

Justus-Liebig-Universität Gießen
Institut für Soziologie

Bachelorthesis

**„Akademischer Kapitalismus“ analysiert:
Sprachliche Trends in soziologischen Fachzeitschriften**

Thomas Haase*

13. August 2024

Erstgutachter: Prof. Dr. Elmar Schlüter

Zweitgutachter: Tim Schmidt

** Semester: 6. Fachsemester*

E-Mail 1: thomas.haase@sowi.uni-giessen.de

E-Mail 2: thhaase.soz@gmail.com

Inhaltsverzeichnis

1	Einleitung	1
2	Theorie	3
2.1	Einleitung	3
2.2	Hintergründe des akademischen Kapitalismus	3
2.3	Konsequenzen des Akademischen Kapitalismus	6
2.4	Fragestellungen	9
2.5	Zusammenfassung	10
3	Stand der Forschung	11
3.1	Einleitung	11
3.2	Sentiments in PubMed-Abstracts	11
3.3	Interkulturelle Psychologie und Akademischer Kapitalismus	12
3.4	Sentiment Analyse in der Naturschutzforschung	13
3.5	Positivität in Wissenschaftsdisziplinen	13
3.6	Zusammenfassung	14
4	Methode	15
4.1	Einleitung	15
4.2	Sentiment- und Diktionärsanalysen	15
4.3	Daten	16
4.4	Operationalisierung	18
4.5	Analyseansatz	19
4.6	Zusammenfassung	20
5	Ergebnisse	23
5.1	Einleitung	23
5.2	Datenbereinigung und Inspektion	23
5.3	Jährlich aggregierte Sentimenttrends	26
5.4	Vergleich der Diktionärergebnisse	30
5.5	Lineare Regressionen	32
5.6	Regressionsdiagnostik	35
5.7	Zusammenfassung	38
6	Diskussion und Fazit	41
6.1	Einleitung	41
6.2	Interpretation der Ergebnisse	41
6.3	Implikationen der Ergebnisse	42
6.4	Limitationen	43
6.5	Fazit	44

Anhang	VII
Literatur	XXI

Kurzfassung

Die Bachelorarbeit untersucht die Veränderung des positiven Framings in soziologischen Publikationen vor dem Hintergrund des akademischen Kapitalismus. Basierend auf der Theorie einer zunehmenden Selbstvermarktung wird analysiert, ob sich diese durch verstärkt positive Sprache in wissenschaftlichen Abstracts niederschlägt. Die Arbeit führt eine Sentimentanalyse von 121.789 englischsprachige Abstracts soziologischer Artikel aus der Web of Science Datenbank (1993-2023) mit zwei Diktionären für zusätzliche Validität durch. Mittels Intervallgrafiken, linearen Regressionen und Monte-Carlo-Simulationen werden zeitliche Trends positiver und negativer Sentiments untersucht. Die Ergebnisse zeigen einen sehr kleinen, statistisch nicht signifikanten Anstieg positiver Sentiments über die Zeit, während für negative Sentiments kein klarer Trend feststellbar ist. Die Arbeit offenbart methodische Herausforderungen bei der Sentimentanalyse wissenschaftlicher Texte, insbesondere hinsichtlich Datenverteilung und der Signifikanzinterpretation großer Stichproben. Entgegen den Ergebnissen bisheriger Forschung liefert die Untersuchung keine Evidenz für einen substantiellen Anstieg positiven Framings.

1 | Einleitung

Im Zuge der Globalisierung gewann die Ökonomisierung des gemeinschaftlichen Zusammenlebens neue Aufmerksamkeit. Der globale Markt erreicht trotz nationaler Regularien den Alltag der in der Gesellschaft lebenden Individuen. Das überrascht wenig, denn Max Weber zeigte bereits 1904, durch welche Logik protestantisch-gesellschaftliche Normen selbst die Wurzel des Kapitalismus bilden (Weber, 1934). Außerdem reflektiert Weber auch bereits 1919 die Verknüpfung von Wissenschaft und Ökonomie auf den ersten Seiten seines Aufsatzes “Wissenschaft als Beruf” (Weber, 1919). Die moderne Wissenschaftsforschung begann mit ihrer “praxeologischen Wende” 1980 bis 2000 die Interaktionen von Wissenschaft mit der Politik und Ökonomie zu untersuchen (Kaldewey, 2023, S. 8). Seit einigen Jahren veröffentlichen Bibliometriker und Forschende unterschiedlichster Disziplinen Ergebnisse, welche einen Anstieg positiver Formulierungen in wissenschaftlichen Publikationen zeigen. Dies geschieht bis auf einzelne Ausnahmen ungeachtet der gesellschaftlichen Einbettung von Wissenschaft und der daraus resultierenden Interaktion zwischen Wissenschaft, Ökonomie, Politik und Gesellschaft. Diese Arbeit soll einen Beitrag zum besseren Verständnis der Veränderung akademischer Sprache in wissenschaftlichen Publikationen leisten. Dabei untersucht diese Arbeit eine Veränderung des Framings soziologischer Publikationen und setzt die Ergebnisse in Beziehung zu der zunehmenden Ökonomisierung der Wissenschaft.

Zunächst fasst die Arbeit die Theorie des akademischen Kapitalismus zusammen. Das erste Kapitel stellt die Ursachen des akademischen Kapitalismus anhand zentraler Werke vor. Im Anschluss geht das Kapitel auf die Konsequenzen des akademischen Kapitalismus ein, welcher auf die im Wissenschaftssystem forschenden Individuen wirkt. Die Fragestellung wird auf Basis der beschriebenen Theorie dargestellt und eingeordnet. Das zweite Kapitel stellt die relevantesten Studien für die Untersuchung der Fragestellung dar. Anschließend erläutert das dritte Kapitel die methodischen Grundlagen der Untersuchung. Dabei wird genauer auf die Methode der Sentiment- und Diktionsanalysen sowie auf die Operationalisierung und den Analyseansatz eingegangen. Im vorletzten Kapitel werden die Ergebnisse der Analyse beider Diktionäre dargestellt, verglichen und durch lineare Regressionen reflektiert. Im letzten Kapitel werden alle Ergebnisse miteinander in Beziehung gesetzt und ein Fazit gezogen. Die Arbeit leistet einen neuen Beitrag, indem sie die Analyse des Framings wissenschaftlicher Publikationen explizit auf die Disziplin der Soziologie ausrichtet. Zusätzlich bringt diese Arbeit auch neue methodische Anregungen in das beforschte Themengebiet ein, indem durch Simulationen Verzerrungen der statistischen Signifikanz aufgrund hoher Fallzahlen mit in die Interpretation einbezogen werden. Um besonders valide Ergebnisse zu erzielen wird eine Vollerhebung aller soziologischen Publikationen von Web of Science durchgeführt und die Messung des Framings mit zwei Diktionären durchgeführt.

2 | Theorie

2.1 Einleitung

Das folgende Theoriekapitel behandelt zentrale Inhalte der Theorie des “Akademischen Kapitalismus”. Der Fokus liegt auf den Werken von Slaughter und Leslie (1999) und Münch (2011), aufgrund ihrer fundamentalen Bedeutung für die Theorie des “Akademischen Kapitalismus”. Das Kapitel beginnt mit einer Darstellung der ursprünglichen politischen Entscheidungen, welche als Ursache des akademischen Kapitalismus gesehen werden. Die eingeführten Policies veränderten das Wissenschaftssystem grundlegend. Im Anschluss werden die Reaktionen der unternehmerischen Universitäten auf die zuvor beschriebenen Policyänderungen dargestellt. Dazu wird mithilfe des Garbagecan Modells sowie der Ressourcenabhängigkeitstheorie das Verhalten der Unternehmen plausibel gemacht. Danach wird darauf eingegangen, wie der akademische Kapitalismus aus dem Verhalten der Universitäten als überliegende Struktur durch Konstruktionsmechanismen emergiert und durch Erhaltungsmechanismen bestehen bleibt. Der zweite Abschnitt beschäftigt sich mit den Konsequenzen des akademischen Kapitalismus sowohl für die im Wissenschaftssystem tätigen Individuen, als auch für die Struktur wissenschaftlicher Disziplinen. Anschließend fokussiert sich die Analyse auf die Reaktionen einzelner Forscher, welche auf die durch den akademischen Kapitalismus geänderten Umstände reagieren. In der letzten thematischen Ausführung werden die Konsequenzen des veränderten Handelns der Forscher für wissenschaftliche Disziplinen dargestellt und einzelne Aspekte disziplinärer Unterschiede knapp aufgegriffen. Zum Schluss des Kapitels wird die Fragestellung dieser Arbeit vorgestellt, in den Kapitelkontext eingeordnet und danach eine Zusammenfassung des gesamten Kapitels erstellt.

2.2 Hintergründe des akademischen Kapitalismus

Die politischen Entscheidungen und deren Hintergründe, welche als Ursache für strukturelle Veränderungen des Wissenschaftssystems gelten, werden zu Beginn dargestellt. Anschließend erfolgt eine Erläuterung der daraus resultierenden strukturellen Veränderungen im Wissenschaftssystem. In der ersten Hälfte des 20. Jahrhunderts waren Akademiker, insbesondere Fakultätsmitglieder, weitgehend vom Markt abgeschirmt und ihre Arbeit staatlich finanziert (Slaughter & Leslie, 1999). Ab Mitte des 20. Jahrhunderts begann jedoch eine allmähliche Einbeziehung der Professoren in den Markt, die durch die Globalisierung noch beschleunigt wurde (Slaughter & Leslie, 1999). Weber (1919) beschreibt zwar bereits “Die großen Institute medizinischer oder naturwissenschaftlicher Art sind Staatskapitalistische Unternehmungen.”, merkt aber dennoch an “Es besteht eine außerordentlich starke Kluft [...] zwischen dem Chef eines solchen großen kapitalistischen Universitätsunternehmens und dem gewöhnlichen Ordinarius alten Stils.”. Die 1980er Jahre markierten für diese Entwicklung einen Wendepunkt (Slaughter & Leslie, 1999).

Zu dieser Zeit initiieren Ronald Reagan und Margaret Thatcher maßgeblich ein Marktparadigma voran, welches zu einer globalen Hegemonie eines New Public Management führte (Münch, 2011; Slaughter & Leslie, 1999). Insbesondere ist hier der “Managerialismus” zu nennen, welcher unternehmerische Entscheidungslogiken in politischen Institutionen legitimierte (Münch, 2011). Somit wurde auch die Wissenschaft durch den Staat zunehmend nach marktorientierten Kriterien bewertet (Münch, 2011). Diese Bewertung untergrub den impliziten Gesellschaftsvertrag, der zuvor das wissenschaftliche Arbeiten legitimierte (Münch, 2011). Der zuvor existierende, implizite Vertrag wird von Münch (2011) als eine Art “harmonische Idealform” beschrieben. Diese beschreibt, wie die Wissenschaft eine Art Treuhänderschaft mit anderen Gesellschaftsbereichen eingeht. Dabei wird durch die Wissenschaft wissenschaftliches Kapital bereitgestellt und im Gegenzug soziales und ökonomisches Kapital sowie Legitimation für Grundlagenforschung erhalten. Die Studie Slaughter und Leslie (1999) dokumentiert die Intensivierung des universitären Wettbewerbs in den 80er Jahren in den USA, Kanada, Großbritannien und Australien. Fördermittelverknappung und selektive Unterstützung erfolgreicher Institutionen fungierten dabei als treibende Kräfte. Ziel der Politik war die nationale Wettbewerbsfähigkeit zu steigern und sich an globale Marktanforderungen anzupassen (Slaughter & Leslie, 1999). Ähnliche Dynamiken lassen sich auch in Deutschland beobachten (Münch, 2011).

Ein Beispiel, wie durch Policyänderungen Universitäten unabhängiger von staatlicher Förderung gemacht werden sollten und zugleich den Zwischenuniversitären Wettbewerb zu steigern, war der Bayh-Dole Act von 1980 in den USA (Münch, 2011). Dieser erlaubte Universitäten, Forschungsergebnisse zu patentieren und somit direkt in privates ökonomisches Kapital umzuwandeln (Münch, 2011). Somit wurde wissenschaftliches Wissen zunehmend von einem öffentlichen zu einem privaten Gut umgewandelt (Münch, 2011). Dieses private Wissen ist besonders nützlich für die Industrie, welche verstärkt nach neuen Produkten suchte und auf staatlich finanzierte, kommerzielle Forschung drängte (Slaughter & Leslie, 1999). Die Näherung der Universitäten an die Industrie führte zur, teils staatlich geförderten, Einrichtung kooperativer Forschungszentren (Slaughter & Leslie, 1999). Zudem wurde die Erhöhung der Studierendenzahlen zu einem politischen Ziel, oft ohne Berücksichtigung von Unterschieden der Bildungssysteme verschiedener Länder (Münch, 2011). Dieses Ziel wurde bei dem Bolognaprozess verfolgt, welcher die Angleichung des Bildungssystems zur Verbesserung der OECD-Rankingposition anstrebte (Münch, 2011). Zuvor übertrafen allerdings deutsche Facharbeiter die US-Bachelorabsolventen hinsichtlich ihrer Qualifikation (Münch, 2011).

Universitäten reagieren als handelnde Institutionen auf die bisher beschriebenen Policies. Anhand von zwei theoretischen Modellen lassen sich die universitären Reaktionen verstehen. Das Garbagecan-Modell bietet einen Erklärungsansatz für die Anpassungsfähigkeit von Universitäten in komplexen und unsicheren Umgebungen (M. D. Cohen et al., 1972). Es beschreibt Universitäten als organisierte Anarchien, in denen Entscheidungsprozesse oft zufällig und nicht-linear verlaufen (M. D. Cohen et al., 1972). Eine Organisation ist eine organisierte Anarchie, wenn (1) Zielsetzungsproblematiken, (2) kein Verständnis über interne Prozesse und (3) eine hohe Variation des aufgebrachtten Aufwandes zwischen den handelnden Personen besteht (M. D. Cohen et al., 1972). Unter diesen Bedingungen, welche alle an Universitäten gegeben sind, sind

Probleme von den Lösungen abgekoppelt (M. D. Cohen et al., 1972). Probleme werden nicht direkt bearbeitet und es existieren bereits Lösungen für noch nicht aufgekommene Probleme (M. D. Cohen et al., 1972). Aufkommende Probleme können nur gelöst werden, wenn zufällig die richtige Kombination aus Problem, Lösung und Entscheidungsträger mit dem notwendigen Einflussbereich und Aufmerksamkeit zusammenfallen (M. D. Cohen et al., 1972). Statt expliziten Problemlösestrategien zeichnet sich eine Garbagecan-Organisation durch hohe Flexibilität aus (M. D. Cohen et al., 1972). “[A Garbagecan Organisation] does enable choices to be made and problems resolved, even when the organization is plagued with goal ambiguity and conflict, with poorly understood problems that wander in and out of the system, with variable environment and with decision makers who may have other things on their minds.” (M. D. Cohen et al., 1972, S. 16).

Neben dem Garbagecan Modell gibt es die Ressourcenabhängigkeitstheorie mit der auch Slaughter und Leslie (1999) arbeiten. Nach dieser kann die Macht, die Geldgeber über Organisationen auswirken, in zwei Dimensionen betrachtet werden (Slaughter & Leslie, 1999). Mehr Macht übt der Geldgeber aus, der im Vergleich zu anderen Geldgebern mehr Geld investiert (Slaughter & Leslie, 1999). Außerdem übt der Geldgeber mehr Macht aus, dessen Finanzierung besondere Kritikalität für die Organisation besitzt, wobei mit dem Entfernen von besonders kritischer Finanzierung die Organisation nicht mehr funktionieren würde (Slaughter & Leslie, 1999). Organisationen streben nach Stabilität und erzeugen diese durch Autonomie (Slaughter & Leslie, 1999). Ungleichgewichte und Destabilisierung der finanziellen Grundlagen führen zu Turbulenzen in der Organisation, die diese schädigen können (Slaughter & Leslie, 1999). Dabei erbringen Organisationen zusätzlichen Aufwand, um erneut Stabilität herzustellen (Slaughter & Leslie, 1999). Mit der Reduzierung kritischer, staatlicher Mittel werden Universitäten destabilisiert, wohingegen diese über industrielle Geldgeber und geänderte Ausgabemuster versuchen Stabilität herzustellen (Slaughter & Leslie, 1999). Die Industrie stellt allerdings im Gegenzug neue Anforderungen an die Universität, weshalb die Stabilität nicht vollständig erzeugt werden kann (Slaughter & Leslie, 1999). Folglich erfordert eine Verringerung von Geldmitteln aufseiten der Universität strategische Ressourcenallokation, die nur mithilfe einer mächtigeren Universitätsverwaltung möglich ist, die tief in die Organisation herein kontrollieren und wirken kann (Münch, 2011).

Aus den präsentierten Modellen geht hervor, dass Hochschulen auf die neuen wettbewerbsfördernden Policies reagieren. Die erläuterten Anpassungsprozesse führen in der Gemeinschaft aller Universitäten zu Veränderungen, wodurch ein Wissenschaftssystem entsteht, welches durch den akademischen Kapitalismus geprägt ist. Handlungen akademischer Akteure wie Universitäten, Fakultäten oder Forscher, die das Ziel haben Profit zu generieren, werden als “Akademischer Kapitalismus” bezeichnet (Slaughter & Leslie, 1999). Häufige Ziele sind das Bemühen um externe Förderungen als auch profitorientierte Aktivitäten wie Patentierungen, Spin-off-Unternehmen und Industriepartnerschaften mit Gewinnkomponente (Slaughter & Leslie, 1999). Notwendig sind entsprechende Handlungen aufgrund der verlagerten Hochschulfinanzierung von Blockzuschüssen zu projektbezogenen Mitteln (Slaughter & Leslie, 1999). Diese Verschiebung reduziert effektiv die Ressourcen für Lehre und stellt dagegen mehr Ressourcen der Forschung zur Verfügung, fordert

allerdings ebenfalls die Drittmittelakquise ein (Slaughter & Leslie, 1999). Universitäten reagierten darauf mit neuen Einnahmestrategien wie beispielsweise höheren Studiengebühren, Verträgen und privaten Spenden (Slaughter & Leslie, 1999). Das Triple-Helix-Modell beschreibt die Entstehung unternehmerischer Universitäten als Resultat der Vereinigung politischer, industrieller und universitärer Anforderungen (Baumeler, 2009). Diese Entwicklung führt zu einer verstärkten Fokussierung auf die Kommerzialisierung von Wissen, wobei akademische Kernaufgaben wie Forschung und Lehre zunehmend vom “Geist der Unternehmung” durchdrungen werden (Baumeler, 2009). Unternehmerische Universitäten legitimieren sich politisch, indem sie ihren Nutzen für die wissensbasierte Industrie hervorheben (Baumeler, 2009).

In der Folge verlagert sich die Ausrichtung der Forschung von der Anerkennung innerhalb der eigenen wissenschaftlichen Gemeinschaft hin zur Lösung industrieller und sozialer Probleme. Diese Entwicklung steht im Kontrast zu dem von Münch (2011) beschriebenen historischen Ideal der Wissenschaft. Dieses zeichnet sich durch das Streben nach Anerkennung durch Peers und Studierende, die Akzeptanz von Scheitern als Teil des Fortschritts und gegenseitige Anerkennung für die Beteiligung in der wissenschaftlichen Gemeinschaft aus. Zudem betont Münch (2011) die Bedeutung von Qualität und Originalität einzelner Beiträge gegenüber quantitativen Performanzprofilen, mit denen das Marketing der eigenen Person im Wettbewerb mit anderen Wissenschaftlern optimiert wird (Münch, 2011). Statt des Ideals werden Gemeinschaften durch Wettbewerb gespalten. In diesem Wettbewerb werden Statushierarchien durch Mechanismen der Sichtbarkeit, auf wenigen Kriterien basierende Entscheidungen und Konsekrationseffekte konstruiert (Münch, 2011). Erhalten werden die Statushierarchien durch den Austausch von Ressourcen gegen Anerkennung, Schließungseffekte durch Oligopole und den Matthäus-Effekt (Münch, 2011). Auf Kosten der universitären Gemeinschaft werden in Deutschland Exzellenzuniversitäten und nicht-universitäre Forschungszentren als Pendant zu den Ivy League Universitäten aufgebaut, um die internationale Sichtbarkeit zu erhöhen (Münch, 2011). Dies geschieht auf Kosten des Erkenntnisfortschritts, der durch einen Austausch vieler kleinerer Zentren begünstigt werden kann (Münch, 2011).

2.3 Konsequenzen des Akademischen Kapitalismus

In dem vorangegangenen Abschnitt wurde thematisiert, wie nationale Policies Wettbewerb unter Universitäten fördern, um unter anderem Sichtbarkeit zu steigern. Diese Veränderungen wirken sich auf die Beziehungen zwischen Universitäten aus und verstärken akademischen Kapitalismus. Der nächste Abschnitt betrachtet die Auswirkungen des neuen Systems auf die Individuen. Die Umsetzung der Ziele von Universitäten wie unter anderem Effizienzgewinne, eine höhere Anzahl tieferer Durchbrüche sowie erfolgreiche Absolventen werden durch einen wachsenden Kontroll- und Managementapparat überwacht (Münch, 2011). Paradox ist, dass die eigentliche Qualität von Wissenschaft nicht messbar ist. Goodhart (1981, S. 116) beschreibt, “[...] that any observed statistical regularity will tend to collapse once pressure is placed upon it for

control purposes [...]”. Dies gilt für die oft zur Bewertung von Wissenschaft genutzten Zitationsmetriken und Artikelmetadaten, welche zudem sehr selektiv sind (Münch, 2011). Mit diesen Maßen sind motivierte Handlungen jeglicher Art daher anfällig für unintendierte Nebenfolgen. Der Fokus auf Kontrolle vernachlässigt außerdem die Bildung (Münch, 2011). Um das Erreichen der zu Beginn genannten Zielmetriken zu optimieren, werden nur Studierende mit bereits hohen Kompetenzen und Forschende mit bereits einflussreicher Forschung eingestellt/angeworben. Damit wird die Ausbildung in den Hintergrund gestellt. Die Vernachlässigung äußert sich somit in der Produktion von Abschlüssen, die ohne fachliches Wissen erworben werden und setzt Lehrer-Schüler-Beziehungen unter Druck (Münch, 2011). Es werden hauptsächlich “Normalwissenschaftler” hervorgebracht, “die zu keinem kühnen Gedanken mehr fähig sind” (Münch, 2011, S. 146). Eine verstärkte Konformität erzeugt Angst vor Fehlern, welche zur Bevorzugung risikoarmer Projekte führt (Münch, 2011). Diese defensive Haltung äußert sich in der Forschungspraxis durch das Zurückhalten von Datensätzen und die Produktion minimaler Publikationen (Münch, 2011). Entsprechende Entwicklungen gefährden langfristig die Innovationskraft und Qualität der Wissenschaft und verstärken vorherrschende Wettbewerbssituationen. Insbesondere die Industrie bietet Forschern sicherere Alternativen und profitiert bei Kooperationen oder Abwerbungen selbst durch spezialisiertes Expertenwissen.

Den erhöhten Druck durch stärker begrenzte Ressourcen geben Universitäten zum Teil an ihre Angestellten weiter. In der Folge müssen diese als “staatlich subventionierte Unternehmer” agieren (Slaughter & Leslie, 1999). Die zunehmende Bedeutung der Drittmittelwerbung beeinflusst das Publikationsverhalten maßgeblich (Münch, 2011). Dazu werden Strategien entwickelt, um das eigene Publizieren an die neuen Umstände anzupassen. Immer häufiger werden Veröffentlichungen mit immer größeren Autorengruppen getätigt (Münch, 2011). Somit können Fragestellungen tiefer bearbeitet werden und mehr Artikel zur eigenen Publikationsliste hinzugefügt werden. Durch die Hierarchien zwischen Universitäten und Forschungszentren (oder Exzellenzclustern) wird an den Universitäten ein Mittelbau aufgezo-gen, der sich stärker auf Lehre konzentrieren muss und daher kaum Zeit für Publikationen hat. Wer forschen möchte, muss an die Forschungszentren, um von der Lehre befreit zu sein, was zu einer Trennung von Forschung und Lehre führt (Münch, 2011).

Auch auf der Individualebene entsteht ein Wettbewerb. Um sich in diesem durchzusetzen, müssen Wissenschaftler sich selbst bewerben. Für diese Werbung agieren sie “wie Marketingpersonen” (Ylijoki, 2003). Ein Geschichtswissenschaftler beschreibt, wie Geld durch “some kind of trickery” beschafft werden muss, bevor die eigentliche Forschung beginnen kann (Ylijoki, 2003). Die Erstellung von Demonstratoren und Prototypen als Anschauungsobjekte dienen der Legitimation gegenüber Forschungsgeldgebern und der Hochschule (Baumeler, 2009). Diese Prototypen werden meist nur aus dem Glauben angefertigt, die Industrie und restliche Gesellschaft fordere praktische Demonstration als Legitimation ein. Im von Baumeler (2009) begleiteten Projekt wurde der Demonstratorenbau von der Forschung abgekoppelt, da Ingenieurstätigkeiten eigentlich “Industriesache” sind und für die Forschung notwendige Ressourcen nicht verschwendet werden sollten. Die Beschreibung der Abkopplung deckt sich mit Aussagen anderer Universitätsangestellter aus Välimaa (2001) und Ylijoki (2003). Demnach findet ein ineffizientes, oberflächliches Marketing Einzug

in den akademischen Bereich (Välimaa, 2001). Dieses führt zur Verschwendung von Steuergeldern zu unwissenschaftlichen Marketingzwecken und einer Fokusverschiebung von Forschungsqualität zu Wettbewerbsvorteilen durch Sichtbarkeit (Baumeler, 2009; Ylijoki, 2003). Statt wissenschaftliches Arbeiten durch Markthandlungen komplett zu ersetzen, entsteht ein Dualismus, unter dem sich die Logik wissenschaftlichen Handelns aber nicht verändert (Ylijoki, 2003). Neben dem klassischen wissenschaftlichen Arbeiten muss ein großer Teil an zeitlichen und monetären Ressourcen in Sichtbarkeitserhöhende und Forschungslegitimierende Maßnahmen investiert werden. Unter den ohnehin schon verknappten Ressourcen leidet wissenschaftliche Qualität, aber insbesondere die Lehre, da diese weniger leicht sichtbar gemacht werden kann (Münch, 2011; Ylijoki, 2003). In der Lehre wird ebenfalls durch Marketing Sichtbarkeit hergestellt, wie in etwa mit der Einführung neuer “Hybridstudiengänge” verbreitet (Münch, 2011).

