

WISTA

Wirtschaft und Statistik

K. Lorentz | M. Zwick

S. Hoffmann | T. Laick | L. Syme

M. Bellinghen

K. Hofmann

K. Pötzsch | M. Rotsche

R. Sturm | A. Urban

I. Herrle | L. Rohleder

J. A. Czepek | N.-S. Schneider

C. Canan | N. Eppers |

F. Czerniak | R. Mach

E. Stäger

R. Münnich

Die Anschriftenstichprobe des Zensus 2022

Das Lehramtsstudium im Spiegel der Hochschulstatistik

Fakten zur Waldfläche in Deutschland – unterschiedliche
Daten für unterschiedliche Fragestellungen

Personenverkehrsstatistik im Wandel

Large Cases Unit: Erkenntnisse aus fünf Jahren Arbeit

Auswertungen über Unternehmensgruppen im statistischen
Unternehmensregister

Erweiterte Aufkommens- und Verwendungstabellen
– eine Darstellung der Aktivitäten unterschiedlicher
Unternehmenskategorien

Der Data Governance Act

Integrationsberichterstattung im Dashboard Deutschland –
die neue Plattform Dashboard Integration

Neuronale Netze in der Baustatistik: Automatisiertes
Erkennen von Baustellen anhand von Luftbildern

Wissenschaftlicher Nachwuchspreis „Statistical Science
for the Society“ 2024

2 | 2025

ABKÜRZUNGEN

D	Durchschnitt (bei nicht addierfähigen Größen)
Vj	Vierteljahr
Hj	Halbjahr
a. n. g.	anderweitig nicht genannt
o. a. S.	ohne ausgeprägten Schwerpunkt
Mill.	Million
Mrd.	Milliarde

ZEICHENERKLÄRUNG

–	nichts vorhanden
0	weniger als die Hälfte von 1 in der letzten besetzten Stelle, jedoch mehr als nichts
.	Zahlenwert unbekannt oder geheim zu halten
. . .	Angabe fällt später an
X	Tabellenfach gesperrt, weil Aussage nicht sinnvoll
I oder —	grundsätzliche Änderung innerhalb einer Reihe, die den zeitlichen Vergleich beeinträchtigt
/	keine Angaben, da Zahlenwert nicht sicher genug
()	Aussagewert eingeschränkt, da der Zahlenwert statistisch relativ unsicher ist
	Abweichungen in den Summen ergeben sich durch Runden der Zahlen.
	Tiefer gehende Internet-Verlinkungen sind hinterlegt.

INHALT

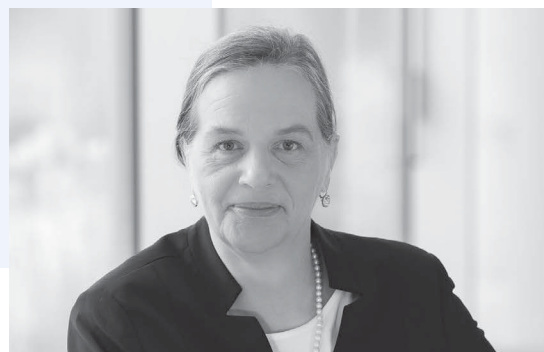
3	Editorial
4	Kennzahlen und Indikatoren
8	Aktuelle Informationsangebote
10	Kurznachrichten
17	Kai Lorentz, Markus Zwick Die Anschriftenstichprobe des Zensus 2022 Ermittlung der Einwohnerzahl in den Sampling Points <i>The sample of addresses in the 2022 Census. Determining the population of the sampling points</i>
30	Stefanie Hoffmann, Tobias Laick, Lorna Syme Das Lehramtsstudium im Spiegel der Hochschulstatistik <i>Teacher training reflected in higher education statistics</i>
43	Marius Bellinghen Fakten zur Waldfläche in Deutschland – unterschiedliche Daten für unterschiedliche Fragestellungen <i>Facts on the forest area in Germany – Different data for different issues</i>
55	Karl Hofmann Personenverkehrsstatistik im Wandel <i>Developments in passenger transport statistics</i>
71	Klaus Pöttsch, Maximilian Rotsche Large Cases Unit: Erkenntnisse aus fünf Jahren Arbeit <i>Large Cases Unit: key findings from five years of work</i>

INHALT

85	Roland Sturm, Adrian Urban Auswertungen über Unternehmensgruppen im statistischen Unternehmensregister <i>Analyses of enterprise groups in the statistical business register</i>
98	Isabel Herrle, Lucas Rohleder Erweiterte Aufkommens- und Verwendungstabellen – eine Darstellung der Aktivitäten unterschiedlicher Unternehmenskategorien <i>Extended supply and use tables – A detailed look at the activities of various enterprise categories</i>
111	Judith Anna Czepek, Nicholas-Sebastian Schneider Der Data Governance Act Auswirkungen auf die vertrauenswürdige Nutzung von Daten öffentlicher Einrichtungen: Rolle eines Datentreuhänders <i>The Data Governance Act. Impact on the trustworthy use of data held by public sector bodies: the role of data trustee</i>
121	Coşkun Canan, Nina Eppers, Frederik Czerniak, Rolf Mach Integrationsberichterstattung im Dashboard Deutschland – die neue Plattform Dashboard Integration <i>Dashboard Integration – New platform for integration reporting on Dashboard Germany</i>
136	Elena Stäger Neuronale Netze in der Baustatistik: Automatisiertes Erkennen von Baustellen anhand von Luftbildern <i>Neural networks in construction statistics: automatic identification of building sites using aerial photos</i>
148	Ralf Münnich Wissenschaftlicher Nachwuchspreis „Statistical Science for the Society“ 2024 <i>Price for junior scientists “Statistical Science for the Society” 2024</i>

EDITORIAL

Dr. Ruth Brand



LIEBE LESERIN, LIEBER LESER,

in dieser Zeitschrift haben sich bereits mehrere Beiträge mit dem Design und der Hochrechnung der Anschriftenstichprobe des Zensus 2022, der daraus folgenden Ermittlung der Bevölkerungszahl sowie mit der Qualität der ermittelten Ergebnisse auseinandergesetzt. Ein Artikel in dieser Ausgabe erörtert nun die Einzelheiten der Hochrechnung sowie Qualitätsbetrachtungen auf der Grundlage der jeweiligen Standardfehler. Er legt dar, dass die Einwohnerzahlermittlung im Zensus 2022 auf einer bewährten und anerkannten Methodik beruht, welche das Bundesverfassungsgericht in seinem Urteil von 2018 bereits als grundsätzlich adäquat bezeichnet hatte.

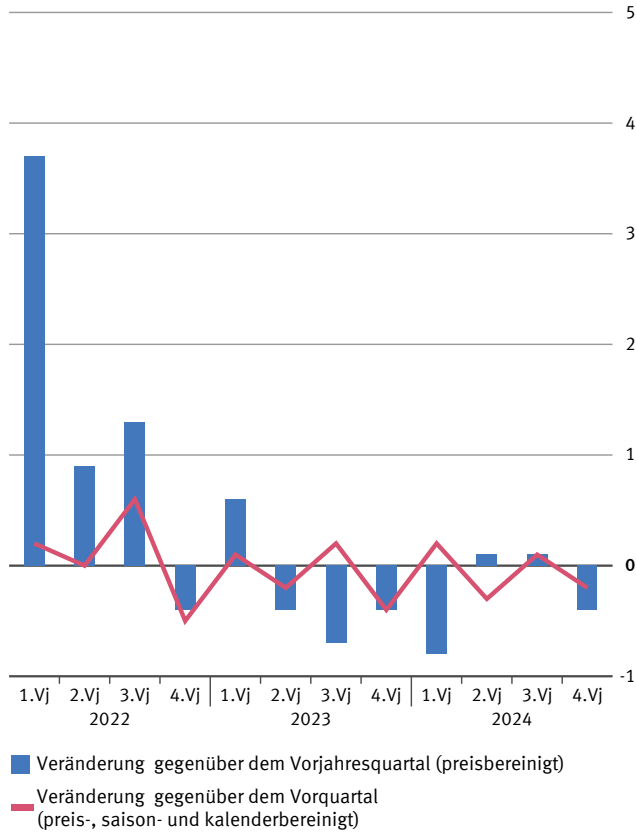
Der große Nutzen der amtlichen Unternehmensstatistiken für die Beobachtung des Wirtschaftsgeschehens ist unbestritten und die Sicherung ihrer Qualität hat hohe Priorität. So untersucht die Large Cases Unit (LCU) der Statistischen Ämter des Bundes und der Länder seit fünf Jahren große multinationale Unternehmensgruppen und ihre Untergliederungen auf Kohärenz der vorliegenden Daten. Ein Aufsatz informiert über die wichtigsten Erkenntnisse der bisherigen LCU-Arbeit. Der Fokus eines weiteren Artikels liegt auf der Darstellung der wirtschaftlichen Bedeutung von Unternehmensgruppen in Deutschland mithilfe ausgewählter Auswertungen aus dem statistischen Unternehmensregister. Auch hier gingen qualitätssichernde Maßnahmen einher mit erweiterten Nutzungsmöglichkeiten.

Einige weitere Themen in dieser Ausgabe unseres Wissenschaftsmagazins WISTA sind Auswertungen zum Lehramtsstudium vor dem Hintergrund des Lehrkräftemangels, Fakten zur Waldfläche in Deutschland sowie Entwicklungsmöglichkeiten der Personenverkehrsstatistik – abwechslungsreicher Lesestoff ist garantiert.

Ruth Brand

Präsidentin des Statistischen Bundesamtes

Bruttoinlandsprodukt
in %

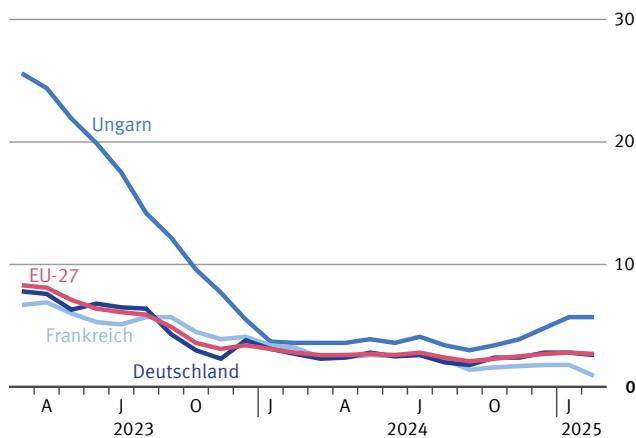


Verbraucherpreisindex
2020 = 100

2024		2025	
Januar	117,6	Januar	120,3
Februar	118,1	Februar	120,8
März	118,6	März	121,2
April	119,2		
Mai	119,3		
Juni	119,4		
Juli	119,8		
August	119,7		
September	119,7		
Oktober	120,2		
November	119,9		
Dezember	120,5		

121,2 **2,2 %**
Veränderung zum Vorjahresmonat

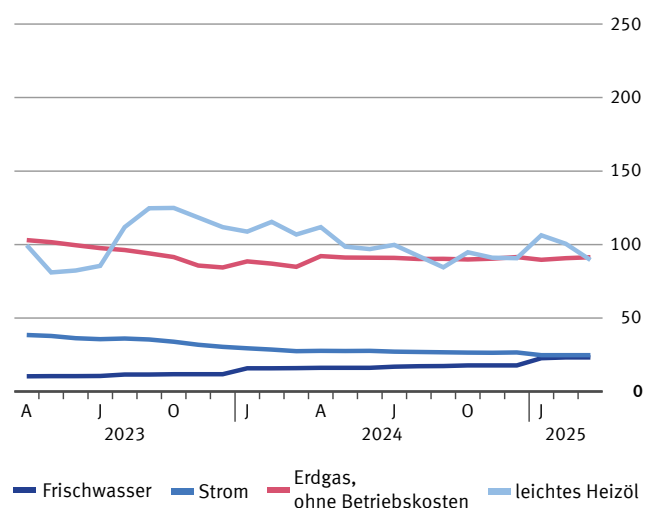
Harmonisierter Verbraucherpreisindex insgesamt
Veränderung gegenüber dem Vorjahresmonat in %



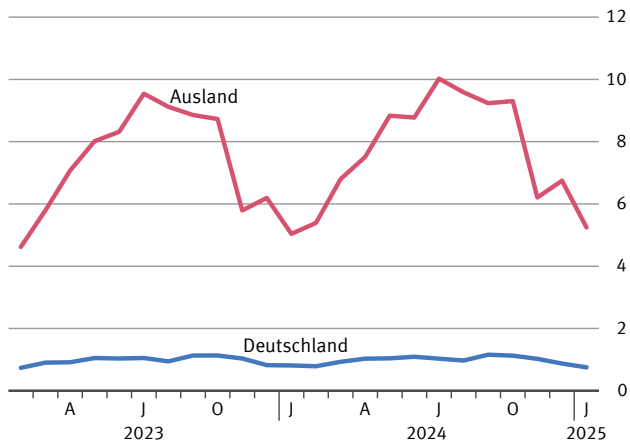
Dargestellt sind neben Deutschland und der Europäischen Union insgesamt (EU-27) die Länder mit der höchsten und der niedrigsten Veränderungsrate innerhalb der EU.

Stand: 11.04.2025

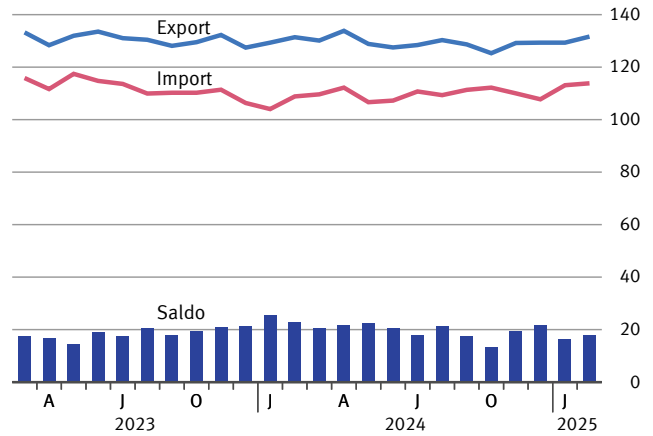
Entwicklung der Verbraucherpreise für Energie und Wasser
Preisabstand in % gegenüber dem Jahr 2020



Einsteigende Flugpassagiere auf deutschen Flughäfen nach dem Streckenziel in Mill.

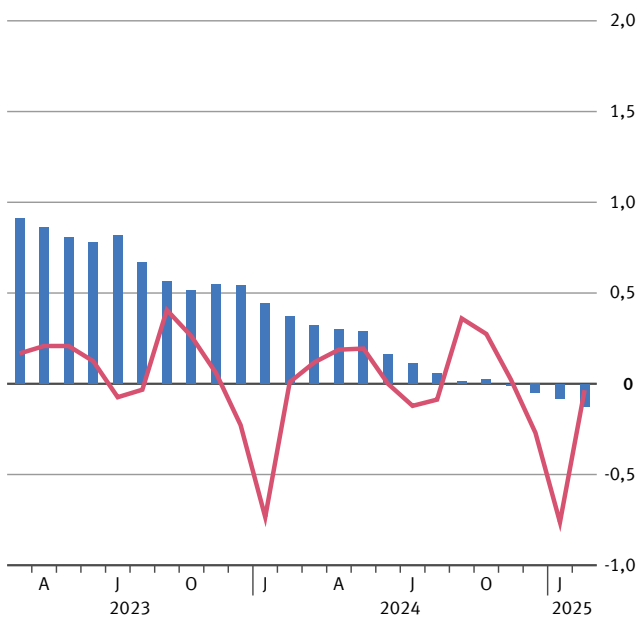


Außenhandel in Mrd. EUR



Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

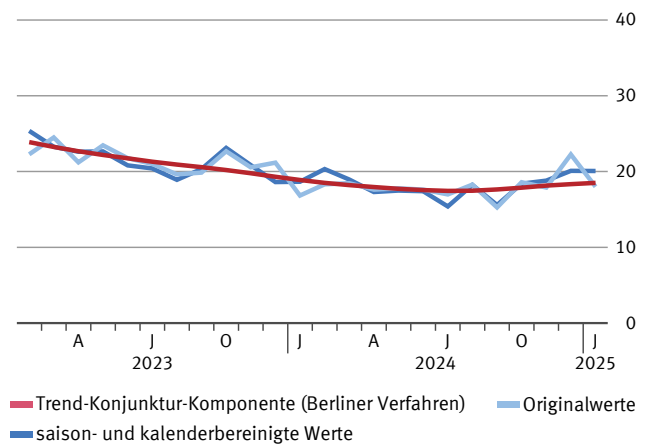
Erwerbstätige mit Wohnort in Deutschland Veränderungen in %



■ Veränderung gegenüber dem Vorjahresmonat
— Veränderung gegenüber dem Vormonat

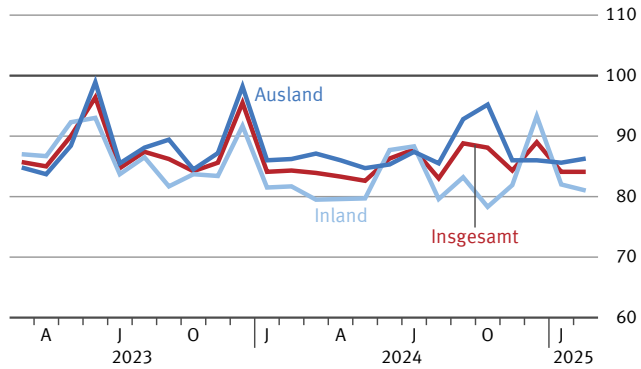
Stand: 11.04.2025

Genehmigte Wohnungen Wohn- und Nichtwohngebäude, in 1 000



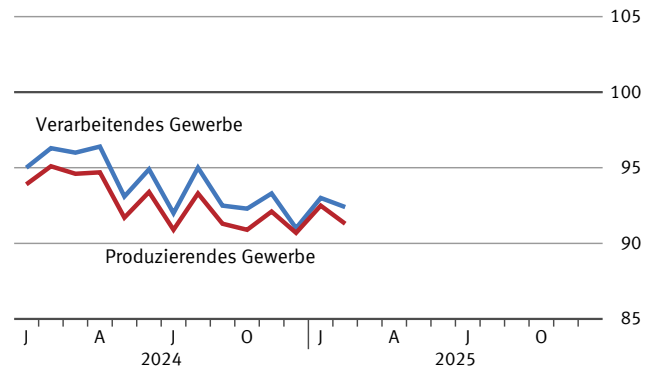
Kennzahlen und Indikatoren

Auftragseingang im Verarbeitenden Gewerbe
Volumenindex 2021 = 100



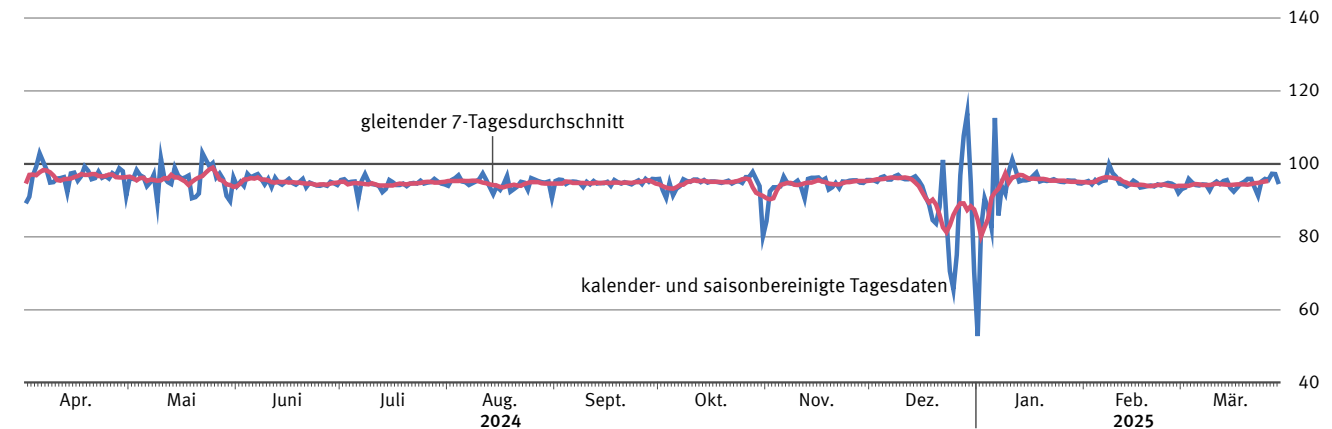
Kalender- und saisonbereinigter Wert nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

Produktion im Produzierenden und Verarbeitenden Gewerbe
Index 2021 = 100



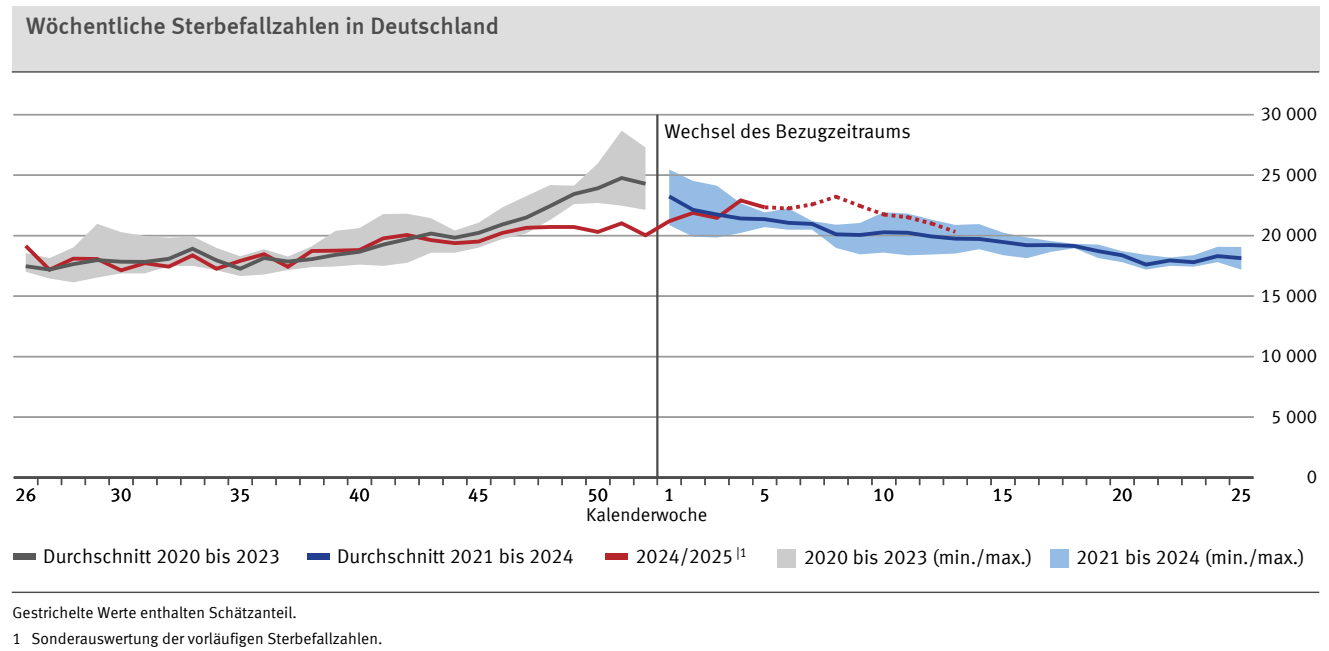
Kalender- und saisonbereinigte Werte nach dem Verfahren X13 JDemetra+. – Vorläufiges Ergebnis.

Lkw-Maut-Fahrleistungsindex
2021 = 100



Quellen: Bundesamt für Logistik und Mobilität, Deutsche Bundesbank, Statistisches Bundesamt

Stand: 11.04.2025



Stand: 11.04.2025



Ukraine

Der Angriff Russlands auf die Ukraine und die damit verbundenen Sanktionen haben starke Auswirkungen auf Wirtschaft und Bevölkerung sowie den Energie-sektor. Auf einer Sonderseite zum Thema stellt das Statistische Bundesamt relevante Daten zur Verfügung. Über die Seite gelangt man auch zum [zentralen Hilfs-portal](#) der Bundesregierung für Geflüchtete aus der Ukraine.

➤ www.destatis.de/Im-Fokus/Ukraine



Dashboard Deutschland

Das 2020 vom Statistischen Bundesamt entwickelte Datenportal wurde im Dezember 2024 erweitert: Es besteht nun aus dem Dashboard Integration, das über 60 Indikatoren zum Integrationsgeschehen in Deutschland umfasst, und dem Dashboard Konjunktur, das hochaktuelle und hochfrequente Zahlen zur Wirtschaftslage in Deutschland bietet. Der integrierte Pulsmesser Wirtschaft ermöglicht intuitives und einfaches Vergleichen von Daten sowie das Erkennen von konjunkturellen Entwicklungen und Zusammenhängen mithilfe täglicher, wöchentlicher, monatlicher und vierteljährlicher Indikatoren. Das Dashboard Deutschland trägt damit zu einem faktenbasierten demokratischen Diskurs der Öffentlichkeit und zur evidenzbasierten Entscheidungsfindung durch Politik und Verwaltung bei.

➤ www.dashboard-deutschland.de



EXSTAT – Experimentelle Statistiken

In der Rubrik „EXSTAT – Experimentelle Statistiken“ veröffentlicht das Statistische Bundesamt regelmäßig neue, innovative Projektergebnisse. Sie entstehen auf der Grundlage neuer Datenquellen und Methoden. Im Reifegrad und in der Qualität unterscheiden sie sich von amtlichen Statistiken, insbesondere in Bezug auf Harmonisierung, Erfassungsbereich und Methodik. Dennoch sind es Ergebnisse der Statistischen Ämter des Bundes und der Länder, die interessante, neue Perspektiven auf verschiedene Themenfelder der Statistik bieten.

➤ www.destatis.de/exstat



Ergebnisse des Zensus 2022

Wie viele Menschen leben in Deutschland? Wie wohnen und arbeiten sie? Die Ergebnisse des Zensus 2022 liefern aktuelle Antworten auf Fragen zur Bevölkerung und zu den Wohnverhältnissen in Deutschland. Alle Daten sind kostenlos und frei zugänglich in der [Zensusdatenbank](https://www.zensus2022.de/zensusdatenbank) verfügbar. Der [Zensusatlas](https://www.zensus2022.de/zensusatlas) bietet eine anschauliche Darstellung.

➤ www.zensus2022.de



Klima

Der Klimawandel ist eine der größten Herausforderungen der heutigen Zeit, alle Bereiche der Gesellschaft sind betroffen. Wie beeinflusst unsere Lebens- und Wirtschaftsweise das Klima? Wie wirkt sich die Umstellung hin zu mehr Klimaschutz gesamtgesellschaftlich aus? Was bedeutet sie für unseren Alltag – vom Weg zur Arbeit bis zum aktuellen Strompreis? Wo zeigen sich die Folgen des Klimawandels? Daten und Fakten zum Thema Klima, Klimawandel und Klimaschutz sind gebündelt unter

➤ www.destatis.de/klima



Fachkräfte

Fachkräftemangel und Arbeitskräftebedarf sind zunehmend wichtige Faktoren für die wirtschaftliche Entwicklung in Deutschland. Daten und Fakten dazu bündelt das Statistische Bundesamt auf einer eigenen Sonderseite. Das Angebot umfasst die Bereiche Demografie, Erwerbstätigkeit, Bildung und Zuwanderung – und wird sukzessive erweitert.

➤ www.destatis.de/fachkraefte

KURZNACHRICHTEN

AUS ALLER WELT

56. Sitzung der Statistischen Kommission der Vereinten Nationen

Die Statistische Kommission der Vereinten Nationen (UNSC) spielt eine zentrale Rolle bei der Weiterentwicklung, Harmonisierung und Implementierung internationaler statistischer Methoden und Verfahren. Als Gremium des Wirtschafts- und Sozialrats der Vereinten Nationen (ECOSOC) repräsentiert sie die statistischen Institutionen der Mitgliedstaaten auf Leitungsebene. Mit derzeit noch 24 nach einem Regionalschlüssel gewählten Mitgliedern begleitet die UNSC die Arbeit der Statistikabteilung der Vereinten Nationen (UNSD). Bis zum Jahr 2028 erfolgt ein schrittweises Anheben der Mitgliederzahl auf 54 Mitglieder. Das Statistische Bundesamt vertritt gemäß § 19 Bundesstatistikgesetz die deutsche amtliche Statistik in diesem Gremium und war bis Ende 2024 gewähltes Mitglied; eine erneute Kandidatur wird voraussichtlich für das Jahr 2026 angestrebt.

Wichtige Ergebnisse der Sitzung vom 4. bis 7. März 2025 sind:

- › Verabschiedung der ECOSOC-Resolution zu einem revidierten System der Volkswirtschaftlichen Gesamtrechnungen (SNA 2025)
Die Mitgliedstaaten des Europäischen Statistischen Systems begrüßten die Berücksichtigung neuer Entwicklungen im SNA 2025. Sie betonten, dass weitere methodische Arbeiten in den kommenden Jahren notwendig sind, um die internationale Vergleichbarkeit und die Belastbarkeit der Ergebnisse der Volkswirtschaftlichen Gesamtrechnungen (VGR) künftig sicherzustellen. Die Resolution wurde einstimmig ver-

abschiedet. Die UNSD wird Bedenken hinsichtlich der Implementierung neuer Indikatoren berücksichtigen.

- › Vorbereitung der kommenden Zensusrunde 2030

Der Resolutionsentwurf unterstrich die Notwendigkeit, administrative Daten stärker zu nutzen, eine bessere Georeferenzierung zu erreichen und nationale statistische Bevölkerungsregister aufzubauen; er wurde leicht revidiert nach weiteren Abstimmungsgesprächen final verabschiedet. Die UNSC-Sitzungen der kommenden Jahre werden die nächste Zensusrunde weiter inhaltlich behandeln.

- › Intensive Diskussion zur Reform der Arbeitsweise der UNSC

Ein Vorschlag betraf die Einrichtung einer Arbeitsgruppe zur Data Governance, um gemeinsame Terminologien und Antworten auf die gestiegenen Anforderungen an nationale Statistikämter im Bereich Datenökosysteme zu entwickeln. Deutschland betonte die Wichtigkeit, den fachlichen Charakter der UNSD zu erhalten und warnte davor, dass die Öffnung des Gremiums für weitere „Datenexperten“ möglicherweise die Unabhängigkeit der amtlichen Statistik beeinträchtigen könnte. Die Statistische Kommission stimmte der Einrichtung einer Arbeitsgruppe zu und wird die Diskussionen in den nächsten Monaten fortsetzen.

- › Einsetzung eines unabhängigen Beirats bei den Fundamental Principles of Official Statistics (FPOS)

Der Beirat soll über die Einhaltung der Empfehlungen der FPOS wachen, auf die Nichteinhaltung der FPOS hinweisen und Maßnahmen zur Besserung vorschlagen. Der Vorschlag der Mitgliedstaaten des Europäischen Statistischen Systems, die Zusammensetzung des Gremiums um Personen zu ergänzen, die mit der Weiterentwicklung von nationalen statistischen

Ämtern vertraut sind, wurde in der überarbeiteten Fassung berücksichtigt und verabschiedet. Zudem soll das UNSC-Büro das Auswahlverfahren für die Mitglieder des Beirats einleiten; die dann neuen Mitglieder sollen in der Sitzung der UNSC 2026 bestätigt werden.

- › Jährlicher Sachstandsbericht zu den Indikatoren der Nachhaltigen Entwicklung (SDG)

Seit der Verabschiedung der UN-Agenda 2030 im Jahr 2016 werden im jährlichen Turnus die Qualität und Brauchbarkeit einzelner Indikatoren überprüft und Anpassungen am Indikatoren-Programm vorgenommen. Deutschland betonte den allgemeinen Nutzen der Entwicklung der Berichterstattung zur Agenda 2030 für die amtliche Statistik und kündigte die Vorstellung eines freiwilligen nationalen Berichts auf dem High Level Political Forum 2025 an.

- › Bericht der sogenannten Friends of the Chair Group zu den Sozial- und Bevölkerungsstatistiken

Ziel ist die Konzeption eines Rahmenwerks, um die Kohärenz und Harmonisierung der Statistiken sicherzustellen. Die Gruppe soll ihren Abschlussbericht auf der 57. Sitzung der UNSC 2026 vorlegen.

- › Bericht der UN-Expertengruppe Migrationsstatistiken für 2023 und 2024

Deutschland unterstützte die Klärung der Begriffe Migration, vorübergehende Mobilität und Tourismus und betonte die Herausforderungen bei der Erhebung spezifischer Daten. Beim 4. Internationalen Forum für Migrationsstatistik im Juni 2025 in Schweden werden die Herausforderungen bei der Erhebung und Nutzung von Migrationsdaten sowie entsprechende Lösungen und erfolgreiche Praktiken erörtert.

- › Bericht des UN Network of Economic Statisticians zur Harmonisierung der Empfehlungen für wirtschaftliche Statistiken

Deutschland begrüßte die Bemühungen, KI-Anwendungen zur Unterstützung der Erstellung von Wirtschaftsstatistiken zu nutzen, betonte jedoch, dass KI-Modelle gegenwärtig noch nicht die Genauigkeit der ARIMA-Modelle erreichen, die derzeit in Deutschland für die Nowcasts des Bruttoinlandsprodukts verwendet werden.

- › Bericht des Expertenkomitees zur Aktualisierung des Systems für umweltökonomische Rechnungslegung (SEEA)

Deutschland unterstützt aktiv den Aktualisierungsprozess und betont die Notwendigkeit einer engen Zusammenarbeit zwischen dem UN Committee of Experts on Environmental-Economic Accounting und der UN London Group zur Umweltstatistik. Die Aktualisierung soll sicherstellen, dass das SEEA weiterhin auf neue politische Anforderungen reagieren kann. Dazu gehören die Themen Klimawandel, Kreislaufwirtschaft und Biodiversität. Der Aktualisierungsprozess wird durch konkrete Beiträge und direkte Finanzierung von Maßnahmen (Aufbau von Kapazitäten) getragen. Das dann revidierte SEEA wird voraussichtlich 2028 der UNSC zur Genehmigung vorgelegt.

- › Bericht des Ausschusses der Experten für Unternehmens- und Handelsstatistiken zu den Fortschritten bei der Integration von Unternehmens- und Handelsstatistiken

Die UNSC unterstützte das geänderte Mandat des Ausschusses zur Stärkung der Integration mit anderen statistischen Bereichen. Diese Entwicklungen und Herausforderungen zeigen die dynamische Natur der internationalen Statistik und die Bedeutung der Zusammenarbeit zwischen den Mitgliedstaaten und internationalen Organisationen, um robuste qualitativ hochwertige und zuverlässige statistische Daten zu gewährleisten. Es wurden erhebliche Fortschritte bei der Überarbeitung der Handbücher für Handelsstatistiken erzielt, die voraussichtlich 2025 abgeschlossen werden.

Die Sitzungsunterlagen stehen zur Verfügung unter:

➤ unstats.un.org

AUS EUROPA

58. Sitzung des AESS

Die 58. Sitzung des Ausschusses für das Europäische Statistische System (AESS) fand am 6. Februar 2025 in Luxemburg statt. Insbesondere strategische Themen der Weiterentwicklung des Europäischen Statistischen Systems (ESS) standen auf der Tagesordnung:

- › Das Statistische Amt der Europäischen Union (Eurostat) hat für die 56. Sitzung der Statistischen Kommission der Vereinten Nationen (Bericht siehe oben) den Entwurf einer gemeinsamen Stellungnahme des Europäischen Statistischen Systems (ESS) für die Tagesordnungspunkte National Accounts und Fundamental Principles of Official Statistics (FPOS) vorgelegt.

- › Vorschlag für eine Verordnung über die europäische Fischerei- und Aquakulturstatistik (EFAS)
Nach dem vorgelegten Entwurf sollen die Fischereistatistiken künftig unter der gemeinsamen EFAS-Rahmenverordnung zusammengefasst werden. Hauptziel ist, die davon betroffenen Statistiken zu harmonisieren (Definitionen, Merkmale), zu modernisieren (Erhebungsmethoden) und neue Datenanforderungen der Europäischen Union (EU) zu decken, insbesondere zum Green-Deal und zum EU-Aktionsplan für Organische Produktion. Deutschland kritisierte die zu erwartenden Mehrbelastungen und das Ausklammern wichtiger Aspekte aus der Diskussion in den statistischen Gremien. Einige EU-Mitgliedstaaten äußerten zudem rechtliche Bedenken hinsichtlich des Datenaustauschs mit internationalen Organisationen.

- › Vision 2030

Der AESS behandelte das Fazit eines dem AESS vorgeschalteten Workshops, in dem die geplante Vision 2030 eingehend erörtert wurde. Die Vision hat zum Inhalt, Prioritäten im Arbeitsprogramm entlang des europäischen Mehrjahresprogramms zu setzen. Darüber hinaus werden Aspekte der Weiterentwicklung aufgegriffen, wie der Einsatz von Künstlicher Intelligenz und die Möglichkeiten der Nutzung privater Daten in der amtlichen Statistik.

- › Implementierung der überarbeiteten EU-Statistikverordnung Nr. 223/2009

Eurostat startet eine entsprechende Initiative und legte dieser Sitzung einen Aktionsplan vor, der konkrete Umsetzungsmaßnahmen mit Zeitplänen vorsieht. Die Umsetzungsmaßnahmen beziehen sich unter anderem auf einen nachhaltigen Zugang zu administrativen, öffentlich zugänglichen und privat gehaltenen Daten, auf statistische Antworten in Krisensituationen, auf die Förderung des Datenaustauschs zwischen dem Europäischen Statistiksistem der Zentralbanken (ESZB) und dem ESS sowie auf die neuen Rollen der Nationalen Statistischen Ämtern in den nationalen Governance-Rahmen. Das von der polnischen EU-Ratspräsidentschaft im April 2025 ausgerichtete High-Level-Meeting wird sich dem vorliegenden Entwurf des Aktionsplans eingehend widmen. Ein detaillierter Zeitplan wird dem AESS Ende Juni 2025 vorgelegt. Der AESS stimmte dem Vorgehen unter dem Hinweis zu, dass für die Umsetzung der Implementierung den Nationalen Statistischen Ämtern keine zusätzlichen personellen Ressourcen zur Verfügung stehen.

- › Sachstandsbericht zu den Common European Data Spaces (CEDS) der EU-Kommission

Die CEDS gibt es unter anderem in den Bereichen Mobilität, Gesundheit, Tourismus, Skills, Green Deal, Landwirtschaft und Energie (digital-strategy.ec.europa.eu). Aktuell ist kein CEDS in dem Stadium, um tatsächlich einen Datenaustausch vorantreiben zu können. Um sicherzustellen, dass Eurostat über aktuelle Entwicklungen informiert bleibt, werden eine Beobachterrolle im European Data Innovation Board (EDIB) sowie eine Mitgliedschaft im Stakeholder-Netzwerk des Data Spaces Support Centre (DSSC) angestrebt. Zudem werden die Direktorengruppen und der AESS regelmäßig über die weiteren Entwicklungen bei den CEDS informiert. Damit wird eine vom Statistischen Bundesamt – zusammen mit den Statistikämtern Österreichs und der Schweiz – vorgebrachte Forderung umgesetzt.

- › 20 Jahre Code of Practice

Den 20. Jahrestag des erstmals 2005 verabschiedeten EU-Verhaltenskodex für europäische Statistiken (Code of Practice – CoP) wird eine zusammen mit der dänischen EU-Ratspräsidentschaft organisierte und im September/Oktober 2025 in Brüssel stattfindende Konferenz angemessen würdigen. Teilnehmen sollen

insbesondere Vertreterinnen und Vertreter der EU, des Europäischen Beratungsgremiums für die Statistische Governance (ESGAB) und der Europäischen Zentralbank. Ein inhaltlicher Schwerpunkt wird die professionelle Unabhängigkeit der Nationalen Statistischen Ämter sein. Darüber hinaus soll ein Kommunikationspaket den 20. Jahrestag begleiten. Ziel ist, den Bekanntheitsgrad und das Bewusstsein für den Code of Practice allgemein zu erhöhen und vor allem die professionelle Unabhängigkeit, Integrität und Rechenschaftspflicht der Nationalen Statistischen Ämter zu stärken. Zum 20. Jahrestag des Europäischen Verhaltenskodex in der amtlichen Statistik soll es zudem bei der diesjährigen Deutschen Statistischen Woche, die vom 2. bis 7. September im Statistischen Bundesamt in Wiesbaden stattfinden wird, eine Panelsitzung geben.

- › Die Konferenz der Amtsleitungen der Nationalen Statistischen Ämter (DGINS) wird in diesem Jahr von Statistics Croatia ausgerichtet und am 22. und 23. Oktober 2025 in Split stattfinden. Die DGINS-Konferenz 2025 ist als Diskussionsforum über die wesentlichen Richtungsvorgaben der künftigen europäischen Tourismusstatistik geplant.

Europäischer Statistikwettbewerb 2025

Am Europäischen Statistikwettbewerb 2025 nehmen zeitgleich Teams aus 21 EU-Staaten teil. In Deutschland hat das Statistische Bundesamt zusammen mit der Bundesagentur für Arbeit die inzwischen abgeschlossene nationale Phase durchgeführt.

Zwei Schulen aus Nordrhein-Westfalen (Kategorie A, Klassenstufen 11, 12 und 13) und Sachsen (Kategorie B, Klassenstufen 8, 9 und 10) haben die nationale Phase für sich entschieden und sind somit automatisch für die europäische Phase qualifiziert. In dieser Runde wird ein Video zum Thema „Is Gen Z changing the world?“ gedreht. Zwei weitere Startplätze werden unter den Finalistenteams von einer Jury vergeben. Insgesamt werden somit vier Videos aus Deutschland für das europäische Finale nominiert.

➤ www.destatis.de

AUS DEM INLAND

Empfehlungen des Statistischen Beirats für die 21. Legislaturperiode

Der Statistische Beirat berät das Statistische Bundesamt in Grundsatzfragen und vertritt dabei die Belange der Nutzenden, Befragten und Produzierenden der Bundesstatistik. Für die 21. Legislaturperiode hat er richtungsweisende Empfehlungen zur Fortentwicklung der amtlichen Statistik formuliert.

Das Beratungsgremium des Statistischen Bundesamtes unterstreicht dabei die zentrale Rolle der amtlichen Statistik als unverzichtbares Fundament für faktenbasierte Entscheidungen in Politik, Wirtschaft und Gesellschaft. Zunehmende globale Krisen erfordern die rasche Verfügbarkeit präziser und belastbarer Daten als verlässliche Informationsbasis, um kurzfristige Maßnahmen und langfristige Strategien zu gestalten. Der Statistische Beirat betont daher die Notwendigkeit, die rechtlichen und organisatorischen Rahmenbedingungen der amtlichen Statistik zu modernisieren. Dazu zählen die Anpassung des Bundesstatistikgesetzes und das Vorantreiben weiterer Einzelgesetze, wie etwa des Forschungsdatengesetzes. Auch empfiehlt der Beirat, eine mittelfristige Programm- und Finanzplanung für das Statistische Bundesamt sowie Maßnahmen zur Gewährleistung der Qualität der amtlichen Statistik einzuführen und gleichzeitig die Berichtspflichten der Auskunftgebenden zu reduzieren. Dies soll beispielsweise erreicht werden, indem Statistiken mit vorhandenen Register- oder Verwaltungsdaten verknüpft sowie neue Datenquellen erschlossen werden.

➤ www.destatis.de

Neuer Zugangsweg der Forschungsdatenzentren

Seit Januar 2025 wird das Datenangebot der Forschungsdatenzentren durch Remote-Scientific-Use-Files (Remote-SUF) ergänzt. Datennutzende können so Zugang zu faktisch anonymen Mikrodaten über das neue Remote-Access-System erhalten. Auf das Remote-Access-System darf nur innerhalb der Räume vertraglich festgelegter

wissenschaftlicher Einrichtung mit dienstlicher Hardware zugegriffen werden. Zunächst stehen Remote-SUF für die fallpauschalenbezogene Krankenhausstatistik (DRG-Statistik) und den Mikrozensus zur Verfügung.

➤ www.forschungsdatenzentrum.de

VERANSTALTUNGEN

Call for Papers für die Statistische Woche 2025

Die Statistische Woche 2025 wird von der Deutschen Statistischen Gesellschaft gemeinsam mit dem Verband Deutscher Städtestatistiker vom 2. bis 5. September 2025 veranstaltet. Die Deutsche Gesellschaft für Demographie wird mit weiteren Sektionen vertreten sein. Sie findet im Statistischen Bundesamt in Wiesbaden mit folgenden Schwerpunktthemen statt:

- › Mikro- und makroökonomische Verteilungsfragen zu Einkommen, Vermögen und Konsum
- › Schätzung von Unsicherheit
- › High-dimensional Time Series and Network Data

Im Rahmen der Jahrestagung tagen die Ausschüsse der Deutschen Statistischen Gesellschaft insbesondere zu den Themen:

- › Computational Statistics and Data Science
- › Empirische Wirtschaftsforschung und angewandte Ökonometrie
- › Finanzstatistik
- › Kompetenzentwicklung: Data Literacy und Statistik
- › Methodik Statistischer Erhebungen
- › Regionalstatistik
- › Statistik in Umweltwissenschaften, Naturwissenschaften und Technik
- › Statistische Theorie und Methodik
- › Nonparametric and Robust Statistics

- › Wirtschafts-, Sozial- und Marktstatistik
- › Arbeitsmarkt und Soziale Sicherung
- › Volkswirtschaftliche Gesamtrechnungen, Wohlfahrtsmessung
- › (Systematic) Risk Modeling
- › Chancen, Grenzen und Entwicklungen der Konjunkturstatistiken
- › Joint DStatG and TIES Session
- › Statistics in Industry
- › VDSt: tba
- › DGD: Projektionen im Kontext demografischer und gesellschaftlicher Entwicklungen
- › DGD: Regionale, sozioökonomische und geschlechtsspezifische Unterschiede in Morbidität und Mortalität
- › Poster

Alle Statistikerinnen und Statistiker sind eingeladen, sich mit Vorträgen und/oder Postern zu beteiligen. Die eingereichten Beiträge werden einem Begutachtungsprozess unterzogen. Es ist auch möglich, ganze Sitzungseinheiten (bestehend aus drei bis vier Vorträgen) vorzuschlagen. Nachwuchswissenschaftlerinnen und -wissenschaftler der Deutschen Statistischen Gesellschaft werden besonders ermuntert, ganze Sitzungseinheiten einzureichen. Einreichungen sind noch bis zum 30. April 2025 möglich.

➤ statistische-woche.de

StatistikTage Bamberg|Fürth 2025: Dimensionen der Nachhaltigkeit – Möglichkeiten und Grenzen der Messung in Forschung und Statistik

Das Bayerische Landesamt für Statistik, das Institut für Arbeitsmarkt- und Berufsforschung (IAB) und die Julius-Maximilians-Universität Würzburg organisieren im Rahmen des Statistik Netzwerks Bayern vom 24. bis 25. Juli 2025 die 13. StatistikTage Bamberg|Fürth. Die Veranstaltung findet im Bayerischen Landesamt für Statistik in Fürth statt und widmet sich dem Thema „Dimensionen der Nachhaltigkeit“. Der Fokus wird dabei auf der Messung der Nachhaltigkeitsindikatoren in Statistik und Forschung liegen.

➤ www.statistiknetzwerk.bayern.de

NEUERSCHEINUNGEN

Does Healthcare Deliver? Results from the Patient-Reported Indicator Surveys (PaRIS)

Die Patient-Reported Indicator Surveys (PaRIS) sind eine Initiative der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD), bei der die Länder gemeinsam eine neue Generation von Indikatoren in der Gesundheitsversorgung entwickeln, standardisieren und umsetzen. Der Bericht „Does Healthcare Deliver?“ präsentiert die Ergebnisse der ersten internationalen Umfrage unter Patientinnen und Patienten im Alter von 45 Jahren und älter, wobei der Schwerpunkt auf der größten und am schnellsten wachsenden Gruppe liegt: Menschen mit chronischen Erkrankungen. Durch die Erfassung der von den Patientinnen und Patienten selbst angegebenen Ergebnisse und Erfahrungen bietet diese Erhebung eine unschätzbare Perspektive auf die Gesundheitssysteme.

➤ www.oecd.org

OECD Economic Outlook, Interim Report 2025

Die Weltwirtschaft blieb im Jahr 2024 widerstandsfähig und wuchs in der zweiten Jahreshälfte mit einer soliden Jahresrate von 3,2%. Die jüngsten Konjunkturindikatoren deuten jedoch auf eine Abschwächung der globalen Wachstumsaussichten hin. Die Stimmung der Unternehmen und Verbraucher hat sich in einigen Ländern verschlechtert. Der Inflationsdruck hält in vielen Volkswirtschaften weiter an. Gleichzeitig ist die politische Unsicherheit groß und es bestehen weiterhin erhebliche Risiken.

Der „OECD Economic Outlook, Interim Report 2025“ argumentiert, dass die internationale Zusammenarbeit unter den derzeitigen Umständen von entscheidender Bedeutung ist. Zur Unterstützung des Wirtschaftswachstums wäre eine weitere Senkung der Leitzinsen in den Ländern angebracht, in denen die Inflation voraussichtlich zurückgehen oder niedrig bleiben wird.

➤ www.oecd.org

Economic Implications of Data Regulation: Balancing Openness and Trust

Globale Datenströme sind für die heutigen sozialen und wirtschaftlichen Interaktionen von entscheidender Bedeutung. Es gibt jedoch Bedenken hinsichtlich der Privatsphäre, des Datenschutzes und der Sicherheit, was zu uneinheitlichen nationalen Regelungen geführt hat. Da die rechtlichen Rahmenbedingungen für Daten von Land zu Land unterschiedlich sind, kann dies zu Einschränkungen des globalen Datenflusses führen und negative wirtschaftliche Auswirkungen haben.

„Economic Implications of Data Regulation – Balancing Openness and Trust“ zeigt, wie gute Regulierung des Datenflusses und der Sicherheit aussehen kann. Der OECD-Bericht analysiert unter anderem die wirtschaftlichen Herausforderungen, die sich ergeben, wenn Datenflüsse zu unterschiedlich reguliert werden. Er identifiziert auch Regulierungsformen zur Datenspeicherung, um einen geschützten freien Datenverkehr zu gewährleisten.

➤ www.oecd.org

DIE ANSCHRIFTENSTICHPROBE DES ZENSUS 2022

Ermittlung der Einwohnerzahl in den Sampling Points

Kai Lorentz, Markus Zwick

➤ **Schlüsselwörter:** Stichprobendesign – Einwohnerzahl – Hochrechnung – GREG-Verfahren – Haushaltsstichprobe

ZUSAMMENFASSUNG

Der Beitrag beschreibt die Vorgehensweise zur Ermittlung der Einwohnerzahl der Gemeinden im Sinne des Zensusgesetzes 2022. Die verwendete Methodik ist dabei weitgehend identisch mit derjenigen des Zensus 2011. Der Artikel zeigt im Detail, dass die verwendeten Methoden auf etablierten und wissenschaftlich allgemein akzeptierten Verfahren fußen. Entsprechend erhebt dieser Artikel keinen Anspruch auf Originalität, im Gegenteil. Er legt dar, dass die Einwohnerzahlermittlung im Zensus 2022 auf einer bewährten und anerkannten Methodik beruht, welche das Bundesverfassungsgericht in seinem Urteil von 2018 bereits als grundsätzlich adäquat bezeichnet hatte.

➤ **Keywords:** *sample design – number of inhabitants – extrapolation – GREG method – household sample*

ABSTRACT

The article describes the procedure for determining the number of inhabitants of the municipalities as defined by the 2022 Census Act. The methodology used is largely identical to that of the 2011 Census. The article shows in detail that the methods used are based on established and generally accepted scientific procedures. Therefore, this article makes no claim to originality – quite the contrary. It demonstrates that the population count of the 2022 Census is based on a time-tested and recognised methodology, which the Federal Constitutional Court already found to be adequate, in principle, in its judgment of 2018.



Dr. Kai Lorentz

studierte Mathematik an der Johannes Gutenberg-Universität Mainz. Er leitet das Referat „Mathematisch-statistische Verfahren für Bevölkerungs- und Sozialstatistiken“ des Statistischen Bundesamtes. Er war insbesondere mit Fragen der mathematisch-statistischen Methodik für den Zensus 2022 befasst.



Prof. Dr. Markus Zwick

ist Volkswirt und leitet die Gruppe „Institut für Forschung und Entwicklung in der Bundesstatistik, Forschungsdatenzentrum“ des Statistischen Bundesamtes.

Danksagung

Wir danken verschiedenen Gutachterinnen und Gutachtern für ihre Anmerkungen, die den Beitrag deutlich verbessert haben.

1

Einleitung

Dieser Beitrag ist einer von mehreren, die sich mit dem Design und der Hochrechnung der Anschriftenstichprobe des Zensus 2022, der daraus folgenden Ermittlung der Bevölkerungszahl sowie mit der Qualität der ermittelten Ergebnisse auseinandersetzen. Der Beitrag von Steffen Klink und Kai Lorentz (2022) „Auswahlplan und Stichprobenhauptziehung für den Zensus 2022“ beschreibt dabei das Design des aus dem Zensus 2011 weiterentwickelten Konzepts des Stichprobenplans. Die dort sowie weiter unten dargelegte konzeptionelle Weiterentwicklung der Methodik des Zensus 2022 ist insbesondere eine Konsequenz aus dem Urteil des Bundesverfassungsgerichts zum Zensus 2011 aus dem Jahr 2018.¹

Die Konzeption der Hauptstichprobe des Zensus 2022 baut auf Präzisionszielen für die Ermittlung der Einwohnerzahl auf, diese sind in dem Beitrag von Corinna Bretsch und Kai Lorentz (2019) „Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021“ beschrieben.

Die Publikation von Corinna Bretsch, Steffen Seibel, Ingeborg Vorndran und Miriam Pfahl (2024) „Ermittlung der Einwohnerzahl im Zensus 2022“ beschreibt im Detail, wie die Bevölkerungszahl in den Gemeinden auf der Grundlage der Daten der Einwohnermeldeamtsregister und der Stichproben berechnet wurde. In diesem Beitrag wird insbesondere dargelegt, wie aus dem konsolidierten stichtagsbezogenen Bestand der Melderegister mittels der aus der Stichprobe hochgerechneten Über- und Untererfassung die amtliche Einwohnerzahl in den Sampling Points des Zensus 2022 ermittelt wurde. Dieser Beitrag geht auch auf die Dublettenbereinigung sowie auf die Erhebung in Sonderbereichen ein.

Im vorliegenden Beitrag werden die Einzelheiten der Hochrechnung sowie Qualitätsbetrachtungen auf der Grundlage der jeweiligen Standardfehler erörtert. Kapitel 2 legt dabei die Gründe dar, warum und wie die Methodik des Zensus 2022 auf der Grundlage des Urteils des Bundesverfassungsgerichts weiterentwickelt wurde. Kapitel 3 fasst allgemeine Ausführungen zur Ermittlung

der Einwohnerzahlen zusammen. Kapitel 4 gibt eine Übersicht über die Datenquellen und die verschiedenen Arten von Sampling Points. Die Kapitel 5 bis 7 erläutern die geschichtete Zufallsstichprobe des Zensus 2022, die Problematik der Über- und Untererfassung der Bevölkerung im Melderegisterbestand sowie das Verfahren zur Schätzung der Einwohnerzahlen. Kapitel 8 setzt sich mit den Standardfehlern in den Sampling Points als Präzisionsziel auseinander und gibt erste Hinweise darauf, warum die Standardfehler zum Teil größer ausgefallen sind als erwartet. Ein Fazit schließt den Beitrag ab.

2

Konsequenzen aus dem Urteil des Bundesverfassungsgerichtes zum Zensus 2011

Das Bundesverfassungsgericht hat den Zensus 2011 mit der zugrunde liegenden Methodik geprüft und in seinem Urteil von 2018 als verfassungsgemäß bestätigt.² Eine Weiterentwicklungsnotwendigkeit für den nächsten durchzuführenden Zensus sah das Gericht in der Behandlung der sogenannten kleinen Gemeinden, das heißt der Gemeinden mit weniger als 10 000 gemeldeten Einwohnerinnen und Einwohnern. Für diese Gemeinden gab es beim Zensus 2011 eine andere Herangehensweise hinsichtlich der Einwohnerzahlermittlung im Vergleich zu den Gemeinden mit mehr als 10 000 Einwohnerinnen und Einwohnern. Dort erfolgte bei kleinen Gemeinden eine statistische Korrektur der Melderegister mithilfe der „Befragung zur Klärung von Unstimmigkeiten“ (kurz BKU; siehe Statistische Ämter des Bundes und der Länder, 2015, Seiten 44 bis 46).

Für den Zensus 2022 wurde dieser Punkt seitens des Gesetzgebers aufgegriffen und somit der Empfehlung des Bundesverfassungsgerichts gefolgt. Das Zensusgesetz sieht in seiner gültigen Fassung vom 10. Dezember 2020 vor, dass die auf der Haushaltsstichprobe beruhenden Korrekturen im Zensus 2022 auf die sogenannten kleinen Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern ausgeweitet wer-

1 Bundesverfassungsgericht, Urteil vom 19. September 2018 (2 BvF 1/15; 2 BvF 2/15): www.bundesverfassungsgericht.de

2 www.bundesverfassungsgericht.de

den.¹³ Die Methodik der Ermittlung der amtlichen Einwohnerzahl besteht in einer statistischen Korrektur der stichtagsbezogenen konsolidierten und um unzulässige Dubletten bereinigten Melderegisterdaten mithilfe einer Anschriftenstichprobe, dies für alle Gemeinden beziehungsweise Sampling Points.¹⁴ Zum detaillierten Korrekturverfahren siehe Bretsch und andere (2024, hier: Seite 19 ff.), zur Abgrenzung von Gemeinden und Sampling Points siehe Klink/Lorentz (2022, hier: Seite 16).

Für die Ausgestaltung dieser Anschriftenstichprobe waren verschiedene stichprobenmethodische Aspekte zu berücksichtigen, in erster Linie die Qualitätsvorgaben in den Planungsgrößen zur Ermittlung der Stichprobenumfänge vor dem Hintergrund einer sehr heterogenen Gemeindestruktur in Deutschland. Auch war dem Aspekt der Begrenzung der Befragungsbelastung im Sinne einer möglichst effizienten Informationsermittlung Rechnung zu tragen. Die reine Ausweitung der Methodik zur statistischen Korrektur der Melderegister auf alle Gemeinden hätte bei ansonsten unverändertem Stichprobendesign zu einem deutlich höheren Stichprobenumfang geführt. Hier griff der Auftrag des Bundesverfassungsgerichts, stets Möglichkeiten einer grundrechtsschonenden Datenerhebung zu prüfen und anzuwenden.

Unter Einbeziehung wissenschaftlicher Expertise (zum Beispiel Burgard und andere, 2020) wurde ein Stichprobendesign entwickelt, durch das insbesondere die regionalen Ziehungsgesamtheiten (die sogenannten Sampling Points) den genannten methodischen Herausforderungen adäquat Rechnung getragen haben. Gleichzeitig wurden hierbei Qualitätsanforderungen auf Ebene dieser Sampling Points formuliert, welche bei der Stichprobenplanung zu berücksichtigen waren.

