3.5.2 (S. 73) beschrieben, für 2.656 Autoren im EconLit-Datensatz und 840 Mathematiker aus dem JSTOR-Datensatz erfasst werden.³⁵ Die Untersuchung für die *Volkswirtschaftslehre*, das heißt die Artikel im EconLit-Datensatz, gliedert sich wie bereits gesehen in alle Artikel und Top-Artikel. Entsprechend ist auch die Untersuchung des Alters beim ersten Artikel in das Alter beim ersten Top-Artikel, der vermeintlich besonders bedeutenden Publikation, und das Alter beim ersten Artikel im gesamten EconLit-Datensatz geteilt. Die Abbildungen 4.32 und 4.33 veranschaulichen das mittlere Alter von Autoren in der Volkswirtschaftslehre, die bei ihrem ersten Artikel bzw. Top-Artikel zwischen 25 und 45 Jahre alt waren. Es ist ein deutlicher Aufwärtstrend zu erkennen. Im Jahr 1970 liegt das mittlere Alter beim ersten Artikel noch bei 32,5 Jahren, geht dann in den ersten Jahren auf 30,8 im Jahr 1973 zurück und steigt anschließend über den Beobachtungszeitraum hinweg auf 37,2 im Jahr 2014 (Abbildung 4.32).

Das mittlere Alter für Autoren bei ihrem ersten Top-Artikel ist in Abbildung 4.33 dargestellt. Auch hier ist im Beobachtungszeitraum ein deutlicher Anstieg zu beobachten. 1970 liegt das mittlere Alter beim ersten Top-Artikel bei 32,6 Jahren und steigt

³⁵Als Mathematiker werden hier Autoren bezeichnet, die in einer der Zeitschriften publiziert haben, die im JSTOR-Datensatz zur *Mathematik* gezählt werden.

bis zum Ende des Beobachtungszeitraums um viereinhalb Jahre auf 37,3 im Jahr 2014. Auch wenn man das etwas niedrigere Durchschnittsalter in den Jahren 2010–2012 zum Vergleich heranzieht, bleibt ein Anstieg von mehr als 1,6 Jahren bis zum Jahr 2012.

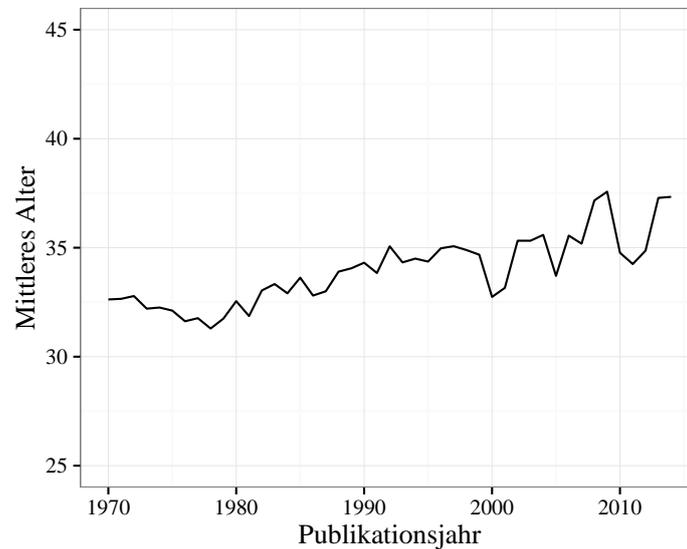


Abbildung 4.33: Mittleres Alter bei erstem Top-Artikel in der Volkswirtschaftslehre.

Die grafischen Zeittrends hinsichtlich des Alters beim ersten Artikel spiegeln sich auch in den Regressionen wider (Tabelle 4.13). In den Modellen (1), (2), (4) und (5) werden Autoren betrachtet, die bei ihrem ersten Artikel 25–45 Jahre alt sind, während das Alter bei Erstpublikation der Autoren in den Modellen (3) und (6) zwischen 25 und 35 Jahren liegt. Unter diversen Kontrollen ist der zeitliche Effekt bzgl. des Publikationsjahres positiv und hoch signifikant, wobei er in den Modellen (1)–(3) für Top-Artikel stärker ausfällt als in den Modellen (4)–(6) für alle Artikel. Beispielsweise ist der geschätzte Koeffizient für den Zeiteffekt bei Top-Artikeln im Modell (2) 0,101, während er im entsprechenden Modell (5) für alle Artikel bei 0,055 liegt. Dieser Unterschied ist vermutlich Ausdruck des höheren Wettbewerbs um und der höheren Qualitätsanforderungen für eine Publikation in einer der Top-Zeitschriften. Junge Wissenschaftler gehen bei ihrer ersten Publikation daher ggf. strategisch vor und publizieren zunächst in weniger renommierten Zeitschriften anstatt zunehmend lange auf eine Publikation in einer der Top-Zeitschriften zu warten. Auch wenn man die Regressionen getrennt für einzelne Zeitschriften oder Geschlechter durchführt, bleibt die Effektrichtung erhalten. Die einzige Ausnahme für die Top-Zeitschriften ist das *Journal of Political Economy*, für das der geschätzte Koeffizient bzgl. des Publikationsjahres negativ ist. Dieses Ergebnis ist allerdings statistisch nicht signifikant. Sofern es für eine Zeitschrift genug

Beobachtungen gibt, sind die Trends auch statistisch signifikant. Dies spricht für die Robustheit der beobachteten Zeittrends.

Tabelle 4.13: Lineare Regressionen zum Alter eines Autors bei seinem ersten Artikel in einer volkswirtschaftlichen Zeitschrift.

Abhängige Variable: Alter bei erstem Artikel						
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,098*** (0,010)	0,101*** (0,012)	0,049*** (0,008)	0,084*** (0,008)	0,055*** (0,013)	0,031*** (0,009)
Teamgröße		0,240 (0,168)	0,060 (0,112)		0,177 (0,160)	−0,107 (0,109)
Autorin		0,454 (0,357)	0,631** (0,240)		0,425 (0,333)	0,417+ (0,222)
Top-Inst		−1,568*** (0,261)	−0,967*** (0,173)		−1,313*** (0,251)	−0,631*** (0,162)
Nicht-Aka		1,108* (0,501)	0,879* (0,367)		1,590*** (0,413)	0,707* (0,294)
Migration		−0,066 (0,513)	0,257 (0,336)		−0,546 (0,519)	−0,510 (0,346)
Nicht-US		−0,018 (0,272)	0,165 (0,181)		0,201 (0,262)	0,138 (0,173)
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Altersspanne	25–45	25–45	25–35	25–45	25–45	25–35
Beobachtungen	1.635	1.599	1.115	2.656	2.597	1.942
R ²	0,055	0,124	0,101	0,037	0,359	0,308

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Für alle volkswirtschaftlichen Artikel hat die Teamgröße keinen signifikanten Effekt auf das Alter beim ersten Artikel. Der Effekt ist in den meisten Modellen zwar positiv, allerdings ist der Effekt nicht signifikant. Für den positiven, wenn auch nicht signifikanten, Effekt der Teamgröße sind zwei Erklärungen denkbar: Einerseits kann es sein, dass Autoren, die ohnehin produktiver sind und ihren ersten Artikel bereits in jüngeren

Jahren veröffentlichen, tendenziell auch in kleineren Teams arbeiten, da sie weniger Zuarbeit von anderen Wissenschaftlern benötigen bzw. sich die Arbeit mit weniger Koautoren teilen müssen. Andererseits könnte das größere Team auch ein Indikator für den Umfang und die Komplexität bzw. Schwierigkeit des zugrundeliegenden Problems sein, für das es tendenziell mehr Zeit braucht, um es zu lösen. Die zugehörigen Ergebnisse werden entsprechend zu einem späteren Zeitpunkt in einem Artikel veröffentlicht, zu welchem die beteiligten Autoren im Mittel älter sind.

Autorinnen sind bei ihrem ersten Artikel älter als ihre männlichen Kollegen. Dieser Effekt ist allerdings nur für die Altersspanne 25–35, das heißt in den Modellen (3) und (6), signifikant. Eine mögliche Erklärung hierfür könnten Schwangerschaften sein, die ggf. eine zeitweilige Unterbrechung oder zumindest Verzögerung der Arbeits- bzw. Forschungstätigkeit zur Folge haben. Hierfür spricht auch, dass der Effekt für die Altersspanne 25–35, in der sich immer noch die meisten Schwangerschaften abspielen (z. B. Jolly u. a. 2000; Ventura u. a. 2012) und die aufgrund diverser gesundheitlicher Risiken aus medizinischer Sicht für eine Schwangerschaft empfohlen wird (z. B. Fretts u. a. 1995; Jolly u. a. 2000; Callaghan und Berg 2003), signifikant ist, während dies für die Altersspanne 25–45 nicht zutrifft.³⁶

Die Zugehörigkeit zu einer Top-Institution wirkt sich sowohl für Top-Autoren als auch alle übrigen positiv auf das Alter bei Erstpublikation aus, in dem Sinne, dass solche Autoren im Mittel jünger sind als ihre Kollegen von weniger renommierten Institutionen. Autoren die von einer nicht-akademischen Institution stammen, sind bei ihrem ersten Artikel älter als ihre Kollegen an Universitäten und anderen Forschungseinrichtungen. Dieser Effekt ist für Top-Artikel auf dem 5 %-Niveau signifikant und für alle Artikel hoch signifikant. Autoren mit Migrationshintergrund sind bei ihrer Erstpublikation, bis auf das Modell (3) bei Top-Artikeln (d. h. für die Altersspanne 25–35) jünger als ihre Kollegen. Allerdings ist dieser Effekt in keinem der betrachteten Modelle statistisch signifikant. Schließlich zeigen die Regressionstabellen, keine eindeutigen Effekte hinsichtlich des Besitzes einer US-Staatsbürgerschaft. Für alle Artikel ist der geschätzte Koeffizient für nicht US-Bürger positiv, allerdings nicht signifikant.

Das mittlere Alter beim ersten Artikel von Mathematikern, die zu diesem Zeitpunkt zwischen 25 und 45 Jahre alt sind, ist in Abbildung 4.34 dargestellt.³⁷ Trotz der Schwan-

³⁶Tatsächlich hat sich die Verteilung der Schwangerschaften in der jüngeren Vergangenheit verschoben, sodass heute mehr Frauen im höheren Alter ab 35 schwanger werden als noch vor einigen Jahren (z. B. Ventura u. a. 2012). Allerdings ist der Großteil der Frauen bei ihrer Geburt immer noch jünger als 35 und etliche Studien warnen vor den Risiken einer Schwangerschaft ab dem 35. Lebensjahr (z. B. Fretts u. a. 1995; Jolly u. a. 2000; Callaghan und Berg 2003).

³⁷Die entsprechende Abbildung B.17 für die Altersspanne 25–35 findet sich im Anhang (S. 220).

kungen zwischen den Publikationsjahren, ist insgesamt ein steigender, in den späteren Jahren abflachender Trend zu beobachten. Man beachte, dass die Altersdaten für Mathematiker weiter zurückreichen als für Volkswirte. Der Beobachtungszeitraum beginnt hier bereits 1950. Allerdings ist der Alterstrend auch steigend, wenn man Erstpublikationen ab 1960 bzw. 1970 betrachtet. Im Mittel sind Mathematiker bei ihrem ersten Artikel etwas jünger als Volkswirte. Für das Jahr 1950 ergibt sich ein Altersschnitt von 28,3. In den 1960er Jahren und Anfang der 1970er Jahre liegt das mittlere Alter Anfang dreißig und steigt auf 35,4 bzw. 33,4 in den Jahren 2012 bzw. 2013.

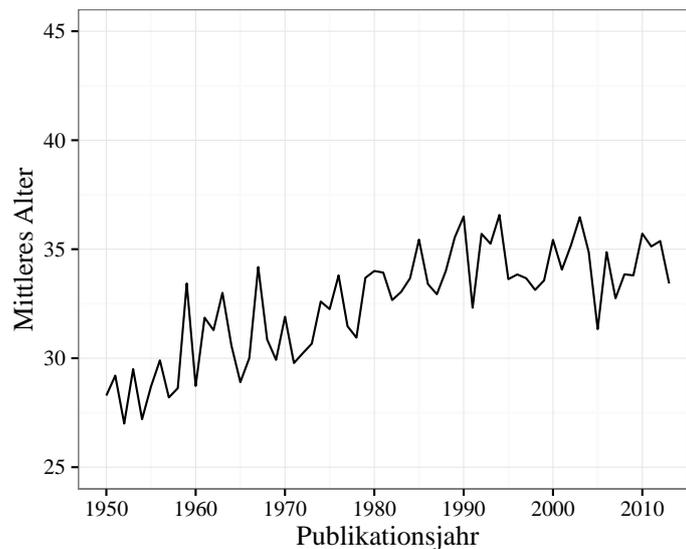


Abbildung 4.34: Mittleres Alter bei erstem Artikel in der Mathematik.

In Tabelle 4.14 finden sich die Regressionen zum in Abbildung 4.34 dargestellten Trend. In den Modellen (1) bis (4) wird der gesamte Zeitraum ab 1950 untersucht. Die Modelle (5) und (6) betrachten Autoren, die ihren ersten Artikel frühestens 1960 publizierten. Für Autoren in der Mathematik sind weniger Informationen bekannt als für Autoren in der Volkswirtschaftslehre (vgl. Abschnitt 3.5, S. 71). Daher finden sich in den Regressionsmodellen weniger Kontrollvariablen. Wie man sieht, ist der Zeittrend in allen Modellen positiv und hoch signifikant. Autorinnen sind bei ihrem ersten Artikel jünger als ihre männlichen Kollegen, allerdings ist dieser Effekt nur auf dem %-Niveau signifikant. Die Teamgröße wirkt sich in dem Sinne negativ auf das Alter beim ersten Artikel aus, dass der Autor zu diesem Zeitpunkt älter ist, je größer das Team ist, in dem er den Artikel verfasst hat. Dieses Ergebnis deckt sich mit dem für die Volkswirtschaftslehre. Mögliche Erklärungsansätze für diesen Effekt wurden bereits für die Volkswirtschaftslehre diskutiert (S. 132). Im Gegensatz zu den Ergebnissen für die Autoren volkswirtschaftlicher Artikel ist der geschätzte Effekt bzgl.

der Teamgröße hoch signifikant. Zwischen den einzelnen Zeitschriften bestehen zwar Unterschiede, allerdings sind diese statistisch nicht signifikant. Den Kontrollen für die Zeitschrift in Modell (4) zufolge sind Autoren bei ihren Erstpublikationen im *American Journal of Mathematics* und dem *SIAM Review* am jüngsten und in den Zeitschriften *Mathematics of Operations Research* und *SIAM Journal on Applied Mathematics* am ältesten. Der positive Zeittrend für die Altersspanne 25–45 aus Tabelle 4.14 wird auch für die Altersspanne 25–35 bestätigt.^{38,39}

Tabelle 4.14: Lineare Regressionen für das Alter beim ersten Artikel eines Autors in der Mathematik.

	Abhängige Variable: Alter bei erstem Artikel					
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,106*** (0,010)	0,108*** (0,010)	0,090*** (0,011)	0,095*** (0,011)	0,077*** (0,013)	0,082*** (0,014)
Autorin		-1,047 (0,658)	-1,158+ (0,647)	-1,090+ (0,650)	-1,146+ (0,658)	-1,061 (0,662)
Teamgröße			1,248*** (0,218)	1,206*** (0,225)	1,209*** (0,225)	1,184*** (0,233)
Zeitschrift (Dummys)				Ja		Ja
Beginn	1950	1950	1950	1950	1960	1960
Beobachtungen	840	840	840	840	768	768
R ²	0,110	0,113	0,146	0,157	0,101	0,112

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

4.5 Spezialisierung

In diesem Abschnitt wird untersucht, ob und wie sich die Spezialisierung von Autoren in der Volkswirtschaftslehre im Laufe der Zeit verändert hat. Wie in Abschnitt 3.6 (S. 76) beschrieben, wird der Grad der Spezialisierung anhand der relativen Häufigkeit bemessen, dass Autoren zwischen ihren ersten beiden Artikeln ihre JEL-Felder wechseln. Die Länge des Zeitraums zwischen diesen beiden Artikeln kann die Wahrscheinlichkeit

³⁸Eine entsprechende Regressionstabelle findet sich in Tabelle B.3 (S. 220) im Anhang.

³⁹Der Zeittrend bleibt auch für den Zeitraum ab 1970 in allen betrachteten Modellen steigend. Für die Altersspanne 25–45 sind auch diese Ergebnisse signifikant.

für einen solchen Feldsprung beeinflussen. Daher müssen Verzerrungen vermieden werden, die der Konstruktion dieses Maßes entspringen. Innerhalb des ersten Jahres ist der größtmögliche Zeitraum zwischen zwei Artikeln null, da stets vom Publikationsjahr ausgegangen und nicht unterjährig gerechnet wird. Im letzten Beobachtungsjahr können zwei Artikel eines Autors bis zu 45 Jahre auseinander liegen. Um mögliche Verzerrungen durch teilweise sehr unterschiedlich lange Zeiträume zwischen zwei Artikeln zu vermeiden, werden lediglich Zeiträume von maximal 3 Jahren zwischen zwei Artikeln betrachtet. Im ersten Jahr des Beobachtungszeitraums von 1970–2014 können maximal Zeiträume der Länge null beobachtet werden. Im zweiten Jahr, also 1971, sind die Zeiträume zwischen zwei Artikeln maximal ein Jahr lang. Entsprechend werden in den ersten drei Jahren Zeiträume von maximal zwei Jahren zwischen zwei aufeinander folgenden Artikeln beobachtet. Um hierdurch entstehende Verzerrungen zu vermeiden, wird der Beobachtungszeitraum für dieses Maß um die ersten drei Jahre gekürzt und reicht somit von 1973 bis 2014. Für den zeitlichen Abstand zwischen den beiden relevanten Artikeln wird in den Regressionen kontrolliert (Kontrollvariable „Zeitraum“). Die folgende Analyse der Feldsprünge ist ferner auf Artikel mit nur einem Autor beschränkt; die grundsätzlichen Trends sind allerdings ähnlich, wenn man auch Artikel mit mehr als einem Autor betrachtet.

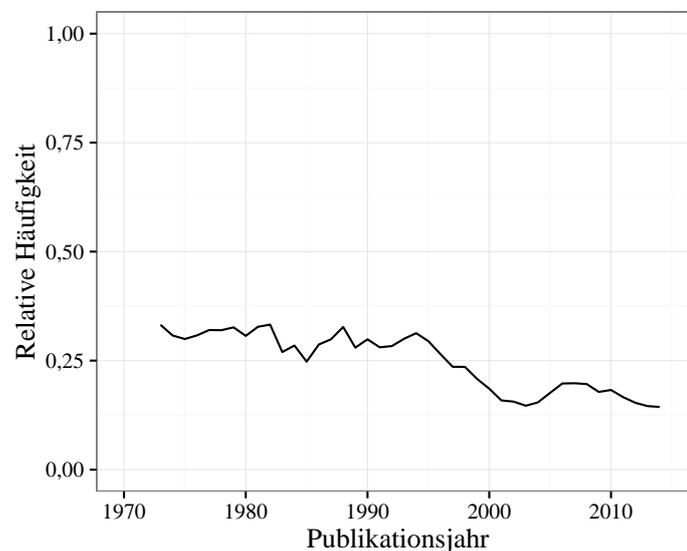


Abbildung 4.35: Relative Häufigkeit eines Feldsprungs für Autoren im gesamten EconLit-Datensatz.

Abbildung 4.35 weist einen deutlichen Abwärtstrend in der relativen Häufigkeit von JEL-Feld-Wechseln zwischen den ersten beiden Solo-Artikeln aus. Insgesamt geht die relative Häufigkeit für einen JEL-Feld-Wechsel von 33,2 % im Jahr 1973 auf 14,4 % in

2014 zurück. Die Regressionen in Tabelle 4.15 bestätigen diesen Trend unter diversen Kontrollen. Der grafische Trend für Top-Artikel fällt sogar noch etwas stärker aus (vgl. Abbildung 4.36). Die relative Häufigkeit fällt von 28,3 % im Jahr 1973 auf 0 % im Jahr 2014. Allerdings sind die Werte in den letzten Jahren des Beobachtungszeitraums mit Vorsicht zu genießen, da hier aufgrund der kleineren Stichprobe und der beschriebenen Einschränkungen die Anzahl der Beobachtungen gering ist. Aus den Regressionen geht hervor, dass die relative Häufigkeit das JEL-Feld zu wechseln für Autorinnen geringer ist, während sie für Autoren einer Top-Institution etwas höher ist.

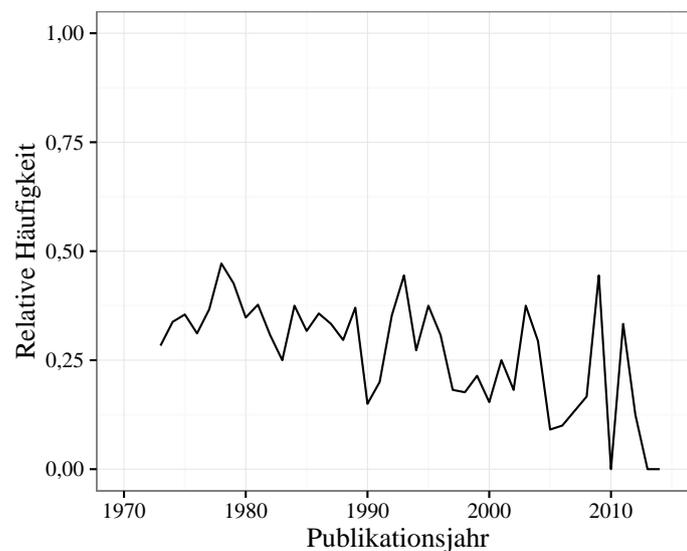


Abbildung 4.36: Relative Häufigkeit eines Feldsprungs für Top-Autoren im EconLit-Datensatz.

Da die beobachtete Variable, also ob ein Autor seine JEL-Felder wechselt oder nicht, in diesem Fall binär ist, wurde hier eine Probit-Regression gerechnet. In Tabelle 4.15 sind die Regressionsergebnisse hinsichtlich der Spezialisierung von Autoren zusammengefasst. Es ist deutlich zu sehen, dass der Zeittrend bzgl. der relativen Häufigkeit von Feldsprüngen zwischen den relevanten Artikeln in allen Modellen negativ und statistisch signifikant ist.⁴⁰ Für Top-Artikel sind die Effekte lediglich für die Kontrollvariable des Publikationsjahrs signifikant. Ein Grund hierfür könnte der geringere Stichprobenumfang sein, da sich die meisten Effektrichtungen mit denen für alle Artikel decken und die Ergebnisse für letztere bis auf die Kontrolle für die Zugehörigkeit zu einer nicht-akademischen Institution („Nicht-Aka“) signifikant sind. Der Effekt des Zeitraums zwischen zwei Publikationen ist sowohl für Top-Artikel als auch für die

⁴⁰Auch für Regressionen ohne die oben beschriebene Kürzung des Beobachtungszeitraums bleiben Effektrichtungen und Signifikanzniveaus der Ergebnisse aus Tabelle 4.15 bestehen.

Gesamtheit der Artikel positiv und für letztere auch signifikant. Dies bedeutet, dass die relative Häufigkeit eines Feldsprungs steigt, je größer der zeitliche Abstand zwischen den beiden relevanten, aufeinander folgenden Artikeln ist. Die relative Häufigkeit von Feldsprüngen ist den Regressionsergebnissen zufolge für Autorinnen geringer und für Angehörige einer Top-Institution höher. Die Ergebnisse für die Zugehörigkeit zu einer nicht-akademischen Institution sind für beide Gruppen von Artikeln nicht signifikant und insofern nicht robust, dass die Effektrichtungen für Top-Artikel und alle Artikel unterschiedlich sind.⁴¹

Tabelle 4.15: Probit-Regressionen für die relative Häufigkeit von Feldsprüngen im EconLit-Datensatz.

	Abhängige Variable: Relative Häufigkeit eines Feldsprungs					
	(1)–(3): Top-Artikel			(4)–(6): Alle Artikel		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	–0,020*** (0,003)	–0,012** (0,004)	–0,013** (0,004)	–0,021*** (0,001)	–0,011*** (0,001)	–0,011*** (0,001)
Zeitraum		0,003 (0,035)	0,001 (0,035)		0,023** (0,007)	0,023** (0,007)
Autorin			–0,195 (0,137)			–0,070** (0,021)
Top-Inst			0,012 (0,102)			0,071** (0,028)
Nicht Aka			0,393+ (0,213)			–0,042 (0,030)
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Beobachtungen	1.374	1.374	1.374	38.674	38.674	38.674

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Wie in Abschnitt 3.1.2 (S. 43) beschrieben, wurde die JEL-Klassifizierung im Jahr 1991 reformiert. In Abbildung 4.35 für den gesamten EconLit-Datensatz sind allerdings keine systematischen Veränderungen der relativen Häufigkeit von Feldsprüngen im Jahr 1991 bzw. in den Jahren danach zu beobachten. Für Top-Autoren bzw. Top-Artikel

⁴¹Streng genommen ist der Effekt der Zugehörigkeit zu einer nicht-akademischen Institution für Top-Artikel auf dem 10 %-Niveau signifikant.

ist in den frühen 1990er Jahren zunächst sogar ein Anstieg der relativen Häufigkeit von Feldsprüngen zu erkennen (vgl. Abbildung 4.36). Die relativen Häufigkeiten von JEL-Feld-Wechseln für alle Artikel in den Jahren vor, während und nach der Reform des JEL-Klassifizierungssystems, liegen bei 26,0 % für Wechsel innerhalb des alten Systems von 1989 zu 1990, bei 32,1 % für Wechsel von 1990 zu 1991 vom alten in das neues System und 30,8 % für Wechsel innerhalb des neuen Systems von 1991 zu 1992. Der unmittelbare Effekt der Änderung der JEL-Klassifizierung scheint somit sogar entgegen den beobachteten Ergebnissen zu wirken, sodass der Rückgang der relativen Häufigkeit von Feldsprüngen nicht per Konstruktion durch die Änderung des Klassifizierungssystems im Jahr 1991 bewirkt wurde. Insgesamt kann also davon ausgegangen werden, dass die von der JEL-Klassifikation abgedeckten ökonomischen Felder, durch die Zuordnung von Card und DellaVigna (2013), korrekt abgebildet werden. Ihre Zuordnung diente auch in dieser Arbeit der Vereinheitlichung der JEL-Felder.

4.6 Zusammenhang zwischen den betrachteten Maßen und dem Wissensinput

Eine Hypothese dieser Arbeit ist, dass die Menge an Wissen mit der Zeit zunimmt und somit die Last des Wissens verstärkt. Daher wurden für die einzelnen Maße bislang Zeittrends untersucht. Allerdings könnten sich verschiedene exogene Faktoren, im Laufe der Zeit entwickelt haben und die Ergebnisse beeinflussen. Vor dem Hintergrund von Analysen technologischer Lebenszyklen argumentiert Brockhoff (1993), dass es angebrachter wäre den direkten Input in technologische Entwicklungen zu betrachten als das Proxy der Zeit. In der vorliegenden Arbeit entspricht der Input am ehesten der Anzahl der Quellenangaben in Form von Wissen. Auch Jones (2009a) untersuchte, wie der Input in ein Patent beispielsweise die Anzahl der Patentanmelder beeinflusst. Nach dem Beispiel von Jones (2009a) wird in dieser Arbeit um Zeittrends korrigiert, indem für jedes Jahr die mittlere Anzahl an Quellenangaben berechnet und die Anzahl der Quellenangaben in jedem Artikel aus diesem Jahr mit dieser Zahl verglichen wird. Das normalisierte, zeitbereinigte Maß für den Wissensinput entspricht der Anzahl Quellenangaben abzüglich der durchschnittlichen Anzahl Quellenangaben im entsprechenden Jahr geteilt durch die Standardabweichung in diesem Jahr. Betrachtet man also einen Artikel aus dem Jahr j mit R Quellenangaben und entspricht die mittlere Anzahl Quellenangaben in diesem Jahr \bar{R}_j sowie die zugehörige Standardabweichung σ_j , so wird

die normalisierte Differenz in der Anzahl Quellenangaben R_{norm} wie folgt berechnet:

$$R_{norm} = \frac{R - \bar{R}_j}{\sigma_j}$$

Dieses normalisierte Maß dient in den folgenden Regressionen als zeitbereinigtes Proxy für den Wissensinput. Im Folgenden wird hierauf auch mit *normalisierter Wissensinput* bzw. in den Regressionstabellen kurz mit *Input* Bezug genommen. Um verbleibende Zeiteffekte abzufangen, wird mittels Dummy-Variablen für die einzelnen Publikationsjahre kontrolliert. An dieser Stelle sei nochmal erwähnt, dass die Anzahl der Quellenangaben gleichzeitig die Tiefe und Breite vorhergehenden Wissens abbildet. Allerdings lässt sich zwischen diesen beiden Ausprägungen bei den gegebenen Daten nicht unterscheiden. Ein Indikator für die Tiefe könnte die Tiefe des vorhergehenden Zitationsbaums sein, die in der Patentanalyse von Jones (2009a) betrachtet wurde. Diese ist allerdings, wie erwähnt, für die vorliegenden Publikationsdaten nicht verfügbar.

Konsistent mit den Hypothesen über die Last des Wissens ist zu beobachten, dass Publikationen mit einer größeren Menge vorhergehenden Wissens (d. h. einem größeren normalisierten Wissensinput) auch von einer größeren Menge von Autoren geschrieben wurden. Der normalisierte Wissensinput hat einen signifikant positiven Effekt auf die Teamgröße. Dieses Ergebnis gilt sowohl für alle Modelle zum EconLit-Datensatz (Tabelle 4.16) als auch für die Disziplinen im JSTOR-Datensatz und Artikel der Zeitschrift *Science* (vgl. Tabelle 4.17). Dabei ist der Effekt für Artikel der Zeitschrift *Science* mit einem Koeffizienten von 0,318 am größten und für Artikel der *Betriebswirtschaftslehre und Finance* sowie volkswirtschaftliche Top-Artikel mit Koeffizienten von 0,013 am geringsten (vgl. Modell „BWL“ in Tabelle 4.17 und Modell (3) in Tabelle 4.16). Man beachte, dass auch eine umgekehrte Kausalität besteht, dass also die Teamgröße einen Einfluss auf die Anzahl der Quellenangaben hat. Ein solcher Zusammenhang wurde bereits in Abschnitt 4.2 (S. 102) bei der Untersuchung der Anzahl Quellenangaben berücksichtigt und ein positiver Zusammenhang gezeigt. Dort wurde einerseits die Abhängigkeit der mittleren Anzahl Quellenangaben von der Teamgröße grafisch veranschaulicht (vgl. z. B. Abbildung 4.17 oder Abbildung 4.19, S. 107 ff.) und andererseits in den Regressionen für die Teamgröße kontrolliert.

Ein größerer Wissensinput drückt sich auch in einem längeren Titel aus, wie man den Tabellen 4.18 und 4.19 entnehmen kann. Für alle Modelle zum EconLit-Datensatz (Tabelle 4.18) ist dieser Effekt positiv und hoch signifikant. Die Teamgröße wirkt sich auch in diesen Modellen signifikant positiv auf die Länge des Titels aus. Diese

Effektrichtung der Teamgröße auf die Titellänge war bereits in Tabelle 4.12 (S. 128) zu sehen. Allerdings stand dort die Untersuchung des zeitlichen Trends im Vordergrund und es wurde auch nicht für den Wissensinput kontrolliert. Auch für die Disziplinen im JSTOR-Datensatz und der Zeitschrift *Science* zeigt sich dieser Zusammenhang (Tabelle 4.19), wobei das Ergebnis für die *Mathematik* nicht signifikant und das für *Geographie und Geologie* lediglich auf dem 5 %-Niveau signifikant ist. Insgesamt sprechen die Ergebnisse dafür, dass die Länge des Titels tatsächlich etwas über den Umfang, die Menge des verarbeiteten Wissens und die Komplexität eines Artikels aussagt. Zu diesem Ergebnis kommen, wie bereits erwähnt, beispielsweise auch White und Hernandez (1991), allerdings unter Verwendung einer anderen Methodik. Zusammen mit dem Ergebnis aus Abschnitt 4.3 (S. 122), dass die Länge der Titel zugenommen hat, lässt sich auch die Titellänge als Maß für die zunehmende Last des Wissens auffassen.