Unterschiedliche Disziplinen und Fachgruppierungen tragen unterschiedlich hohes Potenzial für Markthandlungen von Fakultäten und deren Angestellten. Auf diese Unterschiede soll nun im thematisch letzten Absatz eingegangen werden. Die Unterscheidung zwischen angewandter Forschung und Grundlagenforschung wird vom wachsenden Forschungsmanagement sowie der Industrie nicht beachtet (Slaughter & Leslie, 1999). In der Konsequenz wird die Grundlagenforschung häufig nicht in Fördermaßnahmen beachtet, da sie im Vergleich zur angewandten Forschung weniger Einfluss auf Industrie hat und somit weniger Potenzial für mögliche Kapitaltransaktionen besitzt. Daran wird der entscheidende Vorteil deutlich, mit dem sich marktnahe Forschung gegenüber abstrakterer Forschung durchsetzt (Slaughter & Leslie, 1999). Disziplinäre Strukturen organisieren sich im Umfeld organisationaler Komplexität in und zwischen Universitäten und Fakultäten (Becher, 1989, S. 42). Unterschiede zwischen verschiedenen Disziplinen sowie Grenzen verschiedener Disziplinen lassen sich auch an den Metadaten der produzierten wissenschaftlichen Veröffentlichungen feststellen (Andersen, 2023; Lietz, 2020). Neben den marktnahen, angewandten Naturwissenschaften sind die professionellen, quantitativen, in Journalen veröffentlichen Wissenschaften durch die neuen Finanzierungsstrategien begünstigt (Münch, 2011). So werden innerhalb der Soziologie die policyorientierte-, öffentliche- und grundlagentheoretische Soziologie langsam verdrängt (Münch, 2011). Entsprechende “Survival of the fittest” Mechanismen beschleunigen wissenschaftliche Homogenisierung. Durch Finanzierung wird extern Einfluss auf Forschungsthemen genommen. Diese externen Themensetzungen beschleunigen die Vereinheitlichung von Forschung. Insbesondere in den bereits stärker marktangepassten Naturwissenschaften lässt sich ein hoher Grad an homogenisierten Praktiken und theoretischen Ansätzen feststellen (Llanos et al., 2019). In der Soziologie kann ebenfalls eine Homogenisierung über Publikationen (Zitationsverhalten, Kollaborationen, . . .) festgestellt werden (Volle et al., 2024). In der deutschen und amerikanischen Soziologie bilden sich qualitative und quantitative Forschung mit sukzessive als unabhängige Gruppen, wobei beide Felder weniger theoretisch werden (Volle et al., 2024). Insbesondere in der deutschen Soziologie spielt Theorie eine stärkere Rolle und trägt positiv zu einer Stabilisierung einer Forschungsgemeinschaft bei (Volle et al., 2024).

2.4 Fragestellungen

Die theoretischen Überlegungen werfen die Frage auf, ob sich durch den akademischen Kapitalismus nicht nur Publikationsstrategien, sondern auch der sprachliche Ausdruck in wissenschaftlichen Veröffentlichungen ändert. Der politisch motivierte Wettbewerb führt zu neuen Handlungslogiken im Forschungsalltag. Wissenschaftler müssen ihre Arbeit verstärkt gegenüber Geldgebern aus Industrie und Politik legitimieren. Forschende konkurrieren um Positionen in spezialisierten Forschungszentren, die Vorteile bei der Beschaffung zukünftiger Projekte bieten und eine sichere Marktposition versprechen. Aus diesen Gründen stellt sich die Frage, ob positives Framing in Publikationen als Optimierung eigener Forschung auf die selbstlegitimierende Funktion genutzt wird. Dies könnte zu einer besseren Bewertung der eigenen Forschung bei Reviewprozessen, Lesern und der Industrie motivieren. Auf der anderen Seite ist der akademische Sprachstil grundsätzlich neutral und die übermäßige Wertung eigener Ergebnisse könnte irritieren und unprofessionell wirken. In dieser Arbeit soll daher geprüft werden, ob sich ein Anstieg der Häufigkeit von positivem Framing in soziologischen Publikationen finden lässt. Ein Anstieg der Häufigkeit von positivem Framing ist dann gegeben, wenn mit steigenden Jahren die Häufigkeit von Wörtern, welche die Forschung positiver darstellen, ansteigt. Um auszuschließen, dass eine grundsätzliche Steigerung von geäußerten Emotionen in wissenschaftlichen Publikationen stattfindet, wird geprüft, ob sich die Häufigkeit negativer Formulierungen über die Zeit verringert oder keinen Trend anzeigt.

RQ: Lässt sich in soziologischen Publikationen ein signifikanter Anstieg des Ausdrucks positiver Sentiments und nur ein schwacher bis kein signifikanter Anstieg negativer Sentiments feststellen?

Die Fragestellung thematisiert somit explizit keinen Zusammenhang der zuvor beschriebenen Theorie. Es wird lediglich die Existenz sprachlicher Auswirkungen in Publikationen überprüft und keine Verbindung zwischen möglicherweise auftretendem positivem Framing und Konzepten des akademischen Kapitalismus hergestellt. Der Anspruch der Arbeit und ihrer Methoden ist grundlegend explorativer, und damit induktiver Natur. Explorative Methoden haben insbesondere dann eine Berechtigung, wenn ein Phänomen unzureichend untersucht ist, da durch rein deduktive Analyse eine bestehende Theorie nicht erweitert werden kann (Stebbins, 2001). Die sprachlichen Auswirkungen der Disziplin der Soziologie befinden sich als Konzept auf der Makroebene. Da der akademische Kapitalismus ebenfalls ein Makrophänomen darstellt, eignet sich das Colemanboot um die Beziehung genauer aufzuschlüsseln (Siehe Anhang .1) (Coleman, 2000). Das Colemanboot in Abbildung 2.1 stellt die Beziehung zwischen dem akademischen Kapitalismus und den sprachlichen Veränderungen in wissenschaftlichen Publikationen dar. Dieses Schema wird mit einem zweiten Colemanboot erweitert, welches den akademischen Kapitalismus durch neue Wettbewerbsfördernde Policies erklärt. Die Fragestellung der Arbeit zielt auf die Prüfung, ob sprachliche Optimierung in Form von positivem Framing in soziologischen Publikationen existiert.

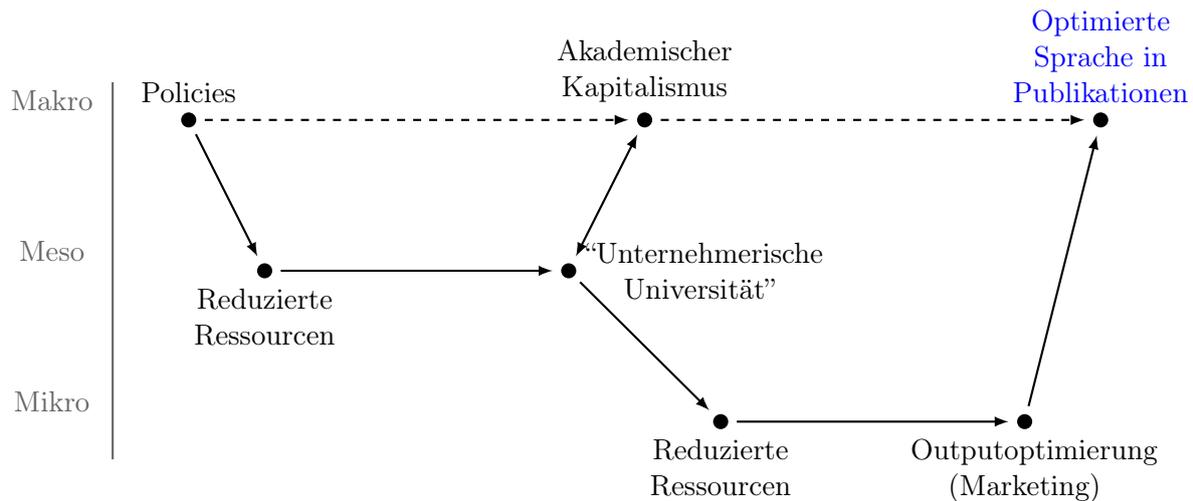


Abbildung 2.1: Colemanboot der im Kapitel beschriebenen Konzepte. Anhand der Forschungsfrage wird die Existenz eines positiven Framings (blau) untersucht.

2.5 Zusammenfassung

In diesem Kapitel wurde die Theorie des akademischen Kapitalismus und seine Auswirkungen auf die Forschung dargestellt. Zum Schluss wurde die Fragestellung der Arbeit präsentiert. Im ersten Abschnitt konnten staatliche Ziele für Wettbewerbsfördernde Gesetze sowie einzelne Ereignisse mit besonderer historischer Relevanz kurz dargestellt werden. Nun unternehmerisch agierende Universitäten reagieren auf die Verknappung von Ressourcen, indem sie versuchen Stabilität herzustellen. Dabei vermarkten sich Universitäten als problemlösungsorientiert, obwohl sie als ineffiziente Garbagecan-Organisation organisiert sind. In dem resultierenden Wettbewerb müssen sie sich durch marktorientiertes Handeln gegenüber anderen Universitäten und Forschungszentren im Kampf um Sichtbarkeit durchsetzen. Dieser Wettbewerb wird unfair durch politisch speziell geförderte Forschungszentren und führt zu ineffizienterem Erkenntnisfortschritt. Die angestellten Individuen werden vom Universitätsmanagement durch uneffektive Maßstäbe bewertet. Aufgrund der innerhalb der Universität weitergegebenen Wettbewerbsprinzipien nimmt die Forschung an risikoarmen Projekten zu. Diese Entwicklung resultiert aus dem erhöhten Risiko, dem Forschende ausgesetzt sind. Im Resultat setzen sich Forschende im Forschungswettbewerb mithilfe forschungslegitimierende Marketingstrategien gegenüber Geldgebern und Gesellschaft durch. Entsprechende Praktiken belasten die ohnehin verknappten Ressourcen zusätzlich, während die eigentliche Forschung unverändert weiterlaufen muss. Marktnähere Disziplinen und Fachgruppierungen besitzen dabei Wettbewerbsvorteile. Angewandte und empirische Forschung wird durch diese Wettbewerbsvorteile begünstigt wobei Grundlagenforschung und Theorie verdrängt werden. Extern begünstigte Themen beschleunigen die Homogenisierung von Wissenschaft, wodurch alternative und klassische Ansätze weniger Beachtung finden oder ganz verschwinden.

3 | Stand der Forschung

3.1 Einleitung

Ziel des folgenden Abschnittes ist die Darstellung bisheriger Forschung in chronologischer Reihenfolge, welche besondere Relevanz bei der Beantwortung der Forschungsfrage besitzt. Hierzu wird eine Auswahl aktueller und klassischer Studien getroffen, um innerhalb des Themas besonders innovative Ansätze zu beleuchten. Der Abschnitt beginnt mit der Arbeit von Vinkers et al. (2015), welche nahezu alle nach ihr folgenden Studien inspirierte und deren Methode und insbesondere das genutzte Diktionär in einer Vielzahl vergleichbarer Studien verwendet wurden. Anschließend folgt die Präsentation der Studie von Holtz et al. (2017). Das Thema der Analyse ist vergleichbar mit dem Inhalt von Vinkers et al. (2015), allerdings wird dieser Ansatz um zwei neue Fragestellungen erweitert. Die zusätzlichen Fragestellungen erlauben eine größere Nähe der Methodik zur Theorie, wodurch die Schlussfolgerungen fundierter begründet werden können (Holtz et al., 2017). Im weiteren Verlauf erfolgt eine Auseinandersetzung mit Lennox et al. (2020). Diese Studie validiert die Erkenntnisse von Vinkers et al. (2015) durch einen innovativen Triangulationsansatz. Dieser ist aufgrund seiner häufigen Adaption (Wen & Lei, 2022; Yuan & Yao, 2022) relevant. Zuletzt stellt der Abschnitt die Untersuchung von Liu und Zhu (2023) vor. Diese ist die einzige Studie, in der unter anderem eine Sentimentanalyse von Abstracts aus der Disziplin der Soziologie durchgeführt wurde, weshalb die Arbeit nicht nur methodisch, sondern auch inhaltlich für die Fragestellung dieser Thesis relevant ist (Liu & Zhu, 2023).

3.2 Sentiments in PubMed-Abstracts

Die Studie von Vinkers et al. (2015) untersucht die Häufigkeit positiver, negativer und neutraler Wörter in wissenschaftlichen Abstracts. Motiviert wird die Untersuchung durch vergangene Erfahrungen der Übertreibung wissenschaftlicher Ergebnisse (Vinkers et al., 2015). Eine Kultur der Produktivität gilt dabei als mögliche Ursache für die Nutzung von positiven und negativen Wörtern (Vinkers et al., 2015). Das verwendete Diktionär ist durch Diskussionen der Forschenden, Analysen zufälliger Abstracts und einem Abgleich mit Thesaurus-Listen erstellt worden (Vinkers et al., 2015). Zusätzliche Validierung des Diktionärs kann durch das Hinzufügen von Wörtern aus der Studie McCarthy (2015) erreicht werden. Neben dem Diktionär kommt eine weitere Wortliste zum Einsatz, welche aus jeweils 50 zufällig gezogenen Adjektiven und Nomen besteht (Vinkers et al., 2015). Die zufälligen und neutralen Wörter helfen dabei im Vergleich mit den Ergebnissen des Diktionärs die Validität der Ergebnisse zu erhöhen (Vinkers et al., 2015). Das Diktionär dient im gesamten Prozess der Durchsichtung aller Abstracts der PubMed Datenbank im Zeitraum von 1974 bis 2014 (Vinkers et al., 2015). Die Studie teilt die Anzahl an Abstracts, welche mindestens ein positives, negatives oder neutrales Wörter enthalten, durch die Anzahl veröffentlichter Abstracts pro Jahr (Vinkers et al., 2015). Mit einer Grafik, welche den

Anstieg der individuellen Wörter des Diktionärs zeigte, wird untersucht, inwiefern die Ergebnisse abhängig von einzelnen Wörtern sind. Das Ergebnis ist ein entdeckter Anstieg von 880% für Abstracts mit mindestens einem positiven Wort (Vinkers et al., 2015). Insbesondere sind die Wörter “robust”, “novel”, “innovative” und “unprecedented” für den Anstieg der Häufigkeit positiv markierter Abstracts verantwortlich (Vinkers et al., 2015). Dieser Effekt kann jedoch ebenso ohne die Beachtung der aufgelisteten Wörter nachgewiesen werden (Vinkers et al., 2015). Auch für die Abstracts, welche mindestens eines der negativen Wörter enthielten, steigt die Häufigkeit über den Zeitraum an (Vinkers et al., 2015). Dieser Anstieg ist im Vergleich zu den positiven Abstracts wesentlich schwächer (Vinkers et al., 2015). Für die neutralen und zufälligen Wörter kann kein Anstieg festgestellt werden (Vinkers et al., 2015).

3.3 Interkulturelle Psychologie und Akademischer Kapitalismus

Ziel von Holtz et al. (2017) ist zu ermitteln, inwieweit Selbstvermarktung zur Logik wissenschaftlichen Schreibens in der Kulturübergreifenden Psychologie geworden ist. Die Studie untersucht, ob es eine Änderung der positiven und negativen Sentiments gab und ob die Häufigkeit von Phrasen zugenommen hat, mit denen marginal signifikante Ergebnisse beschrieben werden (Holtz et al., 2017). Darüber hinaus analysiert Holtz et al. (2017) sprachliche Eigenschaften des sozialwissenschaftlichen Schreibstils, welche auf die Kritik von Billig (2013) zurückgehen. Als Datenbasis verwenden Holtz et al. (2017) Volltexte und Abstracts, welche zwischen den Jahren 1974 und 2014 in zwei Journalen der Kulturübergreifenden Psychologie veröffentlicht wurden (Holtz et al., 2017). Für die Untersuchung positiver und negativer Sentiments werden die positiven und negativen Wörtern aus der Studie von Vinkers et al. (2015) genutzt (Holtz et al., 2017). Zusätzliche Validierung der Ergebnisse geschieht durch die Hinzunahme des umfangreicheren LIWC Diktionärs (Holtz et al., 2017). Für die zweite Fragestellung wurde ein eigenes Diktionärs mit Phrasen erstellt, die dazu genutzt werden marginal signifikanter Ergebnisse zu publizieren (Holtz et al., 2017). Die Analyse des Schreibstils umfasst das Zählen von Wörtern mit mindestens sechs Buchstaben, häufigen Verben und Nominalisierungen sowie die Auszählung der Phrase “more research is needed” (Holtz et al., 2017).

In den Abstracts und Volltexten beider untersuchten Journale zeigt sich ein Anstieg der relativen Häufigkeit positiver Sentiments im Verhältnis zur Gesamtwortzahl über die Zeit (Holtz et al., 2017). Die relativen Häufigkeiten negativer Wörter weisen keinen Trend oder einen Abwärtstrend auf (Holtz et al., 2017). Die relative Häufigkeit von Artikeln mit Phrasen zur Beschreibung marginal signifikanter Ergebnisse stieg zeitlich in beiden untersuchten Journalen (Holtz et al., 2017). Ein, nach Billig (2013), schlechterer wissenschaftlicher Stil nimmt in beiden Journalen über die Zeit zu. Jedoch zeigt sich in nur einem der beiden Journale ein klarer Anstieg von Nominalisierungen (Holtz et al., 2017). Ein verstärkter akademischer Kapitalismus könnte die gemeinsame Ursache des Anstiegs der gemessenen Konzepte sein (Holtz et al., 2017).

3.4 Sentiment Analyse in der Naturschutzforschung

Das zentrale Anliegen der Untersuchung ist die Erforschung zeitlicher Trends von Sentiments in Publikationen der Naturschutzforschung (Lennox et al., 2020). Die Studie überprüft die Vermutung, wie sich negative Entwicklungen im Naturschutz in Form von verstärkt negativen Sentiments in der Naturschutzforschung zeigen (Lennox et al., 2020). Für die Beantwortung der Vermutung wurden Abstracts von Artikeln aus 6 viel zitierte Journalen, die zwischen 1998 und 2017 veröffentlicht worden sind, analysiert (Lennox et al., 2020). Eine Klassifizierung der Artikel danach, ob in Abstracts oder Titeln bedrohte Tierarten genannt werden ermöglichte eine detailliertere Analyse der Sentimentverteilung (Lennox et al., 2020). Die Sentimentanalyse wurde mit den Sentimentdictionären Bing, NRC, AFINN und einem Sentimentalgorithmus von Jockers (2015) durchgeführt (Lennox et al., 2020). Alle Sentimentwerte wurden als Pole einer Skala verwendet (Lennox et al., 2020). Eine mit dem jeweiligen Mittelwert addierte Z-Standardisierung der jeweiligen Diktionärsergebnisse ermöglicht den Vergleich der Ergebnisse zwischen den Diktionären (Lennox et al., 2020). Die Summierung der standardisierten Werte pro Jahr ergab einen durch alle Diktionäre informierten Sentimentwert (Lennox et al., 2020). Eine Bereinigung der Diktionäre auf Wörter mit domänenspezifischen Bedeutungen wie "Shark" oder "parasite" schließt mögliche Verzerrungen weiter aus (Lennox et al., 2020). Alle Sentimentdiktionäre korrelierten schwach, aber signifikant miteinander. Der aggregierte Sentimentwert zeigte einen schwachen, aber signifikanten Anstieg mit steigenden Jahren (Lennox et al., 2020). Abstracts, welche ausgestorbene Spezies nannten, besaßen die niedrigsten Sentimentwerte (Lennox et al., 2020). Entgegen der eingangs beschriebenen Erwartung scheint ein Trend der optimistischen Naturschutzforschung an Gewicht zu gewinnen, in dem der Fokus auf Erfolgsgeschichten und Fortschritte gelegt wird (Lennox et al., 2020).

3.5 Positivität in Wissenschaftsdisziplinen

Liu und Zhu (2023) nutzten als Ausgangspunkt ihrer Studie, dass es durch kompetitive Auswahlmechanismen für Forschende wichtiger wird ihre Forschung zu bewerben und positive Sprache in Veröffentlichungen ein Werkzeug dazu ist. Der interdisziplinäre Vergleich positivem Framings war neuartig und schließt eine damals existierende Forschungslücke (Liu & Zhu, 2023). Untersucht wird unabhängig und gemeinsam die Veränderung positiven Framings über die Zeit, zwischen Disziplinen und der Zusammenhang von positivem Framing und Forschungseinfluss (Liu & Zhu, 2023). Der verwendete Datensatz enthält Abstracts von Artikeln, die zwischen 1997 - 2019 auf Web of Science in den Disziplinen Kommunikation, Linguistik, Politikwissenschaft, Soziologie und Luft- und Raumfahrt, Automatisierung und Steuerung, Softwaretechnik, Transportation veröffentlicht wurden (Liu & Zhu, 2023). Die ersten vier Disziplinen sind Vertreter der weichen Wissenschaften und die letzten vier der harten Wissenschaften (Liu & Zhu, 2023). Die positiven und negativen Wörtern des Diktionärs von Vinkers et al. (2015) bildeten die Basis der Untersuchung (Liu & Zhu, 2023). Zusätzlich wurde die Sentimentsoftware von Jockers (2015) verwendet,

um die Ergebnisse reliabler zu gestalten (Liu & Zhu, 2023). Die Normalisierung der Artikelzitationen geschah mit der Zitationssumme aller Artikel des jeweiligen Jahres und der Summe aller Artikel im gesamten Zeitraum.

In allen Disziplinen gab es einen signifikant positiven Anstieg positiver Sentiments, während die Ergebnisse für die negativen Wörter gemischerter waren (Liu & Zhu, 2023). Die relative Häufigkeit negativer Wörter zeigte nur in der Disziplin Luft- und Raumfahrt, Linguistik, Politikwissenschaft und Soziologie einen leichten signifikanten negativen Trend, in allen anderen Disziplinen konnte kein Trend gezeigt werden (Liu & Zhu, 2023). Insgesamt ergaben sich für die harten Wissenschaften positivere Sentiments als für die weichen Wissenschaften (Liu & Zhu, 2023). Zwischen den normalisierten Zitationszahlen und den standardisierten positiven Sentiments gab es eine sehr kleine, aber dennoch signifikante positive Korrelation (Liu & Zhu, 2023). Die negativen Sentiments korrelierten nicht mit den Zitationszahlen (Liu & Zhu, 2023).

3.6 Zusammenfassung

Dieses Kapitel präsentierte vier Schlüsselstudien, die für die Untersuchung des Anstiegs positiver Sentiments in soziologischen Abstracts bedeutsam sind. Vinkers et al. (2015) entwickelten ein grundlegendes Diktionär zur Sentimentanalyse wissenschaftlicher Abstracts. Ihre Methodik, insbesondere das erstellte Wörterbuch, wurde in zahlreichen nachfolgenden Studien aufgegriffen und bildet auch die Basis für die vorliegende Thesis. Die Arbeit von Holtz et al. (2017) demonstriert, wie das Diktionär aus Vinkers et al. (2015) mit einem umfangreicheren Wörterbuch (LIWC) verglichen werden kann, um die Validität der Ergebnisse zu erhöhen. Dieser Ansatz verwendet relative Häufigkeiten für beide Diktionäre und wird für die geplante Untersuchung soziologischer Abstracts dieser Arbeit besonders relevant. Eine alternative Vorgehensweise zur Integration verschiedener Diktionäre innerhalb einer Analyse verwendet Lennox et al. (2020). Durch die Summe standardisierter Sentimentwerte unterschiedlicher Diktionäre lassen sich verschiedene Sentimentdiktionäre direkt miteinander integrieren. Ihre Methode zur Standardisierung und Aggregation von Sentimentwerten bietet wertvolle Anregungen für die Analyse mit mehreren Diktionären. Die Studie von Liu und Zhu (2023) ist von besonderer Bedeutung, da sie als einzige auch soziologische Abstracts untersucht. Ihre Ergebnisse liefern wichtige Vergleichswerte und methodische Ansätze für die Analyse positiver Sentiments in der Soziologie. Alle vorgestellten Studien bilden zusammen eine solide methodische und inhaltliche Grundlage für die geplante Untersuchung der Fragestellung.

4 | Methode

4.1 Einleitung

Dieses Kapitel widmet sich der Datengrundlage und den methodischen Ansätzen, die dieser Arbeit zugrunde liegen. Der erste Abschnitt beschreibt die praktischen Hintergründe der Sentiment- und Diktionärsanalyse. Beide bilden im späteren Verlauf den Kern der Analyse. Im Anschluss wird die Eingrenzung des Feldes “Soziologie” begründet sowie das Vorgehen der Datenbeschaffung beschrieben. Eine genaue Betrachtung des Suchstrings und der Möglichkeiten zur Ergebnisfilterung folgt. Grundlage der Analyse bilden die positiven Framings sowie die Konstruktion der relativen Werte, welche in diesem Kapitel 4 operationalisiert werden. Der Ansatz der Analyse besteht zentral aus der Untersuchung der aggregierten Mittelwerte pro Jahr und dem Vergleich der Ergebnisse beider Diktionäre. Zur Beantwortung der zentralen Fragestellung wird für die Analyse ein besonderer Fokus auf die Besonderheiten der Interpretation statistischer Signifikanz und die Verwendung inferenzstatistischer Methoden, im Kontext der hohen Fallzahl, gelegt. Weitergehend werden mögliche Trendverläufe durch lineare OLS-Regressionen geprüft. Abschließend folgt eine Zusammenfassung des Kapitels.