Dieses methodische Grundgerüst des Designs der Anschriftenstichprobe ist schließlich in das Zensusgesetz 2022 eingeflossen. In § 11 Zensusgesetz 2022 werden der Begriff der Gemeinde im Zensuskontext konkret definiert sowie die Qualitätsansprüche (als Einflussgröße auf die Stichprobenplanung), welche auf dieser regionalen Ebene anzustreben sind.

Gemäß diesen Vorgaben des Zensusgesetzes 2022 und den oben zitierten wissenschaftlichen Empfehlungen wurde die Anschriftenstichprobe im Statistischen Bundesamt gezogen.¹⁵

3

Allgemeine Erläuterungen zur Ermittlung der amtlichen Einwohnerzahlen im Zensus 2022

Der Zensus 2022 wurde wie der Zensus 2011 registergestützt durchgeführt, das heißt als Datenquellen dienen Register sowie Daten aus einer Stichprobenerhebung. Bei den zur Ermittlung der amtlichen Einwohnerzahlen genutzten Registerdaten handelt es sich um Daten aus den amtlichen Melderegistern. Die Stichprobendaten zur Einwohnerzahlermittlung gehen aus der Anschriftenstichprobe (Haupt- und Nachziehung) des Zensus hervor.¹⁶ Daneben gibt es die Vollerhebung an Anschriften mit Sonderbereichen (Wohnheime und Gemeinschaftsunterkünfte) sowie die Registerauszählung derjenigen Anschriften, welche nach der Nachziehung und vor dem Zensusstichtag in die Melderegister aufgenommen wurden.

➤ Grafik 1 zeigt die für die Einwohnerzahlermittlung maßgeblichen Bestandteile der Hochrechnung und der Auszählungen. Dieser Artikel beschreibt die Schätzung des Einwohnerzahlbeitrages gemäß dieser Grafik, das heißt ohne die übrigen dargestellten additiven Auszählungsteile. Die Schätzung weiterer Erhebungsmerkmale des Zensus 2022, zum Beispiel der Merkmale zur Bildung und Erwerbstätigkeit (sogenannte Ziel2-Merkmale), werden in Klink/Lorentz (2022) im Detail beschrieben.

3 § 11 Zensusgesetz 2022

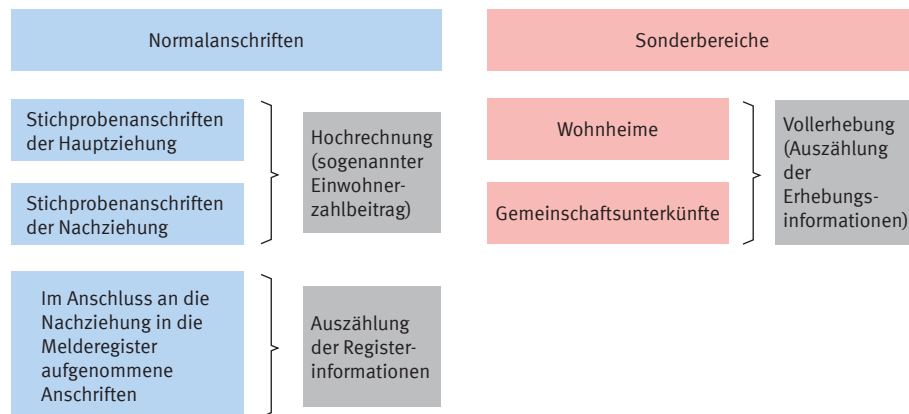
4 Die Auswahlinheit „Anschrift“ wurde genau wie 2011 gewählt, um gemäß des Erhebungskonzeptes Fehlbestände der Melderegister identifizieren zu können (siehe unten im Text).

5 Details zur Stichprobenziehung, insbesondere zu weiteren Designaspekten wie Schichtung und Intervallvorgaben zu Auswahlätzen, finden sich in Burgard und andere (2020) sowie Klink/Lorentz (2022).

6 Zum Stichprobendesign und der Stichprobenauswahl siehe Klink/Lorentz (2022) sowie Kapitel 4 dieses Beitrags.

Grafik 1

Für die Einwohnerzahlermittlung im Zensus 2022 maßgebliche Bestandteile der Hochrechnungen und der Auszählungen



4

Datenquellen der Einwohnerzahlermittlung und Übersicht über die Sampling Points

Die Registerdaten der Einwohnermeldeämter sowie die Stichprobendaten, die zur Einwohnerzahlermittlung¹⁷ zur Verfügung standen, bildeten einen unterschiedlichen Anteil der Bevölkerung ab. Während der konsolidierte Melderegisterbestand (Bretschki und andere, 2024) prinzipiell die Grundgesamtheit der Bevölkerung enthielt, beschränkten sich die Stichprobendaten auf rund 10 % der Anschriften beziehungsweise Personen. In den Melderegisterdaten sind unter anderem die Merkmale Wohngemeinde, Geschlecht, Geburtsdatum und Staatsangehörigkeit enthalten, die auch in der Haushaltsstichprobe erhoben werden. Darüber hinaus ist in

¹⁷ Zum Einwohner- beziehungsweise Bevölkerungsbegriff im Zensus 2022 siehe Bretschki und andere (2024), hier: Seite 18 f.

den Stichprobendaten das Merkmal Wohnungsstatus verfügbar, das angibt, über welche Art von Wohnsitz die angetroffene Person an der Anschrift wohnhaft ist (alleiniger beziehungsweise Haupt- oder Nebenwohnsitz).

➤ **Übersicht 1** gibt einen Überblick über die Merkmale der beiden Datenquellen.

Die Melderegisterdaten stellen Daten zu den angegebenen Merkmalen für alle Personen bereit, die ihren alleinigen Wohnsitz oder ihren Hauptwohnsitz in einem Meldeamt haben registrieren lassen.

Der Gesetzgeber hat in § 11 Absatz 1 Zensusgesetz 2022 den Begriff „Gemeinde“ im Zensuskontext definiert. Wir verwenden hier im Folgenden den Begriff Sampling Point. Dieser sieht konkret folgendermaßen aus:

- › Jeder Stadtteil im Sinne des Zensusgesetzes einer Gemeinde mit mindestens 400 000 gemeldeten Einwohnerinnen und Einwohnern bildet einen eigenen Sampling Point (hierbei haben die betroffenen Städte beziehungsweise die jeweils zuständigen Statistischen Ämter der Länder diese Einteilung eigenstän-

Übersicht 1

Personengruppen und Merkmale in den Datenquellen

Datenquelle	Beschreibung	Enthaltene Merkmale
Melderegisterdaten	Gesamtbevölkerung	Wohngemeinde, Wohnstadtteil, Geschlecht, Geburtsdatum, Staatsangehörigkeit
Haushaltsstichprobe	Repräsentative Stichprobe mit 9,7 % der Anschriften	Wohngemeinde, Wohnstadtteil, Geschlecht, Geburtsdatum, Staatsangehörigkeit, Wohnstatus

Übersicht 2

Arten von Sampling Points

Sampling Point	Verwaltungseinheiten des Sampling Points
Stadtteil	1 definierter Stadtteil einer Gemeinde mit mindestens 400 000 gemeldeten Einwohnerinnen und Einwohnern
Gemeinde	1 Gemeinde mit weniger als 400 000 gemeldeten Einwohnerinnen und Einwohnern
Gemeindeverband	1 Gemeindeverband
Gemeindeverbands-Rest	Vereinigung der Gemeinden eines Gemeindeverbands mit jeweils weniger als 2 000 gemeldeten Einwohnerinnen und Einwohnern

dig vorgenommen, Richtgröße waren 200 000 Einwohnerinnen und Einwohner).

- › Jede verbandsfreie Gemeinde mit weniger als 400 000 gemeldeten Einwohnerinnen und Einwohnern definiert einen eigenen Sampling Point.
- › Die Bundesländer Baden-Württemberg, Bayern, Brandenburg, Sachsen und Sachsen-Anhalt besitzen Gemeindeverbände, hier stellt jede verbandsangehörige Gemeinde einen eigenen Sampling Point dar.
- › In den Bundesländern Niedersachsen, Rheinland-Pfalz und Thüringen, die ebenfalls über Gemeindeverbände verfügen, definiert jeder Gemeindeverband einen Sampling Point.
- › Für die Bundesländer Schleswig-Holstein und Mecklenburg-Vorpommern, in denen auch Gemeindeverbände vorkommen, wird jede verbandsangehörige Gemeinde mit mindestens 2 000 gemeldeten Einwohnerinnen und Einwohnern als eigenständiger Sampling Point behandelt. Die verbandsangehörigen Gemeinden mit weniger als 2 000 gemeldeten Einwohnerinnen und Einwohnern, die zu demselben Gemeindeverband gehören, werden zu einem Sampling Point zusammengefasst. Ein solcher Sampling Point, der also aus mehreren verbandsangehörigen Gemeinden besteht, wird als Gemeindeverbands-Rest bezeichnet.

Mit diesen fünf verschiedenen Arten von Sampling Points wurden so aus den 10 786 Gemeinden in Deutschland (Stand: 31. Dezember 2021) 6 499 Sampling Points gebildet. [↗ Übersicht 2](#)

5

Geschichtete Zufallsstichprobe

Die Stichprobe des Zensus 2022 wurde wie im [Zensus 2011](#) als geschichtete Zufallsstichprobe auf Anscritenebene konzipiert. Die Zerlegung einer Grundgesamtheit in homogenere Schichten vor dem Ziehen und das Ziehen von dann im Grunde unabhängigen Stichproben erhöht die Präzision von hochgerechneten Ergebnissen einer Stichprobe.¹⁸ Eine größere Homogenität in den Schichten der Auswahlgrundlage hat grundsätzlich zur Folge, dass die zu erwartenden Stichprobenzufallsfehler bei der Hochrechnung der Einwohnerzahlen bei festgehaltenem Stichprobenumfang reduziert werden. Umgekehrt bedeutet dies, dass bei festgehaltenem Präzisionsziel der Stichprobenumfang reduziert werden kann.

Im [Zensus 2022](#) stellte die Schichtung eine hierarchische Kombination aus der regionalen Schichtung, hier die Sampling Points als erste Schichtungsvariable, und einer Schichtung in den Sampling Points nach Anscritengrößenklassen, gemäß der Anzahl an Personen, die an einer Anscrit gemeldet waren, dar.

Im Gegensatz zum Zensus 2011, hier wurden je Sampling Point konstant acht Anscritengrößenklassen mit jeweils ungefähr gleich großer Anzahl an Personen (ex ante) gebildet¹⁹, wurden im Zensus 2022 abhängig von den gemeldeten Personen des jeweiligen Sampling Points nach dem gleichen Prinzip bis zu 16 Anscritengrößenklassen konzipiert.

⁸ Siehe hierzu Krug und andere (2001), Seite 100 ff.

⁹ Siehe Berg/Bihler (2011), hier: Seite 322.

Hierzu wurde zunächst für jeden Sampling Point eine Schicht der unbemeldeten Anschriften (sogenannte Nullanschriften) gebildet. Anschließend wurde gemäß [Tabelle 1](#) die Anzahl der Anschriftengrößenklassen je Sampling Point bestimmt. Innerhalb des jeweiligen Sampling Points wurden hierzu die Anschriften der Auswahlgrundlage gemäß ihrer Anzahl an gemeldeten Bewohnerinnen und Bewohnern aufsteigend sortiert und die Schichten derart gebildet, dass in etwa die gleichen Anzahlen an gemeldeten Personen sich in den Schichten befanden.

Tabelle 1**Schichtung der Zufallsstichprobe**

Anzahl gemeldeter Bewohnerinnen und Bewohner von ... bis ...		Anzahl der Anschriften-größenklassen
0	0	,00'
1	49	,01'
50	99	,02'
100	149	,03'
150	199	,04'
200	299	,05'
300	399	,06'
400	599	,07'
600	799	,08'
800	999	,09'
1 000	1 499	,10'
1 500	1 999	,11'
2 000	2 499	,12'
2 500	4 999	,13'
5 000	9 999	,14'
10 000	24 999	,15'
25 000		,16'

Durch die Erhöhung der Anzahl der Anschriftengrößenklassen beim Stichprobendesign des Zensus 2022 gegenüber dem Stichprobendesign des Zensus 2011 wurde eine stärkere Homogenisierung der einzelnen Schichten der Auswahlgrundlage erreicht. Mit anderen Worten, es konnten innerhalb der größeren Anzahl an Schichten Blöcke von Anschriften und korrespondierten gemeldeten Personen gebildet werden, die weniger stark voneinander abwichen, sodass die Varianz der Zielvariablen innerhalb der Schichten geringer war. Da der Standardfehler der hochgerechneten Ergebnisse von der Varianz und dem Auswahlatz abhängt (siehe auch Formel 1) bedeutete dies, dass die im Zensusgesetz angestrebte jeweilige Präzision in den Schichten mit einem geringeren Auswahlatz realisiert werden konnte

(ex ante). Details zu dieser Vorgehensweise finden sich in Klink/Lorentz (2022).

- (1) Standardfehler für die frei hochgerechnete

$$\text{Merkmalssumme} \sim \frac{S_{hY}^2}{n_h} |_{10}$$

für die einfache Zufallsauswahl ohne Zurücklegen, dabei bezeichnet S_{hY}^2 die Varianz des Merkmals Y und n_h den Stichprobenumfang (einer Schicht h)

Diese Änderung war eine notwendige Reaktion auf das Urteil des Bundesverfassungsgerichts. Da im Zensus 2022 auch bei Gemeinden kleiner 10 000 Einwohner über die Stichprobe die Melderegisterdaten korrigiert werden mussten, war die Stichprobe anzupassen. Hier hatte sich ja bekanntlich das 2011 eingesetzte Korrekturverfahren für Gemeinden kleiner 10 000 Einwohner bei der Evaluierung als nicht vollständig ausreichend erwiesen. Eine konzeptionelle Übernahme des Schichtungsschemas des Zensus 2011 hätte eine deutliche Ausweitung des Stichprobenumfangs zur Erreichung der gesetzlichen Präzisionsziele bedeutet, dies wäre dann aber mit einer Verletzung des Gebots des Bundesverfassungsgerichts zur grundrechtsschonenden Datenerhebung einhergegangen, die zusätzlichen Kosten und Aufwände hier noch gar nicht benannt.

Das Konzept zur Bildung der Anschriftengrößenklassen im Zensus 2022 war ein einvernehmliches Ergebnis der Beratungen der Statistischen Ämter des Bundes und der Länder auf der Grundlage einer wissenschaftlichen Begleitstudie der Universität Trier¹¹.

6

Über- und Untererfassung der Bevölkerung im Melderegisterbestand

Um die Konzeption der Einwohnerzahlermittlung für die Sampling Points anschaulich darzulegen, erfolgt hier zunächst eine Zusammenfassung der Zielgrößen.

Bei der Einwohnerzahlermittlung war für jeden Sampling Point die Anzahl der existenten Personen (Einwohnerin-

10 Statistisches Bundesamt (1960), Kapitel 3.4 Beurteilung der Genauigkeit, Seite 97 ff., hier: Formel 46 auf Seite 103.

11 Siehe Burgard und andere (2020), hier: Seite 24.

nen und Einwohner), die Anzahl der Übererfassungen (Karteileichen) und die Anzahl der Untererfassungen (Fehlbestände) aus der Anschriftenstichprobe zu ermitteln. Die hierfür genutzten Datenquellen waren

- › der konsolidierte stichtagsbezogene Melderegisterbestand nach Mehrfachfallprüfung (Wittmaack, 2025) mit den Merkmalen Wohnort (Sampling Point), Geschlecht männlich/weiblich, Alter in Jahren, Staatsangehörigkeit deutsch/nicht deutsch sowie
- › die Haushaltsstichprobe mit den Merkmalen Wohnstatus, Wohnort (Sampling Point), Geschlecht männlich/weiblich, Alter in Jahren, Staatsangehörigkeit deutsch/nicht deutsch.

Hierbei wurde das Merkmal Wohnort auf Sampling-Point-Ebene aus den beiden Merkmalen Wohngemeinde und Wohnstadtteil abgeleitet. Außerdem wurde das Alter zum Zensusstichtag in Jahren mithilfe des Geburtsdatums berechnet. Das Merkmal Staatsangehörigkeit deutsch/nicht deutsch wurde mittels des Merkmals Staatsangehörigkeit auf Staatenebene generiert.

Den Melderegisterdaten wurden für jeden Sampling Point die Anzahl der gemeldeten Personen entnommen. „Gemeldet“ bedeutet im Folgenden stets gemeldet mit alleinigem beziehungsweise Hauptwohnsitz.

Allgemein gilt die Beziehung

- (2) Anzahl an gemeldeten Personen – Anzahl an Karteileichen
= Anzahl an existenten¹² Personen – Anzahl an Fehlbeständen

Die Größe der Fallzahlen des Merkmals „Anzahl der existenten Personen“ und des durch die beiden obenstehenden Differenzen beschriebenen Merkmals sind in der Regel deutlich größer als die Fallzahlen der Merkmale „Anzahl der Karteileichen“ und „Anzahl der Fehlbestände“. Denn zumeist ist die Anzahl an gemeldeten beziehungsweise existenten Personen deutlich größer als die Anzahl an Karteileichen beziehungsweise Fehlbeständen. Der Stichprobenfehler für Zielgrößen mit größeren Fallzahlen von Anschriften mit Werten verschieden von „0“ fällt geringer aus als für Zielgrößen mit kleineren derartigen Fallzahlen. Daher wurden als Zielgrößen nicht direkt die Anzahl an Karteileichen und Fehlbeständen

geschätzt, sondern es wurde der Weg über die Anzahl der existenten Personen und die Anzahl der Personen gegangen, die die obenstehende Gleichung beschreibt. Bei den durch die obige Gleichung beschriebenen Personen handelt es sich um Personen, die gemeldet und existent sind (linke Seite der Gleichung), was äquivalent dazu ist, dass sie existent und im Melderegister erfasst sind (rechte Seite der Gleichung). Bei ihnen handelt es sich also um die Personen, die „gemeldet und wohnhaft“ sind.

Um diese Personen zu beschreiben, ist der Begriff „paarige Person“ geläufig. Eine Person heißt paarig, wenn sie gemeldet und zugleich existent ist.

Für jeden Sampling Point gelten die Beziehungen

- (3) Anzahl der paarigen Personen
= Anzahl der gemeldeten Personen – Anzahl der Karteileichen
= Anzahl der existenten Personen – Anzahl der Fehlbestände,

- (4) Anzahl der Karteileichen
= Anzahl der gemeldeten Personen – Anzahl der paarigen Personen

und

- (5) Anzahl der Fehlbestände
= Anzahl der existenten Personen – Anzahl der paarigen Personen.

Bei der Einwohnerzahlermittlung wurden also für jeden Sampling Point die Anzahl der existenten Personen und die Anzahl der paarigen Personen aus der Hochrechnung ermittelt. Die Anzahl der Karteileichen und die Anzahl der Fehlbestände ergaben sich dann für jeden Sampling Point durch Differenzbildung gemäß den obenstehenden Formeln.

Neben den Melderegisterdaten flossen in die Einwohnerzahlermittlung auch die Informationen aus der Haushaltsstichprobe zu den Merkmalen Wohnort, Geschlecht männlich/weiblich, Alter in Jahren und Staatsangehörigkeit deutsch/nicht deutsch ein.

Die Haushaltsstichprobe ist eine Stichprobe auf Anschriftenebene. Die Klumpung der Personen auf Anschriftenebene ist aufgrund des Erhebungskonzeptes unverzichtbar, da nur auf diese Art die Fehlbestände aufgedeckt werden können.

¹² Existent bedeutet hierbei, dass die betreffende Person mit alleinigem oder Hauptwohnsitz an der Stichprobenanschrift festgestellt wurde.

Notation

Es seien D die Menge der Sampling Points und für jeden Sampling Point $d \in D$ sei U_d die Auswahlgesamtheit für die Haushaltsstichprobe des Sampling Points, also die Menge aller Anschriften, die mit positiver Wahrscheinlichkeit in die Haushaltsstichprobe des Sampling Points gezogen werden konnten, und s_d die Menge der Anschriften, die tatsächlich in die Haushaltsstichprobe des Sampling Points gezogen wurden. Für jeden Sampling Point $d \in D$ und jede Anschrift $k \in U_d$ führen wir die folgenden Bezeichnungen ein:

d_k : Designgewicht (invertierte Inklusionswahrscheinlichkeit) der Anschrift k ,

g_k : Korrekturgewicht der Anschrift k ,

w_k : GREG-Gewicht der Anschrift k ,

\mathbb{x}_k : Hilfsvektor der Anschrift k ,

y_k : Anzahl der existenten Personen an der Anschrift k und

z_k : Anzahl der paarigen Personen an der Anschrift k .

Hierbei sind die Komponenten des Hilfsvektors zur Anschrift k wie folgt definiert:

$\mathbb{x}_{k,1}$: Anzahl der gemeldeten Personen an der Anschrift k insgesamt,

$\mathbb{x}_{k,2}$: Anzahl der gemeldeten männlichen Personen an der Anschrift k ,

$\mathbb{x}_{k,3}, \dots, \mathbb{x}_{k,10}$: Anzahl der gemeldeten Personen an der Anschrift k der Altersklasse 1, ..., 8,

$\mathbb{x}_{k,11}$: Anzahl der gemeldeten deutschen Personen an der Anschrift k .

7

Die GREG-Schätzung der Einwohnerzahlen

Das Schätzverfahren ist – wie im Zensus 2011 – die verallgemeinerte Regressionsschätzung, die kurz GREG-Schätzung (GREG für Generalized REGression) genannt wird.

Im Gegensatz zur freien Hochrechnung werden bei der GREG-Schätzung als Hochrechnungsfaktoren nicht die Kehrwerte der Wahrscheinlichkeit der jeweiligen Anschrift, in die Stichprobe zu gelangen, also die invertierten Inklusionswahrscheinlichkeiten genutzt. Bei der GREG-Schätzung werden diese sogenannten Designgewichte durch die Verwendung von Hilfsinformationen, die hier aus den Registerdaten stammen, angepasst.

Wie im Zensus 2011 wurden für die GREG-Schätzung die folgenden Hilfsvariablen auf Anschriftenebene genutzt (Statistische Ämter des Bundes und der Länder, 2015):

- › Anzahl der gemeldeten Personen insgesamt
- › Anzahl der gemeldeten männlichen Personen
- › Anzahl der gemeldeten Personen im Alter von 0 bis 5 Jahren (Altersklasse 1)
- › Anzahl der gemeldeten Personen im Alter von 6 bis 17 Jahren (Altersklasse 2)
- › Anzahl der gemeldeten Personen im Alter von 18 bis 24 Jahren (Altersklasse 3)
- › Anzahl der gemeldeten Personen im Alter von 25 bis 29 Jahren (Altersklasse 4)
- › Anzahl der gemeldeten Personen im Alter von 30 bis 39 Jahren (Altersklasse 5)
- › Anzahl der gemeldeten Personen im Alter von 40 bis 49 Jahren (Altersklasse 6)
- › Anzahl der gemeldeten Personen im Alter von 50 bis 59 Jahren (Altersklasse 7)
- › Anzahl der gemeldeten Personen im Alter von 60 bis 64 Jahren (Altersklasse 8)
- › Anzahl der gemeldeten deutschen Personen.

Die Anzahl der gemeldeten weiblichen Personen, die Anzahl der gemeldeten Personen im Alter von mindestens 65 Jahren (Altersklasse 9) und die Anzahl der gemeldeten nicht deutschen Personen sind dabei implizit enthalten, weil sie sich durch Differenzbildung aus den vorhandenen Größen ergeben.

Die Formel der GREG-Schätzung für die Bevölkerung (\hat{t}) ist gemäß Särndal und andere (1992) gegeben durch:

$$(6) \quad \hat{t}_{y,d,GREG} = \sum_{k \in S_d} d_k y_k + \sum_{j=1}^J \hat{\beta}_j \left(\sum_{k \in U_d} \mathbb{x}_{kj} - \sum_{k \in S_d} d_k \mathbb{x}_{kj} \right),$$

wobei hier $J = 11$ und $\hat{\beta}_j$ die j -te Komponente des Vektors der geschätzten Regressionskoeffizienten $\hat{\beta}$ bezeichnet:

$$(7) \quad \hat{\beta} = \left(\sum_{k \in S_d} d_k \mathbb{x}_k \mathbb{x}_k^T \right)^{-1} \sum_{k \in S_d} d_k \mathbb{x}_k y_k$$

Mittels Anwendung der GREG-Schätzung auf die Haushaltsstichprobe wurden anschriftenbezogene Hochrechnungsfaktoren ermittelt, die im Weiteren verwendet wurden, um die Anzahl der existenten und paarigen Personen im jeweiligen Sampling Point zu schätzen. Die Anzahl der existenten Personen eines Sampling Points ergab sich dann, indem man für jede Anschrift des Sampling Points das Produkt aus dem zur Anschrift gehörigen Hochrechnungsfaktor und der Anzahl an existenten Personen an dieser Anschrift gemäß Haushaltsstichprobe ermittelte und diese Produkte über alle Anschriften des Sampling Points aufsummierte.

Der Schätzer in Gleichung (6) kann in der folgenden Form mit Hochrechnungsgewicht dargestellt werden, was rechentechnische Vorteile besitzt:

$$(8) \quad \hat{t}_{y,d,GREG} = \sum_{k \in S_d} d_k g_k y_k = \sum_{k \in S_d} w_k y_k,$$

wobei

$$(9) \quad g_k := 1 + (\sum_{i \in U_d} \mathbb{x}_i - \sum_{i \in S_d} d_i \mathbb{x}_i)^T (\sum_{i \in S_d} d_i \mathbb{x}_i \mathbb{x}_i^T)^{-1} \mathbb{x}_k$$

und das GREG-Gewicht der Anschrift k durch

$$(10) \quad w_k := d_k g_k$$

gegeben ist.

Ziel der GREG-Schätzung ist es nun, die Hochrechnungsfaktoren so zu ermitteln, dass die Schätzungen für die Anzahl an existenten und paarigen Personen plausibel sind – in dem Sinne, dass sich die Totalwerte für

bestimmte Merkmalsausprägungen, wie Anzahl an gemeldeten Personen insgesamt und Anzahl an gemeldeten männlichen Personen gemäß den Registerdaten, mithilfe der GREG-Hochrechnungsfaktoren aus der Haushaltsstichprobe berechnen lassen.

Für einen beliebigen, aber festen Sampling Point $d \in D$ ergibt sich für die Anzahl der gemeldeten Personen

$$\begin{aligned} (11) \quad & (\sum_{k \in S_d} w_k \mathbb{x}_{k,1}, \dots, \sum_{k \in S_d} w_k \mathbb{x}_{k,11}) = \sum_{k \in S_d} w_k \mathbb{x}_k^T \\ & = \sum_{k \in S_d} d_k (1 + (\sum_{i \in U_d} \mathbb{x}_i - \sum_{i \in S_d} d_i \mathbb{x}_i)^T \\ & \quad (\sum_{i \in S_d} d_i \mathbb{x}_i \mathbb{x}_i^T)^{-1} \mathbb{x}_k)^T \mathbb{x}_k^T \\ & = \sum_{i \in U_d} \mathbb{x}_i^T + \sum_{k \in S_d} d_k \mathbb{x}_k^T - \sum_{k \in S_d} d_k \mathbb{x}_k^T (\\ & \quad (\sum_{i \in S_d} d_i \mathbb{x}_i \mathbb{x}_i^T) (\sum_{i \in S_d} d_i \mathbb{x}_i \mathbb{x}_i^T)^{-1} \\ & \quad = \sum_{i \in U_d} \mathbb{x}_i^T = (\mathbb{x}_{d,1}, \dots, \mathbb{x}_{d,11}), \end{aligned}$$

wobei $\mathbb{x}_{d,1}$ die Anzahl der gemeldeten Personen im Sampling Point d insgesamt, $\mathbb{x}_{d,2}$ die Anzahl der gemeldeten männlichen Personen im Sampling Point d , $\mathbb{x}_{d,3}$ die Anzahl der gemeldeten Personen im Sampling Point d der Altersklasse 1, ..., $\mathbb{x}_{d,10}$ die Anzahl der gemeldeten Personen im Sampling Point d der Altersklasse 8 und $\mathbb{x}_{d,11}$ die Anzahl der gemeldeten deutschen Personen im Sampling Point d bezeichnet.

Die Schätzer für die Anzahl der existenten und paarigen Personen für die einzelnen Sampling Points erhält man nun, indem man für jede Stichprobenanschrift das zugehörige GREG-Gewicht mit der Anzahl an existenten beziehungsweise paarigen Personen an dieser Anschrift multipliziert und die Ergebnisse an den einzelnen Stichprobenanschriften des jeweiligen Sampling Points aufsummiert. Damit ergibt sich also für jeden Sampling Point d der Schätzer für die Anzahl an existenten Personen \hat{Y}_d :

$$(12) \quad \hat{Y}_d = \sum_{k \in S_d} w_k y_k,$$

und für den Schätzer der Anzahl an paarigen Personen \hat{Z}_d im Sampling Point d :

$$(13) \quad \hat{Z}_d = \sum_{k \in S_d} w_k z_k.$$

Die geschätzte Anzahl der existenten und paarigen Personen auf Sampling-Point-Ebene, die auf diese Weise ermittelt werden, ist im Allgemeinen nicht ganzzahlig.

Das abschließend angewandte Rundungsverfahren liefert schlussendlich ganzzahlige Werte.

Die angewandte Varianzformel der GREG-Schätzung lautet in den obigen Bezeichnungen:

$$(14) \quad \hat{V}(\hat{t}_{y,d,GREG}) = \sum_{h=1}^H \frac{N_h^2}{n_h} \left(1 - \frac{n_h}{N_h}\right) \frac{1}{n_h - 1} \sum_{k \in s_{d,h}} (v_k - \bar{v}_{d,h})^2$$

(sogenannter Residualvarianzschätzer).

Dabei ist

$$(15) \quad v_k := g_k \cdot \varepsilon_k \quad \text{das mit dem Korrekturgewicht } g_k \text{ gewichtete Residuum der Regression,}$$

$$(16) \quad \varepsilon_k := E_k - \mathbb{I}_k^T (\sum_{i \in s_d} d_i \mathbb{I}_i \mathbb{I}_i^T)^{-1} \sum_{i \in s_d} d_i \mathbb{I}_i E_i \quad \text{das Residuum und}$$

$$(17) \quad \bar{v}_{d,h} := \frac{1}{n_h} \sum_{k \in s_{d,h}} v_k \quad \text{das arithmetische Mittel von } v \text{ in der Schicht } h \text{ des Sampling Points } d.$$

Schließlich wird der relative Standardfehler der Einwohnerzahl geschätzt als

$$(18) \quad rse(\hat{t}_{y,d,GREG}) := \frac{(\hat{V}(\hat{t}_{y,d,GREG}))^{\frac{1}{2}}}{\hat{t}_{y,d,GREG}}.$$

8

Präzisionsziele

In Stichprobenerhebungen ist der Standardfehler ein wichtiges Maß für die Präzision und Qualität der hochgerechneten Ergebnisse. Aus diesem Grund hat der Gesetzgeber im Zensusgesetz 2022 unter § 11 „Zwecke und Umfang der Haushaltebefragung auf Stichprobenbasis“ detailliert Präzisionsziele als anzustrebende Planungsgröße vorgegeben. Hiernach soll der einfache relative Standardfehler höchstens 0,5 % betragen, dies gilt für Gemeinden mit mindestens 10 000 Einwohnerinnen und Einwohnern. Für kleinere Gemeinden sieht § 11 Absatz 2 Zensusgesetz 2022 eine Präzisionszielfunktion mit einem gleitenden Übergang und abgestufter Präzision vor. Um den Stichprobenumfang und dadurch verbundene Mehrkosten für die Erhebung zu optimieren, ist dies notwendig, da ansonsten bei heterogener Einwohnerstruktur der Gemeinden diese zur Realisierung des 0,5-%-Präzisionsziels mit sehr großem Auswahlatz oder auch vollständig in der Stichprobe übernommen werden müssen.

Wie schon mit Formel (1) dargelegt, hängt der Standardfehler einer Stichprobe vor allem vom Auswahlatz und der Varianz beziehungsweise Standardabweichung ab. Dies bedeutet, dass bei heterogener Struktur der Einwohner, sprich hoher Varianz, ein höherer Auswahlatz

Tabelle 2

Ermittlung der Einwohnerzahl im Zensus 2022: Gemeinden/Stadtteile mit Standardfehler in den angegebenen Grenzen nach Gemeindegrößenklassen

	Präzisionsziel eingehalten		Standardfehler zwischen dem 1- und einschließlich dem 1,2-Fachen des Präzisionsziels		Standardfehler zwischen dem 1,2- und einschließlich dem 2-Fachen des Präzisionsziels		Standardfehler mit mehr als dem Doppelten des Präzisionsziels		Insgesamt	
	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%	Anzahl	%
Gemeinden mit ... weniger als 10 000 Einwohnerinnen und Einwohnern	1 325	28,5	518	11,1	1 832	39,4	980	21,1	4 655	100
10 000 bis 50 000 Einwohnerinnen und Einwohnern	95	5,9	135	8,4	948	59,0	428	26,7	1 606	100
50 000 bis 100 000 Einwohnerinnen und Einwohnern	31	27,7	31	27,7	47	42,0	3	2,7	112	100
mehr als 100 000 Einwohnerinnen und Einwohnern	96	76,2	19	15,1	11	8,7	0	0,0	126	100
Insgesamt	1 547	23,8	703	10,8	2 838	43,7	1 411	21,7	6 499	100

notwendig ist, um die jeweiligen angestrebten Präzisionsziele zu erreichen. Heterogene Struktur der Einwohner bedeutet hierbei, dass die gemeldeten Personen je Adresse stark variieren. Bei sehr gleicher Struktur der gemeldeten Personen je Adresse in den Schichten, zum Beispiel in einer Gemeinde mit überwiegend Einfamilienhäusern, ist hingegen ein kleinerer Auswahlatz ausreichend, um die gesetzlich benannten Präzisionsziele zu erreichen. Auch die tatsächliche Korrelation zwischen den existenten und den gemeldeten Personen hat einen Einfluss auf die Standardfehler der GREG-Schätzung der Einwohnerzahlen, wie sich aus der Formel (14) ableiten lässt. Je höher dieser statistische Zusammenhang ist, desto niedriger die Standardfehler. Diese Korrelationen stehen in direktem Zusammenhang mit der Qualität der Melderegister.

Diese Gesichtspunkte werden bei der Stichprobenplanung berücksichtigt. Der Gesetzgeber hat für den Zensus 2022 keinen maximalen Stichprobenumfang vorgegeben, sodass zu den angestrebten Präzisionszielen der jeweils erforderliche Stichprobenumfang in den jeweiligen Schichten im Vorfeld, also ex ante, berechnet werden konnte. Grundlage hierzu waren Annahmen zur Varianz, sprich zur Verteilung der Personen und Adressen, die in der notwendigen Tiefengliederung noch bis auf den Zensus 2011 zurückgingen. In der Stichprobenplanung realisierten sich so Auswahlätze in allen 91 812 Schichten der Hauptziehung, die die gesetzlich angestrebten Präzisionsziele umsetzten.

Ex post, also nach der Stichprobenziehung, zeigt sich, dass sich die Annahmen, welche im Rahmen der Stichprobenplanung getroffen wurden, nicht vollumfänglich bewahrt haben. So wurden die angestrebten Präzisionsziele in größerem Ausmaß verfehlt; besonders auffällig ist dabei die Gemeindegrößenklasse von 10 000 bis 50 000 Einwohnerinnen und Einwohner. Darüber hinaus sind Abweichungen feststellbar für Gemeindegrößenklassen unter 10 000 Einwohnerinnen und Einwohnern, für die im Jahr 2011 keine stichprobenbasierte Anpassung der Melderegister erfolgte.

Der Gesetzgeber hat schon vorab berücksichtigt, dass die angestrebten Genauigkeitsziele, unter anderem durch die Weiterentwicklung der Methodik, gegebenenfalls nicht für alle Gemeinden realisierbar sind. Daher sind die Präzisionsziele des § 11 Zensusgesetz 2022 explizit als anzustrebende und nicht als verbind-

lich einzuhaltende Werte festlegt. In der Begründung zum Zensusgesetz 2022 wird bereits konstatiert, es sei absehbar, dass mit dem zugrunde zu legenden Stichprobenumfang die angestrebten Genauigkeitsziele nicht in allen Gemeinden erreicht werden können. Mit dem Satz „Auch bei Nichterreichen der Genauigkeitsziele sind jedoch nach Satz 2 nachträgliche erneute Stichprobenziehungen ausgeschlossen.“ (ZensG, Begründung zu Zensus 2022 | [Häufiges Verfehlen des angestrebten Präzisionsziels bei der Einwohnerzahlermittlung] 2/9 § 11, Absatz 2) erkennt der Gesetzgeber explizit die Möglichkeit des Verfehlers von Präzisionszielen an.

Das Ausmaß der verfehlten Genauigkeitsziele lag höher als im Zensus 2011. Dies war zu erwarten, da für wesentlich mehr, insbesondere kleine Gemeinden Stichprobenergebnisse zu ermitteln waren. Ex ante ist die Stichprobenplanung von einem Präzisionsgewinn durch die Ausweitung der Anschriftengrößenklassen ausgegangen. Dieser Effekt scheint sich ex post in Teilen umgedreht zu haben. Insbesondere durch sogenannte Schichtwechsler, das heißt Anschriften, die aufgrund der Anzahl der tatsächlich am Stichtag dort wohnhaften Personen in einer anderen Schicht hätten verortet werden müssen, hat sich wohl die Gesamtstruktur deutlicher verändert als erwartet. Dies führte in der Summe dazu, dass in einer Reihe von Schichten zu hohe Auswahlätze, im Bezug zum Präzisionsziel, umgesetzt wurden und andere Schichten durch das Fehlen dieser Stichprobeneinheiten das Präzisionsziel verfehlten.

Verschärft haben die Situation im Weiteren auch die Veränderungen der Wohnsituation durch die Flüchtlingsbewegungen sowie die Coronakrise. Insgesamt führten diese Effekte vielfach dazu, dass entweder betroffene Personen bereits melderechtlich erfasst worden waren, sich zum Zensusstichtag aber gar nicht mehr an der Anschrift aufhielten, oder umgekehrt Personen an einer Anschrift wohnhaft waren, ohne bislang jedoch melderechtlich dort erfasst zu sein. Dieser Effekt war in der Stichprobenplanung nur unzureichend antizipierbar gewesen. Hier fehlten ausreichend regional feingegliederte Informationen.

Die genauen Ursachen sowie der Einfluss dieser auf die Besetzungszahlen und Varianzen in den Schichten sind nun im Weiteren detailliert zu untersuchen. Auch die Wahl der Anschriftengrößenklassen ist zu überprüfen.

Die sich dabei ergebenden Ergebnisse und Erkenntnisse werden mit Vorliegen veröffentlicht und gehen dann, abhängig von der künftigen Konzeption, in die Stichprobenplanungen für den Zensus 2031 ein.

9

Fazit

Mit dem Zensus 2022 erfolgte nach 2011 zum zweiten Mal eine regional detaillierte Ermittlung der Einwohnerzahl in Deutschland auf der Grundlage einer Kombination von administrativen und Befragungsdaten. Die grundlegende Methodik basierte auf dem Zensustest von 2001 und wurde laufend, dies insbesondere aufgrund der Erfahrungen mit dem Zensus 2011, weiterentwickelt. Im Kern war die Befragung in Form einer Anschriftenstichprobe weiterhin notwendig, da derzeit die Register noch nicht die Präzision und Vollständigkeit aufweisen, um die Bevölkerungszahl in hinreichender Qualität zu ermitteln, wie dies Eurostat, das Statistische Amt der Europäischen Union, vorgibt.

Der vorliegende Beitrag erläutert die Ermittlung der Einwohnerzahl auf der Grundlage der Anschriftenstichprobe in den Sampling Points durch Korrektur der Einwohnerzahl aus den Melderegistern durch die Schätzung von Karteileichen und Fehlbeständen. Hierbei wurde das schon im Zensus 2011 genutzte Verfahren der GREG-Schätzung angewandt. Diese Methodik wurde 2018 seitens des Bundesverfassungsgerichts als adäquat anerkannt.

Weiterentwickelt wurde das Stichprobenverfahren aufgrund der Anmerkungen des Bundesverfassungsgerichts zur Behandlung der Gemeinden mit weniger als 10 000 Einwohnerinnen und Einwohnern. Der Aufsatz stellt die hierzu notwendige methodische Weiterentwicklung dar. Mit Vorliegen der Ergebnisse des Zensus 2022 zeigt sich, dass die in der Stichprobenplanung angestrebten und umgesetzten Präzisionsziele ex post in Teilen nur unzureichend erreicht werden konnten. Der Aufsatz bietet hier erste Erklärungen, die Analysen hierzu müssen aber noch fortgesetzt werden.

Als Fazit kann festgehalten werden, dass das mit wissenschaftlicher Fundierung von den Statistischen Ämtern des Bundes und der Länder weiterentwickelte Verfahren

des Zensus und die daraus resultierenden Ergebnisse des Zensus 2022 ein sachgerechtes Bild der regionalen Verteilung der Bevölkerung in den Gemeinden darstellt. Dies nicht an allen Stellen mit der angestrebten Präzision, aber auch mit Blick auf andere amtliche Stichprobenerhebungen mit hinreichend guter Qualität.

Die Präzisionsziele des Zensus 2022 sind Gegenstand weiterer auch wissenschaftlicher Untersuchungen, die Ergebnisse hierzu werden so benötigt Eingang in die Methodik zum Zensus 2031 finden. [u](#)

LITERATURVERZEICHNIS

- Berg, Andreas/Bihler, Wolf. [*Das Stichprobendesign der Haushaltsstichprobe des Zensus 2011*](#). In: Wirtschaft und Statistik. Ausgabe 4/2011, Seite 317 ff.
- Bretsch, Corinna/Lorentz, Kai. [*Präzisionsziele für die Ermittlung der Einwohnerzahl bei der Haushaltsstichprobe im Zensus 2021*](#). In: WISTA Wirtschaft und Statistik. Sonderheft Zensus 2021. 2019, Seite 12 ff.
- Bretsch, Corinna/Seibel, Steffen/Vorndran, Ingeborg/Pfahl, Miriam. [*Ermittlung der Einwohnerzahl im Zensus 2022*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2024, Seite 17 ff.
- Burgard, Jan Pablo/Münnich, Ralf/Rupp, Martin. *Qualitätszielfunktionen für stark variierende Gemeindegrößen im Zensus 2021*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 14. Ausgabe 1/2020, Seite 5 ff.
DOI: [10.1007/s11943-019-00256-6](https://doi.org/10.1007/s11943-019-00256-6)
- Klink, Steffen/Lorentz, Kai. [*Auswahlplan und Stichprobenhauptziehung für den Zensus 2022*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2022, Seite 13 ff.
- Krug, Walter/ Nourney, Martin/Schmidt, Jürgen. *Wirtschafts- und Sozialstatistik, Gewinnung von Daten*. 6. Auflage. München, Wien 2001.
- Münnich, Ralf/Gabler, Siegfried/Ganninger, Matthias/Burgard, Jan Pablo/Kolb, Jan-Philipp. *Das Stichprobendesign des registergestützten Zensus 2011*. In: Methoden, Daten, Analysen. Jahrgang 5. Ausgabe 1/2011, Seite 37 ff.
[Zugriff am 11. Dezember 2024]. Verfügbar unter: nbn-resolving.org
- Särndal, Carl-Erik/Swensson, Bengt/Wretman, Jan. *Model assisted survey sampling*. New York 1992.
- Statistische Ämter des Bundes und der Länder. [*Ergebnisse des Zensustests*](#). In: Wirtschaft und Statistik. Ausgabe 8/2004, Seite 813 ff.
- Statistische Ämter des Bundes und der Länder. *Zensus 2011. Methoden und Verfahren*. 2015. [Zugriff am 11. Dezember 2024]. Verfügbar unter: www.zensus2011.de
- Statistisches Bundesamt. *Stichproben in der amtlichen Statistik*. Wiesbaden 1960.
- Wittmaack, Moritz. [*Mehrfachfallprüfung im Zensus 2022 – die neue Strategie zur automatisierten Identifikation und Bewertung von Mehrfachfallfällen*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2025, Seite 127 ff.

RECHTSGRUNDLAGEN

Gesetz zur Durchführung des Zensus im Jahr 2022 (Zensusgesetz 2022 – ZensG 2022) vom 26. November 2019 (BGBl. I Seite 1851), das durch Artikel 2 des Gesetzes vom 3. Dezember 2020 (BGBl. I Seite 2675) geändert worden ist.

Stefanie Hoffmann

hat einen Masterabschluss in Politikwissenschaft und ist seit 2022 wissenschaftliche Mitarbeiterin im Referat „Hochschulen“ des Statistischen Bundesamtes. Ein Schwerpunkt ihrer Arbeit liegt auf der Betreuung und Weiterentwicklung der Studienverlaufsstatistik, insbesondere der Entwicklung neuartiger Quotenberechnungen, die erstmals auf Basis von Verlaufsdaten möglich sind.

Tobias Laick

hat Wirtschaftswissenschaften studiert und ist seit Ende 2018 im Referat „Hochschulen“ des Statistischen Bundesamtes tätig. Er begleitet die Studienverlaufsdatenbank bereits seit der Produktivsetzung im Juli 2019 und ist somit als Experte für die Konzipierung und Durchführung von Verlaufsenauswertungen zuständig.

Dr. Lorna Syme

ist promovierte Wirtschaftswissenschaftlerin und seit 2020 im Statistischen Bundesamt tätig. Seit 2023 ist sie als Referentin im Referat „Hochschulen“ unter anderem für die Studienverlaufsstatistik fachlich zuständig.

DAS LEHRAMTSSTUDIUM IM SPIEGEL DER HOCHSCHULSTATISTIK

Stefanie Hoffmann, Tobias Laick, Lorna Syme

➤ **Schlüsselwörter:** Lehrkräfteausbildung – Studienverlaufsstatistik – Abschlusswechsel – Übergang vom Bachelor- zum Masterstudium – Studienabbruch

ZUSAMMENFASSUNG

Die Ausbildung neuer Lehrkräfte ist eine zentrale Aufgabe des Bildungssystems, um den aktuell diskutierten Lehrkräftemangel an den Schulen in Deutschland zu beheben. Die Lehrkräfteausbildung erfolgt maßgeblich an den Hochschulen, weshalb die Daten der amtlichen Hochschulstatistik besonders gefragt sind. Die Hochschulstatistik liefert seit Jahrzehnten flächendeckende und über die Zeit vergleichbare Aggregatdaten zu Studierenden und Absolvierenden von Lehramtsstudiengängen. Seit 2017 werden diese durch amtliche Daten auf Basis individueller Verläufe ergänzt, die die Analysemöglichkeiten erheblich verbessern. Der Aufsatz gibt einen Überblick über die Ergebnisse zu Studierenden und Absolvierenden im Lehramt und stellt die Auswertungsmöglichkeiten auf Basis von Studienverlaufsdaten vor.

➤ **Keywords:** teacher training – study pathway statistics – change of degree programme – transition from bachelor's to master's degree – student withdrawal

ABSTRACT

The training of new teachers is a key function of the education system and aims to address the teacher shortage in German schools that is currently being discussed. As teacher training primarily takes place at higher education institutions, data from official higher education statistics are in particular demand. For decades, higher education statistics have provided comprehensive and comparable aggregate data on students and graduates of teacher training degree programmes. Since 2017, this information has been supplemented by official data based on individual study pathways, thereby greatly enhancing the opportunities for analysis. The article provides an overview of the results on students and graduates of teacher training degree programmes and presents the opportunities for evaluation on the basis of data on study pathways.

1

Einleitung

In einer Zeit, in der Bildung als wesentlicher Faktor für die gesellschaftliche und wirtschaftliche Entwicklung gilt, ist die Sicherstellung einer ausreichenden Anzahl gut ausgebildeter Lehrkräfte unerlässlich. Lehrkräftemangel führt unter anderem zu Unterrichtsausfall und einer erhöhten Arbeitsbelastung bei aktiven Lehrkräften und stellt eine gewaltige Herausforderung für das Bildungssystem dar.

Die deutsche Lehrkräfteausbildung ist in zwei Phasen gegliedert: das Hochschulstudium, das fachwissenschaftliche und fachdidaktische Kenntnisse vermittelt und laut Regelstudienzeit zwischen 4 und 5 Jahren dauert, und den üblicherweise 1,5 bis 2 Jahre dauernden Vorbereitungsdienst – das sogenannte Referendariat –, in dem die angehenden Lehrkräfte praktische Erfahrungen sammeln und pädagogische Fähigkeiten vertiefen (SWK, 2023). Die Hochschulstatistik bietet einen breiten Kranz amtlicher Daten zur ersten Phase der Ausbildung – dem Hochschulstudium – und kann Hochschulen und Bildungspolitik dabei unterstützen, das Lehramtsstudium zu evaluieren.

Bereits seit Jahrzehnten liefert die amtliche Hochschulstatistik Hochschulen, Verwaltung und Politik relevante Planungsdaten über Zahl und Entwicklung von Studienanfängerinnen beziehungsweise Studienanfängern, Studierenden und Absolvierenden im Lehramtsstudium, über die fachliche Differenzierung, die Altersstruktur und weitere Merkmale. Da die Daten berichtszeitraumbezogen erhoben werden und nicht über die Zeit verknüpft werden konnten, waren die Analysemöglichkeiten bisher im Wesentlichen auf Aggregate beschränkt. Die Einführung einer Studienverlaufsstatistik auf Basis individueller Studienverläufe bietet seit Kurzem die Möglichkeit, den Übergang vom Lehramtsbachelor in den Lehramtsmaster, den Wechsel zwischen Lehramtsstudiengängen und Nicht-Lehramtsstudiengängen sowie einen Abbruch eines Lehramtsstudiums abzubilden (Neumann/Brings, 2021).

Dieser Aufsatz stellt zunächst die Organisation des Lehramtsstudiums in Deutschland vor. Anschließend folgt eine Untersuchung der Studierenden- und Absolventenzahlen in Lehramtsstudiengängen seit Beginn der

1990er-Jahre. Danach werden mithilfe von Auswertungen aus der Studienverlaufsdatenbank Studienverläufe im Lehramt analysiert. Im Fokus der Analyse steht die vergleichende Betrachtung von Lehramtsstudierenden und Nicht-Lehramtsstudierenden. Ein abschließendes Fazit bewertet die zentralen Ergebnisse.

2

Organisation des Lehramtsstudiums in Deutschland

Infolge der Bologna-Erklärung von 1999 zur Schaffung eines Europäischen Hochschulraums (Bologna-Reform) wurde in Deutschland ein System gestufter Studiengänge eingeführt. Ziel dabei war, die Hochschulabschlüsse international vergleichbar zu machen, den internationalen Austausch an den Hochschulen zu fördern und die Wettbewerbsfähigkeit der deutschen Absolventinnen und Absolventen zu erhöhen. Diplom- oder Masterstudiengänge mit einer durchschnittlichen Regelstudienzeit von 4 bis 5 Jahren wurden in zwei gestufte Studienabschnitte mit jeweils separatem Studienbeginn und Abschluss aufgeteilt: ein Bachelorstudium, das wissenschaftliche Grundlagen vermittelt und mit einem ersten berufsqualifizierenden Abschluss endet, sowie ein Masterstudium, das eine anwendungs- oder forschungsorientierte Vertiefung bietet und an das Bachelorstudium angeschlossen werden kann.

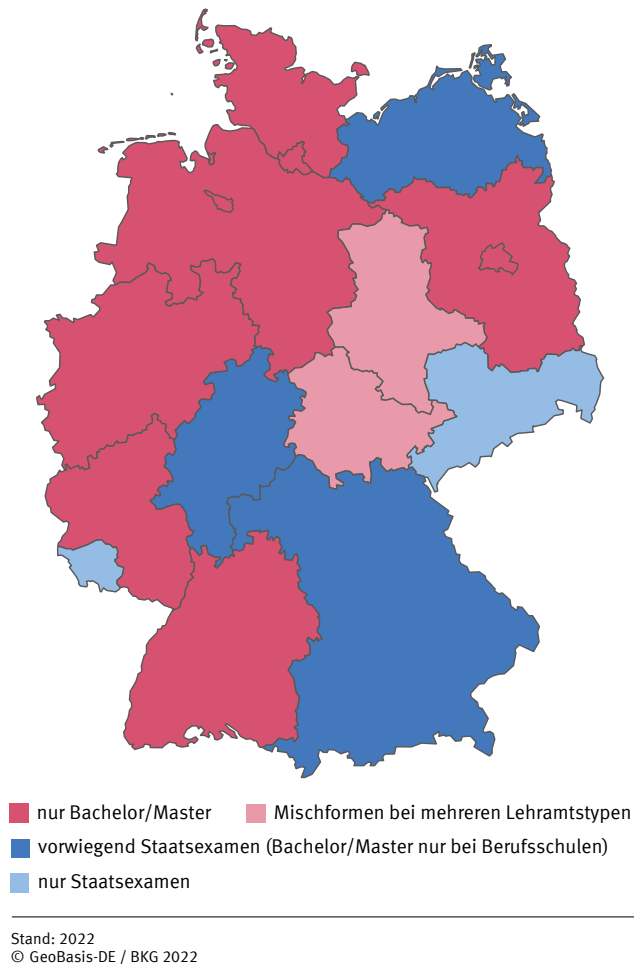
Bis zur Bologna-Reform erfolgte die hochschulische Lehrkräfteausbildung ausschließlich in Staatsexamensstudiengängen. Einige Bundesländer halten bis heute an dieser Struktur des Lehramtsstudiums fest. Andere Länder haben, wenn auch zu unterschiedlichen Zeitpunkten, die hochschulische Lehrkräfteausbildung vollständig auf das Bachelor-Master-System umgestellt, weitere bieten beide Systeme parallel an. Diese Unterschiede resultieren aus der föderalen Struktur Deutschlands, bei der die Bildungspolitik Ländersache ist, und erschweren Vergleiche zwischen den Bundesländern.

➤ Grafik 1

Eine weitere Besonderheit ist bei der Analyse gestufter Studiengänge im Lehramt zu berücksichtigen: Mit dem Lehramtsbachelor ist zwar bereits ein erster berufsqualifizierender Abschluss erreicht, dieser reicht jedoch in der Regel weder für die Aufnahme einer Berufstätigkeit

Grafik 1

Organisation des Lehramtsstudiums in Deutschland



als Lehrerin oder Lehrer, noch als Zugang zum Vorbereitungsdienst aus.

3

Analysemöglichkeiten der amtlichen Hochschulstatistik

Die amtliche Hochschulstatistik basiert auf Verwaltungsdaten der Hochschulen. Gemäß dem Hochschulstatistikgesetz sind alle in Deutschland staatlich anerkannten Hochschulen verpflichtet, Daten an die Statistischen Ämter des Bundes und der Länder zu melden. Somit sind vollständige Aussagen über das Hochschulsystem möglich. Statistikdaten zu Studierenden sind semester-

weise zu liefern. Anhand der Angaben zum Fach- und Hochschulsemester lässt sich die Zahl der Studienanfängerinnen und Studienanfänger eines Studienjahrs bestimmen. Die Absolventinnen und Absolventen eines Prüfungsjahrs¹ können über die Prüfungsstatistik identifiziert werden, zu der die Hochschulen semesterweise für jeden Studiengang Prüfungsteilnehmende und abgelegte Abschlussprüfungen melden müssen.

Studierenden- und Prüfungsstatistik ermöglichen es, Aussagen über Input und Output des Lehramtsstudiums auf Basis aggregierter Zahlen zu treffen. Allerdings lassen sich anhand dieser Daten keine individuellen Studienverläufe abbilden und Veränderungen zwischen Studienbeginn und möglichem Abschluss nicht oder nur unzureichend analysieren. Daher war es bisher nicht möglich, folgende Fragestellungen zur Evaluierung des Lehramtsstudiums in Deutschland zu beantworten:

- › Wie viele Studierende beginnen ein Lehramtsstudium und entscheiden sich in einer frühen Phase ihrer Studienlaufbahn für einen Abschlusswechsel?
- › Wie viele Studierende brechen ihr Lehramtsstudium in einer frühen Phase ihrer Studienlaufbahn ab?
- › Wie viele Lehramtsstudierende nehmen nach einem erfolgreichen Bachelorabschluss ein Lehramtsmasterstudium auf?
- › Wie verhalten sich Nicht-Lehramtsstudierende im Vergleich zu Lehramtsstudierenden hinsichtlich Abschlusswechsel, Studienabbruch und Übergang in weiterführende Studiengänge?

Antworten ermöglicht die noch recht neue Studienverlaufsstatistik. Sie wurde mit der Novelle des Hochschulstatistikgesetzes von 2016 eingeführt und erlaubt erstmals, einzelne Datensätze der Studierenden- und Prüfungsstatistik miteinander sowie über die Zeit zu verknüpfen (Neumann/Brings, 2021). Anhand der Studienverlaufsstatistik können seit dem Sommersemester 2017 somit individuelle Studienverläufe statistisch abgebildet werden. Dadurch lassen sich beispielsweise Aussagen zu Abschlusswechseln, Studienabbrüchen sowie zu Übergängen vom Bachelor- ins Masterstudium treffen.

¹ Das Studienjahr besteht in der Hochschulstatistik aus einem Sommersemester und dem nachfolgenden Wintersemester (Studienjahr 2023 = Sommersemester 2023 + Wintersemester 2023/2024); das Prüfungsjahr besteht aus einem Wintersemester und dem darauf folgenden Sommersemester (Prüfungsjahr 2023 = Wintersemester 2023/2024 + Sommersemester 2024).

3.1 Entwicklung der Studierenden- und Absolventenzahlen im Lehramtsstudium

Studierende

Seit 1993 ist die Zahl der Lehramtsstudierenden gestiegen. Während im Wintersemester 1993/1994 insgesamt 209 000 Studierende im Lehramt immatrikuliert waren, ist diese Zahl bis zum Wintersemester 2023/2024 auf 260 000, also um 25 %, angestiegen. Gleichzeitig ist die Gesamtstudierendenzahl stärker gestiegen, nämlich um 54 %; daher hat sich der Anteil der Lehramtsstudierenden an den Studierenden insgesamt rückläufig entwickelt: Er ist von 11 % im Jahr 1993 auf 9 % im Jahr 2023 gesunken. ➤ Grafik 2

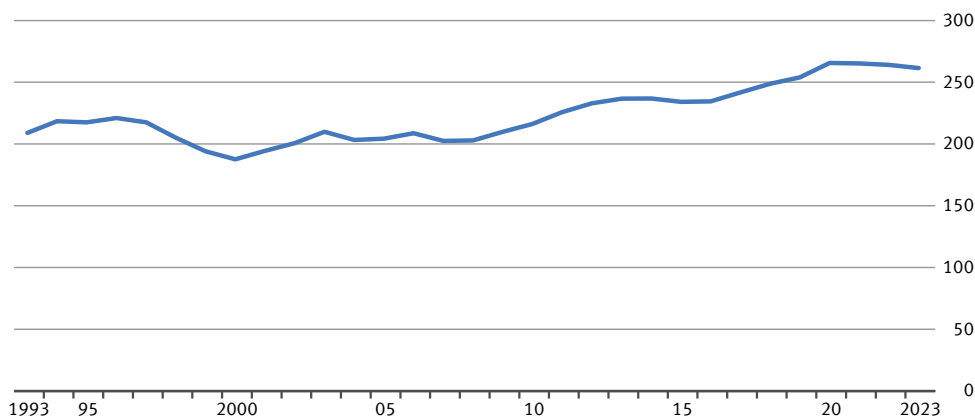
Ein Grund für die stärker gestiegene Gesamtstudierendenzahl ist die Ausweitung stark berufsorientierter Studiengänge. Ein Beispiel für die Akademisierung der Ausbildung sind die Gesundheitsberufe: So wurde in den letzten Jahren die Ausbildung von Hebammen sowie in den Bereichen Logopädie/Sprachtherapie oder Physiotherapie, die zuvor vor allem an beruflichen Schulen erfolgte, an die Hochschulen verlagert.