Tabelle 4.16: Lineare Regressionen für den Zusammenhang zwischen der Teamgröße und dem normalisierten Wissensinput im EconLit-Datensatz.

	Abhängige Variable: Teamgröße					
	(1)–(3): Top-Artikel			(4)–(6): Alle-Artikel		
	(1)	(2)	(3)	(4)	(5)	(6)
Wissensinput	0,024*** (0,005)	0,015** (0,005)	0,013* (0,006)	0,024*** (0,002)	0,029*** (0,002)	0,029*** (0,002)
Autorinnen			–0,638*** (0,024)			–0,640*** (0,007)
Alle Top-Inst			–0,350*** (0,012)			–0,356*** (0,006)
Alle Nicht-Aka			–0,348*** (0,029)			–0,295*** (0,009)
Publikationsjahr (Dummys)	Ja	Ja	Ja	Ja	Ja	Ja
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Beobachtungen	24.904	24.904	16.497	209.603	209.603	185.704
R^2	0,133	0,149	0,215	0,073	0,148	0,192

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Tabelle 4.17: Lineare Regressionen für den Zusammenhang zwischen der Teamgröße und dem normalisierten Wissensinput für den JSTOR-Datensatz und die Zeitschrift *Science*.

	Abhängige Variable: Teamgröße				
	BWL	Geo	Mathe	Psycho	<i>Science</i>
Wissensinput	0,013** (0,005)	0,035 ⁺ (0,018)	0,041*** (0,004)	0,055*** (0,012)	0,318*** (0,017)
Autorinnen	-0,655*** (0,020)	-0,380*** (0,058)	-0,500*** (0,019)	-0,812*** (0,032)	-2,603*** (0,110)
Publikationsjahr (Dummys)	Ja	Ja	Ja	Ja	Ja
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	
Beobachtungen	34.828	2.264	37.954	9.654	57.263
R^2	0,169	0,257	0,175	0,175	0,087

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Tabelle 4.18: Lineare Regressionen für den Zusammenhang zwischen der Titellänge und dem normalisierten Wissensinput für den EconLit-Datensatz.

	Abhängige Variable: Titellänge					
	(1)–(3): Top-Artikel			(4)–(6): Alle-Artikel		
	(1)	(2)	(3)	(4)	(5)	(6)
Wissensinput	0,099*** (0,022)	0,096*** (0,022)	0,050* (0,022)	0,314*** (0,008)	0,309*** (0,008)	0,152*** (0,009)
Teamgröße		0,135*** (0,030)	0,097** (0,030)		0,210*** (0,010)	0,173*** (0,010)
Publikationsjahr (Dummys)	Ja	Ja	Ja	Ja	Ja	Ja
Zeitschrift (Dummys)			Ja			Ja
JEL-Feld (Dummys)			Ja			Ja
Beobachtungen	24.590	24.590	24.590	208.481	208.481	208.481
R^2	0,003	0,004	0,044	0,025	0,027	0,113

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Tabelle 4.19: Lineare Regressionen für den Zusammenhang zwischen der Titellänge und dem normalisierten Wissensinput für den JSTOR-Datensatz und die Zeitschrift *Science*.

	Abhängige Variable: Titellänge				
	BWL	Geo	Mathe	Psycho	<i>Science</i>
Wissensinput	0,065*** (0,019)	0,211* (0,087)	0,018 (0,018)	0,227*** (0,037)	0,239*** (0,012)
Teamgröße	0,203*** (0,022)	0,533*** (0,101)	0,025 (0,023)	0,260*** (0,029)	0,118*** (0,003)
Publikationsjahr (Dummys)	Ja	Ja	Ja	Ja	Ja
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	
Beobachtungen	34.589	2.233	37.634	9.635	55.655
R^2	0,074	0,281	0,057	0,085	0,069

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

Für die JEL-Feldsprünge und das Alter bei Erstpublikation zeigen sich bei dieser Betrachtung allerdings keine signifikanten Effekte. Diese Nullresultate könnten auf Unterschiede zwischen dem hier verwendeten Maß und dem Maß von Jones (2009a) hindeuten. Jones maß lediglich die Tiefe von Wissen, indem er die Tiefe des Zitationsbaums heranzog. Das hier verwendete Maß beinhaltet sowohl die Tiefe, als auch die Breite von Wissen. Sind in einem Artikel also mehr Quellen angegeben als in einem anderen, so kann dies einerseits daran liegen, dass erster umfangreicher oder komplexer ist oder dass es sich um ein spezielleres Thema handelt und entsprechend mehr Wissen zu dessen Vertiefung notwendig ist. Autoren mit einem breiteren Hintergrund, die vermeintlich auch mehr Quellen in ihren Artikeln angeben, könnte es ggf. sogar leichter fallen ihr JEL-Feld zu wechseln. Somit kann der hier betrachtete normierte Wissensinput entsprechende Feldsprünge nicht adäquat erklären.

4.7 Vergleich der Ergebnisse für JEL-Felder und Disziplinen

4.7.1 Vergleich von JEL-Feldern im EconLit-Datensatz

In diesem Abschnitt wird auf die Unterschiede zwischen den 13 JEL-Feldern (inklusive „Andere“) bzgl. der untersuchten Maße eingegangen.⁴² Tabelle 4.20 gibt einen Überblick über das durchschnittliche Alter bei Erstpublikation, die Teamgröße beim ersten Artikel eines Autors, die relative Häufigkeit eines Feldsprungs und die Anzahl angegebener Quellen beim ersten Artikel eines Autors, für die einzelnen JEL-Felder.⁴³ Die entsprechenden Spalten zu den Maßen sind jeweils zweigeteilt, mit der Anzahl der Beobachtungen als erster und dem Mittelwert als zweiter Information.

Die zweite Spalte in Tabelle 4.20 weist Informationen zum Alter der Autoren bei ihrem ersten Artikel aus. Die Werte liegen für die unterschiedlichen Felder erstaunlich nahe beieinander. Der Altersschnitt beim ersten Top-Artikel ist mit einem Wert von 32,5 Jahren im JEL-Feld „Geschichte“ am geringsten. Gleichzeitig weist das Feld „Geschichte“ eine geringe Beobachtungszahl und die geringste mittlere Teamgröße für alle Artikel auf sowie eine der geringsten für Top-Artikel. Dies könnte ein möglicher Hinweis darauf sein, dass dieses Feld weniger kompetitiv ist und junge Autoren darin entsprechend früher zum ersten Mal publizieren. Beim mittleren Alter weisen die Felder „Arbeitsmarktökonomik“ und „Theorie“ die geringsten Mittelwerte auf. Zu den Feldern, mit den höchsten Altersschnitten gehören „Entwicklung“, „Finance“, „Gesundheits- und Stadtökonomik“ (GSÖ) sowie „Andere“. Die mittlere Teamgröße beim ersten Artikel ist für alle Felder ähnlich. Sie reicht von 1,9 bis 2,3 Autoren für Top-Artikel und von 2,0 bis 2,4 für alle Zeitschriften (dritte Spalte in Tabelle 4.20). Das Feld „Geschichte“ stellt dabei eine Ausnahme dar, da die mittlere Teamgröße für alle Artikel mit Abstand am geringsten ist (\bar{x} 1,5 Autoren).

Hinsichtlich der relativen Häufigkeit eines Feldsprungs lassen sich für Top-Artikel zwei Ausreißer ausmachen (vierte Spalte in Tabelle 4.20). In den Feldern „Geschichte“ und „Theorie“ ist die relative Häufigkeit mit 0,57 bzw. 0,49 am höchsten. Allerdings bestimmt sich dieser Wert für das Feld „Geschichte“ aus lediglich 37 Beobachtungen. Für die übrigen Artikel reicht die relative Häufigkeit von 0,31 bis 0,44. Betrachtet man

⁴²Aufgrund der geringen Beobachtungszahlen ist das JEL-Feld „Experimentalökonomik“ unter „Andere“ inbegriffen.

⁴³Die Teamgröße und die Anzahl der Quellenangaben beziehen sich auf die Artikel-Ebene.

Feldsprünge bei allen Artikeln im EconLit-Datensatz, so liegen die relativen Wechselhäufigkeiten näher beieinander. Für das JEL-Feld „Theorie“ ist die relative Häufigkeit mit 0,39 am höchsten, wobei sie für die übrigen Felder von 0,26 bis 0,34 reicht.

Für die meisten Felder reicht die mittlere Anzahl angegebener Quellen beim ersten Top-Artikel von 23,7 bis 28,1 (fünfte Spalte in Tabelle 4.20). Dieser Wert ist für die JEL-Felder „Entwicklungsökonomik“ und „Finance“ mit 31,9 bzw. 31,0 etwas höher. Am höchsten ist die mittlere Anzahl angegebener Quellen beim ersten Artikel allerdings für Top-Artikel aus dem Feld „Geschichte“ (\bar{x} 36,3 Quellenangaben). Auch für alle Artikel ist der Durchschnitt in diesem JEL-Feld am größten (46,4). Insgesamt liegen die Durchschnitte für alle Artikel weiter auseinander. Die wenigsten Quellen beim ersten Artikel werden in den Feldern „Theorie“ (25,9) und „Ökonometrie“ (26,3) angegeben und die meisten, mit Ausnahme des Felds „Geschichte“, in den Feldern „Entwicklungsökonomik“ (35,8) sowie „Gesundheits- und Stadtökonomik“ (35,6). Wie bereits grafisch gesehen, sind die mittleren Teamgrößen (vgl. Abbildung 4.1, S. 84) und die mittlere Anzahl der Quellenangaben (vgl. Abbildung 4.14, S. 103) für Top-Artikel geringer als für alle Artikel.

Die Länge des Titels ist in Tabelle 4.20 aus Platzgründen nicht abgebildet. Auch wenn die Werte für die einzelnen JEL-Felder nahe beieinander liegen, sind Unterschiede zwischen den Feldern erkennbar. Wie bereits in Abbildung 4.27 zu sehen war, sind die Titel von Top-Artikeln kürzer als die für alle Artikel. Die kürzesten Titel weist dabei jeweils das JEL-Feld „Theorie“ auf (\bar{x} 7,8 Wörter für Top-Artikel und \bar{x} 9,3 Wörter für alle Artikel) und die längsten das JEL-Feld „Geschichte“ (\bar{x} 9,9 bzw. \bar{x} 13,1 Wörter). Für alle Artikel weist mit durchschnittlich 12,6 Wörtern die zweitlängsten Titel das JEL-Feld „Entwicklungsökonomik“ auf. Abgesehen von diesen drei Feldern liegen die restlichen relativ nahe beieinander, wobei die mittlere Länge von Titeln für Top-Artikel 8,3–9,5 Wörter beträgt und die für alle Artikel 10,3–12,2.

Insgesamt sind sich die JEL-Felder, bis auf einzelne Ausnahmen, bzgl. der betrachteten Maße ähnlich. Dies zeigt sich einerseits in den relativ geringen Abweichungen bei den Mittelwerten für die verschiedenen Felder (vgl. Tabelle 4.20) und darin, dass die Ergebnisse der Regressionen für einzelne Felder konsistent sind mit denen aus den Untersuchungen über alle Felder hinweg (vgl. Tabellen 4.1, 4.4, 4.6, 4.9 und 4.13).

Tabelle 4.20: Vergleich der JEL-Felder hinsichtlich der Mittelwerte der betrachteten Maße.

JEL-Feld	Alter bei 1. Artikel		Team- größe		Feld- sprung		Quellen- angaben	
	N	Ø	N	Ø	N	Ø	N	Ø
13								
Top-Artikel								
Arbeitsmarktökonomik	267	33,2	2.955	2,1	309	0,31	2.698	25,6
Außenwirtschaft	185	33,1	1.708	1,9	269	0,37	1.394	24,3
Entwicklungsökonomik	122	34,1	1.508	2,2	135	0,42	1.389	31,9
Finance	187	34,3	3.154	2,2	218	0,35	2.981	31,0
Geschichte	29	32,5	234	1,9	37	0,57	212	36,3
GSÖ	78	34,4	1.163	2,3	105	0,40	1.063	28,1
Industrieökonomik	144	34,0	2.171	2,1	224	0,39	1.974	27,4
Makroökonomie	277	33,2	2.617	1,9	432	0,41	2.198	23,9
Mikroökonomie	371	33,3	4.202	2,1	589	0,44	3.715	26,8
ÖW	130	33,4	1.375	2,1	218	0,44	1.199	23,7
Ökonometrie	134	33,5	1.534	2,0	200	0,37	1.258	24,9
Theorie	134	32,8	1.163	2,0	220	0,49	1.043	24,3
Andere	174	34,9	2.477	2,1	326	0,42	2.158	26,1
Alle Artikel								
Arbeitsmarktökonomik	344	31,9	57.664	2,2	10.905	0,31	13.251	33,6
Außenwirtschaft	333	32,6	39.670	2,0	10.530	0,32	7.603	34,0
Entwicklungsökonomik	240	34,2	89.735	2,3	13.219	0,26	16.560	35,8
Finance	270	33,6	61.770	2,3	10.419	0,29	16.910	30,7
Geschichte	60	32,4	7.458	1,5	2.318	0,33	1.453	46,4
GSÖ	198	33,6	62.846	2,3	9.972	0,30	11.385	35,6
Industrieökonomik	260	32,9	81.942	2,4	12.548	0,31	21.825	33,2
Makroökonomie	511	32,7	35.627	2,0	11.791	0,34	6.752	26,9
Mikroökonomie	460	32,2	58.894	2,2	13.582	0,34	15.378	34,2
ÖW	224	32,7	28.556	2,1	8.213	0,35	5.734	29,6
Ökonometrie	169	32,0	23.547	2,4	4.225	0,34	9.451	26,3
Theorie	134	31,9	7.939	2,2	3.329	0,39	2.670	25,9
Andere	537	33,3	144.422	2,4	21.126	0,31	34.325	33,2

Bem.: (i) GSÖ = Gesundheits- und Stadtökonomik, ÖW = Öffentliche Wirtschaft.
(ii) Die Maße Teamgröße und Quellenangaben beziehen sich jeweils auf den ersten Artikel eines Autors.

4.7.2 Vergleich der untersuchten Disziplinen

Im vorangehenden Abschnitt 4.7.1 wurden die Ergebnisse für die *Volkswirtschaftslehre* für einzelne JEL-Felder verglichen und Unterschiede sowie Gemeinsamkeiten aufgezeigt. In diesem Abschnitt werden die Ergebnisse für Maße, die für alle in der vorliegenden Arbeit untersuchten Disziplinen verfügbar sind, zusammengefasst und ein Vergleich zwischen den Disziplinen gezogen.⁴⁴ Diese Maße sind die Teamgröße, die Anzahl der Quellenangaben und die Titellänge. Dabei sei angemerkt, dass der Vergleich zwischen den Disziplinen nicht ganz einfach ist, da einerseits Unterschiede in der Anzahl der erhobenen Zeitschriften und andererseits in den Beobachtungszeiträumen bestehen. Beispielsweise umfasst der gesamte EconLit-Datensatz eine Vielzahl volkswirtschaftlicher Zeitschriften und deckt diese Disziplin somit breiter ab als die vier Zeitschriften aus *Geographie und Geologie*. Die Beobachtungszeiträume umfassen 45 Jahre zwischen 1970 und 2014 für die *Volkswirtschaftslehre*, 50 Jahre (1960–2009) für die Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* sowie *Psychologie* und 52 Jahre (1960–2011) für die überwiegend naturwissenschaftlichen Artikel aus der Zeitschrift *Science*. Dennoch erscheint ein solcher Vergleich sinnvoll und interessant, um einen Eindruck von den Unterschieden zwischen den Disziplinen zu bekommen.

Um eine bessere Vergleichbarkeit zu ermöglichen und mit den unterschiedlich langen Beobachtungszeiträumen umzugehen, werden für alle Disziplinen relative Veränderungen der einzelnen Maße pro Dekade errechnet. In den folgenden, zusammenfassenden Tabellen sind die Werte der Maße zu Beginn und am Ende eines jeden Beobachtungszeitraums, die absolute und relative Differenz dieser Werte sowie die absolute und relative Veränderung pro Dekade dargestellt.

Tabelle 4.21 fasst die Ergebnisse für die mittlere Teamgröße zusammen. Die obere Hälfte der Tabelle bezieht sich auf alle Artikel in den Datensätzen, während sich die untere Hälfte auf Artikel mit mindestens einem debütierenden Autor bezieht. Zunächst werden die Ergebnisse für alle Artikel verglichen. Die Teamgrößen für *Geographie und Geologie* sind sowohl zu Beginn ($\bar{\emptyset}$ 1,1 Autoren) als auch am Ende ($\bar{\emptyset}$ 1,4 Autoren) des Beobachtungszeitraums die geringsten der untersuchten Disziplinen. Der größte absolute und relative Anstieg pro Dekade in der Teamgröße ist für die Zeitschrift *Science* zu beobachten (21,8 %). In dieser sind Teams über den Beobachtungszeitraum

⁴⁴Vereinfachend wird hier von Disziplinen gesprochen, auch wenn die Zeitschrift *Science* an sich keine Disziplin darstellt. Die Artikel aus *Science* stehen allerdings gewissermaßen stellvertretend für die Naturwissenschaften.

von 52 Jahren im Mittel um 3,3 Autoren größer geworden. Die *Psychologie* weist zwar den zweitgrößten absoluten Zuwachs in der Teamgröße auf, allerdings reiht sie sich bei der relativen Veränderung pro Dekade mit 11,7 % hinter der *Betriebswirtschaftslehre und Finance* (15 %) und den Top-Artikeln der *Volkswirtschaftslehre* (13,5 %) ein. Die Teamgröße ist für Artikel aus den renommiertesten Zeitschriften der *Volkswirtschaftslehre* stärker gestiegen als für alle Artikel dieser Disziplin. Die Reihung der einzelnen Disziplinen hinsichtlich der jeweiligen mittleren Teamgröße und der Stärke des Anstiegs bleibt auch für erste Artikel von Autoren weitestgehend bestehen. An den mittleren Teamgrößen zu Beginn und am Ende des Beobachtungszeitraums ist ferner zu erkennen, dass Teams bei ersten Artikeln größer sind als für alle Artikel der jeweiligen Disziplinen.

Tabelle 4.21: Vergleich der Disziplinen hinsichtlich der Entwicklung der mittleren Teamgrößen.

Disziplin	Mittlere Teamgröße		Veränderung im ges. Zeitraum		Veränderung je Dekade	
	Anfang	Ende	absolut	relativ	absolut	relativ
Alle Artikel						
BWL	1,2	2,3	1,2	101,4 %	0,24	15,0 %
Geo	1,1	1,4	0,2	20,6 %	0,05	3,8 %
Mathematik	1,2	2,0	0,8	67,9 %	0,16	10,9 %
Psychologie	1,7	3,0	1,3	74,2 %	0,25	11,7 %
<i>Science</i>	1,9	5,2	3,3	179,1 %	0,64	21,8 %
VWL	1,3	2,1	0,8	63,8 %	0,18	11,6 %
VWL (Top)	1,3	2,3	1,0	76,9 %	0,22	13,5 %
Erste Artikel						
BWL	1,2	2,5	1,3	107,2 %	0,26	15,7 %
Geo	1,2	1,6	0,4	28,4 %	0,07	5,1 %
Mathematik	1,3	2,1	0,8	63,8 %	0,16	10,4 %
Psychologie	1,9	3,2	1,3	71,3 %	0,27	11,4 %
<i>Science</i>	2,0	7,4	5,4	265,9 %	1,04	28,3 %
VWL	1,3	2,4	1,1	86,1 %	0,25	14,8 %
VWL (Top)	1,3	2,5	1,2	94,7 %	0,27	16,0 %

Bem.: Die Spalten „Anfang“ und „Ende“ beziehen sich auf das erste bzw. letzte Jahr des jeweiligen Beobachtungszeitraums.

Tabelle 4.22: Vergleich der Disziplinen hinsichtlich der Entwicklung der mittleren Anzahl Quellenangaben.

Disziplin	Mittlere Anzahl Quellenangaben		Veränderung im ges. Zeitraum		Veränderung je Dekade	
	Anfang	Ende	absolut	relativ	absolut	relativ
Alle Artikel						
BWL	10,9	42,5	31,6	291,2 %	6,3	31,4 %
Geo	19,8	42,8	23,0	116,0 %	4,6	16,7 %
Mathematik	6,0	23,7	17,7	294,2 %	3,5	31,6 %
Psychologie	12,1	60,1	47,9	394,8 %	9,6	37,7 %
<i>Science</i>	8,4	22,8	14,4	172,4 %	2,8	21,3 %
VWL	14,7	37,9	23,2	157,6 %	5,2	23,4 %
VWL (Top)	13,2	38,6	25,4	193,0 %	5,6	27,0 %
Erste Artikel						
BWL	10,1	43,5	33,4	331,7 %	6,7	34,0 %
Geo	14,5	42,1	27,6	190,5 %	5,5	23,8 %
Mathematik	5,6	23,8	18,2	322,4 %	3,6	33,4 %
Psychologie	12,4	63,6	51,3	414,2 %	10,3	38,7 %
<i>Science</i>	7,0	26,8	19,8	283,0 %	3,8	29,5 %
VWL	14,6	39,3	24,7	168,7 %	5,5	24,6 %
VWL (Top)	13,5	37,2	23,6	174,8 %	5,3	25,2 %
Erste Artikel (Solo)						
BWL	11,5	39,6	28,1	245,4 %	5,6	28,1 %
Geo	14,5	42,6	28,1	193,7 %	5,6	24,0 %
Mathematik	5,7	18,8	13,1	230,0 %	2,6	27,0 %
Psychologie	13,3	37,6	24,4	183,8 %	4,9	23,2 %
<i>Science</i>	8,8	13,2	4,3	49,1 %	0,8	8,0 %
VWL	14,1	40,9	26,8	190,4 %	6,0	26,7 %
VWL (Top)	12,3	39,8	27,5	223,6 %	6,1	29,8 %

Bem.: (i) Die Spalten „Anfang“ und „Ende“ beziehen sich auf das erste bzw. letzte Jahr des jeweiligen Beobachtungszeitraums. (ii) Die Werte für „Alle Artikel“ beziehen sich auf Artikel-Ebene, während sich die Werte für „Erste Artikel“ auf Autoren-Ebene beziehen.

In Tabelle 4.22 sind die Ergebnisse für die Anzahl der Quellenangaben zusammengefasst. Diese Tabelle ist dreigeteilt und weist die Veränderungen in der mittleren Anzahl Quellenangaben für alle Artikel (oberes Drittel), für erste Artikel (mittleres Drittel) und für erste Artikel, bei denen der Autor alleiniger Autor war (unteres Drittel), aus. Zunächst werden wie für die Teamgröße die Ergebnisse für alle Artikel verglichen. Ähnlich wie bei der Teamgröße, weist die *Geographie und Geologie* auch für dieses

Maß die kleinste Wachstumsrate pro Dekade auf (3,8 %). Für alle Disziplinen ist ein deutlicher Anstieg zu erkennen. Den höchsten relativen Anstieg pro Dekade weist die *Psychologie* auf (37,7 %), während er für *Geographie und Geologie* mit 16,7 % am geringsten ausfällt. Für *Betriebswirtschaftslehre und Finance* sowie *Mathematik* sind die relativen Veränderungen pro Dekade mit 31,4 % bzw. 31,6 % sehr ähnlich, auch wenn sich die jeweiligen Mittelwerte am Anfang bzw. Ende des Beobachtungszeitraums deutlich unterscheiden. Auch die relativen Veränderungen pro Dekade für die Zeitschrift *Science* und die *Volkswirtschaftslehre* liegen nah beieinander. Es wird ferner deutlich, dass die Anzahl der Quellenangaben pro Artikel für Top-Artikel der *Volkswirtschaftslehre* stärker gestiegen ist als für alle Artikel dieser Disziplin. Aus Tabelle 4.22 geht auch hervor, dass sich die Reihung der Disziplinen anhand der relativen Veränderungen ändert, wenn man von der Veränderung über den gesamten Beobachtungszeitraum zur Veränderung pro Dekade übergeht. So ist die mittlere Anzahl der Quellenangaben in der Zeitschrift *Science* im Beobachtungszeitraum um 172,4 % gestiegen und damit stärker als für die *Volkswirtschaftslehre*, für die „nur“ ein Anstieg von 157,6 % auszumachen ist. Allerdings ist der Beobachtungszeitraum für *Science* sieben Jahre länger als der für die *Volkswirtschaftslehre*, weshalb sich die Reihung der beiden umkehrt, wenn man die Veränderungen pro Dekade vergleicht (21,3 % zu 23,4 %).

Die qualitativen Ergebnisse für erste Artikel (vgl. mittleres Drittel in Tabelle 4.22) stimmen weitestgehend mit denen für alle Artikel (oberes Drittel) überein und auch die Reihung der einzelnen Disziplinen bleibt weitestgehend bestehen. Allerdings sind die beiden Bereiche der Tabelle nicht direkt miteinander vergleichbar, da sich die Ergebnisse für alle Artikel im oberen Drittel der Tabelle auf Artikel-Ebene und die im mittleren Drittel auf Autoren-Ebene beziehen. Auf diesen Unterschied sind auch die etwas größeren relativen Veränderungen pro Dekade für erste Artikel zurückzuführen. Betrachtet man nämlich auf Autoren-Ebene, so geht die Anzahl der Quellenangaben in einem Artikel mit der Anzahl der Autoren des Artikels in den jeweiligen Jahresmittelwert ein. Da die mittleren Teamgrößen in den Beobachtungszeiträumen zugenommen haben und die Anzahl der Quellenangaben positiv mit der Teamgröße korreliert, ergeben sich zum Ende des Beobachtungszeitraums entsprechend größere Mittelwerte.⁴⁵ Deutliche Anstiege der mittleren Anzahl Quellenangaben sind auch für erste Artikel zu beobachten, bei denen der Autor alleiniger Autor war (vgl. unteres Drittel in Tabelle 4.22). Eine Ausnahme stellen Artikel aus der Zeitschrift *Science* dar, die einen

⁴⁵Für die volkswirtschaftlichen Top-Artikel in Tabelle 4.22 ist die relative Veränderung pro Dekade bei ersten Artikeln nicht größer als für alle Artikel. Ein Grund hierfür ist, dass sich, wie in Abschnitt 4.2.2 (S. 115) erläutert, in ersten Top-Artikeln mit lediglich einem Autor im Mittel sogar mehr Quellenangaben finden als in Artikeln mit mehreren Autoren und die beschriebenen Effekte somit hier nicht greifen.

relativ geringen Zuwachs (8 % pro Dekade) im Vergleich zu den übrigen Disziplinen aufweisen. In Abbildung 4.26 (S. 121) war zu sehen, dass der Mittelwert für diese zwar bis 1991 steigt, danach allerdings fallend ist. Erstaunlich ist, dass erste Artikel in der *Volkswirtschaftslehre*, bei denen der Autor alleiniger Autor war, am Ende des Beobachtungszeitraums (d. h. im Jahr 2014) im Mittel mehr Quellenangaben aufweisen als alle Artikel und alle ersten Artikel.

Tabelle 4.23: Vergleich der Disziplinen hinsichtlich der Entwicklung der mittleren Anzahl der Wörter im Titel (Titellänge).

Disziplin	Mittlere Titellänge		Veränderung im ges. Zeitraum		Veränderung je Dekade	
	Anfang	Ende	absolut	relativ	absolut	relativ
BWL	7,2	9,7	2,4	33,6 %	0,49	6,0 %
Geo	7,6	9,3	1,7	22,0 %	0,34	4,1 %
Mathematik	7,7	9,0	1,3	16,2 %	0,25	3,0 %
Psychologie	9,5	12,1	2,6	26,9 %	0,51	4,9 %
<i>Science</i>	7,3	8,0	0,8	10,6 %	0,15	2,0 %
VWL	9,4	11,9	2,5	26,4 %	0,55	5,3 %
VWL (Top)	8,2	8,8	0,6	7,5 %	0,14	1,6 %

Bem.: Die Spalten „Anfang“ und „Ende“ beziehen sich auf das erste bzw. letzte Jahr des jeweiligen Beobachtungszeitraums.

Zuletzt werden noch die mittleren Längen der Titel in den jeweiligen Disziplinen verglichen (vgl. Tabelle 4.23). Artikel aus den neun renommiertesten Zeitschriften der *Volkswirtschaftslehre* weisen den geringsten relativen Zuwachs pro Dekade auf (1,6 %) und nach der Zeitschrift *Science* insgesamt die zweitkürzesten Titel. Somit fällt der absolute Zuwachs hier besonders gering aus (0,14 Wörter pro Dekade), da die Titel im Vergleich zu den anderen Disziplinen ohnehin kurz sind und den geringsten relativen Zuwachs erfahren (vgl. hierzu auch Abbildung 4.27, S. 123). Im Gegensatz dazu ist die Länge der Titel für alle volkswirtschaftlichen Artikel im EconLit-Datensatz deutlich stärker gestiegen (5,3 %), sodass lediglich *Betriebswirtschaftslehre und Finance* einen höheren relativen Anstieg pro Dekade aufweist. Auch für die Länge der Titel zeigt sich die Wichtigkeit eines relativen Vergleichs über einen gleich langen Zeitraum. So weist die *Psychologie* zwar die längsten Titel (\varnothing 12,1 Wörter) sowie den größten absoluten und relativen Anstieg in der Anzahl der Wörter über den Beobachtungszeitraum auf, reiht sich allerdings beim Vergleich pro Dekade mit einem Zuwachs von 4,9 % hinter der *Betriebswirtschaftslehre und Finance* (6,0 %) sowie der *Volkswirtschaftslehre* (5,3 %) ein.

Kapitel 5

Umgang mit der Last des Wissens

In der Einleitung der vorliegenden Arbeit (Kapitel 1) und bei der Vorstellung von Ergebnissen verwandter Arbeiten (Kapitel 2) wurden bereits diverse Studien genannt, die eine stetig steigende Informations- und Wissensmenge (z. B. in Form von veröffentlichten Patenten, Büchern oder wissenschaftlichen Artikeln) belegen (z. B. Price 1963; Archibald und Line 1991; Lyman und Varian 2003; Bornmann und Mutz 2015). Es scheint Einigkeit darüber zu bestehen, dass Individuen und Organisationen mit zunehmenden Wissensmengen umgehen und zunehmend komplexe Probleme lösen müssen. Beispielsweise führt Aldermann (2001) aus: „in recent years the nature of many engineering projects has subtly yet substantially changed to create a greater degree of complexity, which demands learning and the acquisition and management of new sources of knowledge“ (S. 211). Daher sollte Wissensmanagement den Einfluss zunehmender Komplexität auf die Generierung impliziten und expliziten Wissens berücksichtigen (Sacchetti und Sugden 2009, S. 25).

Hinsichtlich der technologischen Leistungsfähigkeit von Wissensmanagement-Systemen gab es in den vergangenen Jahrzehnten beträchtlichen Fortschritt. Wie Drucker und Maclariello (2008) darlegen, erlaubt es “information technology [...] knowledge to spread nearly instantly and making it accessible to everyone” (S. 37). Allerdings lösen die Möglichkeiten der Informationsverarbeitung und -bereitstellung aus Nutzersicht nicht immer das Problem der Last des Wissens. Im Gegenteil können technische Helfer dieses sogar verstärken. In einem Artikel im Wall Street Journal diskutiert Davenport (2015) Gründe dafür, dass das Interesse an Wissensmanagement in den letzten Jahren abgenommen hat und schreibt: „It was too time-consuming to search for and digest stored knowledge. Even in organizations where a lot of knowledge

was contributed to KM systems—consulting firms like Deloitte and Accenture come to mind—there was often too much knowledge to sort through. Many people didn't have the patience or time to find everything they needed. Ironically, the greater the amount of knowledge, the more difficult it was to find and use".