4.2 Sentiment- und Diktionärsanalysen

Computergestützte Textanalysen basieren auf linguistischen Konzepten. Ein Korpus aus Texten oder Dokumenten bildet die Datenbasis (Grimmer & Stewart, 2013). Diese Texte können zum Beispiel Tweets, Bücher, Kapitel oder wissenschaftliche Abstracts darstellen. Die Unterteilung von Texten in Tokens als “semantisch bedeutungsvolle Einheiten” stellt die Grundlage dieser Analysen dar (Silge & Robinson, 2017). Tokens werden in der Regel als Wörter festgelegt, können aber je nach Fragestellung auch Morpheme oder Paragraphen sein. Die Diktionärsanalyse ist eine Methode zur Textklassifikation (Grimmer & Stewart, 2013). Zur Klassifikation werden die Häufigkeiten vorkommender Tokens in den jeweiligen Texten ausgezählt. Tokens werden zuvor bestimmten Schlüsseln, die Kategorien repräsentieren, zugeordnet. Einem Textdokument werden am Ende Häufigkeiten der auftretenden Tokens nach Schlüsseln gruppiert zugewiesen (Ribeiro et al., 2016). Die Qualität der Ergebnisse hängt demnach maßgeblich von der Güte des verwendeten Diktionärs ab. Diktionäre werden immer für bestimmte Textdomänen angelegt, denn die Bedeutung der Wörter ändert sich je nach Textdomäne. Sentimentanalysen zielen darauf ab, Gefühle in Texten mittels computergestützter Methoden zu identifizieren (Thelwall, 2022). Die Klassifikationsansätze variieren zwischen dichotomen (z.B. positiv/negativ) und mehrstufigen Skalen (z.B. von stark negativ bis stark positiv). Polytome Klassifikationen erfassen eine breitere Emotionspalette, indem sie ebenfalls beispielsweise Trauer oder Überraschung als Emotionen klassifizieren (Mohammad & Turney, 2013).

Die Diktionsanalyse ist durch die fehlende Berücksichtigung von Kontext, wie etwa Ironie sowie die Abhängigkeit von vordefinierten Wortlisten limitiert. Erweiterte Ansätze wie in etwa Klassifikations-Transformermodelle, die auf maschinellem Lernen basieren, versuchen diese Einschränkungen teilweise zu kompensieren. Der Vorteil von Diktionsklassifikationen gegenüber Methoden des maschinellen Lernens ist die höhere Transparenz und Nachvollziehbarkeit der Klassifikation, da für jedes Wort die genauen Schlüssel und Häufigkeiten eingesehen werden können. Für Sentimentanalysen muss begründet werden können, warum ein Text geäußerte Sentiments enthalten sollte. Eine Besonderheit wissenschaftlicher Texte liegt darin, dass diese vorwiegend neutral formuliert sind. Wörter, die normalerweise für Sentimentausdrücke verwendet werden, haben im wissenschaftlichen Kontext häufig andere oder anderswirkende Bedeutungen. Aufgrund dessen ist der zentrale Ansatz dieser Untersuchung, die Bewertung der jeweiligen Forschung durch die AutorInnen zu erfassen. Beispielsweise könnten Forschende ihre Ansätze als "creative" beschreiben. Die Sentimentanalyse stellt somit eine Art Bewertungsanalyse dar. Bei einer Bewertung werden Urteile gefällt, bei denen Forschende einen Sachverhalt deuten und sich gegenüber diesem positionieren (positiv, negativ). Methodisch unterscheidet sich die Datenaufbereitung einer korpusbasierten Analyse von den Aufbereitungstechniken numerischer Daten. Üblicherweise werden alle Tokens in ihre kleingeschriebene Version transformiert, um auch Wörter am Satzanfang zu berücksichtigen. Ausgehend von einer Bag-of-Word Analyse werden in der Regel Stoppwörter entfernt. Bei dieser Analysemethode spielt die Reihenfolge der Tokens keine Rolle, wobei die entfernten Stoppwörter ("ä", "än", "be") die Textbedeutung nicht wesentlich beeinflussen. Zudem erfolgen häufig die Löschung von Satzzeichen und Nummern sowie eine Trennung von Bindestrich-Wörtern. Die angewendete Datenaufbereitungs-Technik der Untersuchung wird in Kapitel 5 erläutert.

4.3 Daten

Die Analyse basiert ausschließlich auf Daten aus der Web of Science Datenbank (Clarivate Analytics, 2024). Neben den formalen Metadaten wie Autoren, Publikationsjahr, Titel, Ort und DOI sind über Web of Science auch inhaltliche Metadaten wie beispielsweise Abstracts zur Verfügung gestellt. Die Disziplin der Soziologie wird mithilfe der von Web of Science bereitgestellten Journalkategorien eingegrenzt. Zur Disziplin der Soziologie zählen demnach alle Artikel, welche in einem Journal veröffentlicht wurden, welches von Web of Science zur Journalkategorie "Sociology" zugeordnet wird. Alle Journals dieser Kategorie sind Teil des "Social Sciences Citation Index" oder des "Emerging Sources Citation Index" (Clarivate Analytics, n. d. a). Beide Zitationsdatenbanken sind Teil der "Web of Science Core Collection" (Clarivate Analytics, n. d. b). Das Verfahren zur Kategorisierung von Journals durch Web of Science ist dokumentiert (of Science Group, 2019). Die Kategorisierung umfasst vier Stufen bei dem 28 Kriterien getestet werden (of Science Group, 2019). Trotz Kritik von Milojević (2020) und Norris und Oppenheim (2007) gewähren die Journalkategorien tiefe Einblicke innerhalb der Wissenschaftsforschung (Boyack et al., 2005; Leydesdorff et al., 2013). Kritisiert wird unter anderem, dass themenbasierte Klassifizierungen akkurater

als Journalbasierte seien und multidisziplinäre Forschung, sowie in Büchern veröffentlichte Forschung nicht abgebildet werden könnte (Milojević, 2020; Norris & Oppenheim, 2007). Dennoch können über themenklassifizierung durch Journale Überblicke über die Wissenschaftslandschaft gewonnen werden (Boyack et al., 2005; Leydesdorff & Rafols, 2009). Alternative Ansätze um die Grenzen von Wissenschaftsfeldern zu bestimmen finden sich innerhalb der Netzwerkforschung und bauen auf Zitationsnetzwerken auf (Lietz, 2020; Mejia et al., 2021). Heruntergeladen werden Artikel, die innerhalb des Zeitraums 1974-2023 veröffentlicht wurden. Grundlage für diese Entscheidung bildete die in Kapitel 3 vorgestellte Literatur. Bei späteren deskriptiven Analyse des Datensatzes ergab sich jedoch, dass nur Publikationen nach 1993 über Abstractdaten verfügen. Aufgründessen konnte für die Analyse lediglich der Zeitraum 1993-2023 beachtet werden. Mit dem folgenden Suchstring wurde die Web of Science Core Collection über die “Advanced Query” durchsucht und alle Artikel heruntergeladen:

$$(WC=(Sociology)) \text{ AND } DOP=(1974-01-01/2024-01-01)$$

Die “Fast5000” Download Option über die Webanzeige von Web of Science erlaubte das Abspeichern der Metadaten in Abschnitten von 5000 Artikeln. Web of Science blockiert das manuelle Herunterladen von mehr als 100000 Artikeln pro Suche. Um diese Blockade zu umgehen wurde die Suche in vier Zeitblöcke unterteilt damit die Anzahl der Gesamtergebnisse pro Suche unter 100000 gehalten werden kann. Die gewählten Zeitabschnitte waren: 1970-1985, 1986-2000, 2001-2014, 2015-2023. Insgesamt konnten so Metadaten zu 345.019 Artikeln aus insgesamt 309 Journals heruntergeladen werden. Methodisch stellt der Datensatz eine Vollerhebung der Artikelmetadaten aller Artikel dar, die zwischen 1993 und 2023 in Journals mit der Web of Science Journalkategorie “Sociology” auf Web of Science veröffentlicht wurden. Die textcat-Erweiterung für R kategorisierte die gespeicherten Abstracts nach Sprache (Hornik et al., 2023). Durch das anschließende Filtern aller nichtenglischsprachigen Artikel werden sprachbasierte Fehlklassifizierungen der englischsprachigen Diktionäre vermieden. Es kann davon ausgegangen werden, dass die textcat-Erweiterung reliable Ergebnisse liefert, da sie denselben N-Gram basierten Klassifikator aus sehr bekannten Programmen wie beispielsweise LibreOffice verwendet (Cavnar & Trenkle, 2001; McNamara et al., 2024).

Um die Wahl von Web of Science als Datenbank zu testen, wurde ein Testdatensatz über OpenAlex heruntergeladen. (Priem et al., 2022). OpenAlex ermöglichte mit denselben Disziplinären-Einschränkungen trotz größerer Trefferzahlen Zugriff auf weniger Abstracts. Das spricht nicht nur für eine kleinere Datenmenge, sondern auch eine unschärfere Auswahl der Abstracts. Die Analyse basiert demnach auf der Datenbank Web of Science, da diese Zugang zu einer höheren absoluten Anzahl an Abstracts bei einer insgesamt kleineren Artikeltrefferzahl ermöglichte.

4.4 Operationalisierung

Um positive und negative Bewertungen der eigenen Forschung Abstracts zu erfassen, verwende ich den grundsätzlichen Ansatz diktionärsbasierter Sentimentanalysen. Die Analyse wird mit den positiven und negativen Begriffen aus des Diktionärs von Vinkers et al. (2015) durchgeführt. Die Ergebnisse werden zusätzlich mit der Open Source Sentimentsoftware VADER validiert (Hutto & Gilbert, 2014). Das Diktionär von Vinkers et al. (2015) ist speziell für die Analyse wissenschaftlicher Artikel erstellt worden, beläuft sich im Umfang allerdings lediglich auf 50 Wörter (Siehe Anhang .2). Die Ergebnisse sind daher anfällig für Änderungen der Häufigkeiten einzelner Wörter. Vinkers et al. (2015) stellen selbst fest, dass ihr nachgewiesener Anstieg positiver Wörter maßgeblich von nur 4 Worten “robust”, “novel”, “innovative” und “unprecedented” verursacht wurde. VADER schneidet im Vergleich zu anderen etablierten Sentimentdiktionären in verschiedenen Vergleichen als eines der akkuratesten Diktionären ab (Hutto & Gilbert, 2014; Ribeiro et al., 2016). Obwohl VADER nicht wie Vinkers et al. (2015)’s Diktionär speziell auf für die Analyse wissenschaftlicher Texte erstellt wurde kann VADER auch für wissenschaftliche Texte valide Ergebnisse erzeugen. VADER’s Algorithmus implementiert die Beachtung von Punkt und Zeichensetzung, Groß- und Kleinschreibung sowie Wörter die Sentiments Invertieren, verstärken oder abschwächen (Negationen, “but”, “very”, etc.) (Edlinger et al., 2023; Hutto & Gilbert, 2014). Die Validierung wird durch die Kontrolle der Interraterreliabilität und anschließender Kontrolle der jährlichen Übereinstimmung durchgeführt.

Die Diktionärsanalysen werden mit dem Textanalyseframework `quanteda`, welches für die Statistiksprache R entwickelt wurde, durchgeführt (Benoit et al., 2018; {R Core Team}, 2024). `Quanteda` ermöglicht das Erstellen und Anwenden eigener Diktionäre. Über diese Funktion wurde das Diktionär von (Vinkers et al., 2015) in die Analyse integriert. VADER besitzt ebenfalls eine Schnittstelle für die Verwendung in R (Hutto & Gilbert, 2014). Die Übersicht der Häufigkeiten aller Wörter des Diktionärs von Vinkers et al. (2015) wurde aufgrund der echtzeitkompilierenden Vorteile mit Julia erstellt (Bezanson et al., 2017). Alle über die Analyse der Jahresmittelwerte der Sentiments hinausgehenden Verfahren werden aufgrund der arbeitsspeichersparenderen Umgebung in Python statt R durchgeführt (Van Rossum et al., 2009) .

Beim Anwenden beider Diktionäre wird jedem Wort eines Abstracts eine von drei Kategorieausprägungen zugewiesen. Die Kategorieausprägungen $K = \{k_1, k_2, k_3\}$ jeder Sentimentanalyse sind $K_{\text{vinkers}} = \{\text{positiv, negativ, unmatched}\}$ und $K_{\text{VADER}} = \{\text{positiv, negativ, neutral}\}$ (Hutto & Gilbert, 2014; Vinkers et al., 2015). Für alle Abstracts werden relative Häufigkeiten der absoluten Häufigkeiten jeder Kategorie berechnet.

$$h(k_i) = \frac{H(k_i)}{H(k_1) + H(k_2) + H(k_3)}$$

VADER gewichtet die Intensität der Sentiments pro Wort allein zwischen -4 und +4 (Hutto & Gilbert, 2014). Diese Sentiments werden erneut gewichtet auf Basis von Sentimentintensitätsändernden Wörtern wie “very”, “but”, “not” oder “!”. Die final berechneten relativen Häufigkeiten

geben für jedes Abstract einen Wert der Positivität bzw. Negativität eines Sentiments an. Das so durch die Wortzahl eines Abstracts normalisierte Maß erlaubt den Vergleich von auf Jahre aggregierten Sentimentwerten.

4.5 Analyseansatz

Die Darstellung von Mittelwerten und Konfidenzintervallen der jährlichen Sentimentanteile in Intervallplots bildet den Ausgangspunkt der Datenanalyse. Diese Visualisierung aggregierter Daten ermöglicht einen ersten Einblick in die zeitliche Entwicklung der Sentimentanteile. Bei den beschriebenen Abbildungen spielt die Darstellung des Konfidenzintervalls eine wichtige Rolle. Die signifikante Abweichung vom Mittelwert zeigt eine Überzufälligkeit an. Im Folgenden wird knapp begründet, ob inferenzstatistische Methoden bei nahezu Vollerhebungen legitimiert werden können. Im Gegensatz zu deskriptiven, rein beschreibenden Statistiken erlauben Inferenzstatistische Modelle die Erklärung eines Effekts durch die Aufteilung in einen systematischen und einen stochastischen Teil (Broscheid & Gschwend, 2005). Beispielsweise wird bei einer linearen Regression der stochastischen “Zufallsfaktor” ϵ als reiner Zufall oder Sammelmaß aller nichtbeachteten Einflussfaktoren aufgefasst (Broscheid & Gschwend, 2005). In dem Fall, dass die Annahmen des Modells verletzt werden stehen zwei Handlungsoptionen zur Verfügung. Entweder können “Verletzungen der Annahmen als Petitessen betrachtet und statistisch korrigiert werden” oder “diagnostizierte Verletzungen der Modellannahmen über die ϵ_i als substantiell interessant betrachtet werden.” (Broscheid & Gschwend, 2005, S. 20). Die weitergehende statistische Analyse der stochastischen Abweichungen ermöglicht den Gewinn neuer Erkenntnisse mit welchen das statistische Modell und ebenfalls die Theorie verbessert werden kann. Auch wenn diese Ziele unabhängig von der Stichprobengröße funktionieren ist zu beachten, dass bei besonders großen Fallzahlen der Standardfehler artifiziiell verkleinert wird:

$$\lim_{n \rightarrow \infty} \frac{\sigma}{\sqrt{n}} = \sigma \cdot \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} = \sigma \cdot 0 = 0 \quad (4.1)$$

Bei besonders großen Fallzahlen werden somit vernachlässigbare Effekte artifiziiell statistisch signifikant (Tabachnick & Fidell, 2014). Die Nichtinterpretierbarkeit der statistischen Signifikanz erschwert die Abschätzung der tatsächlichen Relevanz des Effektes.

Im Anschluss an die Betrachtung der aggregierten Sentiments werden die Ergebnisse beider Diktionäre miteinander verglichen. Die Übereinstimmung der Ergebnisse beider Diktionäre wird mit einer Pearsonkorrelation überprüft. Dies ermöglicht einen Schluss auf die Effektivität der zusätzlichen Validierung durch das VADER Diktionär. Die Pearsonkorrelation ist mit ihren Eigenschaften als justiertes Maß für Interraterübereinstimmungen für den Vergleich gut geeignet (Lehrstuhl Psychologische Methodenlehre, SS24). Justierte Maße integrieren Konsistenz aber nicht die Strenge der Bewertungen. Beide Diktionäre bewerten aufgrund ihrer unterschiedlich großen Wortlisten und unterschiedlichen Beachtung der Umgebungsterme eines Satzes auf unterschiedlichen Skalen. Eine Beachtung der Strenge würde daher zu bedeutungslosen Ergebnissen führen.

Ergänzend zur Korrelation wird für den direkten Vergleich die Differenz der z-standardisierten positiven Sentimentanteile pro Jahr sowie die Differenz der z-standardisierten negativen Sentimentanteile pro Jahr berechnet. Die positiven pro Jahr aggregierten Sentimentanteile beider Diktionäre sind allerdings fehlerbehaftete Werte. Die Fehlerfortpflanzung nach Gauß (4.2) ermöglicht die Berechnung des Standardfehlers der Differenz beider Werte. Die Standardfehler der zu vergleichenden Sentimentanteile bilden die Basis dieses Verfahrens (Tipler & Mosca, 2015). Der theoretische Hintergrund des Verfahrens ist, dass sich statistische Fehler im Vergleich zu systematischen Fehlern gegenseitig aufheben (Tipler & Mosca, 2015).

$$\Delta Y = \sqrt{\sum_{i=1}^m \left(\frac{\partial Y}{\partial X_i} \Delta X_i \right)^2} \quad (4.2)$$

Dabei ist $\Delta Y = \frac{\sigma}{\sqrt{n}}$ der Standardfehler. Das Einsetzen der Sentimentdifferenz in Gleichung 4.2 liefert das Ergebnis 4.3 (Siehe Anhang .4). Dabei sind s_{vader} und s_{vink} die Sentimentanteile der Abstracts.

$$d_s = s_{\text{vader}} - s_{\text{vink}} \quad (4.3)$$

$$\Delta d_s = \sqrt{\Delta s_{\text{vader}}^2 + \Delta s_{\text{vink}}^2} \quad (4.4)$$

Die lineare Regression nach dem OLS-Verfahren ermöglicht die Spezifikation möglicher Trendverläufe, die durch reine Darstellung der Mittelwerte nicht sicher bestimmt werden können. Die Anwendung eines parametrischen Verfahrens wird durch die Formulierungen von vermuteten Trendverläufen im Anschluss an die Analyse der aggregierten Daten legitimiert. Ebenfalls beeinflusst die hohe Stichprobengröße maßgeblich die Interpretation der statistischen Signifikanz in der linearen Regressionsanalyse. Lin et al. schlugen 2013 vor, mithilfe von Montecarlo-Simulationen sogenannte CPS-Charts (**C**oefficient/**P**-value/**S**amplesize) zu erstellen. In diesen Grafiken wird der p-Wert in Abhängigkeit der Fallzahl aufgetragen. Dazu werden bei $n = 0$ beginnend größer werdende Stichproben gezogen, für jede das lineare Modell geschätzt und der p-Wert abgespeichert (Lin et al., 2013). Dies ermöglicht eine transparente Darstellung der p-Wert Verzerrungen durch die Stichprobengröße.

4.6 Zusammenfassung

Der erste Abschnitt des Kapitels beschreibt die methodischen Hintergründe einer Diktionärsbasierten Sentimentanalyse. Diese ordnet mit Diktionären einzelne Tokens bestimmten Kategorien zu um Textdokumente in Emotions oder Bewertungsschemata zu klassifizieren. Die Qualität der Ergebnisse hängt dabei von der Güte und Kongruenz des verwendeten Diktionärs zur Textdomäne

ab. Als Datengrundlage für die Analyse wurden 345.019 Abstracts aus Publikationen heruntergeladen, die zwischen 1974 und 2023 in Journals mit der Web of Science Kategorie “Sociology” erschienen sind. Die Abstracts werden nach Sprache gefiltert, sodass nur die englischsprachigen Abstracts bestehen bleiben. Zur Operationalisierung positiver und negativer Bewertungen werden das Diktionär von Vinkers et al. (2015) sowie VADER verwendet. Dieses Vorgehen erhöht die Validität der Ergebnisse. Für beide Diktionäre werden die absoluten Häufigkeiten der in den Texten erkannten Tokens in relative Häufigkeiten zur Wortanzahl des jeweiligen Abstracts umgerechnet. Die Summe der relativen Sentiments aller in einem Jahr veröffentlichten Abstracts ermöglicht den Vergleich zwischen verschiedenen Jahren. Der Analyseansatz umfasst die Darstellung von Mittelwerten und Konfidenzintervallen jährlicher Sentimentanteile. Im dazugehörigen Abschnitt wurde außerdem auf die Bedeutung von Inferenzstatistik im Kontext großer Fallzahlen eingegangen. Besonders große Fallzahlen erhöhen die Statistische Signifikanz artifiziell, was die Abschätzung der tatsächlichen Relevanz erschwert. Eine Pearson Korrelation der Ergebnisse beider Diktionäre ermöglicht eine Einschätzung der Ergebnisvalidität. Zusätzlich wird die Differenz zwischen den positiven Sentiments beider Diktionäre und den negativen Sentiments beider Diktionäre berechnet. Die lineare Regression nach dem OLS-Verfahren ermöglicht die genauere Untersuchung vermuteter Trendverläufe.

5 | Ergebnisse

5.1 Einleitung

Das folgende Kapitel präsentiert die Analyseergebnisse. Zu Beginn wird die durch deskriptive Analysen gestützte Datenbereinigung erklärt und ein erster Eindruck der Datenstruktur gewonnen. Im Anschluss folgt die Analyse der Jahresmittelwerte positiver und negativer Sentimentanteile mittels Intervallplots. Diese gehen bereits durch die Darstellung von Konfidenzintervallen über rein deskriptive Analysen hinaus. Folgend wird die Übereinstimmung der Diktionärergebnisse zunächst durch eine Pearsonkorrelation kontrolliert. Zusätzlich wird die Differenz der jeweils positiven und negativen Sentiments der zuvor berechneten Jahresmittelwerte beider Diktionäre gebildet und erneut mit Konfidenzintervallen dargestellt. So können überzufällige Unterscheidungen der Diktionäre für einzelne Jahre identifiziert werden. Um zu überprüfen, ob die Sentimenttrends nichtlinear verlaufen, werden im Anschluss lineare Regressionen mit nichtlinearen Termen ausgewertet. Die hohe Fallzahl des Datensatzes verursacht Verzerrungen der statistischen Signifikanz. Simulationen, in denen die Regression für verschieden große Stichproben berechnet wird, ermöglichen eine genauere Einschätzung der jeweiligen Effekte. Zuletzt bietet die knappe Kontrolle ausgewählter Annahmen der linearen Regression einen Einblick in die Datenstruktur. Infolgedessen kann die Legitimität der Regressionsergebnisse kritisch hinterfragt werden. Am Ende des Kapitels folgt eine Zusammenfassung der Analyseergebnisse.

5.2 Datenbereinigung und Inspektion

Um Entscheidungen bei der Datenbereinigung informationsgeleitet zu treffen, werden zunächst einige explorativ-deskriptive Analysen angestellt. Die Abbildung 5.1 zeigt, dass fast alle Abstracts im Datensatz nach 1992 veröffentlicht wurden. Daher kann die Analyse nicht wie geplant mit Abstracts seit 1974 durchgeführt werden. Die Beschränkung auf Publikationen ab 1993 dient dem Ausschluss von Verzerrungen durch niedrige jährliche Abstract-Anzahlen. Zusätzlich wurden alle Artikel entfernt, die entweder keine Abstracts enthielten oder deren Abstracts nicht auf Englisch verfasst waren. Die Identifizierung der nicht-englischen Artikel wird in Abschnitt 4.3 erläutert. Fünf Abstracts wiesen eine auffällig extreme Länge auf. Die Entfernung aller Texte mit mehr als 2000 Zeichen, was circa einer Seite Text entspricht, vermeidet die Beeinflussung der Ergebnisse durch Ausreißer.

Die Heatmap 5.2 kontrolliert, ob der nahezu exponentielle Anstieg der Artikelzahl auf einzelne Journale zurückzuführen ist. Die Grafik bildet jedes Journal auf der y-Achse über alle auf der x-Achse aufgezeichneten Jahre ab. Die Anzahl der veröffentlichten Artikel eines Journal in einem Jahr ist farblich gekennzeichnet.