➤ Grafik 3 zeigt die Anzahl der Studierenden, die im Zeitraum von 1993 bis 2023 ein Lehramtsstudium aufgenommen haben, sowie die Anzahl derjenigen, die in diesem Zeitraum ein Lehramtsstudium erfolgreich abgeschlossen haben. Dabei ist zu beachten, dass sowohl die Anfängerinnen und Anfänger als auch die Absolventinnen und Absolventen in der Hochschulstatistik gemäß dem Bologna-Gedanken für jeden Studienab-

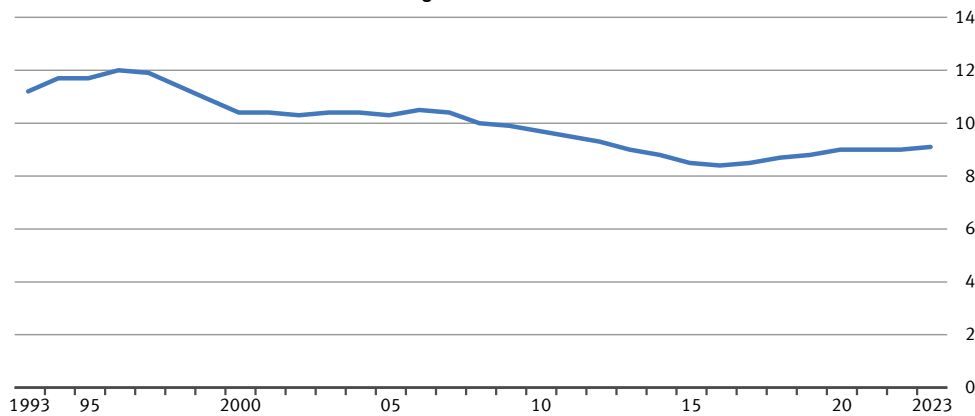
Grafik 2

Entwicklung der Studierenden im Lehramt

Anzahl Lehramtsstudierende in 1 000



Anteil Lehramtsstudierende an Studierenden insgesamt in %

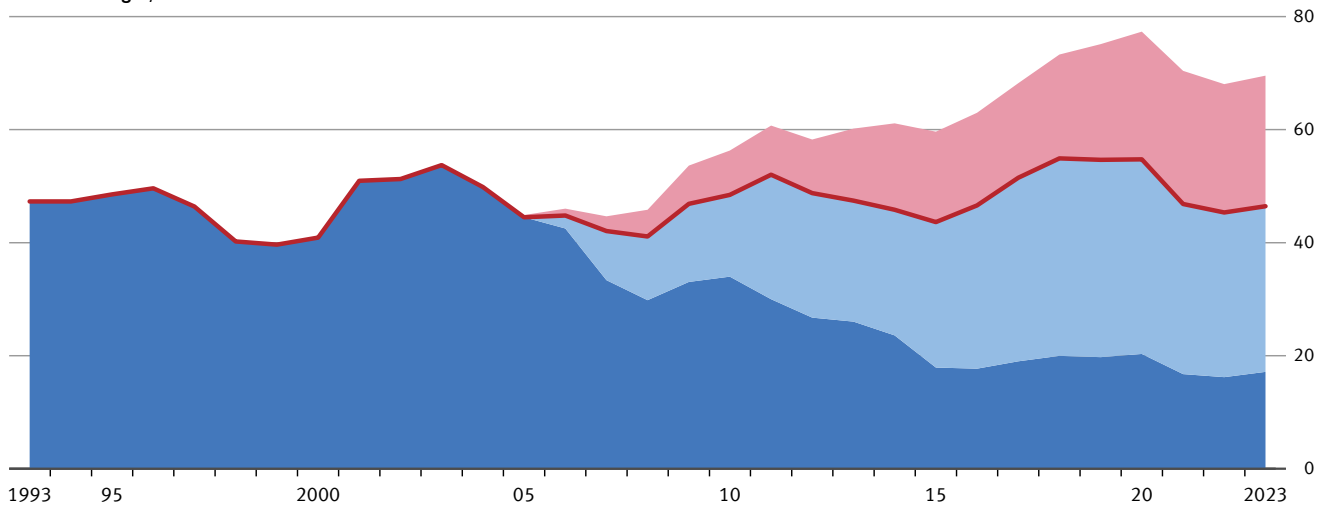


Studierende im Wintersemester.

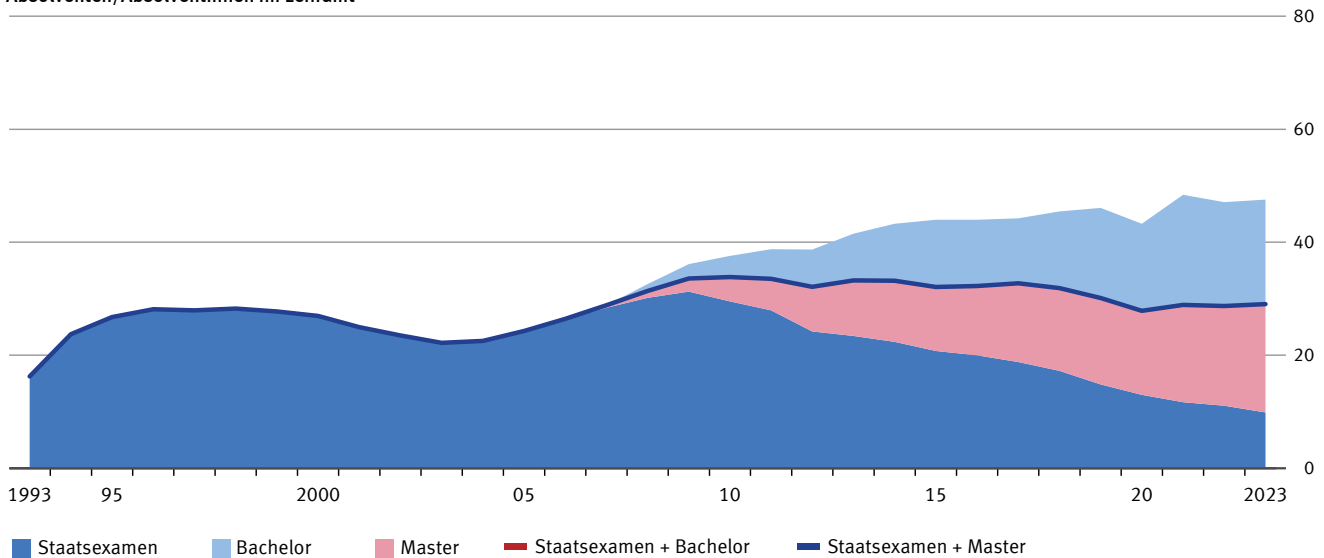
Grafik 3

Aufnahmen und Abschlüsse von Lehramtsstudiengängen
in 1 000

Studienanfänger/-innen im Lehramt



Absolventen/Absolventinnen im Lehramt



Studienanfänger/-innen im Lehramt im 1. Fachsemester im Studienjahr sowie Absolventen/Absolventinnen im Lehramt im Prüfungsjahr.

schnitt separat gezählt werden. Lehramtsstudierende, die zunächst ein Bachelorstudium und anschließend das Masterstudium absolvieren, werden in der Statistik auf ihrem hochschulischen Ausbildungsweg somit bis zum Referendariat je zweimal als Studienanfängerin oder Studienanfänger sowie als Absolvierende gezählt, Staatsexamensstudierende dagegen nur jeweils einmal. Um den Effekt der teilweisen Umstellung der hochschulischen Lehramtsausbildung vom Staatsexamen auf das Bachelor-Master-System seit 2006 auf die Entwicklung

von Studienanfänger- und Absolventenzahlen zu visualisieren, kennzeichnet die durchgezogene Linie in der Zeitreihe zu Studienanfängerinnen und Studienanfängern im Lehramt die Teilmenge mit Start eines grundständigen Studiums (Staatsexamens- und Bachelorstudiengänge) und in der Zeitreihe zu Absolvierenden im Lehramt die Teilmenge mit Abschluss eines zum Vorbereitungsdienst berechtigenden Studiums (Staatsexamens- und Masterstudiengänge). Für alle weiteren Analysen in diesem Abschnitt werden als Studienanfän-

gerinnen und Studienanfänger sowie Absolventinnen und Absolventen nur diese Teilmengen betrachtet.

Studienanfängerinnen und -anfänger

Im Gegensatz zum Gesamtanstieg bei den Lehramtsstudierenden in den letzten 30 Jahren lag die Zahl der Anfängerinnen und Anfänger eines grundständigen Studiums im Lehramt 2023 um 1,8 % niedriger als 1993. Während im Studienjahr 1993 noch 47 300 Personen ein Lehramtsstudium begannen, waren es im Studienjahr 2023 nur noch 46 400 (Bachelor oder Staatsexamen). Demgegenüber ist die Zahl der Absolvierenden im Lehramt (Staatsexamen und Master) von 16 200 Personen im Jahr 1993 auf 29 000 Personen im Jahr 2023 gestiegen.

Die zahlenmäßige Entwicklung bei den Studienanfängerinnen und Studienanfängern im Lehramt seit 1993 verlief wellenförmig. Eine Ursache für die zyklische Entwicklung dürften die Veränderungen beim tatsächlichen oder wahrgenommenen Bedarf an neuen Lehrerinnen und Lehrern sein. Dabei führen eine hohe Nachfrage nach Lehrkräften und die guten Berufsaussichten zu mehr Studienanfängerinnen und Studienanfängern und vice versa. Da sich die Lehrkräfteausbildung über einen längeren Zeitraum erstreckt, in dem sich die Arbeitsmarktsituation bis zum Studienabschluss wieder verschlechtern kann, passen Nachfrage und Angebot nicht immer zusammen.

Die Zahl der Studienanfängerinnen und Studienanfänger im Lehramt verlief zwischen 1996 und 1999 zunächst stark rückläufig, stieg dann bis 2003 wieder an. Danach folgte ein neuerlicher Rückgang bis 2008, gefolgt von einem Anstieg bis 2011. Dieses Muster setzt sich bis zum aktuellen Rand (2023) fort, wobei die jeweiligen Tiefpunkte kontinuierlich über denen der vorherigen Zyklen lagen.

Grafik 3 zeigt die über alle Schularten summierte zahlenmäßige Entwicklung bei Anfängerinnen und Anfängern sowie Absolvierenden eines Lehramtsstudiums. Die hochschulische Lehramtsausbildung erfolgt aber nicht generalistisch, sondern spezifisch für die Anforderungen und Inhalte in den einzelnen Schularten. Die Kultusministerkonferenz unterscheidet insgesamt sechs Lehramtstypen (KMK, 1999). [↗ Übersicht 1](#) Dabei können zwischen den Lehramtstypen nicht nur die spezifischen Studieninhalte und die Regelstudienzeiten variieren, sondern auch die spezifische Lehrkräftenachfrage.

Übersicht 1

Lehramtstypen in der Unterscheidung der Kultusministerkonferenz

	Definition
Lehramtstyp 1	Lehrämter der Grundschule beziehungsweise der Primarstufe
Lehramtstyp 2	Übergreifende Lehrämter der Primarstufe und aller oder einzelner Schularten der Sekundarstufe I
Lehramtstyp 3	Lehrämter für alle oder einzelne Schularten der Sekundarstufe I
Lehramtstyp 4	Lehrämter der Sekundarstufe II (allgemeinbildende Fächer) oder für das Gymnasium
Lehramtstyp 5	Lehrämter der Sekundarstufe II (berufliche Fächer) oder für die beruflichen Schulen
Lehramtstyp 6	Sonderpädagogische Lehrämter

Absolventinnen und Absolventen

Die zahlenmäßige Entwicklung der Absolventinnen und Absolventen eines Lehramtsstudiums im Zeitraum 2008 bis 2023 nach den sechs Lehramtstypen ist in [↗ Grafik 4](#) dargestellt. Dabei sind deutliche Unterschiede in der Verteilung festzustellen. Die größte Gruppe von Absolvierenden findet sich im Lehramtstyp 4, der für die Sekundarstufe II und das Gymnasium ausbildet und im abgebildeten Zeitraum zwischen 38 und 46 % der Absolvierenden umfasst, gefolgt vom Lehramtstyp 1, der Grundschulen beziehungsweise die Primarstufe abdeckt. Der Anteil des Lehramtstyps 1 an allen Lehramtsabsolvierenden lag zwischen 2008 und 2023 zwischen 23 und 28 %.

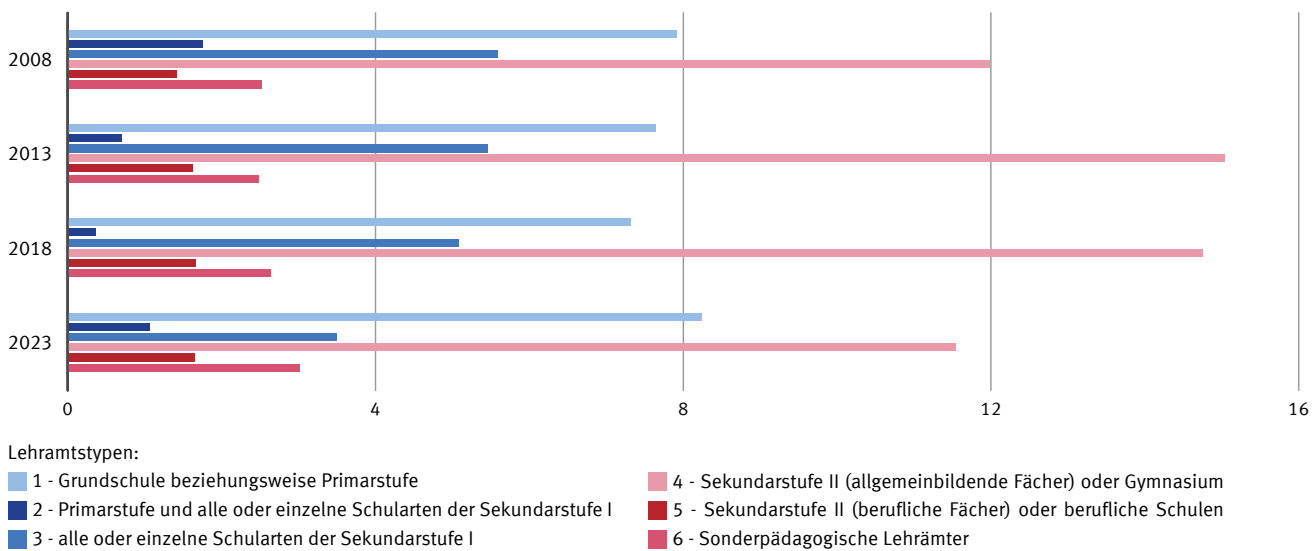
In der Zeitverlaufsbetrachtung von 2008 bis 2023 zeigt sich, dass sich die Anteile in den Lehramtstypen 1 und 4 gegenläufig entwickelt haben. In den Jahren 2013 und 2018, als die Anteilswerte für den Lehramtstyp 1 überdurchschnittlich hoch waren, wies gleichzeitig der Lehramtstyp 4 unterdurchschnittliche Anteilswerte auf.

Der Anteil des Lehramtstyps 2 „Übergreifende Lehrämter der Primarstufe und aller oder einzelner Schularten der Sekundarstufe I“ an allen Lehramtsabschlüssen sank zwischen 2008 und 2018 von knapp 6 auf 1 %. Dieser quantitative Bedeutungsverlust ist Folge des Auslaufens vieler Studiengänge in diesem Bereich. Für 2023 hat sich hier ein Anstieg auf 4 % ergeben; dieser ist auf Studierende zurückzuführen, die aufgrund von auslaufenden Prüfungsordnungen zum Abschluss kommen mussten.

Seit 2008 liegt der Anteil der Absolventinnen konstant zwischen 72 und 75 %. Mit Blick auf die Geschlechterverteilung innerhalb der Lehramtstypen zeigt sich, dass der Anteil der Absolventinnen im Lehramtstyp 1 (Grund-

Grafik 4

Absolventen/Absolventinnen (Master und Staatsexamen) nach Lehramtstyp
in 1 000



Nicht nachgewiesen sind sonstige, nicht eindeutig den oben genannten Lehramtstypen zuordenbare Lehrämter.

schule) mit Werten zwischen 88 und 90% erwartungsgemäß am höchsten ist und über die Zeit stabil auf einem hohen Niveau bleibt. Im Gegensatz dazu haben die Lehramtstypen 4 (Gymnasium) mit Werten zwischen 34 und 37 % und 5 (Berufsschulen) mit Werten zwischen 37 und 44 % den höchsten Anteil an Absolventinnen und Absolventen.

3.2 Studienverlauf von Lehramtsstudierenden

Analysen über Lehramtsstudierende oder zum Studium allgemein waren in der amtlichen Hochschulstatistik bis zur Novellierung des Hochschulstatistikgesetzes 2016 auf die berichtszeitraumbezogenen Ergebnisse der Studierenden- und Prüfungsstatistik beschränkt. Wie im Abschnitt 3.1 dargestellt, ermöglichen diese Daten seit langer Zeit eine zuverlässige Betrachtung von Struktur und Entwicklung auf Aggregatebene. Mit Aggregatdaten lässt sich der individuelle Studienverlauf jedoch nicht beziehungsweise nur unzureichend abbilden. Die Einführung der Studienverlaufsstatistik auf Individualdatenbasis, die spezifische Analysen zum Wechsel oder zum Abbruch eines Studiums erlaubt, verspricht auch für die statistische Evaluierung des Lehramtsstudiums eine große Qualitätsverbesserung.

Wechsel zwischen Lehramts- und Nicht-Lehramtsstudiengängen

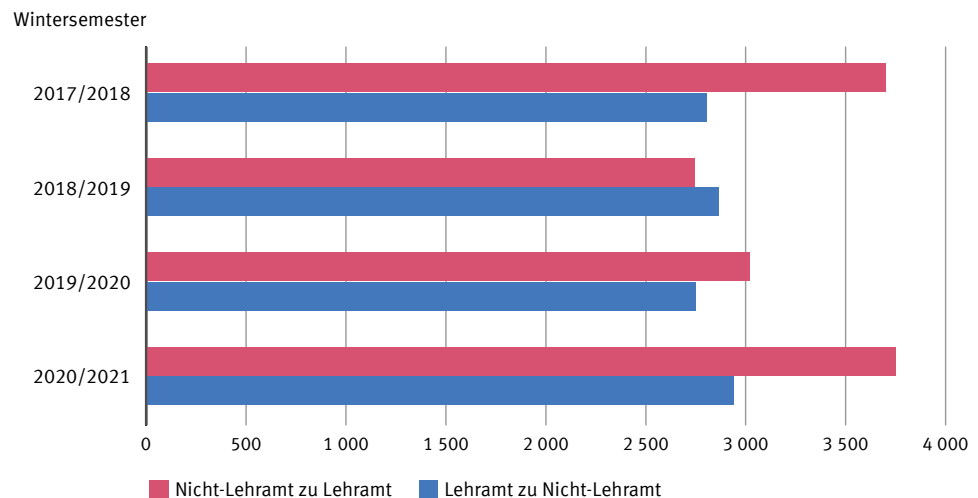
Auf Basis der Studienverlaufsstatistik können Abschlusswechsel beobachtet werden. Somit kann eine Aussage darüber getroffen werden, wie viele Studienanfängerinnen und Studienanfänger (1. Fachsemester) innerhalb der ersten drei Semester von einem Lehramtsstudium in ein Nicht-Lehramtsstudium wechseln und wie viele den entgegengesetzten Wechsel vornehmen. Ein Überschuss der aus dem Lehramtsstudium heraus Wechselnden könnte aus früh enttäuschten Erwartungen an die Studieninhalte beziehungsweise -anforderungen oder einer schlechten Organisation des Lehramtsstudiums resultieren.

➤ **Grafik 5** quantifiziert die Entwicklung der Wechselbewegungen über die Zeit für insgesamt vier Anfängerkohorten, jeweils unterschieden nach den beiden Vergleichsgruppen: Studierende, die ein Lehramtsstudium beginnen, und solche, die ein Nicht-Lehramtsstudium aufnehmen.

Im Wintersemester 2020/2021 haben 54 500 Studierende ein Studium mit einer angestrebten Abschlussprüfung im Lehramt (Bachelor oder Staatsexamen) begonnen. Von diesen Studienanfängerinnen und Studienanfängern sind 2 900 innerhalb von drei Semestern

Grafik 5

Abschlusswechselnde innerhalb von drei Semestern aus dem und ins Lehramt



Studienanfänger/-innen im ersten Fachsemester im Wintersemester mit Wechsel des Abschlusses und ohne abgelegte Abschlussprüfung innerhalb von drei Semestern.

zu einer Nicht-Lehramt-Abschlussart gewechselt. Umgekehrt sind von den 580 000 Studierenden, die im Wintersemester 2020/2021 das Studium mit angestrebter Abschlussprüfung im Nicht-Lehramt begonnen haben, 3 700 ins Lehramt gewechselt. Diese Zahlen weisen in allen vier Anfängerkohorten eine ähnliche Tendenz auf: Lediglich bei der Anfängerkohorte im Wintersemester 2018/2019 überwiegt die Anzahl von Studierenden, die aus dem Lehramt weg wechselten.

Studienabbruch

Das Abschlusswechselerhalten erlaubt eine Aussage darüber, ob während des Studiums eine fachliche Umorientierung stattfindet. Unzufriedenheit mit dem Studium, Überforderung und Ähnliches können aber auch

zu einem Abbruch der hochschulischen Ausbildung führen. Die Studienabbrecherquote misst, wie viele Studierende in das Hochschulsystem in Deutschland eintreten und ein grundständiges Studium beginnen, die Hochschulkarriere aber innerhalb von drei Semestern ohne Abschluss beenden. [Tabelle 1](#) stellt die Quoten für frühen Studienabbruch bei Lehramts- und Nicht-Lehramtsstudierenden in der Anfängerkohorte 2019² vor.

- 2 Studienanfängerinnen und Studienanfänger, die sich innerhalb von drei Semestern ohne Abschluss exmatrikulieren, werden für einen Karenzzeitraum von drei weiteren Semestern beobachtet. Sollten sie innerhalb dieser Karenzzeit eine Abschlussprüfung ablegen oder sich neu einschreiben, zählen sie nicht zu den Abbrecherinnen und Abbrechern. Angesichts dieser Karenzzeit und des begrenzten Datenbestands in der Studienverlaufsdatenbank ist die aktuellste Kohorte, für die eine Studienabbrecherquote berechnet werden kann, die Anfängerkohorte aus dem Studienjahr 2019.

Tabelle 1

Studienanfänger/-innen (1. Hochschulsesemester) im Erststudium grundständiger Studiengänge nach Art der angestrebten Abschlussprüfung im Studienjahr 2019

	Insgesamt	Darunter: Studienabbrecher/-innen
	Anzahl	Quote in %
Grundständige Studiengänge insgesamt	422 651	10,8
Lehramt zusammen	35 474	7,9
Bachelorstudiengänge	23 846	8,6
Staatsexamen	11 628	6,5
Nicht-Lehramt zusammen	387 177	11,0
Bachelorstudiengänge	347 110	11,7
Sonstige Studiengänge zusammen (unter anderem Staatsexamen)	40 067	5,5

Von den insgesamt 420 000 Studierenden, die im Studienjahr 2019 ein grundständiges Studium begannen, haben 11 % das Studium innerhalb von drei Semestern abgebrochen. Im Vergleich von Lehramts- und Nicht-Lehramtsstudierenden unterscheidet sich diese Quote deutlich. Während bei Nicht-Lehramtsstudiengängen ebenfalls 11 % der Studierenden das Studium innerhalb von drei Semestern abgebrochen haben, lag die Studienabbrecherquote bei Lehramtsstudiengängen bei nur 8 %. Dieser Unterschied gilt auch bei der Betrachtung des Bachelorstudiums, wo weniger Studierende im Lehramt das Studium abbrechen als im Nicht-Lehramt.

Übergang vom Bachelor- ins Masterstudium

Weder das Abschlusswechselverhalten noch die Studienabbrecherquote innerhalb von drei Semestern nach Studienbeginn deuten auf übermäßige Schwächen zu Beginn des Lehramtsstudium hin. So gehen in den Lehramtsstudiengängen anteilig weniger Studierende in einer frühen Studienphase verloren als in den Nicht-Lehramtsstudiengängen. Mit der Studienverlaufsstatistik lässt sich auch der Übergang vom Bachelor- ins Masterstudium abbilden. Wie in Kapitel 2 ausgeführt, reicht ein abgeschlossenes Lehramts-Bachelorstudium allein in der Regel nicht aus, um die zweite Ausbildungsphase – das Referendariat – zu beginnen und später als Lehrkraft tätig zu sein. Um sich für das Referendariat zu qualifizieren, müssen Bachelorabsolventinnen und -absolventen zunächst noch einen Lehramtsmaster absolvieren.

↪ **Tabelle 2** zeigt die Bachelor-Master-Übergangsquoten bei Lehramts- und Nicht-Lehramtsbachelorabschlüssen, berechnet als Anteil aller Bachelorabsolventinnen und -absolventen des Prüfungsjahrs 2021, die im Zeitraum vom Sommersemester 2020 bis Wintersemester 2022/2023 ein Masterstudium begonnen haben. Der

gewählte Analysezeitraum berücksichtigt einerseits, dass einige Studierende ein Masterstudium beginnen, bevor der formale Abschluss ihres Bachelorstudiums vorliegt. Andererseits werden mit einer Karenzzeit von etwa 1,5 Jahren zwischen Abschluss des Bachelorstudiums und Aufnahme des Masterstudiums auch diejenigen Studierenden berücksichtigt, die zwischen den Studienabschnitten eine kurze Lücke aufweisen.

Wie Tabelle 2 zeigt, setzten 86 % der Absolventinnen und Absolventen mit Lehramtsbachelorabschluss im Prüfungsjahr 2021 bis zum Wintersemester 2022/2023 ihr Studium in einem Lehramtsmaster fort. Im Vergleich dazu nehmen innerhalb des gleichen Zeitraums nur 54 % der Absolvierenden mit einem Nicht-Lehramtsbachelorabschluss ein Nicht-Lehramtsmasterstudium auf. Damit weist das Lehramtsstudium eine deutlich höhere Übergangsquote vom Bachelor- zum Masterstudium auf als die herkömmlichen gestuften Studiengänge. Um Verzerrungen infolge des Studienangebots zu vermeiden, bleiben beim Vergleich – und in den folgend präsentierten Ergebnissen – Fachhochschulen und Verwaltungsfachhochschulen unberücksichtigt, da fast ausschließlich Universitäten und gleichrangige Hochschulen das Lehramtsstudium anbieten.

Während sich die Übergangsquoten vom Bachelor- ins Masterstudium zwischen Lehramts- und Nicht-Lehramtsstudiengängen unterscheiden, ähneln sie sich hinsichtlich des Zeitraums zwischen Bachelorabschluss und Masterbeginn. Die semesterweise Betrachtung der Übergangsquoten für die Bachelorabsolventinnen und -absolventen aus dem Prüfungsjahr 2021 zeigt, dass die meisten Bachelorabsolvierenden von Lehramts- wie von Nicht-Lehramtsstudierenden das Masterstudium bis Ende eines Jahres nach dem Bachelorabschluss begonnen haben. Dabei beginnen Absolventinnen und Absolventen

Tabelle 2

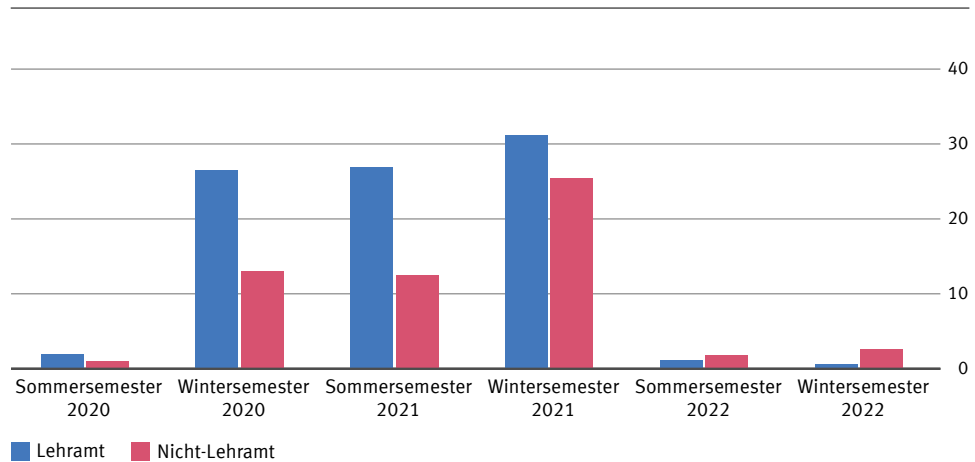
Übergangsquote von Bachelorabsolventen/-absolventinnen im Prüfungsjahr 2021 mit Masterbeginn von Sommersemester 2020 bis Wintersemester 2022

	Lehramt		Nicht-Lehramt	
	Anzahl	%	Anzahl	%
Bachelorabsolventen/-absolventinnen insgesamt	19 176	100	107 771	100
Lehramtsmasterstudium begonnen	16 457	85,8	4 224	3,9
Nicht-Lehramtsmasterstudium begonnen	455	2,4	58 045	53,9
Sonstiges Studium begonnen	148	0,8	4 076	3,8
Kein weiteres Studium begonnen	2 116	11,0	41 426	38,4

Fachhochschulen und Verwaltungsfachhochschulen sind ausgeschlossen.

Grafik 6

Semesterweise Übergangsquoten vom Bachelor- ins Masterstudium in %



Fachhochschulen und Verwaltungsfachhochschulen sind ausgeschlossen. Anteil (%) Lehramtsbachelorabsolventen/-absolventinnen des Prüfungsjahrs 2021 mit Lehramtsmasterbeginn in den jeweiligen Semestern, sowie Anteil Nicht-Lehramtsbachelorabsolventen/-absolventinnen des Prüfungsjahrs 2021 mit Nicht-Lehramtsmasterbeginn in den jeweiligen Semestern. Bachelorabsolventen/-absolventinnen, die gegebenenfalls nacheinander mehrere Masterstudiengänge aufgenommen haben, sind hier mehrfach gezählt. Dadurch ergibt sich in der Summe gegebenenfalls eine Abweichung zu den Übergangsquoten in Tabelle 2.

mit Bachelorabschluss im Lehramt in Relation etwas früher mit dem Lehramtsmasterstudium als Bachelorabsolvierende eines Nicht-Lehramtsstudiengangs mit einem Nicht-Lehramtsmasterstudium. [↗ Grafik 6](#)

Die oben genannte Übergangsquote von 86 % für die Personen, die im Prüfungsjahr 2021 ihr Bachelor-Lehramtsstudium erfolgreich abgeschlossen haben, gilt nicht einheitlich für alle Lehramtstypen, wie [↗ Tabelle 3](#) zu entnehmen ist. Besonders hohe Übergangsquoten von 88 bis 90 % zeigten sich in den Lehramtstypen, die sich auf die Lehrämter für die Grundschule/Primarstufe,

Sekundarstufe I und Sonderpädagogik konzentrieren. Dagegen lagen die Übergangsquoten in den Lehramtstypen, die Lehrämter für Gymnasien, die Sekundarstufe II und berufliche Schulen abdecken, mit 81 bis 82 % deutlich niedriger.

Mit der Bachelor-Master-Übergangsquote wird nicht nur der Teil der Absolventinnen und Absolventen eines Lehramtsbachelors bestimmt, der das Lehramtsstudium im Master fortsetzt. Dabei wird auch untersucht, wer die hochschulische Laufbahn mit dem Bachelorabschluss bereits beendet hat und wer sie außerhalb des Lehramts

Tabelle 3

Übergangsquote von Lehramtsbachelorabsolventen/-absolventinnen nach Lehramtstyp im Prüfungsjahr 2021

	Bachelorabsolventen/-absolventinnen im Prüfungsjahr 2021		
	insgesamt	davon Lehramtsmasterstudium begonnen	
		Anzahl	%
Insgesamt	19 176	16 457	85,8
Lehramtstyp 1 – Lehrämter der Grundschule beziehungsweise der Primarstufe	4 129	3 698	89,6
Lehramtstyp 2 – Übergreifende Lehrämter der Primarstufe und aller oder einzelner Schularten der Sekundarstufe I	1 700	1 537	90,4
Lehramtstyp 3 – Lehrämter für alle oder einzelne Schularten der Sekundarstufe I	1 374	1 237	90,0
Lehramtstyp 4 – Lehrämter der Sekundarstufe II (allgemeinbildende Fächer) oder für das Gymnasium	7 042	5 783	82,1
Lehramtstyp 5 – Lehrämter der Sekundarstufe II (berufliche Fächer) oder für die beruflichen Schulen	1 273	1 033	81,1
Lehramtstyp 6 – Sonderpädagogische Lehrämter	1 743	1 540	88,4
Lehramtsprüfungen Bachelor (soweit keine Aufteilung möglich)	1 915	1 629	85,1

Bachelorabsolventen/-absolventinnen im Prüfungsjahr 2021 im Lehramt mit Lehramtsmasterbeginn von Sommersemester 2020 bis Wintersemester 2022.

fortsetzt. Die Aufnahme eines Masterstudiums im Lehramt nach einem abgeschlossenen Lehramtsbachelor ist nicht zwangsläufig. Die Konzeption der hochschulischen Lehramtsausbildung im System gestufter Studiengänge ermöglicht es genauso, nach einem Lehramtsbachelor einen Master ohne Ausrichtung auf einen Lehramtsabschluss anzuschließen. So knüpften 2,4 % der Absolvierenden im Lehramt im Prüfungsjahr 2021 an ihr abgeschlossenes Lehramtsbachelorstudium ein herkömmliches Nicht-Lehramtsmasterstudium an (siehe Tabelle 2). Umgekehrt beginnen 3,9 % der Absolvierenden eines Nicht-Lehramtsbachelorstudiums ein Lehramtsmasterstudium. Da die Grundgesamtheit der Absolvierenden mit herkömmlichem Bachelor deutlich größer ist als die mit Lehramtsbachelor, wechseln damit auch absolut gesehen mehr Studierende beim Übergang vom Bachelor zum Master in ein Lehramtsstudium hinein als aus dem Lehramtsstudium heraus.

Von den Personen, die im Prüfungsjahr 2021 ein Bachelorstudium im Lehramt erfolgreich beendeten, haben 11 % bis einschließlich zum Wintersemester 2022/2023 kein weiteres Studium begonnen. Demgegenüber war der entsprechende Anteil bei Personen, die ein Bachelorstudium eines Nicht-Lehramtsstudiums abgeschlossen haben, mit 38 % deutlich höher. Mit Blick darauf, dass ein Lehramtsbachelor in der Regel noch nicht für den Start des Vorbereitungsdienstes ausreicht, ist auch der deutlich niedrigere Anteil von bereits nach dem Bachelorabschluss beendeten Hochschulkarrieren bei Lehramtsstudierenden eine relevante Größe. Denn soweit diese Personen ihre hochschulische Laufbahn auch später nicht fortsetzen, stehen sie auch künftig nicht als Lehrkräftenachwuchs zur Verfügung.

4


Fazit und Ausblick

Ein vorherrschendes Thema in der aktuellen bildungspolitischen Diskussion ist der Lehrkräftemangel. Die Ausbildung und Anwerbung von qualifiziertem Lehrkräftenachwuchs in ausreichender Zahl hat daher oberste Priorität auf der bildungspolitischen Agenda. Dies insbesondere, da andere denkbare Gegenmaßnahmen wie eine Verringerung des Teilzeitanteils bei den aktiven Lehrkräften, eine Erhöhung der Klassengrößen

oder ein verstärkter Einsatz von Quereinsteigerinnen und Quereinsteigern sehr umstritten sind. Vor diesem Hintergrund kommt der Hochschulstatistik mit ihren Daten zur hochschulischen Ausbildung des Lehrkräftenachwuchses eine besondere Bedeutung zu. Dabei haben sich die Analysemöglichkeiten der Hochschulstatistik durch die Einführung der Studienverlaufsstatistik erheblich erweitert. Sie ermöglicht seit Kurzem die Quantifizierung etwa von Übergängen von einem Lehramtsbachelor- in ein Lehramtsmasterstudium oder zum frühzeitigen Abbruch eines Lehramtsstudiums auf Basis individueller Verlaufsdaten.

Die aktuellen Ergebnisse der Hochschulstatistik zum Lehramtsstudium zeigen seit über 15 Jahren weitgehend stabile absolute Zahlen beim Abschluss eines hochschulischen Lehramtsstudiums, das für das Referendariat qualifiziert. Dem steht ein steigender Bedarf an Lehrkräften gegenüber: Steigende Geburtenzahlen in den Jahren 2016 bis 2021 und die fluchtbedingten Wanderungen seit 2015 führen zum Anstieg der Zahl der Schülerinnen und Schüler (Autor:innengruppe Bildungsberichterstattung, 2024). Die bisher vorliegenden Ergebnisse der Studienverlaufsstatistik geben keine Hinweise auf spezifische Probleme des Lehramtsstudiums, die die Anwerbung des Lehrkräftenachwuchses zusätzlich erschweren könnten. So brechen anteilmäßig weniger Lehramtsstudierende das Studium frühzeitig ab als dies bei Nicht-Lehramtsstudierenden der Fall ist. Zudem überwiegt die Zahl der Studierenden, die aus einem Nicht-Lehramtsstudium ins Lehramtsstudium wechseln, die Zahl der Personen, die umgekehrt in einer frühen Studienphase aus einem Lehramtsstudiengang in einen Nicht-Lehramtsstudiengang wechseln.

Soweit die hochschulische Lehrkräfteausbildung in den Bundesländern im gestuften Bachelor-Master-System erfolgt, lässt sich dafür bisher auf Basis der Studienverlaufsstatistik mit 86 % eine hohe Übergangsquote ins Masterstudium feststellen. Diese übertrifft die Übergangsquote bei den Nicht-Lehramtsstudiengängen (an Universitäten und gleichrangigen Hochschulen) deutlich. Die bisher vorliegenden Daten zeigen zudem, dass der Übergang vom Bachelor- zum Masterstudium bei Lehramtsstudierenden zügiger erfolgt, das Masterstudium nach dem Bachelorabschluss also schneller begonnen wird.

Die bisher vorliegenden Daten der Studienverlaufsstatistik zeigen keine Hinweise auf strukturelle Defizite des Lehramtsstudiums, die zu einer übermäßigen Reduzierung des Lehrkräftepotenzials führen. Dabei ist zu berücksichtigen, dass die Hochschulstatistik ausschließlich die erste und hochschulische Phase der Lehrkräfteausbildung bis zur Erlangung der Qualifikation für das Referendariat als der zweiten Ausbildungsphase abbildet. Über Zahl und Anteil der Absolvierenden eines Lehramtsstudiums, die das Referendariat beginnen, über den Anteil derer, die es erfolgreich abschließen, sowie derer, die anschließend für den Start in den Schuldienst zur Verfügung stehen, liegen keine vergleichbaren und bundesweiten Daten vor. Um das künftige Potenzial qualifizierter Lehrkräfte verlässlich abschätzen zu können, sind jedoch neben den Daten der Hochschulstatistik auch Informationen zum Vorbereitungsdienst sowie zu den Übergängen zwischen den beiden Ausbildungsphasen und in die Berufstätigkeit notwendig. 

LITERATURVERZEICHNIS

Autor:innengruppe Bildungsberichterstattung. *Bildung in Deutschland 2024. Ein indikatorengestützter Bericht mit einer Analyse zu beruflicher Bildung*. 2024. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.bildungsbericht.de

KMK (Kultusministerkonferenz). *Gegenseitige Anerkennung von Lehramtsprüfungen und Lehramtsbefähigungen, Beschluss der Kultusministerkonferenz*. 1999. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.kmk.org

Neumann, Maximilian/Brings, Stefan. *Die neue Studienverlaufsstatistik: Hintergründe, Aufbau, Methodik und erste Ergebnisse*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2021, Seite 67 ff.

SWK (Ständige Wissenschaftliche Kommission der Kultusministerkonferenz). *Lehrkräftegewinnung und Lehrkräftebildung für einen hochwertigen Unterricht. Gutachten der Ständigen Wissenschaftlichen Kommission der Kultusministerkonferenz*. 2023. DOI: [10.25656/01:28059](https://doi.org/10.25656/01:28059)

Statistisches Bundesamt. *Statistischer Bericht – Statistik des Studienverlaufs – 2023*. 2024. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.destatis.de

RECHTSGRUNDLAGEN

Gesetz über die Statistik für das Hochschulwesen sowie für die Berufsakademien (Hochschulstatistikgesetz – HStatG) vom 2. November 1990 (BGBl. I Seite 2414), zuletzt geändert durch Artikel 3 des Gesetzes vom 7. Dezember 2016 (BGBl. I Seite 2826).

FAKTEN ZUR WALDFLÄCHE IN DEUTSCHLAND – UNTERSCHIEDLICHE DATEN FÜR UNTERSCHIEDLICHE FRAGESTELLUNGEN

Marius Belling

📌 **Schlüsselwörter:** Flächenstatistik – Waldgesamtrechnung – Ökosystemrechnungen – Umweltökonomische Gesamtrechnungen – SEEA

ZUSAMMENFASSUNG

Welche Daten stellt das Statistische Bundesamt zur wirtschaftlichen und ökologischen Bedeutung des Waldes zur Verfügung? Warum sind die Ergebnisse der Flächenerhebungen nach Art der tatsächlichen Nutzung, der Waldgesamtrechnung und der Ökosystemrechnungen bezogen auf den Wald nicht immer identisch, obwohl es sich um einen vermeintlich gleichen Sachverhalt handelt? Der Aufsatz beschreibt die Berichtssysteme hinsichtlich des Monitorings der Waldfläche und vergleicht deren Ziele, rechtliche Grundlagen, Methoden und Definitionen.

📌 **Keywords:** Land use statistics by type of actual use – Forest Accounts – Ecosystem Accounts – Environmental Economic Accounting – SEEA

ABSTRACT

What data does the Federal Statistical Office provide on the economic and ecological importance of forests? Why do the results of the surveys of areas by type of actual use, the forest accounts and ecosystem accounts sometimes differ in terms of the forest area, despite referring to what appears to be the same subject matter? This article describes the reporting systems in place for monitoring the forest area and compares their objectives, legal basis, methods and definitions.



Marius Belling

studierte Geographie und Management natürlicher Ressourcen. Er ist als wissenschaftlicher Mitarbeiter im Referat „Monetäre Umweltökonomische Gesamtrechnungen, Nachhaltigkeitsberichterstattung“ des Statistischen Bundesamtes für die Ökosystemrechnungen und die Waldgesamtrechnung zuständig.

1

Einleitung

Etwa ein Drittel der Landfläche Deutschlands ist mit Wald bedeckt. Damit ist der Wald sowohl landschaftsprägend als auch ein wichtiger ökonomischer Faktor. Darüber hinaus erbringt der Wald viele essenzielle Ökosystemleistungen. Die Veränderungen der Waldfläche zu erfassen und im Zeitverlauf zu beobachten, ist daher eine wichtige Grundlage für politische Maßnahmen. Hierzu sind im Laufe der Zeit verschiedene Berichtssysteme entstanden. Die amtliche Statistik in Deutschland weist die Waldfläche in der Flächenerhebung nach tatsächlicher Art der Nutzung, in der Waldgesamtrechnung und in den Ökosystemrechnungen aus. Das Monitoring der Waldfläche ist relevant, um beispielsweise drohende Schädigungen von Ökosystemen oder Veränderungen durch Renaturierungsmaßnahmen zu beobachten – schließlich haben die Vereinten Nationen (UN) die 2020er-Jahre zur [UN-Dekade für die Wiederherstellung von Ökosystemen](#) erklärt. National können die Daten der drei vom Statistischen Bundesamt veröffentlichten Erhebungen vielfältig genutzt werden, beispielsweise im Aktionsprogramm „Natürlicher Klimaschutz“ (BMUV, 2023). Dieses Programm ist Teil der sogenannten Nature Restoration Regulation¹ der Europäischen Union (EU) und auf nationaler Ebene mit dem Bundes-Klimaschutzgesetz verwoben, welches darauf abzielt, bis 2045 Klimaneutralität in Deutschland zu erreichen.

Dieser Artikel vergleicht die Berichtssysteme der Flächenerhebung nach tatsächlicher Art der Nutzung, der Waldgesamtrechnung und der Ökosystemrechnungen hinsichtlich des Monitorings der Waldfläche und erklärt deren Unterschiede. Die Abweichungen zwischen den aus den unterschiedlichen Berichtssystemen veröffentlichten Waldflächen in Deutschland sind im Wesentlichen auf drei Faktoren zurückzuführen, die sich teilweise gegenseitig bedingen:

- › unterschiedliche Definitionen des Begriffs Waldfläche aufgrund unterschiedlicher Gesetzesgrundlagen, Handbücher und Durchführungsverordnungen,
- › die damit einhergehenden Anforderungen an die Datengrundlagen, um zum Beispiel Ergebnisse für

Gemeinden, Länder und den Bund zu erhalten oder einen Detailgrad zu bedienen, der Feuerschneisen als explizite Fläche im Wald berücksichtigt, sowie

- › die unterschiedlich intendierte Verwendung der Ergebnisse für ökologische und/oder ökonomische Fragestellungen sowie für die Belange der Raumordnung.

Die nächsten drei Kapitel beschreiben detailliert jeweils eines der Berichtssysteme zur Flächenerhebung. Kapitel 5 stellt die Unterschiede in den Berichtssystemen heraus. Eine abschließende Zusammenfassung und ein Ausblick folgen in Kapitel 6.

2

Flächenerhebungen nach Art der tatsächlichen Nutzung

Die Flächenerhebungen nach Art der tatsächlichen Nutzung – kurz Flächenstatistik – bilden nicht den Wald im Speziellen ab, vielmehr stellen sie eine Gesamtübersicht über die Nutzung der gesamten Fläche in Deutschland dar. Damit liefert die Flächenstatistik wichtige Ergebnisse für raumordnungsrelevante Entscheidungen auf der jeweiligen administrativen Ebene, zum Beispiel für die Nachhaltigkeitsstrategie der Bundesregierung.

Die Flächenstatistik basiert auf den Angaben aus den Liegenschaftskatastern. Im Vergleich mit den Berichtssystemen der Waldgesamt- und Ökosystemrechnungen ist die Flächenstatistik die Quelle mit der längsten verfügbaren Zeitreihe an Daten. Die Flächenstatistik erfasste ein Flurstück ursprünglich vollständig als Wald, wenn dessen Hauptnutzung Wald war. Das amtliche Liegenschaftskatasterinformationssystem (ALKIS) löste im Jahr 2016 das Automatisierte Liegenschaftsbuch ab. Damit ist nicht mehr das Flurstück Maßstab zur Beurteilung der Flächennutzung, sondern die konkrete Geometrie einer Fläche (Statistische Ämter des Bundes und der Länder, 2019).

Die Flächenstatistik weist bis zum Berichtsjahr 2022 die Fläche Deutschlands in insgesamt 28 Nutzungsarten aus. Ab dem Berichtsjahr 2023 nutzt die Flächenstatistik eine erweiterte Klassifikation und unterscheidet 54 Nutzungsarten (Statistisches Bundesamt, 2024). Dabei

¹ Verordnung (EU) 2024/1991.

steht das Erhebungsmerkmal in der Belegung der Bodenflächen nach Art der tatsächlichen Nutzung² im Vordergrund und nicht die Ausweisung der Fläche nach planerischer oder rechtlicher Zuordnung. Die Flächenstatistik weist von der Gemeindeebene bis zur Bundesebene Flächennutzungsarten mit den dazugehörigen Flächengrößen beziehungsweise -anteilen aus. Die Periodizität der Flächenstatistik ist jährlich.

Die rechtlichen Grundlagen für die Flächenstatistik sind das Bundesstatistikgesetz und das Agrarstatistikgesetz in der jeweils geltenden Fassung. Dabei bildet das Bundesstatistikgesetz den rechtlichen Rahmen und damit die Rechtsgrundlage, auf die sich das Agrarstatistikgesetz bezieht. Die §§ 3 und 4 Agrarstatistikgesetz enthalten spezifische Regelungen zu den Erhebungseinheiten, der Erhebungsart, der Periodizität, dem Berichtszeitpunkt und den Erhebungsmerkmalen der Flächenerhebung.

Die Datenquelle der Flächenstatistik ist ALKIS, re-klassifiziert durch die Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV). Diese Datenquelle ist gesetzlich vorgeschrieben und fokussiert auf die Flächennutzung nach Art der tatsächlichen Nutzung. Die Nutzungsart Wald unterteilt sich in der Flächenstatistik in die Untergliederungen Laubholz, Nadelholz sowie Laub- und Nadelholz. Wald ist dabei definiert als „eine Fläche, die mit Forstpflanzen (Waldbäume und Waldsträucher) bestockt ist“

² Mischklassifikation aus Landbedeckung und Landnutzung (Arnold und andere, 2020).

Übersicht 1

Nutzungsartengliederung der Flächennutzung nach Art der tatsächlichen Nutzung für die Nutzungsartengruppen Wald und Gehölz

Nutzungsartenbereich	Nutzungsartengruppe	Nutzungsart	Erste Untergliederung
30000 Vegetation	32000 Wald	32100 Laubholz	
		32200 Nadelholz	
		32300 Laub- und Nadelholz	32310 Laubwald mit Nadelholz
			32320 Nadelwald mit Laubholz
			323_R Laub- und Nadelholz – nicht weiter untergliedert
		32_R Wald – nicht weiter untergliedert ¹	
	33000 Gehölz	33_R Gehölz – nicht weiter untergliedert	

Quelle: Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland (AdV) – Nutzungsartenkatalog, erweitert um statistische Restpositionen (in hierarchischer Struktur) für die Nutzungsartengruppen Wald und Gehölz

¹ Restposition.

(AdV, 2022). Flächen wie Feuerschneisen oder Waldwege, auf denen sich keine Forstpflanzen befinden, zählen nicht zur Nutzungsart Wald. Gehölz ist definiert als Fläche, „die mit einzelnen Bäumen, Baumgruppen, Büschen, Hecken und Sträuchern bestockt ist“ (AdV, 2022). [↗ Übersicht 1](#)

3

Waldgesamtrechnung

Als Teil der Umweltökonomischen Gesamtrechnungen (UGR) hat die Waldgesamtrechnung die Aufgabe, die physischen Bestände und Stoffströme im Wald zu erfassen und in Teilen monetär zu bewerten. Dabei liegt der Fokus entsprechend auf Wald- und Forstwirtschaft. Diese Prämisse bestimmt Methode und Datenauswahl. Die Waldgesamtrechnung hat als Bundesstatistik nicht das Ziel, räumlich hochaufgelöste Ergebnisse zu erzielen, sondern ein bundesweites Gesamtbild der Wald- und Forstwirtschaft im Zeitverlauf zu präsentieren. Die Waldgesamtrechnung enthält Tabellen mit physischen und monetären Einheiten. So finden sich für die physischen Tabellen der Waldflächenbilanz und Holzvorratsbilanz auch jeweils monetäre Tabellen. Darüber hinaus umfasst die Waldgesamtrechnung Informationen zur Produktion der Forstwirtschaft nach Art der Verwendung, zur Gesamtrechnung für Forstwirtschaft und Holzerte, eine Aufkommens- und Holzverwendungsbilanz,

gegliedert nach Produkteinheiten, sowie physische Kohlenstoffbilanzen. Damit bildet die Waldgesamtrechnung wichtige Kernelemente des Waldes aus vornehmlich ökonomischer Sicht ab. Die Periodizität der Waldgesamtrechnung ist jährlich.

Die Tabellen und Methoden der heutigen Waldgesamtrechnung sind Teil der Umweltökonomischen Gesamtrechnungen und deren Weiterentwicklung. Der Ursprung der Waldgesamtrechnung liegt in den 1990er-Jahren mit einer Initiative des Statistischen Amtes der Europäischen Union (Eurostat), sogenannte Umweltsatellitenkonten zu erstellen. Diese sollten die Volkswirtschaftlichen Gesamtrechnungen (VGR) um relevante Umweltaspekte ergänzen. Teil dieser Initiative waren bis zum Jahr 2006 die Volkswirtschaftlichen Gesamtrechnungen für die Forstwirtschaft (EAF). Ab dem Jahr 2006 wurden diese durch den Tabellensatz „European Framework for Integrated Environmental and Economic Accounting for Forests“ (IEEAF) abgelöst. Dieser enthält insgesamt 19 Tabellen zu physischen Eigenschaften des Waldes und dessen ökonomischer Wertschöpfung. Ab dem Jahr 2014 wurde der IEEAF-Tabellensatz umgestellt und die heutige Waldgesamtrechnung entstand. Die rechtliche Grundlage der Waldgesamtrechnung ist die Erweiterung der EU-Verordnung Nr. 691/2011 durch Anhang VII.

Die Waldgesamtrechnung fokussiert sich auf den Bereich Wald- und Forstwirtschaft, der zwischen Waldfläche verfügbar/nicht verfügbar für die Rohholzproduktion unterscheidet. Generell sind für die Rohholzproduktion verfügbare Waldflächen jene Flächen, in denen ökologische, soziale oder wirtschaftliche Einschränkungen keine wesentlichen Auswirkungen auf die aktuelle oder potenzielle Holzversorgung haben. Bei nicht für die Rohholzproduktion verfügbaren Waldflächen verhindern diese Beschränkungen eine nennenswerte Holzversorgung. Beschränkungen bestehen zum einen durch gesetzliche Vorschriften, wie einer Deklaration als Schutzgebiet, zum anderen durch Entscheidungen der Besitzenden selbst oder durch andere Gründe, wie einer unwirtschaftlichen räumlichen Lage (Rosenkranz und andere, 2018). Bestimmte Flächenkategorien, die theoretisch für die Rohholzproduktion verfügbar sind, werden innerhalb der Waldgesamtrechnung ausgeschlossen. Dies gilt zum Beispiel für Weihnachtsbaumplantagen, sofern diese auf Agrarflächen und nicht auf

Forstflächen bestehen. Es ist nicht das Ziel der Waldgesamtrechnung, alle Flächen ohne Bestockung³ auszuschließen, sondern solche Gebiete gesondert auszuweisen (Eurostat, 2024). Der ökonomische Hintergrund der Waldgesamtrechnung führt dazu, dass auch bestimmte Flächen ohne Bestockung in den Anwendungsbereich „Verfügbar für Rohholzproduktion“ fallen. So sind Waldwege und Straßen hier im Gegensatz zur Flächenstatistik als verfügbar für die Rohholzproduktion klassifiziert, da diese unter anderem für den Transport von Holz und somit wirtschaftlich relevant sind. Dagegen klassifiziert die Waldgesamtrechnung Waldflächen in Schutzgebieten als nicht verfügbar für die Rohholzproduktion, es sei denn, die Schutzkategorie lässt eine ökonomische Bewirtschaftung zu. Flächen, die mit Bäumen bestockt sind, jedoch dem Agrarsektor zugewiesen werden, sind von den Kategorien der Rohholzproduktion ausgeschlossen. Sie sind somit nicht in den Daten der Waldgesamtrechnung enthalten. Ebenfalls ausgeschlossen sind Hecken. [↗ Übersicht 2](#) stellt einen Auszug der Tabelle innerhalb des Waldkontenhandbuchs von Eurostat dar und fokussiert auf den nationalen Kontext.

Die Waldgesamtrechnung ermittelt die physische Waldflächenbilanz basierend auf den Ergebnissen der Bundeswaldinventur. Gesetzliche Grundlage der Bundeswaldinventur ist das Bundeswaldgesetz, insbesondere § 41. Die Bundeswaldinventur wird in einem zehnjährigen Turnus erstellt und ist eine terrestrische Großrauminventur auf Stichprobenbasis. Über ganz Deutschland sind dabei Probepunkte in einem systematischen Raster verteilt. Dazu entsenden die Bundesländer Expertinnen und Experten in sogenannten Inventurtrupps zu allen Probepunkten. Die ermittelten Daten werden anschließend an das Thünen-Institut für Waldökosysteme übermittelt und dort ausgewertet. Die jüngste Bundeswaldinventur mit dem Referenzjahr 2022 ist im Oktober 2024 veröffentlicht worden (BMEL, 2024).

Für die Waldgesamtrechnung werden jährliche Änderungen zu Erstaufforstungen und Waldumwandlungen hauptsächlich der Treibhausgas-Berichterstattung, genauer gesagt dem Sektor „Landnutzung, Landnutzungsänderung und Forstwirtschaft“, entnommen. Die Daten dieses Sektors der Treibhausgas-Berichterstattung ermitteln die Thünen-Institute für Agrarklimaschutz, Wald-

³ Bestockung meint hier die Bedeckung der Fläche durch Wald- und/oder Forstpflanzen.

Fakten zur Waldfläche in Deutschland – unterschiedliche Daten für unterschiedliche Fragestellungen

Übersicht 2

Einteilung der Flächenkategorien in der Waldgesamtrechnung

Flächenkategorie	Verfügbar für Rohholzproduktion	Nicht verfügbar für Rohholzproduktion	Ausgeschlossen
Waldwege und Straßen	X		
Feuerschneisen	X		
Lichtungen	X		
Windschutzbestand	X		
Bestand zum Schutz vor Erosion oder Überschwemmung	X		
Baumkorridore	X		
Aufforstungen	X		
Kahlschlagflächen und gerodete Flächen	X		
Naturschutzrechtlich geschützte Flächen	X (abhängig vom Produktionskontext)	X	
Sukzessionsbestände auf Ackerflächen	X		
Weihnachtsbaumplantagen	X (sofern Fläche forstlich genutzt)		X (sofern Fläche agrarisch genutzt)
Flächen außerhalb gesetzlich ausgewiesener Waldflächen, die der Definition von „Wald“ entsprechen	X (es sei denn die Fläche fällt unter die Definition von „Nicht verfügbar für Rohholzproduktion“)		
Bestände in agrarischen Produktionssystemen (zum Beispiel Obstbäume)			X
Baumschulen	X (sofern Fläche forstlich genutzt)		X (sofern Fläche agrarisch genutzt)
Samenplantagen			X
Hecken			X
Haine			X

Quelle: Eurostat, [European forest accounts handbook](#), Ausgabe 2024

Ökosysteme und Holzforschung. Die Treibhausgas-Berichterstattung nutzt dabei ein Stichprobennetz, das alle 100 m die Landnutzung im Probenpunkt ermittelt. Dazu werden verschiedene Datenquellen wie das Basis-DLM (Basis-Digitales Landschaftsmodell), welches die Grundlage des Amtlich Topographisch-Kartographischen Informationssystems (ATKIS) bildet, das [Landbedeckungsmodell für Deutschland](#) oder OpenStreetMap zur Ableitung von Straßen genutzt. Das Ziel dieses flexiblen Erfassungssystems ist es, aus der Menge an Informationen die zuverlässigsten Landnutzungsinformationen zu erfassen, zuzuweisen und mögliche Unsicherheiten und Fehlerquellen zu eliminieren (Umweltbundesamt, 2023).