Nonaka und Takeuchi (1995) bauen auf dem Verständnis von Wissen nach Polanyi (1969) auf, der zwischen explizitem und implizitem Wissen unterscheidet. Ihr Buch „The Knowledge Creating Company“ greift dieses Verständnis von Wissen auf und stellt einen der zentralen Beiträge des Wissensmanagements der letzten Jahrzehnte dar. Im Zentrum des darin beschriebenen Modells und folgender Modelle, wie dem von Davenport und Prusak (1998), stehen Menschen, die Informationen verarbeiten müssen, um Wissen zu generieren und die als hauptsächlicher Engpass bei der weiteren Schaffung von Wissen gesehen werden können. Sie sind dabei natürlichen Grenzen ausgesetzt, die beispielsweise ihre kognitiven Leistungen beeinflussen (Miller 1955; Kotovsky, Hayes und Simon 1985; Sweller 1994; Marois und Ivanoff 2005). Jones (2009a) stellt fest, dass Humankapital (d. h. Menschen und deren individuell gebundenes Wissen) sich insofern von anderen Bestandsgütern unterscheidet, als dass es nicht einfach transferiert werden kann und merkt an: „the vessel of human capital—the individual—is born with little knowledge and absorbs information at a limited rate“ (S. 283).

In diesem Kapitel liegt der Fokus auf der Diskussion von Ansätzen (z. B. Methoden oder Techniken), die dabei helfen, mit den zunehmend großen und komplexen Problemen bei der Wissensschaffung, -aneignung und -umwandlung umzugehen. Ferner werden die grundlegenden Prinzipien hinter den Ansätzen herausgearbeitet und ihre Anwendung auf verschiedene Bereiche illustriert. Der Fokus liegt dabei auf den strukturellen Implikationen natürlicher, menschlicher Grenzen. Es wird von selbstverschuldeten Herausforderungen abgesehen, die durch menschliches Verhalten, Interessenskonflikte und andere Bedenken entstehen können. Beispielsweise könnten Wissensarbeiter unwillig sein, ihr Wissen mit anderen zu teilen (Malhotra 2003; Wasko und Faraj 2005). Auch wenn diese Aspekte eine wichtige Rolle im Wissensmanagement spielen, liegt der Fokus auf den Charakteristiken von Prinzipien zur Lösung großer (*large-scale*) Probleme, die auf technologische und soziale Bereiche übertragen werden können.

Die ermittelten Prinzipien sind nicht nur auf die Last des Wissens, wie sie bereits in Abschnitt 1.2 (S. 4) erläutert wurde, anwendbar. Sie können grundsätzlich auch auf Probleme übertragen werden, die nicht auf die Ansammlung von Wissen zurückzuführen sind, sondern auf die Ansammlung anderer Ressourcen wie Menschen, Maschinen oder

Software. Sie werden daher auch für verwandte Probleme diskutiert, die ähnlichen Grundsätzen des Wachstums und der Akkumulation von Ressourcen folgen. Während die einzelnen Prinzipien entweder selbsterklärend oder in der bestehenden Literatur etabliert sind, werden sie hier in ein einfaches und doch übergreifendes Rahmenwerk eingeordnet und durch diverse Beispiele konkreter Ansätze veranschaulicht, die die Prinzipien implementieren.

Methodisch wird hier die Perspektive der *Design Science* (Hevner und Chatterjee 2010; Vaishnavi und Kuechler 2008) eingenommen. Konkret wird den Schritten von Peffers u. a. (2008) gefolgt, zunächst ein Problem zu identifizieren, Ziele bzgl. dessen Lösung darzulegen, ein Rahmenwerk zu entwickeln und dieses Rahmenwerk mittels Beispielen zu veranschaulichen. Insofern ist der Beitrag dieses Kapitels konzeptioneller sowie exploratorischer Natur und stellt einen Startpunkt für weitere Diskussionen und zukünftige Forschung dar, die beispielsweise in der Sammlung empirischer Evidenz und der Ableitung konkreter betrieblicher Maße und Empfehlungen bestehen könnte.

5.1 Prinzipien zum Umgang mit der Last des Wissens

Ausgangspunkt dieses Kapitels bzw. der folgenden Überlegungen ist eine Situation, in der die verfügbaren Methoden oder Ressourcen nicht ausreichen, um ein gegebenes Problem zu lösen. Der zur Lösung eines Problems aufgewandte Input besteht aus den eingesetzten Methoden und Ressourcen. Diese werden im Folgenden auch abstrakt als *Werkzeug* bezeichnet. Die Ausgangssituation ist in Abbildung 5.1 jeweils auf der linken Seite veranschaulicht, indem das gegebene Problem größer ist als das zur Verfügung stehende Werkzeug. Die im Folgenden vorgestellten Prinzipien beschreiben Wege, die Lösung des gegebenen Problems durch geeignete Transformation zu ermöglichen oder zumindest anzunähern. Die drei grundlegenden Prinzipien lauten:

- 1) Erhöhung eines Inputfaktors,
- 2) paralleler Einsatz von Inputfaktoren,
- 3) Problemreduktion.

Die Prinzipien 1) und 2) erhöhen die Leistungsfähigkeit oder die Anzahl der eingesetzten Ressourcen, um einer Aufgabe gegebener Größe zu begegnen. Prinzip 3) reduziert die Problemstellung, sodass sie mit den gegebenen Ressourcen gelöst werden kann. Bevor die einzelnen Prinzipien in den folgenden Abschnitten näher erläutert werden,

soll das folgende einfache Beispiel als erste Illustration dienen: Angenommen das vorliegende Problem besteht darin, in einem vorgegebenen Zeitraum einen Literaturüberblick zu einer vorgegebenen Anzahl wissenschaftlicher Arbeiten zu geben. Ferner seien der Zeitraum und die Anzahl der Arbeiten so gewählt, dass die Aufgabe von einem Wissenschaftler nicht im Rahmen seiner Arbeitszeiten zu bewerkstelligen ist. Das gegebene Problem ist somit größer als die zunächst zur Verfügung stehenden Mittel.

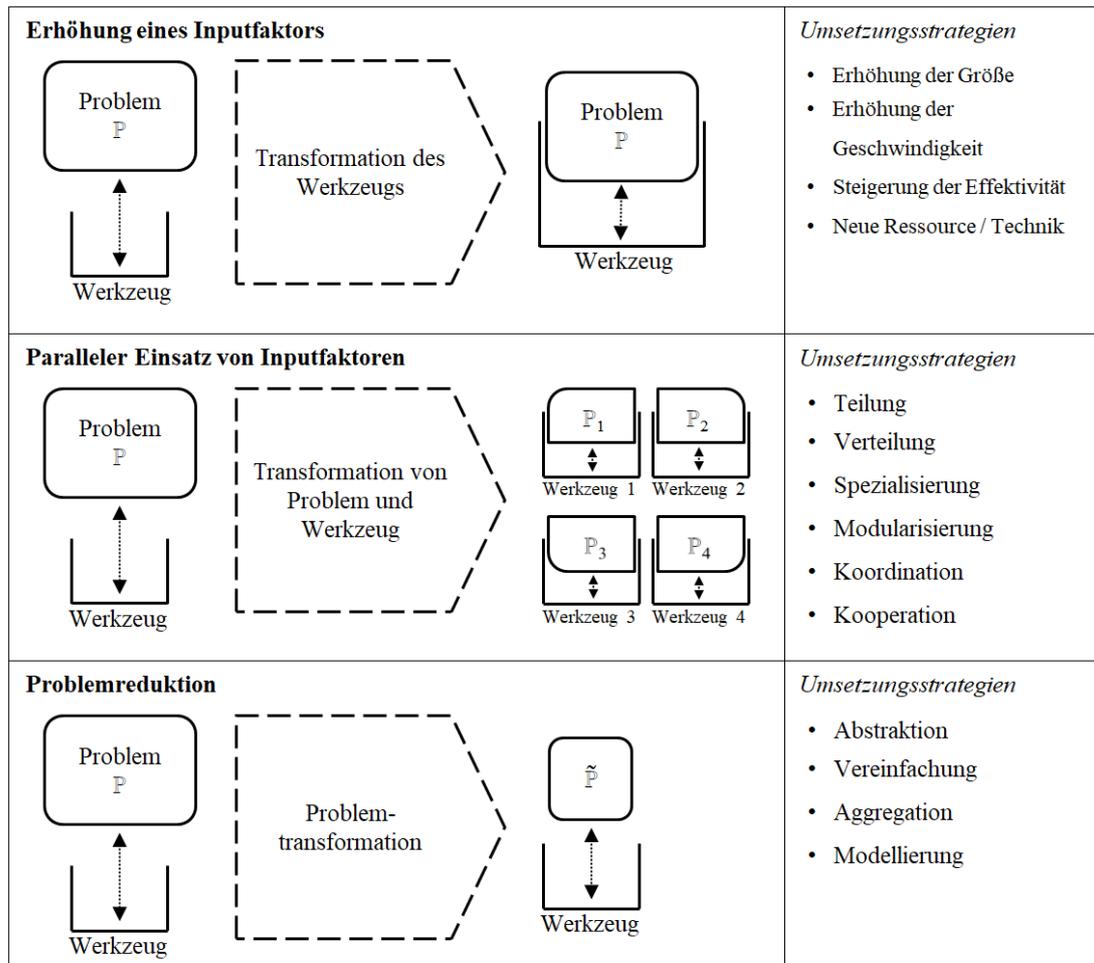


Abbildung 5.1: Die drei allgemeinen Prinzipien zum Umgang mit und der Lösung von komplexen Problemen sowie beispielhafte Strategien ihrer Umsetzung.

Das Prinzip 1) setzt nun bei der Person an, die den Literaturüberblick ausarbeiten soll und versucht dessen Leistungsfähigkeit zu steigern (Erhöhung des Inputs bzw. Transformation des Werkzeugs). Dies ist beispielsweise dadurch möglich, dass der Wissenschaftler mehr als seine tägliche Arbeitszeit investiert und durch diese Erhöhung des Ressourceneinsatzes das Problem doch in der gegebenen Zeit bewältigen kann. Prinzip 2) sieht hingegen vor die zu verarbeitenden Arbeiten aufzuteilen und mehrere Wissenschaftler gleichzeitig an der Erstellung des Literaturüberblicks arbeiten zu lassen,

um diesen in der gegebenen Zeit fertigzustellen (gleichzeitiger Einsatz mehrerer Werkzeuge). Prinzip 3) setzt nicht bei der eingesetzten Ressource bzw. dem eingesetzten Werkzeug an, in diesem Fall also dem Wissenschaftler, sondern am Problem selbst und reduziert dieses. Dies führt zwar nicht zur vollständigen Lösung des ursprünglichen Problems, allerdings lassen sich hierdurch gegebenenfalls die wichtigsten Artikel verarbeiten und ein Teil der Aufgabe im gegebenen Rahmen zufriedenstellend lösen. Abbildung 5.1 illustriert die Prinzipien und zählt einige Strategien ihrer Umsetzung auf.

5.1.1 Das Prinzip der Erhöhung eines Inputfaktors

Ein natürlicher Ansatz, ein größeres Problem zu lösen, ist hierfür ein „größeres Werkzeug“ zu benutzen. Wenn ein Bolzen nicht mit einem gewöhnlichen Hammer ins Material geschlagen werden kann, dann löst möglicherweise ein Vorschlaghammer diese Aufgabe. Im Wissensmanagement sind die Schaffer, Boten und Verarbeiter von Wissensaufgaben häufig einzelne Individuen. Um die Leistungsfähigkeit eines Individuums zu erhöhen, kann diese Person härter, schneller oder länger arbeiten oder mehr üben, um effizienter zu werden. Ein neuer, vielversprechender Trend hinsichtlich der Aufnahme von Wissen ist das multimediale Lernen, das helfen kann, die kognitiven Limitationen von Individuen zu umgehen bzw. zu überwinden (vgl. Abschnitt 5.2.1 für eine weitere Diskussion).

Die Leistungsfähigkeit eines Werkzeugs zu erhöhen erfordert typischerweise dessen Transformation und damit zusätzlichen Aufwand. Die Erhöhung der Leistungsfähigkeit stößt allerdings an einem bestimmten Punkt an ihre Grenzen. Zum Beispiel kann ein Individuum zwar lernen schneller zu lesen oder mehr Zeit investieren, allerdings gibt es heutzutage in vielen Bereichen derart viel Wissen, dass ein Individuum sich dieses nicht vollständig innerhalb eines gegebenen Zeitraums oder sogar im Laufe seines Lebens aneignen kann. Derartige Begrenzungen bestehen prinzipiell für alle Inputfaktoren in allen Bereichen. Beispielsweise, unterliegt die Entwicklung von Halbleitern physikalischen Gesetzen und erfährt Limitationen bei der Verkleinerung, der Geschwindigkeit und dem Energieverbrauch. Die Entwicklung von Maschinen unterliegt physikalischen Beschränkungen, die durch die Charakteristiken der verwendeten Materialien und die Mechaniken der Konstruktionen bestimmt sind.

Ferner kann die Beschränkung eines Inputfaktors in einem interdependenten System das Wachstum des gesamten Systems verlangsamen. Solche Begrenzungen bezeichnet man auch als Nadelöhre (engl. *bottle necks*). Beispielsweise lässt sich die Höchst-

geschwindigkeit eines Automobils steigern, indem man das Fahrzeug leichter und aerodynamischer baut sowie einen leistungsfähigeren Motor einsetzt. Entwickelt man allerdings gleichzeitig nicht die Karosserie oder einzelne Bauteile wie die Reifen derart, dass sie den erhöhten Kräften standhalten, dann limitieren diese letztlich die maximal erreichbare Geschwindigkeit. Entsprechend kann auch ein Wissensarbeiter seine Arbeitszeiten erhöhen und gleichzeitig seine Lesegeschwindigkeit steigern. Ist allerdings das ganze System der Wissensverarbeitung durch den kognitiven Verarbeitungsprozess in seinem Gehirn limitiert, dann sind seiner Wissensaufnahme trotz der Erhöhung zweier Inputfaktoren Grenzen gesetzt.

Letztlich skaliert das Erhöhen eines einzelnen Inputfaktors nur bis zu einem gewissen Punkt. Stattdessen müssen andere Ansätze die Beschränkungen überwinden, die einem Individuum oder einer Maschine innewohnen, um Aufgaben zu lösen, die hinter dem Punkt liegen, der mit maximaler Erhöhung eines einzelnen Faktors erreicht werden kann.¹ Abbildung 5.1 illustriert zwei Alternativen. Die Idee hinter der ersten Alternative ist es, das größere Problem in kleinere Probleme aufzuteilen und diese unter Einsatz mehrerer Werkzeuge bzw. Lösungsinstanzen separat anzugehen. Die zweite Alternative reduziert das gegebene Problem auf ein kleineres bzw. einfacheres Problem und löst dieses mit den zur Lösung zur Verfügung stehenden Mitteln. Die zweite Alternative liefert somit nicht zwingend eine Lösung des ursprünglichen Problems, sondern eines verwandten Problems, dessen Lösung ausreichend nah an die des ursprünglichen Problems herankommt.

5.1.2 Das Prinzip des parallelen Einsatzes von Inputfaktoren

Beim parallelen Einsatz von Inputfaktoren wird die Anzahl der zur Lösung eines Problems eingesetzten Werkzeuge erhöht, um der Größe der Aufgabe gerecht zu werden. Hierfür muss das zugrunde liegende Problem in mehrere separate Teilprobleme bzw. Teilaufgaben aufgespalten werden können (z. B. Smith und Davis 1981). Anschließend wird jedes dieser Teilprobleme mit Hilfe jeweils eines Werkzeugs einzeln angegangen. In den Computerwissenschaften ist dieses Prinzip auch unter dem Begriff *Divide & Conquer* bekannt (z. B. Kuck 1980; Smith 1985). Die Teilaufgaben können dabei parallel verarbeitet werden, um schneller zu Ergebnissen zu gelangen (z. B. Horowitz und Zorat 1983; Atallah, Cole und Goodrich 1989). Falls die Teilaufgaben verschiedene,

¹Dabei wird auch der Begriff „Lösungsansatz“ verwendet, wobei er sich nicht auf eine übergreifende „Lösung“ für besagte Last bezieht, sondern eher auf die gewählten Instrumente zur Lösung spezifischer Probleme.

spezifische Aspekte der ursprünglichen Aufgabe repräsentieren, kann jedes Werkzeug auf die spezifischen Anforderungen optimiert werden. Daher erlaubt die Aufteilung einer Aufgabe in Teilaufgaben die Spezialisierung auf Teillösungen. Das Prinzip umfasst also auch die spezialisierte Arbeitsteilung, wie sie beispielsweise von Adam Smith (1776) für die Produktion beschrieben wurde. Es ist allerdings nicht auf diese beschränkt, sondern umfasst ebenso die homogene Aufteilung eines teilbaren Problems. Ferner war eines der Ziele von Smith den Durchsatz in der Produktion zu erhöhen. Das Ziel der in diesem Kapitel vorgestellten Prinzipien ist allerdings größere und komplexere Probleme überhaupt zu lösen und nicht die Anzahl der gelösten Probleme pro Zeiteinheit zu steigern. Die gesteigerte Produktivität kann dabei ein Nebeneffekt der Implementierung der drei Prinzipien sein.

Das prototypische Beispiel für den parallelen Einsatz von Inputfaktoren ist der Einsatz von Teamarbeit. Teammitglieder können größere Aufgaben aufteilen und diese Teilaufgaben unter den verschiedenen Mitgliedern verteilen (vgl. z. B. McDowell und Melvin 1983; Barnett, Ault und Kaserman 1988; Jones 2009a). Indem ein Teammitglied sich auf das Wissen beschränkt, welches notwendig ist, um die ihm zugeteilte Teilaufgabe zu lösen, bleibt die Menge des Wissens, das es sich aneignen muss, auf diesen Bereich beschränkt. Aus diesem Grund hält Jones (2009a) Teamarbeit für ein hervorragendes Mittel, um der Last des Wissens zu begegnen. Melero und Palomeras (2015) fügen dieser Idee einen weiteren Aspekt an und argumentieren, dass manche Teams neben der Verteilung von und Spezialisierung auf Teilaufgaben gegebenenfalls eine weitere Komponente benötigen. Sie zeigen empirisch die Rolle und Wichtigkeit von Generalisten für den Teamerfolg. Generalisten können demnach die Koordination übernehmen, die notwendig wird, wenn die Teilaufgaben und entsprechenden Teillösungen nicht unabhängig voneinander sind.

Wie erwähnt ist eine Voraussetzung für den parallelen Einsatz von Inputfaktoren, dass das zugrunde liegende Problem teilbar ist. Die einzelnen Lösungen der Teilprobleme müssen, nachdem das Problem geteilt wurde und die Teilaufgaben verteilt sowie gelöst wurden, anschließend zu einer Gesamtlösung der ursprünglichen Aufgabe kombiniert werden. Im Falle, dass es Abhängigkeiten zwischen den Teilaufgaben gibt, ist die Koordination zusätzlich erschwert. Idealerweise sollten Interdependenzen daher vermieden werden, da Koordination zusätzlichen Aufwand bedeutet (für den Aufwand der Koordination in Forschungsteams vgl. z. B. Hudson 1996; Cummings und Kiesler 2007; Bercovitz und Feldman 2011). Falls alle Teilaufgaben unabhängig voneinander behandelt werden können und daher keine Koordination notwendig ist, spricht man von modularen Teilaufgaben. Da es den parallelen Einsatz von Inputfaktoren vereinfacht

und erleichtert, kann Modularität als ein wichtiges und angestrebtes Design-Prinzip angesehen werden.

5.1.3 Das Prinzip der Problemreduktion

Das Prinzip der Problemreduktion zielt darauf ab, die Größe des ursprünglichen Problems zu reduzieren, um es der Leistungsfähigkeit zur Verfügung stehender Werkzeuge, Methoden bzw. Ressourcen entsprechend anzupassen. Strategien zur Problemreduktion beinhalten die Vereinfachung durch Abstraktion, das Weglassen nicht-essentieller Komponenten oder die Aggregation. Eine spezielle Form der Abstraktion ist das Aufstellen eines formalen Modells, das anstelle des ursprünglichen Problems betrachtet wird. In der Wissenschaft reduzieren formale Modelle regelmäßig größere und komplexere Probleme auf wenige mathematische Gleichungen. Wie Hermann Hesse einst sagte: „Jede Wissenschaft ist, unter anderem, ein Ordnen, ein Vereinfachen, ein Verdaulichmachen des Unverdaulichen für den Geist.“ (Hesse 1970, S. 179).² Diese Verwertbarkeit geht allerdings häufig auf Kosten der Details des ursprünglichen Problems. In den Computerwissenschaften gibt es die Unterscheidung zwischen einer Komprimierung mit (engl. *lossy*) und ohne (engl. *lossless*) Informationsverlust (z. B. Howard 1996; Sheng u. a. 1998), vor allem im Kontext der Speicherung von Texten oder Bildern. Bis zu einem gewissen Punkt ist es gegebenenfalls möglich, ohne Informationsverlust zu abstrahieren. Allerdings führt eine Abstraktion darüber hinaus dazu, dass das reduzierte Problem zunehmend vom originären Problem abweicht. Daher beinhaltet die Problemreduktion häufig einen Kompromiss (engl. *trade-off*) zwischen der Lösbarkeit eines vereinfachten Problems und dem Erhalt aller Aspekte des originären Problems.

Die Vereinfachung und Abstraktion von Problemen ermöglicht es außerdem, diese in Relation zu einer allgemeineren Klasse von Problemen zu setzen und diese mit den bereits entwickelten Methoden zu lösen, die gegebenenfalls in einem anderen Kontext entwickelt wurden. Daher ist Abstraktion häufig eine Voraussetzung für den Einsatz einer neuen Methode oder den Transfer einer bestehenden Methode auf die Lösung eines Problems.

²Die Erstausgabe dieses Werkes von Hesse stammt aus dem Jahr 1943.

5.1.4 Implikationen

Auf Basis der oben diskutierten Prinzipien können drei Implikationen gezogen werden. Erstens, da die Erhöhung eines einzelnen Inputfaktors schwer auf größere Probleme skalierbar ist, sind die Prinzipien des parallelen Einsatzes von Inputfaktoren und die Problemreduktion vorzuziehen. Zweitens, kann der parallele Einsatz von Inputfaktoren lediglich auf solche Probleme angewandt werden, die eine Aufteilung des Problems erlauben und bei denen die Teillösungen zu einer Lösung des ursprünglichen Problems zusammengesetzt werden können. Auch die Anzahl von Instanzen, die zur Lösung mit diesem Lösungsprinzip notwendig ist, steigt mit der Problemgröße. Außerdem ist für die Zerlegung des ursprünglichen Problems in Teilprobleme und das Wiederausammen-setzen der Teillösungen zusätzlicher Aufwand notwendig.

Drittens, kann Problemreduktion nur auf solche Probleme angewandt werden, die in sinnvoller Weise reduziert werden können. In manchen Fällen steigt der Aufwand des zweistufigen Prozesses, ein Problem zunächst zu reduzieren und anschließend zu lösen, langsamer als der Aufwand zur Lösung des ursprünglichen Problems mit der Problemgröße zunimmt. In diesen Fällen ist Problemreduktion als eine kosteneffiziente und skalierbare Lösung vorzuziehen. Dabei sollte allerdings beachtet werden, dass Problemreduktion immer mit einem Kompromiss zwischen der Lösbarkeit eines vereinfachten Problems und dem möglichen Verlust an Informationen einhergeht. Diese Gegenpositionen sollten aufmerksam abgewogen werden. Ferner ist auch eine Kombination der Prinzipien denkbar, um einen größeren Effekt zu erzielen.

5.2 Beispiele für die Umsetzung der drei Prinzipien

In Abbildung 5.2 finden sich konkrete Beispiele für die Umsetzung der obigen drei Prinzipien in den Bereichen *Wissensmanagement und IT*, *Forschung und Entwicklung*, *Management* sowie *organisatorische Strukturen*. In den folgenden Abschnitten werden diese Beispiele, das heißt konkrete Methoden und Techniken, für jedes der drei Prinzipien näher erläutert.

5.2.1 Erhöhung eines Inputfaktors

In den letzten Jahren wurden verschiedene Ansätze entwickelt, um die Wissensverarbeitung von Individuen zu verbessern. Beispielsweise wurden Entscheidungsunterstüt-

zungssysteme (engl. *Decision Support Systems*) und Arbeitsflusssysteme entwickelt, um Individuen und Gruppen das Treffen von Entscheidungen zu erleichtern (hierzu geben beispielsweise Er (1988) oder Power (2002) einen Überblick). Crossland, Wynne und Perkins (1995) berichten, dass der Einsatz von Entscheidungsunterstützungssystemen in der schnelleren Lösung von Problemen und weniger Fehlern resultiert.

Elektronisches oder multimediales Lernen haben eine bedeutende Innovation im Bereich der Lerntechniken angestoßen, die das Potenzial haben, den Wissenstransfer zu verbessern. Diese Formate unterstützen die kognitiven Fähigkeiten von Studierenden besser und machen den Wissenstransfer daher effektiver als die traditionelle Lehre (Wegner, Holloway und Garton 1999; Joy und Garcia 2000; McDonald 2002). Online-Lehre ermöglicht eine Kombination aus kurzen Videosequenzen, Sprachaufnahmen und Animationen. Eingebaute Aufgaben und Quizze halten die Aufmerksamkeit der Studenten aufrecht, sorgen dafür, dass diese die präsentierten Inhalte direkt reflektieren und ermöglichen eine sofortige Überprüfung des Verständnisses. Diverse Studien und Meta-Analysen belegen die erhöhte Leistung der Online-Lehre gegenüber traditioneller Lehre (Maki u. a. 2000; Shachar und Neumann 2003; Lapsley u. a. 2008). Ein Beispiel solcher Online-Lehre sind die sogenannten *Massive Open Online Courses* (MOOCs). Neben den erwähnten Vorteilen digitaler Lehre erlauben es die hohen Teilnehmerzahlen (z. B. Pappano 2012) von MOOCs sehr spezifische Kurse anzubieten, für die sich innerhalb einzelner Organisationen zu wenige Teilnehmer gefunden hätten. Da solche Kurse permanent verfügbar und wiederholbar sind, sobald sie einmal aufgezeichnet wurden, spart dieser Ansatz auch Ressourcen (z. B. Zeit oder Geld), die anderweitig für Forschung oder die Verbesserung bestehender Kurse eingesetzt werden können. Zwei weitere Vorteile der Online-Lehre erleichtern die Skaleneffekte von MOOCs. Studierende können diese Kurse von jedem beliebigen Ort auf der Welt zu jeder beliebigen Zeit verfolgen, solange sie eine Internetverbindung haben. Das bedeutet, dass Studierende, die einen MOOC verfolgen, temporär und geografisch unabhängig sind von den Lehrpersonen und anderen Studierenden. Sie genießen somit eine erhöhte Flexibilität, was sich gegebenenfalls positiv auf den Lernerfolg auswirkt (Allen und Seaman 2006; Means u. a. 2009).

Im Idealfall hilft multimediales Lernen Individuen schneller zu lernen, mehr in der gleichen Zeit zu lernen und möglicherweise sogar etwas zu lernen, das bei einer anderen Form der Wissensvermittlung ihre kognitiven Fähigkeiten überstiegen hätte. Beispielsweise schreibt Mayer (2005): „people learn more deeply from words and pictures than from words alone“ (S. 31). Fletcher und Tobias (2005) beziehen sich auf die Philosophie des Bischofs George Berkeley aus dem 18. Jahrhundert, der suggerierte,

	Prinzip		
	Erhöhung eines Inputfaktors	Paralleler Einsatz von Inputfaktoren	Problemreduktion
Wissensmanagement und IT	Üben und Lernen Multimediales Lernen Erhöhung des Zeitaufwands Erhöhung des Speichers und der Rechenleistung Optimierung von Algorithmen	Groupware Systeme Online Kurse (MOOCs) Verteilte Datenbanken Normalisierung von Datenbanken Divide & Conquer Algorithmen Grafikkarten	Templates and Views Objektorientierte Programmierung (Schnittstellen)Standards (USB etc.) Komprimierungsalgorithmen APIs
Forschung und Entwicklung	Längere Ausbildungs- bzw. Entwicklungszeiten Verbesserungen in der Bildung und Lehre Verbesserte Ausstattung (z. B. im Labor)	Koautorenschaften Forschungsk Kooperationen Forschungsaufenthalte	Abstracts bei wiss. Artikeln Qualitätskontrolle durch Peer-Review-Prozess und komprimierte Veröffentlichung in renommierten Zeitschriften
Management im Allgemeinen	Einstellungsverfahren, Leistungsfähigkeit von Angestellten, Maschinen und Informationssystemen Tätigkeitsdauer Anreizsysteme Prozessoptimierung	Modulares Management Projektmanagement Geschäftsreisen, Telefonkonferenzen Corporate Venture Capital (CVC)	Management Scorecard TRIZ Unternehmenskennzahlen (z. B. Umsatz, Gewinn, ROI)
Anwendungsbereich Organisationale Strukturen (Firmen, Staaten)	Anreizsysteme	Abteilungsstrukturen Föderalismus Hierarchien	Politische Parteien Wahlen

Abbildung 5.2: Beispiele für die konkrete Umsetzung der drei Prinzipien zum Umgang mit und der Lösung von komplexen Problemen.

dass Worte andere Denkprozesse anstoßen als Bilder. Während das „Haus“ einen an das abstrakte Konzept eines Hauses denken lässt, erweckt die Fotografie eines Hauses das entsprechende Bild im Kopf. Schweitzer (2012) führt aus, wie verschiedene Komponenten elektronischer Anweisungs- und Lernsysteme die kognitiven Restriktionen von Individuen überwinden können. Durch die Verbesserung der Effektivität von Lehrinstrumenten können technologische Systeme dazu beitragen, die Grenzen individueller Wissensarbeiter zu verschieben.

Es gibt verschiedene Wege einen einzelnen Inputfaktor in technologischen Bereichen und im Management zu erhöhen. Im Rahmen der Data Science und Computerwissenschaft umfassen technische Prozeduren zur Erhöhung der Verarbeitungsgeschwindigkeit die Ausweitung des Speichers, die Erhöhung der Rechenoperationen pro Verarbeitungszyklus und die Ausweitung der Bandbreite zwischen den einzelnen Komponenten (Hennessy und Jouppi 1991). Im Management steckt man mehr Ressourcen in einen Prozess, stellt mehr qualifizierte Arbeitskräfte ein, besorgt bessere Ausstattung und erhöht das Budget, um größere Aufgaben effektiver zu lösen. Zusätzlich kann die Leistungsfähigkeit durch zielgerichtete Anreizsysteme weiter gesteigert werden, die Mitarbeiter dazu motivieren, ihre optimale Leistung abzurufen. Entsprechend können Unternehmen in leistungsfähigere und besser ausgebildete Mitarbeiter investieren, die für sie größere und komplexere Aufgaben und Probleme lösen (Huselid 2013). Erhöhte Leistung in der Forschung kann dadurch erreicht werden, dass man mehr Ressourcen wie Ausstattung und Laborzeit in ein Forschungsthema steckt. Auch Investitionen in die Ausbildung von Forschern sollten in höherer Pro-Kopf-Leistung resultieren.