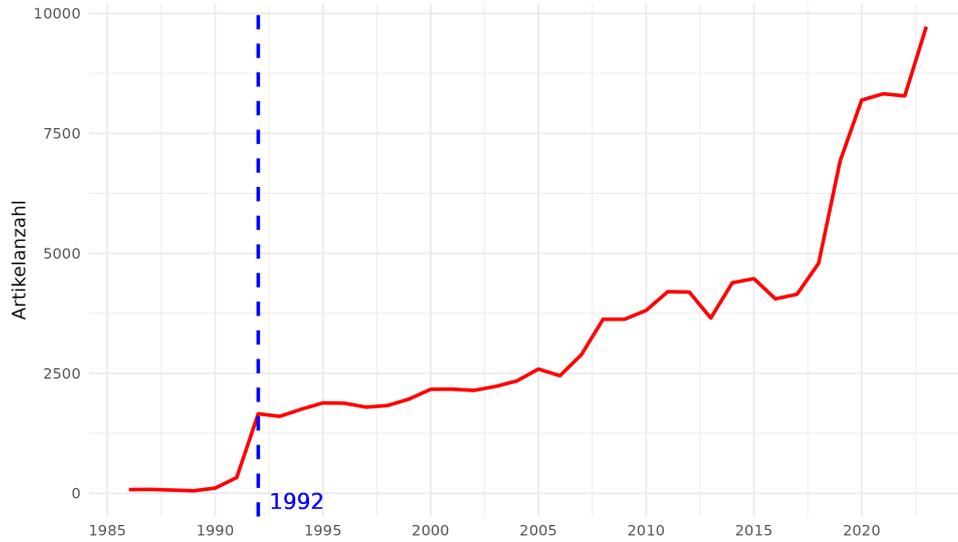


Abbildung 5.1: Im Datensatz enthaltene veröffentlichte Artikel mit Abstract pro Jahr

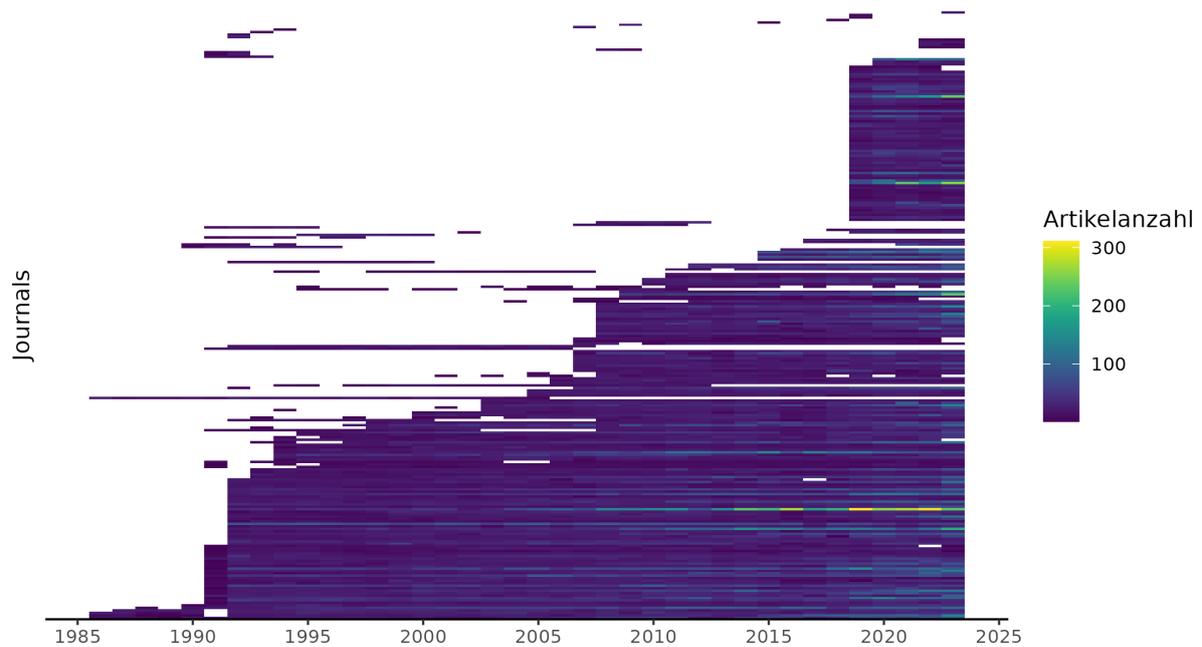


Abbildung 5.2: Heatmap der veröffentlichten Artikel jedes Journal pro Jahr. Je heller die Felder, desto mehr Artikel wurden in einem Jahr in dem jeweiligen Journal veröffentlicht

Die Abbildung 5.2 zeichnet sich durch eine auffällige Treppenform aus. Es wurden eine ganze Reihe Journale im Jahr 1992 zu Web of Science hinzugefügt, sowohl 2008 als auch 2019. Gleichzeitig lassen sich auch einige Journale identifizieren, die seit explizit diesen Jahren nicht mehr auf Web of Science vertreten waren. Die Publikationsfrequenzen jedes Journals beläuft sich in den meisten Fällen auf unter 100, oder sogar unter 50 Artikel pro Jahr. Das gilt hauptsächlich für die frühe-

ren Jahre des Datensatzes. Insbesondere bei den über den längsten Zeitraum veröffentlichenden Journalen existiert ein sichtbarer Anstieg der Anzahl pro Jahr veröffentlichter Artikel. Der in Abbildung 5.1 exponentiell starke Anstieg an veröffentlichten Artikeln pro Jahr ist auf die 2008 und 2019 neu hinzugefügten Journale zurückzuführen. In Abbildung 5.2 sind 3 - 5 Journale zu erkennen, die durchgehend eine wesentlich höhere Artikelanzahl veröffentlichen. Die Identifikation der publikationsstärksten Journale erfolgt mittels der Häufigkeitstabelle 5.1, welche die sechs meistpublizierenden Zeitschriften aufführt. Damit lässt sich das in der Abbildung 5.2 als "gelbe Linie" hervorstechende Journal als "Social Indicators Research" identifizieren.

Journal	Abstracts
Social Indicators Research	3900
Ethnic and Racial Studies	2232
Journal of Marriage and Family	2134
Social Science Quarterly	2131
Social Science Research	2014
Society & Natural Resources	1972

Tabelle 5.1: Journale mit den meisten verfügbaren Abstracts

Zuletzt stellt sich die Frage, wie gravierend sich die Anzahl veröffentlichter Artikel der einflussreichsten Journale, von den Artikelzahlen der übrigen Journale unterscheidet. Lagemaße geben einen ersten Einblick in die grundsätzliche Verteilung der von einem Journal veröffentlichten Artikel. Ein Journal veröffentlichte über den gesamten Zeitraum im Mittel 503 Artikel, der Median liegt allerdings bloß bei 373,3. Social Indicators Research veröffentlichte somit mehr als 10 Mal zu viele Artikel im Vergleich zu einem im Median liegenden Journal. Diese Werte deuten zudem auf eine stark rechtsschiefe Verteilung hin. Die Abbildung 5.3 stellt dazu ein Balkendiagramm der pro Journal veröffentlichten Artikelanzahl dar.

Die sortierte Verteilung der Journale in 5.3 zeigt die für das Gesetz von Bradford typische Verteilung (Bradford, 1985). Bradfords Gesetz besagt, dass wenn Journale absteigend nach ihren enthaltenen Artikeln sortiert werden, die Verteilung einer geometrischen Reihe $1 : n^2 : n^3$ folgt (Bradford, 1985). Bradford teilte die Verteilung ihrem Anstieg nach in Zonen, die durch ihren Anstieg gemäß der geometrischen Reihe nach identifiziert werden. Diese Zonen sind auch in der Abbildung klar zu sehen. Beispielsweise ist die Kernzone, welche dem linearen Anstieg folgt, nach den ersten beiden Journalen deutlich erkennbar. Ungefähr ein Drittel aller Artikel ist Teil dieser Zone und definiert den Kern des Feldes (Bradford, 1985). Der Datensatz wird aus zwei Gründen nicht auf die Artikel aus der Kernzone eingegrenzt. Erstens ändert sich die Reihenfolge der Journale über die Zeit. Um einen zeitlichen Trend zu untersuchen, ist es daher wichtig alle – und nicht nur die nach aktuellen Maßstäben relevanten Journale in die Analyse aufzunehmen. Aus demselben Grund werden die Journale nicht nach Impactfaktor gefiltert. Der Impactfaktor wird auf Basis der letzten N Jahre berechnet. Die Informationen, welche Impactfaktoren die Journale die letzten 30 Jahre besaßen, liegen nicht vor. Zweitens merkt Bradford (1985) selbst an, dass die Artikel, welche in den Journalen des Long-Tails veröffentlicht werden, in keiner Disziplin Teil

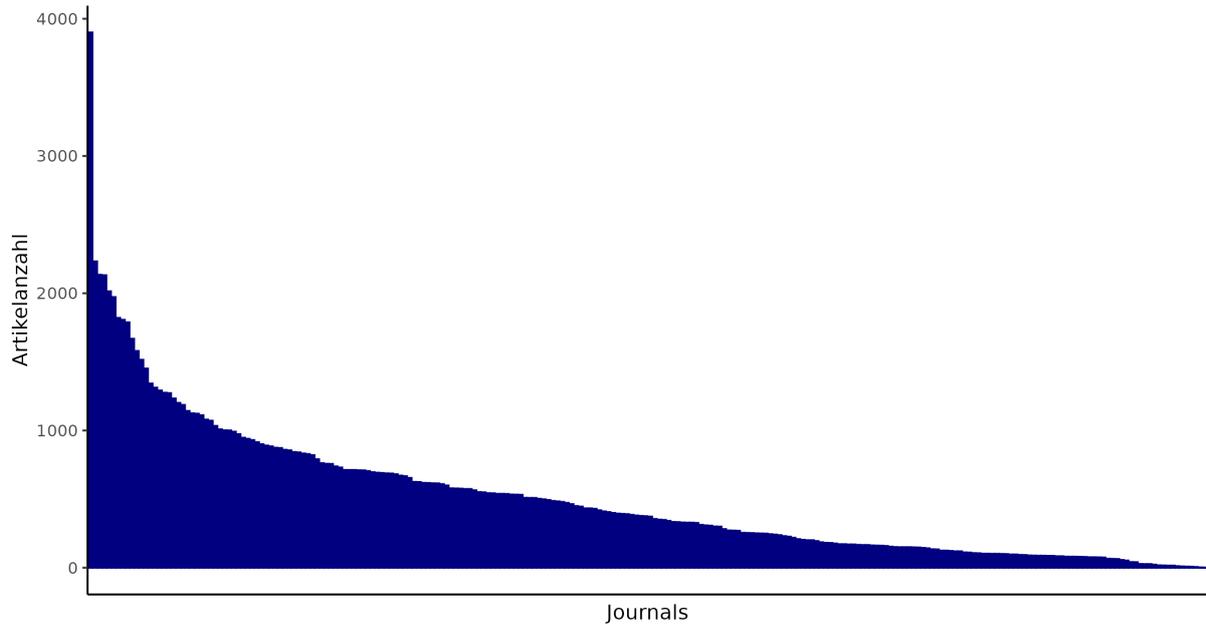


Abbildung 5.3: Veröffentliche Artikel pro Journal

des Kerns sind. Trotz fehlender Kernrelevanz gelten Journale und ihre Beiträge als integrale Bestandteile eines Feldes, weshalb deren Einbeziehung in gründliche Untersuchungen unerlässlich ist.

Im Folgenden wird die Vorbereitung der Textdaten knapp beschrieben. Um das Diktionär von Vinkers et al. (2015) auf die Abstracts anzuwenden, wurden Bindestriche entfernt. Quanteda beachtet bei Diktionärsanalysen automatisch Groß- und Kleinschreibung, weswegen in diesem Schritt die Transformation aller Wörter in kleine Schreibweise nicht nötig war (Benoit et al., 2018). Die Anwendung von Diktionären erfordert im Vergleich zu anderen computergestützten Textanalysen eine geringere Textanpassung, da Diktionäre den Rohtext durchsuchen sollen. Für die Anwendung von VADER muss der Text nicht extra vorbereitet werden. VADERS Algorithmus beachtet die Groß- und Kleinschreibung selbst (Hutto & Gilbert, 2014). Es können für 31 Abstracts keine Sentiments von VADER gebildet werden.

5.3 Jährlich aggregierte Sentimenttrends

Die folgende Analyse der zeitlichen Trends wird mit Intervallplots durchgeführt. Die Jahre werden dabei als Gruppen behandelt, für welche die aggregierten Sentiments mit Mittelwerten und 95% Konfidenzintervallen dargestellt sind.

Die Abbildung zeigt, dass der Anteil positiver Sentiments in den Abstracts im gesamten Zeitraum nahezu konstant und fast linear zugenommen hat. In den frühen 1990er Jahren, liegen die Jahresmittelwerte der Anteile positiver Wörter meist unter 0,07% und besitzen weite Konfidenzintervalle. Im Vergleich zu den restlichen Jahren steigen die Mittelwerte der letzten fünf Jahre kaum

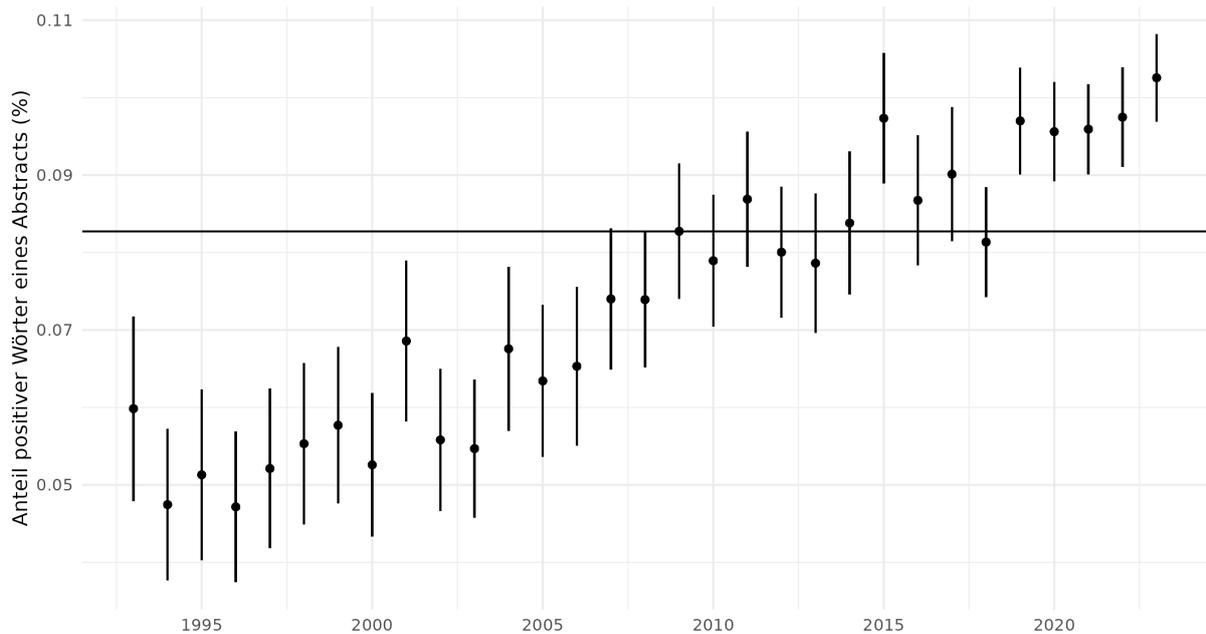


Abbildung 5.4: Diktionär von Vinkers et al. 2015: Jahresmittelwerte des prozentualen Anteils der positiven Sentiments pro Abstract mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

an. Die erhöhte Unsicherheit in den frühen Jahren könnte auf die geringere Anzahl an Abstracts zurückzuführen sein (Abbildung 5.1). Die Verkleinerung der Konfidenzintervalle ist besonders sichtbar ab 2006 und 2017. In diesen Jahren sind ebenfalls in der bereits besprochenen Grafik 5.1 stärkere Anstiege der Artikelanzahl zu verzeichnen, was die Vermutung der Artikelanzahl als Ursache bestätigt. Die erhöhte Variabilität der Werte zwischen 2000 - 2004 und 2012 - 2018 ließe sich vielleicht mit der Abbildung zur Journalaktivität (5.2) erklären, allerdings finden sich in dieser keine Auffälligkeiten in beiden Zeiträumen. Die Unregelmäßigkeiten im Anstieg positiver Sentiments müssten daher auf inhaltliche Trends zurückzuführen sein.

Bei näherer Betrachtung des Einflusses der einzelnen Wörter (Siehe Anhang: .5) fallen “encouraging”, “innovative”, “novel”, “supportive” und “unique” durch ihre jeweils sehr unterschiedlichen Jahresmittelwerte im Zeitraum 2000 bis 2004 auf. Die Abweichung des Jahresmittelwertes von 2000 in Anhang .5 ist nahezu vollkommen auf Unregelmäßigkeiten in der Häufigkeit des Wortes “supportive” zurückzuführen. Die Häufigkeit des Wortes “supportive” stieg in den Jahren vor 2000 fast konstant an, fiel allerdings im Jahr 2000 abrupt ab. Insgesamt werden die Ergebnisse am meisten von den Wörtern “unique”, “creative” und “novel” geprägt. Allerdings haben auch “innovative”, “prominent”, “robust” und “supportive” einen hohen Einfluss auf das Ergebnis. Zwischen 2012 und 2018 ist das sonst sehr starke Wachstum von “novel” unterbrochen. Die Worte “unique”, “innovative” und “robust” nehmen in ihrer Häufigkeit ab, während “creative” sich für 4 Jahre fast verdoppelt, bevor es wieder auf seinen Ursprungswert sinkt.

Für die in Abbildung 5.5 dargestellten negativen Sentiments zeichnet sich kein klarer Trend ab. Im Vergleich mit den Anteilen positiver Sentiments lässt sich festhalten, dass der Mittelwert

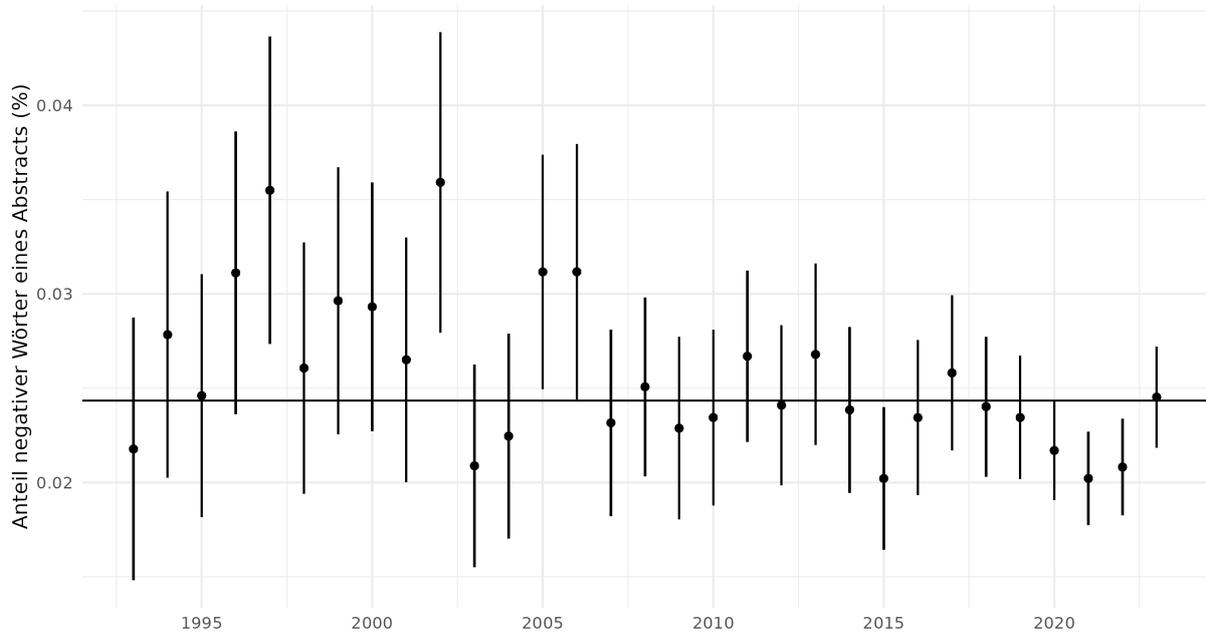


Abbildung 5.5: Diktionär von Vinkers et al. 2015: Jahresmittelwerte des prozentualen Anteils der negativen Sentiments pro Abstract mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

der prozentualen Anteile aller positiven Sentiments nahezu dreimal so hoch, wie der Mittelwert der prozentualen Anteile aller negativen Sentiments ist. Die Jahresmittelwerte zwischen den Jahren 1993 - 1997 steigen steil linear an, wobei das Konfidenzintervall des Jahresmittelwertes von 1997 als einziges komplett über dem Durchschnitt aller Sentimentanteile liegt. Die darauffolgenden Jahre, bis zum Jahresmittelwert 2001, befinden sich wieder näher am Gesamtdurchschnitt der Sentimentanteile. Zwischen 2001 und 2007 springen die Jahresmittelwerte der negativen Sentimentanteile weit um den Mittelwert aller negativen Sentiments. Danach befinden sich allerdings bis 2020 alle Jahresdurchschnitte sehr nah am Mittelwert aller negativen Sentiments. Insbesondere ab 2017 lässt sich ein negativer Trend feststellen, der sich 6 Jahre hält, 2023 aber durch einen Ausreißer aufzulösen scheint. Die Konfidenzintervalle verkleinern sich ab 2006 und 2017 wie bereits in der Abbildung der positiven Sentiments aufgrund der größeren Stichprobenzahl.

Die Jahresmittelwerte der positiven Sentiments des VADER Diktionärs, welche in Abbildung 5.6 dargestellt sind, zeigen allesamt höhere Werte im Vergleich zu den in Abbildung 5.4 dargestellten Sentiments. Die Skalendifferenzen der Sentimentmethoden ergeben sich aus den ungleichen Diktionärsumfängen und VADERS ergänzender Regeln der Tokenklassifizierung. In den Anfangsjahren 1993 bis 2000 zeichnet sich für die gegebenen Werte kein Trend ab. Im genannten Zeitintervall liegen die Sentimentmittelwerte um den Wert 7.1% und damit unter dem Gesamtmittelwert (7.6%) aller positiven Sentimentwerte. Zwischen den Jahren 2002 und 2015 steigen die Sentimentmittelwerte nahezu linear bis auf circa 7.9% über den Gesamtmittelwert an. Die Sentimentwerte der letzten Jahre sinken 2017 und 2018 abrupt bis knapp über den Mittelwert. Dieses Niveau halten folgenden Jahre und zeigen keinen klaren Trend. Allerdings steigt der Anteil

positiver Sentiments im Jahr 2023 wieder leicht an. Wie bereits bei dem Diktionär von Vinkers et al. (2015) treten auch bei den positiven Sentiments von VADER in den Jahren nach 2000, insbesondere 2001, sowie 2016 bis 2019 hohe Unregelmäßigkeiten und kein klarer Verlauf der positiven Sentimentwerte auf. Ebenfalls lassen sich Parallelen in der Weite der Konfidenzintervalle erkennen. Auch in Abbildung 5.6 verkleinern sich die Konfidenzintervalle ab 2006 und 2017 aufgrund der höheren Stichproben sichtbar.

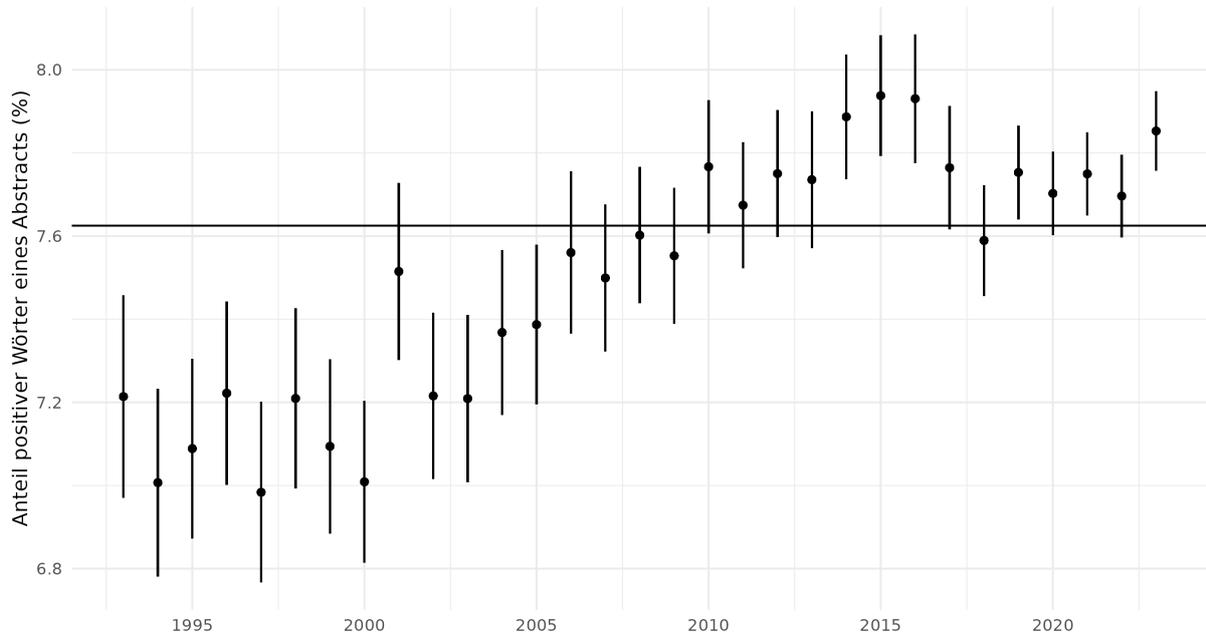


Abbildung 5.6: VADER: Jahresmittelwerte des prozentualen Anteils der positiven Sentiments pro Abstract mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

Die Verteilung der negativen Sentiments ist in Abbildung 5.7 dargestellt. Der Durchschnittswert der negativen Sentiments beträgt 5,1. Dieser Wert entspricht ungefähr einem Viertel des Durchschnittswertes der positiven Sentiments. Die negativen Sentimentanteile sind im Durchschnitt etwa viermal schwächer ausgeprägt als die positiven Stimmungen in der untersuchten Datenmenge. Bis auf zwei Jahre (2011, 2014) schneiden alle Konfidenzintervalle den Gesamtmittelwert der Sentimentanteile. Von einem statistisch signifikanten Trend ist daher nicht auszugehen. Dennoch ist der Verlauf der Jahresmittelwerte deutlicher als negativ zu erkennen, als bei den negativen Sentiments des Diktionärs von Vinkers et al. (2015). In den Jahren 1994 bis 1999 befinden sich alle Jahresmittelwerte über dem Durchschnitt der Anteile aller negativen Sentiments. Zwischen 2000 und 2015 befinden sich, bis auf den Jahresmittelwert von 2013, alle Jahresmittelwerte unterhalb des Gesamtmittelwertes. Zwischen 2005 und 2015 lässt sich kein klarer Trend erkennen. Danach steigen die Mittelwerte rapide bis 2017 über den Mittelwert aller Sentimentanteile an, nur um dann wieder leicht zu sinken, sich dabei aber über dem Gesamtmittelwert zu halten.