Die Kategorien verfügbar und nicht verfügbar für die Rohholzproduktion werden innerhalb der Waldgesamtrechnung tabellarisch weiter unterteilt. Die Unterteilung enthält jeweils Anfangs- und Endbestand sowie Angaben zu Änderungen aufgrund von Bewirtschaftungsmaßnahmen, wie Erstaufforstung, Waldumwandlung/Rodung, und sonstigen Änderungen. ➡ Tabelle 1 zeigt die Flächenänderungen der physischen Waldflächen der Waldgesamtrechnung im Zeitverlauf von 2018 bis 2021. Gegenwärtig werden die Daten der Waldgesamtrechnung für das Berichtsjahr 2022 entsprechend der neuen Daten der Bundeswaldinventur aktualisiert.

Tabelle 1
Physische Waldflächenbilanz bezogen auf die gesamte deutsche Waldfläche

	2018	2019	2020	2021 ¹
	1 000 ha			
	Insgesamt			
Anfangsbestand	11 443	11 451	11 460	11 468
Änderungen aufgrund von Bewirtschaftungsmaßnahmen				
Erstaufforstung	14	14	14	13
Waldumwandlung/Rodung	– 6	– 6	– 6	– 6
Sonstige Änderungen
Endbestand	11 451	11 460	11 468	11 475
	verfügbar für Rohholzproduktion			
Anfangsbestand	10 769	10 778	10 771	10 716
Änderungen aufgrund von Bewirtschaftungsmaßnahmen				
Erstaufforstung	14	14	14	13
Waldumwandlung/Rodung	– 6	– 6	– 6	– 6
Sonstige Änderungen	– 0	– 15	– 63	– 9
Endbestand	10 778	10 771	10 716	10 714
	nicht verfügbar für Rohholzproduktion			
Anfangsbestand	674	674	689	752
Änderungen aufgrund von Bewirtschaftungsmaßnahmen				
Erstaufforstung
Waldumwandlung/Rodung
Sonstige Änderungen	0	15	63	9
Endbestand	674	689	752	761

1 Vorläufiges Ergebnis.

4

Ökosystemrechnungen

Lange Zeit hat das Statistische Bundesamt den Umweltzustand nur über die Entwicklung der Landnutzung im Rahmen der auf Katasterdaten basierenden Flächenstatistik dargestellt. Mit dem Aufbau der Ökosystemrechnungen und der rechtlichen Legitimation durch Anhang IX der EU-Verordnung 691/2011 liegen seit dem Berichtsjahr 2015 detaillierte Informationen zur Fläche der Ökosysteme in Deutschland sowie deren zeitlicher Veränderung vor. Die Zustandsbilanz erweitert die Ökosystemrechnungen. Flächen- und Zustandsbilanz sind die Grundlage für die künftige gesetzlich vorgeschriebene Berechnung bestimmter Ökosystemleistungen wie die Luftfilterung von Schadstoffen durch die Vegetation, die Kühlung durch Vegetation an Hitzetagen in urbanen Räumen oder die Sequestrierung und Speicherung von

Kohlenstoff. Die Ökosystemrechnungen orientieren sich in ihrer Ausgestaltung hauptsächlich an ökologischen Charakteristiken. Dies bedeutet, dass nicht ausschließlich die Nutzung der Flächen, sondern auch deren Bedeckung und somit systematisch die ökologische Relevanz zur Erbringung von Ökosystemleistungen im Vordergrund stehen. Definiert ist ein Ökosystem als ein dynamischer Komplex von Gemeinschaften aus Pflanzen, Tieren und Mikroorganismen sowie deren nicht lebender Umwelt, die als funktionelle Einheit in Wechselwirkung stehen.⁴ Das bedeutet im Einzelfall, dass aufgrund des holistischen Ansatzes beispielsweise Waldwege und Bäche als Teil des Waldökosystems angesehen werden und die Fläche nicht weiter unterteilt wird.

Zur Erstellung der Flächenbilanz wird eine detaillierte und räumlich hochaufgelöste Datengrundlage ange-

4 Definition gemäß dem Übereinkommen über die Biologische Vielfalt: www.bfn.de

legt, die wiederum jede Fläche Deutschlands durch eine jeweils spezifische Ökosystemklasse charakterisiert. Dies bildet die Voraussetzung, um Ökosystemflächen zur Flächenbilanz zu aggregieren (Statistisches Bundesamt, 2021a). Die theoretische Basis der Flächenbilanz ist die nationale Ökosystemklassifikation (Statistisches Bundesamt, 2021b), welche hierarchisch strukturiert ist. Die Klassifikation unterteilt Flächen nach ökologischen und strukturellen Merkmalen und umfasst 74 Klassen, 21 Gruppen, 6 Abteilungen und 2 Abschnitte (terrestrisch und marin)⁵. Die Ökosystemabteilung Wälder und Gehölz enthält insgesamt 13 Ökosystemklassen, gegliedert in die Ökosystemgruppen Laubwälder, Nadelwälder, Mischwälder und Wald-Strauch-Übergangsflächen. Für jede Klasse sind eindeutige und leicht nachvollziehbare Zuteilungskriterien definiert, die diese Eigenschaften widerspiegeln. Die Datengrundlage, die eine vollständige Klassifizierung aller Flächen des nationalen

Territoriums Deutschlands ermöglicht, besteht aus einer Vielzahl von räumlich hochaufgelösten Datensätzen, wie dem aus Sentinel 2 abgeleiteten Datensatz zu Uferflächen zur Ableitung von Auenwäldern mit einer räumlichen Auflösung von 10 m je Pixel⁶. Die Datensätze werden mit der Programmiersprache Python und mittels eines Geographischen Informationssystems (GIS) aufbereitet und verarbeitet (Statistisches Bundesamt, 2021a).

Die Flächenbilanz wird in einem dreijährlichen Turnus erstellt. Damit können flächenhafte Veränderungen innerhalb der Ökosysteme zwischen den Berechnungsjahren nachvollzogen werden. [Tabelle 2](#) zeigt die Flächenbilanz auf Ebene der Ökosystemgruppen innerhalb der Abteilung Wälder und Gehölz für die Jahre 2018 und 2021 sowie die Veränderung.

5 Die Ökosystemabschnitte C (subterran) und D (mittlere und obere Biosphäre) werden nach dem internationalen Rahmenwerk SEEA EA (System of Environmental Economic Accounting – Ecosystem Accounting) derzeit nicht erfasst, sind aber gegebenenfalls für künftige Erweiterungen angelegt.

6 Sonderauswertung des Bundesamtes für Kartographie und Geodäsie.

Tabelle 2

Flächenbilanz der Ökosysteme auf Gruppenebene für die Abteilung Wälder und Gehölz

	Fläche 2018	Fläche 2021	Nettoveränderung zwischen 2018 und 2021	Veränderung des Bestands 2018 zu 2021
	ha			%
Laubwälder	3 731 353	3 729 093	– 1 403	0
Montane und Subalpine Laubwälder	194 274	193 560	– 690	0
Moorwälder mit Laubbäumen	29 151	30 440	+ 1 280	+ 4
Auenwälder	115 380	107 751	– 7 406	– 6
Sonstige Laubwälder	3 392 549	3 397 343	+ 5 414	0
Nadelwälder	5 505 615	5 225 057	– 280 650	– 5
Montane und Subalpine Nadelwälder	872 800	839 283	– 33 516	– 4
Moorwälder mit Nadelbäumen	7 145	7 175	+ 29	0
Sonstige Nadelwälder	4 625 671	4 378 598	– 247 163	– 5
Mischwälder	1 614 352	1 632 043	+ 17 724	+ 1
Montane und Subalpine Mischwälder	246 998	248 369	+ 1 371	+ 1
Moorwälder mit Laub- und Nadelbäumen	10 052	10 108	+ 54	+ 1
Sonstige Mischwälder	1 357 302	1 373 567	+ 16 299	+ 1
Wald-Strauch-Übergangsflächen	718 997	978 616	+ 260 309	+ 36
Aufforstung	411 583	396 455	– 15 138	– 4
Lichtungen und Naturverjüngungsflächen	10 280	247 498	+ 237 154	+ 2 307
Sonstige Flächen mit Büschen und Sträuchern	297 134	334 663	+ 38 293	+ 13

Quelle: Statistischer Bericht der Flächenbilanz – Berichtszeitraum 2018-2021. www.destatis.de

Ziel der Ausweisung ist die Erfassung von realen Veränderungen. Daher werden Veränderungen unter einer Mindestgröße (bezogen auf Fläche und Kompaktheit) nicht explizit ausgewiesen und als technische Ungenauigkeit der Geodatenanalyse interpretiert.

5

Unterschiede der Berichtssysteme

Die drei beschriebenen Berichtssysteme bedienen unterschiedliche Blickwinkel auf den Wald. Durch die jeweilige Rechtsgrundlage bedarf jedes Berichtssystem einer individuellen Datengrundlage und eines gesonderten methodischen Vorgehens. Folgend werden die Unterschiede der Berichtssysteme anhand der Waldfläche noch einmal zusammengefasst.

Grundsätzlich werden die Waldgesamtrechnung und die Ökosystemrechnungen als Rechensysteme bezeichnet. Der Fokus liegt hierbei auf einer kohärenten und integrierten Darstellung der Daten innerhalb der Umweltökonomischen Gesamtrechnungen. Die Umweltökonomischen Gesamtrechnungen ermöglichen so die Analyse von Wechselwirkungen zwischen Menschen und Umwelt. Die Flächenstatistik wiederum ist kein Rechensystem, sondern eine Statistik mit Erhebung, Aufbereitung und Veröffentlichung nach den beschriebenen gesetzlichen Vorgaben.

Flächenstatistik, Waldgesamtrechnung und Ökosystemrechnungen definieren Wald jeweils unterschiedlich. In der Waldgesamtrechnung und der Flächenstatistik steht die Nutzung beziehungsweise die Art der tatsächlichen

Nutzung für die Berechnung der Waldfläche im Vordergrund. In den Ökosystemrechnungen wird der Wald als Ökosystem zur Erbringung spezifischer Ökosystemleistungen verstanden.

Die Flächenstatistik verwendet die Definition des ALKIS für Waldflächen, in der die Bestockung ausschlaggebend ist. Dagegen definiert die Waldgesamtrechnung auch Flächen ohne Bestockung als Waldflächen. Während bei den Ökosystemrechnungen auch Gehölz dem Wald zugehörig ist, trennt die Flächenstatistik zwischen Wald und Gehölz als jeweilige Nutzungsarten.

Neben den Definitionen unterscheiden sich auch der Anwendungsbereich, die Periodizität, die Datengrundlagen sowie die administrativen Veröffentlichungsebenen. ➔ **Übersicht 3**

Durch die Unterteilung der Waldfläche in verfügbar beziehungsweise nicht verfügbar für die Rohholzproduktion legt die Waldgesamtrechnung den Schwerpunkt auf die ökonomische Berichterstattung zu Wald- und Forstwirtschaft. Dies führt zu inhaltlichen Unterschieden zu den Ökosystemrechnungen und zur Flächenstatistik. Denn für die Ökosystemrechnungen liegt der Fokus auf ökologischen Prinzipien und dem Potenzial einer Fläche, bestimmte Ökosystemleistungen zu erbringen. Für die Flächenstatistik liegt der Anwendungsbereich ausschließlich auf der tatsächlichen Flächennutzung.

Übersicht 3

Übersicht der Statistiken zu Waldflächen in Deutschland

Statistik/Rechensystem	Anwendungsbereich	Periodizität	Administrative Einheit	Datengrundlage
Physische Waldflächenbilanz der Waldgesamtrechnung	Fokus liegt auf Forstwirtschaft mit Aufteilung in Flächen verfügbar oder nicht verfügbar für die Rohholzproduktion und der monetären Bewertung der Waldfläche	jährlich	Bund	Bundeswaldinventur, Treibhausgasberichterstattung, Fernerkundung
Waldfläche innerhalb der Flächenbilanz der Ökosystemrechnungen	Fokus liegt auf ökologischen Prinzipien und dem Potenzial, bestimmte Ökosystemleistungen zu erbringen	alle drei Jahre	Gemeinden, Länder, Bund	Landbedeckungsmodell für Deutschland, Tatsächliche Überflutungsflächen nach Delineation of Riparian Zones – Copernicus, Digitales Geländemodell, Klassifikationsprozess der Flächenbilanz
Wald innerhalb der Flächenstatistik	Fokus liegt auf der tatsächlichen Flächennutzung	jährlich	Gemeinden, Länder, Bund	Amtliches Liegenschaftskatasterinformationssystem ALKIS

Fakten zur Waldfläche in Deutschland – unterschiedliche Daten für unterschiedliche Fragestellungen

Weitere Unterschiede ergeben sich aus der gesetzlich vorgeschriebenen Periodizität der Durchführung sowie der Veröffentlichung der Ergebnisse auf unterschiedlichen administrativen Ebenen. So veröffentlichen die Ökosystemrechnungen in einem dreijährlichen Turnus Daten zur Waldfläche, Flächenstatistik und Waldgesamtrechnung hingegen jährlich. Für die Waldgesamtrechnung erfolgt die Veröffentlichung der Daten auf Bundesebene, für Flächenstatistik und Ökosystemrechnungen auf Gemeinde-, Länder- und Bundesebene.

Die genutzte Datengrundlage unterscheidet sich ebenfalls zwischen den Berichtssystemen. Die jeweilige Datengrundlage muss den Bedingungen hinsichtlich des Anwendungsbereiches, der Periodizität und der räumlichen Veröffentlichungsebene entsprechen. Um bestimmte Flächen innerhalb der Waldgesamtrechnung zur Ermittlung der Waldfläche zu berücksichtigen oder auszuschließen, bedarf es einer Datengrundlage, die beispielsweise zwischen Waldwegen und Straßen, Weihnachtsbaumplantagen oder Feuerschneisen unterscheidet. Einen solchen Detailgrad bietet nur die Bundeswaldinventur. Die Flächenstatistik stützt sich ausschließlich auf die tatsächliche Art der Flächennutzung und ermöglicht daher keine derartige Flächenunterteilung. Der holistische Ansatz der Ökosystemrechnungen mit Fokus auf das Ökosystem an sich ermöglicht weder, die tatsächliche Flächennutzung wie in der Flächenstatistik festzustellen, noch, die Flächen gemäß der Waldgesamtrechnung zu unterteilen. Aktuell existiert keine Datengrundlage, die alle drei genannten Berichtssysteme vollständig „aus einer Hand“ bedienen kann.

↪ **Tabelle 3** stellt die Unterschiede in den Veröffentlichungsergebnissen dar. Auf der administrativen Ebene des Bundes und bedingt durch die Periodizität der Berichtssysteme ist ein Vergleich für die Jahre 2018 und 2021 möglich.

Die Summe aus den Nutzungsartengruppen Wald und Gehölz der Flächenstatistik für das Jahr 2021 ist rund 3 % geringer als die Waldfläche der Waldgesamtrechnung und rund 4 % kleiner im Vergleich zur Waldfläche in den Ökosystemrechnungen. Da Waldwege und Feuerschneisen (Waldgesamtrechnung) oder kleine Gewässer (Ökosystemrechnungen) in den beiden Berichtssystemen der Waldfläche zugerechnet werden, ist die größere Waldfläche im Vergleich zur Flächenstatistik plausibel. Im Gegensatz zur Flächenstatistik ist der Unterschied in der Waldfläche zwischen der Waldgesamtrechnung und den Ökosystemrechnungen mit einem Unterschied von rund 0,8 % im Jahr 2021 deutlich geringer. Die Abweichung ist durch die unterschiedlichen Datenquellen bedingt. Es liegt keine systematische Unter- beziehungsweise Überschätzung im Vergleich beider Datenquellen vor.

Hierbei ist es wichtig zu erwähnen, dass die Flächenstatistik die Daten des ALKIS nutzt, welches die Vermessungsämter der Länder aktualisieren. Eine Aktualisierung kann anlassbezogen oder periodisch (meist zwischen zwei und drei Jahren) erfolgen. Ab dem Berichtsjahr 2023 kommt ein erweiterter Nutzungsartenkatalog der tatsächlichen Nutzung zum Einsatz. Dieser enthält nun 54 anstatt bisher 28 Nutzungsarten und ermöglicht eine tiefere Untergliederung innerhalb der vier Hauptnutzungsarten Siedlung, Verkehr, Vegetation und Gewässer. Die Waldgesamtrechnung nutzt für die jährliche Fortschreibung der Bundeswaldinventur Daten der Treibhausgas-Berichterstattung, hingegen basieren die Waldflächen der Ökosystemrechnungen auf dem Landbedeckungsmodell für Deutschland. Mit der Veröffentlichung der Bundeswaldinventur für das Referenzjahr 2022 wird deutlich, wie gut die Fortschreibung der vorangegangenen Bundeswaldinventur mittels der Treibhausgas-Berichterstattung funktioniert hat. Die Differenz zu den Waldflächen der Ökosystemrechnungen sowie zur Flächenstatistik kann sich durch eine Revision der Waldgesamtrechnung noch einmal verändern.

Tabelle 3

Ergebnisse der Waldflächenberechnung von Flächenstatistik, Waldgesamtrechnung und Ökosystemrechnungen auf Bundesebene

	Flächenstatistik			Waldgesamt- rechnung	Ökosystem- rechnungen	Waldgesamt- rechnung zu Flächenstatistik Wald und Gehölz	Ökosystem- rechnungen zu Flächenstatistik Wald und Gehölz	Ökosystem- rechnungen zu Waldgesamt- rechnung
	Wald	Gehölz	Wald und Gehölz					
	ha							
2018	10 654 595	412 142	11 066 737	11 451 388	11 570 318	+ 3,48	+ 4,55	+ 1,04
2021	10 669 894	440 797	11 110 691	11 474 874	11 564 808	+ 3,28	+ 4,09	+ 0,78

6

Zusammenfassung und Ausblick

Der Wald ist vielseitig und aus verschiedenen Blickwinkeln zu betrachten – das zeigt der Vergleich zwischen den Berichtssystemen in diesem Artikel. So erklärt sich, dass für einen auf den ersten Blick vermeintlich gleichen Sachverhalt, hier: die Waldfläche, unterschiedliche Werte existieren und keiner davon eine falsche Aussage trifft. Oft bestimmt die (zweckgebundene) rechtliche Grundlage, aus welchem Blickwinkel der Wald betrachtet wird. Weitere Gründe für die Unterschiede zwischen den Berichtssystemen sind die teils gesetzlich vorgeschriebenen Datenquellen sowie der Anwendungsbereich mit ökologischem, ökonomischem oder raumordnungsrelevantem Schwerpunkt. Auch die Art der Datenerhebung, repräsentative Stichprobe für die Waldgesamtrechnung im Vergleich zu Fernerkundungsdaten in den Ökosystemrechnungen, führt zu Unterschieden. Dabei besteht die Notwendigkeit, für die jeweiligen Berichtssysteme passende Datenquellen zu nutzen, um den jeweiligen Anforderungen wie Periodizität oder räumlicher Veröffentlichungsebene gerecht zu werden. Seitens der Flächenstatistik ist die Datenquelle gesetzlich bestimmt.


In der Konsequenz stehen für den gleichen Sachverhalt unterschiedliche Zahlen zur Verfügung. Ohne Kenntnis der unterschiedlichen Methoden und Datenquellen führt dies zu Unsicherheiten bezüglich der Interpretation der Ergebnisse des jeweiligen Berichtssystems. Um dem zu begegnen und um eine hohe Transparenz in der Beschreibung der Methoden und Ergebnisse zu erreichen, ergreift die amtliche Statistik unterschiedliche Maßnahmen. Dazu zählen die Veröffentlichung von Qualitätsberichten und die Katalogisierung von Revisionen.

Die Entwicklung der Berichtssysteme im Laufe der Zeit verdeutlicht, dass unterschiedliche Sichtweisen auf den Wald notwendig sind, um relevante Daten für politische Entscheidungen bereitzustellen. So erfuhr die Flächenstatistik seit dem Inkrafttreten des Agrarstatistikgesetzes 1989 zahlreiche Erweiterungen und Novellierungen, um dem wachsenden Informationsbedarf gerecht zu werden. Wurde einst die Flächennutzung über das Flurstück bestimmt, sind heute die tatsächlichen Geometrien aus ALKIS Datengrundlage, beispielsweise um den Indikator [Anstieg der Siedlungs- und Verkehrsfläche](#)

für die Nachhaltigkeitsstrategie der Bundesregierung berechnen zu können.

Die Datenreihe der Waldgesamtrechnung beginnt mit dem Referenzjahr 2014, basierend auf den Daten der Bundeswaldinventur, die auf Grundlage des Bundeswaldgesetzes durchgeführt wird. Mit der Erweiterung der EU-Verordnung 691/2011 muss Deutschland künftig regelmäßig die Daten der Waldgesamtrechnung an die EU übermitteln. Der Blick auf die Waldfläche unter ökonomischen Gesichtspunkten zeigt das politische Interesse an Daten zur wirtschaftlichen Nutzung von Wäldern. Neben der Waldgesamtrechnung sind auch die Ökosystemrechnungen in der gleichen EU-Verordnung geregelt, was die zunehmende Bedeutung von Ökosystemleistungen und deren Integration in politische Entscheidungsprozesse unterstreicht. So können die Ökosystemrechnungen beispielsweise für das Aktionsprogramm Natürlicher Klimaschutz (BMUV, 2023) genutzt werden.

Trotz der verschiedenen Ansätze und Blickwinkel bleibt es wichtig, bestehende Datenquellen – soweit rechtlich möglich und inhaltlich sinnvoll – besser zu nutzen und auf parallele Systeme zu verzichten. Künftig werden beispielsweise die Daten des Monitorings „Landnutzung, Landnutzungsänderung und Forstwirtschaft“ – welches ein Dateninput für die Waldgesamtrechnung ist – auch in die Berechnung der Ökosystemleistung der Globalen Klimaregulierung innerhalb der Ökosystemrechnungen einfließen. Ein anderes Beispiel erfolgreicher Integration bestehender Daten ist die Position des Nettozuwachses von Holz aus der Waldgesamtrechnung für die Ökosystemleistung „Bereitstellung von Holz“ innerhalb der Ökosystemrechnungen. Die Angleichung der Systeme bietet eine Perspektive für eine bessere Interpretation der Daten. Eine gemeinsame Datengrundlage, die alle Bedingungen der jeweiligen Berichtssysteme erfüllt, existiert bislang nicht und ist auch nicht absehbar.

Letztlich wird der Wald aus unterschiedlichen Blickwinkeln betrachtet. Die Vielfalt der Berichtssysteme spiegelt die unterschiedlichen Aspekte der Nutzung des Waldes wider und schafft die Grundlagen für spezifische Analysen: Für ökosystematische Untersuchungen sind die Ökosystemrechnungen geeignet, für die tatsächliche Flächennutzung die Flächenstatistik und für den ökonomischen Blick auf die Forstwirtschaft ist die Waldgesamtrechnung die Datenquelle der Wahl. Das Ziel, verlässliche Daten rund um das Thema Wald zu produzieren, bleibt essenziell. 

LITERATURVERZEICHNIS

AdV (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland). *Dokumentation zur Modellierung der Geoinformationen des amtlichen Vermessungswesens (GeoInfoDok) – Ausleitung des Objektartenkataloges für das AAA-Anwendungsschema. Version 7.1.2. Stand: 01.11.2022*. [Zugriff am 28. Februar 2025]. Verfügbar unter: www.adv-online.de

AdV (Arbeitsgemeinschaft der Vermessungsverwaltungen der Länder der Bundesrepublik Deutschland). *ALKIS-Objektartenkatalog Version 6.0. Stand: 11.04.2008*. [Zugriff am 25. Februar 2025]. Verfügbar unter: www.adv-online.de

Arnold, Stephan/Lucas, Christian/Pauly, Ralf. *Der neue Nutzungsartenkatalog zur erweiterten tatsächlichen Nutzung in der amtlichen Flächenstatistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2020, Seite 44 ff.

Eurostat. *European forest accounts handbook*. Ausgabe 2024. [Zugriff am 18. Februar 2025]. Verfügbar unter: ec.europa.eu

BMEL (Bundesministerium für Ernährung und Landwirtschaft). *Vierte Bundeswaldinventur 2022*. 2024. [Zugriff am 25. Februar 2025]. Verfügbar unter: www.bundeswaldinventur.de

BMUV (Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz). *Aktionsprogramm Natürlicher Klimaschutz. Kabinettsbeschluss vom 29. März 2023*. 2023. [Zugriff am 12. Februar 2025]. Verfügbar unter: www.bmu.de

Rosenkranz, Lydia/Englert, Hermann/Jochem, Dominik/Seintsch, Björn. *Methodenbeschreibung zum Tabellenrahmen der European Forest Accounts und Ergebnisse der Jahre 2014 und 2015*. 2. revidierte Fassung 2018. [Zugriff am 12. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Methode der Zustandsbilanzierung der Ökosysteme*. 2023. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Qualitätsbericht: Flächenerhebung nach Art der tatsächlichen Nutzung 2023*. 2024. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Umweltökonomische Gesamtrechnungen. Methode der Flächenbilanzierung der Ökosysteme*. 2021a. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Umweltökonomische Gesamtrechnungen. Ökosystemgesamtrechnungen. Nationale Ökosystemklassifikation für Deutschland*. 2021b. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.destatis.de

Statistische Ämter des Bundes und der Länder. *Methodenbericht zur Flächenerhebung Ausgabe 2018*. 2019. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.destatis.de

LITERATURVERZEICHNIS

Umweltbundesamt. *Berichterstattung unter der Klimarahmenkonvention der Vereinten Nationen und dem Kyoto-Protokoll 2023. Nationaler Inventarbericht zum Deutschen Treibhausgasinventar 1990 – 2021*. 2023. [Zugriff am 18. Februar 2025]. Verfügbar unter: www.umweltbundesamt.de

RECHTSGRUNDLAGEN

Bundes-Klimaschutzgesetz (KSG) vom 12. Dezember 2019 (BGBl. I Seite 2513), das zuletzt durch Artikel 1 des Gesetzes vom 15. Juli 2024 (BGBl. I Nr. 235) geändert worden ist.

Gesetz über Agrarstatistiken (Agrarstatistikgesetz – AgrStatG) in der Fassung der Bekanntmachung vom 17. Dezember 2009 (BGBl. I Seite 3886), das zuletzt durch Artikel 1 des Gesetzes vom 14. November 2022 (BGBl. I Seite 2030) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Gesetz zur Erhaltung des Waldes und zur Förderung der Forstwirtschaft (Bundeswaldgesetz – BWaldG) vom 2. Mai 1975 (BGBl. I Seite 1037), das zuletzt durch Artikel 112 des Gesetzes vom 10. August 2021 (BGBl. I Seite 3436) geändert worden ist.

Verordnung (EU) Nr. 525/2013 des Europäischen Parlaments und des Rates vom 21. Mai 2013 über ein System für die Überwachung von Treibhausgasemissionen sowie für die Berichterstattung über diese Emissionen und über andere klimaschutzrelevante Informationen auf Ebene der Mitgliedstaaten und der Union und zur Aufhebung der Entscheidung Nr. 280/2004/EG (Amtsblatt der EU Nr. L 165, Seite 13).

Verordnung (EU) 2024/1991 des Europäischen Parlaments und des Rates vom 24. Juni 2024 über die Wiederherstellung der Natur und zur Änderung der Verordnung (EU) 2022/869 (Amtsblatt der EU Reihe L vom 29. Juli 2024).

Verordnung (EU) 2024/3024 des Europäischen Parlaments und des Rates vom 27. November 2024 zur Änderung der Verordnung (EU) Nr. 691/2011 in Bezug auf die Einführung neuer Module für die umweltökonomischen Gesamtrechnungen (Amtsblatt der EU Reihe L vom 6. Dezember 2024, Seite 1).

PERSONENVERKEHRSSTATISTIK IM WANDEL

Karl Hofmann

➤ **Schlüsselwörter:** ÖPNV – COVID-19-Pandemie – Deutschlandticket –
Automatische Fahrgastzählsysteme – Mobilitätsdaten

ZUSAMMENFASSUNG

Die amtliche Personenverkehrsstatistik informiert über die Entwicklung und Struktur des öffentlichen Personenverkehrs. Sie basiert auf den Angaben der Verkehrsunternehmen. Aufgrund struktureller Änderungen ist es für diese jedoch zunehmend schwieriger, zuverlässige und genaue Daten zu liefern, was sich auf die Qualität der vierteljährlichen Ergebnisse auswirkt. Zudem sind regionale Ergebnisse der jährlichen Personenverkehrsstatistik nicht hinreichend genau. Der Artikel erläutert zwei Lösungsansätze aus einem EU-Projekt, um die Datengrundlage der Statistik zu verbessern: einen experimentellen Index sowie die Verwendung von Mobilitätsdaten.

➤ **Keywords:** *public short-distance passenger transport – COVID-19-pandemic – Germany ticket – automatic passenger counting systems – mobility data*

ABSTRACT

Official passenger transport statistics provide information on the development and structure of public passenger transport. They are based on data from transport enterprises. Due to structural changes, however, it has become increasingly difficult for these enterprises to report reliable and accurate data, which in turn affects the quality of the quarterly results. In addition, regional results of annual passenger transport statistics are not accurate enough. This article discusses two approaches that were developed in an EU-funded project to enhance the data basis of the statistics: an experimental index and the use of mobility data.



Karl Hofmann

ist Volkswirt und Wissenschaftlicher Mitarbeiter im Referat „Personenverkehr, Verkehrsunfälle“ des Statistischen Bundesamtes. In einem von der Europäischen Union finanzierten Projekt entwickelt er die Datengrundlage der Personenverkehrsstatistik weiter. Dafür untersucht er die Nutzung von Mobilitätsdaten und konzipiert einen experimentellen Index.

1

Einleitung

Der öffentliche Personennahverkehr¹ (ÖPNV) spielt sowohl gesellschaftlich als auch wirtschaftlich eine wichtige Rolle. So erleichtert er die Anreise zu Arbeitsstätten und Bildungseinrichtungen. Ein für viele Menschen zugänglicher, gut ausgebauter und bezahlbarer öffentlicher Personennahverkehr kann soziale Gerechtigkeit und Teilhabe unterstützen.¹² Angaben zur ÖPNV-Nutzung liefert die amtliche Personenverkehrsstatistik (PVS). Sie erhebt dazu quartalsweise und jährlich Angaben bei Verkehrsunternehmen, differenziert nach den Verkehrsmitteln Eisenbahn, Straßenbahn und Omnibus. Die Personenverkehrsstatistik informiert damit über konjunkturelle und strukturelle Entwicklungen.

Gemessen am gesamten Fahrgastaufkommen repräsentiert der öffentliche Personennahverkehr etwa 99 % des gesamten Linienverkehrs in Deutschland.¹³ Von 2004 bis 2019 wurden im öffentlichen Personennahverkehr stetig mehr Personen befördert und sie legten mehr Kilometer zurück. Die COVID-19-Pandemie und Maßnahmen zu ihrer Eindämmung führten im Jahr 2020 zu starken Rückgängen.

Mit der Pandemie verstärkte sich der Trend zum Home-office. Viele Menschen pendeln nicht mehr jeden Tag zu ihrer Arbeitsstätte (Statistisches Bundesamt, 2024a). Außerdem vereinfachte das [Deutschlandticket](#) als neues, preiswertes monatliches Abonnement die ÖPNV-Nutzung ab Mai 2023 erheblich.

Die Pandemie und das Deutschlandticket führten bei einem Teil der in der amtlichen Personenverkehrsstatistik befragten Unternehmen zu erheblichen Schwierigkeiten, genaue und verlässliche Meldungen zu ihrer Fahrgastbeförderung abzugeben. Des Weiteren zeigte sich, dass Unternehmen Probleme haben, ihre erbrachten Fahrleistungen regional aufzuteilen und ihre Linien-

längen nach Bundesländern zu gliedern. Gleichzeitig stieg das Informationsbedürfnis über den öffentlichen Personennahverkehr während und nach der Pandemie sowie mit Einführung des Deutschlandtickets.

Der Aufsatz beschäftigt sich mit zwei Ansätzen aus einem Projekt der Europäischen Union (EU) zur Verbesserung der Datengrundlage der vierteljährlichen und jährlichen Personenverkehrsstatistik im Bereich des öffentlichen Personennahverkehrs. Kapitel 2 erläutert die Systematik der Personenverkehrsstatistik, beschreibt die Entwicklung der Fahrgastzahlen und diskutiert die Ursachen mangelnder Genauigkeit und Zuverlässigkeit. Anschließend stellt Kapitel 3 zwei freiwillige Erhebungen und ihre Ergebnisse dar. Kapitel 4 schildert die Entwicklung eines experimentellen Index, der die Veränderung der Fahrgastzahlen und der Beförderungsleistung abbildet. Kapitel 5 stellt die Nutzung von statischen Fahrplan- und Haltestellendaten aus der „Mobilithek“ vor. Der Aufsatz schließt mit einer Zusammenfassung und einem Ausblick in Kapitel 6.

2

Überblick

2.1 Gesetzesgrundlage

Das Verkehrstatistikgesetz ordnet in § 1 die Durchführung von Bundesstatistiken „zur Beurteilung der Struktur und Entwicklung [...] des Schienenverkehrs und des gewerblichen Straßen-Personenverkehrs“ an. Die Angaben zur Personenverkehrsstatistik erheben die Statistischen Ämter der Länder. Auskunftspflichtig sind Unternehmen, die als Betriebsführer oder beauftragte Beförderer [öffentlichen Personennahverkehr](#) mit Eisenbahnen oder Straßenbahnen (Schienennahverkehr) oder Personennah- oder Personenfernverkehr mit Omnibussen betreiben (§ 16 Absatz 1 Verkehrstatistikgesetz).¹⁴ Zur Personenverkehrsstatistik werden alle Unternehmen herangezogen, die Personenbeförderung betreiben, unabhängig davon, ob dies ihre Haupttätigkeit ist (funktionaler Ansatz).

1 Der öffentliche Personennahverkehr besteht aus dem Schienenpersonennahverkehr und dem Nahverkehrsteil des öffentlichen Straßenpersonennahverkehrs. Der restliche Linienverkehr entfällt auf den Fernverkehr mit Bussen und Bahnen.

2 Zum Thema Mobilität in Verbindung mit Armut und sozialer Teilhabe siehe Aberle und andere (2025), Stark und andere (2023) sowie Dangschat (2022) und die Literatur darin.

3 Dies galt sowohl 2019 vor der COVID-19-Pandemie als auch 2023 (BMDV, 2024, Kapitel B5).

4 Der Personenfernverkehr mit Omnibussen wird in diesem Beitrag nicht betrachtet.

Das Verkehrstatistikgesetz regelt in § 16 Absatz 1 die Durchführung

- › einer **vierteljährlichen Erhebung** bei Unternehmen, die mindestens 250 000 Fahrgäste im Jahr befördert haben, sowie bei allen Unternehmen mit gewerblichem Schienennahverkehr (§ 17 Absatz 1 Nr. 1 Verkehrstatistikgesetz),
- › einer **jährlichen Erhebung** bei allen Unternehmen, die vierteljährlich befragt werden, sowie bei einer geschichteten Stichprobe von 2 500 Unternehmen, die weniger als 250 000 Fahrgäste im Jahr befördert haben, und
- › einer **fünfjährlichen Erhebung** bei allen Unternehmen.⁵

Die in diesem Aufsatz betrachtete vierteljährliche Erhebung erfasst im Liniennahverkehr die Anzahl der Fahrgäste und die Beförderungsleistung insgesamt (Unternehmensfahrten) sowie aufgeteilt auf die drei Verkehrsmittel Eisenbahnen, Straßenbahnen und Omnibusse (Verkehrsmittelfahrten).

2.2 Entwicklungen im öffentlichen Personennahverkehr

Basierend auf Quartalssummen hat der öffentliche Personennahverkehr im Jahr 2019 rund 11,5 Milliarden Fahrgäste befördert und rund 108 Milliarden Personenkilometer Beförderungsleistung erbracht.⁶ Damit wurden die höchsten Werte seit Beginn der Zeitreihe im Jahr 2004 gemessen. Im ersten Pandemiejahr 2020 brachen die Fahrgastzahlen und die Beförderungsleistung gegenüber 2019 um jeweils rund ein Drittel ein. Somit repräsentiert das Jahr 2019 das Vorkrisenniveau und dient als Vergleichsgrundlage für die Analyse von Fahrgastdynamiken vor, während und nach der Pandemie.⁷

➤ **Grafik 1** zeigt die unterjährliche Entwicklung absoluter Fahrgastzahlen. Die Zahl der insgesamt beförderten Fahrgäste im Liniennahverkehr halbierte sich zwischen dem vierten Quartal 2019 (= Hoch) und dem zweiten Quartal 2020 (= historisches Tief). Dies ist hauptsächlich auf den ersten Lockdown ab März 2020 zurückzuführen. Nach kurzer Erholung ließen zwei weitere Lockdown-Maßnahmen ab November 2020 und das Recht auf Homeoffice ab Ende Januar 2021 die Fahrgastzahlen erneut sinken (Statistisches Bundesamt, 2022). Erst

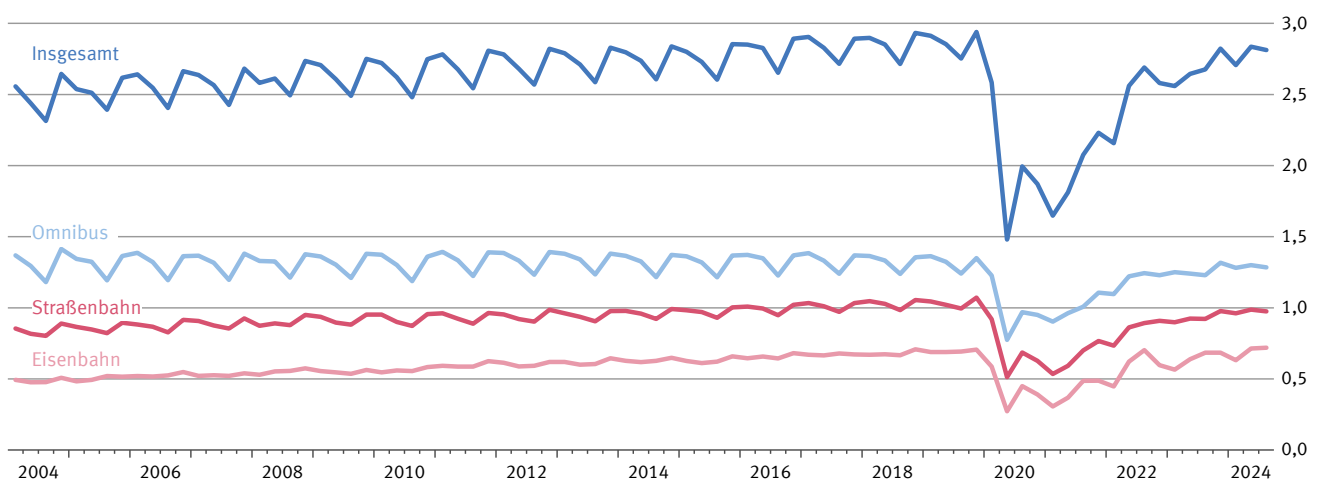
5 Für weitere Informationen zur Personenverkehrsstatistik siehe Bierau/Reim (2004).

6 Die Beförderungsleistung ist das Produkt aus der Anzahl der Fahrgäste und der mittleren Reiseweite.

7 Für vergleichbare Analysen für den Flugverkehr siehe Haas (2023).

Grafik 1

Entwicklung der Fahrgastzahlen im öffentlichen Personennahverkehr
Mrd. Fahrgäste



nach dem Auslaufen der Eindämmungsmaßnahmen ab März 2022 nahmen die Fahrgastzahlen schrittweise wieder zu.

Mit dem Wegfall der Pandemie-Maßnahmen zeigen sich strukturelle Veränderungen. Ergebnisse des Mikrozensus belegen, dass sich das Arbeiten von zu Hause in Deutschland in bestimmten Wirtschaftsbereichen etabliert hat (Statistisches Bundesamt, 2023). Zudem soll das Deutschlandticket mit seinem einheitlichen und günstigen Preis sowie seiner deutschlandweiten Gültigkeit die ÖPNV-Nutzung vereinfachen und attraktiver machen. Im Juni 2024 wurden nach Berechnungen des Verbands Deutscher Verkehrsunternehmen (VDV) 13,8 Millionen Deutschlandtickets genutzt. Der häufigste Nutzungsgrund ist das Pendeln (49 %), gefolgt von Alltagsfahrten (27 %) sowie Freizeit- und Urlaubsfahrten (24 %). Betrachtet man alle Fahrten im öffentlichen Personennahverkehr, so wurden im Juni 2024 bereits 58 % davon mit dem Deutschlandticket durchgeführt.¹⁸

2.3 Datenquellen

Die Personenverkehrsstatistik beruht auf den Angaben der Verkehrsunternehmen. Anfälligkeiten dieses Systems waren in der Vor-Corona-Zeit nicht offenkundig, da die Entwicklung im öffentlichen Personennahverkehr über Quartale und Jahre stabilen Mustern folgte. Die COVID-19-Pandemie deckte methodische Einschränkungen in der Personenverkehrsstatistik auf. Den Verkehrsunternehmen fehlten valide Erfahrungswerte zum Fahrgastaufkommen und -verhalten unter Pandemiebedingungen. So häuften sich 2020 im zweiten und dritten Quartal unplausible Meldungen: Die Angaben zu Fahrgästen und Beförderungsleistung im Liniennahverkehr passten nicht zu den zu erwartenden Rückgängen durch Ausgangsbeschränkungen und Lockdowns.¹⁹ Die amtliche Personenverkehrsstatistik hat für den Liniennahverkehr mit Omnibussen und Straßenbahnen in beiden Quartalen keine Angaben veröffentlicht, da es den befragten Unternehmen nicht möglich war, Schätzungen fundiert und pünktlich zu erstellen. Das 9-Euro-Ticket,

das im Sommer 2022 für drei Monate erhältlich war, und das ab Mai 2023 eingeführte Deutschlandticket bringen ebenso Sondereffekte mit sich, die die Statistik möglicherweise nicht zuverlässig widerspiegelt. Da weniger Einzelfahrten gebucht werden und es an aktuellen Informationen zu Nutzungshäufigkeiten der neuen Zeitkarten mangelt, ist es für die Unternehmen schwieriger, die Fahrgastzahlen zu ermitteln.

Den Verkehrsunternehmen stehen folgende Informationsquellen zur Verfügung, um Fahrgastzahlen und die Beförderungsleistung zu ermitteln:

- › **Automatische Fahrgastzählsysteme** (AFZS) sind eine wichtige interne Informationsquelle für Verkehrsunternehmen und -verbünde. Sie können Beförderungsleistungen und Verkehrsmittelfahrten direkt ermitteln. Die Systeme enthalten eine Sensorik, die kontinuierlich und systematisch ein- und aussteigende Fahrgäste zählt. Da Anschaffung, Einsatz und Auswertung kostspielig sind, werden die Systeme noch nicht flächendeckend eingesetzt. Dies gilt insbesondere für Omnibusse.
- › Eine zweite wichtige interne Informationsquelle sind **manuelle Fahrgasterhebungen** durch Zählung und Befragung. Letztere ermittelt neben der Beförderungsleistung auch die Fahrgast- und Umsteigerzahlen und die genutzten Ticketarten. Manuelle Erhebungen sind jedoch teuer und komplex. Es fehlt zudem an geschultem Personal. Zwischen 2020 und 2022 waren manuelle Erhebungen während der Lockdowns und Ausgangsbeschränkungen stark eingeschränkt oder nicht möglich.
- › Eine weitere Informationsquelle sind **Verkaufszahlen zu Einzeltickets und Zeitfahrkarten** je Unternehmen und Verbund. Einzeltickets ermöglichen präzise Angaben zur Zahl der Fahrten, sie repräsentieren jedoch nur einen Teil des Fahrgastaufkommens. Aus den verkauften Zeitfahrkarten ergeben sich hingegen keine direkten Kenntnisse über die Fahrgastzahlen. Diese müssen anhand angenommener Fahrhäufigkeiten geschätzt werden. Das Deutschlandticket hat diese Problematik erheblich verschärft: Es wird nicht nur von den Verkehrsverbünden selbst vertrieben, sondern auch von zentralen Anbietern. Alle Anbieter melden ihre Verkäufe mit Angabe der Postleitzahl der kaufenden Person an die **Deutschland-Ticket-Clearing-Stelle**. Der Deutschlandticket-Monitor stellt

8 Quelle: Evaluation Deutschlandticket | Interpretationsbericht 1. Hj. 2024 (pdf). [Zugriff am 17. März 2025]. Verfügbar unter www.vdv.de

9 Das Bundesministerium für Digitales und Verkehr erweiterte im Jahr 2020 die „Gleitende Mittelfristprognose“ um das „Kurzfristige Monitoring des Personenverkehrs während der Corona-Pandemie“ (BMDV, 2021).

den Tarifparteien – oft Verkehrsverbünde – regionale Verkaufszahlen seit Mai 2024 bereit. Nachteil ist, dass die Unternehmen als Meldeeinheiten der Personenverkehrsstatistik keinen direkten Zugriff auf die Informationen haben.

- › Daher verwendet ein Teil der Verkehrsunternehmen regelmäßige **Empfehlungen** des Verbands Deutscher Verkehrsunternehmen **zu Fahrtenhäufigkeiten** von Zeitkarten und dem Deutschlandticket als eine externe Quelle.¹⁰ Die Empfehlungen erarbeiten Unternehmen mit guten Datengrundlagen wie automatischem Fahrgastzählsystem gemeinsam mit dem Verband Deutscher Verkehrsunternehmen. Hierdurch ergibt sich ein zeitlicher Verzug, sodass die Empfehlungen den Auskunftspflichtigen zum Zeitpunkt der Meldung für die Vierteljahresstatistik noch nicht zur Verfügung stehen. Im Omnibusverkehr fließen außerdem Ergebnisse der amtlichen Statistik in die Empfehlungen ein, sodass sich ein Zirkelschluss ergibt, wenn die Unternehmen wiederum auf Basis der Empfehlungen ihre Fahrgastzahlen schätzen.

Die Auflistung verdeutlicht, dass es für Verkehrsunternehmen, die Fahrgäste weder automatisch noch manuell zählen, bei zunehmend steigendem Anteil von Zeitkarten und insbesondere dem Deutschlandticket schwierig ist, valide Angaben zu Fahrgastzahlen und Beförderungsleistungen zu ermitteln. Für die amtliche Statistik stellte sich damit die Frage, wie sie trotz dieser Rahmenbedingungen die Qualität der Personenverkehrsstatistik verbessern kann. Hierfür sammelte sie in zwei freiwilligen Erhebungen weitere Informationen.

3

Freiwillige Zusatzerhebungen

Um wissenschaftlich-methodische Fragestellungen zu klären, wurden zwei freiwillige Erhebungen nach § 7 Bundesstatistikgesetz durchgeführt. Zunächst erfolgte im Dezember 2022 eine Befragung der Verkehrsverbünde zu den ihnen vorliegenden Daten. Anschlie-

ßend – und im Rahmen eines EU-Projekts – wurden im dritten Quartal 2023 die Verkehrsunternehmen des vierteljährlichen Berichtskreises befragt. Folgende Themen standen bei den Erhebungen im Mittelpunkt:

- › Qualität der Datengrundlagen bei den Verkehrsverbünden,
- › Zusammenarbeit zwischen Unternehmen und Verkehrsverbünden,
- › Einsatz von automatischen Fahrgastzählssystemen in Unternehmen und
- › Angebot von On-Demand-Verkehren¹¹ durch Unternehmen.

3.1 Erhebung bei Verkehrsverbünden

Die von der amtlichen Statistik erhobenen Angaben liegen nicht nur den Verkehrsunternehmen, sondern auch den Verkehrsverbünden vor. Daher stellte sich die Frage, ob die Quartalerhebung der Personenverkehrsstatistik zumindest teilweise von Unternehmen auf Verkehrsverbünde verlagert werden kann.

Von 83 Verkehrsverbünden nahmen 58 an der freiwilligen Erhebung 2022 teil, davon machten 53 verwertbare Angaben. Insgesamt lagen bei 30 von 53 Verbünden Daten zur Gesamtzahl beförderter Fahrgäste vor. Den Verbünden fehlten jedoch Informationen über die Aufteilung der Leistungsmerkmale auf die Verkehrsmittel. Auch die Beförderungsleistung je Verkehrsmittel ist vielen Verkehrsverbünden unbekannt. Somit ist eine flächendeckende Umstellung der Erhebung nicht möglich.

Die Datenlage bei den großen Verbünden¹² war dabei besser. Von 30 Verkehrsverbünden wurden 10 als groß eingestuft; sie verfügen über verkehrsmittelgenaue Daten zu den Fahrgastzahlen. Könnte eine Teilumstellung der Erhebung bei Verbünden mit guter Datenlage die Unternehmen entlasten? Dafür war zu klären, ob Unternehmen in mehreren Verbundgebieten tätig sind und ob sie Verkehre in Eigenregie durchführen.

10 Die Empfehlungen des Verbands Deutscher Verkehrsunternehmen sind primär für Mitgliedsunternehmen vorgesehen. Diese beförderten im Jahr 2023 rund 88 % aller Fahrgäste im öffentlichen Personennahverkehr (www.vdv.de). Nicht-Mitgliedsunternehmen können einen möglicherweise eingeschränkten Zugang zu den Empfehlungen erhalten.

11 On-Demand-Verkehre sind Mobilitätsangebote auf Bestellung. Darunter fallen Einzel- und Sammelfahrten mit Personenkraftwagen und Kleinbussen, die man telefonisch oder per App buchen kann. So sollen Lücken im öffentlichen Personennahverkehr in dünn besiedelten Gebieten und zu Tagesrandzeiten verringert werden (BeNaMo, 2025).

12 Große Verbünde befördern 10 % oder mehr der Fahrgäste in einem Bundesland.

Dies mündete in eine EU-finanzierte freiwillige Erhebung der Unternehmen des vierteljährlichen Berichtskreises im dritten Quartal 2023, an der 326 von 798 Meldeeinheiten teilnahmen. Aufgegliedert beförderten 64 % der antwortenden Unternehmen Fahrgäste für einen Verkehrsverbund, 21 % für zwei oder mehr Verbünde und 15 % unabhängig von einem Verbund. Hierbei zeigte sich, dass knapp die Hälfte der Unternehmen ausschließlich Leistungen für Verkehrsverbünde erbringt, während die andere Hälfte parallel dazu auch Verkehre in Eigenregie durchführt. Somit ist eine teilweise Übertragung der Auskunftspflicht auf große Verkehrsverbünde nicht möglich, weil die Unternehmen weiterhin einen Teil ihrer Verkehre melden müssten, um eine Untererfassung zu vermeiden.

3.2 Einsatz von automatischen Fahrgastzählssystemen

Automatische Fahrgastzählssysteme (AFZS) liefern nachfrageseitige, statistisch gesicherte Daten. Diese benötigen Unternehmen, Aufgabenträger und Verkehrsverbünde in Entscheidungs-, Kontroll- und Nachweisprozessen, zum Beispiel für die Finanzierung von Verkehrsleistungen (VDV, 2023).

Von den 326 an der freiwilligen Erhebung teilnehmenden Unternehmen setzen 90 automatische Fahrgastzählssysteme in ihren Verkehrsmitteln ein. Diese beförderten 59 % aller Fahrgäste, die in Deutschland im Berichtsjahr 2023 mit Eisenbahnen fuhren, 54 % der Fahrgäste mit Straßenbahnen und 37 % der Fahrgäste mit Omnibussen.

Unternehmen, die noch keine automatischen Zählssysteme in ihren Verkehrsmitteln einsetzen (siehe auch Tabelle 1), wurden befragt, ob sie dies künftig planen. Von den zehn Unternehmen mit Eisenbahnen planen drei und von den sechs Unternehmen mit Straßenbahnen plant eines 2024 oder in späteren Jahren den Einsatz von automatischen Fahrgastzählssystemen (siehe Abschnitt 4.2). Bei den Unternehmen mit Omnibusverkehr planen rund 50 von 255 in den nächsten Jahren den ersten Einsatz von automatischen Fahrgastzählssystemen. Die Hälfte davon befördert Fahrgäste in Baden-Württemberg.¹³

13 Baden-Württemberg fördert den AFZS-Einsatz im Rahmen des [Gemeindeverkehrsfinanzierungsgesetzes](#).

3.3 Befragung zu On-Demand-Verkehren

On-Demand-Verkehre wurden bei der letzten Novellierung des Personenbeförderungsgesetzes neu geregelt. Linienbedarfsverkehr nach § 44 mit Fahrzeugen, die bis zu neun Sitzplätze aufweisen, sowie der gebündelte Bedarfsverkehr nach § 50 erhebt die Personenverkehrsstatistik bisher nicht.

Nur 40 der 326 an der freiwilligen Erhebung teilnehmenden Verkehrsunternehmen boten im dritten Quartal 2023 Linienbedarfsverkehr oder gebündelten Bedarfsverkehr an. Insgesamt haben 37 der 40 Unternehmen rund 1,5 Millionen Fahrgäste mit On-Demand-Verkehren befördert.¹⁴ Dies sind rund 0,1 % der von den 326 Unternehmen beförderten Quartalsfahrgästen. Da der Anteil der On-Demand-Verkehre bei den Quartalsunternehmen der Personenverkehrsstatistik verschwindend gering ist, untersucht die amtliche Statistik dieses Beförderungsangebot weiterhin nicht.

4

Entwicklung eines experimentellen AFZS-Index

Um belastbarere Aussagen zur Veränderung von Fahrgastzahlen und Beförderungsleistung treffen zu können, entwickelte das Statistische Bundesamt in einem EU-Projekt einen [experimentellen AFZS-Index](#) für die Quartalerhebung. Ein Index gewährt einen zeitlichen Vergleich relativer Veränderungen in Bezug auf eine Basisperiode. So ist eine einfache Interpretation für Nutzende der Personenverkehrsstatistik möglich. Des Weiteren ist ein Index relativ robust gegenüber Verzerrungen, da sowohl systematische als auch stochastische Fehler (wie Messfehler) nicht so stark ins Gewicht fallen wie bei Absolutwerten.

4.1 Konzept

Ein AFZS-Index informiert die Nutzenden zusätzlich zur herkömmlichen Personenverkehrsstatistik. Der AFZS-Index soll Sonderentwicklungen im Personennahverkehr

14 Die restlichen drei Unternehmen machten keine Angaben zur Zahl der Fahrgäste.

genau und zuverlässig abbilden, und zwar vor allem dann, wenn unvorhergesehene Ereignisse (wie Pandemien) oder gänzlich neue Trends (wie das Deutschlandticket) eintreten und zu besonderen Dynamiken im Mobilitätsgeschehen führen. Für den Einsatz von automatischen Fahrgastzählsystemen in den Verkehrsunternehmen gibt es detaillierte Handlungsempfehlungen zum Aufbau einer Messfahrtenplanung mit Stichprobenplan sowie zur Hochrechnung der Messwerte auf die Gesamtheit der Fahrten und eine anschließende Fehlerrechnung (VDV, 2023).

In den experimentellen AFZS-Index fließen Werte der Quartalsmeldeeinheiten je Verkehrsmittelart ein. Er berücksichtigt nur jene Unternehmen, die automatische Zählsysteme laut der freiwilligen Erhebung im dritten Quartal 2023 in mindestens 10% ihrer Fahrzeuge einsetzen und AFZS-Daten zumindest teilweise für die Statistikmeldung nutzten (AFZS-Index-Bedingung).

Unternehmen gehen in den AFZS-Index gleichgewichtet ein. Der Index stellt die Veränderungen der Fahrgastzahlen und der Beförderungsleistung im Vergleich zum ersten Quartal 2023 dar. Dieses wurde als Basisquartal gewählt, weil es frei von Effekten durch Corona oder Son-

dertickets ist. Für die Erstellung der Indexverläufe wurde angenommen, dass die Angaben der Unternehmen zu automatischen Fahrgastzählsystemen bereits im ersten Quartal 2022 galten. Vorherige Quartale waren aufgrund der Pandemiebedingungen mit größerer Unsicherheit behaftet und wurden nicht einbezogen.

Parallel zum AFZS-Index wurde ein Index erstellt, der die Aktivitäten aller Quartalsmeldeeinheiten je Verkehrsmittelart abbildet (PVS-Index). Die folgenden Abschnitte stellen Analysen zur Repräsentativität und zu den Ergebnissen beider Indizes vor. Auf einen „Nicht-AFZS-Index“, der die Veränderungsraten für Quartalseinheiten ohne AFZS-Einsatz oder ohne AFZS-Datennutzung widerspiegelt, wurde aus methodischen Gründen verzichtet. Die dafür infrage kommenden Einheiten aus der freiwilligen Zusatzerhebung haben einen zu geringen Anteil an der Fahrgastbeförderung.

4.2 Repräsentativität

Von den 90 Unternehmen, die in der freiwilligen Erhebung angaben, automatische Fahrgastzählsysteme einzusetzen, erfüllen 38 Unternehmen die AFZS-Index-

Tabelle 1

Abdeckung des AFZS-Index im öffentlichen Personennahverkehr im Berichtsjahr 2023

	Unternehmen		Fahrgäste		Beförderungsleistung	
	Anzahl	%	Anzahl	%	Personen-kilometer	%
Eisenbahnverkehr						
Keine Teilnahme an freiwilliger Erhebung	28	51,9	1 009 012 483	39,1	16 382 918 376	28,9
Kein AFZS-Einsatz	10	18,5	63 263 250	2,5	1 902 684 330	3,4
AFZS-Einsatz > 0%, keine Auswertung für Statistik	2	3,7	40 177 800	1,6	2 221 921 268	3,9
AFZS-Einsatz > = 10%, Auswertung für Statistik	14	25,9	1 470 292 019	56,9	36 135 887 947	63,8
Zusammen	54	100	2 582 745 552	100	56 643 411 921	100
Straßenbahnverkehr						
Keine Teilnahme an freiwilliger Erhebung	38	66,7	1 424 811 853	38,3	5 766 969 419	36,5
Kein AFZS-Einsatz	6	10,5	295 551 872	8,0	1 552 349 995	9,8
AFZS-Einsatz > 0%, keine Auswertung für Statistik	4	7,0	197 985 852	5,3	911 612 912	5,8
AFZS-Einsatz > = 10%, Auswertung für Statistik	9	15,8	1 799 600 534	48,4	7 569 359 736	47,9
Zusammen	57	100	3 717 950 111	100	15 800 292 062	100
Omnibusverkehr						
Keine Teilnahme an freiwilliger Erhebung	440	59,5	2 554 261 166	50,8	20 984 089 941	59,5
Kein AFZS-Einsatz	225	30,5	601 500 549	12,0	5 892 901 920	16,7
AFZS-Einsatz > 0%, keine Auswertung für Statistik	51	6,9	682 394 907	13,6	4 252 539 450	12,1
AFZS-Einsatz > = 10%, Auswertung für Statistik	23	3,1	1 191 604 163	23,7	4 115 678 518	11,7
Zusammen	739	100	5 029 760 785	100	35 245 209 829	100

AFZS: Automatisches Fahrgastzählsystem.