5.2.2 Paralleler Einsatz von Inputfaktoren

Es gibt verschiedene Wege Inputfaktoren parallel einzusetzen. Eine klassische Illustration dieser Idee ist Adam Smiths Beschreibung der Arbeitsteilung. Smith (1776, Kapitel 1) erzählt die Geschichte eines Stecknadel-Herstellers: „a workman not educated to this business (which the division of labor has rendered a distinct trade), nor acquainted with the use of the machinery employed in it (to the invention of which the same division of labor has probably given occasion), could scarce, perhaps, with his utmost industry, make one pin in a day, and certainly could not make twenty. But in the way in which this business is now carried on, not only the whole work is a peculiar trade, but it is divided into a number of branches, of which the greater part are likewise peculiar trades.“

Im Wissensmanagement wurden Groupware Systeme entwickelt, um Individuen zu helfen, gemeinsam Aufgaben zu bearbeiten (vgl. beispielsweise Ellis, Gibbs und Rein (1991) für einen Überblick). Auch beim Wissenstransfer ermöglichen Informationssysteme den parallelen Einsatz von Input. Abgesehen von den positiven Effekten in Bezug auf die Effizienz des Lernens, kann quasi eine unbeschränkte Anzahl von Studierenden an den oben erwähnten MOOCs teilnehmen. Pappano (2012) zufolge meldeten sich mehr als 150.000 Teilnehmer für den ersten Kurs der MOOC-Plattform Udacity „Introduction to Artificial Intelligence“ an. Dies sind Teilnehmerzahlen, die die räumlichen Kapazitäten einer klassischen Vorlesung an einer Universität übersteigen. Insofern teilen Informationssysteme das Problem einer sehr hohen Teilnehmerzahl in viele einzelne Lerngruppen und ermöglichen die Inhalte solcher Online-Vorlesungen einer breiten Zuhörerschaft zu vermitteln.

Ein Beispiel für Modularität in akademischer Forschung ist die Zusammenarbeit von Koautoren. Die Zusammenarbeit mehrerer Autoren und Interdisziplinarität haben in den vergangenen Jahren zugenommen (z. B. Laband und Tollison 2000; Adams u. a. 2005; Singh und Fleming 2010; Hamermesh 2015), während Artikel, die lediglich von einem Autor geschrieben werden, seltener geworden sind (Glänzel und Schubert 2005). Bis zu einem gewissen Ausmaß kann diese Zunahme externen Faktoren, wie der Globalisierung, verbesserten und vereinfachten Möglichkeiten der Kommunikation, als auch Publikationsstrategien der Autoren zugeschrieben werden. Die offensichtlichste Funktion einen Artikel gemeinsam zu schreiben und über Disziplinen hinweg zusammenzuarbeiten ist allerdings die Auf- und Verteilung der Aufgaben in einem Forschungsprojekt unter mehreren Forschern mit spezifischen Kompetenzen in verschiedenen Bereichen. Ferner besitzen größere Teams tendenziell heterogeneres Wissen und heterogenere Perspektiven auf denselben Sachverhalt. Burt (2004) zeigt, dass verschiedene Perspektiven die Kreativität und die Problemlösungskompetenz steigern können, wie dies beispielsweise im Kontext von Innovationen von Hargadon und Sutton (1997) gezeigt wurde.

Informationsverarbeitung befasst sich häufig mit mehreren Formen der Modularisierung. Auf Seiten der Hardware kommen routinemäßig Techniken der Modularisierung zum Einsatz. Beispiele hierfür sind die Bereitstellung spezieller Prozessoren für spezifische Aufgaben (z. B. Grafik, Mathematik, Sprach- und Bewegungsverarbeitung), der simultane Einsatz mehrerer Prozessorkerne und die Nutzung von Netzstrukturen, um Aufgaben zu verteilen und auf mehreren, separaten Recheneinheiten zu lösen. Software unterstützt nicht nur die eben erwähnten Hardware-Funktionen, sondern kennt ebenfalls modulare Konzepte wie *Divide & Conquer* und das objektorientierte Programmieren.

Heutzutage ist Arbeitsteilung in der Produktion gängige und wohlbekannte Praxis und Modularisierung gewinnt zunehmend an Bedeutung (Langlois 2002). Ferner wird dieses Konzept bei anderen Aufgaben und im Management selbst angewandt. Elemente der Modularisierung werden vom Projektmanagement eingesetzt, welches spezifische Aufgaben herausbricht, die von einem entsprechenden Team gelöst werden sollen. Ein weiteres Beispiel der Modularisierung sind die Einstellung und der Einsatz spezialisierter Fachkräfte für spezifische Aufgaben.

In organisatorischen Strukturen wird Modularisierung häufig nach Phasen des Wachstums beobachtet. Wenn Unternehmen eine bestimmte Größe erreichen, werden sie in Abteilungen und Gruppen mit spezifischen Funktionen bzw. Aufgabenbereichen wie Marketing, Forschung und Entwicklung, Produktion und Einkauf strukturiert. Schließlich können Unternehmen manche ihrer bisherigen Funktionen auslagern, die sie nicht mehr zu ihren Kernkompetenzen zählen. In ähnlicher Weise werden auch Länder bzw. Nationen nicht in allen Bereichen von einer zentralen Regierung regiert, sondern sind, wie im Fall von Deutschland, auf der Ebene der Bundesländer, Regionen und Kreise organisiert. Die Tatsache, dass sogar Länder wie China oder Russland, die für ihre starke zentrale Regierung bekannt sind, in Regionen unterteilt sind, zeigt, wie allgegenwärtig das Prinzip der Modularisierung in großen Strukturen wie Ländern ist.

5.2.3 Problemreduktion

Es gibt verschiedene Methoden und Techniken, die das Prinzip der Problemreduktion durch Abstraktion umsetzen. Heutzutage ist es üblich, ein wissenschaftliches Papier mit einem *Abstract* zu beginnen, einem Paragraphen, der die Motivation, Forschungsfrage und Hauptergebnisse zusammenfasst. Die Praxis ein *Abstract* anzugeben, erleichtert und beschleunigt die Einschätzung des Inhalts und damit der relevanten Informationen eines Artikels, seiner Einordnung in die einschlägige Literatur und die sorgfältige Literaturrecherche. Ähnlich wie Schnittstellen bei der Programmierung erfüllt das *Abstract* gewissermaßen die Funktion einer Schnittstelle zum Artikel und erlaubt die schnelle Erfassung seiner Inhalte und die Nutzung des Artikels für weitere Forschung. Ein weiteres Beispiel der Abstraktion bzw. Reduktion in diesem Kontext sind *Highlights*, die über den Abstract hinaus eine weitere Ebene der Zusammenfassung darstellen. Dabei werden Autoren dazu aufgefordert bei einer Einreichung, neben dem zusammenfassenden Abstract, eine Liste mit wenigen Stichpunkten (in der Regel drei bis fünf) anzugeben, die die wichtigsten Inhalte bzw. Ergebnisse des Artikels wiedergeben.

Ein nahezu ideales Beispiel für Abstraktion und den Transfer von Technologien in die technologische Entwicklung ist Genrich Altshullers *Theorie der Erfinderschen Problemlösung* (TRIZ). Altshuller (1994) analysierte Patentdaten, um generische Probleme und deren Lösungen zu identifizieren, die sich aus in Konflikt stehenden Zielen ergeben. Wenn es beispielsweise einen Konflikt zwischen der Masse und der Länge eines sich bewegenden Objekts gibt (Widersprüche 1 und 3 in der TRIZ-Matrix), identifizierte Altshuller dafür die Lösungen Gewichtsreduktion, Dynamisierung, Pneumatik und Hydraulik. Die Stärke von TRIZ liegt in der Abstraktion konkreter technologischer Probleme, um generische Probleme und deren Merkmale zu erhalten. Der Vorteil liegt darin, dass es weniger generische als spezifische Probleme gibt. Löst man ein generisches Problem, so hat man einen Bauplan zur Lösung unzähliger weiterer spezieller Instanzen dieses Problems.

Programmierer und Computer-Wissenschaftler haben eine systematische Herangehensweise an Aggregation und Abstraktion. Neben der regelmäßigen Formulierung von Problemen in abstrakter Sprache beim Programmieren, übertragen Computer-Wissenschaftler Abstraktion auf ihre Lösungsansätze. Bei der objektorientierten Programmierung reicht es zu einem gewissen Grad aus, dass Lösungen Blackboxen sind. Um Quellcode zu entwickeln, der eine spezifische Aufgabe löst, ist kein Wissen über die Funktionsweise dieser Blackboxen nötig. Stattdessen ist der Code in Form von Objekten verkapselt, die lediglich eine dokumentierte Schnittstelle für Input und Output aufweisen. In gleicher Weise ermöglichen standardisierte Hardware-Schnittstellen (z. B. USB-Anschluss) eine „plug and play“-Nutzung von Hardware-Komponenten.

Eine weitere allgemein bekannte Technik der Problemreduktion ist die Vereinfachung durch Aggregation. Die Natur von Entscheidungsproblemen ist, dass sie in einfachen, binären Entscheidungen resultieren müssen. Manager basieren ihre Entscheidungen auf Zahlen, die komplizierte Realitäten auf einfache Punktwerte (z. B. Umsatz, Gewinn oder *Return on Investment*) aggregieren. In der Politik dienen Parteien dazu, die vielen Meinungen in einem Land zu aggregieren und deren Anzahl signifikant zu reduzieren. Darüber hinaus dient das Parteiprogramm dazu, Meinungen vieler Individuen in einem einzigen Dokument abzubilden. Auf ähnliche Weise aggregieren die Stimmen bei Wahlen die öffentliche Meinung auf einfache binäre Entscheidungen.

Kapitel 6

Zusammenfassung, Schlussfolgerungen und Ausblick

6.1 Zusammenfassung und Diskussion der Ergebnisse

Die Anfangshypothese der vorliegenden Arbeit war, dass junge Wissenschaftler eine zunehmende Menge an Wissen verarbeiten müssen, um selbst zu wissenschaftlichen Erkenntnissen zu gelangen. Dieser zunehmend intensive Verarbeitungsprozess spiegelt sich einerseits in der Anzahl der Quellenangaben bei den ersten Artikeln der Autoren wider. Diese hat für alle in dieser Arbeit betrachteten Disziplinen in den untersuchten Zeiträumen (1970–2014 für die *Volkswirtschaftslehre*, 1960–2009 für *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* als auch *Psychologie* sowie 1960–2011 für die Zeitschrift *Science*) deutlich zugenommen (Abschnitt 4.2.2, S. 115). Andererseits tragen diese jungen Wissenschaftler erst zu einem späteren Zeitpunkt zu ihrem Fachbereich bei, was sich im mittleren Alter der Autoren bei ihrem ersten Artikel äußert. Sowohl für Wissenschaftler der Volkswirtschaftslehre als auch der Mathematik ist dieses im Beobachtungszeitraum gestiegen (Abschnitt 4.4, S. 129).

Unter der Annahme, dass Wissenschaftler immer mehr Wissen verarbeiten müssen, sind ferner mögliche Strategien hinsichtlich des Verarbeitungsprozesses und des Verfassens von Artikeln zu beobachten. Diese betreffen einerseits die zunehmende Spezialisierung der Autoren (vgl. Abschnitt 4.5, S. 135), die sich darin äußert, dass Autoren zwischen ihren ersten beiden Solo-Artikeln seltener ihr Wissenschaftsfeld wechseln. Um der zunehmenden Menge an Wissen zu begegnen, spezialisieren sich Wissenschaftler also zunehmend auf kleinere Teilgebiete ihres Fachbereichs, für die entsprechend

auch die Menge des zu verarbeitenden Wissens geringer ist. Zudem ermöglicht der Verbleib innerhalb eines Teilgebiets dem Autor, sein bereits erarbeitetes Wissen und die Bekanntschaften (z. B. zu Redakteuren oder Gutachtern relevanter Zeitschriften), die er im Umfeld des Teilgebiets gesammelt hat, zu nutzen. Bestehende Bekanntschaften können unter Umständen im Publikationsprozess behilflich sein (Colussi 2018). Andererseits ist zu beobachten, dass ein Artikel von einer zunehmenden Anzahl an Autoren verfasst wird. Hierin könnten sich zwei Effekte widerspiegeln: Einerseits müssen zunehmend spezialisierte Wissenschaftler zusammen arbeiten, wenn sie ein etwas allgemeineres Problem lösen wollen. Andererseits können sich die Wissenschaftler eine zunehmende Menge anfallender Arbeit an der Verfassung eines Artikels teilen. Der Aufwand zur Lösung eines gegebenen Forschungsproblems und zur Veröffentlichung des entsprechenden Artikels in einer wissenschaftlichen Zeitschrift umfasst dabei nicht nur die Abdeckung des benötigten Wissens. Er umfasst auch Aspekte, die nur indirekt mit der Forschungsarbeit verwandt sind, so wie die Einwerbung von Mitteln, die Arbeit an der Textqualität, die Erfüllung formeller Vorgaben oder die Kommunikation mit den Redakteuren und Gutachtern.

Die empirischen Ergebnisse der vorliegenden Arbeit sind weitestgehend im Einklang mit einer Last des Wissens wie sie Jones (2009a) beschreibt und deuten auf einen allgemeineren Trend hin, der verschiedene Bereiche des technologischen und wissenschaftlichen Fortschritts betrifft. Offen bleibt, warum die mittlere Anzahl der Quellenangaben für Artikel der renommiertesten volkswirtschaftlichen Zeitschriften, die von lediglich einem Autor verfasst wurden, in den letzten Beobachtungsjahren teilweise höher ist als für Artikel mit mehreren Autoren und warum die mittlere Anzahl der Quellenangaben für Artikel der Zeitschrift *Science* mit einem oder zwei Autoren ab 1991 fallend ist. Diese Ergebnisse stimmen nicht mit den Hypothesen einer Last des Wissens überein, denen zufolge Wissenschaftler zunehmend viel Wissen verarbeiten müssen (ein Proxy vorliegender Arbeit hierfür ist die Anzahl der Quellenangaben) und sich diese Arbeit ggf. mit Koautoren teilen. Vor dem Hintergrund zunehmend großer Teams, weisen solche Autoren bzw. die zugehörigen Artikel in diesen renommierten Zeitschriften, im Vergleich zu Autoren in und Artikeln von größeren Teams, möglicherweise bestimmte Charakteristiken auf, die nicht aus den vorliegenden Datensätzen hervorgehen.

Die Last des Wissens, wie sie Jones beschreibt, entsteht aus der Natur des Wissens selbst. Daher sind Ansätze zum adäquaten Umgang mit dieser Last zentral für die zukünftige Schaffung neuen Wissens. In Kapitel 5 wurde ein allgemeines und umfassendes Rahmenwerk vorgestellt, das helfen kann, verschiedene Ansätze einzuordnen

und ihre Wirkungsweise zu verstehen. Das Rahmenwerk umfasst drei grundsätzliche Prinzipien zur Lösung größerer und komplexer Probleme. Das Prinzip der *Erhöhung eines einzelnen Inputfaktors* und der *parallele Einsatz von Inputfaktoren* erhöhen die zur Verfügung stehenden Mittel in einem Maße, dass diese der Größe des zu lösenden Problems entsprechen. Das Lösungsprinzip der *Problemreduktion* wiederum reduziert ein gegebenes Problem, sodass dieses mit den zur Verfügung stehenden Mitteln gelöst werden kann. Die Wahl des jeweiligen Prinzips und der entsprechenden Methode bzw. Technik sollte durch deren Anwendbarkeit und Skalierbarkeit motiviert sein. Das Prinzip der *Erhöhung eines Inputfaktors* ist typischerweise dadurch limitiert, dass Inputfaktoren gegebenenfalls nicht beliebig skalierbar sind, wie beispielsweise Menschen oder Maschinen. Daher muss das Wissensmanagement ab einer gewissen Größe auf besser skalierbare Lösungen zurückgreifen. Dies kann beispielsweise durch den *parallelen Einsatz von Inputfaktoren* oder durch *Problemreduktion* geschehen. Die Lösung von Problemen ist beim Einsatz von Methoden bzw. Techniken, die diese Prinzipien implementieren, unabhängig von den Limitationen einer einzelnen Person oder eines Inputfaktors.

Die Hypothese einer Last des Wissens stellt nur eine mögliche Erklärung für die beobachteten empirischen Trends dar. Abgesehen vom eigentlichen Zuwachs an Wissen könnten auch andere Faktoren eine Rolle spielen und zu den beobachteten Trends beitragen. Das Alter eines Wissenschaftlers bei seiner ersten Publikation könnte beispielsweise auch durch längere Ausbildungszeiten getrieben sein, wobei sich auch hierin der Verarbeitungsprozess bestehenden Wissens manifestiert und Ausbildungszeiten somit Teil der Last des Wissens sind. Einen weiteren Einfluss auf das Alter beim ersten Artikel hat der zunehmende Wettbewerb um Publikationen in den renommierten Zeitschriften (z. B. in Bezug auf gängige Rankings). Um den Effekt des zunehmenden Wettbewerbs zu mindern, wurde in der vorliegenden Arbeit das Alter von Autoren bei ihrem ersten Artikel in der Volkswirtschaftslehre auch für eine Vielzahl von Zeitschriften betrachtet, die nicht zu den renommiertesten Zeitschriften gehören. Für diese breitere Stichprobe zeigt sich ebenfalls, dass das Alter beim ersten Artikel zunimmt (vgl. Abbildung 4.32, S. 130), auch wenn der Anstieg hier nicht so stark ausfällt wie für die renommiertesten Zeitschriften. Eine gestiegene Lebenserwartung (z. B. Manton, Gu und Lamb 2006) könnte eine weitere Erklärung dafür sein, warum das mittlere Alter beim ersten Artikel zugenommen hat. Durch die höhere Lebenserwartung verschieben bzw. verlängern sich bestimmte Lebensabschnitte (z. B. Schulzeit, Studium, Berufsleben), sodass darauffolgende später einsetzen (z. B. Meissner und Stockfisch 2012; Lee und Goldstein 2003).

Die Veröffentlichung eines Artikels wird durch das *Peer-Review-Verfahren* verzögert. Diese Verzögerung kann mitunter mehrere Jahre betragen (z. B. Ellison 2002, S. 951). Da das Peer-Review-Verfahren allerdings bereits etabliert ist und seit geraumer Zeit praktiziert wird, kann dieses nicht per se den Anstieg des Durchschnittsalters von Autoren zur Folge haben. Allerdings hat sich die Dauer des Peer-Review-Prozesses in den vergangenen Jahren erhöht (Ellison 2002), weil beispielsweise Gutachter eine gestiegene Zahl an Einreichungen prüfen müssen (z. B. Gad-el-Hak 2004).¹ Somit könnte auch diese Entwicklung einen Effekt auf das Alter von Autoren bei der Publikation eines Artikels haben.

Auch die gestiegene Anzahl der Quellenangaben in Artikeln kann neben der gestiegenen Menge des zu verarbeitenden Wissens weitere Ursachen haben, wobei die verschiedenen Effekte nur schwer auseinanderzuhalten sind. Beispielsweise tragen Fortschritte in Informations- und Kommunikationstechnologien zu dieser Entwicklung bei. Das Auffinden relevanter Arbeiten ist durch Online-Suchmaschinen wie *Google Scholar*, Online-Datenbanken wie EconLit, JSTOR und *Web of Science* oder diverse wissenschaftliche Plattformen wie *ResearchGate* vermeintlich einfacher geworden. So lassen sich zu einer gegebenen Arbeit bei *Google Scholar* oder dem *Web of Science* beispielsweise Arbeiten anzeigen, die diese zitieren. Auch eine personenbezogene Suche nach Literatur ist möglich, bei der man sich weitere Arbeiten der auf einer Arbeit angegebenen Autoren anzeigen lassen kann. Betrachtet man bei *Google Scholar* eine wissenschaftliche Arbeit, so bekommt man sogar Vorschläge für verwandte Arbeiten angezeigt. Ähnliches gilt zum Beispiel auch für die Plattform *ScienceDirect* des Verlags *Elsevier*, auf der ebenfalls Empfehlungen für verwandte Literatur gemacht werden. Neben dem Auffinden relevanter Arbeiten wird auch deren Verwaltung durch Literaturverwaltungsprogramme wie *Citavi* oder *Mendeley* erleichtert. Mit wenigen Klicks lassen sich gefundene Arbeiten und die zugehörigen Metadaten in die eigene Literaturdatenbank einpflegen, um sie bei Bedarf in wissenschaftlichen Arbeiten zu zitieren. Diese Programme erleichtern ferner die Erstellung von Literaturverzeichnissen. Neben der Literaturverwaltung ist *Mendeley* auch beim Auffinden relevanter bzw. verwandter Literatur behilflich und schlägt Nutzern auf Basis der Arbeiten in ihrer Datenbank potenziell interessante Artikel vor (Hoey 2015).

¹Beispielsweise hat sich die Anzahl der Einreichungen im *Review of Economic Studies* in den vergangenen zehn Jahren mehr als verdoppelt und ist von 700 auf 1.600 angestiegen, weshalb die Anzahl der Redakteure von vier auf sieben erhöht wurde (vgl. <http://www.restud.com/submissions/>; zuletzt besucht am 31.05.2018).

Weitere Ursachen für eine mit den Jahren gestiegene Anzahl der Quellenangaben sind systemimmanent und betreffen beispielsweise die bestehenden Maße zur Bewertung wissenschaftlicher Arbeiten. Ein Qualitätskriterium einer wissenschaftlichen Arbeit ist zum Beispiel wie häufig diese zitiert wird (Gross und Gross E. M. 1927; Garfield 1972; Nicolaisen 2002; Bornmann und Daniel 2008). Ähnliches gilt für die Relevanz eines Autors bzw. einer wissenschaftlichen Zeitschrift. Zum Beispiel wird das Lebenswerk eines Autors anhand des *h-Index* (vgl. Hirsch 2005) und die Relevanz einer Zeitschrift anhand des *Impact Factors* (Garfield 1955, 2006) bewertet, die wiederum beide auf der Anzahl der Zitationen wissenschaftlicher Artikel beruhen. Entsprechend bestehen für Wissenschaftler im Wettbewerb um Fördergelder, Drittmittel, Forschungsstellen sowie Professuren starke Anreize zu publizieren, zu zitieren und zitiert zu werden. Diese Anreize können unter Umständen verzerrende Effekte auf die wissenschaftliche Praxis haben, sodass gegebenenfalls nicht nur zum Zweck der Anerkennung wissenschaftlicher Leistung bzw. dem Beleg von Aussagen zitiert wird. Eine Übersicht und Diskussion möglicher anderer Motive findet sich zum Beispiel in Egghe und Rousseau (1990). So werden vermehrt auch eigene Arbeiten (*self-citation*) oder Arbeiten von Kollegen bzw. Koautoren (*co-citation*) zitiert (z. B. Phelan 1999; Lawrence 2007; Bethard und Jurafsky 2010; Lin und Huang 2012; Teodorescu und Andrei 2014). Simkin und Roychowdhury (1996) liefern empirische Evidenz dafür, dass Autoren teilweise nicht alle Artikel lesen, die sie in ihren Arbeiten zitieren. Eine Übersicht weiterer Studien, die Hinweise für solche Praktiken liefern, findet sich beispielsweise in Bornmann und Daniel (2008). Trotz der angesprochenen verzerrenden Effekte stützen sich zahlreiche Studien auf die Anzahl der Quellenangaben (z. B. Yitzhaki und Ben-Tamar 1991; Ucar u. a. 2013; Bornmann und Mutz 2015).

Ähnlich wie für die Anzahl der Quellenangaben ist der Anstieg der Teamgröße vermutlich nicht nur der steigenden Menge an Wissen und der zunehmenden Spezialisierung von Wissenschaftlern geschuldet, sondern auch den erleichterten Bedingungen für Kooperationen. So sind die Kosten der Zusammenarbeit von Autoren insgesamt vermeintlich gesunken. Informations- und Kommunikationstechnologien erleichtern den Austausch von Dokumenten, die gemeinsame Arbeit an Dokumenten sowie die Absprache und Diskussion in persönlichen Gesprächen. Ferner sind aufgrund der Globalisierung und technischer Fortschritte bei Fortbewegungsmitteln wie dem Auto oder dem Flugzeug auch die Reisekosten und -zeiten gesunken (z. B. Katz und Martin 1997, S. 8), was sich wiederum positiv auf geografisch entfernte Kooperationen auswirkt. Neben den vereinfachenden Rahmenbedingungen für die Zusammenarbeit von Wissenschaftlern bestehen außerdem Anreize zur Arbeit in Teams hinsichtlich der Maße, die

wissenschaftliche Produktivität und Relevanz messen. Bei gleichem Aufwand pro Artikel wird ein Autor, der seine Artikel stets in Teams verfasst, mehr Artikel publizieren, als wenn er diese alleine verfassen würde (Barnett, Ault und Kaserman 1988; Hollis 2001; Lee und Bozeman 2005). Ferner steigt unter Kollaboration auch sein *h-Index*, da er einerseits mehr Artikel veröffentlicht und andererseits diverse Studien zeigen, dass Artikel mit mehr Autoren häufiger zitiert werden (z. B. Wuchty, Jones und Uzzi 2007; Hirsch 2010; Larivière, Gingras, Sugimoto u. a. 2015; Mukherjee u. a. 2017). Teilweise könnte allerdings auch die zunehmende Teamgröße selbst zu ihrem eigenen Anstieg beitragen. Mit zunehmender Teamgröße steigt auch der Aufwand bei der Abstimmung bzw. Koordination innerhalb solcher Teams (Hudson 1996; Cummings und Kiesler 2007; Bercovitz und Feldman 2011), weshalb es gegebenenfalls zusätzlicher Teammitglieder wie eines Teamleiters bedarf, die die Koordination übernehmen.

Auch das Ergebnis der zunehmenden Spezialisierung muss gegebenenfalls relativiert werden. Die abnehmende relative Häufigkeit dafür, dass Wissenschaftler ihr Forschungsfeld wechseln und damit die zunehmende Spezialisierung könnten damit zusammenhängen, dass die einzelnen Forschungsfelder breiter werden. Somit kann es durchaus sein, dass Wissenschaftler weiterhin ihren Forschungsbereich zwischen Publikationen wechseln, allerdings innerhalb ihres Forschungsfeldes. Solche Wechsel werden aller Voraussicht nach nicht von aktuellen Klassifizierungssystemen wie dem der JEL-Codes (vgl. Abschnitt 3.1.2, S. 43) als solche erkannt, da diese die einzelnen Forschungsbereiche ggf. nicht mehr ausreichend detailliert klassifizieren. Wie in Abschnitt 3.1.2 (S. 43) bereits erwähnt und von Cherrier (2017) beschrieben, ist auch das JEL-Klassifikationssystem, das in dieser Arbeit zur Bemessung der Spezialisierung von Autoren verwendet wurde, kein starres und unveränderbares System. Es unterlag seit seiner Einführung bereits mehreren Überarbeitungen, die aufgrund der Entwicklungen in der Volkswirtschaftslehre notwendig waren, um das Feld und seine einzelnen Forschungsbereiche adäquat abzubilden.

6.2 Limitationen der vorliegenden Arbeit

Unausweichlich unterliegt die vorliegende Arbeit gewissen Limitationen. Diese betreffen unter anderem die zugrunde liegenden Daten. Bei der Aufbereitung des Datensatzes wurden einige Heuristiken angewandt, beispielsweise zur Identifikation von Personen aus den Autorennennungen (Abschnitt 3.4, S. 66) oder der Bestimmung des Geschlechts (Abschnitt 3.5, S. 71). Erstere wirkt sich unter anderem auch auf die Bestimmung des

ersten Artikels eines Autors aus. Beispielsweise könnte ein Artikel fälschlicherweise als erster Artikel eines Autors identifiziert werden, wenn der Autor im Laufe seiner Karriere seinen Namen geändert hat. Zwar umfassen die Publikationsdaten zur *Volkswirtschaftslehre* alle in EconLit verfügbaren Zeitschriften und damit eine umfassende Stichprobe volkswirtschaftlicher Artikel, allerdings ist für die Disziplinen *Betriebswirtschaftslehre und Finance*, *Geographie und Geologie*, *Mathematik* sowie *Psychologie* lediglich eine Auswahl renommierter Zeitschriften verfügbar. Entsprechend bezieht sich der erste Artikel eines Autors auf den ersten Artikel innerhalb dieses Datensatzes. Es ist nicht auszuschließen, dass der Autor bereits zuvor einen Artikel verfasst hat, der nicht im Datensatz enthalten ist.

Auch wenn diverse andere Studien ihre Untersuchung ebenfalls auf die renommiertesten Zeitschriften eines Fachbereichs einschränken (z. B. Card und DellaVigna 2013; Hamermesh 2015), ist diese Auswahl nicht ganz unkritisch. Da der *Impact Factor* von Zeitschriften heutzutage als Bewertungskriterium für darin veröffentlichte Artikel herangezogen wird (Starbuck 2005; Baum 2011), besteht ein intensiver Wettbewerb um Publikationen innerhalb der renommiertesten Zeitschriften. Entsprechend wird in diesen, nach Durchlaufen des Begutachtungsprozesses, eine Auswahl von Arbeiten höchster Qualität veröffentlicht, die lediglich einen kleinen und sehr speziellen Teil wissenschaftlicher Publikationen einer Disziplin darstellen. Insofern unterliegt die Untersuchung dieser Zeitschriften gewissen Verzerrungen, derer man sich bewusst sein sollte.

Wie bereits weiter oben angesprochen führt die Schaffung von Maßen zur Bewertung einzelner wissenschaftlicher Arbeiten, Autoren oder Zeitschriften (z. B. Anzahl Publikationen, *Impact Factor* oder *h-Index*) dazu, dass Wissenschaft gegebenenfalls aus den falschen Motiven betrieben wird. Im Vordergrund stehen möglicherweise nicht mehr der Erkenntnisgewinn und das Erreichen bzw. Überzeugen eines möglichst großen Publikums, sondern das Publizieren möglichst vieler Artikel in Zeitschriften mit möglichst hohen Rankings oder das Erreichen eines möglichst hohen *h-Index*. Dieses Phänomen ist in den Sozialwissenschaften auch als *Reaktivität* bekannt. In diesem Fall richten Wissenschaftler und ganze Institutionen ihre Forschungsinteressen und -arbeiten nach dem aus, was in den Top-Zeitschriften publiziert wird (Adler und Harzing 2009; Willmott

2011; Frey 2017) bzw. was die gängigen Erfolgsmaße beeinflusst.² Diese Anpassungen der Wissenschaftler und Institutionen an die bestehenden Maße beeinflussen wiederum die Maße selbst und deren Entwicklung. Insofern generieren diese Erfolgsmaße eine andere Form der Last und ein erhöhtes Maß an „Rauschen“, das sich in zunehmend vielen Publikationen, für die der Erkenntnisgewinn gegebenenfalls fraglich ist, und Zitationen, die nicht aus inhaltlichen Motiven gemacht werden, äußert. Aus den genannten Gründen wurde in der vorliegenden Arbeit zumindest für die Volkswirtschaftslehre neben den neun renommiertesten Zeitschriften auch eine deutlich breitere Auswahl an volkswirtschaftlichen Zeitschriften betrachtet. Die Ergebnisse der beiden Untersuchungsstränge decken sich weitestgehend und sprechen für die Robustheit der Ergebnisse. Der einzige größere Unterschied ergab sich für die Länge der Titel von Artikeln. Für die renommiertesten Zeitschriften sind die Titel im Beobachtungszeitraum nur unwesentlich länger geworden, während die Länge der Titel für die breitere Auswahl volkswirtschaftlicher Zeitschriften deutlich zugenommen hat.

Ein zentrales Ergebnis dieser Arbeit bezieht sich auf das Alter von Autoren bei ihrem ersten Artikel. Informationen zum Geburtsjahr sind allerdings lediglich für einen Teil der Autoren der Volkswirtschaftslehre und der Mathematik verfügbar. Entsprechend sind die Ergebnisse bzgl. dieses Maßes nicht auf die anderen untersuchten Disziplinen übertragbar. Es sind außerdem Verzerrungen hinsichtlich der Verfügbarkeit dieser Informationen denkbar. So könnten die Autoren, zu denen man das Geburtsjahr aus öffentlich zugänglichen Quellen ermitteln kann, eine Eigenschaft teilen, die sie von den übrigen Autoren unterscheidet.