Beide Diktionäre stellen eine insgesamt stärkere Nutzung positiver Sprache im Vergleich zur Nutzung negativer Sprache dar. Dieser Effekt ist in der Forschung bereits unter der Bezeichnung

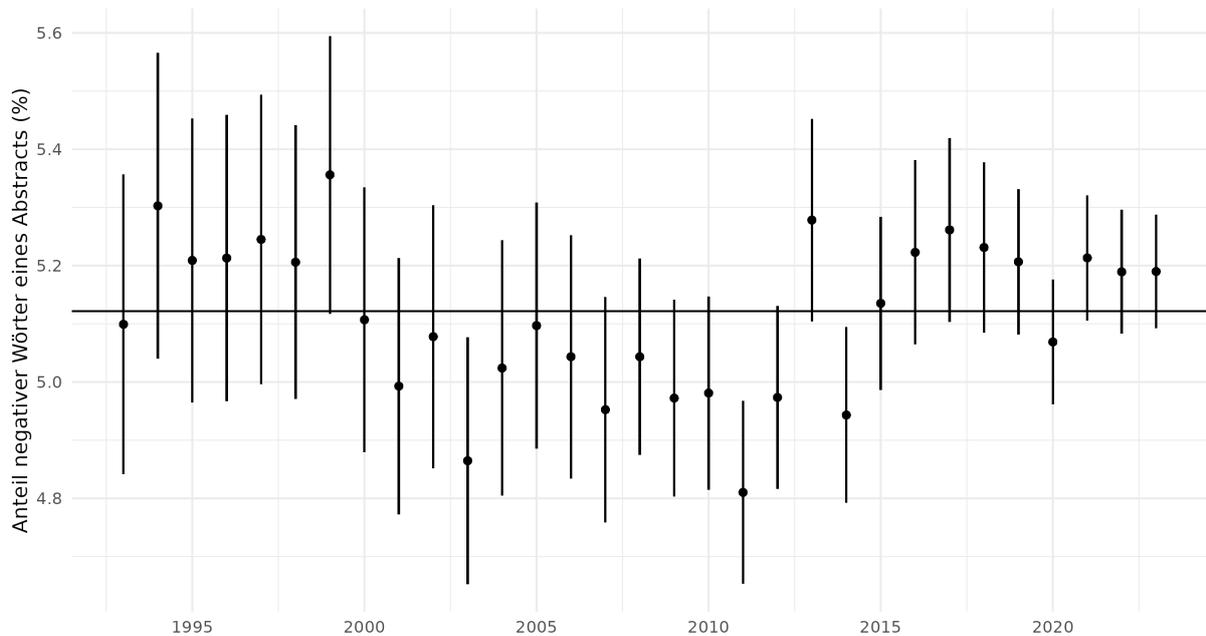


Abbildung 5.7: VADER: Jahresmittelwerte des prozentualen Anteils der negativen Sentiments pro Abstract mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

Polyanna-Hypothese bekannt. Ihren Ursprung besitzt die Polyanna-Hypothese in der Studie Boucher und Osgood (1969), welche eine Reihe kleiner Studien auflistet, um zu zeigen, dass positive Wörter mehr Bedeutung tragen und diverser eingesetzt werden. Dodds et al. (2015) zeigten, dass der von Boucher und Osgood (1969) gezeigte Effekt sich mit 24 Korpora für 10 Sprachen bestätigen lässt.

Die bisherige Analyse stellt einen überzufällig positiven Trend der Abweichungen der Jahresmittelwerte positiver Sentiments in den frühen und späten Jahren auf Basis der Konfidenzintervalle fest. Für die negativen Sentiments kann kein klarer überzufälliger Trend festgestellt werden.

5.4 Vergleich der Diktionärergebnisse

Im Folgenden werden die Ergebnisse beider Diktionäre für positive und negative Sentiments verglichen. Die Pearsonkorrelation der gemessenen Sentimentanteile für die rohen Werte sowie die Jahresmittelwerte sind in Tabelle 5.2 abgebildet.

Die Bewertung einzelner Abstracts unterscheidet sich mit nur 15% und 20% sehr stark voneinander. Die hohen Abweichungen sind plausibel, da viele Journale keines der positiven Wörter des Diktionärs von Vinkers et al. (2015), aufgrund seines geringen Umfangs, enthalten. VADER klassifiziert viele Begriffe als positiv, die statt für eine positive Bewertung der eigenen Forschung eher die Forschung einem bestimmten Themengebiet zuordnen würden. Beispielsweise wertet

Sentiment	Einzelne Abstracts	Jahresebene
Positive Sentiments	0.20	0.91
Negative Sentiments	0.15	0.22

Tabelle 5.2: Pearsonkorrelation der Sentimentanteile, die durch VADER und Vinkers et al. (2015)'s Diktionär festgestellt wurden.

VADER die folgenden Begriffe als positiv: “trust”, “interest”, “satisfaction”, “ethical”, “freedom”. Auf der individuellen Ebene sind VADERs Ergebnisse daher weniger valide als die Ergebnisse des Diktionärs von Vinkers et al. (2015). Das Diktionär von Vinkers et al. (2015) unterschätzt dementsprechend die Sentimentanteile, während VADER diese überschätzt. Auf der Jahresebene ist die Übereinstimmung positiver Sentiments mit 91% sehr hoch. Die aggregierten negativen Sentiments weisen allerdings auch auf der Jahresebene lediglich eine Korrelation von 22% auf. Die niedrige Korrelation der negativen Sentiments beider Diktionäre auf Jahresebene lässt sich durch das Fehlen eines klaren Trends erklären. Mit derselben Argumentation wird der Eindruck eines positiven Trends durch die hohe Übereinstimmung der Jahressentiments beider Diktionäre bekräftigt.

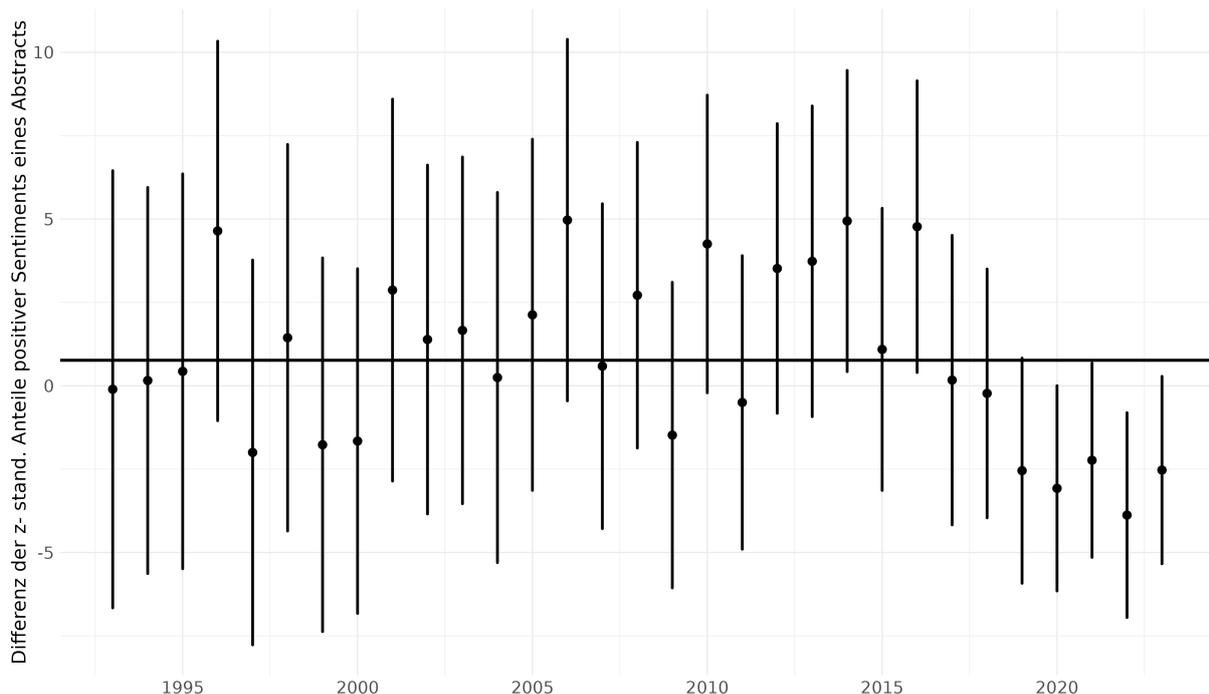


Abbildung 5.8: Differenz der z-standardisierten positiven Sentimentanteilen der Sentimentsoftware VADER minus des Diktionärs von Vinkers et al. (2015) mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

Die Abbildung 5.8 zeigt die Differenz der z-standardisierten Jahresmittelwerte der positiven Sentimentanteile beider Diktionäre. Der leicht über null liegende Gesamtmittelwert deutet darauf hin, dass die Z-Werte der positiven Sentiments von VADER leicht über denen von Vinkers et al. (2015)s Diktionär liegen. Dieser Umstand könnte auf die Tokens von VADER zurückzuführen

sein. Diese sind eher mit themengebietanzeigenden Bedeutungen als mit forschungsbewertenden Bedeutungen aufgeladen. In der Folge klassifiziert VADER insgesamt eine höhere relative Wortanzahl als positiv. Insgesamt weichen nur in den letzten 4 Jahren 3 Jahre klar überzufällig vom Mittelwert ab. Für die vorigen Jahre kann keine statistisch signifikante Abweichung vom Mittelwert festgestellt werden.

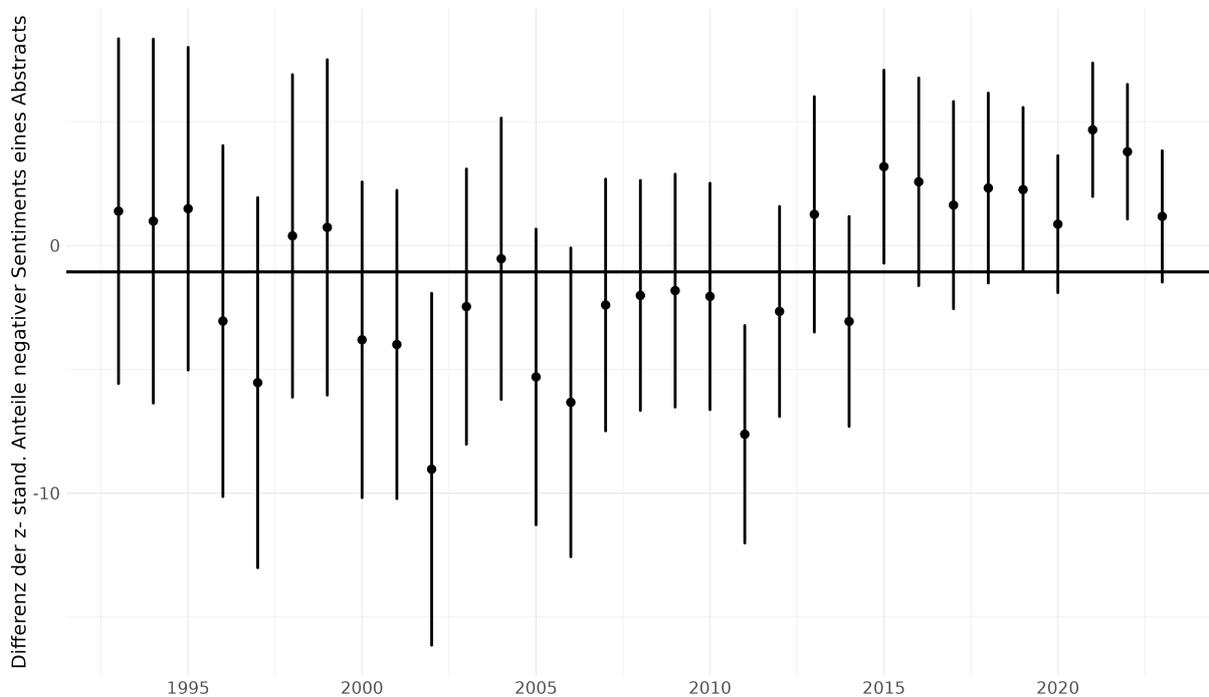


Abbildung 5.9: Differenz der z-standardisierten negativen Sentimentanteile der Sentimentsoftware VADER minus des Diktionärs von Vinkers et al. (2015) mit 95% Konfidenzintervallen; die horizontale Linie markiert den durchschnittlichen Sentimentwert des gesamten Zeitraumes

Die Abbildung 5.9 zeigt die Differenz der z-standardisierten Jahresmittelwerte der negativen Sentimentanteile beider Diktionäre. Der Gesamtmittelwert der Differenzen liegt leicht unter null. VADER bewertete demnach die Journale weniger negativ als das Diktionär von Vinkers et al. (2015). Lediglich die Jahre 2002, 2015, 2021 und 2022 weichen statistisch signifikant vom Mittelwert ab. Die restlichen Mittelwerte zeigen keine überzufällige Abweichung an. Die erstellten Abbildungen der Differenzen zeigen, dass beide Diktionäre eine hohe Übereinstimmung der Bewertung positiver und negativer Sentiments auf der Jahresebene besitzen. Davon ausgenommen sind die Abweichungen der letzten 4 Jahre.

5.5 Lineare Regressionen

In der bisherigen Analyse konnte ein überzufällig positiver Trend der positiven Sentiments festgestellt werden. Die bisher genutzten Methoden lassen allerdings keine Aussagen über die Art des Trends zu. Mit der linearen Regression nach dem bewährten OLS-Verfahren wird im Folgenden der genaue Trendverlauf der positiven Sentimentanteile untersucht.

$$\text{Vink I: } s_{\text{vink}_{\text{positiv}}} = \beta_0 + \beta_1 \cdot \text{Jahr} \quad (5.1)$$

$$\text{Vink II: } s_{\text{vink}_{\text{positiv}}} = \beta_0 + \beta_1 \cdot \text{Jahr} + \beta_2 \cdot \text{Jahr}^2 \quad (5.2)$$

$$\text{Vader I: } s_{\text{vader}_{\text{positiv}}} = \beta_0 + \beta_1 \cdot \text{Jahr} \quad (5.3)$$

$$\text{Vader II: } s_{\text{vader}_{\text{positiv}}} = \beta_0 + \beta_1 \cdot \text{Jahr} + \beta_2 \cdot \text{Jahr}^2 \quad (5.4)$$

In Modell 1 werden die Jahre (Variable “Jahr”) als Prädiktorvariable genutzt. Im Anschluss wird das Modell um die quadrierte Form der Jahre als Prädiktorvariable (Variable “Jahr²”) erweitert. Dies ermöglicht zu überprüfen, ob die visuell erkennbaren geringen Abweichungen eines linearen Trends der relativen Häufigkeiten positiver Sentiments als statistisch überzufällig interpretiert werden sollten. Die Gleichungen (5.1) - (5.4) explizieren die zu berechnenden Regressionsgleichungen. Für die Regressionen werden nicht, wie in der bisherigen Analyse, die pro Jahr aggregierten Daten verwendet. Der Grund ist, dass die durch aggregierte Daten reduzierten Informationen über die Streuung in den Daten ausschlaggebend sein könnten, um entsprechend geringe Abweichungen beim Hinzufügen eines nichtlinearen Terms zu identifizieren.

	vink I	vink II	vader I	vader II
Intercept	-3.5618*** (0.1866)	-6.2191 (46.6747)	-41.7167*** (3.3142)	-4523.4624*** (828.7231)
Jahr	0.0018*** (0.0001)	0.0045 (0.0464)	0.0245*** (0.0016)	4.4841*** (0.8246)
Jahr ²		-0.0000 (0.0000)		-0.0011*** (0.0002)
N	121789	121789	121758	121758
R ²	0.003	0.003	0.002	0.002

Tabelle 5.3: Lineare Regressionen beider Diktionäre auf Abstractebene. Die Standardfehler sind in Klammern dargestellt.

* p<.1, ** p<.05, ***p<.01

Der Intercept ist in keinem der in 5.3 Modelle interpretierbar, da der Beginn der Zeitreihe beliebig nach der Verfügbarkeit der Daten gewählt wurde. Daher gibt das Jahr 1993 keinen bedeutungsvollen Bezugsrahmen, um eine sich vom Intercept unterscheidende Variable mit diesem zu vergleichen. Im Modell “Vink I” ist der Koeffizient mit einem Anstieg von 0,0018 Anteilen positiver Sentiments pro Jahr sehr klein. Umgerechnet steigen die Anteile positiver Sentiments um 1% in circa 500 Jahren an. Der Effekt ist statistisch signifikant mit $p < 0.01$. Nach dem Hinzufügen des quadratischen Terms im Modell “Vink II” steigt der Koeffizient der Variable “Jahr” auf 0,0045 Sentimentanteile pro Jahr an. Der Koeffizient der Variable Jahr² ist sehr klein und negativ. Beide Variablen sind aufgrund hoher Kollinearität nicht statistisch signifikant (J. Cohen et al., 2015, S. 450). Hohe Multikollinearität führt im Extremfall zu Veränderungen der Koeffizienten, weshalb die angegebenen Ergebnisse mit hoher Vorsicht zu betrachten sind (J. Cohen et al., 2015,

S. 450). Beide Variablen sind ihrer Definition nach hoch korreliert, da ein Anstieg von "Jahr" um eine Einheit die Variable "Jahr²" immer um einen proportional größeren Wert ansteigen lässt. Das Modell "Vader I" bestätigt den sehr kleinen, statistisch signifikanten Effekt des Modells "Vink I". Der Koeffizient zeigt an, dass pro Jahr die positiven Sentimentanteile der Abstracts um 0.0245% ansteigen. Nach dem Hinzufügen des quadratischen Terms im Modell "Vader II" steigt der Koeffizient der Variable "Jahr" auf 4,5% an und ist mit $p < 0.01$ statistisch signifikant. Der quadratische Term ist mit $-0,001$ sehr klein, aber dennoch statistisch signifikant mit $p < 0.001$. Auch in diesem Modell verzerrt der Einfluss der hohen Multikollinearität die Prädiktoren. Alle Modelle besitzen ein sehr kleines R^2 . Demnach übersteigt die durchschnittliche Abweichung der Messwerte von den geschätzten Werten die durchschnittliche Abweichung der Messwerte vom Mittelwert um ein Vielfaches. Die Ursache sind die innerhalb der Jahre herrschenden sehr hohen Standardabweichungen. In der Tabelle 5.4 sind die Mittelwerte und Standardabweichungen der positiven Sentiments beider Diktionäre abgebildet. Die Standardabweichung der positiven Sentimentanteile des Diktionärs Vinkers et al. (2015) ist dreimal so groß wie der dazugehörige Mittelwert. Dagegen ist die Standardabweichung der positiven Sentimentanteile von VADER 35% kleiner als der dazugehörige Mittelwert.

positive Sentiments	Mittelwert	Standardabweichung
Vinkers et al. (2015)	0,0827	0,2768
VADER	7,6251	4,9120

Tabelle 5.4: Vergleich der Lage und Streumaße der positiven Sentiments von Vinkers et al. (2015) Diktionär und VADER

Bis auf das Modell "Vink II" waren alle Modelle mit einem sehr kleinen Standardfehler statistisch signifikant. Im Folgenden werden daher die in Kapitel 4 dargelegten Simulationen sowie deren Ergebnisse behandelt. Zur Generierung der Stichproben diente ein schrittweises, zufälliges Ziehen ohne Zurücklegen. Die resultierenden Stichprobenumfänge erstreckten sich von 100 bis 10.000 Datenpunkte, mit einer Erhöhung um jeweils 100 Einheiten. Für jede Stichprobengröße wurden 50 Zufallsstichproben gezogen und die lineare Regression berechnet. Die resultierenden p-Werte für den Koeffizienten der Variable "Jahr" wurden für jede Stichprobengröße und Wiederholung gespeichert und für jede Stichprobengröße ein mittlerer p-Wert aller 50 Wiederholungen jeder Stichprobengröße gerechnet. Für das Modell "Vader II" wurden die Stichprobengrößen und Abstände der Zufallsziehungen abgewandelt, sodass mit einer Stichprobe von $n = 5000$ beginnend die Stichprobe in der folgenden Iteration um 5000 Fälle erhöht wurde. In der Tabelle 5.5 sind die Stichprobengrößen der erstmaligen Überschreitungen der typischen Signifikanzgrenzen eingetragen. Die CPS-Grafiken sind im Anhang .6 zu finden.

Das Modell "Vink I" benötigt im Vergleich zu den anderen Modellen, mit minimal 1900 Fällen, die kleinste Stichprobe um gerade noch signifikant zu sein. Für alle Variablen sind sehr hohe Fallzahlen notwendig um ein statistisch signifikantes Ergebnis zu erhalten. Insbesondere die quadratische Variable im Modell "Vader II", welche den nichtlinearen Effekt abbilden soll, ist erst ab circa 30000 Fällen statistisch signifikant. Diese Ergebnisse zeigen, dass für übliche Stichpro-

p-Wert	Vink I Variable "Jahr"	Vader I Variable "Jahr"	Vader II Variable "Jahr ² "
p < 0.05	1900	4100	30000
p < 0.01	3100	6300	45000
p < 0.001	5300	8700	70000

Tabelle 5.5: Ergebnisse der Monte-Carlo Simulationen: Stichprobengrößen ab denen typische Signifikanzgrenzen überschritten werden

bengrößen kein Effekt statistisch Signifikant wäre. Daher ist auf Basis der linearen Modelle von keinem relevanten Anstieg positiver Sentiments auszugehen.

5.6 Regressionsdiagnostik

Der letzte Abschnitt des Kapitels widmet sich einer knappen Regressionsdiagnostik. Das Ziel der linearen Regressionen war die Prüfung des Einflusses potenzieller nichtlinearer Effekte. Im Folgenden wird durch die Untersuchung einiger Annahmen linearer Regressionen geprüft, wie viel Wert den Modellergebnissen zugeteilt werden kann. Ein Verstoß der Annahmen kann zu verzerrten Standardfehlern oder verzerrten Koeffizienten führen (J. Cohen et al., 2015, S. 118) Die Streuung der Residuen sollte homogen um den Schätzer verteilt sein. Wenn die Residuen ihre Streuung abhängig der unabhängigen Variablen verändern liegt, Heteroskedastizität vor. Dieser Umstand führt zur Verzerrung der Konfidenzintervalle und beeinflusst damit die statistische Signifikanz (J. Cohen et al., 2015). In Abbildung 5.10 sind die Residuendiagramme der vier Regressionsmodelle dargestellt. Diese bilden die Residuen im Bezug zu den geschätzten Werten des Modells ab.

Alle Modelle weisen heteroskedastizid verteilte Residuen auf. Speziell für die Modelle "Vink I" und "Vink II" zeichnet sich eine besonders hohe Heteroskedastizität ab, welche von starken Ausreißern für die höheren geschätzten Werte begleitet wird. Kein Modell zeigt die typische Verteilung linearer Verteilungen, dennoch liegt die LOWESS Kurve nahe null. Weitergehend ist ebenfalls die Lage der niedrigen Messwerte der Vink I & II Modelle markant. Der Wert null scheint sehr viel häufiger aufzutreten als aller darüberliegenden Werte, da die LOWESS Kurve nahezu direkt auf der Nullgeraden liegt. Werte über null kommen erst nach einer "kleinen Lücke" wieder vor. Das Modell "Vader II", welches bereits einen etwas höheren signifikanten nichtlinearen Term besaß, zeigt eine Zunahme der Dichte der geschätzten Werte in den höheren Sentimentanteilen. Der folgende Abschnitt widmet sich zunächst der erhöhten Wertedichte in den höheren geschätzten Werten von "Vader II". Danach wird die Verteilung der Messwerte aufgrund der besonderen kleinen Werte von Vink I & II untersucht. Abbildung 5.11 zeigt die Regressionslinien der vier Modelle, zusammen mit den bereits aus den Intervallplots bekannten Jahresmittelwerten. Der nichtlineare Koeffizient des Modells "Vader II" ist verhältnismäßig klein, seine Auswirkung auf die Trajektorie ist allerdings deutlich sichtbar. Die letzten Jahresmittelwerte beeinflussen auch

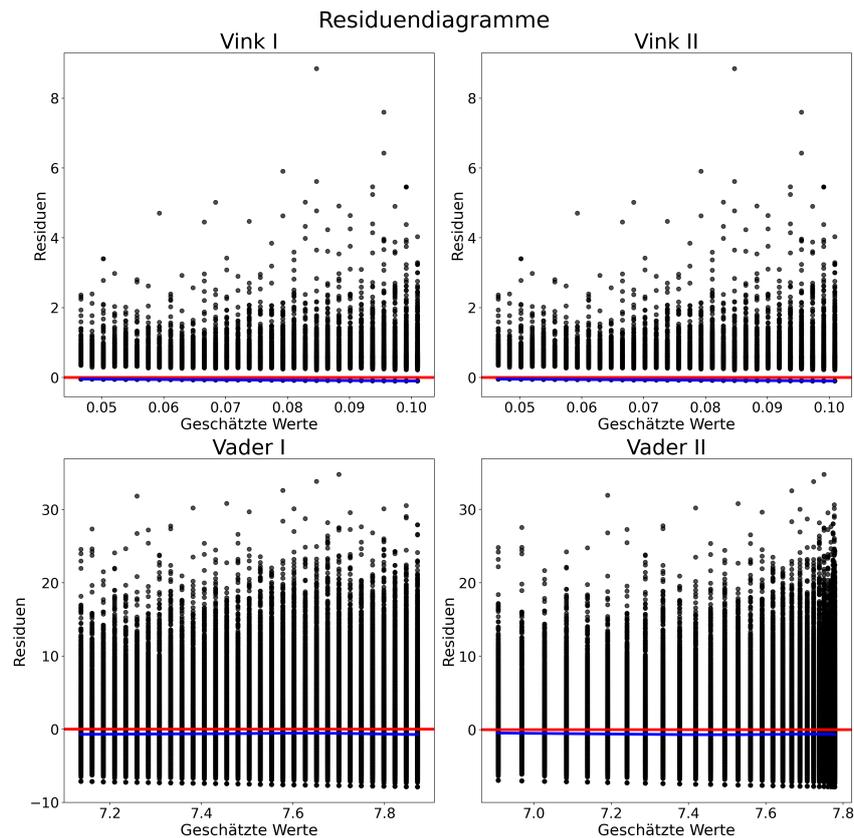


Abbildung 5.10: Residuendiagramme der vier linearen Regressionen. Der Wert null wurde rot markiert. Die lokal gewichtete Regression (LOWESS) ist in blau dargestellt.

die Steigung der Geraden in “Vader I”. Die beiden “Vink I & II” Modelle unterscheiden sich nicht sichtbar voneinander. “Vader II’s Nichtlinearität ist demnach das Resultat der Bewertungsunterschiede der letzten Jahre 2019-2023, welche bereits beim Vergleich der Differenzen identifiziert wurden.