Bedingung. Im Jahr 2023 beförderten 36 von ihnen 1 Million oder mehr Fahrgäste. Bei denjenigen Unternehmen, die die AFZS-Bedingung nicht erfüllten, beförderte nur rund die Hälfte 1 Million Fahrgäste oder mehr. Mehr als die Hälfte der AFZS-Unternehmen haben ihren Sitz in kreisfreien Großstädten.¹⁵ Bei den Unternehmen, die nicht Teil des AFZS-Index sind, war dies nur gut ein Zehntel. Liniennahverkehr betreibende Unternehmen aus Großstädten und solche mit hohen Fahrgastzahlen sind somit im AFZS-Index überrepräsentiert.

Da ein Unternehmen bis zu drei Verkehrsmittelarten im Einsatz haben kann, ergeben sich aus den AFZS-Index-Unternehmen 14 mit Eisenbahn-, 9 mit Straßenbahn- und 24 mit Omnibusverkehren. Trotz der geringen Anzahl erreichen diese Unternehmen im Berichtsjahr 2023 eine für den experimentellen AFZS-Index gute Abdeckung für die Zahl der beförderten Fahrgäste und die Beförderungsleistung bei Eisenbahnen und Straßenbahnen. Im Omnibusverkehr waren die Fahrgastabdeckung sowie der Anteil an der Beförderungsleistung durch automatische Zählsysteme hingegen sehr gering. [↗ Tabelle 1](#)

Die Fahrleistung nach Kreisen ist ein weiteres Kriterium, um die Repräsentativität der Unternehmen je Verkehrsmittel einzuschätzen. Bei dieser Kennziffer steht der Ort der Leistungserbringung im Fokus.¹⁶ Bei Eisen-

bahn- und Straßenbahnverkehren sind die prozentualen Unterschiede zwischen den Verkehrsmitteln aus der Grundgesamtheit und dem AFZS-Index gering. Für die AFZS-Indizes der schienengebundenen Verkehrsmittel wird daher von einer guten Abdeckung nach Kreistypen ausgegangen. Bei Omnibusverkehren hingegen sind Großstädte als Ort der Leistungserbringung im AFZS-Index mit 80 % stark überrepräsentiert. In der Grundgesamtheit der Quartalsunternehmen erbringen nur 30 % der Unternehmen mit Busverkehr ihre Leistung in Großstädten. [↗ Tabelle 2](#)

Weiterhin wurde der AFZS-Index auf seine Abdeckung des Ausbildungsverkehrs überprüft. Auszubildende verfügen über ein geringes Einkommen sowie weniger Möglichkeiten, auf andere Verkehrsmittel außerhalb des öffentlichen Personennahverkehrs auszuweichen. Sie nutzen in der Regel günstige Zeitfahrkarten und neuerdings das Deutschlandticket. Im Eisenbahnnahverkehr ist der Ausbildungsverkehr im AFZS-Index gut abgebildet. Im Straßenbahnverkehr ist er hingegen leicht und im Omnibusverkehr mit 20 % sogar stark unterrepräsentiert.¹⁷ [↗ Tabelle 3](#)

Die bisherigen Untersuchungen ergeben, dass bei den AFZS-Indizes für Eisen- und Straßenbahnen eine gute Abdeckung und Repräsentativität gegeben ist. Eine Ausnahme ist der Ausbildungsverkehr, welcher im AFZS-Index der Straßenbahnen leicht unterrepräsentiert ist.

15 Der Sitzort wurde den [siedlungsstrukturellen Kreistypen](#) des Bundesinstituts für Bau-, Stadt- und Raumforschung zugeordnet.

16 Der Ort der Leistungserbringung wurde den [siedlungsstrukturellen Kreistypen](#) des Bundesinstituts für Bau-, Stadt- und Raumforschung zugeordnet.

17 Bezogen auf den gesamten Verkehr spielt der Ausbildungsverkehr von allen drei Verkehrsmittelarten im Omnibusverkehr mit einem Anteil von 42 % die größte Rolle.

Tabelle 2

Fahrleistung des öffentlichen Personennahverkehrs nach siedlungsstrukturellen Kreistypen 2022

	Eisenbahnverkehr				Straßenbahnverkehr				Omnibusverkehr			
	Unternehmen des vierteljährlichen Berichtskreises		Unternehmen, die AFZS-Bedingungen erfüllen		Unternehmen des vierteljährlichen Berichtskreises		Unternehmen, die AFZS-Bedingungen erfüllen		Unternehmen des vierteljährlichen Berichtskreises		Unternehmen, die AFZS-Bedingungen erfüllen	
	km	%	km	%	km	%	km	%	km	%	km	%
Kreisfreie Großstadt	188 513 495	28	111 707 243	25	273 372 078	88	99 211 963	95	738 903 367	30	242 825 738	80
Städtischer Kreis	230 649 752	35	166 587 069	38	29 171 175	9	3 665 819	4	932 378 744	38	42 442 108	14
Ländlicher Kreis mit Verdichtungsansätzen	134 048 058	20	80 915 107	18	4 468 185	1	954 874	1	442 098 561	18	12 289 845	4
Dünn besiedelter ländlicher Kreis	114 947 791	17	78 904 029	18	3 685 252	1	561 900	1	361 098 695	15	5 461 401	2
Insgesamt	668 159 096	100	438 113 448	100	310 696 690	100	104 394 556	100	2 474 479 367	100	303 019 092	100

AFZS: Automatisches Fahrgastzählsystem.

AFZS-Bedingung: Im AFZS-Index werden nur jene Unternehmen berücksichtigt, die automatische Fahrgastzählsysteme in mindestens 10 % ihrer Fahrzeuge einsetzen und AFZS-Daten zumindest teilweise für die Statistikmeldung nutzen.

Tabelle 3

Fahrgäste im Ausbildungsverkehr 2022

	In Unternehmen des vierteljährlichen Berichtskreises		In Unternehmen, die AFZS-Bedingungen erfüllen	
	Anzahl	%	Anzahl	%
Eisenbahnverkehr	409 163 509	17	269 892 345	19
Straßenbahnverkehr	848 869 143	25	287 061 599	18
Omnibusverkehr	1 970 926 777	42	213 049 276	20

AFZS: automatisches Fahrgastzählsystem.

AFZS-Bedingung: Im AFZS-Index werden nur jene Unternehmen berücksichtigt, die automatische Fahrgastzählsysteme in mindestens 10 % ihrer Fahrzeuge einsetzen und AFZS-Daten zumindest teilweise für die Statistikmeldung nutzen.

Der AFZS-Index für Omnibusverkehre ist dagegen über alle Ebenen zu wenig repräsentativ, um mit ihm belastbare Aussagen zu treffen. Daher wird ein Index für Omnibusverkehre im Folgenden nicht weiter diskutiert.

4.3 Ergebnisse

➤ **Grafik 2** stellt die Verläufe des experimentellen AFZS-Index sowie die des PVS-Index gegenüber. Beide Indizes zeigen je Verkehrsmittel die Veränderung der Fahrgastzahlen im Vergleich zum Basisquartal.

Im dritten Quartal 2022 veränderte sich der AFZS-Index im Eisenbahnnahverkehr deutlich stärker als der PVS-

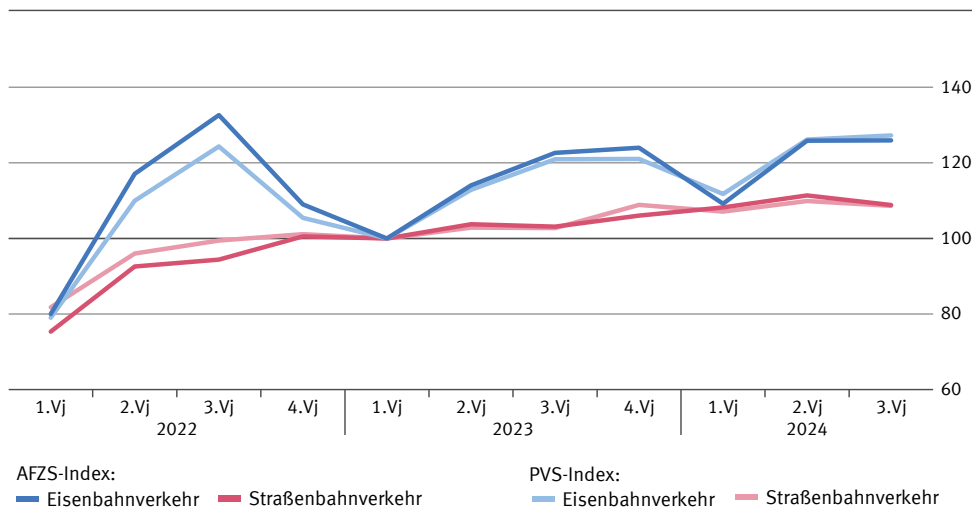
Index. Dies könnte darauf zurückzuführen sein, dass der AFZS-Index auf die gestiegene Nachfrage durch das niedrigpreisige 9-Euro-Ticket sensibler reagierte, während die Personenverkehrsstatistik die Fahrten mit einem 9-Euro-Ticket im Eisenbahnnahverkehr möglicherweise nicht vollständig oder mit gewissen Verzögerungen erfasste. Im Jahr 2023 zeichnet sich mit der Einführung des Deutschlandtickets ein ähnliches, jedoch spürbar abgeschwächtes Muster ab.

Der AFZS-Index für Straßenbahnverkehre startet im ersten Quartal 2022 auf einem niedrigeren Ausgangsniveau als der PVS-Index. Dies hängt möglicherweise damit zusammen, dass die Personenverkehrsstatistik die pandemiebedingten Rückgänge der Fahrgastzahlen im Winter 2021/2022 nicht vollständig oder verzögert darstellt. So löste die Omikron-Variante des Coronavirus ab Ende Dezember 2021 eine Welle mit hohen Infektionszahlen aus (BMG, 2023). Im gesamten Jahr 2022 zeigen sowohl der AFZS-Index als auch der PVS-Index für Straßenbahnen deutliche Zunahmen. Die Einführung des Deutschlandtickets wird von beiden Indizes für Straßenbahnen unterschiedlich abgebildet, doch wird die nachfragesteigernde Wirkung zwischen dem zweiten Quartal 2023 und dem zweiten Quartal 2024 deutlich.

Grafik 2

Fahrgastzahlen je Verkehrsmittel im AFZS-Index und im PVS-Index

1. Quartal 2023 = 100



AFZS: automatisches Fahrgastzählsystem. PVS: amtliche Personenverkehrsstatistik.

Im AFZS-Index werden nur jene Unternehmen berücksichtigt, die automatische Fahrgastzählsysteme in mindestens 10 % ihrer Fahrzeuge einsetzen und AFZS-Daten zumindest teilweise für die Statistikmeldung nutzen.

Zusammenfassend zeigt sich für den AFZS-Index bei Eisenbahnen und Straßenbahnen, dass dieser die Preiseffekte des 9-Euro-Tickets (2022) und des Deutschlandtickets (ab Mai 2023) jeweils gut abbildet. Ob ein AFZS-Index das Fahrgastgeschehen zuverlässiger abbildet als der PVS-Index, in den auch Daten von Unternehmen einfließen, die überwiegend oder ausschließlich externe Informationsquellen zur Ermittlung ihrer Angaben nutzen, ist noch nicht abschließend geklärt. Doch bergen automatische Fahrgastzählssysteme großes Potenzial: Sie sind Pulsmesser im Personenverkehr und erfassen auch besondere Schwankungen im Tagesgeschäft. Der AFZS-Index bietet eine Chance, Nutzende der Statistik zuverlässiger zu informieren. Dazu müsste der AFZS-Index auf eine umfassendere Datengrundlage gestellt werden.

5

Nutzung von Fahrplandaten

Ein zweiter Teil des EU-Projekts untersuchte, wie Fahrplandaten zur Ermittlung von Fahrleistungen auf Kreisebene und Linienlängen auf Bundeslandebene nutzbar sind. Die Delegierte Verordnung (EU) 2024/490 verpflichtet Anbieter des öffentlichen Personennahverkehrs, Daten zu Fahrplänen, Routen und Haltestellen georeferenziert an einen sogenannten nationalen Zugangspunkt (in Deutschland: [Mobilithek](#) des Bundesministeriums für Digitales und Verkehr) zu melden. Auf nationaler Ebene werden die EU-Vorgaben durch §§ 3a, 3b und 3c des Personenbeförderungsgesetzes und durch die Mobilitätsdatenverordnung umgesetzt. Die Daten stehen der Öffentlichkeit kostenlos mit einer Creative-Commons-Attribution-Lizenzierung zur Verfügung.

Gleichzeitig wird die Fahrleistung in der Personenverkehrsstatistik auf Kreisebene jährlich bei den Quartalsunternehmen und alle fünf Jahre bei allen Unternehmen erhoben. Die Linienlänge auf Länderebene wird ebenfalls fünfjährlich abgefragt. Da sich Fahrleistung und Linienlänge in den statischen Fahrplandaten widerspiegeln, entsteht eine doppelte Belastung der Unternehmen. Zudem hat ein Teil der Unternehmen Schwierigkeiten, die Merkmale Fahrleistung und Linienlänge auf Kreis- und auf Länderebene aufzuteilen. Die genannten Strukturmerkmale sind für die Unternehmen schwierig

zu ermitteln, da sie in ihrer Buchhaltung keine regionale Aufteilung führen. Dies spiegelt sich in der Qualität regionaler Ergebnisse wider.

Das bis Ende Mai 2025 laufende EU-Projekt prüft daher, ob die doppelte Befragungsbelastung reduziert und gleichzeitig die regionale Datenqualität erhöht werden kann. Dafür sollen die planmäßigen Fahrleistungen auf Kreisebene trennscharf ermittelt werden. Das Vorgehen wird nachfolgend beschrieben.

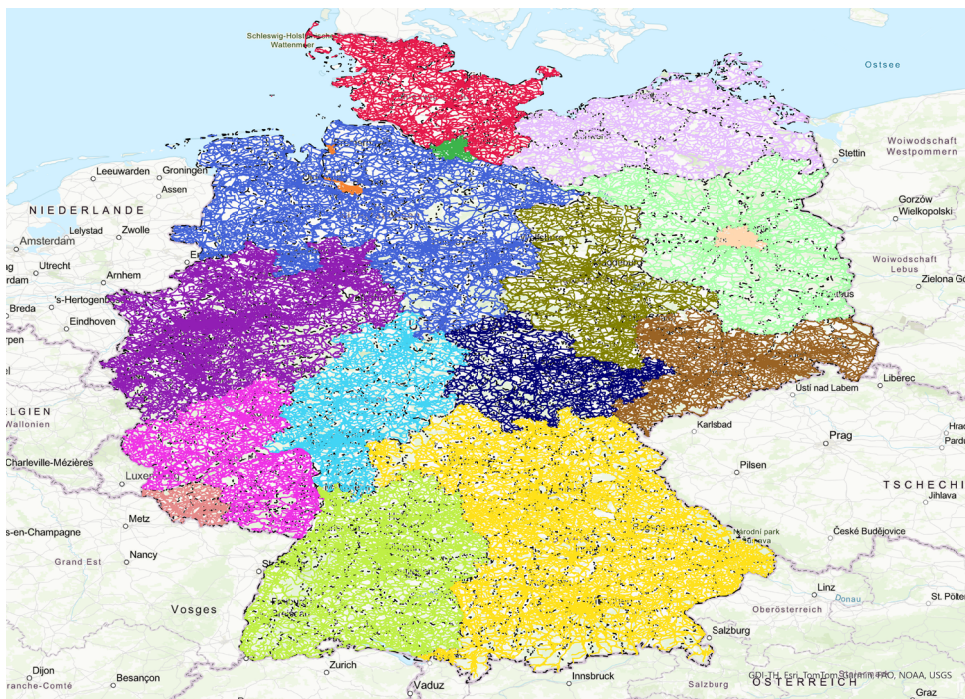
5.1 Statische Daten und Linien-Routings

Für die Übermittlung statischer Fahrplan-, Routen- und Haltestellendaten sind in den genannten europäischen und nationalen Gesetzesgrundlagen Datenformate, -standards und -modelle festgelegt. Die von den Unternehmen zu liefernden Merkmale sind in § 3a des Personenbeförderungsgesetzes definiert. Mit den Bundesländern stellt der [DELFI e.V.](#) (Durchgängige Elektronische Fahrgastinformation) eine IT-Infrastruktur für den Datentransfer bereit. Die Verkehrsunternehmen liefern ihre Fahrplandaten an ihren Verkehrsverbund. Über einen Landesdatenpool gelangen die Daten dann in die DELFI-Integrationsplattform, werden dort zusammengeführt, qualitätsgesichert und anschließend in der deutschlandweiten Open-Data-Plattform hinterlegt. Die Plattform ist mit dem Datenkatalog des nationalen Zugangspunkts verknüpft. In der Praxis bietet die Mobilithek den sogenannten statischen DELFI-Datensatz an (BMDV, 2023).

Für die Verarbeitung des DELFI-Datensatzes zu Verkehrsmittel-Linienverläufen (Linien-Routings) hat das Statistische Bundesamt einen externen Dienstleister beauftragt. Dieser wendet das Map-Matching-Verfahren von Bast/Brosi (2018) an. Basierend auf den Routings berechnet er im Anschluss regionale Strecken- und Leistungsmerkmale je Linie und je PVS-Verkehrsmittel auf Ebene der Kreise, Bundesländer und für Deutschland. Hierbei wird die fahrplanmäßige Leistung der Verkehrsunternehmen ermittelt. Da Linien je nach Uhrzeit und Wochentag unterschiedlich häufig und mit variierender Strecke befahren werden, wird für jede Variante eine Linie gezählt. Um die Linienlänge zu ermitteln, wird dann der fahrtenhäufigkeitsgewichtete Durchschnitt der Linienlänge berechnet. Daraus ergibt sich die Fahrleistung (Produkt aus Fahrtenhäufigkeit und Linienlänge).

Grafik 3

Umfang des Linien-Routings basierend auf dem statischen Fahrplan für das Jahr 2024



Quelle: Kartenreferenzsystem WGS84 (World Geodetic System 1984)

Die aus statischen Fahrplandaten berechneten Linien-Routings zeigt [Grafik 3](#) auf Ebene der Bundesländer. Linien-Routings für den Bund und auf Ebene der Kreise liegen ebenso vor.

5.2 Qualitätsprüfung der Linienroutings

Um als Datenquelle in der Personenverkehrsstatistik verwendet zu werden, wird die Vollständigkeit der Daten der Mobilithek und die Qualität der daraus abgeleiteten Linien-Routings untersucht. Sind die gerouteten Daten verwendbar, würde das nicht nur die Personenverkehrsstatistik verbessern, sondern auch die meldenden Unternehmen entlasten. Die Untersuchung zur Qualität ist zum aktuellen Zeitpunkt noch nicht abgeschlossen, jedoch liegen vorsichtig optimistische Hinweise vor.

Um Datenlücken im Soll-Fahrplan vorab zu reduzieren, wurden für die Berichtsjahre 2023 und 2024 mehrere Datenabrufe aus der Mobilithek unterjährig jeweils vor und nach dem Fahrplanwechsel (in der Regel zweites Wochenende im Juni beziehungsweise Dezember)

durchgeführt. In der Praxis stellte sich heraus, dass Datenlücken im Jahresverlauf abnehmen. Die Prüfung wird daher mit der Jahresendlieferung für ein Berichtsjahr durchgeführt. Datenlücken entstehen, wenn Unternehmen (oder Verkehrsverbünde) Liniengebiete, Linien oder Haltestellen nicht oder unvollständig melden. Als Vergleichsdaten dienen das Zentrale Haltestellenverzeichnis sowie OpenStreetMap.

Im DELFI-Datensatz zeigt sich eine insgesamt hohe und detaillierte Abdeckung. In geringem Umfang werden zwischen 2023 und 2024 Datenlücken geschlossen. Auftretende Datenlücken betreffen relativ oft Informationen zu Linien von Omnibussen. Teilweise handelt es sich bei potenziellen Lücken auch um Haltestellen, die nicht mehr im Regelverkehr des öffentlichen Personenverkehrs angefahren werden. Eine finale Bewertung der Datenlücken steht zum aktuellen Zeitpunkt noch aus.

Einige wenige Verkehrsunternehmen und -verbünde übermittelten selbst erstellte Linien-Routings auf freiwilliger Basis, um die Qualität zu überprüfen. Des Wei-

teren werden öffentlich verfügbare und vollumfängliche Linien-Routings für einen großen Verkehrsverbund für die Berichtsjahre 2023 und 2024 genutzt. Ergebnisse des Abgleichs liegen zum Veröffentlichungszeitpunkt noch nicht vor.

5.3 Abgleich der regionalen Fahrleistung und der Linienlängen

Die Fahrleistungen und Linienlängen aus den Linien-Routings liegen bis auf Kreisebene für das Berichtsjahr 2023 vor. Für eine Einschätzung der Qualität wurden die beiden Merkmale mit Angaben aus der Personenverkehrsstatistik verglichen. Für das Merkmal der regionalen Fahrleistung ist ein passgenauer Vergleich möglich. Linienlängen werden in der Personenverkehrsstatistik allerdings ausschließlich für Straßenbahnen und Omnibusse auf Ebene der Bundesländer in der fünfjährigen Statistik erhoben, aktuell stehen sie für das Berichtsjahr 2019 zur Verfügung.

Auf Bundesebene zeigt sich, dass die Fahrleistung auf Basis der Linien-Routings mit den Angaben der Personenverkehrsstatistik nahezu übereinstimmt.

Beim Merkmal Linienlänge zeigen sich hingegen Abweichungen. So liegen die aufsummierten Linien-Kilometer auf Bundesebene nach Angaben der Personenverkehrsstatistik um bis zu 28 % unter den Gesamtergebnissen der Linien-Routings. Es ist anzunehmen, dass dies nicht einzig auf die zeitliche Differenz zwischen den verfügbaren Berichtsjahren zurückgeführt werden kann. Finale Analysen und Ergebnisse stehen noch aus.

Bei Vergleichen auf Ebene der Bundesländer zeigt sich, dass die Angaben zu Fahrleistungen überwiegend im ein- bis niedrigen zweistelligen Prozentbereich auseinanderliegen. Im Durchschnitt unterschreiten die Landeswerte für Bahnen und Omnibusse aus der amtlichen Statistik die ermittelten Fahrleistungen aus dem Linien-Routing. Bei Straßenbahnen hingegen sind die relativen Abweichungen im Durchschnitt leicht größer Null. Ob die amtliche Personenverkehrsstatistik den statischen Fahrplan unterschätzt, ist noch zu klären. Die Abweichungen nehmen auf Kreisebene weiter zu. Sowohl die Höhe als auch die Streuung der relativen Abweichungen nehmen auf Kreisebene weiter zu. Aufgrund der geografisch genauen Aufteilung beim Linien-Routing liegt die Vermu-

tung nahe, dass dieses zuverlässiger als die regionalen Meldungen zur Personenverkehrsstatistik ist. Auch hier stehen finale Ergebnisse noch aus.

6


Zusammenfassung und Ausblick

Neue Einflussfaktoren wie die Pandemie, die dauerhafte Änderung des Pendlerverhaltens durch die Verbreitung des Homeoffice sowie das Deutschlandticket erschweren es einem Teil der Verkehrsunternehmen und Verkehrsverbünden, ihre Fahrgastzahlen und die Beförderungsleistung verlässlich zu schätzen. Daher ist für die amtliche Statistik der Rückgriff auf Angaben von Unternehmen mit automatischen Fahrgastzählsystemen aussichtsreich. Basierend auf einer freiwilligen Erhebung konnten experimentelle AFZS-Indizes für Eisenbahnnah- und Straßenbahnverkehre entwickelt werden. Es zeigte sich, dass der AFZS-Index die Einführung des 9-Euro-Tickets und des Deutschlandtickets zuverlässig abbildet. Eine klare Stärke eines AFZS-Index ist seine Funktion als Pulsmesser, insbesondere in unsicheren Zeiten und bei weitreichenden Neuerungen im Personennahverkehr.

Seit dem dritten Berichtsquartal 2024 wird der [experimentelle AFZS-Index für Eisenbahnen und Straßenbahnen](#) als zusätzliches Produkt angeboten. Um auch künftig verlässlich und umfassend informieren zu können, soll der neue Index als ein langfristiges Produkt etabliert werden. Um dies zu erreichen, sind weitere Entwicklungsarbeiten und rechtliche Anpassungen vorzunehmen, damit eine Abfrage zum AFZS-Einsatz in das reguläre Erhebungsprogramm der Personenverkehrsstatistik eingehen kann.

In Hinblick auf die Linien-Routings sowie die regionale Ableitung von Linienlängen und planmäßigen Fahrleistungen wurde bereits vor Abschluss der finalen Auswertungen eine hohe Qualität in Umfang und Detailgrad festgestellt. Die ermittelten Linien-Routings und die statischen Fahrplandaten (DELFI, Mobilithek) bieten nach aktuellem Stand das Potenzial, die Qualität zu verbessern sowie die Befragten mit Blick auf regionale Meldungen zu entlasten. Weitere Analysen sind hier jedoch notwendig.

Einen Entwurf eines Gesetzes zur Bereitstellung und Nutzung von Mobilitätsdaten (Mobilitätsdatengesetz) und zur Änderung des Personenbeförderungsgesetzes (Drucksache 20/13954) hat die Bundesregierung im November 2024 vorgelegt. Eine erste Beratung im Deutschen Bundestag fand im Dezember 2024 noch vor den vorgezogenen Bundestagswahlen statt.¹⁸ Werden die beiden Gesetzesvorhaben umgesetzt, so könnten diese zu einer Steigerung der Datenqualität beitragen. Wann und in welcher Form die Gesetze implementiert werden, ist allerdings gegenwärtig offen.

Bisher war es nicht möglich, dynamische Fahrplandaten mit Störungen, Verspätungen und Betriebsausfällen zu betrachten. Da die Delegierte Verordnung (EU) 2024/490 ab Dezember 2028 die Meldung dynamischer Fahrplandaten für den öffentlichen Personennahverkehr vorsieht, könnten künftige Untersuchungen durch das Statistische Bundesamt weitere Potenziale in der Nutzung von Mobilitätsdaten anzeigen. 

18 Entwurf eines Gesetzes zur Bereitstellung und Nutzung von Mobilitätsdaten und zur Änderung des Personenbeförderungsgesetzes (Deutscher Bundestag, 2024a; Deutscher Bundestag, 2024b)

LITERATURVERZEICHNIS

Aberle, Christoph/Daubitz, Stephan/Schwedes, Oliver/Gertz, Carsten. *Measuring transport poverty with a mixed-methods approach. A comparative case study of the German cities Berlin and Hamburg*. In: Journal of Transport Geography. Volume 125 (2025). 104140. Mai 2025. DOI: [10.1016/j.jtrangeo.2025.104140](https://doi.org/10.1016/j.jtrangeo.2025.104140)

Bast, Hannah/Brosi, Patrick. *Sparse Map-Matching in Public Transit Networks with Turn Restrictions*. SIGSPATIAL '18: Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. November 2018, Seite 480 ff. DOI: [10.1145/3274895.3274957](https://doi.org/10.1145/3274895.3274957)

BeNaMo (Begleitforschung Nachhaltige Mobilität). *On-Demand-Verkehr*. [Zugriff am 8. April 2025]. Verfügbar unter: www.zukunft-nachhaltige-mobilitaet.de

Bierau, Dieter/Reim, Uwe. *Novellierung des Verkehrsstatistikgesetzes*. In: Wirtschaft und Statistik. Ausgabe 3/2004, Seite 259 ff.

BMDV (Bundesministerium für Digitales und Verkehr). *DELFI sorgt für Echtzeitdaten*. 24. August 2023. [Zugriff am 27. Februar 2025]. Verfügbar unter: mobilithek.info/blog/delfi-daten-mobilithek

BMDV (Bundesministerium für Digitales und Verkehr). *Verkehr in Zahlen 2024/2025*. 53. Jahrgang. 2024. [Zugriff am 27. Februar 2025]. Verfügbar unter: bmdv.bund.de

BMG (Bundesministerium für Gesundheit). *Coronavirus-Pandemie: Was geschah wann? Chronik aller Entwicklungen im Kampf gegen COVID-19 (Coronavirus SARS-CoV-2) und der dazugehörigen Maßnahmen des Bundesgesundheitsministeriums*. 14. Februar 2023. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.bundesgesundheitsministerium.de

Dangschat, Jens S. *Verkehrswende – sozial und räumlich ausgewogen*. In: Journal für Mobilität und Verkehr. Ausgabe 14/2022. DOI: [10.34647/jmv.nr14.id87](https://doi.org/10.34647/jmv.nr14.id87)

Deutscher Bundestag. Dokumentations- und Informationssystem für Parlamentsmaterialien. Vorgang – Gesetzgebung. *Gesetz zur Bereitstellung und Nutzung von Mobilitätsdaten und zur Änderung des Personenbeförderungsgesetzes*. 20. Wahlperiode. 1. Durchgang 11. Oktober 2024. [Zugriff am 10. Februar 2025]. Verfügbar unter: dip.bundestag.de

Deutscher Bundestag. Kurzmeldungen. Verkehr – Gesetzentwurf – hib 822/2024. *Regierungsentwurf für ein Mobilitätsdatengesetz liegt vor*. 28. November 2024. [Zugriff am 10. Februar 2025]. Verfügbar unter: www.bundestag.de

Haas, Paul William. *Aspekte des Luftverkehrs an deutschen Flughäfen seit 2019*. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2023, Seite 72 ff.

Stark, Kerstin/Kehlbacher, Ariane/Mattioli, Giulio. *Mobilitätsarmut und soziale Teilhabe in Deutschland*. Studie für Agora Verkehrswende 2023. Deutsches Institut für Luft- und Raumfahrt. [Zugriff am 8. April 2025]. Verfügbar unter: www.agora-verkehrswende.de

LITERATURVERZEICHNIS

Statistisches Bundesamt. *Fahrgastzahl im Linienverkehr mit Bussen und Bahnen 2021 auf neuem Tiefststand*. Pressemitteilung Nr. 152 vom 7. April 2022. [Zugriff am 26. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Knapp ein Viertel aller Erwerbstätigen arbeitete 2022 im Homeoffice*. Zahl der Woche Nr. 28 vom 11. Juli 2023. [Zugriff am 26. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Fahrgastzahl im Linienverkehr mit Bussen und Bahnen 2023 um 7 % gestiegen*. Pressemitteilung Nr. 142 vom 8. April 2024. 2024a. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Homeoffice 2023 ähnlich verbreitet wie im Vorjahr, wird jedoch an weniger Tagen genutzt*. Pressemitteilung Nr. N032 vom 26. Juni 2024. 2024b. [Zugriff am 24. Februar 2025]. Verfügbar unter: www.destatis.de

VDV (Verband Deutscher Verkehrsunternehmen). *Automatische Fahrgastzählsysteme – Handlungsempfehlungen zur Anwendung von AFZS im öffentlichen Personenverkehr, Version 2.2*. VDV-Schrift 457, Ausgabe 4/2023. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.vdv.de

VDV (Verband Deutscher Verkehrsunternehmen). *Interpretierende Zusammenfassung – Kernaussagen der bundesweiten Marktforschung zum Deutschland-Ticket*. Berichtszeitraum 1. Halbjahr 2024. [Zugriff am 27. Februar 2025]. Verfügbar unter: www.vdv.de

RECHTSGRUNDLAGEN

Delegierte Verordnung (EU) 2017/1926 der Kommission vom 31. Mai 2017 zur Ergänzung der Richtlinie 2010/40/EU des Europäischen Parlaments und des Rates hinsichtlich der Bereitstellung EU-weiter multimodaler Reiseinformationsdienste (Amtsblatt der EU Nr. L 272, Seite 1).

Delegierte Verordnung (EU) 2024/490 der Kommission vom 29. November 2023 zur Änderung der Delegierten Verordnung (EU) 2017/1926 zur Ergänzung der Richtlinie 2010/40/EU des Europäischen Parlaments und des Rates hinsichtlich der Bereitstellung EU-weiter multimodaler Reiseinformationsdienste (Amtsblatt der EU Nr. L, Seite 1).

Gesetz über die Statistik der See- und Binnenschifffahrt, des Güterkraftverkehrs, des Luftverkehrs sowie des Schienenverkehrs und des gewerblichen Straßen-Personenverkehrs (Verkehrsstatistikgesetz – VerkStatG) vom 20. Februar 2004 (BGBl. I Seite 318), das zuletzt durch Artikel 1 des Gesetzes vom 16. August 2023 (BGBl. I Nr. 218) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Gesetz über Finanzhilfen des Bundes zur Verbesserung der Verkehrsverhältnisse der Gemeinden (Gemeindeverkehrsfinanzierungsgesetz – GVFG) in der Fassung der Bekanntmachung vom 28. Januar 1988 (BGBl. I Seite 100), das zuletzt durch Artikel 323 der Verordnung vom 19. Juni 2020 (BGBl. I Seite 1328) geändert worden ist.

Mobilitätsdatenverordnung (MDV) vom 20. Oktober 2021 (BGBl. I Seite 4728), die zuletzt durch Artikel 1 der Verordnung vom 1. Juli 2022 (BGBl. I Seite 1039) geändert worden ist.

Personenbeförderungsgesetz (PBefG) in der Fassung der Bekanntmachung vom 8. August 1990 (BGBl. I Seite 1960), das zuletzt durch Artikel 7 Absatz 4 des Gesetzes vom 11. April 2024 (BGBl. I Nr. 119) geändert worden ist.

Verordnung (EG) Nr. 1370/2007 des Europäischen Parlaments und des Rates vom 23. Oktober 2007 über öffentliche Personenverkehrsdienste auf Schiene und Straße und zur Aufhebung der Verordnung (EWG) Nr. 1191/69 und (EWG) Nr. 1107/70 des Rates (Amtsblatt der EU Nr. L 315, Seite 1).

LARGE CASES UNIT: ERKENNTNISSE AUS FÜNF JAHREN ARBEIT

Klaus Pötzsch, Maximilian Rotsche

➤ **Schlüsselwörter:** Multinationale Unternehmensgruppen – Volkswirtschaftliche Gesamtrechnungen – Wirtschaftsstatistiken – Kohärenz – Qualitätssicherung

ZUSAMMENFASSUNG

Die Large Cases Unit (LCU) der Statistischen Ämter des Bundes und der Länder hat im Jahr 2020 ihre Arbeit aufgenommen. Sie untersucht seitdem große multinationale Unternehmensgruppen und ihre Untergliederungen auf Kohärenz der in der amtlichen Statistik vorliegenden Daten. Einbezogen in die Prüfungen werden auch externe Datenquellen. Die Aufklärung erkannter Inkohärenzen dient dazu, die Qualität der Volkswirtschaftlichen Gesamtrechnungen und der zugrunde liegenden Wirtschaftsstatistiken zu sichern. Der Artikel beschreibt Fallauswahl, Datenquellen und Kohärenzprüfungen und informiert über die wichtigsten Erkenntnisse der bisherigen LCU-Arbeit.

➤ **Keywords:** multinational enterprise groups – national accounts – economic statistics – consistency – quality assurance

ABSTRACT

The Large Cases Unit (LCU) of the statistical offices of the Federation and the Länder commenced work in 2020. Since then, the LCU has been analysing large multinational enterprise groups (MNEs) and their subdivisions to check the consistency of the data available in official statistics. The LCU also uses data from external data sources to perform the checks. The clarification of inconsistencies that are identified helps ensure the quality of the national accounts and the underlying economic statistics. This article describes how cases are selected, the data sources used and the consistency checks performed, and provides information on the key findings of the work of the LCU so far.



Klaus Pötzsch

ist Diplom-Ökonom und leitet seit 2020 das Referat „Large Cases Unit, Globalisierung“ des Statistischen Bundesamtes.



Maximilian Rotsche

studierte Betriebswirtschaftslehre (M. Sc.) an der Rheinisch-Westfälischen Technischen Hochschule Aachen. Im Referat „Large Cases Unit, Globalisierung“ des Statistischen Bundesamtes umfassen seine Arbeitsschwerpunkte die Bearbeitung der Fallanalysen sowie die methodische Weiterentwicklung der Large Cases Unit.

1

Einleitung

Die Aufgabe der deutschen Large Cases Unit (LCU) ist, die Qualität der Volkswirtschaftlichen Gesamtrechnungen und der zugrunde liegenden Wirtschaftsstatistiken hinsichtlich der konzepttreuen und kohärenten Einbeziehung statistischer Daten multinationaler Unternehmensgruppen zu sichern. Die LCU beschäftigt sich ausschließlich mit solchen Unternehmensgruppen, deren Bedeutung so groß ist, dass potenzielle Fehler in den Daten bedeutsame Auswirkungen auf volkswirtschaftliche Aggregate haben (Ahlborn und andere, 2021).

Die deutsche LCU hat ihre Arbeit im Januar 2020 aufgenommen. Vorausgegangen waren eine Machbarkeitsstudie sowie eine sogenannte Systemsimulation, die ausgewählte multinationale Unternehmensgruppen untersuchte, um festzustellen, welche rechtlichen, organisatorischen, fachlich-methodischen und technischen Voraussetzungen für eine dauerhafte LCU-Arbeit erforderlich sind. Parallel zur ständigen Ausweitung der Fallanalysen¹ wurden seitdem folgende Voraussetzungen geschaffen:

- › rechtlich: Das im Juni 2021 verabschiedete Gesetz zur Prüfung von Daten multinationaler Unternehmensgruppen zur Sicherung der Qualität der Volkswirtschaftlichen Gesamtrechnungen und der Wirtschaftsstatistiken regelt die Lieferung von Mikrodaten an das Statistische Bundesamt, die Mikrodatenverknüpfung beim Statistischen Bundesamt, den Austausch verknüpfter Mikrodaten zwischen dem Statistischen Bundesamt, den Statistischen Ämtern der Länder und der Deutschen Bundesbank sowie die Mitwirkungspflicht von Unternehmensgruppen bei der Aufklärung festgestellter Inkohärenzen.
- › organisatorisch: Ein 2024 aktualisiertes Organisationsmodell beschreibt die Zusammenarbeit zwischen den Statistischen Ämtern des Bundes und der Länder. Die Kooperation der LCU mit der Deutschen Bundesbank ist in einer Zusatzvereinbarung zum bestehenden Memorandum of Understanding festgelegt.

- › fachlich-methodisch: Theoretische Grundlagen, Beschreibungen der den LCU-Arbeiten zugrunde liegenden Basisstatistiken sowie die Methodik der Kohärenzprüfungen enthält ein LCU-Leitfaden der Fallbearbeitung. Daneben wurde ein Manual zu Grundlagen der betrieblichen und Konzern-Rechnungslegung erstellt.
- › technisch: In der durch die LCU erstellten IT-Fachanwendung werden in einer SAS-Datenbank die Daten der einzelnen Basisstatistiken gesammelt, verknüpft und auf der Ebene von Unternehmensgruppen zusammengeführt. Mithilfe einer Excel-Anwendung werden die Daten je Unternehmensgruppe für die Fallbearbeitung zur Verfügung gestellt.

Die deutsche LCU ist eine Gemeinschaftsaufgabe des Statistischen Bundesamtes und der Statistischen Ämter der Länder. Derzeit arbeiten die Landesämter für Hamburg und Schleswig-Holstein (Statistikamt Nord), von Nordrhein-Westfalen (IT.NRW), Baden-Württemberg sowie Bayern direkt an der Fallbearbeitung mit; an der Aufklärung festgestellter Inkohärenzen sowie an der Durchführung von Unternehmensgruppenkontakten sind die Fachbereiche aller Statistischen Landesämter (neben den Fachbereichen im Statistischen Bundesamt) beteiligt.

Der Aufsatz beschreibt in Kapitel 2 zunächst einige methodische Grundlagen der LCU-Arbeit. Kapitel 3 stellt wesentliche Erkenntnisse der bisherigen LCU-Arbeit vor, und zwar sowohl quantitative Ergebnisse als auch wichtige methodische Schlussfolgerungen. Das letzte Kapitel gibt einen kurzen Ausblick über anstehende Weiterentwicklungen.

¹ Zu den Fortschritten bei den Fallanalysen siehe zum Beispiel Hörner und andere (2022).

2

Methodische Grundlagen der LCU-Arbeit

Untersuchungsgegenstand der LCU sind multinationale Unternehmensgruppen und ihre assoziierten Einheiten (Organschaften, Rechtliche Einheiten, Niederlassungen)². Das Prinzip der LCU-Arbeit besteht darin, alle verfügbaren Daten zusammenzuführen, zu verknüpfen, auf Kohärenz zu untersuchen, festgestellte Inkohärenzen aufzuklären und die Erkenntnisse an die Volkswirtschaftlichen Gesamtrechnungen sowie die datenhaltenden Fachbereiche zurückzuspielen. ➤ Grafik 1

Eine Beschreibung von Analyse, Aufklärung und Verwertung erfolgt hier nicht, die Abläufe sind bei Ahlborn und andere (2021, Seite 34 ff.) dargestellt. Dieser Artikel geht lediglich auf Fallauswahl, Beschreibung der Datenquellen und der Kohärenzprüfungen näher ein.

² Erläuterungen zu den einzelnen Einheiten enthält zum Beispiel der [Qualitätsbericht des Statistischen Unternehmensregisters](#), Punkt 2.1.3 (Statistisches Bundesamt, 2024).

2.1 Fallauswahl

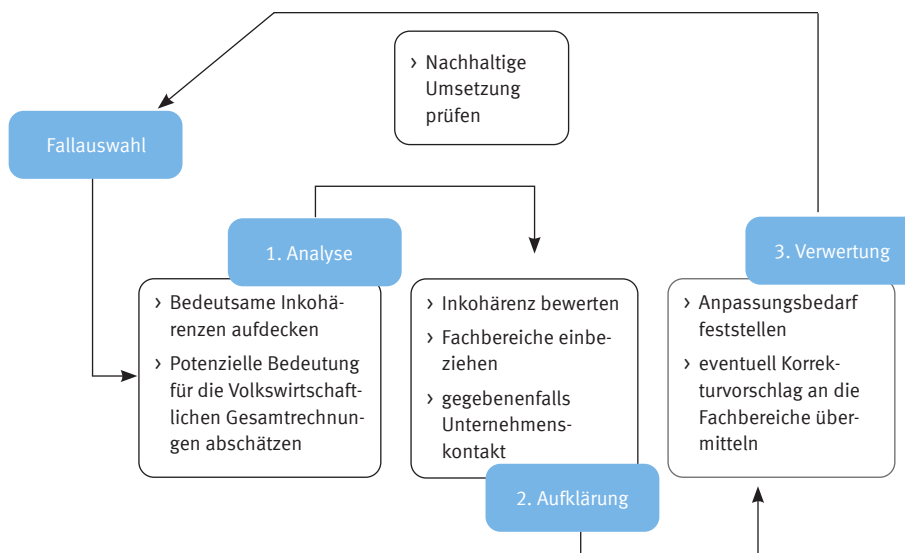
Das statistische Unternehmensregister enthält rund 52 000 multinationale Unternehmensgruppen. Da es nicht möglich ist, all diese Gruppen zu untersuchen, ist eine Priorisierung notwendig. Grundlage für die Auswahl ist ein risikobasierter Ansatz: Bei welchen Unternehmensgruppen besteht das größte Risiko, dass

- › Veränderungen (zum Beispiel Umstrukturierungen wie Abspaltungen, Verschmelzungen, Anwachsungen; Verlagerungen von wirtschaftlichen Tätigkeiten; Änderungen in den Geschäftsmodellen) nicht oder nicht rechtzeitig adäquat in den statistischen Meldedaten berücksichtigt werden?
- › Fehler in den statistischen Meldedaten zu bedeutenden Fehlern in volkswirtschaftlichen Aggregaten führen?

Um die Unternehmensgruppen mit dem größten Risikopotenzial zu ermitteln, wurde der sogenannte LCU-Score entwickelt. Die Grundgesamtheit besteht dabei zunächst aus Rechtlichen Einheiten im statistischen Unternehmensregister, die einer multinationalen Unternehmensgruppe angehören (das heißt, denen ein Gruppenoberhaupt zugeordnet werden kann). Diesen Einheiten werden statistische Daten zugespielt, die auf Gruppenebene aggregiert werden. Die Daten stammen

Grafik 1

Arbeitsweise der deutschen Large Cases Unit



aus dem statistischen Unternehmensregister selbst oder aus statistischen Erhebungen. Sie repräsentieren drei Dimensionen mit insgesamt sieben Kennziffern:

- › Gruppengröße – Zahl der Rechtlichen Einheiten, tätige Personen
- › Wirtschaftsleistung – Umsatz, Umsatzdominanz innerhalb eines WZ-2-Stellers³, Bruttowertschöpfung

› Außenwirtschaftliche Aktivitäten – Außenhandelsvolumen, Volumen der für ausländische Auftraggeber erbrachten Dienstleistungen aus den Umsatzsteuer-Voranmeldungen

Um Ausreißer in einzelnen Kennziffern auszugleichen, werden die aggregierten Werte mit einer Sigmoid-Funktion⁴ geglättet. Dadurch erhält man für jede Unterneh-

3 Abteilungen der Klassifikation der Wirtschaftszweige, Ausgabe 2008, künftig Ausgabe 2025.

4 Zur mathematischen Beschreibung der Sigmoid-Funktion siehe zum Beispiel: mathworld.wolfram.com

Übersicht 1

Wichtigste Datenquellen der Large Cases Unit

	Abkürzung	EVAS-Nr. ¹	Merkmalsträger	Datenhalter
Amtliche Statistiken²				
Statistisches Unternehmensregister	URS	52111	Unternehmensgruppen, Rechtliche Einheiten, Niederlassungen	Statistisches Bundesamt, Statistische Ämter der Länder
Jahreserhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe	JBU	42221	Rechtliche Einheiten	Statistische Ämter der Länder
Investitionserhebung im Verarbeitenden Gewerbe	IE	42231	Rechtliche Einheiten, Niederlassungen	Statistische Ämter der Länder
Monatliche und Vierteljährliche Produktionserhebung im Verarbeitenden Gewerbe	MVP	42121 42131	Niederlassungen	Statistische Ämter der Länder
Monatsbericht für Betriebe im Verarbeitenden Gewerbe	MB	42111	Niederlassungen	Statistische Ämter der Länder
Kostenstrukturerhebungen im Verarbeitenden Gewerbe, in der Energieversorgung und im Baugewerbe	KSE	42251 43221 44253/54	Rechtliche Einheiten	Statistisches Bundesamt
Strukturerhebung im Handel und Dienstleistungsbereich (ab 2021) bis 2020: Jahreserhebung im Handel Strukturstatistik im Dienstleistungsbereich	SHD JEG SID	47410 45341 47415	Rechtliche Einheiten	Statistisches Bundesamt, Statistische Ämter der Länder
Umsatzsteuerstatistik (Voranschläge)	UVA	73311	Umsatzsteuerpflichtige Unternehmen (Rechtliche Einheiten oder steuerliche Organschaften)	Statistische Ämter der Länder
Außenhandelsstatistik (Intra-/Extrahandel)	AH	51141 51231	Importeure/Exporteure (Rechtliche Einheiten oder steuerliche Organschaften)	Statistisches Bundesamt
Externe Datenquellen				
Zahlungsbilanzstatistik ³	ZB	–	Institutionelle Einheiten nach § 63 Außenwirtschaftsverordnung	Deutsche Bundesbank
Erhebung zu Forschung und Entwicklung ⁴	FE	–	Rechtliche Einheiten oder Gruppen von Rechtlichen Einheiten	Stifterverband der deutschen Wissenschaft
Country by Country Reporting ⁵	CbCR	–	Unternehmensgruppen	Bundeszentralamt für Steuern
Mehrwertsteuer-Informationsaustauschsystem ⁶	VIES	–	Rechtliche Einheiten	Bundeszentralamt für Steuern
Geschäftsberichte von Konzernen und Rechtlichen Einheiten	GB	–	Unternehmensgruppen (Konzerne), Rechtliche Einheiten	öffentlich

1 Das „Einheitliche Verzeichnis aller Statistiken des Bundes und der Länder“ (EVAS) enthält alle Statistiken (Erhebungen, Rechenwerke und Register), die vom Statistischen Bundesamt und den Statistischen Ämtern der Länder betreut werden.

2 Für alle amtlichen Statistiken existieren standardisierte [Qualitätsberichte](http://www.bundesbank.de), die unter anderem die Erhebungsmerkmale und methodische Beschreibungen enthalten.

3 Informationen zur Methodik der Zahlungsbilanzstatistik: www.bundesbank.de

4 Informationen zur Erhebung zu Forschung und Entwicklung: www.stifterverband.org

5 Informationen zum Country-by-Country-Reporting: www.bzst.de

6 Informationen zur Erhebung der Zusammenfassenden Umsatzsteuermeldungen, die den VIES-Daten zugrunde liegen: www.bzst.de

mensgruppe und jede der sieben Kennziffern einen Teilscore mit dem Maximalwert 1 000 für die jeweils bedeutendste Gruppe. Die sieben Teilscores werden in einem letzten Schritt gewichtet und anschließend zu einem Gesamtscore addiert. Der Gesamtscore in absteigender Sortierung priorisiert die Unternehmensgruppen.

Neben dem LCU-Score können andere Kriterien ausschlaggebend sein. So können die Fachstatistiken oder die Deutsche Bundesbank Unternehmensgruppen, die bei Erhebungen immer wieder größere Schwierigkeiten bereiten, für eine prioritäre Bearbeitung durch die LCU vorschlagen. Auch äußere Anforderungen, wie im Zusammenhang mit dem Bruttonationaleinkommen-Vorbehalt zur Globalisierung (Draken und andere, 2023), können entscheidend sein.

2.2 Datenquellen der LCU

Datenquellen der LCU sind in erster Linie Einzeldaten amtlicher Statistiken. Hierzu zählen Informationen aus der Unternehmensgruppenverarbeitung (Andom und andere, 2025) und dem Profiling von Unternehmen (Redecker und andere, 2021), die die LCU insbesondere bei der Gruppenstrukturanalyse verwendet.

Daneben werden auch Daten anderer Statistikproduzenten verarbeitet, die entweder dem Statistischen Bundesamt bereits als Sekundärdatenquelle zur Verfügung stehen oder diesem speziell für LCU-Zwecke geliefert werden. Eine herausragende Bedeutung haben außerdem die Geschäftsberichte der Konzerne und Einzelunternehmen. Die darin enthaltenen allgemeinen Informationen (beispielsweise zum Geschäftsmodell oder zu Umstrukturierungen), die Bilanz, die Gewinn- und Verlustrechnung sowie die Erläuterungen dazu in den Anhängen der Berichte fließen maßgeblich in die Kohärenzprüfungen der LCU ein. ➤ Übersicht 1

Vor den eigentlichen LCU-Analysen werden die Mikrodaten je Unternehmensgruppe zusammengeführt und miteinander verknüpft. Dieser – inzwischen weitgehend automatisierte – Prozess ist besonders anspruchsvoll, da den Datenquellen unterschiedliche Einheitenkonzepte zugrunde liegen (siehe Spalte „Merkmalsträger“ in Übersicht 1). Angestrebt wird zumeist, verknüpfte Datensätze auf der Ebene der Rechtlichen Einheiten zu generieren. Dies kann beispielsweise durch Aggregation von Niederlassungsdaten geschehen. Liegen Daten nur

auf einer höheren Aggregationsstufe vor (zum Beispiel Außenhandelsdaten oder Daten der Umsatzsteuer-Voranmeldung auf der Ebene von Organschaften), werden die Kohärenzprüfungen auch auf dieser Ebene durchgeführt. Dabei lassen Zusatzinformationen (zum Beispiel aus der Produktionsstatistik oder aus den Geschäftsberichten in Verbindung mit Informationen aus dem statistischen Unternehmensregister zur Zusammensetzung der Organkreise) häufig belastbare Rückschlüsse auf Rechtliche Einheiten zu.

2.3 Überblick über wichtige Kohärenzprüfungen

Während der Analysephase werden die verknüpften Datensätze dahingehend überprüft, ob zwischen den aus unterschiedlichen Quellen stammenden einzelnen Kennziffern Widersprüche bestehen und ob diese Inkohärenzen bedeutsam sind, das heißt, ob Fehler in einer der Quellen spürbare Auswirkungen auf Gesamtergebnisse hätten. Dafür wurde eine Reihe von standardisierten Kohärenzprüfungen entwickelt. Es ist jedoch wichtig zu betonen, dass die Analyse einer Unternehmensgruppe weit über die bloße Abarbeitung dieser Prüfungen hinausgeht. Entscheidend ist immer die ganzheitliche Betrachtung der Gruppe, ihrer Geschäftsmodelle und Strukturen sowie deren Veränderungen im Zeitverlauf.

Vor den eigentlichen Kohärenzchecks ist es immer erforderlich, die Gruppenstruktur im Zeitverlauf zu prüfen. Was ist das generelle Geschäftsmodell einer Unternehmensgruppe oder einzelner Glieder und welche Implikationen hat das auf die statistischen Meldungen? Gab es Abspaltungen von Gruppenteilen oder Teilen von Rechtlichen Einheiten, die zwangsläufig zu einer Veränderung einzelner Kennziffern führen? Gab es Zukäufe oder Verschmelzungen? Gab es Änderungen in den Geschäftsmodellen, die zu einem Wechsel des Wirtschaftszweiges und damit gegebenenfalls zu einer veränderten Berichtspflicht zu einzelnen Statistiken führen? Wurde dies im statistischen Unternehmensregister und den Statistiken umgesetzt?

Nach dem Abschluss der Gruppenstrukturanalyse beginnen die eigentlichen Kohärenzprüfungen.¹⁵ Die Themen

5 Die einzelnen Prüfungen sowie die methodischen Grundlagen dazu beschreibt ein Leitfaden der LCU-Fallbearbeitung.

Übersicht 2

Wichtige Kohärenzprüfungen der LCU-Analysen nach Datenquellen¹

Prüfinhalt	URS	JBU	IE	MVP	MB	KSE	SHD	UVA	AH	ZB	FE	CbCR	VIES	GB
Umsatzhöhe	X	X			X	X	X	X				X		X
Umsatz (Struktur)		X		X	X	X	X							X
Umsatz und Produktion		X		X	X	X								X
Bruttowertschöpfung						X	X							X
Beschäftigte	X	X				X	X							X
Lohnveredelung				X	X	X	X	X	X	X		X	X	X
Transithandel						X	X	X		X			X	X
Geistiges Eigentum			X			X	X			X	X	X		X
Leasing			X			X	X		X	X		X		X

¹ Die Langfassungen der hier abgekürzten Datenquellen sind Übersicht 1 zu entnehmen.

und Datenquellen wichtiger Kohärenzprüfungen zeigt [Übersicht 2](#).

Sind Daten inkohärent zueinander, schlagen in der Regel mehrere Kohärenzchecks an. Deshalb sind beispielsweise die Prüfungen zur Umsatzhöhe und -struktur, zu Umsatz und Produktion, zur Bruttowertschöpfung und gegebenenfalls zur Lohnveredelung immer im Zusammenhang zu betrachten.

Im Folgenden werden drei Kohärenzchecks exemplarisch näher beschrieben.

Beispiel 1: Prüfung der Umsatzhöhe einer Rechtlichen Einheit

Zum Gesamtumsatz (U) einer Rechtlichen Einheit stehen der LCU verschiedene Quellen zur Verfügung.⁶ Grundsätzlich sind die Zahlen zueinander kohärent, wenn gilt:

Verarbeitendes Gewerbe: $U(\text{JBU}) \approx U(\text{KSE}) \approx U(\text{UVA}) \approx U(\text{GB}) (\approx U(\text{URS}))$
 Handel/Dienstleistungen: $U(\text{SHD}) \approx U(\text{UVA}) \approx U(\text{GB}) (\approx U(\text{URS}))$

Ist dies nicht erfüllt, liegt eine Inkohärenz vor. In der Aufklärungsphase wird zunächst festgestellt, ob die Inkohärenz aufgrund unterschiedlicher Merkmalsdefinitionen in den verschiedenen Quellen erklärbar ist. So kann es sein, dass die eine Quelle (KSE) den Umsatz entsprechend dem nationalen Handelsgesetzbuch-Abschluss meldet und die nächste Quelle (JBU) nach internationalen Rechnungslegungsvorschriften (IFRS – International

Financial Reporting Standards). Im Geschäftsbericht wiederum können die Umsatzangaben der Gewinn- und Verlustrechnung Umsätze ausländischer Zweigniederlassungen enthalten, die in den Meldungen zur amtlichen Statistik (KSE, JBU und SHD) auszuschließen sind. Im Zuge der Aufklärung wird letztlich ermittelt, welche Daten entsprechend der definitorischen Vorgaben der einzelnen Quellen richtig und welche Daten falsch und damit korrekturbedürftig sind. Bei dieser Kohärenzprüfung wird auch untersucht, ob die im statistischen Unternehmensregister verwendete Quelle für das Merkmal Umsatz geändert werden sollte.

Beispiel 2: Prüfung von Umsatz und Produktion einer Rechtlichen Einheit im Verarbeitenden Gewerbe

Eine Rechtliche Einheit im Verarbeitenden Gewerbe kann Umsätze aus unterschiedlichen Aktivitäten generieren: aus eigenen Erzeugnissen (im eigenen Produktionsprozess hergestellte Güter), aus Handelsware (Güter, die eingekauft und in unverändertem Zustand weiterverkauft werden) oder aus sonstigen Tätigkeiten (zum Beispiel Umsätze aus Vermietung, Lizenzerlöse, Erlöse aus nicht industriellen Dienstleistungen). Die LCU prüft, ob der Umsatz aus eigenen Erzeugnissen und der zur Produktionsstatistik gemeldete (über die einzelnen Berichtszeiträume und Niederlassungen aggregierte) Absatzproduktionswert unter Berücksichtigung von Bestandsveränderungen an fertigen und unfertigen Erzeugnissen zueinander passen. Grundsätzlich wird von einer Kohärenz ausgegangen, wenn gilt:

Umsatz eigene Erzeugnisse (KSE, JBU) \approx Absatzproduktionswert (MVP) \pm Bestandsveränderungen (KSE)

⁶ Hier und im weiteren Verlauf steht in der Klammer immer die Datenquelle (siehe Übersicht 1).