Ferner weisen die Ergebnisse dieser Arbeit keine Kausalität zwischen der beschriebenen, möglichen Last des Wissens und den beobachteten Indikatoren nach. Vielmehr sind die Ergebnisse konsistent mit den Hypothesen über die Folgen einer solchen Last des Wissens. Auch wenn die vorliegende Arbeit und die entsprechenden Ergebnisse eine Last des Wissens nicht nachweisen können, ist zu vermuten, dass die Ergebnisse ähnlich aussehen würden, wenn diese Last tatsächlich existierte. Die Ergebnisse unterliegen ferner Endogenitätsproblemen. Diese bestehen darin, dass nicht für alle denkbaren Variablen kontrolliert wurde (*omitted variable bias*). Weiter oben wurde diskutiert, welche anderen Faktoren als die Zeit und die Menge des zu verarbeitenden Wissens die

²Interessant sind in diesem Zusammenhang zwei *Science*-Artikel: Bhattacharjee (2011) berichtet, dass saudi-arabische Universitäten ihr Ranking verbessern, indem sie renommierte Wissenschaftler (z. B. von Elite-Universitäten der USA) dafür bezahlen, dass diese eine Zugehörigkeit zu einer saudi-arabischen Universität auf ihren Artikeln angeben. Hvistendahl (2013) beschreibt, wie Autorenschaften in China verkauft werden. Dabei wird Wissenschaftlern angeboten sie gegen Geld als Autoren fertiger und zur Publikation freigegebener Artikel anzugeben.

gefundenen Trends beeinflussen könnten. Für die Anzahl der Quellenangaben bzw. die Teamgröße sind dies beispielsweise Fortschritte in Informations- und Kommunikationstechnologien. Diese könnten schwierig zu quantifizieren sein sowie selbst von der Zeit abhängen, was wiederum zu methodischen und inhaltlichen Schwierigkeiten führen könnte. Zudem sind für einzelne Modelle auch umgekehrte Kausalitäten denkbar. Beispielsweise wurde in den Modellen zum Trend für die Anzahl der Quellenangaben für die Teamgröße kontrolliert und ein positiver Zusammenhang festgestellt. Allerdings ist auch denkbar, dass die Wirkung umgekehrt ist und die Teamgröße durch die Anzahl der benötigten Quellen beeinflusst wird. Eine solche Perspektive wurde für die Teamgröße in Abschnitt 4.6 (S. 139) eingenommen und die entsprechenden Wechselwirkungen analysiert.

Die Anzahl der Quellenangaben wurde in der vorliegenden Arbeit als ein Maß für die Menge des Wissens, das in einem Artikel verarbeitet wurde, herangezogen. Sie stellt allerdings lediglich ein grobes Maß hierfür dar. Dieses Maß macht beispielsweise keine Unterschiede zwischen der Breite und der Tiefe des verarbeiteten Wissens oder der Qualität und dem Gehalt der zitierten Werke. Es bildet zum Beispiel nicht ab, ob die angegebenen Quellen konsultiert wurden, um eine Thematik möglichst breit abzudecken oder dazu, um einen spezifischen Aspekt zu vertiefen. Zusätzlich wird es, wie oben beschrieben, durch verschiedene Entwicklungen in der Wissenschaft verzerrt.

6.3 Ausblick und zukünftige Forschung

Die in der vorliegenden Arbeit untersuchten wissenschaftlichen Disziplinen bzw. die untersuchte Auswahl an Zeitschriften aus diesen, machen lediglich einen kleinen Teil der gesamten wissenschaftlichen Aktivität aus. Nichtsdestotrotz ähneln die beobachteten Trends beispielsweise denen aus dem Bereich der Patente (Jones 2009a). Sollten ähnliche Mechanismen auch in anderen Disziplinen (z. B. Biologie, Physik, Ingenieurwissenschaften und Medizin) greifen, könnten die Implikationen weitgreifend sein. Eine grundsätzliche Last des Wissens könnte nicht nur den Output wissenschaftlicher Arbeit beeinflussen, sondern auch kommerzielle Anwendungen und den Wohlstand einer Gesellschaft. Um eine grundlegende Existenz einer Last des Wissens zu untersuchen, erscheint die Untersuchung weiterer Disziplinen, speziell im Hinblick auf das Alter von Autoren, als ein Erfolg versprechendes Unterfangen für zukünftige Forschung.

Die beobachteten Trends werden in der vorliegenden Arbeit und in Jones (2009a) gemäß der Hypothesen bezüglich einer Last des Wissens vor allem mit der zunehmend großen Menge bestehenden Wissens erklärt, das Wissenschaftler verarbeiten müssen, um zum Wissenspool in ihrem Fachbereich beizutragen. Außer Acht gelassen wurde allerdings ein zweiter Aspekt, der sich auf die noch zu lösenden Probleme bezieht. Denn wenn Wissenschaftler einmal an der Wissensgrenze angelangt sind, sehen sie sich gegebenenfalls zunehmend schweren Problemen gegenüber. Dies folgt aus der Überlegung, dass wenn Wissenschaftler aus einer gegebenen Menge von Problemen wählen können, sie zunächst die Probleme angehen, die die höchste Rendite bei gegebenem Einsatz versprechen (Segerstrom 1998). Entsprechend könnte das Alter beim ersten Artikel auch ein weiteres Element widerspiegeln: nachdem Wissenschaftler die Grenze des bestehenden Wissens in ihrem Forschungsbereich erreichen und bevor sie ihren ersten Artikel veröffentlichen, müssen sie zunächst ein ungelöstes Problem lösen, um einen neuen Beitrag zur Wissenschaft zu leisten. Daher ist das Alter beim ersten Artikel gleichzeitig ein Indikator für die Dauer der Verarbeitung des nötigen Wissens *und* die Schwierigkeit der verbleibenden Grenzprobleme. Die beiden Effekte lassen sich bei den gegebenen Daten nur schwer trennen. Mögliche zukünftige Forschung könnte sich damit auseinandersetzen, diese beiden Effekte zu isolieren bzw. Maße und Indikatoren für die zunehmende Schwierigkeit von Grenzproblemen zu entwickeln. Die Ausrichtung dieser Forschung kann sowohl die Wissenschaft betreffen, aber auch wirtschaftliche Errungenschaften, wie die Entwicklung neuer Produkte.

Ein Maß für die zunehmenden Grenzprobleme könnte die Länge des Zeitraums zwischen zwei Publikationen sein. Unter der Annahme, dass ein Wissenschaftler mit seiner ersten Publikation auch ein ungelöstes Problem löst und sich somit an der Wissensgrenze befindet, stellt die Zeit zwischen dieser Publikation und seiner darauffolgenden die Zeit dar, die er für die Lösung des zweiten Problems benötigt. Ein Indikator aus der Wirtschaft könnten die Zeiträume zwischen Produkteinführungen sein. Bei der Entwicklung eines neuen Produkts wird häufig auf mehrere bestehende Patente zurückgegriffen. Folgt man der Argumentation von oben, verlängert sich der Zeitraum zwischen zwei Produkteinführungen, da Patente zunehmend viel Zeit benötigen und somit die Entwicklung des Produkts verlängern. Eine analoge Argumentation greift auch bei Software, bei der sich die Entwickler einer zunehmenden Menge an Code-Zeilen gegenübersehen oder mehrere bereits bestehende Software-Produkte nutzen und verbinden, die zusammen ein neues, aber noch komplexeres Produkt ergeben. Somit könnte man die Zeit zwischen zwei Produkteinführungen untersuchen, um die Schwierigkeit praxis- bzw. industrienaher Probleme zu messen.

In der vorliegenden Arbeit wurden Maße betrachtet, die zumeist direkt mögliche Folgen einer Last des Wissens identifizieren (z. B. spezifische Probleme oder Inputfaktoren). In Kapitel 5 wurden allgemeine Prinzipien und konkrete Ansätze zum Umgang mit der Last des Wissens vorgestellt. Die Methoden selbst könnten Ausgangspunkt für eine alternative Herangehensweise an die Messung der Last des Wissens sein. Anstatt die Probleme selbst zu messen, könnte man versuchen die Lösungsansätze, die typischerweise angewandt werden, um mit der Last des Wissens umzugehen, bzw. deren Einsatz zu quantifizieren. Denn die Lösungsansätze sind nicht nur aufgrund ihres Wertes für das Management wichtig. Sie reflektieren auch die Probleme zu deren Lösung sie eingeführt wurden. Beispielsweise könnte man beobachten, welche Methoden wie häufig eingesetzt werden, um daraus auf die zugrunde liegende Komplexität bzw. Schwierigkeit der Probleme zu schließen. Zukünftige Forschung könnte ferner untersuchen, inwiefern und in welcher Form Lösungen zwischen verschiedenen Einsatzbereichen übertragen und miteinander kombiniert werden können. Es wäre interessant zu verstehen, was die Kompromisse (engl. *trade-offs*) bei der Wahl zwischen den Prinzipien und den entsprechenden Methoden bzw. Techniken sind, welche Voraussetzungen erfüllt sein müssen, um die Prinzipien miteinander zu kombinieren und ob Heuristiken ein Element des Lösungsprozesses sein könnten, beispielsweise in der Form, dass man in einem *Trial-&-Error*-Ansatz zwischen den drei Prinzipien wechselt. Ein weiterer Schritt zur Fortführung dieser Forschung könnte die Sammlung von Fallstudien sein, die detailliertere Beispiele für die Lösung größerer Probleme und den Einsatz der identifizierten Prinzipien darstellen.

Ferner könnte es interessant sein zu untersuchen, inwiefern bestimmte Entwicklungen des menschlichen Fortschritts die Last des Wissens reduzieren können. Beispielsweise könnten eine erhöhte Lebenserwartung und ein gesteigertes geistiges Leistungsvermögen im Alter der zunehmenden Menge an zu erlernendem Wissen entgegenwirken. Auch das Lernen selbst hat sich in den letzten Jahren, unter anderem durch Fortschritte im Bereich der Informations- und Kommunikationstechnologien, verändert. Wissen ist durch digitale Datenbanken und Verbesserungen im Bereich der Suchmaschinen zugänglicher geworden. Online-Kurse (z. B. sogenannte MOOCs, ein Akronym für Massive Open Online Courses) bieten Bildung höchster Qualität, die zu geringen Kosten, in einem individualisierten Tempo und zu ausgewählten Themen in Anspruch genommen werden kann. Zusammengenommen vereinfachen diese Entwicklungen die Aufnahme von Wissen. Darüber hinaus bieten neue Technologien verbesserte Möglichkeiten größere Aufgaben zu modularisieren und in größeren Teams zu verteilen.

Wie in der Einleitung dieser Arbeit bereits angedeutet, ist eine bildungsorientierte und forschungsstarke Gesellschaft eine zwingende Voraussetzung für ein innovatives wirtschaftliches Umfeld. Dabei gab es in den vergangenen Jahren bildungspolitische Bestrebungen zu einer (internationalen) Vereinheitlichung von Bildungsangeboten und -abschlüssen, wie der Umstellung auf Bachelor- und Master-Studiengänge in Deutschland. Die Schaffung von Brückenstudiengängen (z. B. Wirtschaftsingenieurwesen) ermöglicht ferner die Verknüpfung zunehmend spezialisierter Fachbereiche. Bildung und Forschungsergebnisse alleine reichen allerdings noch nicht aus Innovationen hervorzubringen und so Fortschritt zu generieren. Bercovitz und Feldmann (2006) beschreiben die Rolle von Universitäten im Innovationsprozess. Demnach unterliegen Universitäten bei der Wissensschaffung sowie dem Wissenstransfer und damit dem Beitrag zu Wirtschaftswachstum den Einflüssen der Gesellschaft, Industrie und Politik. Dabei spielt der Transfer von Forschungsergebnissen eine zentrale Rolle für wirtschaftliches Wachstum (z. B. Mueller 2006). Furman und Stern (2011) liefern Anhaltspunkte dafür, wie Institutionen den Transfer von Wissen vereinfachen und dabei eine praktische Rolle spielen können. Bei der Schaffung von Innovationen und Anreizen zu innovativem Verhalten sind also Austauschprozesse sowie das Zusammenspiel zwischen Wissenschaft, Industrie und Politik wichtig und notwendig. Unterschiedliche Zielsetzungen und eine fehlende Abstimmung können den Transfer von Wissen in neue Produkte, Wirtschaftswachstum sowie einen höheren Lebensstandard erschweren oder gänzlich verhindern.

Eines der Ergebnisse der vorliegenden Arbeit ist, dass die Teamgröße von Artikel zugenommen hat. Es sind also immer mehr Personen gemeinsam am Lösen eines spezifischen Problems beschäftigt. Ungeachtet der möglichen Gründe für diese Entwicklung, die weiter oben bereits diskutiert wurden, scheint unter Wissenschaftlern eine Notwendigkeit zur Zusammenarbeit zu bestehen. Entsprechend sollten zukünftige wissenschaftspolitische Entscheidungen Möglichkeiten zur Kooperationen weiter verbessern. Einen vielversprechenden Forschungsstrang stellt in diesem Zusammenhang die Teamzusammensetzung dar. Fruchtbar könnte die weitere Erforschung der Zusammenstellung von Teams sein und inwiefern sich verschiedene Konstellationen von Personen in einem Team auf dessen Output auswirken. Die Zusammenstellung eines Teams geht einher mit der Heterogenität von Teams und kann beispielsweise das Alter, das Geschlecht, den Forschungsbereich, den kulturellen Hintergrund oder auch die Geographie der einzelnen Teammitglieder betreffen.

In der Literatur finden sich bereits diverse Arbeiten zu den Vor- und Nachteilen heterogener bzw. homogener Teams, die sich vielfach auf Teams im industriellen

Kontext beziehen und die zu unterschiedlichen Ergebnissen kommen. Beispielsweise untersuchen Melero und Palomeras (2015), wie die Zusammensetzung von Teams von Erfindern den Output in Form von Patenten beeinflusst. Ihr Untersuchungsschwerpunkt liegt dabei auf der Spezialisierung der Autoren und der Frage, wie Generalisten, das heißt Individuen mit etwas breiteren Interessens- und Tätigkeitsgebieten, den Erfolg eines Teams beeinflussen. Arbeiten, wie Ucbasaran u. a. (2003) oder Meves (2013) befürworten Heterogenität, da sich die Teammitglieder in diesem Fall hinsichtlich ihrer Kompetenzen, Persönlichkeiten (z. B. hinsichtlich der Risikofreudigkeit) und Wertvorstellungen ergänzen. Ruef, Aldrich und Carter (2003) argumentieren, dass auf der sozialen Ebene in einem Team Homogenität zu bevorzugen ist, da es seltener zu Missverständnissen kommt und Differenzen leichter zu überwinden sind. Ähnlich argumentieren auch Shepperd (1993) sowie Amason und Sapienza (1997), die zwar einräumen, dass größere Teams Zugang zu mehr Ressourcen (z. B. Zeit, Geld oder Know-how) haben und damit leichter schwierigere Aufgaben lösen können. Allerdings betonen sie, dass es leichter zu Koordinationsproblemen und, aufgrund der Verteilung von Verantwortung auf mehrere Personen, zum Motivationsverlust kommen kann. Analoge Untersuchungen für Teams aus dem Bereich der wissenschaftlichen Forschung und wissenschaftlicher Publikationen könnten Aufschluss darüber geben, ob solche Beobachtungen verallgemeinert werden können.

In der Einleitung zur vorliegenden Arbeit und in Abschnitt 2.2 wurde deutlich, dass die Frage, ob sich Wachstum und Fortschritt im Allgemeinen bzw. das Wachstum von Wissen im Speziellen beschleunigt oder verlangsamt umstritten und schwierig zu greifen ist. Ein Grund dafür könnte das Fehlen verlässlicher direkter Maße sein. Entsprechend betrifft ein weiterer interessanter Bereich für zukünftige Forschung alternative Maße zur Messung des Wachstums von Wissen, die beispielsweise auch die Dimensionen der Breite und Tiefe von Wissen abbilden. Bestehenden Maßen, die beispielsweise die Anzahl wissenschaftlicher Werke (z. B. Artikel oder Bücher) zählen oder die Anzahl der Quellenangaben betrachten, fehlt diese Tiefe der Betrachtung. Das Verständnis über die Entwicklung des Wissens in Tiefe und Breite könnte helfen zukünftige Herausforderungen in der Wissenschaft zu erkennen und Lösungsansätze zu entwickeln. Auch fehlen bislang Maße für den tatsächlichen Beitrag eines Artikels. Dieser wird derzeit beispielsweise über die Relevanz des Beitrags gemessen, die wiederum aus der Anzahl der Zitationen oder dem Ranking der Zeitschrift, in der ein Artikel veröffentlicht wurde, abgeleitet wird.

In der bisherigen Debatte zu Wachstum und Fortschritt (vgl. Kapitel 1 und Abschnitt 2.2) scheinen sowohl die pessimistische, als auch die optimistische Sicht durch

empirische Evidenz und valide Argumente gestützt zu sein. Diese Unsicherheit bezüglich der tatsächlichen Entwicklung erschwert Vorhersagen bezüglich zukünftiger technologischer Fortschritte und längerfristige Planung. Diverse Wachstumsprozesse folgen der Form einer S-Kurve, die den Verlauf logistischen Wachstums bzw. exponentieller Wachstumsprozesse in Systemen mit Ressourcenbeschränkungen beschreibt. Dabei beginnt der Prozess typischerweise mit einer Periode langsamen Wachstums bzw. Fortschritts, auf die eine Periode folgt, in der Wachstum bzw. Fortschritt mit zunehmender Geschwindigkeit geschieht. Allerdings greifen im Verlauf dieses Prozesses bestimmte Restriktionen zunehmend und resultieren in einer Abflachung der Kurve. Tatsächlich werden solche oder ähnliche Prozesse in vielen Kontexten beschrieben und beobachtet. Hierzu gehören Beispiele aus den Naturwissenschaften, wie das Wachstum von Populationen von Lebewesen in natürlichen Habitaten (Cook 1965; Zwietering u. a. 1990; Cohen 1995). Aber auch in wirtschaftswissenschaftlichen Kontexten wird ein solcher Wachstumsverlauf beobachtet und modelliert. Beispielsweise für die Adaption von Innovationen (Griliches 1957) oder die Akzeptanz von und Nachfrage nach neuen Produkten (Bass 1969; Mahajan und Muller 1979; Aoki und Yoshikawa 2002).

Vor dem Hintergrund eines bis dato starken Wachstums der Wissenschaft, prognostiziert Price (1963) langfristig einen logistischen Wachstumsverlauf und begründet: „we should otherwise reach absurd conditions at the end of another century“ (S. 21). Ein charmantes Element logistischer Wachstumsprozesse ist, dass sie sowohl die optimistische Sichtweise, die von einer zunehmenden Geschwindigkeit des Wachstums bzw. Fortschritts ausgeht als auch die pessimistische Sichtweise, die eine Verlangsamung des Wachstums bzw. Fortschritts erwartet, erklärt. Diese beiden Situationen entsprechen unterschiedlichen Bereichen der Kurve, nämlich der ersten bzw. zweiten Hälfte. Der tatsächliche Lauf technischen und wissenschaftlichen Fortschritts könnte sogar weniger ideal sein. Beispielsweise könnte die S-Kurve des technologischen Gesamtfortschritts aus mehreren S-Kurven einzelner Technologien bestehen. Wie Christensen (1997) für disruptive Innovationen argumentiert, könnte eine neue Kurve hinsichtlich der Gesamtleistung sogar unterhalb einer bestehenden Kurve beginnen und letztere erst im Zeitverlauf überholen. Eine solche Vorstellung vom Gesamtfortschritt der Wissenschaft könnte eine Erklärung dafür sein, warum man sowohl Beispiele für zunehmende als auch für abnehmende Geschwindigkeit des Wachstums bzw. Fortschritts findet.

Literaturverzeichnis

- AAAS (2017). *Science Media KIT 2017*. Techn. Ber., S. 1–20.
- Ackoff, Russel L. (1989). »From data to wisdom«. In: *Journal of Applied Systems Analysis* 16.1, S. 3–9.
- Adams, James D., Grant C. Black, J. Roger Clemmons und Paula E. Stephan (2005). »Scientific teams and institutional collaborations: Evidence from U.S. universities, 1981-1999«. In: *Research Policy* 34.3, S. 259–285.
- Adler, Nancy J. und Anne-Wil Harzing (2009). »When knowledge wins: Transcending the sense and nonsense of academic rankings«. In: *Academy of Management Learning & Education* 8.1, S. 72–95.
- Aghion, Philip und Peter Howitt (1992). »A model of growth through creative destruction«. In: *Econometrica* 60.2, S. 323–351.
- Aldermann, Neil (2001). »Distributed knowledge in complex engineering project networks: Implications for regional innovation systems«. In: *Knowledge, complexity and innovation systems*. Springer Verlag, S. 209–223.
- Allen, I. Elaine und Jeff Seaman (2006). *Growing by degrees: Online education in the United States, 2005*. Techn. Ber. Sloan Consortium (NJ1).
- Altshuller, Genrich (1994). *And suddenly the inventor appeared: TRIZ, the theory of inventive problem solving*. Worcester, Massachusetts: Technical Innovation Center, Inc.
- Amason, Allen C. und Harry J. Sapienza (1997). »The effects of top management team size and interaction norms on cognitive and affective conflict«. In: *Journal of Management* 23.4, S. 495–516.
- Aoki, Masanao und Hiroshi Yoshikawa (2002). »Demand saturation-creation and economic growth«. In: *Journal of Economic Behavior & Organization* 48.2, S. 127–154.
- Aral, Sinan, Erik Brynjolfsson und Marshall Van Alstyne (2012). »Information, technology, and information worker productivity«. In: *Information Systems Research* 23.3, S. 849–867.
- Arbesman, Samuel (2011). »Quantifying the ease of scientific discovery«. In: *Scientometrics* 86.2, S. 245–250.
- Archibald, G. und M. B. Line (1991). »The size and growth of serial literature 1950-1987, in terms of the number of articles per serial«. In: *Scientometrics* 20.1, S. 173–196.
- Ash, Mitchell, Martin Carrier, Olaf Dössel, Ute Frevert, Siegfried Großmann, Martin Grötschel, Reinhold Kliegl, Alexander Peukert, Hans-Jörg Rheinberger, Eberhard

- Schmidt-Aßmann, Uwe Schimank, Volker Stollorz, Niels Taubert und Peter Weingart (2015). *Empfehlungen zur Zukunft des wissenschaftlichen Publikationssystems*. Techn. Ber. Berlin-Brandenburgische Akademie der Wissenschaften.
- Atallah, Mikhail J., Richard Cole und Michael T. Goodrich (1989). »Cascading divide-and-conquer: A technique for designing parallel algorithms«. In: *SIAM Journal on Computing* 18.3, S. 499–532.
- Atrostic, B. K. (1982). »The demand for leisure and nonpecuniary job characteristics«. In: *The American Economic Review* 72.3, S. 428–440.
- Audretsch, David B. und Max Keilbach (2008). »Resolving the knowledge paradox: Knowledge-spillover entrepreneurship and economic growth«. In: *Research Policy* 37, S. 1697–1705.
- Ayres, Robert (1996). »Technology, progress and economic growth«. In: *European Management Journal* 14.6, S. 562–575.
- Bak, Peter Michael (2016). *Wie man Psychologie als empirische Wissenschaft betreibt*. Wiesbaden: Springer.
- Ball, Rafael (2009). »Wissenschaftskommunikation im Wandel - Bibliotheken sind mitten drin«. In: *Zeitschrift für Bibliothekswesen und Bibliografie* 96, S. 39–54.
- Barnett, A. H., R. W. Ault und D. L. Kaserman (1988). »The rising incidence of co-authorship in economics: Further evidence«. In: *The Review of Economics and Statistics* 70.3, S. 539–543.
- Bass, Frank M. (1969). »A new product growth for model consumer durables«. In: *Management Science* 15.5, S. 215–227.
- Bates, Marcia J. (2005). »Information and knowledge: An evolutionary framework for information science«. In: *Information Research* 10.4.
- Baum, Joel A. C. (2011). »Free-riding on power laws: Questioning the validity of the impact factor as a measure of research quality in organization studies«. In: *Organization* 18.4, S. 449–466.
- Beaudry, Catherine und Vincent Larivière (2016). »Which gender gap? Factors affecting researchers' scientific impact in science and medicine«. In: *Research Policy* 45.9, S. 1790–1817.
- Becker, Gary S. (1962). »Investment in human capital: A theoretical analysis«. In: *Journal of Political Economy* 70.5, S. 9–49.
- Becker, Gary S. und Kevin M. Murphy (1992). »The division of labor, coordination costs, and knowledge«. In: *The Quarterly Journal of Economics* 107.4, S. 1137–1160.
- Bercovitz, Janet und Maryann Feldman (2011). »The mechanisms of collaboration in inventive teams: Composition, social networks, and geography«. In: *Research Policy* 40.1, S. 81–93.
- Bercovitz, Janet und Maryann Feldmann (2006). »Entrepreneurial universities and technology transfer: A conceptual framework for understanding knowledge-based economic development«. In: *Journal of Technology Transfer* 31, S. 175–188.
- Bethard, Steven und Dan Jurafsky (2010). »Who should I cite? Learning literature search models from citation behavior«. In: *Proceedings of the 19th ACM International Conference on Information and Knowledge Management*, S. 609–617.

- Bhattacharjee, Yudhijit (2011). »Saudi universities offer cash in exchange for academic prestige«. In: *Science* 334.6061, S. 1344–1345.
- Bikhchandani, Sushil (1992). »A bargaining model with incomplete information«. In: *The Review of Economic Studies* 59.1, S. 187–203.
- Bird, P. R. und M. A. Knight (1975). »Word count statistics of the titles of scientific papers«. In: *The Information Scientist* 9.2, S. 67–69.
- Bloom, Nicholas, Charles Jones, John Van Reenen und Michael Webb (2017). *Are ideas getting harder to find?* URL: <http://www.nber.org/papers/w23782.pdf> (besucht am 30. 11. 2017).
- Bobtcheff, Catherine, Jérôme Bolte und Thomas Mariotti (2017). »Researcher's dilemma«. In: *Review of Economic Studies* 84.3, S. 969–1014.
- Bodendorf, Freimut (2006). *Daten- und Wissensmanagement*. Berlin Heidelberg: Springer-Verlag, S. 1–223.
- Bornmann, Lutz und Hans-Dieter Daniel (2007). »Multiple publication on a single research study: Does it pay? The influence of numbers of research articles on total citation counts in biomedicine«. In: *Journal of the American Society for Information Science and Technology* 58.8, S. 1100–1107. arXiv: 0803.1716.
- Bornmann, Lutz und Hans-Dieter Daniel (2008). »What do citation counts measure? A review of studies on citing behavior«. In: *Journal of Documentation* 64.1, S. 45–80.
- Bornmann, Lutz und R Mutz (2015). »Growth rates of modern science: A bibliometric analysis based on the number of publications and cited references«. In: *Journal of the Association for Information Science and Technology* 66.11, S. 2215–2222.
- Boschini, Anne und Anna Sjögren (2007). »Is team formation gender neutral? Evidence from coauthorship patterns«. In: *Journal of Labor Economics* 25.2, S. 325–365.
- Broad, William J. (1981). »The publishing game: Getting more for less«. In: *Science* 211.4487, S. 1137–1140.
- Brockhaus (1999). *Wissenschaft*. Leipzig Mannheim: F. A. Brockhaus.
- Brockhoff, Klaus (1993). »Technologiemanagement - Das S-Kurven Konzept«. In: *Zu einer Realtheorie der Unternehmung*. Stuttgart, S. 327–354.
- Brynjolfsson, Erik (1993). »The productivity paradox of information technology«. In: *Communications of the ACM* 36.12, S. 66–77.
- Brynjolfsson, Erik und Andrew McAfee (2011). *Race against the machine: How the digital revolution is accelerating innovation, driving productivity, and irreversibly transforming employment and the economy*. Lexington, Massachusetts: Digital Frontier Press.
- Buchmueller, Thomas C., Jeff Dominitz und W. Lee Hansen (1999). »Graduate training and the early career productivity of Ph. D. economists«. In: *Economics of Education Review* 14, S. 65–77.
- Burt, Ronald S. (2004). »Structural holes and good ideas«. In: *American Journal of Sociology* 110.2, S. 349–399.
- Buxton, A. B. (1987). »Title revisited«. In: *Journal of Documentation* 43.1, S. 65–68.
- Buxton, A. B. und A. J. Meadows (1977). »The variation in the information content of titles of research papers with time and discipline«. In: *Journal of Documentation* 33.1, S. 46–52.

- Callaghan, William M. und Cynthia J. Berg (2003). »Pregnancy-related mortality among women aged 35 years and older, United States, 1991-1997«. In: *Obstetrics & Gynecology* 102.5, S. 1015–1021.
- Card, David und Stefano DellaVigna (2012). *Nine facts about top journals in economics (online appendix)*. URL: https://assets.aeaweb.org/assets/production/articles-attachments/jel/app/51%7B%5C_%7D1%7B%5C_%7DNineFacts%7B%5C_%7Dapp.pdf (besucht am 19. 03. 2018).
- Card, David und Stefano DellaVigna (2013). »Nine Facts about Top Journals in Economics«. In: *Journal of Economic Literature* 51.1, S. 144–161.
- Castelvecchi, Davide (2015). *Physics paper sets record with more than 5,000 authors*. URL: <http://www.nature.com/doifinder/10.1038/nature.2015.17567> (besucht am 23. 05. 2018).
- Cherrier, Beatrice (2017). »Classifying economics: A history of the JEL codes«. In: *Journal of Economic Literature* 55.2, S. 545–579.
- Christensen, Clayton M. (1997). *The innovator's dilemma*. Harvard Business Review Press.
- Cohen, Joel E. (1995). »Population growth and earth's human carrying capacity«. In: *Science* 269.5222, S. 341–346.
- Colussi, Tommaso (2018). »Social ties in academia: A friend is a treasure«. In: *The Review of Economics and Statistics* 100.1, S. 45–50.
- Conley, John P., Mario J. Crucini, Robert A. Driskill und Ali Sina Önder (2013). »The effects of publication lags on life-cycle research productivity in economics«. In: *Economic Inquiry* 51.2, S. 1251–1276.
- Conley, John P. und Ali Sina Önder (2014). »The research productivity of new PhDs in economics: The surprisingly high non-success of the successful«. In: *Journal of Economic Perspectives* 28.3, S. 205–216.
- Cook, L. M. (1965). »Oscillations in the simple logistic growth model«. In: *Nature* 207, S. 316.
- Cordero, Radames J.B., Carlos M. De León-Rodríguez, John K. Alvarado-Torres, Ana R. Rodríguez und Arturo Casadevall (2016). »Life science's average publishable unit (APU) has increased over the past two decades«. In: *PLoS ONE* 11.6, S. 1–14.
- Cronin, Blaise (2001). »Hyperauthorship: A postmodern perversion or evidence of a structural shift in scholarly communication practices?«. In: *Journal of the American Society for Information Science and Technology* 52.7, S. 558–569.
- Cronin, Blaise, Debora Shaw und Kathryn La Barre (2003). »A cast of thousands: Coauthorship and subauthorship collaboration in the 20th century as manifested in the scholarly journal literature of psychology and philosophy«. In: *Journal of the American Society for Information Science and Technology* 54.9, S. 855–871.
- Crossland, M. D., B. E. Wynne und W. C. Perkins (1995). »Spatial decision support systems: An overview of technology and a test of efficacy«. In: *Decision Support Systems* 14.3, S. 219–235.
- Cummings, Jonathon N. und Sara Kiesler (2007). »Coordination costs and project outcomes in multi-university collaborations«. In: *Research Policy* 36.10, S. 1620–1634.