Die Abbildung 5.12 zeigt die Verteilung der positiven Sentimentanteile beider Diktionäre als Histogramme. Die in Abbildung 5.12 dargestellte Verteilung ähnelt stark den Residuenverteilungen, welche in 5.13 dargestellt sind. Markant ist die große Anzahl des Wertes null der positiven Sentimentanteile des Diktionärs von Vinkers et al. (2015). Dies ist auf die geringe Anzahl an Worten im Diktionär im Kontext wissenschaftlicher Texte mit vorwiegend neutraler wissenschaftlicher Sprache zurückzuführen. Der Platz zwischen null und der kleinen rechtsschiefen Verteilung um 0.5 ist vermutlich der Kombination der minimalen Abstractlänge und den wenigen Wörtern eines Diktionärs geschuldet. Die meisten Abstracts, in denen ein positives Wort gefunden wurde, müssen mindestens ungefähr zweihundert Wörter lang sein. Da Vader insgesamt mehr unterschiedliche Begriffe positiv kategorisiert, fällt auch der Abstand zwischen null und dem Maximum der Wahrscheinlichkeitsverteilung weniger prägnant aus. Die starke Rechtsschiefe beider Verteilungen, ohne die Nullwerte, erinnert an eine poisson- oder logarithmischverteilte Variable. Da die Residuen, in Abbildung 5.13, ebenso verteilt sind, wird eindeutig die Annahme normalverteilter Residuen um den Schätzer missachtet.

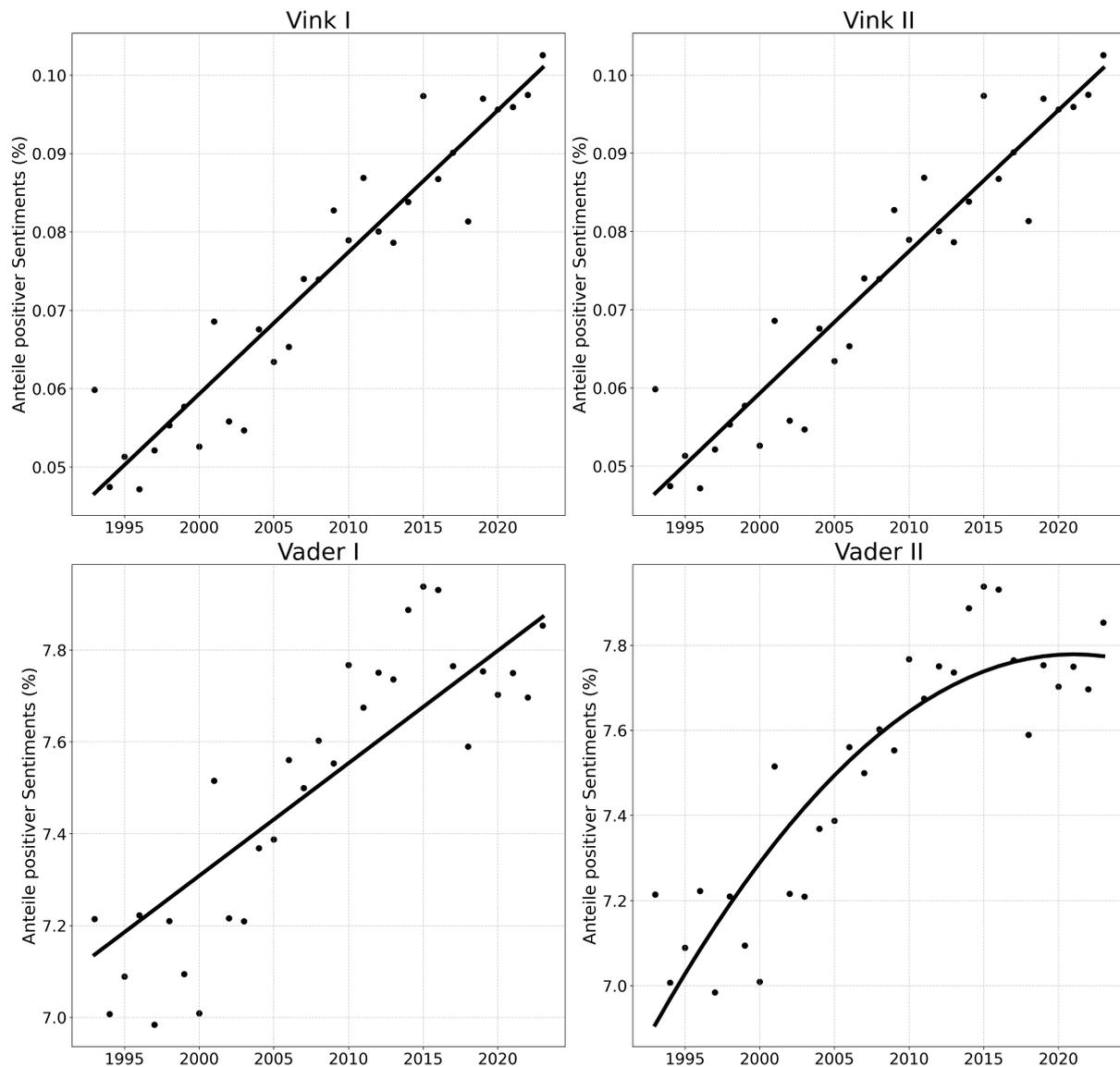


Abbildung 5.11: Vergleich der geschätzten Trajektorien aller vier linearen Regressionen. Zusätzlich abgebildet sind die Jahresdurchschnittswerte, welche bereits in den Intervallplots abgebildet wurden.

Kleine Verstöße gegen die Annahmen der linearen Regression deligitimieren nicht direkt die Anwendung der linearen Regression. Die bislang entdeckten Regelverletzungen haben jedoch ein erhebliches Ausmaß. Nicht nur zeigt eines der beiden Diktionäre einen kleinen nichtlinearen Zusammenhang, sondern es herrschen außerdem große Probleme bei dem anderen Diktionär bezüglich der Struktur der gemessenen Daten. Die festgestellten Verzerrungen betreffen nicht bloß die Standardfehler, sondern auch die Koeffizienten der Modelle (J. Cohen et al., 2015, S. 148–155). Nach den bisherigen Diagnosen stehen Forschenden mehrere Möglichkeiten zur Auswahl, die allerdings nicht weiter im Rahmen der Forschungsfrage dieser Arbeit liegen. Während J. Cohen et al. (2015, S. 447) die Nutzung robuster Regression vorschlägt, um nicht-normalverteilte Daten mit Extremwerten zu modellieren, empfiehlt Chatterjee et al. (2000, S. 163) die Transformation von stark heteroskedastiziden Daten mit \sqrt{Y} vor. Im Anhang .7 sind die Ergebnisse

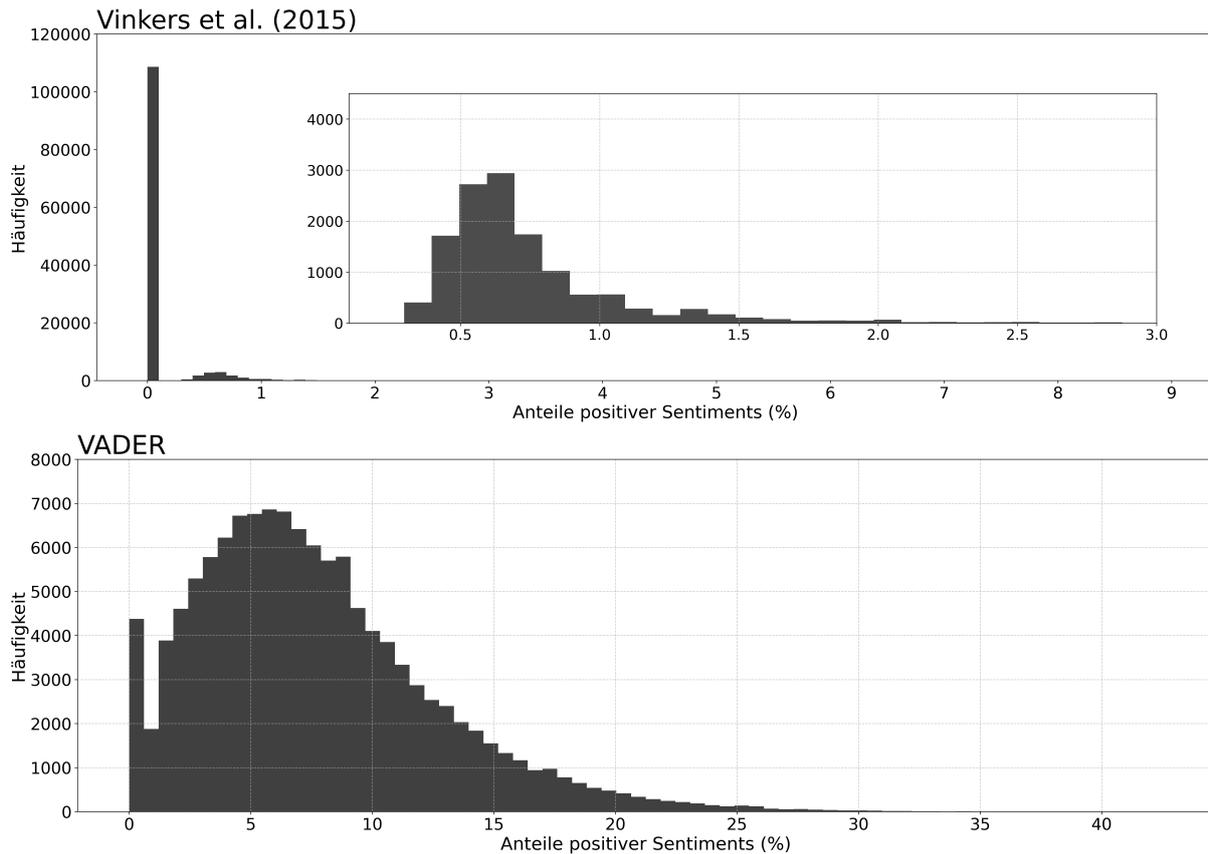


Abbildung 5.12: Histogramme der gemessenen Sentimentanteile des Diktionärs von Vinkers et al. (2015) und VADER

einer entsprechenden Transformation zu finden.

5.7 Zusammenfassung

Zu Beginn des Kapitels zeigten deskriptive Analysen, dass in den Jahren 1993, 2008 und 2019 ein wellenartiger Zuwachs an Journals stattfand. Die Verteilung der in Journal veröffentlichten Artikelzahlen folgt Bradfords Gesetz. Anschließend werden die Zeitreihen der Sentimentanteile, sowohl positiv als auch negativ, mittels Jahresmittelwerten und Intervallplots analysiert. Dabei steigen die Jahresmittelwerte positiver Sentiments überzufällig an, wobei sich für die negativen Sentiments kein klarer Trend beobachten ließ. Die Klassifizierung von Sentiments in Abstracts beider Diktionäre korrelierten auf Abstractebene sehr schwach. Auf der Jahresebene korrelierten die negativen Sentiments sehr schwach, die positiven Sentiments allerdings sehr hoch. Der Anstieg positiver Sentiments konnte so bestätigt werden, mit Ausnahme der Jahre 2019 - 2023. Die linearen Regressionen zeigen einen sehr kleinen Anstieg positiver Sentiments mit einem sehr niedrigen R^2 . Nur das Diktionär VADER besitzt einen kleinen sehr kleinen nichtlinearen Effekt. Die Variablen der Modelle werden erst für Stichproben mit mehreren Tausend Fällen statistisch signifikant. Eine knappe Regressionsdiagnostik identifiziert die heteroskedastizide Datenbasis mit

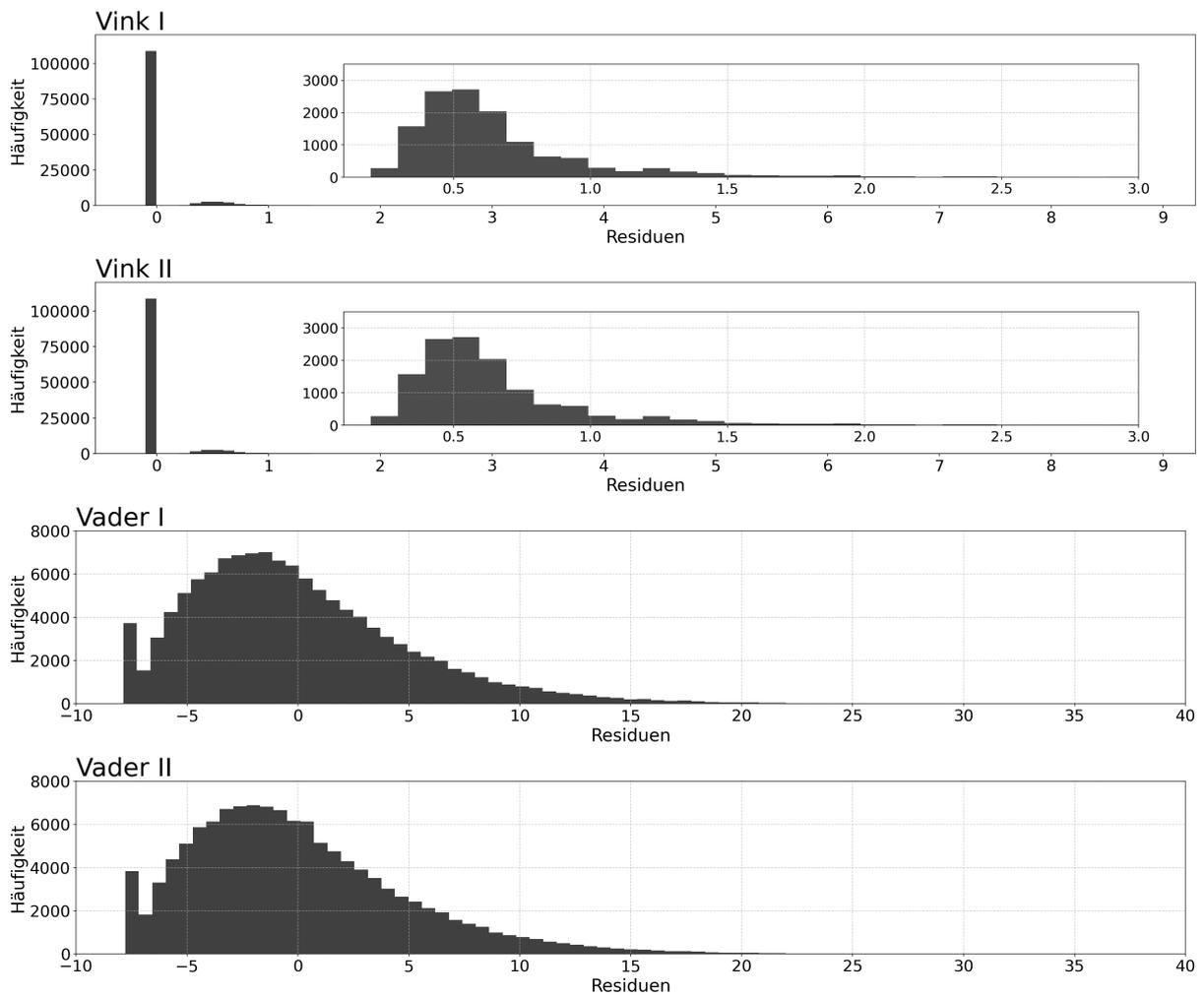


Abbildung 5.13: Histogramme der Residuen aller linearen Modelle

nichtlinearen Residuen. Im Resultat zeigen die Diagnostiken die komplizierte Struktur der gemessenen Daten auf, wodurch Zweifel am linearen Modell entstehen. Zuletzt werden Vorschläge für potenzielle Lösungen der Probleme für weitere Analysen aufgezeigt.

6 | Diskussion und Fazit

6.1 Einleitung

Ziel des Kapitels ist, durch die Diskussion und Interpretation der Analyseergebnisse zu einem Fazit zu gelangen. Dazu werden zunächst die Ergebnisse zusammengefasst und interpretiert, indem Verknüpfungen zwischen den verschiedenen Analyseschritten und der verwendeten Methode hergestellt werden. Anschließend werden Vermutungen über mögliche Ursachen der erkannten Effekte angestellt. Insbesondere integriert der Abschnitt alle angestellten Analysen, um die gemessenen Effekte zu erklären. Dem folgend werden durch Reflexion der bisherigen Forschung und den Erkenntnissen dieser Arbeit, Anmerkungen für zukünftige Forschung formuliert. Außerdem geht der Abschnitt auf die Konsequenzen der Ergebnisse für die Theorie des akademischen Kapitalismus ein. Dabei steht insbesondere der Dualismus unterschiedlicher Handlungsmodi im Vordergrund. Danach werden eine Reihe an Limitationen der durchgeführten Forschung aufgelistet, welche sich insbesondere auf die Methodik wie in etwa die Operationalisierung der verwendeten Konzepte konzentriert. Zuletzt fasst das Fazit die gesamte Arbeit knapp zusammen und kommt zu einem abschließenden Urteil.

6.2 Interpretation der Ergebnisse

Der nachfolgende Abschnitt widmet sich zunächst der Interpretation der Ergebnisse. Im Anschluss werden Vermutungen über die Ursachen der Ergebnisse geäußert und Ideen für zukünftige Analysen des Themas angestellt. In Kapitel 5 konnte ein Anstieg positiver Sentimentmittelwerte mit beiden Diktionären, sowie das Fehlen eines klaren Trends der negativen Sentiments für die gewählte Datenbasis gezeigt werden. Folglich existiert in Artikeln, die zwischen 1993 und 2023 in Journalen mit der von Web of Science vergebenen Kategorie "Sociology" erschienen, ein minimaler Anstieg des positiven Framings. Diese Folgerung jedoch gilt nur, wenn die Streuung der Rohdaten ignoriert wird. Bei der Integration der Streuung durch den Standardfehler, wie beispielsweise in den Intervallplots sowie der linearen Regression geschehen ist, scheint der Anstieg zunächst überzufällig zu sein. Die Simulationen der p-Werte der Regressionsmodelle zeigt allerdings, dass die Signifikanz des Anstiegs auf die hohe Stichprobengröße statt auf einen relevanten Effekt zurückzuführen ist. Mithilfe einer knappen Regressionsdiagnostik kann eine starke Verletzung der meisten Annahmen der linearen Regression gezeigt werden. Zusammenfassend sind die Sentimentmesswerte nicht linear und stark rechtsschief verteilt, da in besonders vielen Abstracts keine positiven Sentiments gefunden werden können. Darüberhinaus existierten einflussreiche Ausreißer in den hohen Wertebereichen. Unter den genannten Verstößen treten starke Zweifel an den Ergebnissen der linearen Regression und dem Trend positiver Sentiments in den Intervallplots auf. Die festgestellten Mittelwertanstiege sind sehr gering und auf Basis der Simulationen

sowie etlicher Modellverstöße nicht statistisch signifikant. Daher ist von keinem nennenswerten Anstieg positiver Sentiments auszugehen.

Im Folgenden werden Vermutungen über mögliche Ursachen der Ergebnisse aufgestellt. Der minimal positive Trend ist allem Anschein nach auf die Häufung der Ausreißer in den Jahren 2019 bis 2023 zurückzuführen. Diese Ausreißer sind in der Abbildung der Residuen 5.10 zu erkennen. Festzustellen ist auch, dass die Ausreißer bei den Messwerten des Diktionärs von Vinkers et al. (2015) stärker ausgeprägt sind als in den Messwerten des VADER Diktionärs. Dieser Unterschied der Ausreißer erklärt die in den Intervallplots auftretenden (scheinbar) überzufälligen hohen Abweichungen in den Jahren 2019 - 2023. Das Diktionär von Vinkers et al. (2015) bewertete in diesen Jahren mehr Begriffe positiv als die VADER Sentimentsoftware, welche einen abrupten Abfall der Anteile positiver Sentiments pro Jahr verzeichnete. Die in der Abbildung 5.11 dargestellten Trajektorien der Regressionsmodelle deuten darauf hin, dass der Messunterschied der Jahressentiments von 2019 bis 2023 verantwortlich für den Einfluss des nichtlinearen Terms im Modell von "Vader II" ist. Qualitative Unterschiede beider Diktionäre wurden bereits während der Präsentation der Ergebnisse dargelegt. Der Abschnitt zum Vergleich der Diktionärergebnisse 5.4 stellt fest, dass das umfangreichere Diktionär von VADER auch Tokens als positiv klassifiziert, welche eher bestimmte Themen statt positive Bewertungen repräsentieren. Für die Klassifikation des Diktionärs von Vinkers et al. (2015) ist ein sprunghafter Anstieg der Begriffe "novel", sowie "innovative" zu verzeichnen (siehe Anhang .5). Wahrscheinlich verursachte die dritte Welle an hinzukommenden Journalen 2019 (siehe Abbildung 5.2) einen Abfall der relativen Häufigkeit der von VADER als positiv kategorisierten Themen. Präziser formuliert: Die von VADER als positiv bewertete Begriffe, welche eher für bestimmte wissenschaftliche Themen stehen, nehmen durch das Hinzufügen neuer Journale in denen diese Themen nicht behandelt werden, in ihrer prozentualen Häufigkeit ab.

6.3 Implikationen der Ergebnisse

Auf Basis der im vorigen Abschnitt angestellten Interpretation lassen sich nun methodische und theoretische Anmerkungen festhalten. Bisherige Studien wählten für ihren Analyseansatz entweder lokale oder lineare Regressionen (siehe Kapitel 3). Die Untersuchung dieser Arbeit der Datenstrukturen von Sentimentwerten mit wissenschaftlichen Texten, sowie der knappen Regressionsdiagnostik zeigen die hohe Bedeutung der Prüfung entsprechender Modellannahmen. Lokale Regressionen sind besonders fehleranfällig für Überanpassung, welche durch starke Ausreißer begünstigt wird. Lineare Regressionen sind insbesondere von den Folgen inhomogener Variablenverteilungen betroffen. Natürlich lassen sich die komplizierten Datenstrukturen relativer Sentimentwerte nicht auf andere Datenbasen und Sentimentskalen übertragen. Dennoch wäre die Vermutung, die Auffälligkeiten seien ein Problem, welches lediglich soziologische Veröffentlichungen betrifft, unrealistisch.

Auf Basis der Ergebnisse lassen sich ebenfalls für die in Kapitel 2 vorgestellte Theorie Schlussfolgerungen ziehen. Das Kapitel 2 zur Theorie des akademischen Kapitalismus, stellte den Dua-

lismus von Markthandlungen und Wissenschaftshandlungen vor (Baumeler, 2009; Ylijoki, 2003). Davon ausgehend wurde die Vermutung geäußert, Marketinghandlungen würden sich ebenfalls im Rahmen von positivem Framing in eigenen wissenschaftlichen Publikationen äußern. Die Ergebnisse der Analyse dieser Arbeit widersprechen der Vermutung. Damit argumentiert diese Arbeit, dass wissenschaftliches Publizieren ungeachtet des Einflusses eines akademischen Kapitalismus als Wissenschaftshandlung gesehen werden kann. Dennoch stellt die Analyse keine allumfassende Untersuchung von positivem Framing in wissenschaftlicher Literatur dar. Beispielsweise konnten die Ursachen der Ausreißer und die Unter- und Überschätzung von Sentiments durch die verwendeten Diktionäre nicht abschließend geklärt werden. Zudem lässt eine einfache Betrachtung von Trends in Zeitreihen und Korrelationen grundsätzlich keine kausalen Schlüsse zu. Neben den zu Beginn genannten methodischen Empfehlungen für die Methodik zukünftiger Forschung lässt sich eine anknüpfende theoretische Anschlussmöglichkeit aufzeigen. Genauer stellt sich die Frage, weshalb akademischer Kapitalismus bestimmte Bereiche der wissenschaftlichen Praxis beeinflussen kann und aus welchen Gründen andere Aspekte davon unberührt bleiben.

6.4 Limitationen

Neben den bisher bereits zum Teil reflektierten Limitationen listet dieser Abschnitt die wichtigsten Grenzen der Analyse auf. Zu Beginn werden insbesondere allgemeine Nachteile quantitativ-explorativer Methodik relevant. Zunächst ist anzumerken, dass die Reichweite der Exploration durch die Eigenschaften des Datensatzes begrenzt ist. Der für die Analyse genutzte Datensatz stellt lediglich Daten, welche über Web of Science verfügbar sind, dar. Ebenfalls konnten keine als Kontrollvariablen relevante Variablen heruntergeladen werden. Darüberhinaus existieren viele die Operationalisierung betreffende Grenzen. Die Eingrenzung der analysierten Publikationen auf die Disziplin der Soziologie wurde über die Journalkategorien von Web of Science festgelegt (of Science Group, 2019). Die Auswahlprozesse der Web of Science Group sind nicht für jedes Journal transparent nachvollziehbar. Diese fehlende Transparenz wurde für die Analyse aufgrund der von of Science Group (2019) kommunizierten Verfahren akzeptiert. Zusätzlich werden im Datensatz lediglich in Journalen veröffentlichte Publikationen abgebildet. Die Soziologie ist allerdings ein hoch-interdisziplinäres Feld in dem eine hohe Diversität an Publikationsformen (Monografien, Sammelbände und Journalartikel) verbreitet ist (Becher, 1989, S. 111). Ebenfalls besitzt das Abstract als Analyseeinheit eine geringere Repräsentationskraft für einen Artikel, als der gesamte Volltext des Artikels. Zudem kamen für Zeitreihenanalysen übliche Methoden nicht zum Einsatz, da durch die grobe Zeitskala die Datenstruktur eher an ein Forschungsdesign mit wiederholten Querschnittsdaten als an eine Zeitreihe erinnerte. Diese Datenstruktur zusammen mit den hohen Varianzen der Sentimentwerte innerhalb der Jahre erschwerte die Nutzung von "Lags" und die Analyse saisonaler Trends. Zuletzt ist die Interpretation der Simulationsergebnisse recht beliebig. Dabei ist allerdings hinzuzufügen, dass die Problematiken bei der frequentistischen Analyse sehr großer Datensätze erst seit der wissenschaftlichen Bearbeitung von Daten sozialer Medien zentral diskutiert werden. In der Vergangenheit limitierte der erschwerte Zugang die Verbreitung sehr großer Datensätze.