Ist der Umsatz deutlich kleiner als Produktion \pm Bestandsveränderungen, handelt es sich in jedem Fall um eine Inkohärenz, das heißt mindestens einer der drei Werte kann nicht korrekt sein. Wenn der Umsatz eigener Erzeugnisse deutlich größer ist, ist zu untersuchen, ob die Rechtliche Einheit passive Lohnarbeit/Lohnveredelung⁷ betreibt, also als Auftraggeber für Lohnarbeit/Lohnveredelung fungiert. In diesem Fall hätte sie den Verkauf der in Lohnarbeit erzeugten Produkte als Umsatz eigener Erzeugnisse zu melden, während keine Meldung zur Absatzproduktion erfolgt. Kann Lohnarbeit ausgeschlossen werden, handelt es sich in der Regel um eine Inkohärenz.

Beispiel 3: Prüfung auf aktive Lohnarbeit/Lohnveredelung einer Rechtlichen Einheit im Verarbeitenden Gewerbe

Die Prüfung der adäquaten Erfassung von Lohnarbeit ist komplex, vor allem wenn die entsprechenden Transaktionen grenzüberschreitend sind. Eigentumsübergänge richtig darzustellen ist jedoch von hoher Bedeutung insbesondere für die Volkswirtschaftlichen Gesamtrechnungen. Als Beispiel dient hier ein Sachverhalt, bei dem der Auftraggeber im EU-Ausland sitzt und der Auftragnehmer (= aktiver Lohnveredeler) eine deutsche Rechtliche Einheit ist. Diese Konstellation kommt in Branchen wie der Mineralölverarbeitung, der chemischen Industrie, der Pharmaindustrie und der Lebensmittelindustrie relativ häufig vor – insbesondere, wenn die deutsche Rechtliche Einheit Teil eines international agierenden Konzerns mit ausländischem Gruppenoberhaupt ist.

Die Prüfung auf aktive Lohnveredelung beginnt mit der Klärung der Frage, ob diese tatsächlich vorliegt. Verschiedene Anhaltspunkte dafür liefern die statistischen Daten selbst, wobei diese oft widersprüchlich sind. Auch weitere Datenquellen (wie UVA, VIES oder CbCR) liefern Hinweise. Verlässliche Informationen sind jedoch in der Regel in den Geschäftsberichten des Konzerns und der Rechtlichen Einheit selbst zu finden. Endgültige Gewissheit gibt oft erst der direkte Kontakt mit der Unternehmensgruppe. Steht zweifelsfrei fest, dass es sich bei der deutschen Rechtlichen Einheit um einen aktiven Lohnveredeler mit einem ausländischen Auftraggeber handelt, müssen Kennziffern aus unterschiedlichen Quel-

len zusammenpassen.⁸ Die statistischen Angaben der deutschen Rechtlichen Einheit ergeben ein kohärentes Bild, wenn folgende Bedingungen erfüllt sind:

- (1) Umsatz eigene Erzeugnisse (KSE) \approx (Absatzproduktionswert (MVP) = Absatzproduktionswert in Lohnarbeit (MVP)) \approx Einnahmen aus Lohnfertigung (ZB)
- (2) Kosten für Roh-, Hilfs- und Betriebsstoffe (KSE) ≈ 0 und Bestandsveränderungen (KSE) ≈ 0
- (3) Importe zur Lohnveredelung (AH) und Exporte nach Lohnveredelung (AH) > 0 ; Importe und Exporte mit Eigentumsübergang ≈ 0
- (4) Nicht steuerbare sonstige Leistungen (UVA) \geq Absatzproduktionswert (MVP)
- (5) Kosten für Einkauf von Handelsware (KSE) \approx Ausgaben im sonstigen Warenverkehr (ZB)
- (6) Umsatz mit Handelsware (KSE) > 0 (in der Regel \geq Kosten für Einkauf von Handelsware (KSE))

Sobald mindestens eine der Bedingungen nicht erfüllt ist, liegt eine Inkohärenz vor.

3

Erkenntnisse der bisherigen LCU-Arbeit

3.1 Fallzahlen

Die LCU untersucht derzeit (Stand: Dezember 2024) rund 110 multinationale Unternehmensgruppen mit etwa 7 100 Rechtlichen Einheiten, in denen fast 3 Millionen Personen beschäftigt sind. Die Rechtlichen Einheiten der von der LCU geprüften Unternehmensgruppen erwirtschafteten im Jahr 2022 einen Umsatz von knapp 2,8 Billionen Euro und ein Außenhandelsvolumen (Summe Exporte und Importe) von 1,1 Billionen Euro. Die Anzahl der von der LCU betrachteten Gruppen erscheint auf den ersten Blick gering. Ihre Bedeutung für die deutsche Wirtschaft insgesamt ist jedoch erheblich.

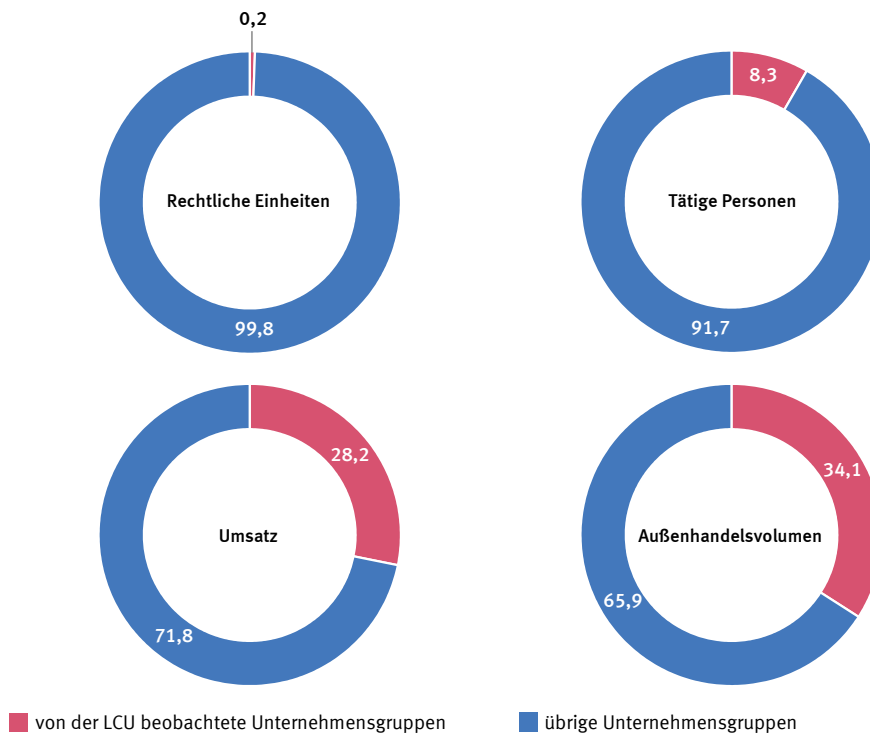
⁷ Zur Definition von Lohnarbeit/Lohnveredelung siehe Hörner und andere (2022; hier: Seite 35).

⁸ Vereinfachend wird hier angenommen, dass die deutsche Rechtliche Einheit ausschließlich aktiver Lohnveredeler ist, das heißt keine Produktion auf eigene Rechnung durchführt. Wie in der Praxis häufig anzutreffen, übernimmt sie jedoch den Vertrieb eines Teiles der in Lohnarbeit erzeugten Waren in Deutschland, der andere Teil wird ins Ausland geliefert. Der Auftraggeber ist ein Unternehmen mit Sitz in einem EU-Land.

Grafik 2

Bedeutung der von der Large Cases Unit (LCU) beobachteten multinationalen Unternehmensgruppen

Anteile an ausgewählten Kennzahlen in %



Die 110 LCU-Unternehmensgruppen machen lediglich 0,03 % der deutschen Unternehmensgruppen aus, ihre Rechtlichen Einheiten nur 0,2 % aller im statistischen Unternehmensregister enthaltenen Rechtlichen Einheiten. In ihnen sind jedoch 8 % aller Beschäftigten tätig, ihr Umsatzanteil beträgt mehr als 28 % und sie stehen für 34 % des deutschen Außenhandelsvolumens.

➡ Grafik 2

Bemerkenswert ist, dass bei der überwiegenden Anzahl der untersuchten Unternehmensgruppen bedeutsame Inkohärenzen festgestellt wurden. Lediglich 30 % der untersuchten Gruppen waren vollständig kohärent oder die Inkohärenzen waren marginal. Es fällt auf, dass in den meisten Fällen die Inkohärenzen nicht intern aufgeklärt werden konnten, sodass ein direkter Kontakt mit der Unternehmensgruppe erforderlich war. Der erste direkte Austausch mit einer multinationalen Unternehmensgruppe hat im März 2021 stattgefunden, seitdem wurden Gespräche mit 50 weiteren geführt. Mit einer Reihe von Gruppen wurden bereits mehrere Termine vereinbart, weil nach dem ersten Gespräch noch Fragen offengeblieben waren. Ein anderer Grund war, dass nach

Anfügen einer neuen Jahresscheibe in die Analysedaten neue Inkohärenzen aufgetreten waren. Allgemein ist die LCU auch daran interessiert, zumindest mit den größten Unternehmensgruppen regelmäßige Kontakte zu pflegen. Durch die Unternehmensgruppenkontakte konnte der Großteil der Inkohärenzen aufgeklärt und eine nachhaltige Veränderung im Meldeverhalten vereinbart werden, was zu einer Qualitätssteigerung der Primärstatistik führt.

Die Bereitschaft der Unternehmensgruppen zur Zusammenarbeit mit der LCU war bisher, von wenigen Ausnahmen abgesehen, sehr hoch. Die Gruppen stellten zum Teil eigene intensive Recherchen zur Aufklärung der Inkohärenzen an. Sie sind insgesamt an einer nachhaltigen Verbesserung ihrer statistischen Meldungen interessiert. Ein wichtiges Ergebnis im Nachgang der Gespräche war, dass viele Unternehmensgruppen ihre internen Abläufe zur Generierung statistischer Meldungen optimiert haben. Das führt tendenziell zu weniger Rückfragen durch die statistischen Ämter und somit langfristig zu einer Reduzierung der Erhebungsbelastung.

3.2 Häufige Inkohärenzen

Fachstatistiken sind häufig für sich allein genommen plausibel; die Inkohärenzen können sich durchaus erst in einer statistik- und institutionenübergreifenden Betrachtung (wie sie die LCU durchführt) ergeben. Im Zuge der Aufklärung von Inkohärenzen gewinnt die LCU Erkenntnisse auf zwei Ebenen: Erstens erkennt sie Fehler in den Erhebungsdaten einzelner Gruppen und ihrer Untergliederungen und entwickelt daraus Korrekturvorschläge, die sie den Volkswirtschaftlichen Gesamtrechnungen und den Fachstatistiken zur Verfügung stellt. Zweitens ergeben sich aus häufig wiederkehrenden Konstellationen methodische Erkenntnisse, die zum Beispiel zu allgemeinen Empfehlungen für die Volkswirtschaftlichen Gesamtrechnungen und Fachstatistiken führen.

Korrekturvorschläge zu Einzeldaten

Diese Korrekturvorschläge umfassen quantitative Änderungen in den Einzeldaten der betrachteten Statistiken. Zu diesem Zweck stellt die LCU Verwertungsberichte zusammen, aus denen hervorgeht, in welcher Einheit, bei welcher Statistik, bei welchen Merkmalen und für

welche Zeiträume Anpassungsbedarf aus Sicht der LCU besteht.

Die Korrekturvorschläge der LCU waren zum Teil beträchtlich. Einen Überblick über den Umfang in ausgewählten Erhebungen und Merkmalen gibt [Tabelle 1](#), in der die aggregierten absoluten Beträge der Korrekturen angeführt sind.

Auch in den Daten der Zahlungsbilanzstatistik der Deutschen Bundesbank wurden Inkohärenzen festgestellt, insbesondere bei den Einnahmen und Ausgaben zum sonstigen Warenverkehr und im Transithandel. Dazu hat die Deutsche Bundesbank bei den Meldern nachgefragt. Einige Fälle führten zu entsprechenden Nachmeldungen, die in die turnusmäßigen Revisionen der Zahlungsbilanzstatistik einfließen.

Häufige Ursachen für Inkohärenzen und daraus folgende Korrekturvorschläge waren in den Statistiken des Verarbeitenden Gewerbes:

- › Gesamtumsatz: Hier handelte es sich oft (insbesondere in der Jahrerhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe) um

Tabelle 1

Umfang der LCU-Korrekturvorschläge für ausgewählte Merkmale (absolute Beträge)

	2018	2019	2020	2021	2022
Mrd. EUR					
Kostenstrukturerhebungen im Verarbeitenden Gewerbe, in der Energieversorgung und im Baugewerbe					
Gesamtumsatz	8,3	2,3	5,6	7,5	3,9
Umsatz aus eigenen Erzeugnissen	65,9	57,3	50,7	87,3	82,9
Umsatz aus Handelswaren	53,1	49,5	45,3	73,5	80,5
Bruttowertschöpfung	4,3	2,7	0,7	0,8	0,2
Jahrerhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe					
Gesamtumsatz	90,9	76,5	51,3	18,7	16,2
Umsatz aus eigenen Erzeugnissen	128,4	120,9	83,3	85,7	68,4
Sonstiger Umsatz	121,7	130,1	91,1	71,2	63,5
Strukturerhebung im Handel und Dienstleistungsbereich (ab 2021); bis 2020: Jahrerhebung im Handel, Strukturstatistik im Dienstleistungsbereich					
Gesamtumsatz	47,9	23,5	43,4	26,1	0,0
Bruttowertschöpfung	17,5	15,1	17,8	30,8	8,4
Außenhandel					
Importe ohne Eigentumsübergang	26,1	30,8	32,2	36,9	49,4
Exporte ohne Eigentumsübergang	29,8	29,4	28,3	36,7	45,1
Statistisches Unternehmensregister					
Umsatz der Rechtlichen Einheiten	114,7	122,1	113,5	54,5	37,6

Stand: Dezember 2024

Untererfassungen. Häufig wurden die Umsatzwerte aus den Angaben des Monatsberichts für Betriebe als Summe der einzelnen Niederlassungen generiert. Niederlassungen außerhalb des Verarbeitenden Gewerbes fehlten dann. Gravierend waren auch fehlende Umsätze, die nicht einzelnen Niederlassungen zuzuordnen waren – beispielsweise wurden Verkäufe von importierten Waren nicht einbezogen. An den Zahlen in Tabelle 1 lässt sich vor allem bei der Jahrerhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe schon der Effekt der LCU-Arbeit erkennen: In den ersten Analysezyklen wurden Inkohärenzen erkannt und aufgeklärt, in den Folgejahren hat sich das Meldeverhalten der betroffenen Einheiten dann nachhaltig gebessert.

- › Umsatzstruktur: Sowohl in den Kostenstrukturerhebungen im Verarbeitenden Gewerbe, in der Energieversorgung und im Baugewerbe als auch in der Jahrerhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe zeigte sich bei den meldenden Rechtlichen Einheiten häufig ein falsches Verständnis der Umsatzkategorien: In vielen Fällen wurden Verkäufe von Erzeugnissen, die von ausländischen Tochtergesellschaften bezogen wurden, als Umsatz aus eigenen Erzeugnissen gemeldet und nicht als Handelsware. Auch Verkäufe von Erzeugnissen, die selbst in (aktiver) Lohnarbeit produziert und danach dem Auftraggeber abgekauft wurden, wurden als eigene Erzeugnisse statt als Handelsware deklariert. Umgekehrt wurden Umsätze mit Waren, die in Lohnarbeit von anderen Rechtlichen Einheiten produziert wurden, als Handelsware und nicht als Umsatz aus eigenen Erzeugnissen angegeben.

In den Handels- und Dienstleistungsstatistiken gab es folgende häufig wiederkehrende Probleme:

- › Gesamtumsatz: In einigen Fällen wurden fälschlicherweise im Ausland erzielte Umsätze (Dienstleistungsexporte) nicht einbezogen. Übererfassungen gab es auch, wenn sonstige betriebliche Erträge wie Verkäufe von Vermögenswerten oder Auflösungen von Rückstellungen in die Umsätze einbezogen wurden. Hier ist der Effekt von LCU-Arbeit ebenfalls sichtbar. Nach Aufklärung großer Fälle gab es (Stand Dezember 2024) beispielsweise für 2022 keinen Korrekturbedarf beim Gesamtumsatz.

- › Vorleistungen (Kosten): Hier bestand der häufigste Fehler darin, dass einige Positionen der sonstigen betrieblichen Aufwendungen in die Meldungen einbezogen wurden, obwohl sie per definitionem auszuschließen sind. Dazu zählen Kursverluste aus Devisenabsicherungen, Verluste durch Umstrukturierungen oder Verluste aus Anlageabgängen.
- › Bestände an Handelswaren: Vereinzelt, aber mit hohen Beträgen, wurden in den Beständen an Handelswaren auch Bestände an Gütern des Anlagevermögens gemeldet. Bestandsveränderungen flossen dann direkt in die Bruttowertschöpfung ein.

In den Außenhandelsmeldungen betrafen die meisten Korrekturvorschläge Änderungen in der Art des Geschäftes. Häufig stellte die LCU fest, dass es sich bei Transaktionen nicht um Käufe und Verkäufe mit Eigentumsübergang handelte, sondern um Vorgänge, bei denen das Eigentum an den Waren bei Grenzübertritt nicht wechselte, insbesondere im Zusammenhang mit aktiver und passiver Lohnveredelung. Dieser Sachverhalt ist für die Volkswirtschaftlichen Gesamtrechnungen besonders wichtig, da die Verwendungsrechnung des Bruttoinlandsproduktes Transaktionen ohne Eigentumsübergang aus dem Außenbeitrag ausschließt.

Die hohen Korrekturvorschläge für das Merkmal Umsatz im statistischen Unternehmensregister haben unterschiedliche Ursachen. Zum einen sind sie Folge von fehlerhaften Umsatzangaben in anderen Statistiken, die als Quelle für die URS-Umsätze dienen. Zu nennen ist hier vor allem die Jahrerhebung für Mehrbetriebsunternehmen im Verarbeitenden Gewerbe. Es handelt sich daher explizit nicht um eine Inkohärenz, die dem statistischen Unternehmensregister zuzurechnen ist. Zum anderen berücksichtigten bis zum Berichtsjahr 2022 die URS-Umsätze, deren Quelle die regelmäßig von der Finanzverwaltung zur Verfügung gestellten Daten der Umsatzsteuer-Voranmeldung sind, lediglich die steuerbaren Umsätze. Nicht steuerbare Umsätze – das sind im Wesentlichen Auslandsumsätze – waren nicht enthalten. Dies führte zu einer Untererfassung insbesondere in den Bereichen Handel und Dienstleistungen. Ab Berichtsjahr 2023 wurde die Berechnungsmethodik im statistischen Unternehmensregister angepasst, nicht steuerbare Umsätze werden nun einbezogen.

Allgemeine methodische Erkenntnisse der LCU

Abweichungen zwischen Umsatz aus eigenen Erzeugnissen und Absatzproduktionswert im Verarbeitenden Gewerbe

Eine häufig auftretende Inkohärenz bei Einheiten des Verarbeitenden Gewerbes besteht darin, dass der jährliche Umsatz aus eigenen Erzeugnissen deutlich höher ausfällt als der über das Jahr kumulierte Absatzproduktionswert aus der Produktionsstatistik. Abgesehen von Vorratsveränderungen ist eine solche Abweichung meist nur dann plausibel, wenn die meldende Einheit als Auftraggeber für Lohnarbeit/Lohnveredelung tätig ist. In vielen LCU-Fällen konnte das ausgeschlossen werden. Ursache war hingegen häufig, dass in der Produktionsstatistik die von den einzelnen Niederlassungen einer Rechtlichen Einheit gemeldete Absatzproduktion nicht wie gefordert zu Marktpreisen bewertet wurde, sondern zu internen Verrechnungspreisen zwischen Niederlassung und Rechtlicher Einheit. Diese decken in der Regel nur die der Niederlassung entstandenen Kosten (gegebenenfalls einschließlich einer Marge) ab. Die Niederlassung kennt häufig die Marktpreise nicht, da die Rechnungslegung an den Kunden durch die Rechtliche Einheit erfolgt. Die Rechnungspreise (Marktpreise) enthalten dann zusätzliche Bestandteile wie Vertriebskosten oder Umlagen für Forschung und Entwicklung.

Bruttowertschöpfung bei Einheiten im Bereich Handel und Dienstleistungen

Bei einer Reihe von Einheiten im Bereich Handel und Dienstleistungen wurden zum Teil dauerhaft negative oder im Zeitverlauf stark schwankende Werte für die Bruttowertschöpfung festgestellt. Die LCU-Analysen zeigten, dass dafür oft die vergleichsweise hohen Vorleistungen verantwortlich waren und dabei insbesondere die Position „Sonstige betriebliche Aufwendungen“. Die gleichlautende Position in den Gewinn- und Verlustrechnungen der Jahresabschlüsse stimmte formal damit überein. Allerdings fiel auf, dass die Jahresabschlüsse Unterpositionen enthalten, die in den statistischen Erhebungen explizit auszuschließen sind, da sie nicht zu den Vorleistungen zählen. Dazu gehören Kursverluste, Verluste aus dem Abgang von Anlagevermögen und andere Aufwendungen, die nicht im Zusammenhang mit der Tätigkeit der Rechtlichen Einheit stehen. Hier sind demnach Erhebungsmerkmal und Position in der Gewinn- und Verlustrechnung wortgleich, definitorisch jedoch unter-

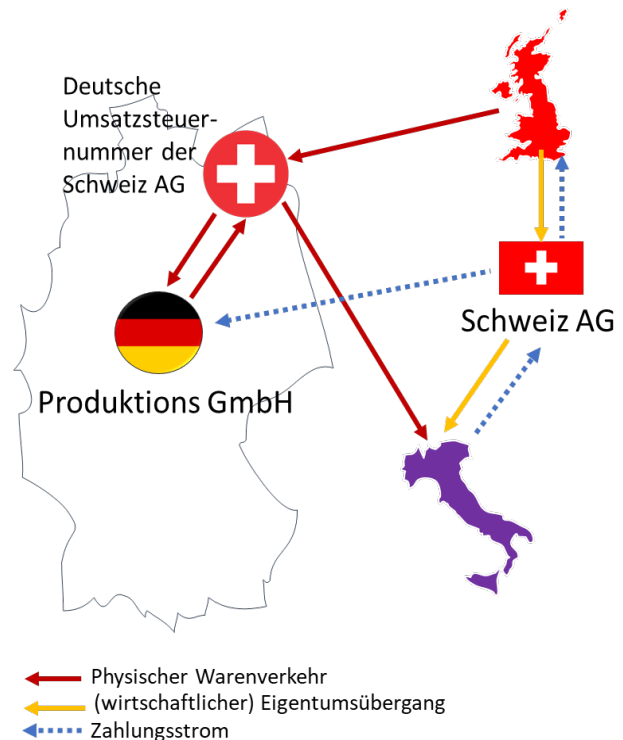
schiedlich abgegrenzt. Das führt offenbar zu Missverständnissen bei der Strukturhebung im Handel und Dienstleistungsbereich. Aufgrund dieser Erkenntnisse der LCU wird geprüft, das Erhebungsmerkmal in den Statistiken umzubenenen.

Ausländische Einheiten mit deutscher Steuernummer

Bei den LCU-Analysen sind in den Daten der Außenhandelsstatistik häufig Transaktionen von ausländischen Einheiten aufgefallen, die in Deutschland mit einer Umsatzsteuernummer registriert sind, hier jedoch weder einen physischen Standort noch tätige Personen oder Produktion haben. Sie sind rechtlich in einem anderen EU- oder Nicht-EU-Land ansässig und ihre Tätigkeit beschränkt sich oft auf Handel oder Lagerung von Waren im Inland. Die Einrichtung solcher Einheiten ist durch die Anforderungen des EU-Mehrwertsteuerrechts notwendig. Diese Value-Added-Tax-Trader (VAT-Trader) wickeln Im- und Exporte ab, bei denen das Eigentum an den gehandelten Waren bei den ausländischen Einheiten bleibt. Ein fiktives Beispiel zeigt [Grafik 3](#).

Grafik 3

Warenströme, Eigentumsübergänge und Zahlungsströme im Zusammenhang mit VAT-Tradern
Fiktives Beispiel



VAT-Trader: Value-Added-Tax-Trader (ausländische Einheiten mit deutscher Steuernummer)

Die Schweiz AG lässt in Deutschland von der Produktions GmbH Erzeugnisse in Lohnarbeit fertigen. Die Rohstoffe kauft die Schweiz AG dafür im Vereinigten Königreich ein und lässt sie direkt nach Deutschland liefern. Die Importe der Rohstoffe werden über eine deutsche Umsatzsteuernummer der Schweiz AG (VAT-Trader) abgewickelt und von dieser zur Außenhandelsstatistik gemeldet. Das Eigentum an den Rohstoffen verbleibt jedoch bei der Schweiz AG. Die Produktions GmbH erzeugt in Lohnarbeit Fertigprodukte und erhält dafür ein Fertigungsentgelt, das sie in der Zahlungsbilanzstatistik als Einnahme aus Lohnfertigung meldet. Die Fertigprodukte sind Eigentum der Schweiz AG und werden an einen italienischen Abnehmer geliefert. Der Export aus Deutschland wird wieder vom VAT-Trader abgewickelt. In dieser Konstellation überschreiten sowohl Rohstoffe als auch Fertigprodukte physisch die deutsche Grenze, ohne dass zu irgendeinem Zeitpunkt das Eigentum auf eine deutsche Einheit übergeht. Sehr häufig wurden zur Außenhandelsstatistik jedoch die Warenbewegungen als Eigentumsübergänge gemeldet, was zu einer fehlerhaften Darstellung der VGR-Exporte und -Importe nach den Regeln des [Europäischen Systems Volkswirtschaftlicher Gesamtrechnungen](#) (ESVG) 2010 führt.

Grundsätzlich handelt es sich bei den grenzüberschreitenden Warenbewegungen der VAT-Trader um Transaktionen ohne Eigentumsübergang. Sie sind also nicht als Ex- und Importe in den Volkswirtschaftlichen Gesamtrechnungen zu buchen. Eine entsprechende Umstellung ist im Zuge der Generalrevision 2024 der Volkswirtschaftlichen Gesamtrechnungen erfolgt (Kuhn und andere, 2024).

3.3 Verwertung der LCU-Erkenntnisse

Für die Implementierung der LCU-Erkenntnisse in den VGR-Ergebnissen sind die Volkswirtschaftlichen Gesamtrechnungen selbst verantwortlich. Die LCU präsentiert dazu die aus ihrer Sicht zu korrigierenden Daten der einzelnen Statistiken als Tabellen und Fallberichte sowie gewonnene methodische Erkenntnisse. Die LCU, alle VGR-Bereiche sowie die Deutsche Bundesbank diskutieren in regelmäßigen Workshops die Ergebnisse der LCU-Analysen und schätzen ein, wie sich mögliche Korrekturen in den Wirtschaftsstatistiken auf die VGR-Ergebnisse auswirken. Die Volkswirtschaftlichen Gesamtrechnungen entscheiden dann, in welcher Form


sie die Korrekturen berücksichtigen. Seit 2021 arbeiten die Volkswirtschaftlichen Gesamtrechnungen jährlich LCU-Erkenntnisse in ihre Ergebnisse ein, zuletzt in größerem Umfang in der Generalrevision 2024 (Kuhn und andere, 2024, Abschnitt 2.6).

Auch die Wirtschaftsstatistiken verarbeiten die LCU-Erkenntnisse. Der Fokus liegt dabei darauf, bei aufgeklärten Inkohärenzen die Erhebungsfehler in den Folgeerhebungen zu vermeiden. Dass dies erfolgreich ist, zeigt sich darin, dass der Umfang der LCU-Korrekturvorschläge im Zeitverlauf bei einigen Kennziffern sinkt (siehe Tabelle 1). Bei systematischen Inkohärenzen wurden in einigen Fällen bereits Definitionen oder Erläuterungen zu einzelnen Merkmalen angepasst. Vergleichbare Revisionsmechanismen für zurückliegende Zeiträume wie in den Volkswirtschaftlichen Gesamtrechnungen gibt es in den meisten von der LCU verwendeten Statistiken bisher jedoch nicht. Zwar werden gravierende Fehler in den Daten nachträglich korrigiert, ein systematischer Ansatz für rückwirkende Korrekturen ist jedoch insbesondere für die Strukturstatistiken noch zu entwickeln. Insgesamt hat die bisherige Arbeit der LCU auch für die Wirtschaftsstatistiken dazu geführt, die Qualität der Ergebnisse zu steigern.

4

Ausblick

Die Weiterentwicklung der LCU in den nächsten Jahren wird sich in mehreren Dimensionen vollziehen. Zunächst soll – abhängig von den zur Verfügung stehenden Ressourcen – der Umfang der Fallbearbeitungen sukzessive ausgeweitet werden. Parallel wird permanent an der Weiterentwicklung der LCU-Methodik gearbeitet. Dazu gehört, neue Datenquellen zu erschließen und neue Kohärenzprüfungen zu entwickeln. So wird ein Schwerpunkt in der nächsten Zeit darin bestehen, die Transaktionen deutscher VAT-Trader im Ausland zu analysieren. Auch die weitere Einführung von Prüfungen im Zusammenhang mit Gütern des geistigen Eigentums und die Entwicklung von Kohärenzchecks zu Investitionen stehen an. Außerdem soll die internationale Zusammenarbeit vertieft werden. Dies betrifft sowohl die Mitarbeit in europäischen Gremien als auch der bilaterale Austausch mit LCUs in anderen Ländern bei der Analyse von Unternehmensgruppen.

In Zusammenarbeit mit den Fachbereichen ist die Verwertung der LCU-Erkenntnisse in den einzelnen Statistiken als künftiger Schwerpunkt zu sehen. Viel wichtiger ist es jedoch, Arbeitsabläufe zu etablieren, die Inkohärenzen und Fehler in den Einzeldaten frühzeitig erkennen und aufklären. Während derzeit die LCU-Analyse meist erst nach Fertigstellung der einzelnen Statistiken erfolgt, soll sie künftig möglichst noch während der Aufbereitungsprozesse stattfinden. Dies würde eine nachträgliche Korrektur der Ergebnisse vermeiden, sowohl in den Fachstatistiken selbst als auch in den Volkswirtschaftlichen Gesamtrechnungen. 

LITERATURVERZEICHNIS

Ahlborn, Markus/Draken, Ferdinand/Schulz, Verena. Qualitätssicherung in der amtlichen Statistik: Large Cases Unit. In: WISTA Wirtschaft und Statistik. Ausgabe 2/2021, Seite 31 ff.

Andom, Senait/Bingel, Johanna/Urban, Adrian. Methodische Weiterentwicklungen für Unternehmensgruppen im statistischen Unternehmensregister. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2025, Seite 91 ff.

Draken-Gädeke, Ferdinand/Kleine, Riepke/Piradashvili, Irina. Untersuchungen zum Einfluss multinationaler Unternehmensgruppen auf das Bruttonationaleinkommen. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2023, Seite 27 ff.

Hörner, Natalie/Rotsche, Maximilian/Söngen, Jens. Fortschritte der Large Cases Unit. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2022, Seite 32 ff.

Kuhn, Michael/Hauf, Stefan/Kuntze, Peter/Janz, Christian/Fries, Claudia/Hater, Arne/Seeger, Daniel/Ammermann, Sascha/Bolleyer, Rita/Schwarz, Norbert/Draken-Gädeke, Ferdinand/Schmidt, Pascal/Appler, Felix. Revision der Volkswirtschaftlichen Gesamtrechnungen 1991 bis 2023. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2024, Seite 15 ff.

Redecker, Matthias/Rommelspacher, Simon/Sturm, Roland. Profiling von Unternehmen im Echtbetrieb. In: WISTA Wirtschaft und Statistik. Ausgabe 5/2021, Seite 99 ff.

Statistisches Bundesamt. *Qualitätsbericht: Statistisches Unternehmensregister 2023*. 2024. [Zugriff am 26. Februar 2025]. Verfügbar unter: www.destatis.de

RECHTSGRUNDLAGEN

Gesetz zur Prüfung von Daten multinationaler Unternehmensgruppen zur Sicherung der Qualität der Volkswirtschaftlichen Gesamtrechnungen und der Wirtschaftsstatistiken (Qualität-VGR und WS-Gesetz – QVWSG) vom 14. Juni 2021 (BGBl. I Seite 1751, 1757).

AUSWERTUNGEN ÜBER UNTERNEHMENSGRUPPEN IM STATISTISCHEN UNTERNEHMENSREGISTER

Roland Sturm, Adrian Urban

📌 **Schlüsselwörter:** Globalisierung – Wirtschaftsstatistik – ökonomische Konzentration – Einfluss aus dem Ausland – Qualitätssicherung

ZUSAMMENFASSUNG

Der Beitrag stellt ausgewählte Auswertungen zu Unternehmensgruppen im statistischen Unternehmensregister dar und beschreibt anhand dieser verschiedene methodische Ansätze zur Erfassung ihrer Strukturen. Mit dem vor Kurzem erfolgten Ausbau der Datenbank des Unternehmensregisters für die Einheit Unternehmensgruppe und der Einführung der „Manuellen Unternehmensgruppenbearbeitung“ als Qualitätssicherung wurden die Grundsteine für bessere Auswertungen zu Unternehmensgruppen gelegt. Der Fokus des Beitrags liegt auf der Darstellung der wirtschaftlichen Bedeutung von Unternehmensgruppen in Deutschland. Zudem thematisiert er die Aktivität von Unternehmensgruppen in den verschiedenen Wirtschaftsbereichen, die Herausforderungen bei der Beschreibung ihrer Aktivitäten sowie die Rolle der internationalen Verflechtung von Unternehmensgruppen.

📌 **Keywords:** globalisation – economic statistics – economic concentration – foreign influence – quality assurance

ABSTRACT

The article presents selected analyses of enterprise groups in the German statistical business register and, based on these analyses, describes various methodological approaches to capturing their structures. The foundation for improved analyses of enterprise groups has been laid with the recent expansion of the database to include the enterprise group as an entity and the introduction of “Manual Enterprise Group Processing” for quality assurance. The article focuses on presenting the economic significance of enterprise groups in Germany. It also addresses the activities of enterprise groups in the various economic sectors, the challenges involved in capturing these activities, and the role of international linkages between enterprise groups.



Roland Sturm

ist Diplom-Volkswirt und leitet das Referat „Unternehmensregister, -gruppen, Erhebungsbetreuung, Profiling“ des Statistischen Bundesamtes. Er vertritt das Statistische Bundesamt in den europäischen Gremien, die sich mit statistischen Einheiten befassen und leitet die Arbeitsgruppen „Profiling“ und „Unternehmensgruppen“ der Statistischen Ämter des Bundes und der Länder.



Adrian Urban

ist Volkswirt (M. Sc.) und Referent für das Sachgebiet Unternehmensgruppen im Referat „Unternehmensregister, -gruppen, Erhebungsbetreuung, Profiling“ des Statistischen Bundesamtes. Er koordiniert die „Manuelle Unternehmensgruppenbearbeitung“ im Statistischen Verbund und entwickelt Konzepte zur Verbesserung der Datenqualität und Analyse von Unternehmensgruppen.

1

Einleitung

Die letzten Veröffentlichungen in dieser Zeitschrift zu Methodik und Auswertungen der Daten des statistischen Unternehmensregisters über Unternehmensgruppen liegen bereits länger zurück (Sturm und andere, 2009; Kleber und andere, 2010). Damals waren Daten zu Unternehmensgruppen im statistischen Unternehmensregister noch Neuland. Die Arbeiten zur Integration und regelmäßigen Aktualisierung von Informationen zu Unternehmensgruppen in das statistische Unternehmensregister sind seitdem deutlich vorangekommen, jedoch konnten noch nicht alle Vorhaben umgesetzt werden. Das betrifft auch den inhaltlichen Ausbau der Registerdatenbank. Die erreichten Fortschritte stellen zwei Beiträge in dieser Zeitschrift vor: Nach dem erreichten Stand und den methodischen Grundlagen zu Unternehmensgruppen im Unternehmensregister (Andom und andere, 2025) folgen in diesem Beitrag erste Auswertungen der Daten. Unter den Unternehmensgruppen finden sich die größten, wirtschaftlich äußerst bedeutenden ökonomischen Akteure des Wirtschaftsgeschehens. Um die häufig komplexen Strukturen dieser Gruppen im statistischen Unternehmensregister korrekt darzustellen, ist ein erheblicher kontinuierlicher Pflegeaufwand notwendig, um das „Wer gehört zu wem?“ der deutschen Wirtschaft zutreffend in den amtlichen Wirtschaftsstatistiken darstellen zu können. Oft ist das nur mit aufwendigen Einzelrecherchen möglich. Die noch im Aufbau befindliche „Manuelle Unternehmensgruppenbearbeitung“ beschäftigt sich mit dieser Aufgabe. Aufgrund dieser Situation haben die Statistischen Ämter des Bundes und der Länder bislang kaum Daten zu Unternehmensgruppen veröffentlicht.

Ein großer Schritt zur besseren Abbildung der Unternehmensgruppen war die Einführung des entsprechenden Entitätstyps in der Datenbank des Unternehmensregisters. Dies stellt für die Datenhaltung, für Qualitätsarbeiten und damit auch für Auswertungen und Veröffentlichungen einen entscheidenden Fortschritt dar. Seit dem Berichtsjahr 2020 enthält die Datenbank des statistischen Unternehmensregisters neben den Entitäten Niederlassung, Rechtliche Einheit und Unternehmen auch die Entität Unternehmensgruppe. Dieser Beitrag stellt erste Daten über Unternehmensgruppen in Deutsch-

land und für deren Verständnis notwendige methodische Erläuterungen gemeinsam dar. Den Auswertungen liegt noch kein abgestimmtes Veröffentlichungskonzept zugrunde; ein solches erstellt derzeit eine Arbeitsgruppe des Statistischen Verbundes¹. Danach sollen solche Daten aus dem Unternehmensregister regelmäßig veröffentlicht werden.

Auf dem Weg zu regelmäßigen Veröffentlichungen präsentiert dieser Beitrag einige erste Auswertungen zum Berichtsjahr 2022. Kapitel 2 definiert die wichtigsten Begriffe, die für die Auswertungen über Unternehmensgruppen im statistischen Unternehmensregister relevant sind. Die Kapitel 3 bis 5 informieren anschließend über die Ergebnisse der ersten Auswertungen und beleuchten die Bedeutung der (größten) Unternehmensgruppen in der deutschen Gesamtwirtschaft sowie in den verschiedenen Wirtschaftsbereichen. Kapitel 6 thematisiert die wirtschaftlichen Verflechtungen innerhalb Deutschlands und mit dem Ausland. Abschließend wird in Kapitel 7 ein kurzes Fazit gezogen.

2

Definitionen

Eine Unternehmensgruppe ist ein Zusammenschluss von Rechtlichen Einheiten, die über rechtliche Kontrollbeziehungen² miteinander verbunden sind. Die Rechtliche Einheit an der Spitze der Kontrollkette wird „Gruppenoberhaupt“ genannt. Darüber hinaus hat jede Unternehmensgruppe eine repräsentative Einheit in Deutschland, die als strategische und operative Steuerungseinheit der Unternehmensgruppe auftritt. Dies ist die „Deutsche Entscheidungseinheit“. Diese Einheit kann mit dem Gruppenoberhaupt übereinstimmen, muss es jedoch nicht. Der Begriff „Unternehmensgruppe“ wurde für die amtliche Statistik bewusst definiert. Unternehmensgruppen müssen nicht zwangsläufig mit der rechtlichen Abgrenzung des Konzerns identisch sein. Die Unterscheidung des Konzepts der Unternehmensgruppe von dem des (rechnungslegenden) Konzerns ist notwendig, um die Vielfalt möglicher Strukturformen darzustellen, die über

1 Den Statistischen Verbund bilden die Statistischen Ämter des Bundes und der Länder.

2 Zur Definition der Kontrolle und für weitere methodische Aspekte bei Unternehmensgruppen siehe Andom und andere (2025).

die klassische Definition eines Konzerns hinausgehen. Eine Unternehmensgruppe kann einem rechnungslegenden Konzern entsprechen, sie kann jedoch auch mehrere Konzerne umfassen, die alle unter der Kontrolle eines gemeinsamen Gruppenoberhaupts stehen. Ebenso ist es möglich, dass eine Unternehmensgruppe in keinem Zusammenhang mit einem rechnungslegenden Konzern steht oder es besteht schlichtweg keine Pflicht zur Erstellung eines Konzernabschlusses. In solchen Fällen liegt zwar eine Unternehmensgruppe vor, jedoch kein Konzern im rechnungslegungsrechtlichen Sinne.³

Alle Rechtlichen Einheiten, die Teil einer Unternehmensgruppe sind, also sowohl beherrschte Einheiten als auch das Gruppenoberhaupt, werden gruppenzugehörige oder gruppenabhängige Rechtliche Einheiten genannt. Sofern die gruppenzugehörigen Rechtlichen Einheiten einer Unternehmensgruppe in mindestens zwei verschiedenen Staaten ansässig sind, handelt es sich um eine multinationale Unternehmensgruppe. Bei ausschließlich im Inland ansässigen Rechtlichen Einheiten handelt es sich um eine nationale Unternehmensgruppe.

Relevant für das statistische Unternehmensregister sind alle Unternehmensgruppen, die mindestens eine deutsche Rechtliche Einheit enthalten. In der Datenbank des Unternehmensregisters wird die Entität Unternehmensgruppe mit ihren eigenen Merkmalen und ihren deutschen gruppenzugehörigen Rechtlichen Einheiten geführt. Zusätzlich sind in der Datenbank auch Informationen zur Identität des Gruppenoberhaupts oder zu ersten ausländischen Müttern oder Töchtern bei den einzelnen Rechtlichen Einheiten abgespeichert. Einige Merkmale, wie in Deutschland Beschäftigte und deutsche Umsätze der Unternehmensgruppe, werden durch die Summierung der Werte der gruppenzugehörigen Rechtlichen Einheiten generiert. Andere Merkmale, wie der Name der Unternehmensgruppe, in großen Teilen die „Deutsche Entscheidungseinheit“ oder insbesondere die wirtschaftliche Haupttätigkeit der Unternehmensgruppe, müssen durch komplexere Algorithmen bestimmt werden.

³ Zur Abgrenzung des Konzernbegriffs im Rechnungslegungsrecht siehe § 290 Handelsgesetzbuch.

3

Erste Zahlen: wirtschaftliche Aktivität der gruppenzugehörigen Einheiten

Das statistische Unternehmensregister weist zum Berichtsjahr 2022 rund 292 000 in Deutschland tätige Unternehmensgruppen mit etwa 938 000 gruppenzugehörigen deutschen Rechtlichen Einheiten nach. Neben Einheiten, die über Umsätze und/oder Beschäftigte verfügen, sind im Kontext von Unternehmensgruppen auch solche Einheiten relevant, die allein durch ihre Position in der Kontrollstruktur der Unternehmensgruppe einen Einfluss auf die Zusammensetzung und Abgrenzung der Unternehmensgruppe haben. Hierbei handelt es sich oft um Holdings und Beteiligungsgesellschaften, die von den Unternehmensgruppen häufig alleine für den Zweck geschaffen werden, Beteiligungsverhältnisse zu organisieren und die Gruppenstruktur zu gliedern.

Eurostat, das Statistische Amt der Europäischen Union (EU), hat wirtschaftliche Aktivität folgendermaßen definiert:

“For the purposes of the European framework for statistical business registers, the following shall be considered to be an economic activity:

- (a) any activity comprising the offer of goods and services on a given market;
- (b) non-market services contributing to the GDP;
- (c) direct and indirect holdings of active legal units.

Holding assets and/or liabilities may also be considered to be an economic activity.”

Für die Operationalisierung und Abgrenzung der wirtschaftlichen Aktivität im statistischen Unternehmensregister gelten folgende Bedingungen:

Eine Rechtliche Einheit ist aktiv im jeweiligen Berichtsjahr nach EU-Definition, sofern sie

- › einen positiven Umsatz für das Berichtsjahr aufweist oder
- › Beschäftigte im Berichtsjahr aufweist oder
- › Mutter einer anderen Rechtlichen Einheit im Berichtsjahr ist oder
- › Teil einer Erhebung ist.

Bei Anwendung der EU-Definition auf die gruppenzugehörigen Rechtlichen Einheiten erhält man folgendes Bild:

Über 614 000 der gruppenzugehörigen Rechtlichen Einheiten gelten als aktiv laut der EU-Definition. Rund 82 % dieser aktiven Einheiten weisen auch Umsätze und Beschäftigte auf. Bei mehr als 110 000 der aktiven Rechtlichen Einheiten (12 %) ist die Aktivität nur auf die Muttereigenschaft der Rechtlichen Einheit zurückzuführen. Diese Einheiten weisen sonst kein weiteres Aktivitätskriterium auf.⁴

4

Bedeutung der Unternehmensgruppen in der deutschen Wirtschaft

Der überwiegende Teil der deutschen Wirtschaft ist in Form von Unternehmensgruppen organisiert. Hier erwirtschaften 68 % der deutschen Beschäftigten⁵ 83 % aller Umsätze der im statistischen Unternehmensregister enthaltenen Wirtschaftsakteure. Demgegenüber sind es nur knapp 15 % aller deutschen Rechtlichen Einheiten, die diese Unternehmensgruppen bilden. Die überwiegende Anzahl der deutschen Wirtschaftsakteure ist also in Form von unabhängigen Einzelgesellschaften organisiert, deren ökonomisches Gewicht aber weit hinter dem der Unternehmensgruppen zurücksteht. [↘ Grafik 1](#)

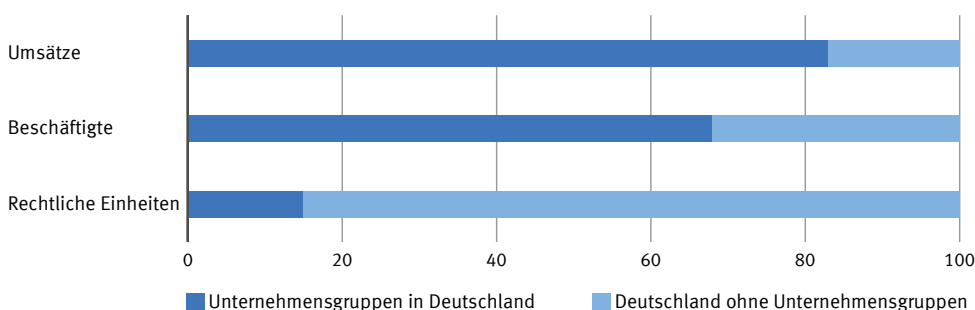
Wird die Verteilung der Unternehmensgruppen nach ihrer Größe betrachtet, so fällt eine ausgeprägte Rechtschiefe auf. Der Versuch, eine Verteilungskurve aller 292 000 Unternehmensgruppen in Deutschland zu erstellen, weist bei üblichen Darstellungsformaten kaum mehr als die Einfärbung der Achsen des Diagramms nach. Erst bei der Betrachtung des obersten Perzentils (rund 2 920 Gruppen) ist eine tatsächliche Verteilungskurve zu erkennen. Die Verteilung der größten 100 Gruppen zeigt [↘ Grafik 2](#).

⁴ Daneben gibt es rund 314 000 Einheiten, bei denen eine Gruppenrelevanz bekannt ist, die aber laut EU-Definition inaktiv sind. Diese Einheiten müssen noch näher analysiert werden.

⁵ Alle in diesem Beitrag enthaltenen Auswertungen zu Beschäftigten beziehen sich ausschließlich auf sozialversicherungspflichtig Beschäftigte.

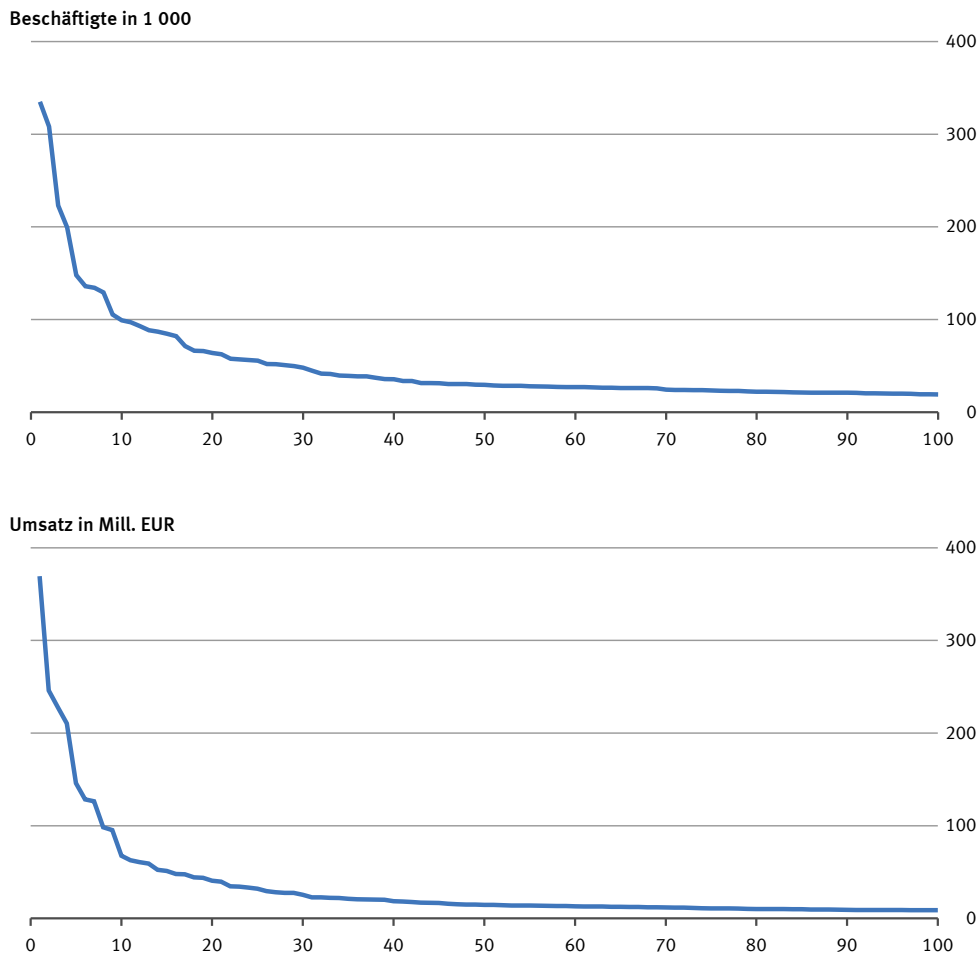
Grafik 1

Bedeutung von Unternehmensgruppen in Deutschland im Berichtsjahr 2022
in %



Grafik 2

Verteilungskurve der 100 größten Unternehmensgruppen nach Beschäftigten und Umsatz im Berichtsjahr 2022



Die nach Beschäftigten 100 größten Unternehmensgruppen in Deutschland beschäftigen hierzulande insgesamt 15 % aller sozialversicherungspflichtig Beschäftigten und erwirtschaften 25 % der Umsätze der deutschen Wirtschaft. Die nach dem Umsatz größten 100 Unternehmensgruppen vereinen sogar 35 % aller Umsätze auf sich und beschäftigen 11 % aller sozialversicherungspflichtig Beschäftigten. Die Analyse der jeweiligen Top-100-Gruppen zeigt auch, dass die größten Unternehmensgruppen überdurchschnittlich hohe Umsätze je Beschäftigten erwirtschaften.

Viele Veröffentlichungen verwenden standardisierte Größenklassen. Angesichts der extrem schiefen Verteilung der Unternehmensgruppen sind diese im oberen Bereich nicht sehr aussagekräftig. Dies zeigt die folgende Dar-

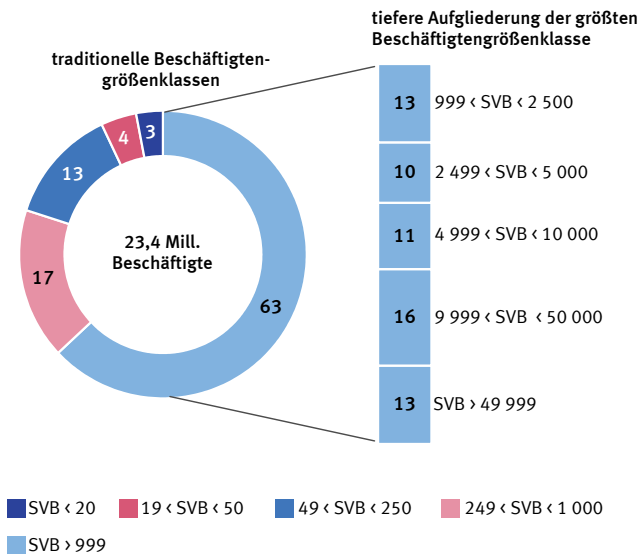
stellung. Das Kreisdiagramm in Grafik 3 verwendet für große Einheiten die gebräuchliche Abgrenzung „mehr als 999 sozialversicherungspflichtig Beschäftigte“.

Die größte Größenklasse wirkt zunächst beeindruckend – in Gruppen mit mehr als 999 Beschäftigten arbeiten 63 % aller Beschäftigten in Unternehmensgruppen. Zugleich geht bei dieser Darstellung als eine Klasse auch viel Information verloren. Daher wird diese größte Klasse weiter unterteilt in fünf Unterklassen – die oberste Kategorie ist dabei „50 000 Beschäftigte und mehr“.

Diese oberste Unterklasse umfasst 28 Unternehmensgruppen, was 0,01 % aller Unternehmensgruppen entspricht. Darin befinden sich allerdings 13 % aller in Unternehmensgruppen Beschäftigten. Auch die zweite

Grafik 3

Beschäftigung in Unternehmensgruppen nach Beschäftigtengrößenklassen im Berichtsjahr 2022 in %



Unterklasse enthält einen relativ hohen Anteil aller Beschäftigten in Unternehmensgruppen (16%) und dies bei einem Kreis von ungefähr 200 Unternehmensgruppen. Insgesamt verdeutlicht diese Darstellung eine erhebliche ökonomische Bedeutung eines Bruchteils der Unternehmensgruppen. [↗ Grafik 3](#)

Als Nächstes wird die Teilpopulation der rund 5 000 größten Unternehmensgruppen betrachtet. Diese knapp 2 % aller Unternehmensgruppen sind die Zielpopulation der sogenannten Manuellen Unternehmensgruppen-

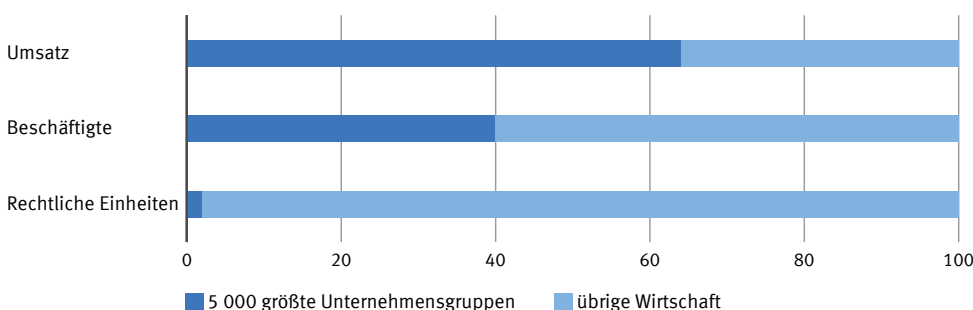
bearbeitung im statistischen Unternehmensregister. Sie wurden ermittelt, indem für verschiedene Verwendungszwecke⁶ die jeweils bedeutendsten Unternehmensgruppen identifiziert wurden und daraus eine Abgrenzung aus Nutzungssicht erstellt wurde. Sie sollen durch manuelle Bearbeitung einer höheren Qualitätspflege unterliegen als die übrigen 98% der Unternehmensgruppen, für die maschinelle Pflegeprozesse erfolgen. Der sich ergebende Umfang von rund 5 000 Gruppen stellt einen notwendigen Kompromiss zwischen Aufwand und Nutzen dar: Eine manuelle Bearbeitung aller Unternehmensgruppen wäre ressourcenintensiv und nicht angemessen, während die Konzentration auf die wirtschaftlich bedeutendsten Gruppen eine zielgerichtete und aufwandseffiziente Form der Registerarbeit darstellt.

Diese 5 000 bedeutendsten Unternehmensgruppen haben zwar nur einen Anteil von 2 % an den deutschen Rechtlichen Einheiten, jedoch 40 % an den Beschäftigten und 64 % an den Umsätzen der deutschen Wirtschaft. Dies verdeutlicht die wirtschaftliche Relevanz dieser Gruppen und wie notwendig eine besonders sorgfältige Qualitätssicherung ist. Auch im europäischen Vergleich zeigt sich, dass eine gezielte Priorisierung der größten und wirtschaftlich bedeutendsten Unternehmensgruppen ein gängiges Vorgehen ist. Die Statistischen Ämter des Bundes und der Länder arbeiten daran, hierfür die notwendigen Kapazitäten aufzubauen. [↗ Grafik 4](#)

⁶ Zu den Nutzenden der Daten zu Unternehmensgruppen siehe Andom und andere (2025). Die Bedeutung wird hier nicht allein durch Umsätze und Beschäftigte abgeschätzt, sondern auch durch weitere Kriterien, die ihre Bedeutung für verschiedene Nutzungen zeigen.

Grafik 4

Bedeutung der 5 000 größten Unternehmensgruppen in Deutschland im Berichtsjahr 2022 in %



5

Bedeutung der Unternehmensgruppen in den verschiedenen Wirtschaftsbereichen

Im Folgenden wird wieder die Gesamtheit der Unternehmensgruppen untersucht. In welchen Bereichen der Wirtschaft sind die Unternehmensgruppen angesiedelt? Eine Darstellung, in welchen Branchen Unternehmensgruppen tätig sind, ist oft nicht trivial. Ältere, auch international gebräuchliche Ansätze haben hierfür vereinfachend den Aktivitätscode (WZ-Code auf der 5-Steller-Ebene der Klassifikation der Wirtschaftszweige, Ausgabe 2008 – WZ 2008) einer prominenten Rechtlichen Einheit, etwa des Gruppenoberhauptes oder der „Deutschen Entscheidungseinheit“ (DEE-WZ), als Indikator für die Tätigkeit der Gruppe herangezogen. Dies liefert aber kein valides Bild. Es werden viel zu hohe Anteile an Beschäftigung und Umsatz im Wirtschaftsabschnitt M „Erbringung von freiberuflichen, wissenschaftlichen und technischen Dienstleistungen“ ausgewiesen, in dem die Unternehmensdienstleistungen angesiedelt sind. Diese stellen oft den wirtschaftlichen Schwerpunkt der steuernden Rechtlichen Einheiten dar, nicht jedoch die Wirtschaftstätigkeit der Gruppen selbst. Dieses Vorgehen wurde im Statistischen Bundesamt als wenig belastbar eingeschätzt und verworfen.¹⁷ Bereits für das Berichtsjahr 2021 wurde der Aktivitätscode des bedeutendsten Unternehmens¹⁸ innerhalb der Gruppe verwendet. Ab dem Berichtsjahr 2022 wird der Aktivitätscode gemäß eines im statistischen Unternehmensregister entwickelten WZ-Bestimmungsalgorithmus aus den gewichteten Aktivitäten aller Unternehmen der Unternehmensgruppe ermittelt. Das von der europäischen Wirtschaftsklassifikation NACE beschriebene Vorgehen wird damit umgesetzt (Top-down-Methode). Die Gewichtung erfolgt anhand der Beschäftigten sowie der Umsätze. Für das Berichtsjahr 2022 wird hier eine Auswertung dargestellt.¹⁹ Dabei sind auch die Ergebnisse dargestellt, die sich nach den wie beschrieben verworfenen Verfahren erzeugen lassen. ➡ Grafik 5 auf Seite 92

7 Nur für das Berichtsjahr 2020 finden sich Daten aus dieser Herangehensweise in der Datenbank des statistischen Unternehmensregisters.