- Davenport, Thomas H. (2015). *Whatever happened to knowledge management?* URL: <https://blogs.wsj.com/cio/2015/06/24/whatever-happened-to-knowledge-management/> (besucht am 24. 08. 2017).
- Davenport, Thomas H. und Laurence Prusak (1998). *Working knowledge: How organizations manage what they know*. Harvard Business School Press.
- Depken, Craig A und Michael R Ward (2009). »Sited, Sighted, and Cited: The Effect of JSTOR in Economic Research«.
- Deutsche Forschungsgesellschaft (2013). *Sicherung guter wissenschaftlicher Praxis*. Weinheim: Wiley-VCH Verlag.
- Diener, Richard A. V. (1984). »Informational dynamics of journal article titles«. In: *Journal of the American Society for Information Science* 35.4, S. 222–227.
- Ding, Waverly W., Sharon G. Levin, Paula E. Stephan und Anne E. Winkler (2010). »The impact of information technology on academic scientists' productivity and collaboration patterns«. In: *Management Science* 56.9, S. 1439–1461.
- Diodato, Virgil (1982). »The occurrence of title words in parts of research papers: variations among disciplines«. In: *Journal of Documentation* 38.3, S. 192–206.
- Drucker, Peter F. und Joseph A. Maclariello (2008). *Management (Revised Edition)*. HarperCollins Pub., S. 38–55.
- Economist (2013). *Has the ideas machine broken down?* URL: <https://www.economist.com/news/briefing/21569381-idea-innovation-and-new-technology-have-stopped-driving-growth-getting-increasing> (besucht am 01. 09. 2017).
- Egghe, L. und R. Rousseau (1990). »Citation analysis«. In: *Introduction to informetrics. Quantitative methods in library, documentation and information science*. Amsterdam: Elsevier, S. 203–290.
- Ellis, C. A., S. J. Gibbs und G. Rein (1991). »Groupware: Some issues and experiences«. In: *Communications of the ACM* 34.1, S. 39–58.
- Ellison, Glenn (2002). »The Slowdown of the Economics Publishing Process«. In: *Journal of Political Economy* 110.5, S. 947–993.
- Ellison, Glenn (2011). »Is peer review in decline?« In: *Economic Inquiry* 49.3, S. 635–657.
- Engemann, Kristie M. und Howard J. Wall (2009). »A journal ranking for the ambitious economist«. In: *Federal Reserve Bank of St. Louis Review* 91, S. 127–140.
- Er, M. C. (1988). »Decision support systems: A summary, problems, and future trends«. In: *Decision Support Systems* 4.3, S. 355–363.
- Esen, Sinan, Hakan Tunahan und Davut Takil (2016). »A comparative bibliometric analysis of finance papers published in high impact journals and developing country-addressed journals: The case of turkish journals«. In: *Journal of Economics, Finance and Accounting* 3.1, S. 37–37.
- Evenson, R. E. (1984). »International invention: Implications for technology market analysis«. In: *R & D, Patents, and Productivity*. University of Chicago Press, S. 89–126.
- Evenson, R. E. (1993). »Patents, R & D, and invention potential: International evidence«. In: *The American Economic Review* 83.2, S. 463–468.

- Fernández-Cano, Antonio, Manuel Torralbo und Mónica Vallejo (2004). »Reconsidering Price's model of scientific growth: An overview«. In: *Scientometrics* 61.3, S. 301–321.
- Fletcher, J. und Sigmund Tobias (2005). »The multimedia principle«. In: *The Cambridge Handbook of Multimedia Learning*. Cambridge University Press, S. 117–133.
- Franceschet, Massimo und Antonio Costantini (2010). »The effect of scholar collaboration on impact and quality of academic papers«. In: *Journal of Informetrics* 4.4, S. 540–553.
- Franzoni, Chiara und Henry Sauermann (2014). »Crowd science: The organization of scientific research in open collaborative projects«. In: *Research Policy* 43.1, S. 1–20.
- Franzoni, Chiara, Giuseppe Scellato und Paula Stephan (2011). »Changing incentives to publish«. In: *Science* 333.6043, S. 702–703.
- Freeman, Richard B., Ina Ganguli und Raviv Murciano-Goroff (2014). *Why and Wherefore of Increased Scientific Collaboration*. URL: <http://papers.nber.org/tmp/99857-w19819.pdf>.
- Fretts, Ruth C., Julie Schmittiel, Frances H. McLean, Robert H. Usher und Marlene B. Goldman (1995). »Increased maternal age and the risk of fetal death«. In: *New England Journal of Medicine* 333.15, S. 953–957.
- Frey, Bruno S. (2017). *Reactivity in Economic Science*. URL: <http://www.crema-research.ch/papers/2017-10.pdf> (besucht am 02.05.2018).
- Furman, Jeffrey L. und Scott Stern (2011). »Climbing atop the shoulders of giants: The impact of institutions on cumulative research«. In: *American Economic Review* 101.5, S. 1933–1963.
- Gad-el-Hak, Mohamed (2004). »Publish or perish - An ailing enterprise?«. In: *Physics Today* 57.3, S. 61–62.
- Galor, Oded und Daniel Tsiddon (1997). »Progress, mobility, and economic growth«. In: *The American Economic Review* 87.3, S. 363–382.
- Garfield, E (1972). »Citation analysis as a tool in journal evaluation.« In: *Science (New York, N.Y.)* 178.60, S. 471–479. arXiv: 05218657199780521865715.
- Garfield, Eugene (1955). »Citation indexes for science: A new dimension in documentation through association of ideas«. In: *Science* 122.3159, S. 108–111.
- Garfield, Eugene (2006). »The history and meaning of the journal impact factor«. In: *Journal of the American Medical Association* 295.1, S. 90–93.
- Gerbert, Philipp, Christoph Gauger und Sebastian Steinhäuser (2015). »The double game of digital strategy«. In: *BCG Perspectives by the Boston Consulting Group*.
- Glänzel, Wolfgang und Andrés Schubert (2005). »Analysing scientific networks through co-authorship«. In: *Handbook of Quantitative Science and Technology Research*. Kluwer Academic Publishers: Dordrecht, S. 255–276.
- Gonzalez-Brambila, Claudia und Francisco M. Veloso (2007). »The determinants of research output and impact: A study of mexican researchers«. In: *Research Policy* 36.7, S. 1035–1051.
- Gordon, Robert J. (1999). »U. S. economic growth since 1870: One big wave?«. In: *The American Economic Review* 89.2, S. 123–128.

- Gordon, Robert J. (2012). *Is U.S. economic growth over? Faltering innovation confronts the six headwinds*. URL: <http://www.nber.org/papers/w18315.pdf> (besucht am 24.04.2018).
- Grant, Robert M. (1996). »Toward a knowledge-based theory of the firm«. In: *Strategic Management Journal* 17, S. 109–122.
- Greene, Mott (2007). »The demise of the lone author«. In: *Nature* 450.20, S. 1165.
- Griliches, Zvi (1957). »Hybrid corn: An exploration in the economics of technological change«. In: *Econometrica* 25.4, S. 501.
- Griliches, Zvi (1989). »Patents: Recent trends and puzzles«. In: *Brookings Papers: Microeconomics*, S. 291–330.
- Griliches, Zvi (1990). »Patent statistics as economic indicators: A survey«. In: *Journal of Economic Literature* 28.4, S. 1661–1707.
- Gross, P. L. K. und Gross E. M. (1927). »College libraries and chemical education«. In: *Science* 66.1713, S. 385–389.
- Hall, Bronwyn H. (2005). »Exploring the patent explosion«. In: *Journal of Technology Transfer* 30.1/2, S. 35–48.
- Hall, Bronwyn H., Adam B. Jaffe und Manuel Trajtenberg (2001). *The NBER patent citation data file: Lessons, insights and methodological tools*. URL: <http://www.nber.org/papers/w8498.pdf> (besucht am 08.02.2018).
- Hamermesh, Daniel S. (2013). »Six decades of top economics publishing: Who and how?«. In: *Journal of Economic Literature* 51.1, S. 162–172.
- Hamermesh, Daniel S. (2015). *Age, cohort and co-authorship*. URL: <http://www.nber.org/papers/w20938> (besucht am 27.03.2018).
- Hamermesh, Daniel S. und Sharon M. Oster (2002). »Tools or toys? The impact of high technology on scholarly productivity«. In: *Economic Inquiry* 40.4, S. 539–555.
- Hargadon, Andrew und Robert Sutton (1997). »Technology brokering and innovation in a product development firm«. In: *Administrative Science Quarterly* 42.4, S. 716–749.
- Hayne, Stephen C., C. A. P. Smith und Leo R. Vijayasarathy (2003). »Who wins on eBay: An analysis of bidders and their bid behaviours«. In: *Electronic Markets* 13.4, S. 282–293.
- Hennessy, John L. und Norman P. Jouppi (1991). »Computer technology and architecture: An evolving interaction«. In: *Computer* 24.9, S. 18–29.
- Henriksen, Dorte (2016). »The rise in co-authorship in the social sciences (1980–2013)«. In: *Scientometrics* 107.2, S. 455–476.
- Hesse, Hermann (1970). *Gesammelte Werke, Band 9, Das Glasperlenspiel*. Frankfurt am Main: Suhrkamp.
- Hevner, Alan und Samir Chatterjee (2010). *Design research in information systems*. Springer Publishing: New York.
- Hilbert, Martin und Priscilla López (2011). »The World ' s Technological Capacity«. In: *Science* 332.April, S. 60–65.
- Hirsch, J. E. (2005). »An index to quantify an individual's scientific research output«. In: *Proceedings of the National Academy of Sciences of the United States of America* 102.46, S. 16569–16572. arXiv: 0509048 [physics].

- Hirsch, J. E. (2010). »An index to quantify an individual's scientific research output that takes into account the effect of multiple coauthorship«. In: *Scientometrics* 85.3, S. 741–754. arXiv: 0911.3144.
- Hitt, Lorin und Erik Brynjolfsson (1996). »Productivity , Profit and Consumer Welfare : Three Different Measures of Information Technology ' s Value«. In: *MIS Quarterly* June, S. 1–20.
- Hoey, Sarah (2015). *New research features on mendeley.com!* URL: <https://blog.mendeley.com/2015/11/03/new-research-features-on-mendeley-com/> (besucht am 04.05.2018).
- Hollis, Aidan (2001). »Co-authorship and the output of academic economists«. In: *Labour Economics* 8.4, S. 503–530.
- Holmstrom, Bengt (1982). »Moral hazard in teams«. In: *The Bell Journal of Economics* 13.2, S. 324–340.
- Hornbostel, Stefan, Bernd Klingsporn und Markus von Ins (2009). »Messung von Forschungsleistungen – eine Vermessenheit ?« In: *Publikationsverhalten in unterschiedlichen wissenschaftlichen Disziplinen - Beiträge zur Beurteilung von Forschungsleistungen*. Alexander von Humboldt Stiftung, S. 14–35.
- Horowitz, Ellis und Alessandro Zorat (1983). »Divide-and-conquer for parallel processing«. In: *IEEE Transactions on Computers* 32.6, S. 582–585.
- Howard, Paul G. (1996). »Lossless and lossy compression of text images by soft pattern matching«. In: *Proceedings of the Data Compression Conference*.
- Hudson, John (1996). »Trends in multi-authored papers in economics«. In: *The Journal of Economic Perspectives* 10.3, S. 153–158.
- Huselid, Mark A. (2013). »The impact of human resource management practices on turnover, productivity, and corporate financial performance«. In: *The Academy of Management Journal* 38.3, S. 635–672.
- Huth, Edward J. (1986). »Irresponsible authorship and wasteful publication«. In: *Annals of Internal Medicine* 104.2, S. 257–259.
- Hvistendahl, Mara (2013). »China's publication bazaar«. In: *Science* 342.29, S. 1035–1039.
- Jacob, Brian A. und Lars Lefgren (2011). »The impact of research grant funding on scientific productivity«. In: *Journal of Public Economics* 95.9-10, S. 1168–1177.
- Jolly, M., N. Sebire, J. Harris, S. Robinson und L. Regan (2000). »The risks associated with pregnancy in women aged 35 years or older«. In: *Human Reproduction* 15.11, S. 2433–2437.
- Jones, Benjamin F. (Jan. 2009a). »The Burden of Knowledge and the “Death of the Renaissance Man”: Is Innovation Getting Harder?« In: *Review of Economic Studies* 76.1, S. 283–317.
- Jones, Benjamin F. (2009b). »The Knowledge Trap: Human Capital and Development Reconsidered«. In: *Working Paper* March.
- Jones, Benjamin F. (2010). »Age and Great Invention«. In: *Review of Economics and Statistics* 92.1, S. 1–14.
- Jones, Benjamin F. und Bruce A. Weinberg (2011). »Age dynamics in scientific creativity«. In: *Proceedings of the National Academy of Sciences of the United States of America* 108.47, S. 18910–4.

- Joy, Ernest H. und Federico E. Garcia (2000). »Measuring learning effectiveness: A new look at no-significant-difference findings«. In: *Journal of Asynchronous Learning Networks* 4.1, S. 33–39.
- Katz, J.Sylvan und Ben R. Martin (1997). »What is research collaboration?« In: *Research Policy* 26.1, S. 1–18. arXiv: 0201012v3 [arXiv:physics].
- Kim, E. Han, Adair Morse und Luigi Zingales (2006). »What has mattered to economics since 1970«. In: *The Journal of Economic Perspectives* 20.4, S. 189–202.
- Kodrzycki, Yolanda K. und Pingkang Yu (2006). »New Approaches to Ranking Economics Journals«. In: *Contributions in Economic Analysis & Policy* 5, S. 1–40.
- Köhler, Jens und Alfred Oswald (2009). *Die Collective Mind Methode*. Berlin Heidelberg: Springer-Verlag.
- Kohlstedt, Sally Gregory (1980). »Science: The struggle for survival, 1880 to 1894«. In: *Science* 209.4452, S. 33–42.
- Kortum, Samuel S. (1993). »Equilibrium R&D and the patent-R&D ratio: U.S. evidence«. In: *The American Economic Review* 83.2, S. 450–457.
- Kortum, Samuel S. (1997). »Research, patenting, and technological change«. In: *Econometrica* 65.6, S. 1389–1419.
- Kosnik, Lea (2015). »What have economists been doing for the last 50 years? A text analysis of published academic research from 1960-2010«. In: *Economics: The Open-Access, Open-Assessment E-Journal* 9.2015-13, S. 1–18.
- Kotovsky, K., J. R. Hayes und H. A. Simon (1985). »Why are some problems hard? Evidence of Hanoi from tower«. In: *Cognitive Psychology* 17, S. 248–294.
- Krampen, Günter (2010). »Acceleration of citing behavior after the millennium? Exemplary bibliometric reference analyses for psychology journals«. In: *Scientometrics* 83.2, S. 507–513.
- Kuck, David J. (1980). *The structure of computers and computations*. John Wiley und Sons.
- Kuld, Lukas und John O’Hagan (2018). »Rise of multi-authored papers in economics: Demise of the ‘lone star’ and why?«. In: *Scientometrics* 114.3, S. 1207–1225.
- Kurzweil, Ray (1990). *The age of intelligent machines*. Cambridge, MA: MIT Press.
- Kuznets, Simon (1965). *Economic growth and structure*. New York: W.W. Norton.
- Laband, David N. und Robert D. Tollison (2000). »Intellectual collaboration«. In: *Journal of Political Economy* 108.3, S. 632–662.
- Laband, David N., Robert D. Tollison und Gokhan Karahan (2002). »Quality control in economics«. In: *Kyklos* 55.3, S. 315–334.
- Langlois, Richard N. (2002). »Modularity in technology and organization«. In: *Journal of Economic Behavior and Organization* 49, S. 19–37.
- Lapsley, Ruth, Brian Kulik, Rex Moody und J. B. Arbaugh (2008). »Is identical really identical? An investigation of equivalency theory and online learning«. In: *The Journal of Educators Online* 5.1, S. 1–19.
- Larivière, Vincent, Yves Gingras und Éric Archambault (2006). »Canadian collaboration networks: A comparative analysis of the natural sciences, social sciences and the humanities«. In: *Scientometrics* 68.3, S. 519–533.

- Larivière, Vincent, Yves Gingras, Cassidy R. Sugimoto und Andrew Tsou (2015). »Team size matters: Collaboration and scientific impact since 1900«. In: *Journal of the Association for Information Science and Technology* 66.7, S. 1323–1332. arXiv: 0803.1716.
- Larsen, Peder Olesen und Markus von Ins (2010). »The rate of growth in scientific publication and the decline in coverage provided by science citation index«. In: *Scientometrics* 84.3, S. 575–603.
- Lawrence, Peter A. (2007). »The mismeasurement of science«. In: *Current Biology* 17.15, S. 583–585.
- Lee, Ronald und Joshua R. Goldstein (2003). »Rescaling the life cycle: Longevity and proportionality«. In: *Population and Development Review* 29, S. 183–207.
- Lee, Sooho und Barry Bozeman (2005). »The impact of research collaboration on scientific productivity«. In: *Social Studies of Science* 35.5, S. 673–702.
- Lehman, Harvey C. (1953). *Age and achievement*. Princeton, New Jersey: Princeton University Press.
- Lehman, Harvey C. (1962). »More about age and achievement«. In: *The Gerontologist* 2.3, S. 141–148.
- Letchford, Adrian, Tobias Preis und Helen Susannah Moat (2016). »The advantage of simple paper abstracts«. In: *Journal of Informetrics* 10.1, S. 1–8.
- Leuker, Tobias (1997). »'Zwerge auf den Schultern von Riesen' - Zur Entstehung des berühmten Vergleichs«. In: *Mittellateinisches Jahrbuch* 32.1, S. 71–76.
- Levin, Sharon G. und Paula E. Stephan (1991). »Research productivity over the life cycle: Evidence for academic scientists«. In: *Economic Review* 81.1, S. 114–132.
- Lin, Wen Yau Cathy und Mu Hsuan Huang (2012). »The relationship between co-authorship, currency of references and author self-citations«. In: *Scientometrics* 90.2, S. 343–360.
- Lucas, Robert E (1988). »ON THE MECHANICS OF ECONOMIC DEVELOPMENT* Robert E. LUCAS, Jr.« In: *Journal of Monetary Economics* 22.February, S. 3–42.
- Lyman, Peter und Hal R. Varian (2003). *How much information?* URL: <http://www.sims.berkeley.edu/how-much-info-2003/><http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/> (besucht am 03.05.2017).
- Mabe, Michael A. (2003). »The growth and number of journals«. In: *Serials* 16.2, S. 191–197.
- Mabe, Michael A. und Mayur Amin (2001). »Growth dynamics of scholarly and scientific journals«. In: *Scientometrics* 51.1, S. 147–162.
- Mabe, Michael A. und Mayur Amin (2002). »Dr Jekyll and Dr Hyde: Author-reader asymmetries in scholarly publishing«. In: *Aslib Proceedings* 54.3, S. 149–157.
- Mahajan, Vijay und Eitan Muller (1979). »Innovation diffusion and new product growth models«. In: *Journal of Marketing* 43.4, S. 55–68.
- Maki, Ruth H., Williams S. Maki, Michele Patterson und P. David Whittaker (2000). »Evaluation of a web-based introductory psychology course: I. Learning and satisfaction in on-line versus lecture courses«. In: *Behavior Research Methods, Instruments, & Computers* 32.2, S. 230–239.

- Malhotra, Yogesh (2003). »Why knowledge management systems fail: Enablers and constraints of knowledge management in human enterprises«. In: *Handbook of Knowledge Management 1 - Knowledge Matters*. Springer Verlag, S. 577–599.
- Manton, Kenneth G., Xiliang Gu und Vicki L. Lamb (2006). »Long-term trends in life expectancy and active life expectancy in the United States«. In: *Population and Development Review* 32.1, S. 81–106.
- Marois, René und Jason Ivanoff (2005). »Capacity limits of information processing in the brain«. In: *Trends in Cognitive Sciences* 9.6, S. 296–305.
- Martin, Ben R. (1996). »The use of multiple indicators in the assessment of basic research«. In: *Scientometrics* 36.3, S. 343–362.
- Marx, Werner (2011). »Literaturflut - Informationslawine - Wissensexplosion Wächst der Wissenschaft das Wissen über den Kopf?«. In: *Forschung über Forschung*. Bd. I. Bielefeld: UVW UniversitätsVerlagWebler, S. 96–104.
- Masten, Scott E., James W. Meehan und Edward A. Snyder (1991). »The costs of organization«. In: *Journal of Law, Economics, and Organization* 7.1, S. 1–25.
- Mayer, Richard E. (2005). »Cognitive theory of multimedia learning«. In: *The Cambridge Handbook of Multimedia Learning*. New York: Cambridge University Press, S. 31–48.
- McDonald, Jeannette (2002). »Is 'as good as face-to-face' as good as it gets?«. In: *Journal of Asynchronous Learning Networks* 6.2, S. 10–23.
- McDowell, J.M. und Michael Melvin (1983). »The determinants of co-authorship: An analysis of the economics literature«. In: *The Review of Economics and Statistics* 65.1, S. 155–160.
- Means, Barbara, Yukie Toyoma, Robert Murphy, Marianne Bakia und Karla Jones (2009). *Evaluation of evidence-based practices in online learning*. URL: <http://files.eric.ed.gov/fulltext/ED505824.pdf> (besucht am 24. 08. 2017).
- Meissner, Frank und Christina Stockfisch (2012). *Arbeitszeiten in verschiedenen Lebensphasen gestalten*. Techn. Ber. DGB Bundesvorstand.
- Melero, Eduardo und Neus Palomeras (2015). »The renaissance man is not dead! The role of generalists in teams of inventors«. In: *Research Policy* 44.1, S. 154–167.
- Melin, Göran (2000). »Pragmatism and self-organization, Research collaboration on the individual level«. In: *Research Policy* 29.1, S. 31–40.
- Merton, Robert K. (1957). »Priorities in scientific discovery: A chapter in the sociology of science«. In: *American Sociology Review* 22.6, S. 635–659.
- Meves, Yvonne (2013). *Emotionale Intelligenz als Schlüsselfaktor der Teamzusammensetzung*. Wiesbaden: Springer Gabler.
- Miller, D (1996). »A preliminary typology of organisational learning: Synthesising the literature«. In: *Journal of Management* 22.3, S. 485–505.
- Miller, George (1955). »The magical number seven, plus or minus two: Some limits on our capacity for processing information«. In: *Psychological review* 101.2, S. 343–352.
- Mokyr, Joel (2002). *The gifts of athena: Historical origins of the knowledge economy*. Princeton, NJ: Princeton University Press.
- Moore, Gordon E. (1975). »Progress in digital integrated electronics«. In: *1975 International Electron Devices Meeting* 21, S. 11–13.

- Mueller, Pamela (2006). »Exploring the knowledge filter: How entrepreneurship and university–industry relationships drive economic growth«. In: *Research Policy* 35, S. 1499–1508.
- Mukherjee, Satyam, Daniel M. Romero, Ben Jones und Brian Uzzi (2017). »The nearly universal link between the age of past knowledge and tomorrow’s breakthroughs in science and technology: The hotspot«. In: *Science Advances* 3.4, S. 1–9.
- Nelson, Richard R. und Edmund S. Phelps (1966). »Investment in humans, technological diffusion, and economic growth«. In: *The American Economic Review* 56.1/2, S. 69–75.
- Nelson, Richard R. und Sidney G. Winter (1982). *An evolutionary theory of economic change*. Harvard University Press.
- Nicolaisen, Jeppe (2002). »The J-shaped distribution of citedness«. In: *Journal of Documentation* 58.4, S. 383–395.
- Nonaka, Ikujiro und Hirotaka Takeuchi (1995). *The knowledge-creating company - How Japanese companies create the dynamics of innovation*. Oxford University Press, S. 3–20.
- North, Klaus, Andreas Brandner und Thomas Steininger (2016). *Wissensmanagement für Qualitätsmanager*. Wiesbaden: Springer.
- Nowell, Cliff und Therese Grijalva (2011). »Trends in co-authorship in economics since 1985«. In: *Applied Economics* 43.28, S. 4369–4375.
- Odlyzko, Andrew (1997). *The economics of electronic journals*. URL: <http://firstmonday.org/ojs/index.php/fm/rt/prinFRIENDLY/542/463> (besucht am 12. 02. 2018).
- Önder, Ali Sina und Sascha Schweitzer (2017). »Catching up or falling behind? Promising changes and persistent patterns across cohorts of economics PhDs in German-speaking countries from 1991 to 2008«. In: *Scientometrics* 110, S. 1297–1331.
- Oster, Sharon M. und Daniel S. Hamermesh (1998). »Aging and productivity among economists«. In: *The Review of Economics and Statistics* 80.1, S. 154–156.
- Osterloh, Margit und Bruno S. Frey (2000). »Motivation, knowledge transfer, and organizational forms«. In: *Organization Science* 11.5, S. 538–550.
- Owen, John D. (1971). »The demand for leisure«. In: *Journal of Political Economy* 79.1, S. 56–76.
- Pappano, Laura (2012). *The year of the MOOC*. URL: <http://www.edinaschools.org/cms/lib07/MN01909547/Centricity/Domain/272/The%20Year%20of%20the%20MOOC%20NY%20Times.pdf> (besucht am 27. 03. 2018).
- Peffer, Ken, Tuure Tuunanen, Marcus A. Rothenberger und Samir Chatterjee (2008). »A design science research methodology for information systems research«. In: *Journal of Management Information Systems* 24.3, S. 45–78.
- Péguy, Charles (1913). »L’argent«. In: *Cahiers de la quinzaine* 14.6, S. 9–68.
- Phelan, T. J. (1999). »A compendium of issues for citation analysis«. In: *Scientometrics* 45.1, S. 117–136.
- Phlips, Louis (1978). »The demand for leisure and money«. In: *Econometrica* 46.5, S. 1025–1043.
- Polanyi, Michael (1969). *Knowing and being*. Routledge & Kegan Paul: London, S. 123–181.

- Power, Daniel J. (2002). »Decision support systems: A historical overview«. In: *Handbook on Decision Support Systems I*. Springer, S. 121–140.
- Price, Derek J. de Solla (1963). *Little science, big science ...and beyond*. New York: Columbia University Press.
- Raan, Anthony F. J. van (2005). »Measuring science«. In: *Handbook of Quantitative Science and Technology - The Use of Publication and Patent Statistics in Studies of S&T Systems*. Hrsg. von Henk F. Moed, Wolfgang Glänzel und Ulrich Schmoch. Kluwer Academic Publishers, S. 19–50.
- Rath, Katharina und Klaus Wohlrabe (2016). »Recent trends in co-authorship in economics: Evidence from RePEc«. In: *Applied Economics Letters* 23.12, S. 897–902.
- Ritzberger, Klaus (2008). »A ranking of journals in economics and related fields«. In: *German Economic Review* 9.4, S. 402–430.
- Riviera, Emanuela (2013). »Scientific communities as autopoietic systems: The reproductive function of citations«. In: *Journal of the American Society for Information Science and Technology* 64.7, S. 1442–1453.
- Romer, Paul M. (1986). »Increasing returns and long-run growth«. In: *The Journal of Political Economy* 94.5, S. 1002–1037.
- Romer, Paul M. (1990). »Endogenous technological change«. In: *Journal of Political Economy* 98.5, S71–S102.
- Rost, Friedrich (2012). *Lern- und Arbeitstechniken für das Studium*. Wiesbaden: Springer.
- Rubinstein, Ariel (1985). »A bargaining model with incomplete information about time preferences«. In: *Econometrica* 53.5, S. 1151–1172.
- Ruef, Martin, Howard E. Aldrich und Nancy M. Carter (2003). »The structure of founding teams: Homophily, strong ties, and isolation among U.S. entrepreneurs«. In: *American Sociological Review* 68.2, S. 195–222.
- Sacchetti, Silvia und Roger Sugden (2009). *Knowledge in the development of economies - Institutional choices under globalisation (new perspectives on the modern corporation)*. Edward Elgar Publishing limited.
- Schankerman, Mark und A. Pakes (1986). »Estimates of the value of patent rights in european countries during post-1950 period«. In: *The Economic Journal* 96.384, S. 1052–1076.
- Schultz, Theorore William (1963). *The economic value of education*. New York: Columbia University Press.
- Schweitzer, Sascha (2012). *Large-scale multi-item auctions: Evidence from multimedia-supported experiments*. KIT Scientific Publishing: Karlsruhe.
- Schweitzer, Sascha (2017). *Python package persons 0.2a*. URL: <https://github.com/SaschaSchweitzer/persons> (besucht am 21. 09. 2017).
- Scott, Charles E. und John J. Siegfried (2008). »American economic association universal academic questionnaire summary statistics«. In: *American Economic Review* 98.2, S. 630–633.
- Segerstrom, P S (1998). »Endogenous Growth Without Scale Effects«. In: *The American Economic Review* 88.5, S. 1290–1310.

- Seidenfaden, Lutz, Björn Ortelbach und Svenja Hagenhoff (2005). *Grundlagen und aktuelle Herausforderungen in der Wissenschaftskommunikation*. Techn. Ber.
- Shachar, Mickey und Yoram Neumann (2003). »Differences between traditional and distance education academic performances : A meta-analytic approach«. In: *International Review of Research in Open and Distance Learning* 4.2, S. 1–20.
- Shapin, Steven (1989). »The invisible technician«. In: *American Scientist* 77.6, S. 554–563.
- Sheng, Fang, Ali Bilgin, Philip J. Sementilli und Michael W. Marcellin (1998). »Lossy and lossless image compression using reversible integer wavelet transforms«. In: *Proceedings 1998 International Conference on Image Processing*. S. 876–880.
- Shepperd, James A. (1993). »Productivity loss in performance groups: A motivational analysis«. In: *Psychological Bulletin* 113.1, S. 67–81.
- Simkin, Mikhail V. und Vwani P. Roychowdhury (1996). »Read before you cite«. In: *Lancet* 348.9021, S. 144. arXiv: 0212043 [cond-mat].
- Simonton, Dean Keith (2013). »Scientific genius is extinct«. In: *Nature* 493.7434, S. 602.
- Singh, Jasjit und Lee Fleming (2010). »Lone inventors as sources of breakthroughs: Myth or reality?«. In: *Management Science* 56.1, S. 41–56.
- Smith, Adam (1776). *An inquiry into the nature and causes of the wealth of nations*. Methuen & Co: London.
- Smith, Douglas R. (1985). »Top-down synthesis of divide-and-conquer algorithms«. In: *Artificial Intelligence* 27, S. 43–96.
- Smith, Keith (2005). »Measuring innovation«. In: *The Oxford Handbook of Innovation*, S. 148–177.
- Smith, Reid G. und Randall Davis (1981). »Frameworks for cooperation in distributed problem solving«. In: *IEEE Transactions on Systems, Man and Cybernetics* 11.1, S. 61–70.
- Solow, Robert M. (1962). »Technical progress, capital formation, and economic growth«. In: *The American Economic Review* 52.2, S. 76–86.
- Starbuck, William H. (2005). »How much better are the most-prestigious journals? The statistics of academic publication«. In: *Organization Science* 16.2, S. 180–200.
- Stephan, Paula E. (1996). »The economics of science«. In: *Journal of Economic Literature* 34.3, S. 1199–1235.
- Stuhlhofer, Franz (1983). »Unser Wissen verdoppelt sich alle 100 Jahre«. In: *Berichte zur Wissenschaftsgeschichte* 193.6, S. 169–193.
- Sweller, John (1994). »Cognitive load theory, learning difficulty, and instructional design«. In: *Learning and Instruction* 4, S. 295–312.
- Tabah, Albert N. (1999). »Literature dynamics: Studies on growth, diffusion, and epidemics«. In: *Annual Review of Information Science and Technology* 34, S. 249–286.
- Tague, Jean, Jamshid Beheshti und Lorna Rees-Potter (1981). »The law of exponential growth: Evidence, implications and forecasts«. In: *Library Trends* 30.1, S. 125–149.