6.5 **Fazit**

Die Arbeit orientierte sich an der Fragestellung, ob sich in soziologischen Publikationen ein Anstieg des positiven Framings in Form der Zunahme positiver Sentiments und keiner Zunahme negativer Sentiments feststellen lässt. Beobachtete Reaktionen von Forschenden auf Strukturen des akademischen Kapitalismus motivierten die Fragestellung. Die Forschenden setzen sich im intensivierten wissenschaftlichen Wettbewerb durch, indem sie die eigene Forschung mit unterschiedlichen Strategien gegenüber der Gesellschaft und Geldgebern bewerben (Ylijoki, 2003). Die bisherige Forschung zeigt für verschiedene Wissenschaftsdisziplinen einen Anstieg positiver Sentiments und keinen positiven Trend negativer Sentiments (Holtz et al., 2017; Lennox et al., 2020; Liu & Zhu, 2023; Vinkers et al., 2015). Davon ausgehend analysiert die vorliegende Arbeit die Sentiments einer Vollerhebung von Web of Science aller Abstracts zwischen 1993 und 2023 aus der Disziplin Soziologie mit zwei Diktionären unterschiedlichen Umfangs. Die durchgeführten Simulationen zeigen, dass die Signifikanz des vermeintlichen Anstiegs positiver Sentiments auf die hohe Fallzahl zurückzuführen ist. In Kombination mit der geringen Stärke des Effekts und Beachtung der problematischen Verteilung der gemessenen Sentimentwerte, kann die Existenz eines Anstiegs positiver Sentiments nicht bestätigt werden. Die negativen Sentiments weisen keinen klaren Trend auf. Die Ergebnisse betonen, wie entscheidend es ist, Daten sorgfältig auf ihre Übereinstimmung mit den Modellannahmen zu prüfen.

Anhang

.1 Colemanboot Policies

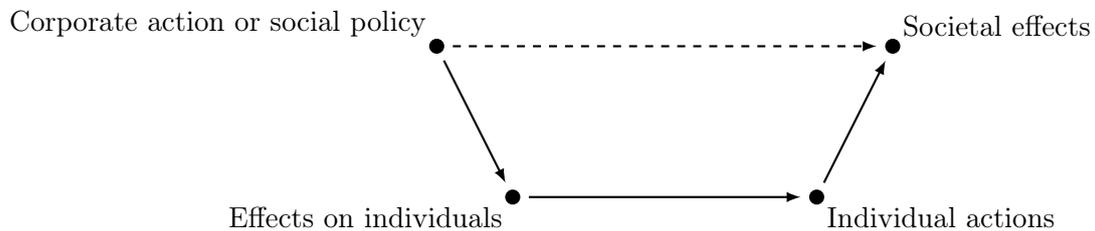


Abbildung 1: Das Colemanboot wird in Coleman (2000) auf Seite 572 eingeführt.

.2 Diktionär: Vinkers et al.

positive Wörter:

Amazing, Assuring, Astonishing, Bright, Creative, Encouraging, Enormous, Excellent, Favourable, Groundbreaking, Hopeful, Innovative, Inspiring, Inventive, Novel, Phenomenal, Prominent, Promising, Reassuring, Remarkable, Robust, Spectacular, Supportive, Unique, Unprecedented

negative Wörter:

Detrimental, Disappointing, Disconcerting, Discouraging, Disheartening, Disturbing, Frustrating, Futile, Hopeless, Impossible, Inadequate, Ineffective, Insignificant, Insufficient, Irrelevant, Mediocre, Pessimistic, Substandard, Unacceptable, Unpromising, Unsatisfactory, Unsatisfying, Useless, Weak, Worrisome

.3 Diktionär: Holtz et al.

Marginal Signifikante Effekte:

marginally significant, $p < .10$, $p < .10$, $p.10$, marginal significant, marginal significance, trend towards significance, trended towards significance, trending towards significance, approaching significance, approached significance

.4 Berechnung der Fehlerfortpflanzung

$$\Delta Y = \sqrt{\sum_{i=1}^m \left(\frac{\partial Y}{\partial X_i} \Delta X_i \right)^2}$$

1. Bilden der partiellen Ableitungen

$$\frac{\partial d_s}{\partial s_{\text{vader}}} = \frac{d}{ds_{\text{vader}}} [s_{\text{vader}} - s_{\text{vink}}] = \frac{d}{ds_{\text{vader}}} [s_{\text{vader}}] + \frac{d}{ds_{\text{vader}}} [-s_{\text{vink}}] = 1 + 0 = 1$$

$$\frac{\partial d_s}{\partial s_{\text{vink}}} = \frac{d}{ds_{\text{vink}}} [s_{\text{vader}} - s_{\text{vink}}] = \frac{d}{ds_{\text{vink}}} [s_{\text{vader}}] - \frac{d}{ds_{\text{vink}}} [s_{\text{vink}}] = 0 - 1 = -1$$

2. Einsetzen in die allgemeine

Fehlerfortpflanzungsformel von Gauß

$$\begin{aligned} \Delta d_s &= \sqrt{(1 \cdot \Delta s_{\text{vader}})^2 + (-1 \cdot \Delta s_{\text{vink}})^2} = \sqrt{\Delta s_{\text{vader}}^2 + (-\Delta s_{\text{vink}})^2} \\ &= \sqrt{\Delta s_{\text{vader}}^2 + \Delta s_{\text{vink}}^2} \end{aligned}$$

5.5 Worthäufigkeiten Vinkers et al. (2015)

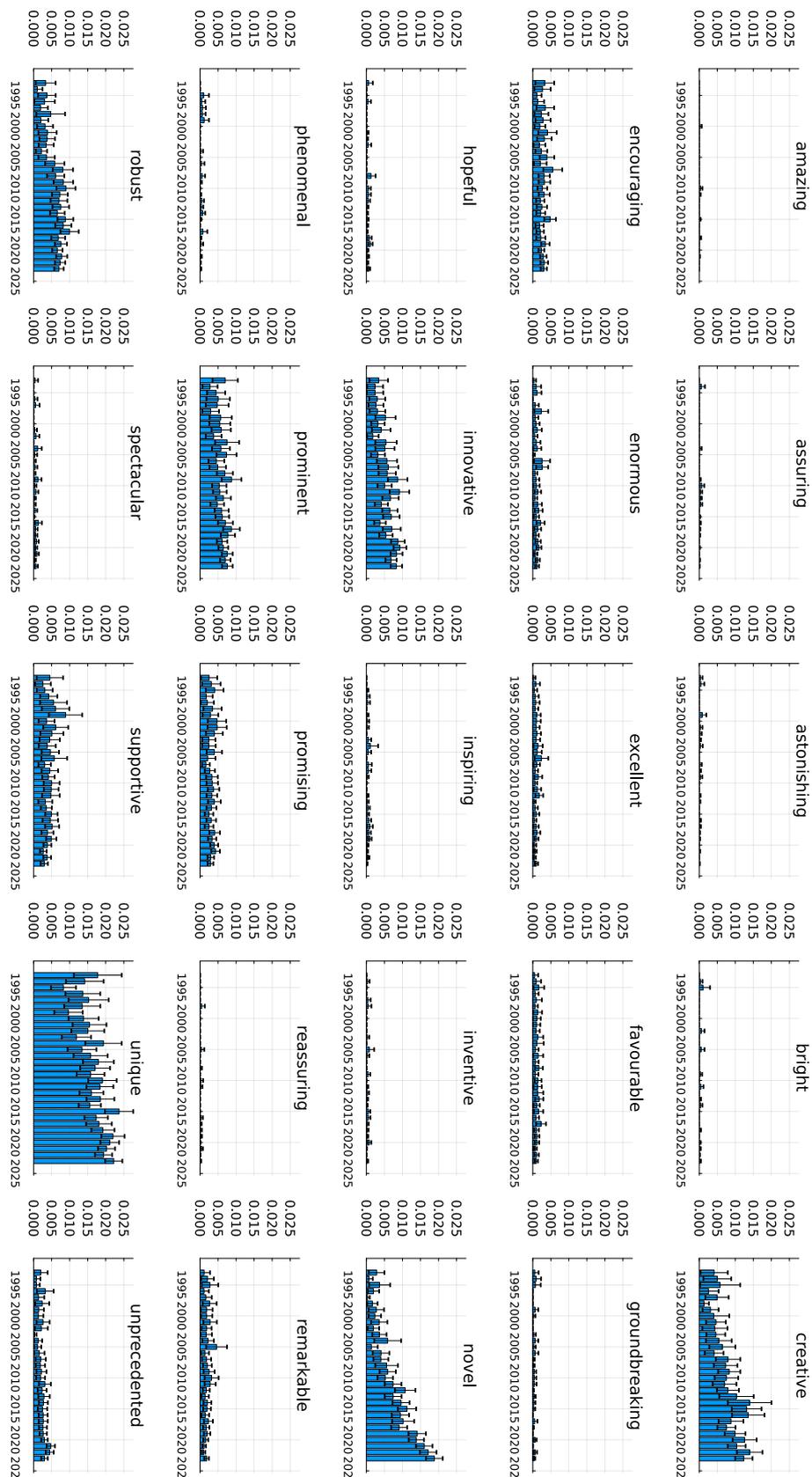


Abbildung 2: Vinkers 2015: Prozentualer Anteil der positive Sentiments pro Wort im Verhältnis zur gesamten Wortmenge des jeweiligen Jahres mit 95% Konfidenzintervallen

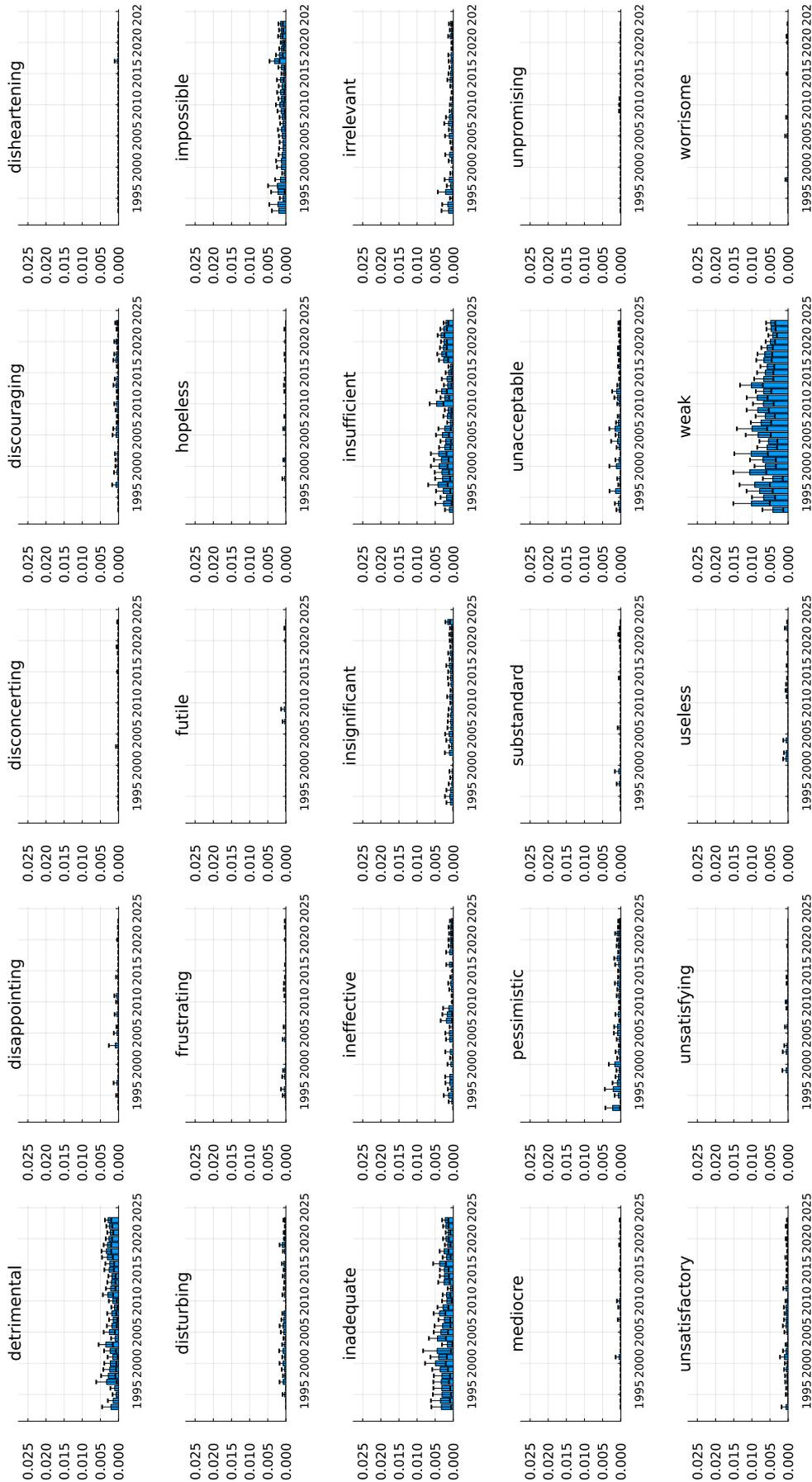


Abbildung 3: Vinkers 2015: Prozentualer Anteil der negativen Sentiments pro Wort im Verhältnis zur gesamten Wortmenge des jeweiligen Jahres mit 95% Konfidenzintervallen

.6 CPS-Plots der Monte-Carlo Simulationen

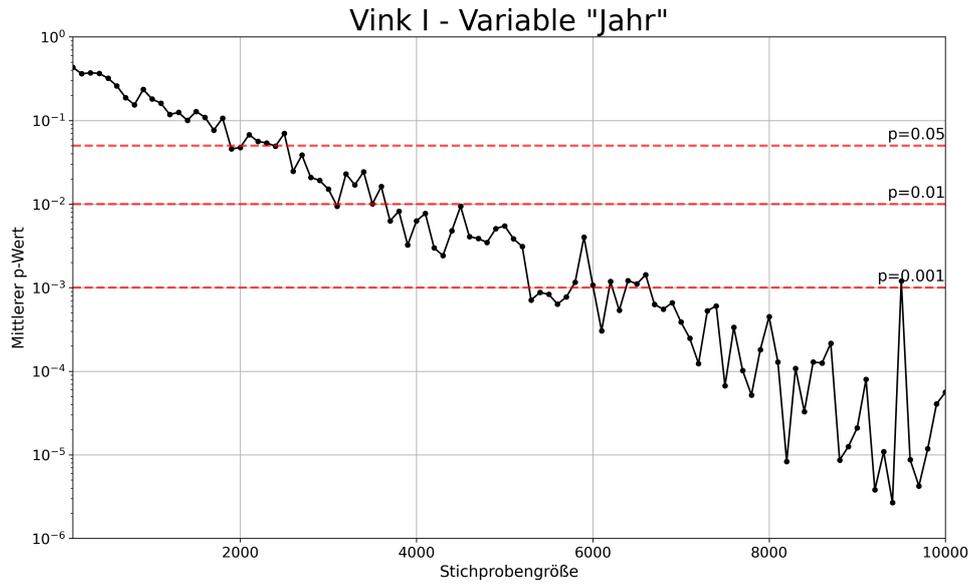


Abbildung 4: CPS-Plot des Modells "Vink I"; p-Werte des Koeffizienten "Jahr" in Abhängigkeit der Stichprobengröße

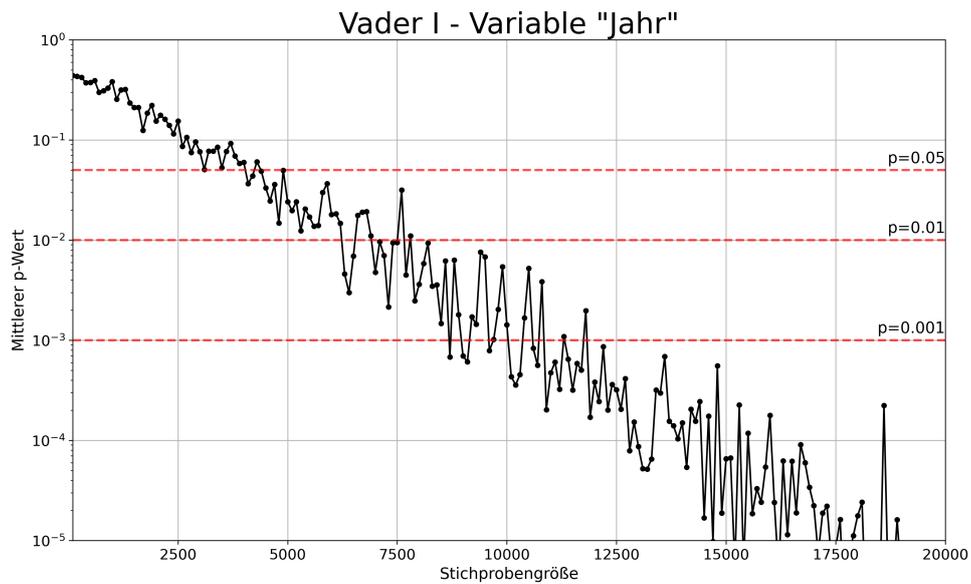


Abbildung 5: CPS-Plot des Modells "Vader I"; p-Werte des Koeffizienten "Jahr" in Abhängigkeit der Stichprobengröße

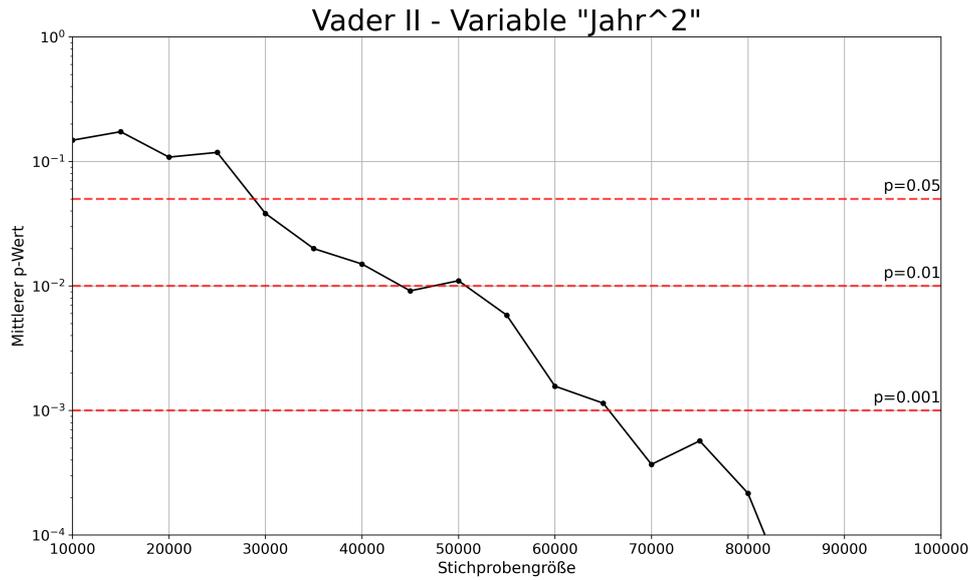


Abbildung 6: CPS-Plot des Modells “Vader II”; p-Werte des Koeffizienten “Jahr²” in Abhängigkeit der Stichprobengröße

.7 Transformation der Zielgröße

	$\sqrt{\text{vink} + 10^{-5}} \text{ I}$	$\sqrt{\text{vink} + 10^{-5}} \text{ II}$	$\sqrt{\text{vader} + 10^{-5}} \text{ I}$	$\sqrt{\text{vader} + 10^{-5}} \text{ II}$
Intercept	-4.2054*** (0.1829)	13.4408 (45.7370)	-8.7060*** (0.6444)	-848.3011*** (161.1437)
Jahr	0.0021*** (0.0001)	-0.0154 (0.0455)	0.0056*** (0.0003)	0.8411*** (0.1603)
Jahr ²		0.0000 (0.0000)		-0.0002*** (0.0000)
N	121789	121789	121758	121758
R ²	0.005	0.005	0.003	0.003

Tabelle 1: Lineare Regressionen beider Diktionäre mit transformierten Daten. Die Standardfehler sind in Klammern dargestellt.
 * p<.1, ** p<.05, ***p<.01

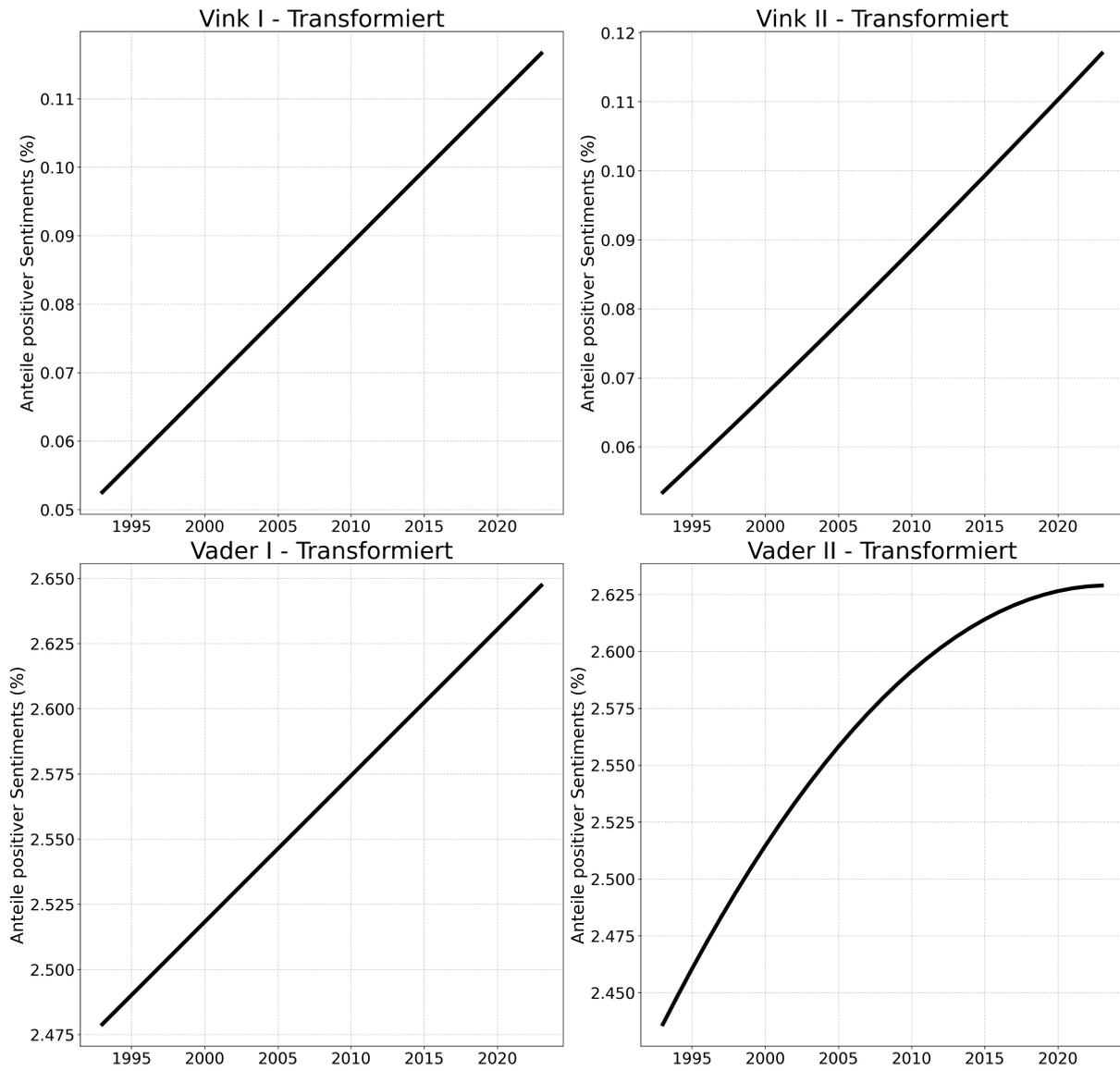


Abbildung 7: Trajektorien der transformierten Modelle

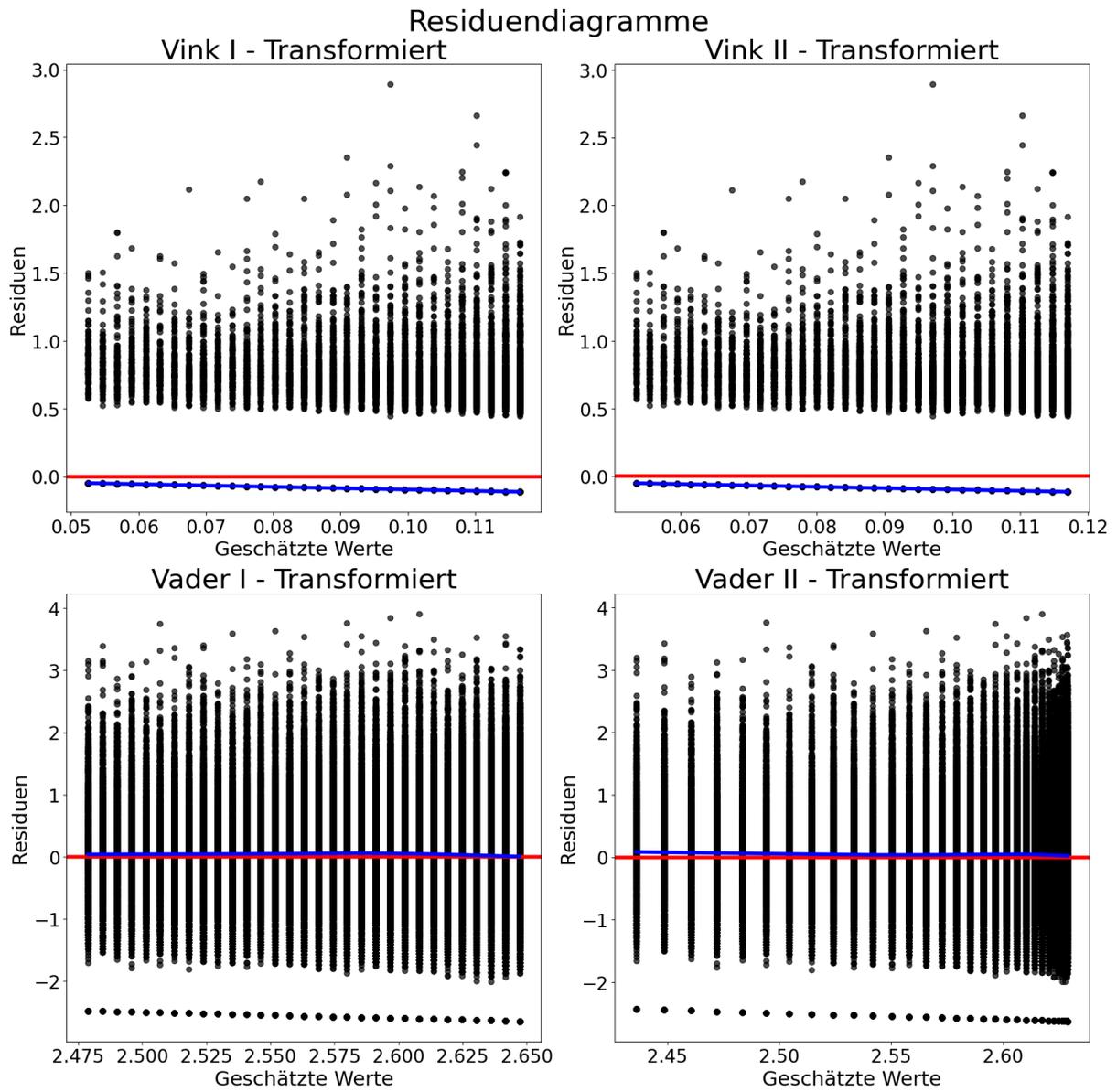


Abbildung 8: Residuendiagramme der Modelle mit transformierten Sentiments. Null wird durch die Rote Linie markiert. Die LOWESS Kurve ist in Blau eingetragen.