8 Die Bedeutung wurde anhand der Anzahl der Beschäftigten gemessen.

9 Eine Auswertung der WZ-Abschnitte T und U erfolgt nicht, da diese Bereiche für die Unternehmensgruppen von geringer Relevanz sind.

Es fällt auf, dass die Anteile gemäß WZ-Bestimmungsalgorithmus in nahezu allen Wirtschaftsabschnitten identisch hoch ausfallen wie bei der Annahme, das jeweils bedeutendste Unternehmen innerhalb der Gruppe wäre ein guter Indikator für den wirtschaftlichen Schwerpunkt der Gruppe. Diese Hypothese wird also durch die Empirie gestützt.

Die bedeutendsten Wirtschaftsabschnitte der Gruppentätigkeit sind die Abschnitte C „Verarbeitendes Gewerbe“ und G „Handel; Instandhaltung und Reparatur von Kraftfahrzeugen“. Hier zeigen sich insbesondere in der Auswertung nach Umsätzen Schwerpunkte der Tätigkeit der Unternehmensgruppen. Die Beschäftigung verteilt sich dagegen weniger stark ausgeprägt.

Ein anderes Bild ergibt der Blick auf die Anzahl der in Unternehmensgruppen angesiedelten Rechtlichen Einheiten. Diese zeigen eine deutlich andere Verteilung als die ökonomischen Variablen der Unternehmensgruppen selbst.

Aber auch mit dieser methodisch korrekten Ermittlung ist die Frage der angemessenen Gliederungstiefe für aussagekräftige Veröffentlichungen noch nicht gelöst. Die deutsche Klassifikation der Wirtschaftszweige (mit einer Gliederungstiefe, die oft Fünfsteller enthält) lässt sich nicht in einfacher Weise nutzen, um die Tätigkeiten von großen, oft diversifiziert tätigen Unternehmensgruppen adäquat einzufangen. Mit der Verwendung eines Vier- oder Fünfstellers würde wohl in vielen Fällen eine Spezifität und Homogenität der ausgewiesenen Tätigkeit suggeriert, die in der Realität nicht gegeben ist. In der hier vorgestellten Darstellung wurde zunächst die Gliederungsebene der Abschnitte der europäischen Wirtschaftszweigklassifikation NACE¹⁰ verwendet.

Des Weiteren wird untersucht, in welchen Wirtschaftsbereichen (abgegrenzt nach der Tätigkeit von Rechtlichen Einheiten) Unternehmensgruppen die höchste ökonomische Bedeutung haben. Diese kann über den Anteil der in Unternehmensgruppen Beschäftigten der WZ-Bereiche gemessen werden.¹¹ Dabei fällt auf, dass im WZ-Abschnitt D „Energieversorgung“ über 90 % der Beschäftigten in Unternehmensgruppen arbeiten. Zusätzlich ist zu erkennen, dass insbesondere die weiteren

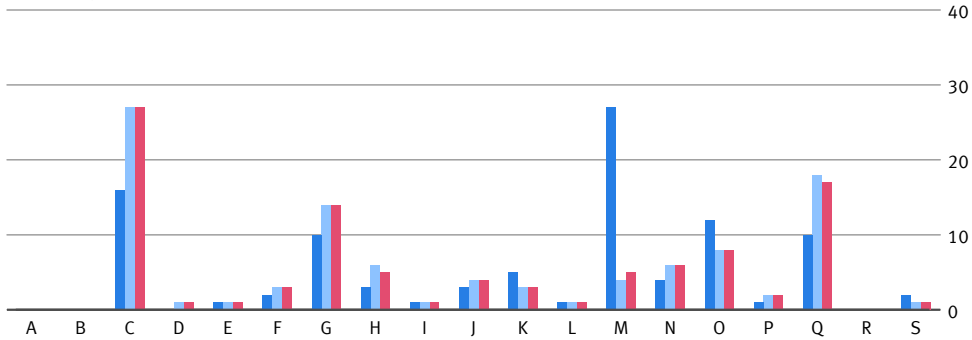
10 Statistische Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft, Rev. 2 (NACE Rev. 2).

11 Auch hier erfolgt keine Auswertung der WZ-Abschnitte T und U.

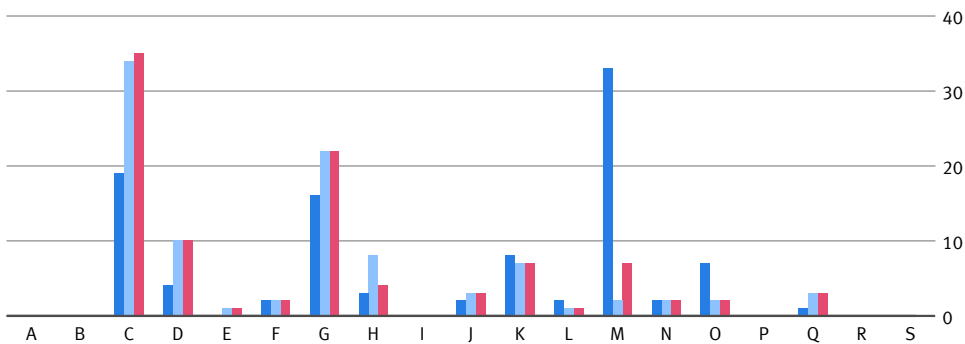
Grafik 5

Vergleich verschiedener Verfahren zur WZ-Bestimmung von Unternehmensgruppen
für das Berichtsjahr 2022
in %

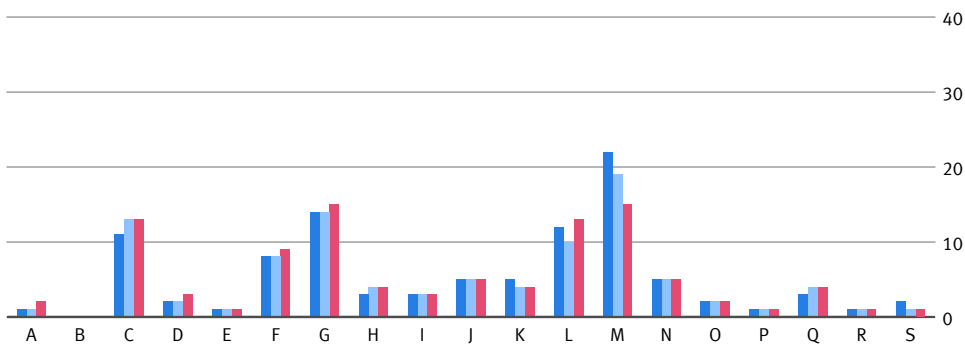
Beschäftigte



Umsätze



Rechtliche Einheiten



■ WZ der Deutschen Entscheidungseinheit ■ WZ des bedeutendsten Unternehmens
■ WZ-Bestimmungsalgorithmus

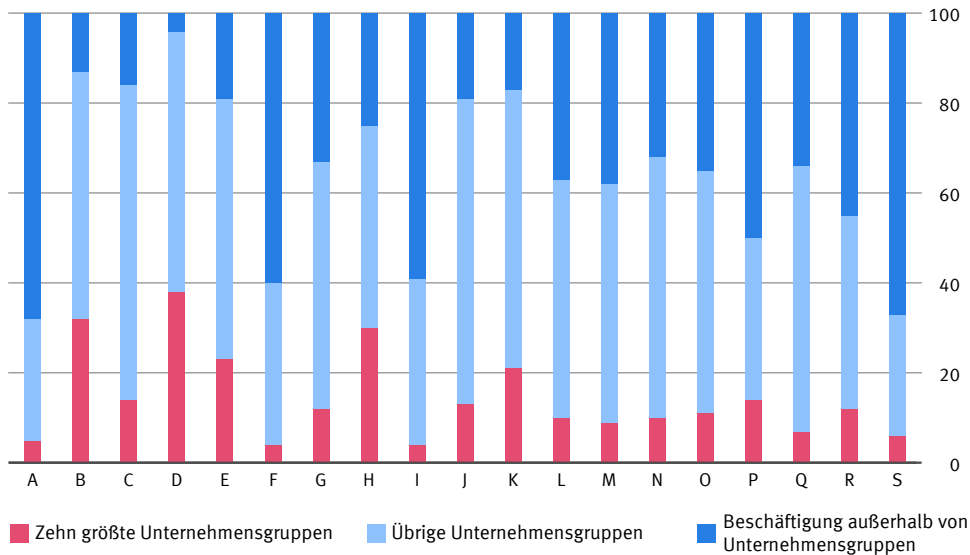
Abschnitte der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

WZ-Bereiche B „Bergbau und Gewinnung von Steinen und Erden“, C „Verarbeitendes Gewerbe“, E „Wasserversorgung; Abwasser- und Abfallentsorgung und Beseitigung von Umweltverschmutzungen“, J „Information und Kommunikation“ und K „Erbringung von freiberuf-

lichen, wissenschaftlichen und technischen Dienstleistungen“ besonders stark durch Unternehmensgruppen bestimmt werden. Die Anteile der Beschäftigten, die hier in Unternehmensgruppen arbeiten, liegen über 80%. Bei näherer Betrachtung der Top-10-Gruppen in jedem

Grafik 6

Bedeutung der zehn größten Unternehmensgruppen nach Beschäftigung in den WZ-Abschnitten im Berichtsjahr 2022
in %



Abschnitte der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

WZ-Abschnitt, die die meisten Beschäftigten umfassen, fällt außerdem auf, dass insbesondere der Abschnitt D „Energieversorgung“ von den Top-Gruppen bestimmt wird. Fast 40 % aller in Deutschland Beschäftigten dieses Abschnittes arbeiten für nur zehn Unternehmensgruppen. [↗ Grafik 6](#)

6

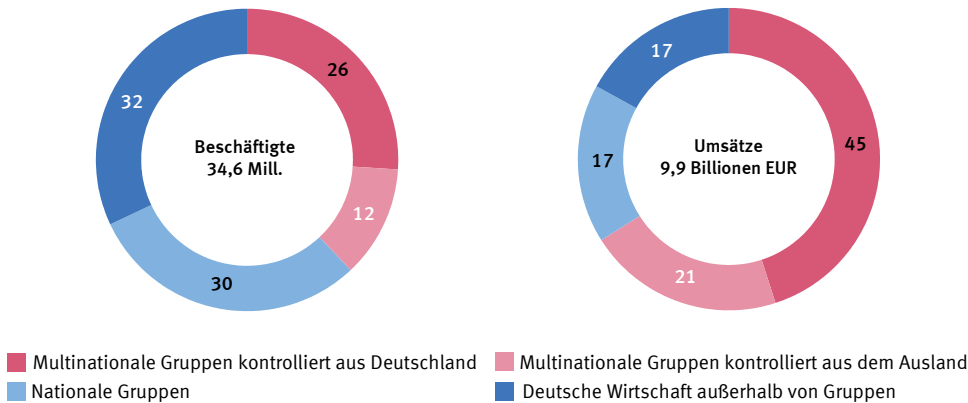
Von wo aus wird die deutsche Wirtschaft gesteuert?

Die deutsche Wirtschaft ist stark mit dem Ausland verflochten. Auch diese These lässt sich anhand der Unternehmensgruppen im statistischen Unternehmensregister empirisch überprüfen. Bei der Betrachtung, ob Unternehmensgruppen sich nur aus deutschen Rechtlichen Einheiten zusammensetzen oder ob sie Teile von multinationalen Gruppen sind, zeigt sich Folgendes: Bei rein inländisch aufgestellten, also nationalen Unternehmensgruppen sind 30 % der deutschen Beschäftigten tätig und diese erwirtschaften 17 % aller deutschen

Umsätze. Dagegen sind bei multinationalen Gruppen, deren Führung in Deutschland angesiedelt ist, 25 % aller Beschäftigten tätig und sie erwirtschaften 45 % der deutschen Umsätze. Bei deutschen Teilen von multinationalen Gruppen, deren Gruppenoberhaupt im Ausland sitzt, sind in Deutschland 12 % der sozialversicherungspflichtig Beschäftigten tätig; diese erwirtschaften 21 % aller Umsätze in Deutschland. An diesen Zahlen erkennt man nicht nur den Umfang der internationalen Verflechtung, sondern auch, dass der zu multinational agierenden Gruppen gehörende Teil der deutschen Wirtschaft überproportional hohe Umsätze je Beschäftigten erwirtschaftet. [↗ Grafik 7 auf Seite 94](#)

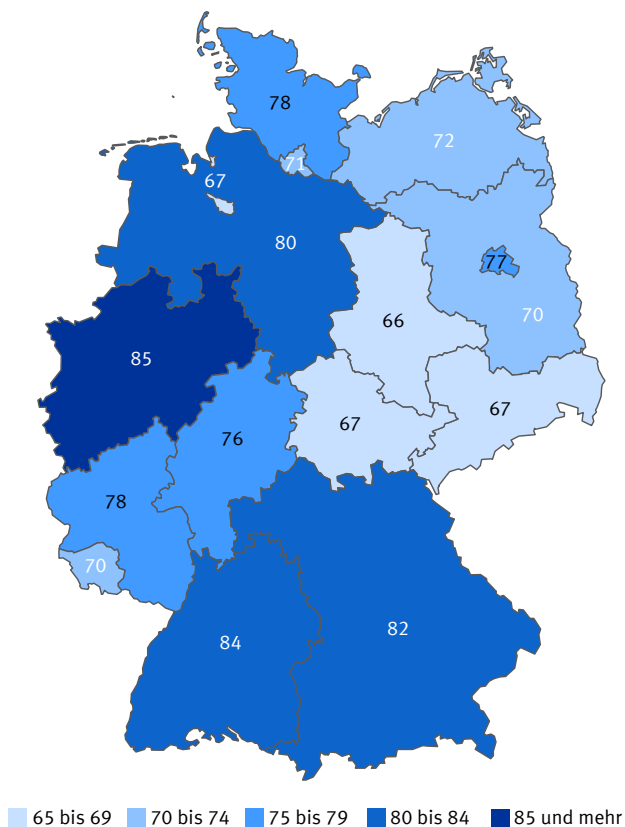
Grafik 7

Multinationalität von Unternehmensgruppen im Berichtsjahr 2022
in %



Grafik 8

Beschäftigung mit Kontrolleinheit im eigenen Bundesland im Berichtsjahr 2022
in %



Stand: 2022
© GeoBasis-DE / BKG 2022

Weitere Untersuchungen betreffen die Standorte der strategischen Steuereinheiten der Unternehmensgruppen in Deutschland. Diese Funktion innerhalb der Unternehmensgruppen nimmt die Rechtliche Einheit wahr, die im statistischen Unternehmensregister als Deutsche Entscheidungseinheit bezeichnet wird. Die folgende Auswertung zeigt¹², welcher Anteil der Beschäftigten in den Bundesländern von Deutschen Entscheidungseinheiten geleitet wird, die im eigenen Bundesland angesiedelt sind.¹³

Dieser Anteil liegt in den meisten Bundesländern über 70%. Lediglich die Bundesländer Bremen, Sachsen, Sachsen-Anhalt und Thüringen weisen geringere Anteile auf. Tiefergehende Analysen – beispielsweise aus welchen anderen Bundesländern die Leitung erfolgt oder in welchen Bundesländern die „eigenen“ Unternehmensgruppen besonders viel Gewicht haben – wären noch durchzuführen. [↗ Grafik 8](#)

Die globalen Leitungen der Unternehmensgruppen, also die Gruppenoberhäupter, können im In- oder im Ausland angesiedelt sein. In der IFATS¹⁴-Statistik (Nahm, 2011)

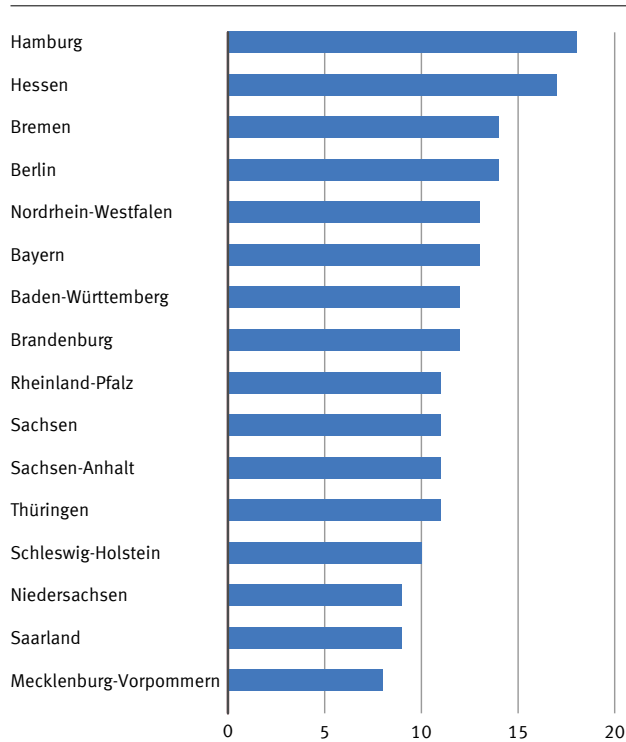
¹² Im Gegensatz zu bisherigen Auswertungen, die sich auf die Beschäftigten der Rechtlichen Einheiten oder Unternehmensgruppen stützen, basieren alle bundeslandspezifischen Analysen auf Beschäftigtenzahlen auf Niederlassungsebene. Dies ermöglicht eine möglichst präzise regionale Darstellung der Ergebnisse.

¹³ Hierbei ist zu beachten, dass bei Beschäftigten, die nicht in Unternehmensgruppen, sondern in meist kleineren eigenständigen Firmen beschäftigt sind, ohnehin oft die Leitung im eigenen Bundesland sitzt.

¹⁴ Inward-Foreign Affiliates Statistics.

Grafik 9

Aus dem Ausland kontrollierte Beschäftigung in den Bundesländern im Berichtsjahr 2022
Anteil in %



des Statistischen Bundesamtes werden die Unternehmen in Deutschland daraufhin untersucht, ob ein Gruppenoberhaupt im Ausland angesiedelt ist und wenn ja, wo. Auf der Basis der Daten des statistischen Unternehmensregisters wurde eine ähnliche Auswertung für die Beschäftigten in den Bundesländern vorgenommen.

➔ Grafik 9

Zu sehen ist, dass der Einfluss aus dem Ausland am größten in den Bundesländern Hamburg und Hessen ist.

Bei der Analyse zu ausländischer Kontrolle durch Unternehmensgruppen kann es auch interessant sein zu identifizieren, woher genau, also aus welchem Staat, die Kontrolle ausgeht. Dies zeigt ➔ Tabelle 1, in der für alle Bundesländer die jeweils drei einflussreichsten Staaten, gemessen an den Beschäftigten der ausländisch beherrschten Unternehmensgruppen, aufgeführt sind. Auch wenn die absoluten Anteile der jeweils betroffenen Beschäftigten gering sind, so fällt auf, dass es wiederkehrend die gleichen Staaten sind, die den größten Einfluss innehaben, und zwar die Vereinigten Staaten, die Schweiz, Frankreich und die Niederlande.

Tabelle 1

Einfluss der jeweils drei einflussreichsten Staaten auf die Beschäftigten der Bundesländer im Berichtsjahr 2022

	Staat TOP 1	Staat TOP 2	Staat TOP 3	Anteil Beschäftigte TOP 1 in %	Anteil Beschäftigte TOP 2 in %	Anteil Beschäftigte TOP 3 in %
Schleswig-Holstein	Vereinigte Staaten	Schweiz	Dänemark	2	1	1
Hamburg	Niederlande	Vereinigte Staaten	Frankreich	3	2	2
Niedersachsen	Vereinigte Staaten	Niederlande	Frankreich	2	1	1
Bremen	Niederlande	Luxemburg	Frankreich	2	2	1
Nordrhein-Westfalen	Vereinigte Staaten	Frankreich	Niederlande	2	1	1
Hessen	Vereinigte Staaten	Frankreich	Schweiz	4	2	2
Rheinland-Pfalz	Vereinigte Staaten	Frankreich	Luxemburg	3	1	1
Baden-Württemberg	Schweiz	Vereinigte Staaten	Frankreich	2	2	1
Bayern	Vereinigte Staaten	Schweiz	Frankreich	3	2	1
Saarland	Vereinigte Staaten	Frankreich	Niederlande	2	1	1
Berlin	Vereinigte Staaten	Niederlande	Schweiz	3	2	1
Brandenburg	Vereinigte Staaten	Frankreich	Luxemburg	2	2	1
Mecklenburg-Vorpommern	Schweiz	Niederlande	Vereinigte Staaten	1	1	1
Sachsen	Schweiz	Vereinigte Staaten	Frankreich	1	1	1
Sachsen-Anhalt	Schweiz	Niederlande	Luxemburg	2	1	1
Thüringen	Niederlande	Vereinigte Staaten	Schweiz	2	1	1
Deutschland	Vereinigte Staaten	Schweiz	Frankreich	2	1	1


Auswertungen aus dem EuroGroups Register

Eurostat verpflichtet die Europäischen Mitgliedstaaten, jährlich Daten zu den nationalen Teilen der in den Registern enthaltenen multinationalen Unternehmensgruppen zu übersenden. Die verschiedenen Lieferungen der Mitgliedstaaten werden hierbei im EuroGroups Register zusammengeführt, um die gesamte multinationale Unternehmensgruppe im europäischen Kontext abzubilden. Das statistische Unternehmensregister erhält hierbei nur Zugriff auf Daten zu den in Europa aktiven multinationalen Unternehmensgruppen, soweit sie mindestens eine deutsche Rechtliche Einheit enthalten. Damit lassen sich zur europäischen Gruppenlandschaft nur bedingt Auswertungen erstellen.

Insgesamt enthält das EuroGroups Register rund 52 000 multinationale Unternehmensgruppen, die in Deutschland aktiv sind. Ungefähr 84 % dieser Gruppen werden aus Europa kontrolliert. Innerhalb dieser Gruppen gibt es 15 600 Gruppen, die aus Deutschland kontrolliert werden. Dies sind 30 % aller in Deutschland tätigen multinationalen Gruppen. Bereits diese ersten Zahlen zeigen, dass Deutschland im europäischen Kontext sehr großes Gewicht besitzt.

Eine detailliertere Betrachtung internationaler Aspekte zu Unternehmensgruppen erfordert noch tiefergehende Analysen und kann daher in diesem Aufsatz noch nicht angeboten werden.

lichen Einfluss auf die deutsche Wirtschaft. Sie weisen hierbei insbesondere überproportional hohe Umsätze je Beschäftigten aus.

Die Abbildung der Unternehmensgruppen im statistischen Unternehmensregister ermöglicht eine detaillierte empirische Untersuchung ihrer wirtschaftlichen Bedeutung. Bereits die ersten hier präsentierten Ergebnisse zeigen, dass eine präzise Erfassung und regelmäßige Qualitätspflege von Daten zu Unternehmensgruppen entscheidend sind, um auf ihrer Basis fundierte wirtschaftspolitische Entscheidungen zu ermöglichen. Gleichzeitig verdeutlichen sie die methodischen Herausforderungen bei der Abbildung dieser oft hochkomplexen Strukturen. Die Weiterentwicklung der Qualitätssicherung im Unternehmensregister, insbesondere durch die „Manuelle Unternehmensgruppenbearbeitung“ (Andom und andere, 2025; hier: Seite 99 ff.), wird daher auch künftig eine zentrale Rolle spielen, um zuverlässige und detaillierte Daten bereitzustellen. 

7

Fazit

Die Analyse der Unternehmensgruppen im statistischen Unternehmensregister zeigt die große wirtschaftliche Bedeutung dieser Einheiten. Obwohl Unternehmensgruppen nur einen geringen Anteil der deutschen Rechtlichen Einheiten ausmachen, erwirtschaften sie mehr als drei Viertel des Umsatzes und stellen mehr als die Hälfte der Beschäftigten in Deutschland. Besonders auffallend ist die hohe Konzentration wirtschaftlicher Leistung bei relativ wenigen großen Unternehmensgruppen. In bestimmten Wirtschaftsbereichen ist dieser Einfluss zudem besonders stark ausgeprägt.

Auch die internationale Verflechtung spielt eine bedeutende Rolle. Multinationale Unternehmensgruppen mit Kontrolle in Deutschland und im Ausland haben erheb-

LITERATURVERZEICHNIS

Andom, Senait/Bingel, Johanna/Urban, Adrian. [*Methodische Weiterentwicklungen für Unternehmensgruppen im Statistischen Unternehmensregister*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2025, Seite 91 ff.

Kleber, Birgit/Sturm, Roland/Tümmler, Thorsten. [*Ergebnisse zu Unternehmensgruppen aus dem Unternehmensregister*](#). In: Wirtschaft und Statistik. Ausgabe 6/2010, Seite 527 ff.

Nahm, Matthias. [*Inward-FATS – Auslandskontrollierte Unternehmen in Deutschland 2008*](#). In: Wirtschaft und Statistik. Ausgabe 9/2011, Seite 899 ff.

Sturm, Roland/Tümmler, Thorsten/Opfermann, Rainer. [*Unternehmensverflechtungen im statistischen Unternehmensregister*](#). In: Wirtschaft und Statistik. Ausgabe 8/2009, Seite 764 ff.

RECHTSGRUNDLAGEN

Handelsgesetzbuch (HGB) in der im Bundesgesetzblatt Teil III, Gliederungsnummer 4100-1, veröffentlichten bereinigten Fassung, das zuletzt durch Artikel 1 des Gesetzes vom 28. Februar 2025 (BGBl. I Nr. 69) geändert worden ist.



Isabel Herrle

hat Volkswirtschaftslehre und Politikwissenschaft studiert und ist seit 2023 Referentin im Referat „Preis- und Volumenmessung, Außenwirtschaftliche Verflechtung, Input-Output-Analysen“ des Statistischen Bundesamtes. Ihre Haupttätigkeiten betreffen die methodische Weiterentwicklung der Exporte und Importe nach dem Konzept der Volkswirtschaftlichen Gesamtrechnungen.



Lucas Rohleder

hat Volkswirtschaftslehre studiert und ist seit 2021 Referent im Referat „Input-Output-Tabellen in jeweiligen Preisen“ des Statistischen Bundesamtes. Seine Haupttätigkeiten betreffen die methodische Weiterentwicklung und Modernisierung der Input-Output-Rechnung.

ERWEITERTE AUFKOMMENS- UND VERWENDUNGSTABELLEN – EINE DARSTELLUNG DER AKTIVITÄTEN UNTERSCHIEDLICHER UNTERNEHMENSKATEGORIEN

Isabel Herrle, Lucas Rohleder

➤ **Schlüsselwörter:** Input-Output-Rechnung – Unternehmensheterogenität – Produktionsstrukturen – Vorleistungen – Volkswirtschaftliche Gesamtrechnungen

ZUSAMMENFASSUNG

Wirtschaftsanalysen erfordern, je nach Fragestellung, detaillierte Untergliederungen der Wirtschaft nach unterschiedlichen Kategorien oder Klassifikationen. Die Volkswirtschaftlichen Gesamtrechnungen stellen dazu Aufkommens- und Verwendungstabellen zur Verfügung, mit Informationen nach Wirtschaftszweigen und Gütergruppen. Die bestehenden Aufkommens- und Verwendungstabellen wurden in einem EU-Projekt nach Größe, Exportintensität und ausländischer beziehungsweise inländischer Kontrolle der enthaltenen Unternehmen weiter unterteilt. Der Artikel beschreibt den internationalen Rahmen, die Methodik und die Datenquellen sowie erste Ergebnisse für das Berichtsjahr 2021. Diese zeigen das Potenzial der erweiterten Aufkommens- und Verwendungstabellen zur Untersuchung von Produktions- und Vorleistungsstrukturen.

➤ **Keywords:** input-output accounts – heterogeneity of enterprises – production structures – intermediate consumption – national accounts

ABSTRACT

Depending on the question being addressed, economic analysis requires a detailed classification or breakdown of the economy by various categories. The national accounts provide supply and use tables for this purpose, which contain information broken down by sector of economic activity and product group. As part of an EU project, the existing supply and use tables were broken down further according to size, export intensity and foreign or domestic control of the enterprises included. This article describes the international framework, the methodology and the data sources, and presents first results for the reference year of 2021.

1

Einleitung

Für makroökonomische Analysen von wirtschaftlichen Strukturen ist es notwendig, Wirtschaftseinheiten gemäß ihrer Merkmale einheitlich zu klassifizieren. In den Volkswirtschaftlichen Gesamtrechnungen erfolgt dies auf unterschiedliche Art und Weise, beispielsweise nach den Aktivitäten von Wirtschaftseinheiten.

In den Aufkommens- und Verwendungstabellen der Volkswirtschaftlichen Gesamtrechnungen wird die Wirtschaft nach Wirtschaftszweigen (WZ) und Gütergruppen gegliedert.¹ Aufkommenstabellen zeigen die Werte der in den jeweiligen Wirtschaftszweigen produzierten Waren und Dienstleistungen, ergänzt um Importe. Die Verwendungstabelle zeigt, wie diese produzierten und importierten Waren und Dienstleistungen entweder als Vorleistungen² nach Wirtschaftszweigen, für den staatlichen und privaten Konsum, Investitionen oder

für Exporte verwendet werden. Anhand dieser Tabellen lassen sich Aussagen über Gruppen von Wirtschaftseinheiten treffen, die sich charakteristische Merkmale und Tätigkeiten teilen.

Um aktuelle Forschungsfragen untersuchen zu können, hat sich in den letzten Jahren das Interesse nach weiter disaggregierten Daten verstärkt. Unterteilt man die Wirtschaftszweige nach bestimmten Unternehmensmerkmalen, so wird deutlich, dass sich die Aktivitäten von Unternehmen auch innerhalb eines Wirtschaftszweiges unterscheiden können und sie deshalb möglicherweise unterschiedlich auf geänderte Rahmenbedingungen oder exogene Schocks reagieren. Relevante Fragestellungen lauten hier beispielsweise: Benötigen große Metallerzeugnishersteller die gleichen Vorleistungen wie kleinere Hersteller? Produzieren exportorientierte Maschinenbauer andere Produkte als die auf den Binnenmarkt konzentrierten Maschinenbauer? Wie unterscheidet sich die Bruttowertschöpfung von ausländisch und inländisch kontrollierten Unternehmen?

Erweiterte Aufkommens- und Verwendungstabellen können einen wichtigen Beitrag zur Beantwortung dieser Fragen leisten, indem sie die Unternehmen in den einzelnen Wirtschaftszweigen weiter nach spezifischen Merkmalen, wie in [Übersicht 1](#) dargestellt, aufgliedern. Im Rahmen eines Projekts des Statistischen Amtes der Europäischen Union (Eurostat) hat das Statistische Bundesamt hierzu eine Machbarkeitsstudie durchgeführt

- 1 Derzeit werden in den Volkswirtschaftlichen Gesamtrechnungen die Statistische Systematik der Wirtschaftszweige in der Europäischen Gemeinschaft (NACE Rev. 2 – [Nomenclature générale des Activités économiques dans les Communautés Européennes](#)) und unter anderem das Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019, angewendet.
- 2 Gemäß dem [Europäischen System Volkswirtschaftlicher Gesamtrechnungen \(ESVG\) 2010](#) umfassen Vorleistungen alle Waren und Dienstleistungen, die für den Produktionsprozess von anderen Produzenten verbraucht, verarbeitet oder umgewandelt werden.

Übersicht 1

Schematische Darstellung einer im Bereich des Bergbaus und des Verarbeitenden Gewerbes nach Exporteuren (X) und Nicht-Exporteuren (nicht-X) erweiterten Aufkommenstabelle (eAT)

		Landwirtschaft	Bergbau und Verarbeitendes Gewerbe				Dienstleistungen (DL)			Summe Produktion	Importe	Summe Aufkommen
Gütergruppen	WZ		WZ A		WZ B		...	WZ X	WZ Y	...		
			X	nicht-X	X	nicht-X	...					
Waren	W1											
	W2											
	W3											
	...											
Dienstleistungen	D1											
	D2											
	D3											
	...											
Gesamt WZ												

und erste experimentelle Tabellen zu erweiterten Aufkommens- und Verwendungstabellen für das Berichtsjahr 2021 (Veröffentlichungsstand der genutzten Daten: 2024) erstellt. Die Wirtschaftszweige wurden dabei in die Unternehmenskategorien Unternehmensgröße nach Beschäftigtenzahlen, Exportintensität und ausländische beziehungsweise inländische Kontrolle unterteilt.

Der Artikel beschreibt die angewendete Methodik, den internationalen Diskurs, die verwendeten Daten sowie erste Ergebnisse.

2

Internationaler Rahmen

In der Forschung über Globalisierung ist Unternehmensheterogenität, also die Unterschiede zwischen Unternehmen beispielsweise hinsichtlich ihrer Größe, Produktion und internationalen Aktivitäten, seit den frühen 2000er-Jahren zu einem zentralen Thema geworden. Wichtige Forschungsergebnisse sind:

- › Exporte sind stark konzentriert: Ein kleiner Teil großer Unternehmen ist für den Großteil der Exporte verantwortlich (Bernard und andere, 2011).
- › Exporteure sind größer, kapitalintensiver, produktiver, beschäftigen höherqualifizierte Fachkräfte und benötigen andere Vorleistungen als Nicht-Exporteure (Bernard/Jensen, 1995).
- › Kleine und mittlere Unternehmen exportieren häufig nicht direkt, sondern beteiligen sich indirekt am Export, entweder durch die Bereitstellung von Vorleistungen für Exporteure oder durch den Export über Großhändler (Chong und andere, 2019).
- › Fast zwei Drittel der weltweiten Exporte entfallen auf multinationale Unternehmensgruppen und ihr Netz ausländischer Tochtergesellschaften (UNCTAD, 2013; Miroudot/Rigo, 2022).

Um diese Heterogenität besser abbilden und erforschen zu können, hat die Expertengruppe für erweiterte Aufkommens- und Verwendungstabellen (EGESUT) der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) verschiedene Möglichkeiten erarbeitet, das bestehende Rahmenwerk der Aufkommens- und Verwendungstabellen nach Unternehmensmerkma-

len aufzuschlüsseln. Folgende Unterteilungen werden im internationalen Diskurs als besonders relevant für Wirtschaftsanalysen angesehen:

- › Unternehmensgröße [kleine und mittlere Unternehmen (KMU) gegenüber großen Unternehmen (GU)],
- › Exportintensität (exportintensive gegenüber binnenmarktorientierten Unternehmen) und
- › Besitzstruktur (inländisch kontrollierte gegenüber ausländisch kontrollierten Unternehmen).

Aufgrund des wachsenden Interesses an diesen Unterteilungen wird im neuen System of National Accounts (SNA) 2025 empfohlen, erweiterte Aufkommens- und Verwendungstabellen zu berechnen.¹³ Mehrere Länder haben bereits Erweiterungen von Aufkommens- und Verwendungstabellen erstellt. Die OECD-Expertengruppe stellte dabei fest, dass es weder notwendig noch möglich ist, dass alle Länder sämtliche Erweiterungen vollumfänglich vornehmen oder dieselbe Methodik anwenden (OECD, 2015). Grund dafür sind unterschiedliche Datenverfügbarkeiten in verschiedenen Ländern, welche eine einheitliche Vorgehensweise erschweren. Die Niederlande und auch Dänemark verfügen beispielsweise über ausreichende Daten, um einige Dienstleistungswirtschaftszweige nach Unternehmensgrößenklassen aufzuschlüsseln. Österreich und Belgien dagegen müssen ihre erweiterten Aufkommens- und Verwendungstabellen aufgrund von Datenlücken auf das Verarbeitende Gewerbe und den Warenhandel beschränken. Neben der Datenverfügbarkeit bestimmen auch unterschiedliche wirtschaftliche Schwerpunkte einzelner Volkswirtschaften die Prioritäten und den individuellen Nutzendenbedarf.

Situation in Deutschland

Deutschland gilt als Exportnation und hat einen starken Mittelstand, der das Rückgrat der deutschen Wirtschaft bildet (Bundesministerium für Wirtschaft und Klimaschutz, 2022). Mittelstandsunternehmen bieten oft anderen Unternehmen maßgeschneiderte Produkte an; ihre Ausgaben für Forschung und Entwicklung sind im Verhältnis zum Umsatz doppelt so hoch wie in anderen Ländern (Venohr und andere, 2014). Viele Mittelständler

3 Das System of National Accounts (SNA) ist ein international standardisiertes Rahmenwerk zur Erfassung, Berechnung und Analyse der wirtschaftlichen Aktivitäten eines Landes und dient als Grundlage für die Volkswirtschaftlichen Gesamtrechnungen.

zählen zur Gruppe der kleinen und mittleren Unternehmen und sind teilweise tief in globale Wertschöpfungsketten integriert. Daher sind in Deutschland Untergliederungen nach Unternehmensgröße und Exportintensität von besonderem Interesse.

Nach einer umfassenden Prüfung der dem Statistischen Bundesamt zur Verfügung stehenden Datenquellen und unter Berücksichtigung der oben genannten Hintergründe werden die deutschen Aufkommens- und Verwendungstabellen derzeit gemäß den von der OECD-Expertengruppe empfohlenen Unterteilungen nach Unternehmensgröße, Exportintensität und inländischer beziehungsweise ausländischer Kontrolle differenziert.

3

Datengrundlage

Ausgangspunkt für die Erstellung von erweiterten Aufkommens- und Verwendungstabellen bilden die regulären Aufkommens- und Verwendungstabellen des Statistischen Bundesamtes. Sie zeigen die gütermäßige Verflechtung der deutschen Volkswirtschaft und sind integraler Bestandteil der Volkswirtschaftlichen Gesamtrechnungen. Die Aufkommenstabelle zeigt den Wert der im Inland produzierten Güter, gegliedert nach Gütergruppen und Wirtschaftsbereichen, sowie die Importe. Die Verwendungstabelle stellt dar, wie die im Inland produzierten und importierten Güter verwendet werden, und zwar nach privatem und staatlichem Konsum, Investitionen und Exporten und nach Vorleistungen der Wirtschaftsbereiche.

Informationen über Produktion und Vorleistungseinsatz auf Unternehmensebene sind für die Berechnung der erweiterten Aufkommens- und Verwendungstabellen elementar. Um die Aufkommenstabelle nach Unternehmenskategorien zu erweitern, wird auf die Produktionserhebung zurückgegriffen. Die notwendigen Informationen über den Vorleistungseinsatz der Unternehmen für die erweiterte Verwendungstabelle sind in der Material- und Wareneingangserhebung enthalten. Die in diesen beiden Statistiken erfassten Unternehmen dienen somit als Grundgesamtheit und bilden den Ausgangspunkt für die Verknüpfung mit weiteren Statistiken. Die Unternehmen in der Produktionserhebung sind allerdings nur teilweise auch in der Material- und Wareneingangserhe-

bung erfasst. Dies schränkt die Vergleichbarkeit beider Statistiken ein.

3.1 Produktionserhebung

Zur vierteljährlichen Produktionserhebung melden Rechtliche Einheiten mit 20 oder mehr Beschäftigten, die entweder selbst zu den NACE-Abschnitten B (Bergbau und Gewinnung von Steinen und Erden) und C (Verarbeitendes Gewerbe) gehören oder Teil von Unternehmen in diesen Bereichen sind.¹⁴ Die vierteljährliche Produktionsstatistik ist eine Vollerhebung mit Abschneidegrenze.

Somit sind qualitativ hochwertige Mikrodaten zur Güterproduktion auf Unternehmensebene für die Wirtschaftsbereiche des Bergbaus und des Verarbeitenden Gewerbes verfügbar. Deshalb werden die erweiterten Aufkommens- und Verwendungstabellen derzeit nur für diese Bereiche erstellt.

3.2 Material- und Wareneingangserhebung

Die Material- und Wareneingangserhebung (MWE) wird alle vier Jahre aus der identischen Stichprobe wie die Kostenstrukturerhebung des Verarbeitenden Gewerbes des jeweiligen Berichtsjahres erhoben; sie enthält Daten zum Material- und Wareneingang von Unternehmen nach Güterarten. Aus der Material- und Wareneingangserhebung lässt sich daher ableiten, welche Güter die einzelnen Unternehmen in ihrer Produktion als Vorleistung verbraucht haben. Die Material- und Wareneingangserhebung ist somit der inputseitige Spiegel der Produktionserhebung. Seit 2006 werden Unternehmen mit 20 bis 49 Beschäftigten nicht mehr befragt, um diese Unternehmen zu entlasten. Die Werte dieser Unternehmen müssen anhand von Daten aus der Kostenstrukturerhebung geschätzt werden.

Wie die Produktionserhebung ist auch die Material- und Wareneingangserhebung nur für die NACE-Abschnitte B (Bergbau) und C (Verarbeitendes Gewerbe) verfügbar. Da die Material- und Wareneingangserhebung nur alle vier Jahre erhoben wird, wurden die erweiterten

¹⁴ In sieben ausgewählten Wirtschaftszweigen liegt die Meldeschwelle bei zehn oder mehr Beschäftigten.

Übersicht 2

Für die erweiterten Aufkommens- und Verwendungstabellen verwendete Datenquellen

Verwendete Statistik	Periodizität/ Art der Statistik	Meldekreis	Verwendete Merkmale
Vierteljährliche Produktionserhebung	<ul style="list-style-type: none"> › monatlich, vierteljährlich › Vollerhebung mit Abschneidegrenzen 	<ul style="list-style-type: none"> › monatlich: Einheiten mit 50 und mehr Beschäftigten › jährlich: Einheiten mit 20 bis unter 50 Erwerbstätigen › Einheiten mit 10 und mehr Beschäftigten aus ausgewählten Wirtschaftszweigen 	<ul style="list-style-type: none"> › Warenproduktion, Reparatur-, Installations- und Verarbeitungsdienstleistungen (CPA-Gütergruppen 01-33)
Material- und Wareneingangserhebung	<ul style="list-style-type: none"> › Erhebung jeweils alle vier Jahre 	<ul style="list-style-type: none"> › Stichprobe aus Grundgesamtheit der Kostenstrukturerhebung › nur Unternehmen mit 50 und mehr Beschäftigten 	<ul style="list-style-type: none"> › Vorleistungen nach Wirtschaftszweigen, Gütergruppen, Anzahl Beschäftigte
AFiD-Panel ¹ Außenhandelsstatistik	<ul style="list-style-type: none"> › Mikrodaten der Intrahandelsstatistik, einschließlich Schätzungen für Unternehmen unterhalb der Meldeschwellen, und der Extrahandelsstatistik 	<ul style="list-style-type: none"> › alle Unternehmen (rechtliche Einheiten), die zur Außenhandelsstatistik (Intra- und Extrahandel) melden, zuzüglich geschätzte Unternehmen unterhalb der Meldeschwellen im Intrahandel 	<ul style="list-style-type: none"> › Export- und Importwert je Unternehmen (Rechtliche Einheiten)
Inward Foreign Affiliates Statistics	<ul style="list-style-type: none"> › Sekundärstatistik, Grundlage sind 12 Strukturstatistiken › Informationen zur ausländischen Kontrolle werden von einem externen Datenanbieter bezogen 	<ul style="list-style-type: none"> › Unternehmen in den Wirtschaftsbereichen NACE B bis S, ohne S94 	<ul style="list-style-type: none"> › Unternehmen unter ausländischer Kontrolle
Unternehmensregister	<ul style="list-style-type: none"> › Register der in Deutschland ansässigen Unternehmen; Pflege unter anderem durch Verwaltungsdaten sowie Rückmeldungen aus statistischen Erhebungen 	<ul style="list-style-type: none"> › alle Unternehmen (Rechtliche Einheiten) mit Sitz in Deutschland 	<ul style="list-style-type: none"> › Anzahl Beschäftigte

Beschäftigte: Tätige Personen gemäß der Definition der amtlichen Statistik.

¹ Das Projekt „Amtliche Firmendaten für Deutschland“ (AFiD) der Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder ermöglicht die flexible und individuelle Zusammenführung von Einzeldaten der Wirtschafts- und Umweltstatistiken.

Aufkommens- und Verwendungstabellen für 2021 mit den Ergebnissen der Material- und Wareneingangserhebung 2022 berechnet. Dieses Vorgehen beruht auf der Annahme, dass sich die Vorleistungsstrukturen innerhalb eines Jahres nicht bedeutsam verändern. Insgesamt 18 076 Unternehmen erfasst die Material- und Wareneingangserhebung 2022.

Zusätzlich zur Produktionserhebung und zur Material- und Wareneingangserhebung werden für die erweiterten Aufkommens- und Verwendungstabellen weitere Datenquellen benötigt, um die Unternehmen nach Größe, Exportintensität und ausländischer beziehungsweise inländischer Kontrolle zu kategorisieren. ➤ Übersicht 2 gibt einen Überblick über die verwendeten Datenquellen, ihren Inhalt und die Untergliederungstiefe.⁵

⁵ In diesem Bericht wird ausschließlich der Begriff Unternehmen verwendet, der aktuell in den Volkswirtschaftlichen Gesamtrechnungen mit der Rechtlichen Einheit gleichzusetzen ist (Stand: 2025).

4

Methodik

4.1 Allgemeiner Ansatz: Mikrodatenverknüpfung

Abhängig von der jeweiligen Datenverfügbarkeit empfiehlt die OECD-Expertengruppe verschiedene Ansätze für die Erstellung von erweiterten Aufkommens- und Verwendungstabellen. Eine wesentliche Berechnungsgrundlage der erweiterten Aufkommens- und Verwendungstabellen in den deutschen Volkswirtschaftlichen Gesamtrechnungen bildet die Verknüpfung von Mikrodaten aus verschiedenen Datenquellen. Diese Mikrodaten werden durch eine eindeutige Unternehmens-ID verknüpft und damit Unternehmensmerkmale aus verschiedenen Statistiken miteinander kombiniert.

Aufgrund der in den Volkswirtschaftlichen Gesamtrechnungen notwendigen konzeptionellen Anpassungen können die Produktionswerte in der Aufkommenstabelle von den Werten der Produktionserhebung abweichen. Um Kohärenz zwischen den erweiterten Aufkommens- und Verwendungstabellen und den Ergebnissen der Volkswirtschaftlichen Gesamtrechnungen sicherzustellen, müssen die aggregierten Werte der erweiterten Aufkommenstabelle mit denen der ursprünglichen Aufkommenstabelle übereinstimmen. Dafür wurden zunächst prozentuale Produktionsstrukturen nach Gütern für die einzelnen Unternehmenskategorien anhand der Produktionserhebung bestimmt und anschließend auf die Werte der regulären Aufkommenstabelle übertragen.

Für die Erstellung der erweiterten Verwendungstabelle wurde ein ähnliches Vorgehen angewendet. Die Werte aus der Material- und Wareneingangserhebung stimmen unter anderem aufgrund von konzeptionellen Anpassungen mit denen der Verwendungstabelle nicht überein. Daher wurden hier, analog zur Produktionsstruktur, die prozentualen Vorleistungsstrukturen aus der Material- und Wareneingangserhebung auf die Werte der Verwendungstabelle übertragen.

4.2 Vorgehen bei kleinen und mittleren sowie großen Unternehmen

Wie die Literaturbeispiele im Kapitel 2 zeigen, ist die Aufschlüsselung der erweiterten Aufkommens- und Verwendungstabellen nach Unternehmensgröße von großem Interesse, da sich kleine und mittlere sowie große Unternehmen beispielsweise häufig in ihrer Integration in den internationalen Handel sowie in ihrem Beitrag zur Bruttowertschöpfung unterscheiden.

Übereinstimmend mit der internationalen Praxis werden Unternehmen mit weniger als 250 Beschäftigten als kleine und mittlere Unternehmen (KMU) und Unternehmen mit 250 oder mehr Beschäftigten als große Unternehmen (GU) klassifiziert. Da in der vierteljährlichen Produktionserhebung keine Beschäftigten erfasst werden, müssen die Unternehmen der Produktionserhebung mit dem statistischen Unternehmensregister verknüpft werden. Durch diese Verknüpfung war es möglich, etwa 98 % der Unternehmen aus der Produktionserhebung einer Beschäftigtengrößenklasse zuzuordnen.

Um die Aufkommenstabelle nach Unternehmenskategorien zu erweitern, wurde erst die Produktion von Gütern je Unternehmensmerkmal und Wirtschaftsbereich berechnet. Dafür wurden zunächst die Unternehmen in den einzelnen Wirtschaftszweigen entweder der Kategorie GU oder der Kategorie KMU zugeordnet. Daraufhin wurde gemäß Formel (1) die Produktion jedes Gutes auf einen KMU- und einen GU-Anteil aufgeteilt, wofür die Produktionswerte aus der Produktionserhebung genutzt wurden. Diese Anteile wurden dann mit den Produktionswerten aus der ursprünglichen Aufkommenstabelle (AT) multipliziert, um den von der jeweiligen Unternehmenskategorie erzeugten Produktionswert zu erhalten – der Rechengang ist Formel (2) zu entnehmen. In den folgenden Formeln steht i für die einzelnen Güter, j für die Wirtschaftsbereiche und k für den Unternehmenstyp (beispielsweise KMU oder GU).

$$(1) \text{ Anteil_Unternehmenstyp}_{ijk} = \frac{\text{Produktion}_{ijk}}{\text{Produktion}_{ij}}$$

$$(2) \text{ Wert_eAT}_{ijk} = \text{Anteil_Unternehmenstyp}_{ijk} \times \text{Wert_AT}_{ij}$$

Für die Berechnung der erweiterten Verwendungstabelle (eVT) wurde das gleiche Vorgehen unter Berücksichtigung der Unternehmen aus der Material- und Wareneingangserhebung als Grundgesamtheit angewendet. Das statistische Unternehmensregister ist bei der Nutzung der Material- und Wareneingangserhebung nicht erforderlich, da Informationen über die Beschäftigtenzahl und den Wirtschaftszweig bereits in der Erhebung enthalten sind. Hier wurde errechnet, welcher Anteil der Vorleistungen aus der Material- und Wareneingangserhebung jeweils von kleinen und mittleren Unternehmen und großen Unternehmen verwendet wird – siehe Formel (3). Die entsprechenden Anteile wurden dann wie in Formel (4) mit den Werten der ursprünglichen Verwendungstabelle (VT) multipliziert, um den Vorleistungseinsatz nach Unternehmenskategorie zu erhalten.

$$(3) \text{ Anteil_Unternehmenstyp}_{ijk} = \frac{\text{Vorleistung}_{ijk}}{\text{Vorleistung}_{ij}}$$

$$(4) \text{ Wert_eVT}_{ijk} = \text{Anteil_Unternehmenstyp}_{ijk} \times \text{Wert_VT}_{ij}$$

4.3 Vorgehen bei Exporteuren beziehungsweise Nicht-Exporteuren

Um eine mögliche Exportabhängigkeit bestimmter Unternehmenstypen zu identifizieren, wurden die Aufkommens- und Verwendungstabellen ebenfalls nach den Kategorien Exporteur/Nicht-Exporteur aufgeschlüsselt. Mithilfe des in Übersicht 2 kurz beschriebenen AfID-Datenpanels Außenhandelsstatistik können hierfür die Exporte mit den in der Produktionserhebung erfassten Produktionswerten von Unternehmen verknüpft werden. International besteht Konsens, Exporteure und Nicht-Exporteure nicht allein danach zu unterscheiden, ob ein Unternehmen überhaupt exportiert oder nicht. Ab wann ein Unternehmen als Exporteur für die Zwecke der erweiterten Aufkommens- und Verwendungstabellen definiert wird, kann je nach Land variieren. Da Unternehmen in Deutschland im Vergleich zu anderen Ländern überdurchschnittlich viel exportieren, wurde ein entsprechend hoher Grenzwert für Exporteure festgelegt:⁶

Nicht-Exporteur, wenn:

$$(5) \frac{\text{Exportwert}}{\text{Umsatz}} < 25\%$$

Exporteur, wenn:

$$(6) \frac{\text{Exportwert}}{\text{Umsatz}} \geq 25\%$$

Bei der Verknüpfung der Datensätze sind Meldepflichten für den Außenhandel und die spezifischen Eigenschaften des Außenhandels-Mikrodaten-Sets zu berücksichtigen. Im Jahr 2021 galt eine allgemeine Meldepflicht für den gesamten Warenverkehr im Extra-EU-Handel, während für den Intra-EU-Handel für Importe eine Meldeschwelle von 800 000 Euro und für Exporte von 500 000 Euro bestand. Für Unternehmen unterhalb dieser Meldeschwelle wurden die Exporte und Importe von der Außenhandelsstatistik geschätzt. So konnte die Mikrodatenverknüpfung mit der Produktionserhebung (für die erweiterten Aufkommenstabellen), mit der Material- und Wareneingangserhebung (für die erweiterten Verwendungstabellen) und mit dem statistischen Unternehmensregister problemlos durchgeführt werden. Die Anteile von Exporteuren und Nicht-Exporteuren an Produktion und Vorleistung je Wirtschaftszweig wurden analog zum Vorgehen der Erweiterung nach Unter-

nehmensgröße berechnet. Mithilfe dieser Anteile werden die Aufkommens- und Verwendungstabellen nach Exporteuren und Nicht-Exporteuren aufgeteilt.

4.4 Vorgehen bei inländisch beziehungsweise ausländisch kontrollierten Unternehmen

Für die Erweiterung nach inländischer beziehungsweise ausländischer Kontrolle werden die Unternehmen in der Statistik über auslandskontrollierte Unternehmen in Deutschland (der sogenannten InwardFATS oder kurz I-FATS)⁷ mit den entsprechenden Unternehmen aus der Produktionserhebung oder der Material- und Wareneingangserhebung verknüpft. Bei der Erstellung der I-FATS werden alle Unternehmen unter ausländischer Kontrolle, die in den Strukturstatistiken in Deutschland erfasst sind, berücksichtigt. Somit war es bei der Berechnung der erweiterten Aufkommens- und Verwendungstabellen nicht erforderlich, Kriterien für inländische oder ausländische Unternehmen gesondert festzulegen. Unternehmen, die beispielsweise in der Produktionserhebung und in I-FATS enthalten sind, werden als ausländisch kontrolliert kategorisiert. Unternehmen, die in der Produktionserhebung, allerdings nicht in I-FATS enthalten sind, werden als inländisch kontrolliert eingestuft. Analog werden die Unternehmen der Material- und Wareneingangserhebung für die erweiterten Verwendungstabellen den Kategorien ausländisch beziehungsweise inländisch kontrolliert zugeordnet.

Nach der Kategorisierung der Unternehmen in der Produktionserhebung sowie der Material- und Wareneingangserhebung wurden die Anteile der Unternehmenskategorie an den Gütern je Wirtschaftszweig ermittelt und auf die Aufkommens- beziehungsweise die Verwendungstabelle übertragen.

⁶ Grenzwerte in anderen Ländern betragen beispielsweise 25 % in Belgien, 10 % in Österreich und 5 % in Dänemark.

⁷ FATS = Foreign Affiliates Statistics – Statistik über die Struktur und Tätigkeit von Auslandsunternehmenseinheiten.

5

Erste Analyseergebnisse

Anhand der oben beschriebenen Methodik wurden die Aufkommens- und Verwendungstabellen für das Berichtsjahr 2021 gemäß der Kriterien Größe, Exporteurstatus und ausländische beziehungsweise inländische Kontrolle untergliedert. Mit diesen Ergebnissen wurden erste Analysen für das Jahr 2021 durchgeführt.

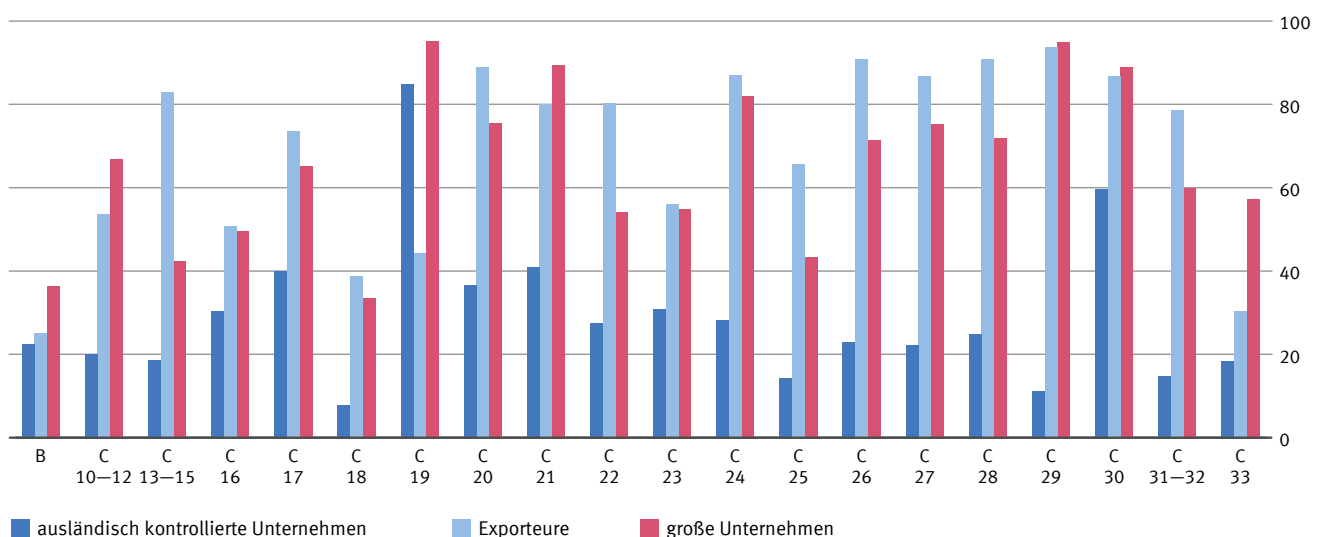
5.1 Struktur der Wirtschaftsbereiche nach Unternehmenskategorien

➤ **Grafik 1** bietet einen Überblick über die Wirtschaftsbereiche des Verarbeitenden Gewerbes nach Unternehmenskategorien. Für jeden Wirtschaftsbereich wurden Anteile an der Produktion von großen Unternehmen, Exporteuren sowie ausländisch kontrollierten Unternehmen dargestellt. Die Differenz der Anteile dieser Kategorien zu 100% ergibt die entsprechenden Anteile für kleine sowie mittlere Unternehmen, Nicht-Exporteure und inländisch kontrollierte Unternehmen.