- Teodorescu, Daniel und Tudorel Andrei (2014). »An examination of 'citation circles' for social sciences journals in eastern european countries«. In: *Scientometrics* 99.2, S. 209–231.
- Tocatlian, Jacques J. (1970). »Are titles of chemical papers becoming more informative?«. In: *Journal of the American Society for Information Science* 21.5, S. 345–350.
- Ucar, Inaki, Felipe Lopez-Fernandino, Pablo Rodriguez-Ulibarri, Laura Sesma-Sanchez, Veronica Urrea-Mico und Joaquin Sevilla (2013). »Growth in the Number of References in Engineering Journal Papers During the 1972-2013 Period«. In: *Scientometrics*, S. 1–11. arXiv: arXiv:1306.4223v1.
- Ucbasaran, Deniz, Andy Lockett, Mike Wright und Paul Westhead (2003). »Entrepreneurial founder teams: Factors associated with member«. In: *Entrepreneurship Theory and Practice* 28.2, S. 107–128.
- Vaishnavi, Vijay und William Kuechler (2008). *Design science research methods and patterns: Innovating information and communication technology*. Auerbach Publications: Boston, MA.
- Van Noorden, Richard (2013). »The true cost of science publishing«. In: *Nature* 495.7442, S. 426–429.
- Ventura, Stephanie J., Sally C. Curtin, Joyce C. Abma und Stanley K. Henshaw (2012). »Estimated pregnancy rates and rates of pregnancy outcomes for the United States, 1990-2008.« In: *National Vital Statistics Reports* 60.7, S. 1–21.
- Verbeek, Arnold, Koenraad Debackere, Marc Luwel und Edwin Zimmermann (2002). »Measuring progress and evolution in science and technology - I: The multiple uses of bibliometric indicators«. In: *International Journal of Management Reviews* 4.2, S. 179–211.
- Wang, Lu (2015). »Graphen-basierte Identitätsambiguierung in Bibliometrie und Social Data«. Masterarbeit. Universität Bayreuth.
- Wasko, Molly McLure und Samer Faraj (2005). »Why should I share? Examining social capital and knowledge contribution in electronic networks of practice«. In: *MIS Quarterly* 29.1, S. 35–57.
- Wegner, Scott B., Ken C. Holloway und Edwin M. Garton (1999). »The effects of internet-based instruction on student learning«. In: *Journal of Asynchronous Learning Networks* 3.2, S. 98–106.
- Weinberger, Cody J., James A. Evans und Stefano Allesina (2015). »Ten simple (empirical) rules for writing science«. In: *PLoS Computational Biology* 11.4, S. 11–13.
- White, Arden und Nelda Rae Hernandez (1991). »Increasing field complexity revealed through article title analyses«. In: *Journal of the American Society for Information Science* 42.10, S. 731–734.
- Willmott, Hugh (2011). »Journal list fetishism and the perversion of scholarship: Reactivity and the ABS list«. In: *Organization* 18.4, S. 429–442.
- Wuchty, Stefan, Benjamin F Jones und Brian Uzzi (2007). »The increasing dominance of teams in production of knowledge.« In: *Science (New York, N.Y.)* 316.5827, S. 1036–1039.

- Yitzhaki, M. und D. Ben-Tamar (1991). »Number of references in biochemistry and other fields; A case study of the Journal of Biological Chemistry throughout 1910-1985«. In: *Scientometrics* 21.1, S. 3–22.
- Yitzhaki, Moshe (1994). »Relation of title length of journal articles to number of authors«. In: *Scientometrics* 30.1, S. 321–332.
- Zuckerman, Harriet (1977). *Scientific Elite: Nobel laureates in the United States*. New York: Free Press.
- Zwietering, M. H., I. Jongenburger, F. M. Rombouts und K. van 't Riet (1990). »Modeling of the bacterial growth«. In: *Applied Environmental Microbiology* 56.6, S. 1875–1881.

Anhang A

Ergänzende Informationen zu den Datensätzen

A.1 Ergänzungen zum EconLit-Datensatz

A.1.1 Anzahl der Zeitschriften pro Jahr im EconLit-Datensatz

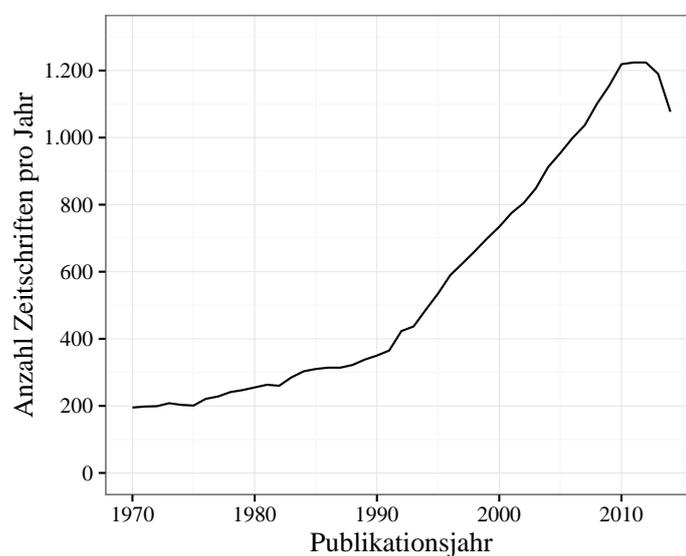


Abbildung A.1: Anzahl der Zeitschriften pro Jahr im EconLit-Datensatz.

A.1.2 Liste der 50 Top-Institutionen

- Australian National University
- Barcelona Graduate School of Economics (GSE)
- Bocconi University
- Boston College
- Boston University
- Brown University
- Centre for Economic Policy Research (CEPR)
- Columbia University
- Cornell University
- Dartmouth College
- Deutsches Institut für Wirtschaftsforschung (DIW), Berlin
- Duke University
- European Central Bank (ECB)
- Federal Reserve Bank of New York
- Federal Reserve Board
- Harvard University
- Hong Kong University of Science and Technology (HKUST)
- International Monetary Fund (IMF)
- Institute of Labor Economics (IZA)
- Leibniz-Institut für Wirtschaftsforschung an der Universität München (ifo)
- London School of Economics (LSE)
- Massachusetts Institute of Technology (MIT)
- Monash University
- National Bureau of Economic Research (NBER)
- National University of Singapore (NUS)
- New York University (NYU)
- Northwestern University
- Organisation de Coopération et de Développement Économiques (OCDE)
- Oxford University
- Paris School of Economics (PSE)
- Peter G. Peterson Institute for International Economics (PIIE)
- Princeton University
- Stanford University
- Stern School of Business, New York University
- Toulouse School of Economics (TSE)
- Universitat Pompeu Fabra
- University College London
- University of British Columbia
- University of California, Berkeley
- University of California, Los Angeles (UCLA)
- University of California, San Diego (UCSD)
- University of Chicago
- University of Michigan
- University of Pennsylvania
- University of Toronto
- University of Warwick
- University of Wisconsin, Madison
- Walter A. Haas School of Business, University of California, Berkeley
- World Bank
- Yale University

A.1.3 Anzahl erfasster Artikel pro Jahr und Zeitschrift

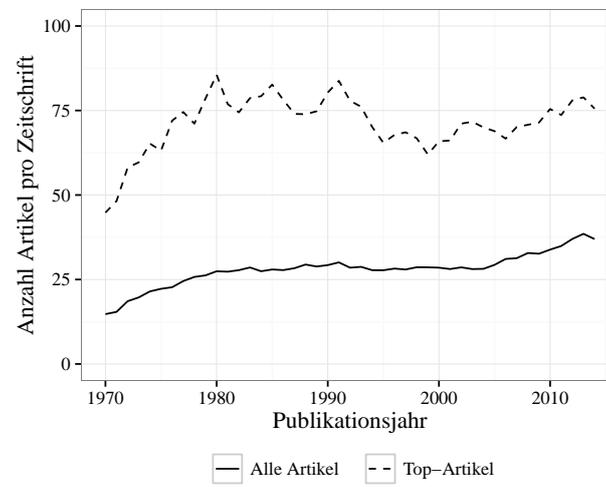


Abbildung A.2: Mittlere Anzahl Artikel pro Zeitschrift und Jahr im EconLit-Datensatz.

A.2 Ergänzungen zum JSTOR-Datensatz

A.2.1 Aufbau einer Webseite mit Artikel-Metadaten bei JSTOR

The screenshot shows a web browser displaying a JSTOR article page. The address bar shows the URL: www.jstor.org/stable/43616927?seq=1#references_tab_contents. The page features the JSTOR logo and search bar at the top. Below the search bar, the journal title "ECONOMETRICA" is displayed, along with the volume and issue information: "Econometrica / Vol. 83, No. 1, January 2015 / TIME AND NO LOTTERIE...". A thumbnail of the journal cover is shown on the left. The main content area includes the article title "JOURNAL ARTICLE TIME AND NO LOTTERIES: AN AXIOMATIZATION OF MAXMIN EXPECTED UTILITY" by Asen Kochov, published in "Econometrica" Vol. 83, No. 1 (January 2015), pp. 239-262. It also lists the publisher "The Econometric Society", the stable URL, and the page count. A list of topics is provided, including Ambiguity, Lotteries, Axioms, Determinism, Mathematical monotonicity, Betting, Utility functions, Economic theory, Consumption, and Axiomatization. There are interactive elements like "Download PDF", "Add to My Lists", "Cite this Item", and "Journal Info". At the bottom, there is an "Abstract" section with the beginning of the text: "This paper axiomatizes an intertemporal version of the maxmin expected-utility model. It employs two axioms specific to a dynamic setting. The first requires that smoothing consumption across states of the world is more".

Abbildung A.3: Beispiel für den Aufbau einer Webseite mit Metainformationen zu einem Beitrag bei JSTOR (Teil 1).

The screenshot shows a web browser window with the URL www.jstor.org/stable/43616927?seq=1#references_tab_contents. The page content is as follows:

Abstract

This paper axiomatizes an intertemporal version of the maxmin expected-utility model. It employs two axioms specific to a dynamic setting. The first requires that smoothing consumption across states of the world is more beneficial to the individual than smoothing consumption across time. Such behavior is viewed as the intertemporal manifestation of ambiguity aversion. The second axiom extends Koopmans' notion of stationarity from deterministic to stochastic environments.

Notes and References

This item contains 29 references.

[Footnotes]

- 2 Gilboa and Schmeidler (1989)
- 4 Strzalecki (2013)
- 5 This reference contains 5 citations:
 - Machina and Schmeidler (1992)
 - Epstein (1999)
 - Epstein (1999),
 - Maccheroni, Marinacci, and Rustichini (2006a)
 - Ghirardato and Marinacci (2002)
- 8 This reference contains 2 citations:
 - Maccheroni, Marinacci, and Rustichini (2006a)
 - Maccheroni, Marinacci, and Rustichini (2006a)

REFERENCES

Anscombe, F. J., AND R. Aumann (1963): "A Definition of Subjective Probability," *The Annals of Mathematical Statistics*, 34, 199-205. [239,240]

Bleichrodt, H., K. I. Rohde, and P. P. Wakker (2008): "Koopmans' Constant Discounting for Intertemporal Choice: A Simplification and a Generalization," *Journal of Mathematical Psychology*, 52 (6), 341-347. [240]

Border, K. C, and U. Segal (1994): "Dynamic Consistency Implies Approximately Expected Utility Preferences," *Journal of Economic Theory*, 63 (2), 170-188. [259]

Cerreia-Vioglio, S., F. Maccheroni, M. Marinacci, and L. Montrucchio (2011): "Un- certainty Averse Preferences," *Journal of Economic Theory*, 146, 1275-1330. [247]

Debreu, G. (1960): "Topological Methods in Cardinal Utility Theory," in *Mathematical Methods in the Social Sciences*. Stanford: Stanford University Press. [253]

Abbildung A.4: Beispiel für den Aufbau einer Webseite mit Metainformationen zu einem Beitrag bei JSTOR (Teil 2). Hinweis: Die Liste der Literaturangaben ist nicht vollständig abgebildet.

A.2.2 Bestimmung von XPath-Ausdrücken

Um die XPath-Ausdrücke einzelner Elemente oder Bereiche einer Webseite zu erhalten, gibt es unterschiedliche Möglichkeiten. Im Rahmen dieser Arbeit wurde die im Browser *Google Chrome* implementierte Funktion genutzt, den Aufbau des HTML-Dokumentes einer Webseite direkt im Browser zu untersuchen. Nach einem Rechtsklick auf die Seite öffnet sich ein Menü, in dem unter anderem der Menüpunkt „untersuchen“ erscheint.

Wählt man diesen aus, wird neben den gerade beobachteten Inhalten ein weiteres Fenster mit dem HTML-Code der Seite angezeigt. In diesem Fenster kann man den Mauszeiger über einzelne Elemente des HTML-Dokumentes bewegen, deren Entsprechung auf der Webseite dann markiert wird. Beispielhaft wurde in Abbildung A.5 der HTML-Tag der Dekade der 2000er Jahre ausfindig gemacht und markiert.¹ Durch Rechtsklick auf diesen HTML-Tag lässt sich wie ebenfalls in Abbildung A.5 zu sehen ist der XPath-Ausdruck für dieses Element kopieren und beispielsweise in das Python-Skript zur Steuerung des Webscrapers einfügen.²

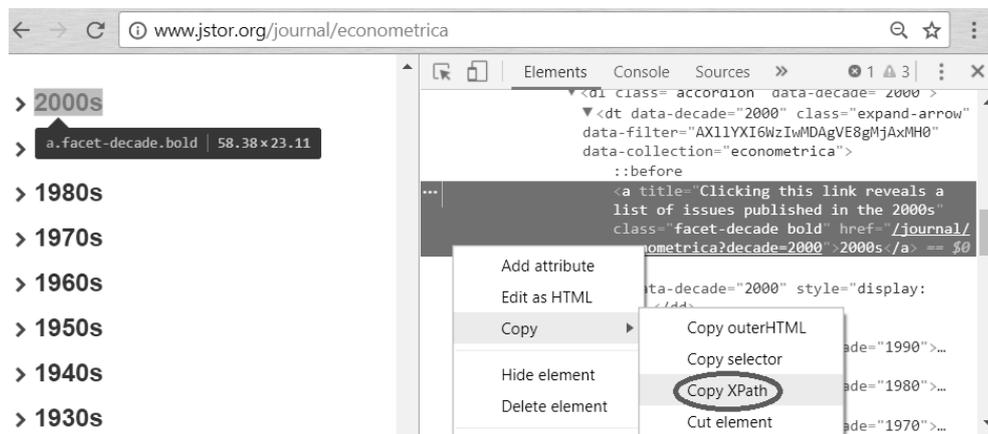


Abbildung A.5: Bestimmung des XPath-Ausdrucks für ein Element der betrachteten Webseite.

A.2.3 Zeitschriften im JSTOR-Datensatz (1878–2014)

Tabelle A.1: Zeitschriften im JSTOR-Datensatz (1878–2014).

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
BWL und Finance				
Academy of Management Journal	1963–2013	2.834	87,7 %	55,6
Administrative Science Quarterly	1956–2014	1.534	75,0 %	26,0
American Economic Review	1911–2014	11.807	65,4 %	113,5
Financial Analysts Journal	1960–2013	3.636	17,9 %	67,3
Financial Management	1972–2012	1.461	92,9 %	35,6

¹HTML-Tags sind Strukturelemente eines HTML-Dokumentes. Beispielweise markieren die Tags `<h1>` bzw. `</h1>` den Anfang bzw. das Ende einer Überschrift, während `` bzw. `` den Beginn und das Ende einer nicht-nummerierten Auflistung markieren.

²Die Ellipse in Abbildung A.5 wurde vom Autor nachträglich zur Hervorhebung eingefügt.

Tabelle A.1: (Fortsetzung)

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
International Economic Review	1960–2013	2.567	96,8 %	47,5
Journal of Business	1954–1997	1.308	40,7 %	29,7
Journal of Consumer Research	1974–2014	2.068	95,1 %	50,4
Journal of Finance	1946–2013	5.269	71,9 %	77,5
Journal of Marketing	1936–2013	3.667	36,4 %	47,0
Journal of Marketing Research	1964–2013	2.795	94,7 %	55,9
Management Science	1954–2014	6.847	93,4 %	112,2
Marketing Science	1982–2013	1.180	66,6 %	36,9
MIS Quarterly	1977–2009	956	28,6 %	29,0
Operations Research	1956–2013	5.135	94,7 %	88,5
Review of Financial Studies	1988–2013	1.328	99,3 %	51,1
Geographie und Geologie				
Geographical Review	1916–2009	3.254	9,2 %	34,6
Geographische Zeitschrift	1895–2010	2.757	2,0 %	23,8
Geography	1929–2010	3.120	38,1 %	38,0
Journal of Geology	1893–1997	4.934	40,1 %	47,0
Mathematik				
American Journal of Mathematics	1878–2013	5.740	44,4 %	42,2
Annals of Mathematics	1884–2009	5.209	10,3 %	41,3
Bulletin of Symbolic Logic	1995–2014	335	71,6 %	16,8
Econometrica	1933–2014	4.634	82,7 %	56,5
Journal of the Am. Mathematical Society	1988–2012	781	59,7 %	31,2
Mathematics of Computation	1960–2013	5.392	85,5 %	99,9
Mathematics of Operations Research	1976–2013	1.835	96,6 %	48,3
Operations Research	1956–2013	5.135	94,7 %	88,5
Proc. of the Am. Mathematical Society	1950–2009	19.081	94,9 %	318,0
SIAM Journal on Applied Mathematics	1966–2013	4.908	43,4 %	102,2
SIAM Journal on Numerical Analysis	1966–2013	4.646	56,9 %	96,8
SIAM Review	1959–2009	2.635	29,0 %	51,7
Psychologie				
Child Development	1930–2009	7.248	96,6 %	90,6
Journal of Business and Psychology	1986–2011	901	97,3 %	34,7

Tabelle A.1: (Fortsetzung)

Disziplin / Zeitschrift	Zeitraum	Anzahl Artikel	Info zu Quellen	Artikel p. a.
Journal of Consumer Psychology	1992–2004	337	97,9 %	25,9
Journal of Occupational Behaviour	1980–1987	188	100,0 %	23,5
Journal of Organizational Behavior	1988–2009	1.137	96,5 %	51,7
Psychological Inquiry	2000–2007	384	98,2 %	48,0
Social Psychology Quarterly	1979–2011	941	64,3 %	28,5
Sociometry	1937–1977	1.367	0,0 %	33,3
VWL				
American Economic Review	1911–2014	11.807	65,4 %	113,5
American Economist	1960–2010	1.206	69,9 %	23,6
Am. Journal of Economics and Sociology	1941–2013	3.495	27,6 %	47,9
Econometric Theory	1985–2013	1.797	82,2 %	62,0
Econometrica	1933–2014	4.634	82,7 %	56,5
Economic Geography	1925–1989	1.643	25,3 %	25,3
Economic History Review	1927–2014	2.382	2,4 %	27,1
Economic Journal	1891–2014	6.398	46,2 %	51,6
Economic Theory	1991–2013	1.864	98,1 %	81,0
International Economic Review	1960–2013	2.567	96,8 %	47,5
Journal of Economic Growth	1996–2013	246	100,0 %	13,7
Journal of Economic Literature	1969–2012	683	82,6 %	15,5
Journal of Economic Perspectives	1987–2014	1.411	93,3 %	50,4
Journal of Finance	1946–2013	5.269	71,9 %	77,5
Journal of Labor Economics	1983–2011	842	59,9 %	29,0
Journal of Political Economy	1892–2011	5.811	39,2 %	48,4
Quarterly Journal of Economics	1886–2012	4.891	31,0 %	38,5
RAND Journal of Economics	1984–2014	1.256	99,9 %	40,5
Review of Economic Studies	1933–2013	2.919	78,1 %	36,0
Review of Economics and Statistics	1919–2013	5.050	68,5 %	53,2

Bem.: (i) Die Abkürzungen Am. bzw. Proc. stehen für American bzw. Proceedings.

(ii) Die Spalte „Info zu Quellen“ gibt an, zu welchem Anteil der Artikel eine Information bzgl. der Anzahl der Quellenangaben vorliegt.

A.2.4 Anzahl erfasster Artikel

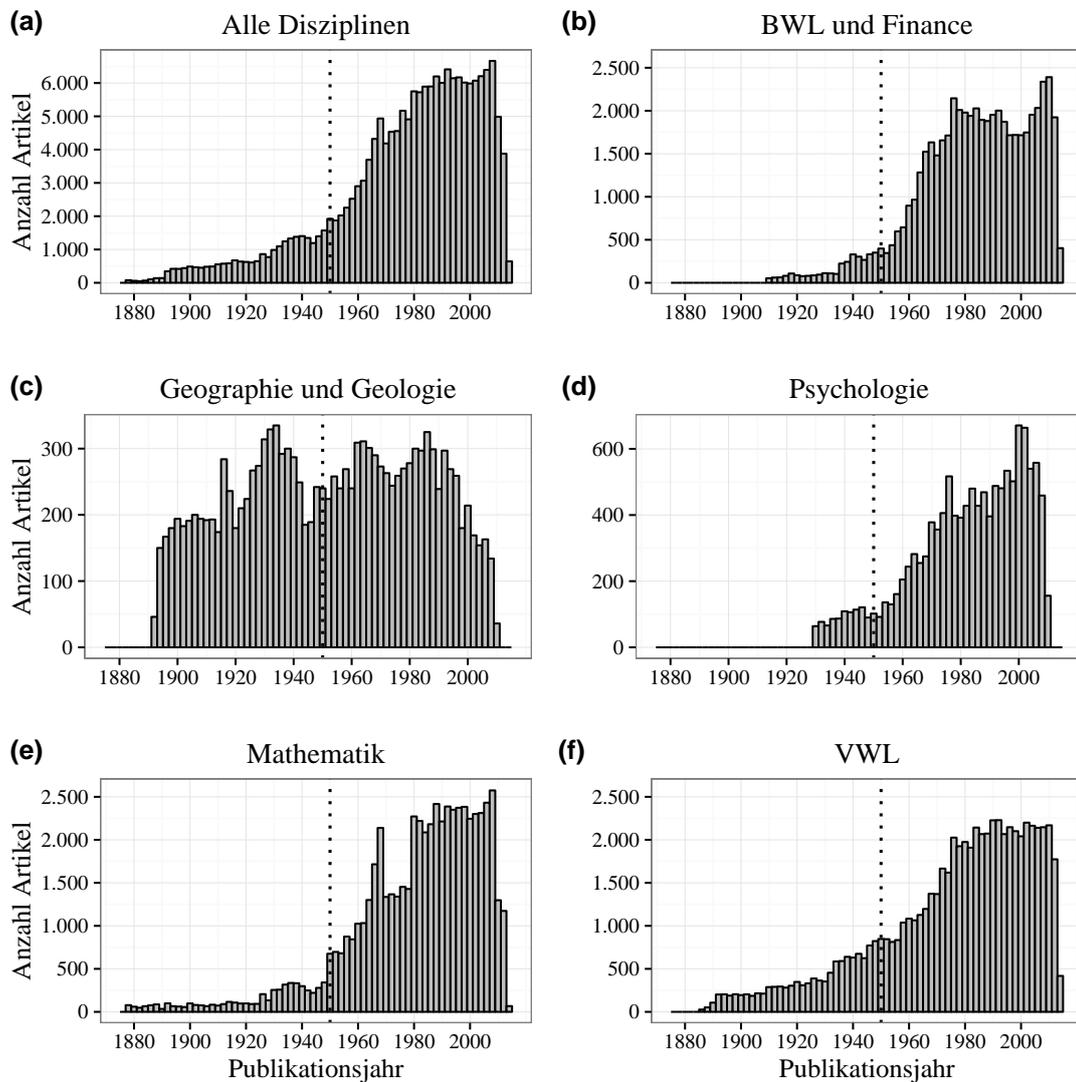


Abbildung A.6: Anzahl erfasster Artikel im JSTOR-Datensatz pro zwei Jahre. Die gepunktete Linie markiert das Jahr 1950.

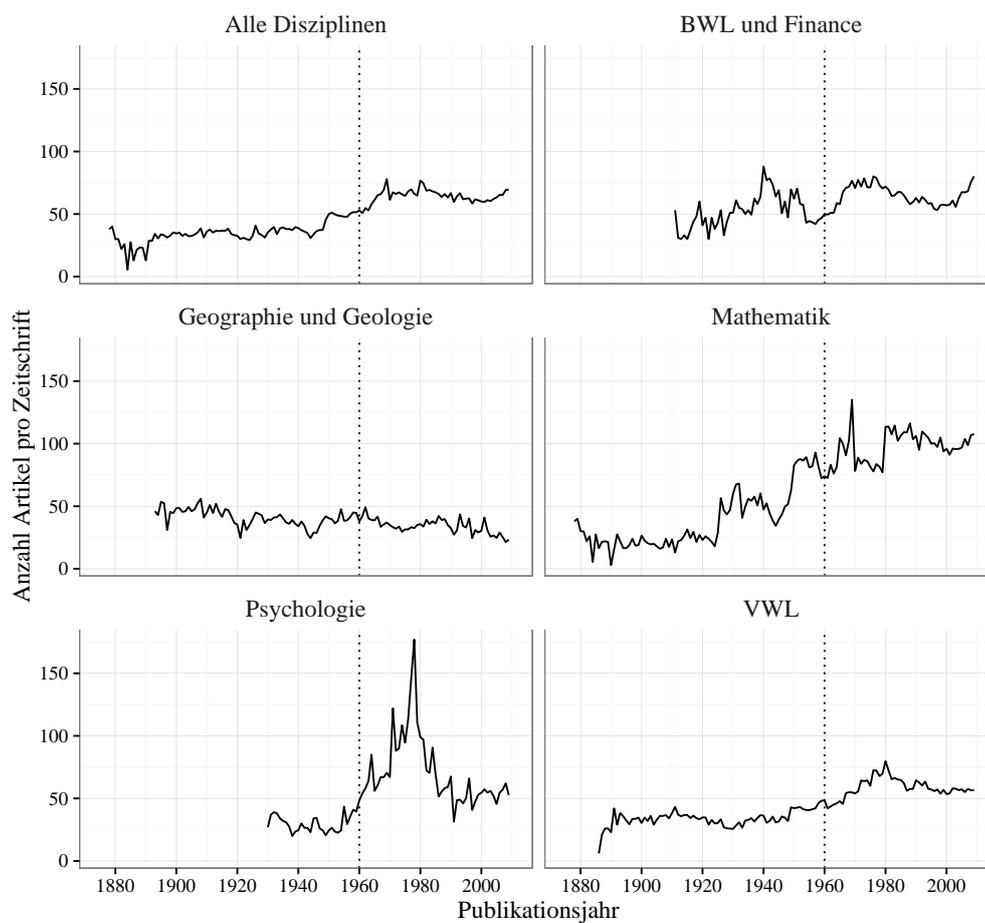


Abbildung A.7: Mittlere Anzahl Artikel pro Zeitschrift und Jahr im JSTOR-Datensatz. Die gepunktete Linie markiert das Jahr 1960.

Anhang B

Ergänzende Auswertungen

B.1 Teamgröße

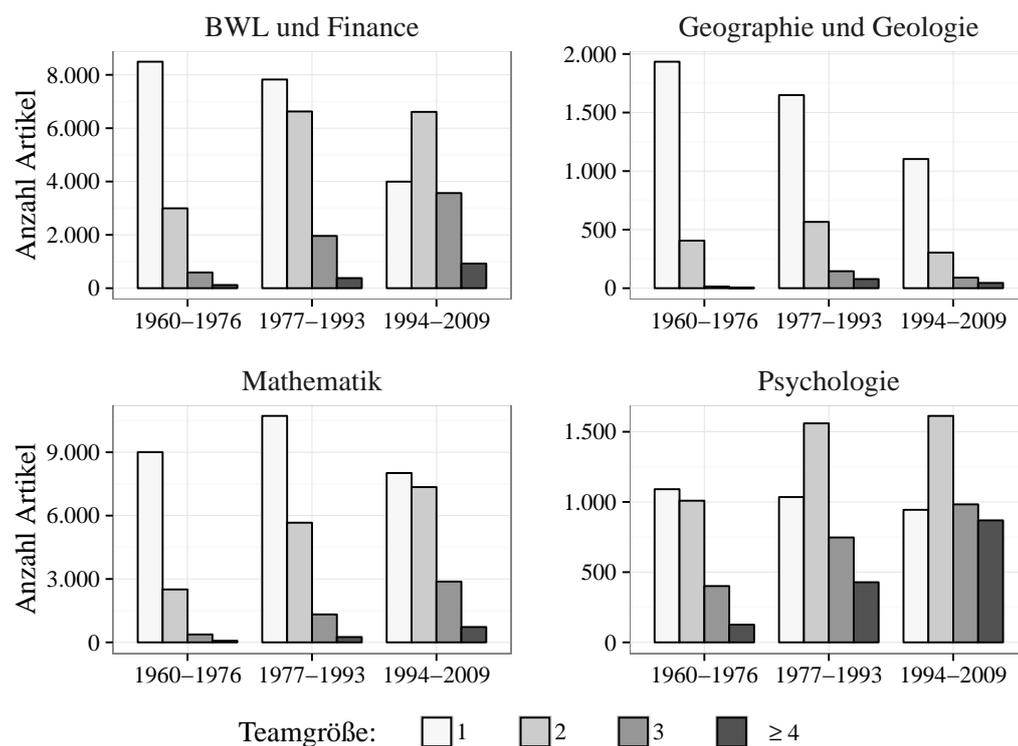


Abbildung B.1: Anzahl der Artikel im JSTOR-Datensatz nach Teamgrößen.

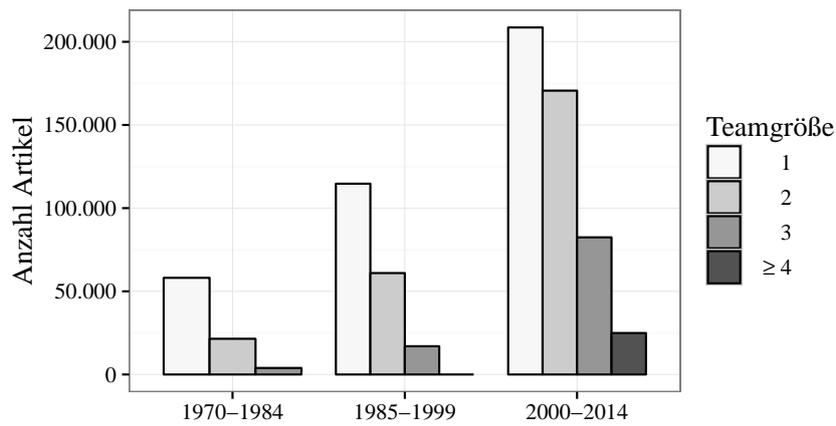


Abbildung B.2: Anzahl der Artikel im EconLit-Datensatz nach Teamgrößen.

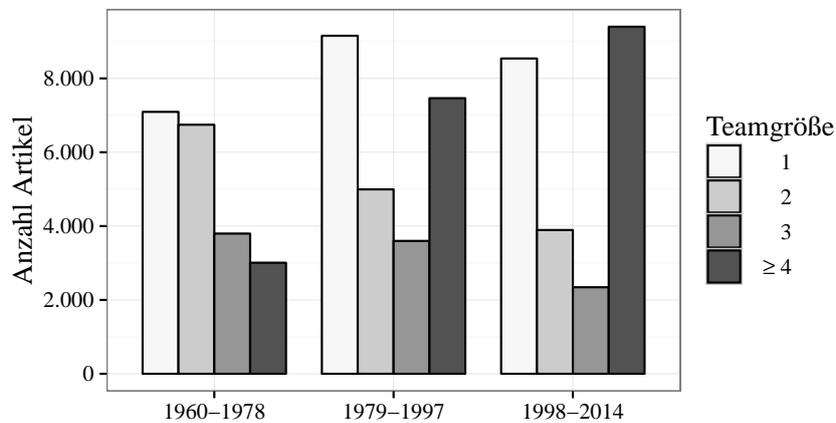


Abbildung B.3: Anzahl der Artikel in *Science* nach Teamgrößen.