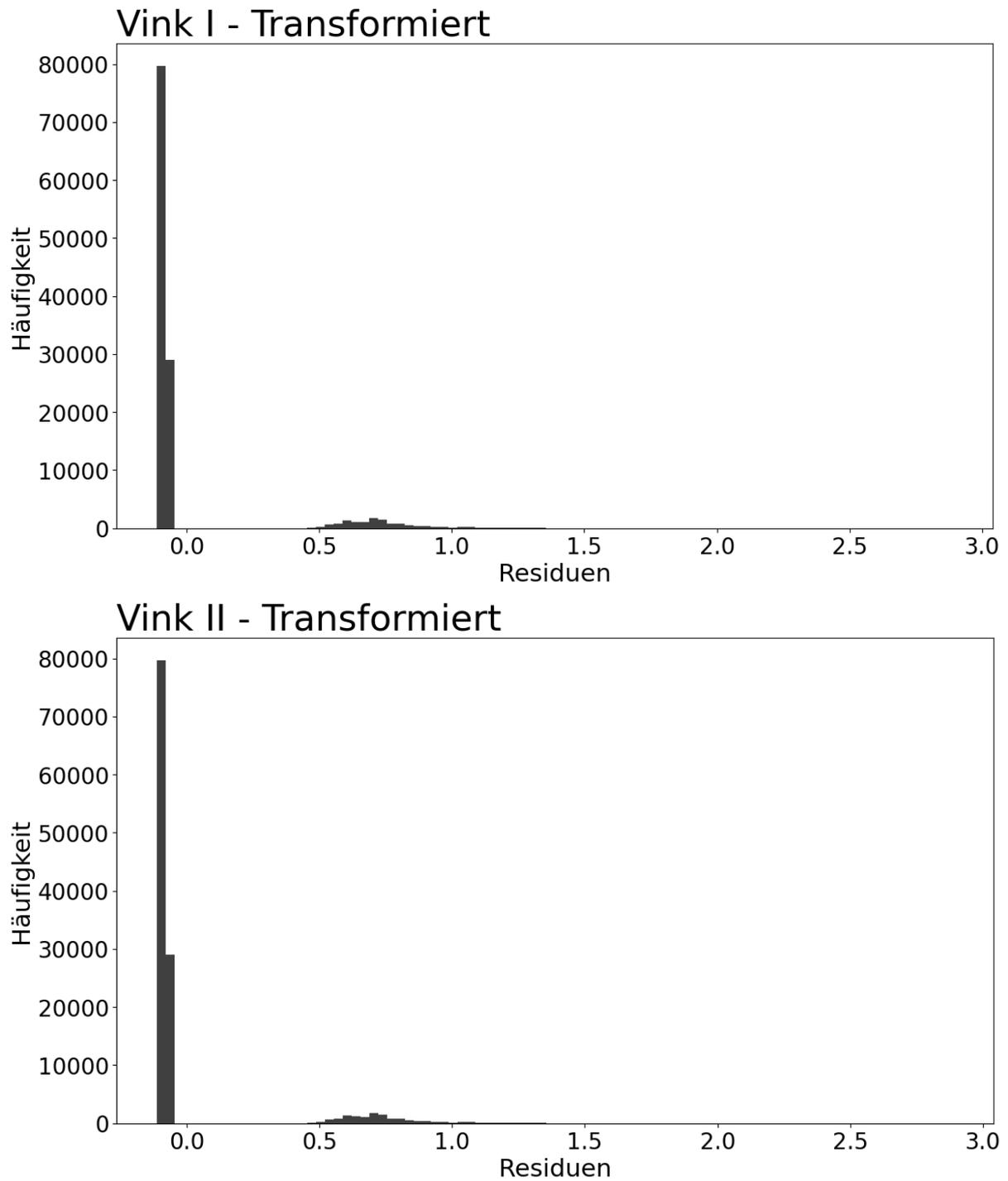


Abbildung 9: Residuenhistogramm der Modelle mit transformierten Sentiments

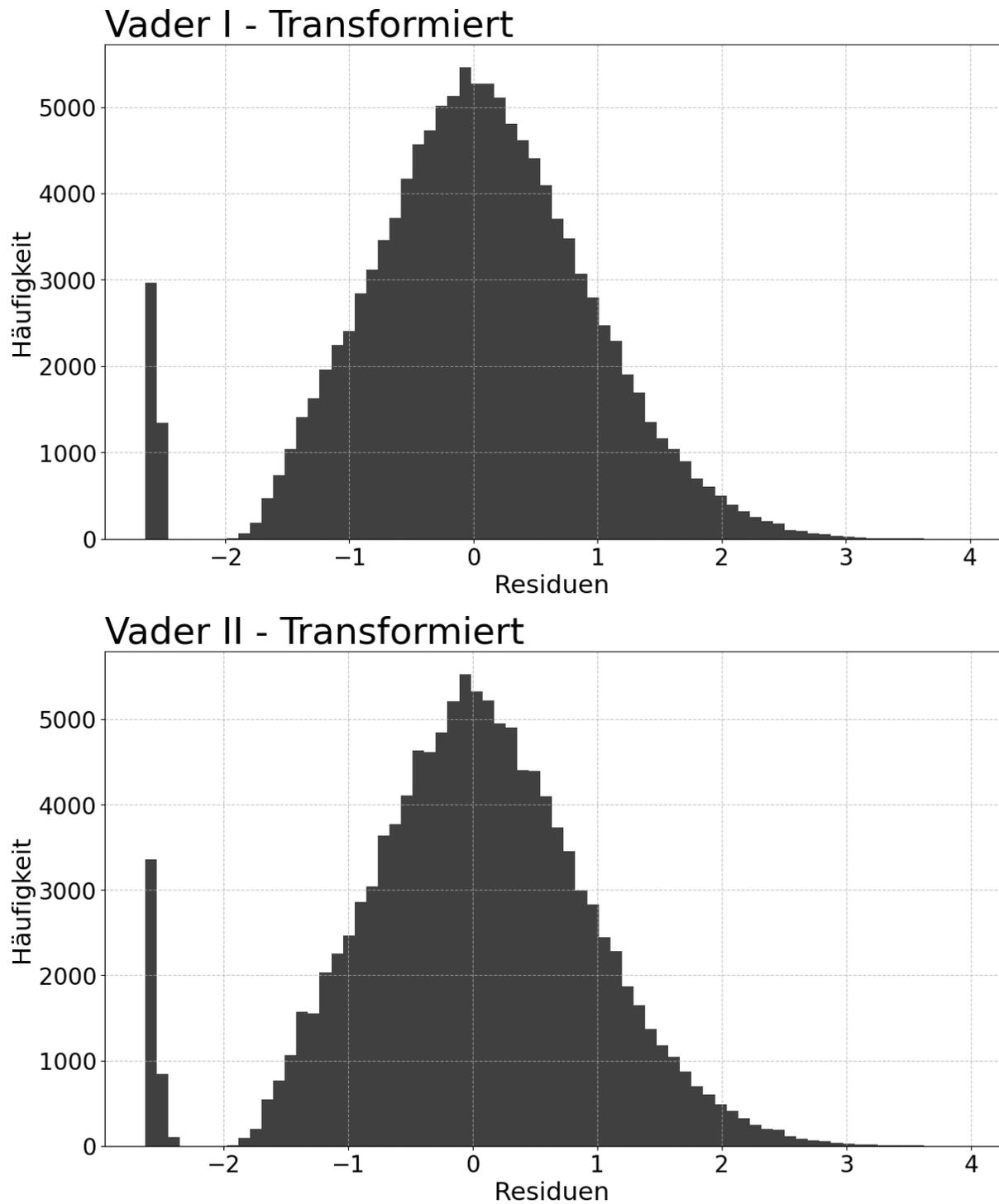


Abbildung 10: Residuenhistogramm der Modelle mit transformierten Sentiments

BA Thesis - "Academic capitalism" analysed: Linguistic trends in sociological journals (#175512)

Created: 05/17/2024 08:37 AM (PT)

Public: 05/29/2024 06:35 AM (PT)

Author(s)

Thomas Haase (Justus-Liebig-Universität Giessen) - thomas.haase@sowi.uni-giessen.de

1) Have any data been collected for this study already?

No, no data have been collected for this study yet.

2) What's the main question being asked or hypothesis being tested in this study?

Can an increase in positive framing, an increase in publications that achieve marginally significant results and an increase in jargon/nominalization be identified in sociological journals as indicators of the increasing influence of academic capitalism?

3) Describe the key dependent variable(s) specifying how they will be measured.

1. The positive framing will be measured through sentiment analysis of abstracts from sociological journals. Two dictionaries will be used for the analysis. The first consists of a small set of custom positive keywords, the second one will be a large sentiment analysis tool similar to LIWC.
2. The frequency of reported marginally significant results will be measured through a dictionary containing phrases which are common to report marginally significant findings.
3. The increase in jargon/nominalization will be measured through counting the use of common verbs as well as text readability scores which measure the complexity of a given text (for example average syllable count per sentence or Flesch's readability score)

4) How many and which conditions will participants be assigned to?

-

5) Specify exactly which analyses you will conduct to examine the main question/hypothesis.

The dataset containing the sociological texts will be split into 3 parts with equally big time intervals. The average measured values will be compared through t-tests. At least for the positive framing measurements the smoothed trend will be plotted to gain deeper insights into the result of the dictionary-based analysis.

6) Describe exactly how outliers will be defined and handled, and your precise rule(s) for excluding observations.

-

7) How many observations will be collected or what will determine sample size? No need to justify decision, but be precise about exactly how the number will be determined.

All articles for the yet-to-be-decided time frame of relevant sociological journals will be used. If data availability is given the time frame of 1970-2020 would be optimal, because the results could be compared to the study of Holtz et al. 2017.

8) Anything else you would like to pre-register? (e.g., secondary analyses, variables collected for exploratory purposes, unusual analyses planned?)

For exploratory purposes additional analysis of the dataset is going to be made (keyness analysis to compare prominent topics of the described subsets of the data for example). This thesis tries to replicate the study "Cross-Cultural Psychology and the Rise of Academic Capitalism: Linguistic Changes in CCR and JCCP Articles, 1970-2014" of Holtz, Deutschmann, Dobewall 2017 for sociological articles with minor modifications in the analysis regarding the questions for sentiments and jargon/nominalization.

Literatur

- Andersen, J. P. (2023). Field-level differences in paper and author characteristics across all fields of science in Web of Science, 2000–2020. *Quantitative Science Studies*, 4(2), 394–422. https://doi.org/10.1162/qss_a_00246 (siehe S. 8)
- Baumeler, C. (2009). Entkopplung von Wissenschaft und Anwendung. Eine neo-institutionalistische Analyse der unternehmerischen Universität / Decoupling Science and Application. A Neo-Institutional Analysis of the Entrepreneurial University. *Zeitschrift für Soziologie*, 38(1), 68–84. <https://doi.org/10.1515/zfsoz-2009-0104> (siehe S. 6 ff., 43)
- Becher, T. (1989). *Academic Tribes and Territories - Intellectual Enquiry and the Cultures of Disciplines*. (Siehe S. 8, 43).
- Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). Quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, 3(30), 774. <https://doi.org/10.21105/joss.00774> (siehe S. 18, 26)
- Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM*, 59, 65–98. <https://doi.org/10.1137/141000671> (siehe S. 18)
- Billig, M. (2013). *Learn to write badly: How to succeed in the social sciences*. Cambridge Univ. Press. (Siehe S. 12).
- Boucher, J., & Osgood, C. E. (1969). The pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 8(1), 1–8. [https://doi.org/10.1016/S0022-5371\(69\)80002-2](https://doi.org/10.1016/S0022-5371(69)80002-2) (siehe S. 30)
- Boyack, K. W., Klavans, R., & Börner, K. (2005). Mapping the backbone of science. *Scientometrics*, 64(3), 351–374. <https://doi.org/10.1007/s11192-005-0255-6> (siehe S. 16 f.)
- Bradford, S. C. (1985). SOURCES OF INFORMATION ON SPECIFIC SUBJECTS (REPRINTED FROM ENGINEERING AN ILLUSTRATED WEEKLY JOURNAL, VOL 137, PG 85-86, 1934). *JOURNAL OF INFORMATION SCIENCE*, 10(4), 176–180. <https://doi.org/10.1177/016555158501000407> (siehe S. 25)
- Broscheid, A., & Gschwend, T. (2005). Zur statistischen Analyse von Vollerhebungen. *PVS*, 46(1), O16–O26. <https://doi.org/10.1007/s11615-005-0241-x> (siehe S. 19)
- Cavnar, W., & Trenkle, J. (2001). N-Gram-Based Text Categorization. *Proceedings of the Third Annual Symposium on Document Analysis and Information Retrieval* (siehe S. 17).
- Chatterjee, S., Hadi, A. S., & Price, B. (2000). *Regression analysis by example* (3. Aufl.). John Wiley & Sons, Inc. (Siehe S. 37).
- Clarivate Analytics. (n. d. a). *Journal Citation Reports - Categories*. Verfügbar 6. August 2024 unter <<https://jcr.clarivate.com/jcr/browse-category-list>>. (Siehe S. 16)

- Clarivate Analytics. (n. d. b). *Web of Science Core Collection*. Verfügbar 6. August 2024 unter <<https://webofscience.help.clarivate.com/en-us/Content/wos-core-collection/wos-core-collection.htm>>. (Siehe S. 16)
- Clarivate Analytics. (2024). *Web of Science*. <<http://www.webofscience.com>>. (Siehe S. 16)
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2015). *Applied multiple regression correlation analysis for the behavioral sciences* (Third edition). Routledge Taylor & Francis Group. (Siehe S. 33, 35, 37).
- Cohen, M. D., March, J. G., & Olsen, J. P. (1972). A Garbage Can Model of Organizational Choice. *Administrative Science Quarterly*, 17(1), 1. <https://doi.org/10.2307/2392088> (siehe S. 4 f.)
- Coleman, J. S. (2000). *Foundations of social theory* (3. print). Belknap Press of Harvard Univ. Press. (Siehe S. 9, VII).
- Dodds, P. S., Clark, E. M., Desu, S., Frank, M. R., Reagan, A. J., Williams, J. R., Mitchell, L., Harris, K. D., Kloumann, I. M., Bagrow, J. P., Megerdooimian, K., McMahon, M. T., Tivnan, B. F., & Danforth, C. M. (2015). Human language reveals a universal positivity bias. *Proc. Natl. Acad. Sci. U.S.A.*, 112(8), 2389–2394. <https://doi.org/10.1073/pnas.1411678112> (siehe S. 30)
- Edlinger, M., Buchrieser, F., & Wood, G. (2023). Presence and consequences of positive words in scientific abstracts. *Scientometrics*, 128(12), 6633–6657. <https://doi.org/10.1007/s11192-023-04864-6> (siehe S. 18)
- Goodhart, C. (1981). Problems of monetary management: The U.K. experience. In *Inflation, depression, and economic policy in the West*. (Siehe S. 6).
- Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Polit. anal.*, 21(3), 267–297. <https://doi.org/10.1093/pan/mps028> (siehe S. 15)
- Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., del Río, J. F., Wiebe, M., Peterson, P., ... Oliphant, T. E. (2020). Array programming with NumPy. *Nature*, 585(7825), 357–362. <https://doi.org/10.1038/s41586-020-2649-2>
- Holtz, P., Deutschmann, E., & Dobewall, H. (2017). Cross-Cultural Psychology and the Rise of Academic Capitalism: Linguistic Changes in *CCR* and *JCCP* Articles, 1970-2014. *Journal of Cross-Cultural Psychology*, 48(9), 1410–1431. <https://doi.org/10.1177/0022022117724902> (siehe S. 11 f., 14, 44)
- Hornik, K., Rauch, J., & Buchta, C. (2023). *Textcat: N-Gram Based Text Categorization*. <<https://CRAN.R-project.org/package=textcat>>. (Siehe S. 17)
R package version 1.0-8

- Hutto, C., & Gilbert, E. (2014). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text. *ICWSM*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550> (siehe S. 18, 26)
- Jockers, M. L. (2015). *Syuzhet: Extract Sentiment and Plot Arcs from Text*. <10.32614/CRAN.package.syuzhet>. (Siehe S. 13)
- Kaldewey, D. (Hrsg.). (2023, 23. Mai). *Wissenschaftsforschung*. De Gruyter. Verfügbar 20. April 2024 unter <<https://www.degruyter.com/document/doi/10.1515/9783110713800-202/html>>. (Siehe S. 1)
- Lehrstuhl Psychologische Methodenlehre. (SS24). *Grundlagen der Diagnostik - Beobachter- und Beurteilerübereinstimmung I* (Lecture). München. Verfügbar 7. August 2024 unter <<https://www.oer.psy.lmu.de/Diag2.html>>. (Siehe S. 19)
- Lennox, R. J., Verissimo, D., Twardek, W. M., Davis, C. R., & Jarić, I. (2020). Sentiment analysis as a measure of conservation culture in scientific literature. *Conservation Biology*, 34(2), 462–471. <https://doi.org/10.1111/cobi.13404> (siehe S. 11, 13 f., 44)
- Leydesdorff, L., Carley, S., & Rafols, I. (2013). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, 94(2), 589–593. <https://doi.org/10.1007/s11192-012-0784-8> (siehe S. 16)
- Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the ISI subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362. <https://doi.org/10.1002/asi.20967> (siehe S. 17)
- Lietz, H. (2020). Drawing impossible boundaries: Field delineation of Social Network Science. *Scientometrics*, 125(3), 2841–2876. <https://doi.org/10.1007/s11192-020-03527-0> (siehe S. 8, 17)
- Lin, M., Jr., H. C. L., & Shmueli, G. (2013). Too Big to Fail: Large Samples and the p-Value Problem. *Information Systems Research*. <https://doi.org/10.1287/isre.2013.04> (siehe S. 20)
- Liu, X., & Zhu, H. (2023). Linguistic positivity in soft and hard disciplines: Temporal dynamics, disciplinary variation, and the relationship with research impact. *Scientometrics*, 128(5), 3107–3127. <https://doi.org/10.1007/s11192-023-04679-5> (siehe S. 11, 13 f., 44)
- Llanos, E. J., Leal, W., Luu, D. H., Jost, J., Stadler, P. F., & Restrepo, G. (2019). Exploration of the chemical space and its three historical regimes. *Proc. Natl. Acad. Sci. U.S.A.*, 116(26), 12660–12665. <https://doi.org/10.1073/pnas.1816039116> (siehe S. 8)
- McCarthy, M. (2015). Superlatives are commonly used in news coverage of cancer drugs, study finds. *BMJ*, 351, h5803. <https://doi.org/10.1136/bmj.h5803> (siehe S. 11)
- McNamara, C., Paci, G., Strba, F., Nabet, J., Tardon, D., Wilk, J., joc-mer, Bergmann, S., Rakic, G., Bailey, D., & Lillqvist, T. (2024, 6. April). *Libexttextcat*. <<https://github.com/LibreOffice/libexttextcat>>. (Siehe S. 17)

- Mejia, C., Wu, M., Zhang, Y., & Kajikawa, Y. (2021). Exploring Topics in Bibliometric Research Through Citation Networks and Semantic Analysis. *Front. Res. Metr. Anal.*, 6. <https://doi.org/10.3389/frma.2021.742311> (siehe S. 17)
- Milojević, S. (2020). Practical Method to Reclassify Web of Science Articles into Unique Subject Categories and Broad Disciplines. *Quantitative Science Studies*, 1(1), 183–206. https://doi.org/10.1162/qss_a_00014 (siehe S. 16 f.)
- Mohammad, S. M., & Turney, P. D. (2013). NRC Emotion Lexicon (siehe S. 15).
- Münch, R. (2011). *Akademischer Kapitalismus : Zur Politischen Ökonomie Der Hochschulreform* (Orig.-Ausg., 1. Aufl.). Suhrkamp. (Siehe S. 3–8).
- Norris, M., & Oppenheim, C. (2007). Comparing Alternatives to the *Web of Science* for Coverage of the Social Sciences' Literature. *Journal of Informetrics*, 1(2), 161–169. <https://doi.org/10.1016/j.joi.2006.12.001> (siehe S. 16 f.)
- of Science Group, W. (Hrsg.). (2019). Web of Science Core Collection Journal Selection Process. *Clarivate Analytics*. Verfügbar 19. Juni 2024 unter <https://clarivate.com/webofsciencegroup/wp-content/uploads/sites/2/2019/08/WS369553747_Fact-Sheet_Core-Collection_V4_Updated1.pdf>. (Siehe S. 16, 43)
- Priem, J., Piwowar, H., & Orr, R. (2022, 16. Juni). *OpenAlex: A Fully-Open Index of Scholarly Works, Authors, Venues, Institutions, and Concepts*. arXiv: 2205.01833 [cs]. <https://doi.org/10.48550/arXiv.2205.01833>. (Siehe S. 17)
- {R Core Team}. (2024). *R: A Language and Environment for Statistical Computing*. Wien, Austria. <<https://www.R-project.org/>>. (Siehe S. 18)
- Ribeiro, F. N., Araújo, M., Gonçalves, P., André Gonçalves, M., & Benevenuto, F. (2016). SentiBench - a benchmark comparison of state-of-the-practice sentiment analysis methods. *EPJ Data Sci.*, 5(1), 23. <https://doi.org/10.1140/epjds/s13688-016-0085-1> (siehe S. 15, 18)
- Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and Statistical Modeling with Python. *9th Python in Science Conference*.
- Silge, J., & Robinson, D. (2017). *Text Mining with R - A Tidy Approach* (1. Aufl.). O'Reilly. (Siehe S. 15).
- Slaughter, S., & Leslie, L. L. (1999). *Academic Capitalism : Politics, Policies, and the Entrepreneurial University* (Paperbacks ed.). Johns Hopkins Univ. Press. (Siehe S. 3–8).
- Stebbins, R. (2001). *Exploratory Research in the Social Sciences*. SAGE Publications, Inc. <https://doi.org/10.4135/9781412984249>. (Siehe S. 9)
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics* (Pearson new international edition, sixth edition). Pearson. (Siehe S. 19).
- Thelwall, M. (2022). Sentiment Analysis. In *The SAGE Handbook of Social Media Research Methods* (2. Aufl., S. 521–530). SAGE Publications. (Siehe S. 15).

- Tipler, P. A., & Mosca, G. (2015). *Physik: für Wissenschaftler und Ingenieure* (J. Wagner, Hrsg.). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-54166-7>. (Siehe S. 20)
- Välimaa, J. (2001). *Finnish higher education in transition: Perspectives on massification and globalisation*. University of Jyväskylä, Institute for Educational Research. (Siehe S. 7 f.)
OCLC: 500352818.
- Van Rossum, Guido and Drake & Fred L. (2009). *Python 3 Reference Manual*. CreateSpace. <10.5555/1593511>. (Siehe S. 18)
- Vinkers, C. H., Tijdkink, J. K., & Otte, W. M. (2015). Use of positive and negative words in scientific PubMed abstracts between 1974 and 2014: Retrospective analysis. *BMJ*, h6467. <https://doi.org/10.1136/bmj.h6467> (siehe S. 11–14, 18, 21, 26, 29–32, 34, 36, 38, 42, 44)
- Volle, J., Schmitz, A., Lietz, H., & Münch, R. (2024). Group Formation in Science between Homogenization and Differentiation: Modeling the Development of U.S. and German Sociology. *International Journal of Sociology*, 54(4), 221–241. <https://doi.org/10.1080/00207659.2024.2357908> (siehe S. 8)
- Weber, M. (1919). *Wissenschaft als Beruf*. (Siehe S. 1, 3).
- Weber, M. (1934). *Die Protestantische Ethik Und Der Geist Des Kapitalismus*. Mohr. (Siehe S. 1).
- Wen, J., & Lei, L. (2022). Linguistic positivity bias in academic writing: A large-scale diachronic study in life sciences across 50 years. *Applied Linguistics*, 43(2), 340–364. <https://doi.org/10.1093/applin/amab037> (siehe S. 11)
- Ylijoki, O.-H. (2003). Entangled in academic capitalism? A case-study on changing ideals and practices of university research (siehe S. 7 f., 43 f.).
- Yuan, Z.-m., & Yao, M. (2022). Is academic writing becoming more positive? A large-scale diachronic case study of Science research articles across 25 years. *Scientometrics*, 127(11), 6191–6207. <https://doi.org/10.1007/s11192-022-04515-2> (siehe S. 11)

Studiengang: _____

Name: _____

Matrikelnummer: _ _ _ _ _

Erklärung zur Abschlussarbeit (Thesis)

Ich erkläre hiermit, dass ich die Thesis selbständig verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, sind durch Angaben und Quellen kenntlich gemacht. Dies gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen und dergleichen.

Weiterhin erkläre ich, dass die Abschlussarbeit (Thesis) noch nicht im Rahmen einer staatlichen oder anderen Prüfung (z. B. als Magister-, Diplom- oder Staatsexamensarbeit) eingereicht wurde.

Mit der Überprüfung meiner Abschlussarbeit mittels einer Anti-Plagiatssoftware bin ich einverstanden und reiche die Abschlussarbeit auch in digitaler Form ein.

(Ort, Datum)

(eigenhändige Unterschrift)