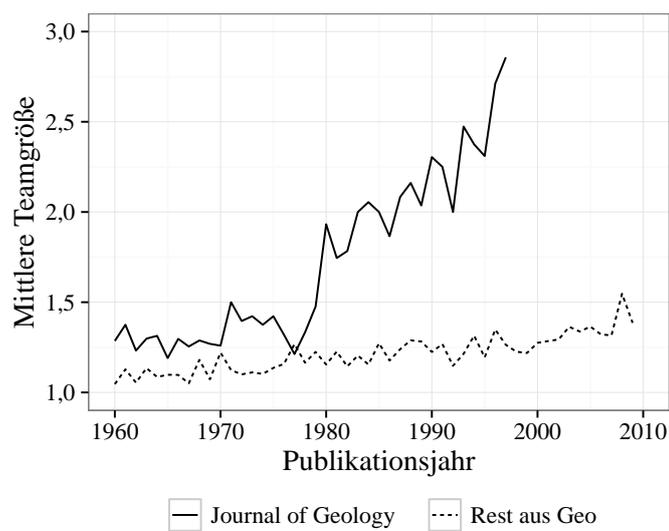


Abbildung B.4: Mittlere Teamgröße in der Zeitschrift *The Journal of Geology* und den restlichen Zeitschriften der Disziplin *Geographie und Geologie*.

B.2 Anzahl der Quellenangaben

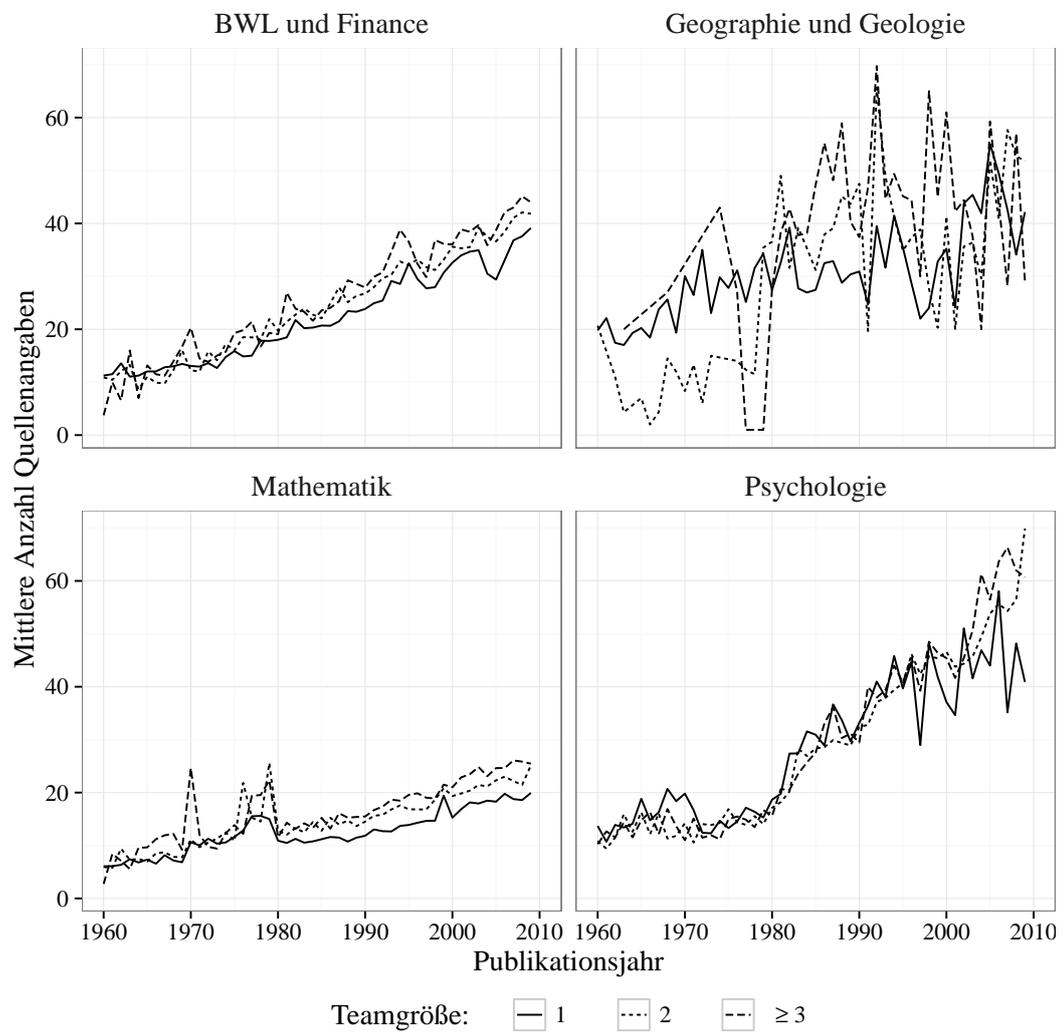


Abbildung B.5: Mittlere Anzahl Quellenangaben pro Artikel im JSTOR-Datensatz für verschiedene Teamgrößen.

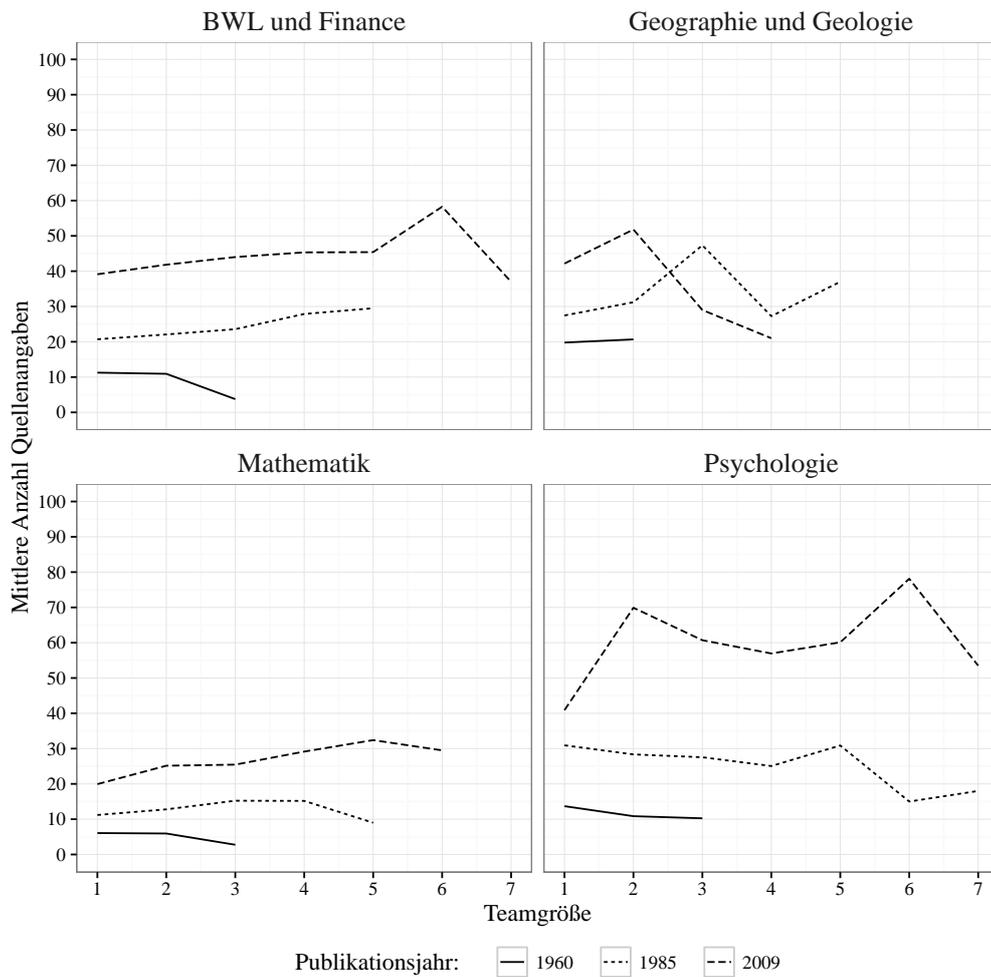


Abbildung B.6: Mittlere Anzahl der Quellenangaben pro Artikel im JSTOR-Datensatz in Abhängigkeit von der Teamgröße.

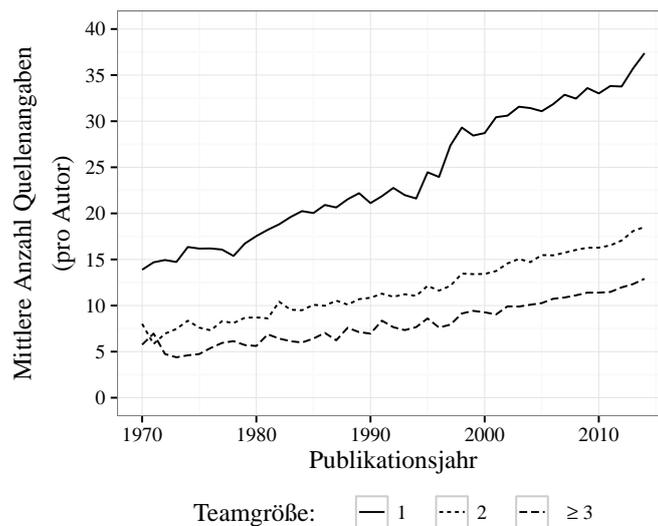


Abbildung B.7: Mittlere Anzahl Quellenangaben pro Artikel und Autor im EconLit-Datensatz für verschiedene Teamgrößen.

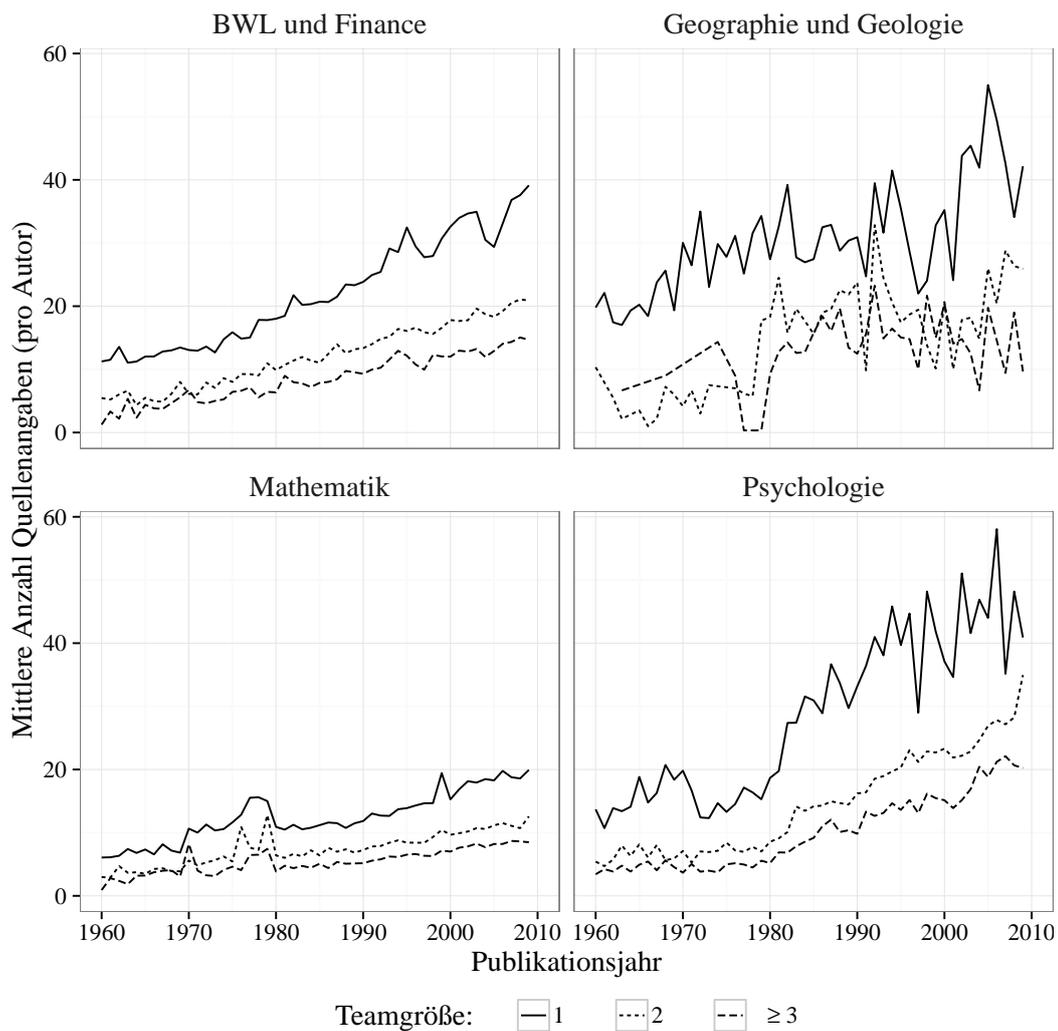


Abbildung B.8: Mittlere Anzahl Quellenangaben pro Artikel und Autor im JSTOR-Datensatz für verschiedene Teamgrößen.

B.3 Anzahl der Quellenangaben bei erstem Artikel

Tabelle B.1: Lineare Regressionen für die Anzahl der Quellenangaben beim ersten Artikel im EconLit-Datensatz (Regressionen auf Artikel-Ebene).

	Abhängige Variable: Anzahl Quellenangaben					
	Top-Artikel: (1)–(3)			Alle Artikel: (4)–(6)		
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,617*** (0,010)	0,569*** (0,012)	0,558*** (0,019)	0,564*** (0,007)	0,600*** (0,010)	0,628*** (0,012)
Teamgröße		0,378* (0,172)	0,342 (0,228)		0,610*** (0,081)	0,605*** (0,088)
Autorinnen			–1,349 ⁺ (0,716)			0,726* (0,289)
Alle Top-Inst			1,796*** (0,425)			0,626* (0,302)
Alle Nicht-Aka			–0,613 (0,852)			–3,512*** (0,401)
Zeitschrift (Dummys)		Ja	Ja		Ja	Ja
JEL-Feld (Dummys)		Ja	Ja		Ja	Ja
Beobachtungen	12.267	12.267	8.286	69.052	69.052	60.206
R^2	0,224	0,296	0,245	0,079	0,328	0,319

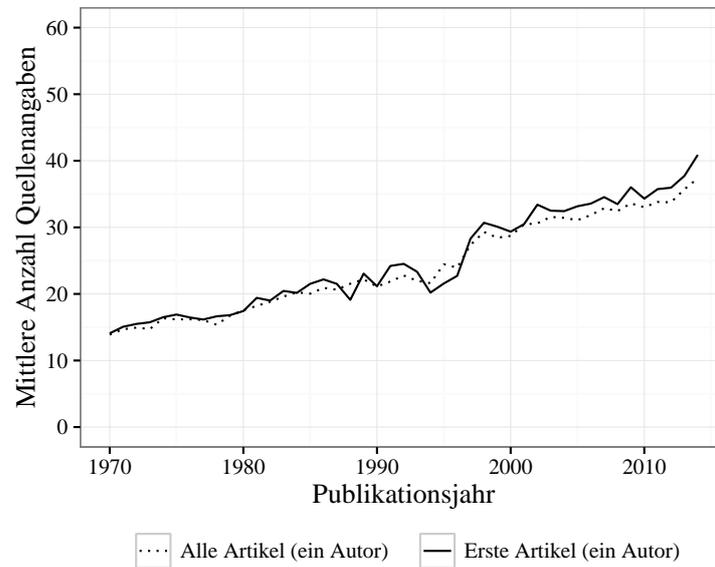


Abbildung B.9: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors im EconLit-Datensatz, wenn er bei diesem alleiniger Autor war.

Tabelle B.2: Lineare Regressionen für die Anzahl der Quellenangaben beim ersten Artikel eines Autors im JSTOR-Datensatz auf Artikel-Ebene.

	Abhängige Variable: Anzahl Quellenangaben				
	Gesamt	BWL	Geo	Mathe	Psycho
Publikationsjahr	0,580*** (0,007)	0,702*** (0,012)	0,455*** (0,064)	0,336*** (0,007)	1,253*** (0,022)
Teamgröße	1,274*** (0,090)	0,065 (0,167)	2,415*** (0,704)	1,269*** (0,117)	1,100*** (0,185)
Autorinnen	0,648* (0,329)	-0,058 (0,656)	-0,281 (2,201)	-0,619 (0,456)	0,649 (0,636)
Zeitschrift (Dummys)	Ja	Ja	Ja	Ja	Ja
Beobachtungen	47.264	18.949	1.522	21.742	7.350
R^2	0,407	0,366	0,233	0,266	0,405

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$

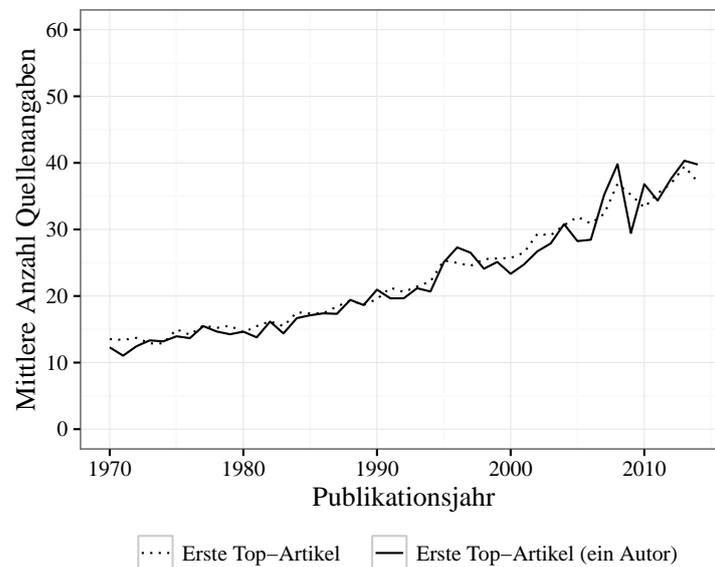


Abbildung B.10: Mittlere Anzahl Quellenangaben beim ersten Top-Artikel eines Autors im EconLit-Datensatz, wenn er bei diesem alleiniger Autor war.

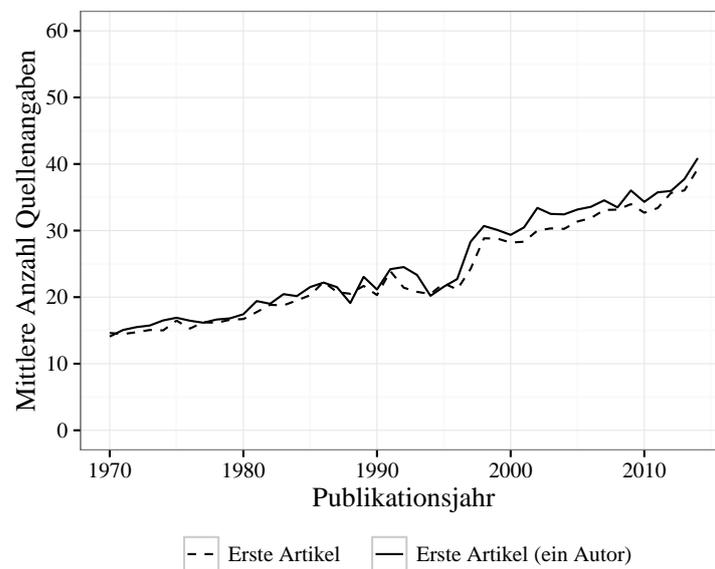


Abbildung B.11: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors im EconLit-Datensatz, wenn er bei diesem alleiniger Autor war.

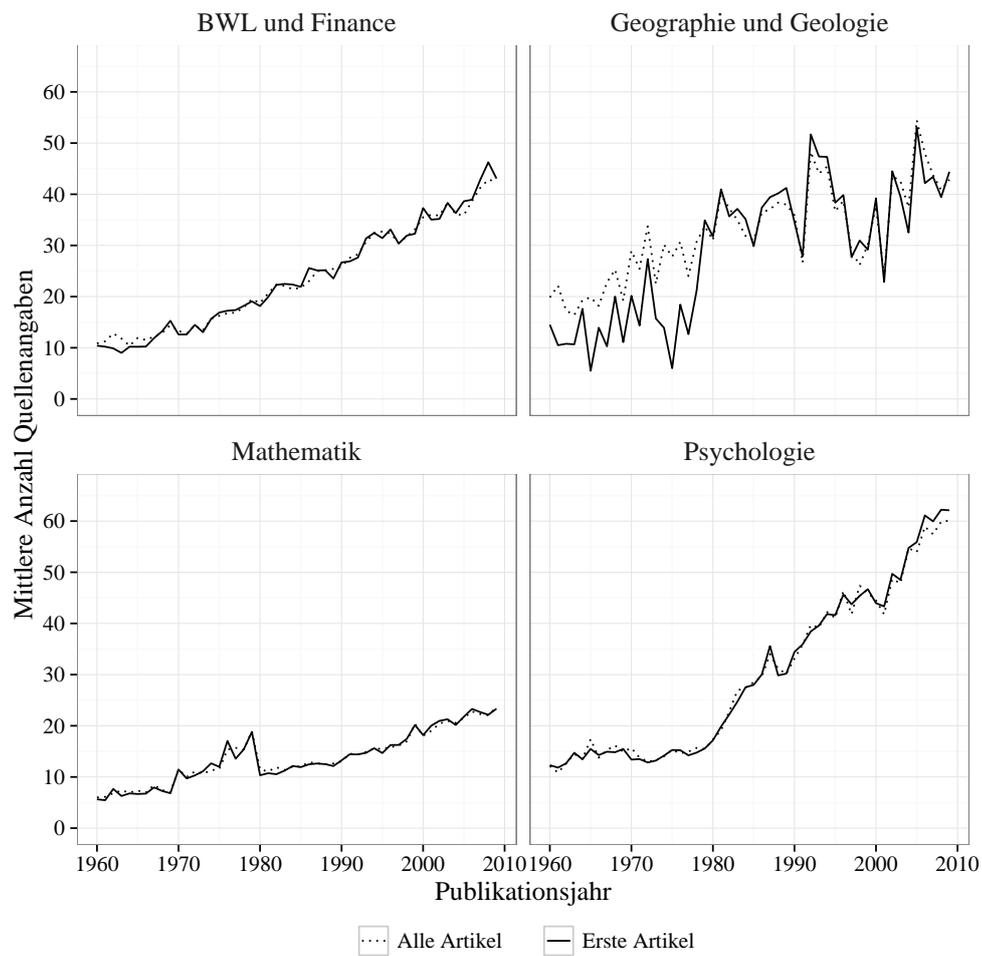


Abbildung B.12: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors im JSTOR-Datensatz.

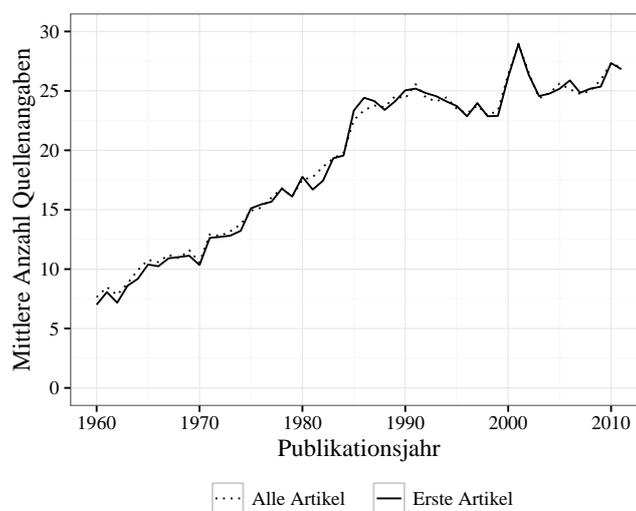


Abbildung B.13: Mittlere Anzahl Quellenangaben beim ersten Artikel eines Autors in der Zeitschrift *Science*.

B.4 Länge von Titeln und Abstracts

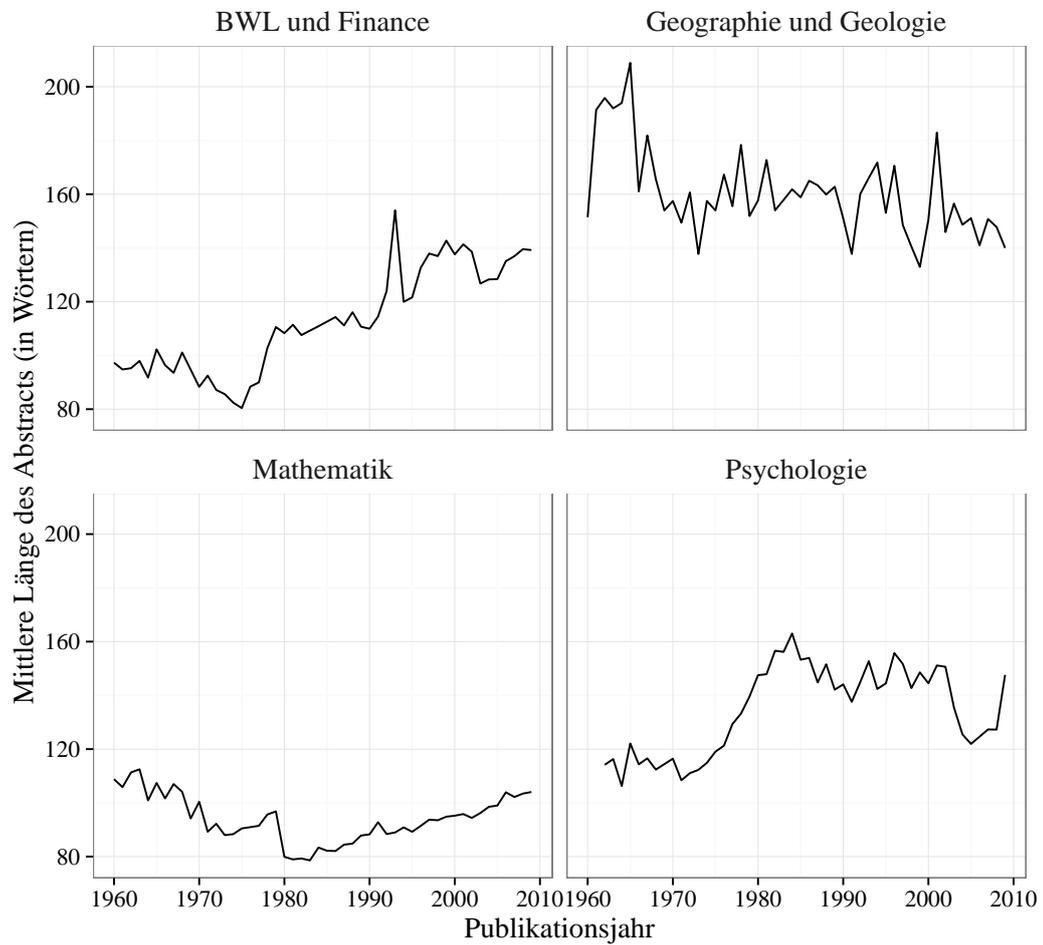


Abbildung B.14: Mittlere Anzahl der Wörter im Abstract von Artikeln im JSTOR-Datensatz.

B.5 Alter bei erstem Artikel

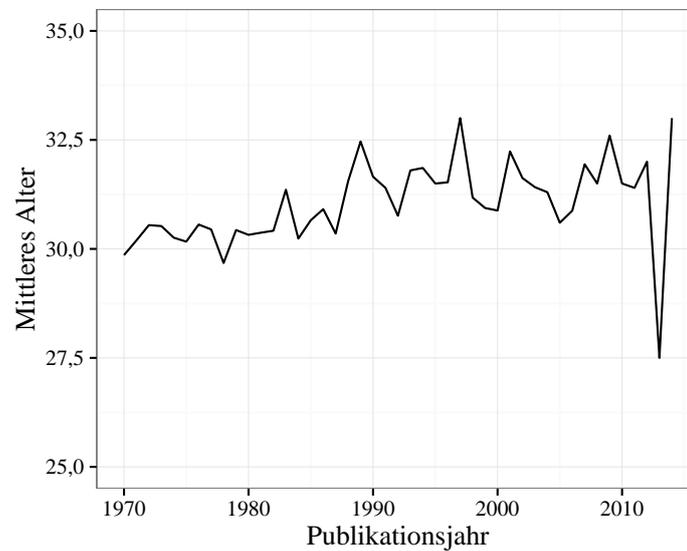


Abbildung B.15: Mittleres Alter beim ersten Top-Artikel eines Autors in der Volkswirtschaftslehre im Alter zwischen 25 und 35 Jahren.

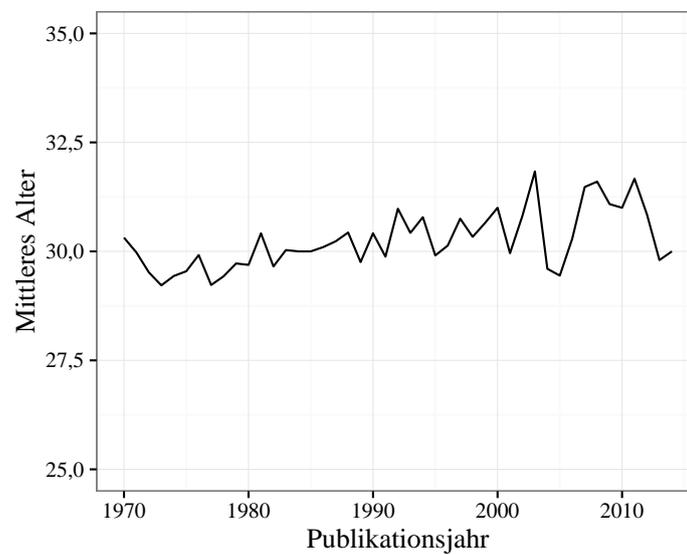


Abbildung B.16: Mittleres Alter beim ersten Artikel eines Autors in der Volkswirtschaftslehre im Alter zwischen 25 und 35 Jahren.

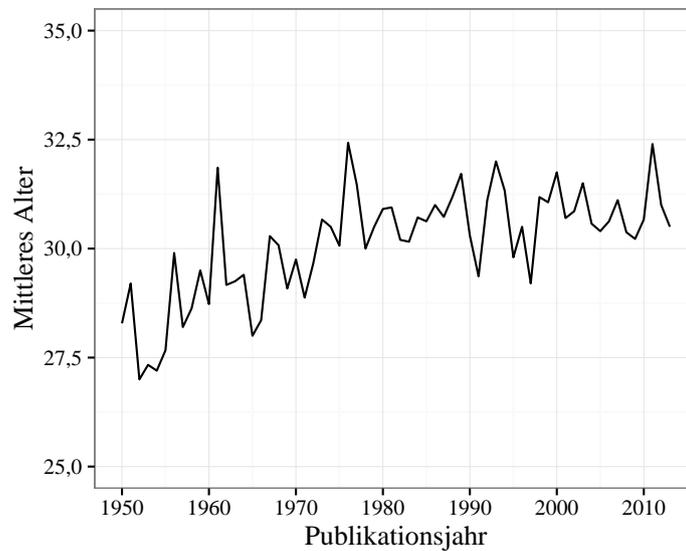


Abbildung B.17: Mittleres Alter beim ersten Artikel eines Autors in der Mathematik im Alter zwischen 25 und 35 Jahren.

Tabelle B.3: Lineare Regressionen für das Alter beim ersten Artikel eines Autors in der Mathematik im Alter zwischen 25 und 35 Jahren.

Abhängige Variable: Alter bei erstem Artikel						
	(1)	(2)	(3)	(4)	(5)	(6)
Publikationsjahr	0,045*** (0,006)	0,047*** (0,007)	0,043*** (0,007)	0,042*** (0,007)	0,033*** (0,008)	0,031*** (0,009)
Autorin		-0,616 (0,428)	-0,681 (0,428)	-0,707 (0,432)	-0,646 (0,420)	-0,671 (0,424)
Teamgröße			0,309* (0,150)	0,238 (0,156)	0,266+ (0,150)	0,216 (0,157)
Zeitschrift (Dummies)				Ja		Ja
Beginn	1950	1950	1950	1950	1960	1960
Beobachtungen	594	594	594	594	527	527
R^2	0,076	0,080	0,086	0,101	0,044	0,059

Signifikanzniveaus: *** $p < 0,001$; ** $p < 0,01$; * $p < 0,05$; + $p < 0,1$