Mittels der erweiterten Aufkommenstabelle kann die Struktur der Produktion der Wirtschaftsbereiche tiefer analysiert werden. Auffallend ist der hohe Anteil an der Produktion, der von Exporteuren stammt. Dieser ist am höchsten in den Wirtschaftsbereichen Chemie (C20), Pharmazie (C21), Datenverarbeitungsgeräte (C26) und Kraftfahrzeuge (C29). Im Kraftfahrzeugbau (C29) stammen etwa 90% der Produktion von großen, exportorientierten und inländisch kontrollierten Unternehmen, was die Exportabhängigkeit inländischer Kfz-Hersteller und deren Zulieferer abbildet. Auch wird deutlich, dass große Unternehmen oft einen hohen Exportanteil aufweisen, während kleine und mittlere Unternehmen eher auf den Inlandsmarkt fokussiert sind. In den Wirtschaftsbereichen Kokerei und Mineralölverarbeitung (C19) und Sonstiger Fahrzeugbau (C30) haben ausländisch kontrollierte Unternehmen den höchsten Anteil an der Produktion. Interessanterweise produzieren in der Kokerei und Mineralölverarbeitung (C19) große, ausländisch kontrollierte Unternehmen vor allem für den inländischen Markt. Ein Grund dafür könnte sein, dass wenige große internationale Unternehmensgruppen die globale Mineralölverarbeitung dominieren und notwendige Rohstoffe importiert werden müssen.

Grafik 1

Erweiterte Aufkommenstabelle nach Wirtschaftsbereichen 2021
Anteile an der Produktion der Unternehmenskategorien in %



Wirtschaftsbereiche der NACE Rev. 2.

5.2 Vorleistungsstrukturen

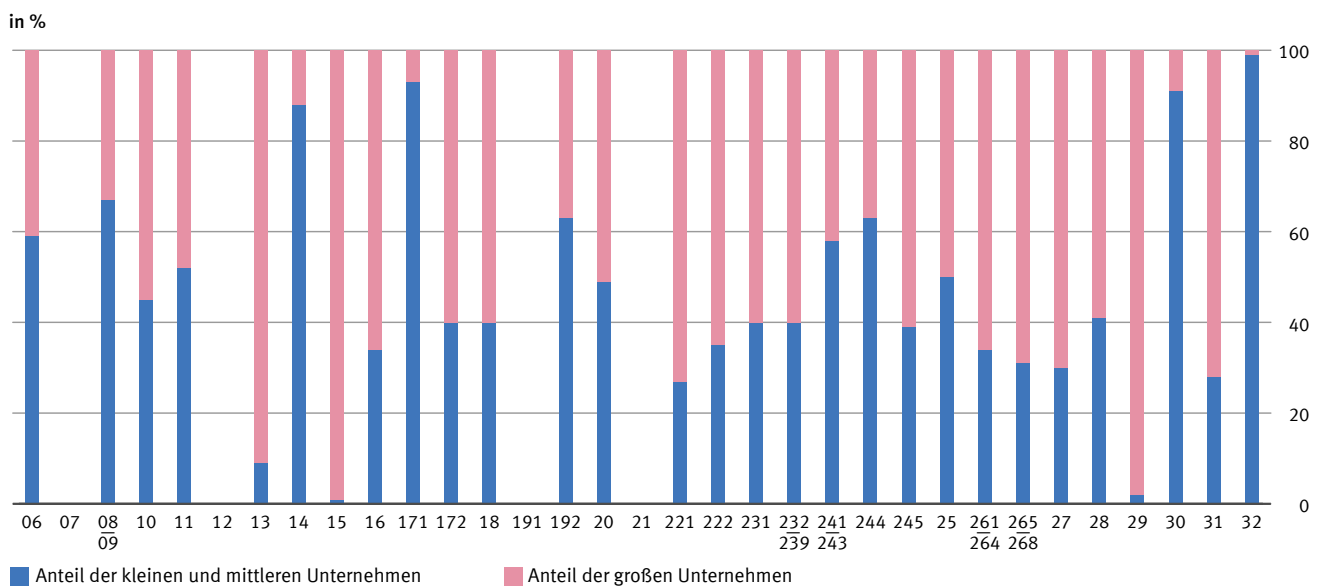
Anhand der erweiterten Aufkommens- und Verwendungstabellen können Unterschiede in Vorleistungsstrukturen nach Unternehmenskategorien untersucht werden. Diese Informationen sind hilfreich, um beispielsweise die Effekte von Schocks auf Lieferketten differenziert nach Unternehmenskategorien zu analysieren. [↗ Grafik 2](#) zeigt, wie sich die Vorleistungsgüter in den Gütergruppen des Wirtschaftsbereichs Herstellung von Metallerzeugnissen (C25) zwischen großen Unternehmen sowie kleinen und mittleren Unternehmen verteilen. Wie hoch der Gesamtverbrauch der jeweiligen Güter in Milliarden Euro ist, ist [↗ Tabelle 1](#) zu entnehmen. Obwohl laut der Material- und Wareneingangserhebung kleine und mittlere Unternehmen mit 52 % etwa gleich viele Vorleistungen verbrauchen wie große Unternehmen (48 %), unterscheiden sich die Vorleistungsstrukturen der Unternehmenskategorien deutlich. Die am Gesamtverbrauch des Wirtschaftszweigs gemessen bedeutendsten Vorleistungen Güter der Metallerzeugung und -bearbeitung (GP241-243)⁸ und Metallerzeugnisse (GP25) werden von kleinen und mittleren Unternehmen

sowie von großen Unternehmen ähnlich viel verwendet. Bei weniger nachgefragten Vorleistungsgütern schwanken die Anteile jedoch beachtlich. Beispielsweise werden 70 % der elektrischen Ausrüstung (GP27) von großen Unternehmen verwendet, obwohl sie nur 48 % der Gesamtnachfrage ausmachen. Dies ist ein Hinweis auf eine erhöhte Abhängigkeit der großen Unternehmen von diesem Gut.

⁸ Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019.

Grafik 2

Verteilung der Vorleistungsgüter in den Gütergruppen des Wirtschaftsbereichs "Herstellung von Metallerzeugnissen" 2021



Gütergruppen des Güterverzeichnisses für Produktionsstatistiken, Ausgabe 2019.

Erweiterte Aufkommens- und Verwendungstabellen – eine Darstellung der Aktivitäten unterschiedlicher Unternehmenskategorien

Tabelle 1

Gesamtverbrauch der Vorleistungsgüter nach Gütergruppen des Wirtschaftsbereichs „Herstellung von Metall-erzeugnissen“ 2021

Gütergruppen des Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019	Vorleistungsgüter
	Mrd. EUR
06	0,406
07	0
08–09	0,003
10	0,011
11	0,007
12	0
13	0,173
14	0,01
15	0,111
16	0,334
171	0,008
172	0,657
18	0,002
191	0
192	0,412
20	2,224
21	0
221	0,21
222	1,785
231	0,71
232–239	0,226
241–243	17,917
244	6,051
245	1,205
25	24,477
26	0,247
27	1,357
28	1,494
29	0,067
30	0,01
31	0,062
32	0,134

Werte aus den Aufkommens- und Verwendungstabellen.

Tabelle 2

Gesamtproduktion nach Gütergruppen 2021

Gütergruppen des Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019	Gesamtproduktion
	Mrd. EUR
052	1 022
06	1 129
07	0
08–09	8 717
10	166 566
11	20 097
12	2 343
13	12 529
14	4 671
15	2 657
16	30 083
171	16 913
172	22 158
18	15 615
191	1 000
192	57 963
20	144 839
21	42 210
221	13 300
222	69 407
231	10 717
232–239	36 057
241–243	51 870
244	46 045
245	12 423
25	139 790
26	84 246
27	103 471
28	233 850
29	348 700
30	43 905
31	22 137
32	34 123
33	58 585

Werte aus den Aufkommens- und Verwendungstabellen.

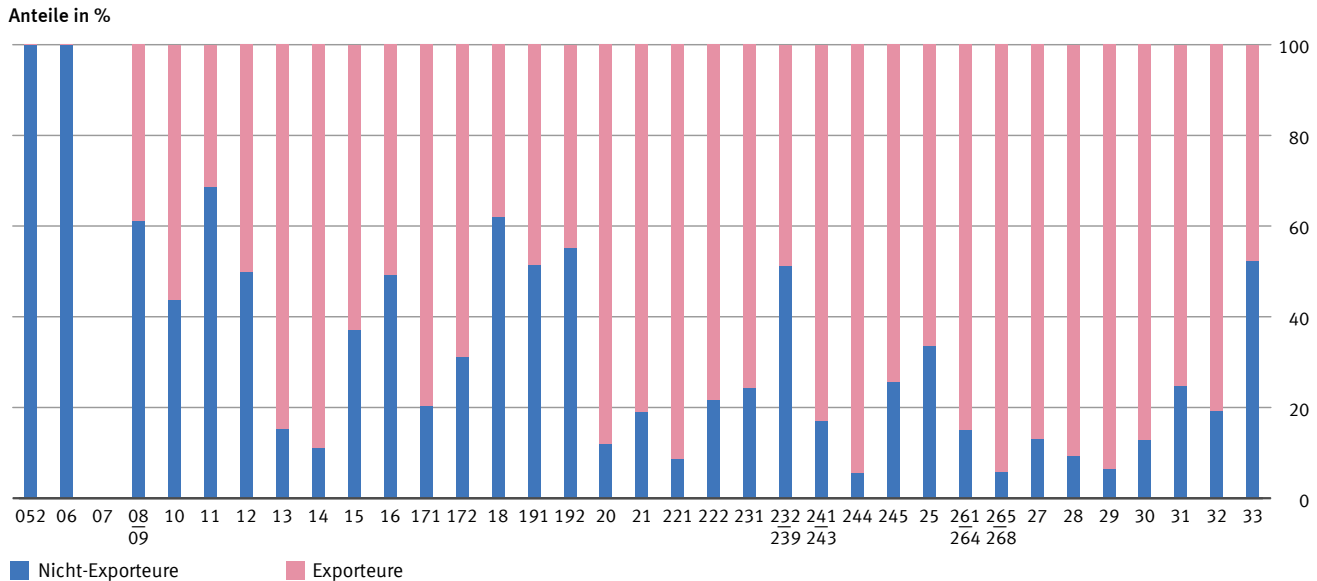
5.3 Produktionsstrukturen

Für weiterführende Analysen sind auch die Produktionsstrukturen der verschiedenen Unternehmenskategorien von Bedeutung. Sie ermöglichen es genau zu bestimmen, welche Unternehmenskategorien welche Güter produzieren. Exemplarisch wurde in [Grafik 3](#) die Gesamtproduktion der verschiedenen Gütergruppen nach Exporteuren und Nicht-Exporteuren dargestellt. Insgesamt ist zu erkennen, dass für die meisten

untersuchten Gütergruppen der Anteil der Exporteure viel höher ist als der Anteil der Nicht-Exporteure, was die Exportabhängigkeit der deutschen Wirtschaft widerspiegelt. [Tabelle 2](#) bestätigt die hohe Relevanz der Produktion der Gütergruppen Maschinen (GP28) und Kraftwagen und Kraftwagenteile (GP29) für die deutsche Wirtschaft. Hier ist in Verbindung mit [Grafik 3](#) deutlich zu erkennen, dass ein Großteil dieser Produktion stark exportabhängig ist.

Grafik 3

Gesamtproduktion in den Gütergruppen sowie nach Exporteuren und Nicht-Exporteuren 2021



Gütergruppen des Güterverzeichnis für Produktionsstatistiken, Ausgabe 2019.

5.4 Bruttowertschöpfung

Detailliertere Informationen zu Produktion und Vorleistungen ermöglichen präzisere Berechnungen der Bruttowertschöpfung. Für die Berechnung der Bruttowertschöpfung nach Unternehmenskategorie wurden die Produktionswerte und Vorleistungen aus den erweiterten Aufkommens- und Verwendungstabellen für ausländisch und inländisch kontrollierte Unternehmen herangezogen. Die Bruttowertschöpfung der jeweiligen Wirtschaftsbereiche ergibt sich, indem die Vorleistungen von den Produktionswerten subtrahiert werden.

(7) *Bruttowertschöpfung = Produktionswert (Herstellungspreise) – Vorleistungen (Anschaffungspreise)*

Aus Grafik 1 lässt sich ableiten, dass der Wirtschaftsbereich Herstellung von Kraftwagen und Kraftwagenteilen (C29) stark von inländischen Unternehmen abhängig ist. Berechnet man die Bruttowertschöpfung dieses Bereichs anhand von Formel (7), so zeigt sich, dass rund 92 % der Bruttowertschöpfung auf heimische Unternehmen entfallen, während etwa 8 % auf ausländisch kontrollierte Unternehmen zurückzuführen sind. Damit ist der Anteil ausländisch kontrollierter Unternehmen an der Bruttowertschöpfung geringer als ihr Anteil an der Produktion (11 %).


6

Fazit

Erweiterte Aufkommens- und Verwendungstabellen sind ein wertvolles Instrument, um wirtschaftliche Strukturen und Zusammenhänge zu untersuchen. Die nach den Unternehmensmerkmalen Größe, Exportintensität und ausländische beziehungsweise inländische Kontrolle erweiterten Aufkommens- und Verwendungstabellen ermöglichen tiefgehende Einblicke in die Vorleistungs- und Produktionsstrukturen dieser Unternehmenskategorien. Solche Erkenntnisse werden besonders wichtig, wenn globale Unsicherheiten zunehmen und Veränderungen eintreten, etwa durch die wirtschaftlichen Folgen der COVID-19-Pandemie oder geopolitische Konflikte. Dies gilt insbesondere für die deutsche Wirtschaft, die stark in weltweite Lieferketten integriert ist.

Datenlücken limitieren allerdings die Anwendungsmöglichkeiten der erweiterten Aufkommens- und Verwendungstabellen. Derzeit können erweiterte Aufkommens- und Verwendungstabellen aufgrund fehlender Mikrodaten nur im Bereich der Waren (GP 05-33) und der Unternehmen des Bergbaus und des Verarbeitenden Gewerbes erstellt werden. Da sich außerdem die

Grundgesamtheiten der Produktionserhebung und der Material- und Wareneingangserhebung hinsichtlich der enthaltenen Unternehmen unterscheiden, können auch direkte Vergleiche der verschiedenen Kategorien zwischen der erweiterten Aufkommenstabelle und der erweiterten Verwendungstabelle aktuell nur grobe Einschätzungen liefern. Außerdem ist für eine Zeitreihenanalyse anhand von erweiterten Aufkommens- und Verwendungstabellen die vierjährige Periodizität der Material- und Wareneingangserhebung nachteilig, da strukturelle Veränderungen über mehrere Berichtsjahre hinweg nur eingeschränkt abgebildet werden können. Besonders in Jahren mit starken wirtschaftsstrukturellen Schocks kann das die Aussagekraft von erweiterten Aufkommens- und Verwendungstabellen verringern.

Wie die aktuellen Auswertungen zeigen, sind die Ergebnisse der experimentellen erweiterten Aufkommens- und Verwendungstabellen vielversprechend. Um das Potenzial von erweiterten Aufkommens- und Verwendungstabellen voll auszuschöpfen, sollten neue Datenquellen erschlossen werden, beispielsweise im Dienstleistungsbereich. Mit robusteren erweiterten Aufkommens- und Verwendungstabellen könnten anschließend erweiterte Input-Output-Tabellen erzeugt werden, welche die Analyse indirekter Effekte nach Unternehmenskategorien ermöglichen. 

LITERATURVERZEICHNIS

Bernard, Andrew B./Jensen, J. Bradford. *Exporters, Jobs, and Wages in U.S. Manufacturing: 1976–1987*. In: Brookings Papers on Economic Activity. Microeconomics. 1995. Seite 67 ff.

Bernard, Andrew B./Jensen, J. Bradford/Redding, Stephen J./Schott, Peter K. *The Empirics of Firm Heterogeneity and International Trade*. In: NBER (National Bureau of Economic Research) Working Paper Series. 2011. Working Paper 17627.

Bundesministerium für Wirtschaft und Klimaschutz. *Aktionsplan zum Dialog- und Arbeitsprozess Mittelstand, Klimaschutz und Transformation*. 2022. [Zugriff am 18. März 2025]. Verfügbar unter: www.bmwk.de

Chong, Stephen/Hoekstra, Rutger/Lemmers, Oscar/Van Beveren, Ilke/Van den Berg, Marcel/Van der Wal, Ron/Verbiest, Piet. *The Role of Small- and Medium-Sized Enterprises in the Dutch Economy: An Analysis Using an Extended Supply and Use Table*. In: Journal of Economic Structures. 2019. Artikel Nr. 8(1), Seite 1 ff.

Miroudot, Sébastien/Rigo, Davide. *Multinational Production and Investment Provisions in Preferential Trade Agreements*. In: Journal of Economic Geography. 2022. Jahrgang 22, Ausgabe 6, Seite 1275 ff. DOI: [10.1093/jeg/lbab031](https://doi.org/10.1093/jeg/lbab031)

OECD. *OECD Expert Group on Extended Supply and Use Tables: Terms of Reference 2015*. [Zugriff am 25. März 2025]. Verfügbar unter: unstats.un.org

Statistische Ämter des Bundes und der Länder – Forschungsdatenzentren. *Metadaten-report – Teil 1 Allgemeine und methodische Informationen zum AfID-Panel Außenhandelsstatistik (EVAS-Nummer 51911), Berichtsjahre 2011-2019*. 2023. [Zugriff am 18. März 2025]. Verfügbar unter: www.forschungsdatenzentrum.de

Statistisches Bundesamt. *Material- und Wareneingangserhebung im Verarbeitenden Gewerbe sowie im Bergbau und in der Gewinnung von Steinen und Erden*. Qualitätsbericht. 2017. [Zugriff am 18. März 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Produktionserhebungen*. Qualitätsbericht. 2024a. [Zugriff am 18. März 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Statistisches Unternehmensregister*. Qualitätsbericht 2023. 2024b. [Zugriff am 18. März 2025]. Verfügbar unter: www.destatis.de

Statistisches Bundesamt. *Statistik über auslandskontrollierte Unternehmen*. Qualitätsbericht. 2024c. [Zugriff am 18. März 2025]. Verfügbar unter: www.destatis.de

UNCTAD (United Nations Conference on Trade and Development). *World Investment Report 2013: Global Value Chains: Investment and Trade for Development*. [Zugriff am 18. März 2025]. Verfügbar unter: unctad.org

Venohr, Bernd/Fear, Jeffrey/Witt, Alessa. *The Best of German Mittelstand – The World Market Leaders*. In: Langenscheidt, Florian/Venohr, Bernd (Herausgeber). *The Best of German Mittelstand*. Köln 2015.

DER DATA GOVERNANCE ACT

Auswirkungen auf die vertrauenswürdige Nutzung
von Daten öffentlicher Einrichtungen:
Rolle eines Datentreuhänders

Judith Anna Czepek, Nicholas-Sebastian Schneider

➤ **Schlüsselwörter:** Data Governance Act (DGA) – Datentreuhänder –
Verwaltungsdaten – Datenökosystem – Datentransparenz

ZUSAMMENFASSUNG

Der Data Governance Act umfasst unter anderem Anfragen mit Interesse an geschützten Daten in öffentlichen Einrichtungen, die Beratung und IT-Serviceangebote für öffentliche Einrichtungen sowie die Pflege einer Metadatenbank. Im Kabinettsbeschluss des nationalen Durchführungsgesetzes ist das Statistische Bundesamt als zuständige Stelle und zentrale Informationsstelle für die Vermittlung von Daten in öffentlicher Hand benannt. Mit einer entsprechenden Aufgabenübertragung würde das Statistische Bundesamt verstärkt neue Zielgruppen wie Wirtschaftsunternehmen, die Verwaltungsdaten für neue Geschäftsmodelle entwickeln wollen, bedienen und sich in beratender und dienstleistender Funktion als erste Anlaufstelle für andere datenhaltende öffentliche Einrichtungen positionieren.

➤ **Keywords:** Data Governance Act (DGA) – data trustee – administrative data – data ecosystem – data transparency

ABSTRACT

The Data Governance Act (DGA) covers enquiries from parties interested in protected data held by public sector bodies, consulting and IT services for public sector bodies, and the management of a meta-database, for example. In the cabinet resolution on the act transposing the DGA provisions into national law, the Federal Statistical Office of Germany is designated as the competent body and single information point for the intermediation of data from the public sector. With this delegation of duties, the Federal Statistical Office would increasingly serve new target groups, such as enterprises wishing to develop administrative data for new business models, and position itself as the go-to point for other data holding public institutions, offering expert advice and services.



Dr. Judith Anna Czepek

ist Soziologin und leitet seit 2020 das Referat „Programm der Bundesstatistik“, welches im Statistischen Bundesamt für die Strategie und Grundsatzfragen zuständig ist. Sie war von 2022 bis 2024 die Projektleiterin zur Umsetzung des Data Governance Acts unter Federführung des Bundesministeriums für Wirtschaft und Klimaschutz.



Nicholas-Sebastian Schneider

ist Politologe und im Referat „Programm der Bundesstatistik“ als wissenschaftlicher Mitarbeiter tätig. Er ist für eine Modernisierung des Rechtsrahmens und in diesem Zusammenhang von Beginn an mit für die erfolgreiche Planung der Umsetzung des Data Governance Acts sowie das Stakeholdermanagement mit den Bundesministerien und mit internationalen Partnern verantwortlich.

1

Einleitung

Der mit der Verordnung 2022/868 in Kraft getretene europäische Daten-Governance-Rechtsakt (Data Governance Act – DGA) gilt ab dem 24. September 2023 und ist unmittelbar geltendes Recht in den Mitgliedstaaten der Europäischen Union (EU). Um die Verpflichtungen aus dem Daten-Governance-Rechtsakt vollständig und bundeseinheitlich zu erfüllen und umzusetzen, sind zusätzliche gesetzliche Durchführungsbestimmungen erforderlich. Diese soll das Daten-Governance-Gesetz (DGG) festlegen, in dem das Statistische Bundesamt als zuständige Stelle für die Unterstützung der öffentlichen Stellen, die geschützte Verwaltungsdaten nutzbar machen wollen, und außerdem als zentrale Informationsstelle vorgesehen ist.

Kapitel 2 beschreibt den Daten-Governance-Rechtsakt als Teil der europäischen Datenstrategie, dessen Zielsetzung und welche Aufgaben das Statistische Bundesamt künftig bei dessen Umsetzung übernehmen könnte. Auf die strategische Bedeutung und die mögliche neue Rolle eines zentralen Datentreuhänders geht Kapitel 3 ausführlich ein. Dazu werden die neuen Aufgaben und die Erschließung neuer Zielgruppen und die damit einhergehenden Mehrwerte dargestellt. Ein Ausblick auf den internationalen Austausch für einen europäischen Datenbinnenmarkt rundet den Beitrag ab.

2

Der Data Governance Act als Teil der europäischen Datenstrategie

2.1 Zielsetzung des Data Governance Acts sowie des Entwurfs eines nationalen Daten-Governance-Gesetzes

Daten gelten als wichtiger Rohstoff einer digitalisierten Welt und die Europäische Union befindet sich in einer globalen Auseinandersetzung um deren Verwertung. Dieser Auffassung pflichtete die ehemalige EU-Vizeprä-

sidentin Margrethe Vestager in einem Interview 2021 wie folgt bei:

“As you say, this is a global matter because we are also in a geopolitical context where we have both systemic rivalry, economic competition, but also the need for partnerships on this planet. So how we deal with data, how we enable it to flow once we trust it, but how we also retain it if it’s our own personal data, well, that is what we have ahead of us both for practical implementation, but also to get the legislation right.” (Vestager, 2021)

Mit der im Jahr 2020 verabschiedeten Datenstrategie nahm die EU die Schaffung von Voraussetzungen für datengetriebene Innovationen unter hohen datenschutzrechtlichen sowie sicherheitstechnischen Standards in den Fokus. Die Steigerung von Vertrauen in die Datennutzung soll mehr Wettbewerb und Transparenz schaffen und die europäische Datensouveränität gewährleisten. Ein wesentlicher Teil der strategischen Umsetzung ist, rechtliche Rahmenbedingungen zu schaffen und anzupassen, welche es Wirtschaft, Politik und Gesellschaft ermöglichen sollen, von den stetig wachsenden Datenmengen zu profitieren. Die dadurch angestrebte Erweiterung des bestehenden europäischen Binnenmarkts um einen einheitlichen europäischen Datenraum, einen echten Binnenmarkt für Daten, ist das formulierte Ziel (Europäische Kommission, 2020, hier: Seite 5).

Ein bedeutender Teil des rechtlichen Rahmenwerks und zudem das erste verabschiedete Gesetz zur Umsetzung der europäischen Datenstrategie ist der Data Governance Act¹, welcher seit dem 24. September 2023 in den EU-Mitgliedstaaten anzuwenden ist. *“The Data Governance Act is an important milestone to boost a data-driven economy in Europe, and the first deliverable following our data strategy (...)”* (Vestager, 2020). Die Verordnung zielt durch einheitliche Vorschriften und grundlegende Standards darauf ab, die Datenverfügbarkeit in einem vertrauenswürdigen und sicheren Umfeld zu erhöhen und damit bisher ungenutzte Datenpotenziale zu heben (siehe Erwägungsgrund 3 zur Verordnung [EU] 2022/868).

Um die Weiterverwendung von Daten durch mehr Rechtssicherheit zu fördern und um Lücken in der bisherigen Rechtssetzung der EU zu schließen, umfasst der

1 Geregelt in der Verordnung (EU) 2022/868.

Data Governance Act im Wesentlichen drei Regelungsbereiche. Dies sind

- (1) die Bedingungen für die Weiterverwendung bestimmter Kategorien geschützter Daten des öffentlichen Sektors in Kapitel II,
- (2) Regelungen für Datenvermittlungsdienste in Kapitel III sowie
- (3) einheitliche Vorgaben für Einrichtungen für daten-altruistische Zwecke in Kapitel IV.

Nachstehend werden die neuen Aufgaben für die Mitgliedstaaten aus Kapitel II betrachtet, da hier die Rahmenbedingungen für die Weiterverwendung von geschützten Daten geregelt werden und beachtliches Potenzial für eine transparente, wettbewerbsfähige Datenlandschaft liegt. Der Data Governance Act zielt darauf ab, eine identifizierte Lücke der Datenverfügbarkeit zu schließen: Die Möglichkeiten, besonders geschützte Verwaltungsdaten weiterzuverwenden, sollen verbessert werden. Insofern werden die bereits bestehenden Open-Data-Regelungen um Datenkategorien mit berechtigten Schutzinteressen von Dritten erweitert, um deren bisher ungenügende Nutzung zu fördern (Specht/Hennemann, 2023, hier: Seite 172 f.). Der Anspruch ist, dass Wirtschaft und Gesellschaft von Daten öffentlicher Einrichtungen noch stärker profitieren sollen, da diese Daten mit öffentlichen Geldern generiert werden (siehe Erwägungsgrund 6 zur Verordnung [EU] 2022/868). Geregelt werden die Weiterverwendung von Geschäftsgeheimnissen (Artikel 3 Absatz 1 Buchstabe a des Daten-Governance-Rechtsakts), von Daten, welche der statistischen Geheimhaltung unterliegen (Artikel 3 Absatz 1 Buchstabe b), von geistigem Eigentum Dritter (Artikel 3 Absatz 1 Buchstabe c) sowie die Weiterverwendung von personenbezogenen Daten (Artikel 3 Absatz 1 Buchstabe d).

Kapitel II des Data Governance Acts befasst sich dabei mit dem „Wie“ und nicht mit dem „Ob“ der Verfügbarmachung beziehungsweise der Weiterverwendung von besonders geschützten Verwaltungsdaten. Das bedeutet, dass der Data Governance Act keine grundsätzliche Pflicht zum Teilen der Daten begründet, sondern der nationale Gesetzgeber den Umfang der durch den Data Governance Act adressierten geschützten Daten selbst regelt (siehe etwa Paschke/Rücker, 2023, Seite 107; Vestager, 2020). Daher bestimmen datenpolitische Vorhaben und die dazugehörige Gesetzgebung der nächs-

ten Jahre maßgeblich den Umfang der für Dateninteressierte zur Verfügung stehenden Daten.

Der Entwurf des Daten-Governance-Gesetzes sollte 2024 die nationale Umsetzung regeln, indem die in Deutschland zuständigen Behörden festgelegt wurden. Um deren Durchsetzungsfähigkeit zu stärken, sollte der Entwurf die Möglichkeit schaffen, künftig aufkommende Regelungsbedarfe, etwa technischer und organisatorischer Standards, durch Rechtsverordnungen zu regeln. Der Entwurf klärt damit vordergründig die durch die EU-Verordnung erforderlichen Zuständigkeiten, bleibt bei der weiteren Ausgestaltung bezüglich der Regelungsmöglichkeiten, die Datennutzung zu erhöhen, zurückhaltend und unbestimmt.

2.2 Aufgaben und relevante Stakeholder

Der Data Governance Act sieht für die Aufgaben in den Kapiteln II, III und IV die Benennung zuständiger nationaler Behörden vor. Dabei war im Entwurf des nationalen Durchführungsgesetzes das Statistische Bundesamt für die Übernahme der neuen Aufgaben aus Kapitel II und die Bundesnetzagentur für Elektrizität, Gas, Telekommunikation, Post und Eisenbahnen (Bundesnetzagentur) für die Übernahme der Aufgaben aus den Kapiteln III und IV Data Governance Act vorgesehen. Das Statistische Bundesamt würde damit die folgenden Aufgaben übernehmen:

- (1) Zuständige Stelle für die Unterstützung der öffentlichen Stellen, die Zugang zur Weiterverwendung von Daten gewähren oder verweigern (Artikel 7 Absatz 1 der Verordnung [EU] 2022/868);
- (2) Zentrale Informationsstelle für die Gewährleistung einschlägiger Informationen und für die Entgegennahme von Anfragen oder Anträgen in Bezug auf die Weiterverwendung von Daten (Artikel 8 Absatz 1 und Absatz 2 der Verordnung [EU] 2022/868).

Mit der Umsetzung einer zuständigen Stelle und zentralen Informationsstelle entstünde eine Servicestelle als erste Anlaufstelle für Datenanfragen sowie für die Unterstützung öffentlicher Einrichtungen zur Weiterverwendung von geschützten Daten. Dazu gehört, die öffentlichen Stellen in verschiedenen (daten- und IT-) technischen und rechtlichen Fragen zu unterstützen sowie eine IT-Infrastruktur mit einer sicheren IT-Verar-

beutungsumgebung bereitzustellen. Eine entsprechende Servicestelle wird die Verfügbarkeit der angefragten Daten prüfen und, sofern verfügbar, die datenhaltende Institution benennen. Sie wird an die Anfragenden und die datenhaltende Organisation melden und über anfallende Gebühren und Nutzungsbedingungen informieren. Im Zuge der Auftragsklärung werden bei den datenhaltenden Stellen Metadaten und die Bedingungen der Weiterverwendung von Daten erfragt. Falls der Auftrag erteilt wird, erfolgt (bei Bedarf) zudem die Durchführung und Umsetzung einer verschlüsselten Datenübermittlung aufbereiteter Daten.

Durch den direkten Kontakt mit den datenhaltenden öffentlichen Einrichtungen kann die Sammlung und Prüfung von Metadaten für den Aufbau der sogenannten Bestandsliste, eines entsprechenden Datenkatalogs geschützter Verwaltungsdaten mit einer Schnittstelle zum EU-Zugangsportaal für einen europaweiten Metadatenkatalog mit den Angaben für die Bedingungen ihrer Weiterverwendung, besonders effizient betrieben werden. Beim Aufbau eines Angebots mit datenhaltenden öffentlichen Stellen ergeben sich vielfältige Möglichkeiten für Standardisierungen. Das ist eine wichtige Voraussetzung für einen effizienten Datenaustausch in den Bereichen der Informationssicherheit, der sicheren Datenübermittlung, -aufbereitung und -bereitstellung.

Neben diesen nationalen ist ein vielschichtiges Netz aus internationalen Stakeholdern an der Umsetzung des Data Governance Acts als EU-Verordnung beteiligt. Vordergründig ist dies die EU-Kommission selbst, die zudem durch die Einrichtung eines Europäischen Dateninnovationsrats (EDIB) eine Gruppe von Expertinnen und Experten berufen hat, welche bei der Umsetzung berät und unterstützt (Artikel 29 und 30). Dem Innovationsrat gehören neben EU-seitigen Mitgliedern die nach nationalem Recht zuständigen Behörden für die Datenvermittlungsdienste und für die Registrierung von datenaltruistischen Organisationen der Mitgliedstaaten an. In Deutschland übernimmt die Bundesnetzagentur somit die Vertretung. Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit (BfDI) ist als Vertretung des Europäischen Datenschutzausschusses Teil des Innovationsrats (BfDI, 2023).

Abgestimmt mit der Generaldirektion „Unterstützung von Strukturreformen“ werden im Rahmen des Technical-Support-Instrument-Förderprogramms der EU verant-

wortliche Einrichtungen nach Artikel 7 und 8 europaweit von erarbeiteten Richtlinien, Handbüchern, multinationalen Antragsformularen sowie vom Wissenstransfer profitieren können.

Die Zahl der durch den Data Governance Act betroffenen Stakeholder wird künftig voraussichtlich deutlich zunehmen, da die Umsetzung der Aufgaben verschiedene Akteure unmittelbar betreffen wird. Dies gilt insbesondere in der Zusammenarbeit vieler wichtiger datenhaltender öffentlicher Einrichtungen wie der Deutschen Bundesbank oder der Statistik der Bundesagentur für Arbeit. Der Austausch in projektspezifischen Arbeitsgruppen wie zu High Value Datasets oder Data Spaces wird zentral für die erfolgreiche Umsetzung der neuen Aufgaben sein und über nationale Grenzen hinaus neue Kooperationsmöglichkeiten im bekannten Datenökosystem mit sich bringen.

3

Strategische Bedeutung und neue Rolle

3.1 Neue Rolle als Datentreuhänder: Herausforderungen und Chancen im Datenmanagement

„Die zunehmende Digitalisierung der Geschäftsprozesse ist ein Grund dafür, dass immer mehr private Akteure in den Markt der Datenanbieter eintreten.“ (Vorgrimler/Decker, 2021, hier: Seite 22). Diese Entwicklung einer zunehmenden Datenproduktion der Unternehmen lässt sich auch auf die Verwaltungen übertragen, deren Modernisierung permanent Daten aus Verwaltungsprozessen hervorbringt, teilweise lediglich als Nebenprodukt. Die Menge an Daten steigt daher weniger durch Erhebungen als vielmehr durch digitale Geschäftsprozesse, welche Daten sehr unterschiedlicher Qualität erzeugen. Die Herausforderung, ein verbessertes Datenangebot zu schaffen, besteht demnach weniger in der Menge an Daten als vielmehr in Informationen über deren Verfügbarkeit und Qualität. Aus den skizzierten Entwicklungen lassen sich wiederum Anforderungen an einen transparenten Datenmarkt ableiten, die sich auf die Zugänglichkeit, Auffindbarkeit und Bereitstellung konzentrieren. Dies ist ein deutlicher Wandel von der Produktion von Daten hin

zur Verfügbarmachung von Daten, die bereits in Unternehmen und Verwaltungen vorliegen.

Um diese bisher unzugänglichen, teilweise prozessproduzierten Daten nutzen zu können, braucht es vertrauenswürdige Vermittler, welche diese Daten auffindbar machen und es mit ihren Kompetenzen beim Aufbereiten und Bereitstellen von Daten ermöglichen, diese zu nutzen. Zudem muss die Datenverarbeitung den hohen Schutzanforderungen der Datenschutz-Grundverordnung genügen – insbesondere, wenn es sich um personenbezogene Daten handelt. Darüber hinaus sollten Nutzende die Datenqualität beurteilen können und Bedingungen für die Weiterverwendung definiert werden. Dies verhindert, dass Daten fehlerhaft verbreitet und interpretiert werden und unterstützt somit die Datenkompetenz der Gesellschaft.

Beim Data Governance Act handelt es sich um ein Ermöglichungsgesetz, da es zwar keinen Rechtsanspruch auf Datenzugang formuliert, aber den Rahmen zur EU-weiten Vernetzung des Datenökosystems vorgibt. Die Artikel 7 und 8 schaffen eine zuständige Stelle und eine zentrale Informationsstelle, welche die Rolle eines Datentreuhänders und -vermittlers auf einem europäischen Datenbinnenmarkt übernehmen. Diese Datentreuhänder sollen eine sichere Bereitstellung von Daten öffentlicher Einrichtungen durch eine sichere IT-Verarbeitungsumgebung ermöglichen und durch eine sogenannte Bestandsliste Metadateninformationen öffentlich machen. Dadurch wird die Auffindbarkeit wesentlich verbessert, da bekannt sein wird, wo welche Daten unter welchen Bedingungen bereitgestellt werden können. Zudem sollen Beratungsleistungen für datenhaltende Einrichtungen angeboten werden.

Datentreuhänder

Ein Datentreuhänder vermittelt Daten zwischen Nutzenden und Anbietern und ist damit ein Intermediär, der den Datenaustausch erleichtern soll. Die Einschaltung einer dritten Instanz soll Vertrauen stiften und Neutralität unter höchsten Datenschutzansprüchen garantieren. Zu den Aufgaben eines Datentreuhänders als Vermittler gehören die Datenübertragung, die Identitätssicherung der Beteiligten, die Zugriffsberechtigungen und die Datenverarbeitung. Ein Datentreuhänder schützt zudem die Daten vor Diebstahl und Manipulation.

In der Datenstrategie der Bundesregierung ist die Einrichtung einer solchen Datentransparenzstelle im Statistischen Bundesamt 2021 mit einer Maßnahme zur „Anwendung anerkannter Datenstandards und Methoden“ verankert (Bundesregierung, 2020). Dazu zählen, für Transparenz im Datenökosystem zu sorgen und „eine qualitative Beurteilung der Angebote nach Maßstäben der amtlichen Statistik“ durchzuführen (Vorgriemer/Decker, 2021). Betrachtet man zudem die oben skizzierten Entwicklungen auf dem Datenmarkt und die Anforderungen der EU-Verordnung, so wäre das Statistische Bundesamt

- › als erfahrener Datenproduzent,
- › als unabhängige Bundesbehörde und
- › als kompetenter Bereitsteller von Daten für die Politik, Wissenschaft und Gesellschaft

der natürliche Partner für die Vermittlung besonders geschützter Daten öffentlicher Einrichtungen und die Anwendung entsprechender europäischer Standards in der Qualität, Methode und IT.

Neutralität, Objektivität und fachliche Unabhängigkeit gehören zu den wichtigsten Fundamenten der bundesdeutschen amtlichen Statistik. Das Statistische Bundesamt ist diesen Grundsätzen gesetzlich sowohl national als auch international verpflichtet. Folglich besitzt es die neutrale Stellung und das Vertrauen, welche für einen fairen Datenzugang durch eine unabhängige Institution und für die Entwicklung eines grenzfreien Datenbinnenmarkts notwendig sind. Hinzu kommt die Expertise im Bereich der Anonymisierung von Datenbeständen durch die Bereitstellung von mehr als 90 Statistiken, durch maschinelle Prüfprozesse auf statistische Geheimhaltung und den Betrieb des Forschungsdatenzentrums seit dem Jahr 2001. Abhängig vom Schutzbedarf der Datenbestände sowie unter Wahrung der höchsten Standards in der IT-Sicherheit und im Datenschutz ermöglicht das Forschungsdatenzentrum verschiedene Datenzugangswege. Diese sind bisher auf die Zielgruppe der Wissenschaft beschränkt und konzentriert auf die Statistiken des Bundes und der Länder (Brenzel/Zwick, 2022). Mit der Verwaltungsdaten-Informationsplattform liefert das Statistische Bundesamt bereits eine Übersicht der öffentlichen Datenbestände, bezogen auf Register und Kataster (Bens/Schukraft, 2018). Neben dieser besonderen Expertise würde die Übertragung der Aufgabe dem Statistischen Bundesamt die Chance eröffnen, seine

technische und methodische Kompetenz in allen Bereichen des Datenmanagements weiterzuentwickeln. Dazu zählen die Modernisierung der Datenverarbeitung, eine digitale Datenarchitektur und der Aufbau einer sicheren cloudbasierten IT-Infrastruktur sowie einer sicheren IT-Verarbeitungsumgebung, in welcher Dritte eine sichere Datenverarbeitung durchführen (lassen) können. Aufgrund der bereits bestehenden Expertise ist die Benennung des Statistischen Bundesamtes im Entwurf des Durchführungsgesetzes Daten Governance Gesetz (DGG) als zuständige Stelle und zentrale Informationsstelle folgerichtig.

Dem Statistischen Bundesamt würde sich mit der Aufgabenübernahme nach dem Gesetzentwurf ein neues Tätigkeitsfeld erschließen, in dem es nicht nur vorwiegend Daten eigener Erhebungen der Wissenschaft bereitstellt oder Register führt, sondern als Datentreuhänder die Position eines neutralen und vertrauenswürdigen Vermittlers von Daten Dritter an berechnigte Anfragende einnimmt.

Die Servicestelle positioniert sich im Datenökosystem als Anbieter für andere datenhaltende öffentliche Einrichtungen, die meist über weniger Erfahrung und Expertise verfügen, aber zunehmend Daten in Geschäftsprozessen produzieren. Die Rolle als Datentreuhänder und -vermittler bringt es zudem mit sich, die Verfügbarkeit von Daten Dritter, das heißt von anderen datenhaltenden öffentlichen Einrichtungen, zu prüfen. Aus dieser Rolle heraus ergeben sich strategische Gewinne, da die Relevanz der Datenvermittlung im vernetzten Datenökosystem zunimmt.

Gleichzeitig steigt die allgemeine Datenverfügbarkeit durch eine solche datentreuhänderische Servicestelle und es wird ein wichtiger Beitrag zur Verwaltungsmodernisierung geleistet. Mit der dadurch notwendigen Modernisierung des Datenmanagements werden einheitliche Datenformate zur Prozessstandardisierung für die effiziente Bildung von digitalen Prozessketten etabliert und mit der Interoperabilität von Komponenten für den Datenaustausch gesorgt.

3.2 Neue Zielgruppen erschließen und Mehrwerte schaffen

Die EU hat ein erhebliches Defizit in der bisherigen Nutzung besonders geschützter Verwaltungsdaten ausgemacht. Der neutrale Zugang zu Daten als Wirtschafts- und Innovationstreiber wird im Data Governance Act bewusst in den Fokus gerückt (Europäische Kommission, 2025a).

Wirtschaft

Als zentrale Zielgruppe benennt die EU-Verordnung die Wirtschaft. Dies entspricht den Zielen des Data Governance Acts, Innovationen zu fördern und den europäischen Wettbewerb durch verbesserte Rahmenbedingungen für die Datenverfügbarkeit zu unterstützen. Durch die Möglichkeit, geschützte Datenbestände anzufragen oder in der Metadatenbank interessante Datenbestände zu sichten, erhält die Wirtschaft niedrigschwellige Zugangswege zu Informationen zu geschützten Verwaltungsdaten und zu den Bedingungen für deren Weiterverwendung. Unternehmen können gezielt geschützte Verwaltungsdaten anfragen und diese perspektivisch, sofern rechtlich geregelt, datenschutzkonform für die Entwicklung oder Weiterentwicklung ihrer Geschäftsmodelle verwenden. Das Verschneiden von Drittdaten kann ermöglichen, neue Potenziale und Geschäftsmodelle zu entdecken und sogar besonders geschützte Daten aus weiteren EU-Mitgliedstaaten zu verwenden. Internationale Trends und nationale Entwicklungen können durch neue Analysemethoden frühzeitig berücksichtigt und mit schnelleren Reaktionszeiten Problemlösungen bereitgestellt werden. Die EU-Verordnung reduziert zudem Marktnachteile, da sie die Rahmenbedingungen verbessert und die Gebühren für kleine und mittlere Unternehmen sowie Startups vergünstigt (siehe Erwägungsgrund 2 zur Verordnung [EU] 2022/868).

Wissenschaft

Mit der Übernahme der neuen Aufgaben werden durch die Servicestelle Bedarfe aus Forschung und Wissenschaft bedient. Sie kann dazu beitragen, dass Forschende Verwaltungsdaten nutzen und somit Wissen generieren können, welches sich aus bisher unerschlossenen Daten öffentlicher Einrichtungen speisen lässt. Dies geschieht bereits vielfach in den Forschungs-

datenzentren großer Verwaltungen mit wissenschaftlich wertvollen Datenbeständen, wie im Bereich der Arbeitsmarktforschung durch Statistiken der Bundesagentur für Arbeit oder die Forschungsdatenzentren des Bundes und der Länder (Brenzel/Zwick, 2022).

Als zentrale Anlaufstelle für Daten aus öffentlichen Einrichtungen, die der Wissenschaft zugänglich gemacht werden, kann die neue Servicestelle die Forschungsdatenzentren und weitere wissenschaftliche Institutionen noch stärker als bisher miteinander vernetzen.

Den Zugriff auf weitere Datenbestände ergänzen die Anforderungen an Interoperabilität und einheitliche Standards, welche durch Anfragen und ein größeres Volumen an bereitgestellten Verwaltungsdaten besser durchzusetzen sind. Einheitliche Standards stellen eine große Herausforderung dar, da weder innerhalb der Forschungsdaten, noch bei Verwaltungsdaten gemeinsame Standards gelten, Schnittstellen kaum existieren und Anfragen stets gesondert bearbeitet werden. Die Servicestelle als erste Anlaufstelle bietet die Chance, solche Anfragen zu bündeln, Rückmeldungen zu konzentrieren und allgemeine Standards zu verbreiten.

Verwaltung

Öffentliche Einrichtungen, die sich aufgrund der Rahmenbedingungen durch den Data Governance Act mit der Frage der Bereitstellung geschützter Verwaltungsdaten befassen, werden vielfach unterschiedlich aufgestellt sein sowie divergierende Datenkompetenzen und Ressourcen für die Bereitstellung aufweisen. Bei einer steigenden Zahl von Anfragen kann die Servicestelle hohe Standards beim bestmöglichen Datenmanagement in die Verwaltungslandschaft über alle föderalen Ebenen hinweg tragen. So kann sie wiederum selbst mittelbar von Verbesserungen für die Statistikproduktion profitieren sowie einen nutzendenfreundlichen Datenzugang fördern. Die Kompetenzen in den öffentlichen Stellen selbst erhöhen sich langfristig zudem über die Arbeiten an einem konkreten angefragten Datensatz. Weiterhin sind Entbürokratisierungseffekte, die Steigerung einer Open-Data-Kultur sowie mehr Offenheit gegenüber KI-Anwendungen in der Verwaltung zu erwarten. Die Verwaltung wird Impulse für das Once-Only-Prinzip oder die Umsetzung des Online-Zugangsgesetzes erhalten.

Durch die Übernahme der Unterstützungsleistungen nach Artikel 7 Data Governance Act wird die konkrete Beratung mittelfristig für anfragende öffentliche Stellen, wie etwa kleinere Kommunalverwaltungen, eine effizientere Datennutzung und einen verbesserten Datenaustausch ermöglichen.

Die damit einhergehende verbesserte Datengrundlage und Datennutzung wird für politische Entscheidungstragende Vorteile mit sich bringen und die Entscheidungsfindung für Projekte oder Gesetzesvorhaben beschleunigen. Auch von der stetig wachsenden Metadatenbank werden Entscheidungsträgerinnen und -träger profitieren können.

Gesellschaft

Eine erfolgreiche Aufgabenumsetzung wird mit verschiedenen positiven Effekten für die Gesellschaft einhergehen. Durch innovative neue Produkte und Geschäftsideen, beispielsweise im Bereich Mobilität, können Bürgerinnen und Bürger perspektivisch von der Nutzung von Verwaltungsdaten profitieren, deren Erstellung durch öffentliche Gelder ermöglicht wurde. Unternehmen, die sich aufgrund von verbesserten Datenzugängen positiv entwickeln, erzeugen zudem wirtschaftliche Erfolge, welche der Gesellschaft zugutekommen. Durch eine verbesserte Datenverfügbarkeit ermöglichte Fortschritte, etwa im Bereich umweltpolitischer Herausforderungen und Nachhaltigkeit, gehen ebenfalls mit positiven gesamtgesellschaftlichen Auswirkungen einher.

Die Möglichkeit, Anfragen zu stellen, und ein transparenteres Verwaltungshandeln auf Grundlage von Daten stärken das Vertrauen von Bürgerinnen und Bürgern in staatliches Handeln und können zu einer erhöhten Datenkompetenz in der Gesellschaft führen.

4

Fazit und Ausblick: Internationaler Austausch für einen europäischen Datenbinnenmarkt

Der Data Governance Act verbessert als EU-Verordnung die Rahmenbedingungen zur Bereitstellung von geschützten Daten und ist in allen Mitgliedstaaten umzusetzen, sodass sich europaweit ein Zugang zu Daten in öffentlichen Einrichtungen eröffnen und stetig erweitern wird. Mit der Umsetzung des Data Governance Acts leistet Deutschland einen wichtigen Beitrag zur Europäischen Datenstrategie mit einem gemeinsamen Datenbinnenmarkt und mehr Datentransparenz in Europa. Hinzu kommen ein besserer Wettbewerb, neue Partnerschaften im Datennetzwerk und ein vertrauenswürdiger Umgang mit geschützten Daten öffentlicher Einrichtungen (Europäische Kommission, 2020).

Um dies noch besser leisten zu können und die Mehrwerte für alle Zielgruppen aus Wirtschaft, Wissenschaft und Gesellschaft sowie Verwaltungen auszubauen, müssen weitere Gesetzesänderungen in Deutschland folgen. Zuerst wäre hier das von der Regierung der 20. Legislaturperiode geplante Forschungsdatengesetz zu nennen. Dieses sollte das Angebot an die Wissenschaft erweitern und den Zugang zu Daten für die Forschung verbessern. Außerdem zielte es darauf ab, Daten miteinander verknüpfen zu können, um die Auswertungspotenziale bestehender Datenbestände verschiedener Datenanbieter in einem Micro Data Center als Datentreuhänder der Wissenschaft zu heben (BMBF, 2024; RatSWD, 2024).

Neben den bereits relativ guten Datenzugangsmöglichkeiten der Wissenschaft aufgrund des Wissenschaftsprivilegs sollten diese auch in Bezug auf Daten öffentlicher Einrichtungen im Sinne einer erweiterten Transparenz für Wirtschaft und Gesellschaft ausgebaut werden. Da der Data Governance Act keine Datenzugangsansprüche enthält, Daten öffentlicher Verwaltungen aber zugleich von der Wirtschaft und den Bürgerinnen und Bürgern finanziert werden, erwachsen hier zunehmend Ansprüche, Daten dieser Verwaltungsprozesse nutzen zu können. Ein Gesetz, welches Datenzugangsansprüche für ein transparentes Verwaltungshandeln enthält, könnte der Politik zudem die Chance

zur Entbürokratisierung bieten. Der Aufbau einer Metadatenbank leistet in dieser Hinsicht bereits nach der gültigen EU-Verordnung einen wichtigen Beitrag.

Perspektivisch gelten diese Vorteile einer verbesserten Datenverfügbarkeit nicht nur für geschützte Daten, sondern auch für nicht geschützte, die vorwiegend aus Gründen der schwierigen Datenverarbeitung und -bereitstellung bisher nicht verfügbar gemacht werden konnten. Hier wird die Servicestelle zur Umsetzung des Data Governance Acts um ungeschützte Daten erweitert werden können: Verbesserte Rahmenbedingungen, die technische und rechtliche Beratung sowie eine sichere IT-Verarbeitungsumgebung und eine transparente Metadatenbank können an einer zentralen Stelle geschützte und ungeschützte Daten öffentlicher Verwaltung vermitteln.

Die Rolle eines Datentreuhänders für Daten öffentlicher Einrichtungen lässt sich zudem auf privat gehaltene Daten von Unternehmen übertragen, die diese aus einer Hand mit Verwaltungsdaten für die Weiterentwicklung von Geschäftsmodellen und für transparente und effiziente Verwaltungs- und Wirtschaftsprozesse nutzbar machen könnten. Auf europäischer Ebene wird die Verfügbarkeit von Daten des Weiteren durch Initiativen wie die europäischen Data Spaces verbessert (Europäische Kommission, 2025b).

Bemängelt wird von der EU-Kommission die bisher unzureichende Umsetzung des Data Governance Acts, da zum Stand Dezember 2024 zehn Mitgliedstaaten die zweite Phase des Vertragsverletzungsverfahrens erreicht haben, darunter Deutschland (Europäische Kommission, 2025c). Ein Grund für die mangelnde Umsetzung sind die weitreichenden Ansprüche an die Leistungen, die durch die zuständigen Behörden zu erbringen sind. Hinzu kommen große nationale Unterschiede in den Datenzugangsansprüchen, die sich auch darauf begründen, dass mehrere EU-Verordnungen bisher nicht optimal ineinandergreifen (siehe Open Data Directive und Data Governance Act). Zudem liegen unterschiedliche Definitionen von Daten in den verschiedenen EU-Verordnungen vor, welche eigentlich ineinandergreifen und sich ergänzen beziehungsweise Lücken schließen sollten (Baloup und andere, 2021). Es könnten daher nicht nur auf nationaler, sondern auch auf EU-Ebene neue Gesetze folgen, um die Datenverfügbarkeit für einen europäischen Datenbinnenmarkt weiter zu verbessern. 

LITERATURVERZEICHNIS

Baloup, Julie/Bayamlıoğlu, Emre/Benmayor, Alik/Ducuing, Charlotte/Dutkiewicz, Lidia/Lalova, Teodora/Miadzvetzkaya, Yuliya/Peeters, Bert. *White Paper on the Data Governance Act*. CiTiP Working Paper Series. 2021. [Zugriff am 12. März 2025]. Verfügbar unter: papers.ssrn.com

Bens, Arno/Schukraft, Stefan. *Registermodernisierung und Verwaltungsdatennutzung in der amtlichen Statistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2018, Seite 11 ff.

BfDI (Die Bundesbeauftragte für den Datenschutz und die Informationsfreiheit). *BfDI wird Vertreter im neuen Europäischen Dateninnovationsrat*. [Zugriff am 12. März 2025]. Verfügbar unter: www.bfdi.bund.de

BMBF (Bundesministerium für Bildung und Forschung). *Eckpunkte BMBF-Forschungsdatengesetz*. 2024. [Zugriff am 12. März 2025]. Verfügbar unter: www.bmbf.de

Brenzel, Hanna/Zwick, Markus. *Eine informationelle Infrastruktur in Deutschland ist erwachsen – das Forschungsdatenzentrum des Statistischen Bundesamtes*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2022, Seite 54 ff.

Bundesregierung. *Datenstrategie der Bundesregierung. Eine Innovationsstrategie für gesellschaftlichen Fortschritt und nachhaltiges Wachstum*. Kabinettsfassung vom 27. Januar 2021. [Zugriff am 20. März 2025]. Verfügbar unter: www.bundesregierung.de

Europäische Kommission. *Eine europäische Datenstrategie. Mitteilung der Kommission an das Europäische Parlament, den Rat, den europäischen Wirtschafts- und Sozialausschuss und den Ausschuss der Regionen*. 2020. COM(2020) 66 final, Seite 5. [Zugriff am 12. März 2025]. Verfügbar unter: eur-lex.europa.eu

Europäische Kommission. *Data Governance Act explained*. 2025a. [Zugriff am 12. März 2025]. Verfügbar unter: digital-strategy.ec.europa.eu

Europäische Kommission. *Common European Data Spaces*. 2025b. [Zugriff am 12. März 2025]. Verfügbar unter: digital-strategy.ec.europa.eu

Europäische Kommission. *Commission calls on 10 Member States to comply with the Data Governance Act*. Pressemitteilung vom 16. Dezember 2024. 2025c. [Zugriff am 12. März 2025]. Verfügbar unter: digital-strategy.ec.europa.eu

Paschke, Anne/Rücker, Daniel. *DGA. Data Governance Act*. Kommentar. München 2024.

RatSWD (Rat für Sozial- und Wirtschaftsdaten). *RatSWD-Stellungnahme zu den BMBF-Eckpunkten zum Forschungsdatengesetz*. 2024. [Zugriff am 12. März 2025]. Verfügbar unter: www.konsortswd.de

LITERATURVERZEICHNIS

Referentenentwurf des Bundesministeriums für Wirtschaft und Klimaschutz. *Entwurf eines Gesetzes zur Durchführung der EU-Verordnung über europäische Daten-Governance* (Daten-Governance-Gesetz – DGG). [Zugriff am 12. März 2025]. Verfügbar unter: www.bmwk.de

Specht, Louisa/Hennemann, Moritz. *Data Governance Act. Article-by-Article Commentary*. Baden-Baden 2023.

Vestager, Margrethe. *Speech by Executive Vice-President Vestager at the Press Conference on the Data Governance Act and the Action Plan on Intellectual Property*. 2020. [Zugriff am 12. März 2025]. Verfügbar unter: ec.europa.eu

Vestager, Margrethe. *Margrethe Vestager explains the EU's position in the global battle for data*. In: Euronews 2021. [Zugriff am 12. März 2025]. Verfügbar unter: www.euronews.com

Vorgrimler, Daniel/Decker, Jörg. *Die Rolle der amtlichen Statistik innerhalb des Daten-ökosystems*. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2021, Seite 17 ff.

RECHTSGRUNDLAGEN

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.

Richtlinie (EU) 2019/1024 des Europäischen Parlaments und des Rates vom 20. Juni 2019 über offene Daten und die Weiterverwendung von Informationen des öffentlichen Sektors (Neufassung) (Amtsblatt der EU Nr. L 172, Seite 56).

Verordnung (EU) 2016/679 des Europäischen Parlaments und des Rates vom 27. April 2016 zum Schutz natürlicher Personen bei der Verarbeitung personenbezogener Daten, zum freien Datenverkehr und zur Aufhebung der Richtlinie 95/46/EG (Datenschutz-Grundverordnung) (Amtsblatt der EU Nr. L 119, Seite 1).

Verordnung (EU) 2022/868 des Europäischen Parlaments und des Rates vom 30. Mai 2022 über europäische Daten-Governance und zur Änderung der Verordnung (EU) 2018/1724 (Daten-Governance-Rechtsakt) (Amtsblatt der EU Nr. L 152, Seite 1).

INTEGRATIONSBERICHTERSTATTUNG IM DASHBOARD DEUTSCHLAND – DIE NEUE PLATTFORM DASHBOARD INTEGRATION

Coşkun Canan, Nina Eppers, Frederik Czerniak, Rolf Mach

📌 **Schlüsselwörter:** Migration – Integrationsindikatoren – Integrationsbericht – Einwanderungsgeschichte – Datenportal

ZUSAMMENFASSUNG

Das Dashboard Integration ist ein neuer eigenständiger Bereich des interaktiven Datenportals Dashboard Deutschland. Es stellt umfassende Informationen zur Situation eingewanderter Menschen und ihrer Nachkommen bereit. Das Dashboard macht alle 60 Indikatoren des 14. Integrationsberichts der Integrationsbeauftragten der Bundesregierung mit interaktiven Grafiken zugänglich. Zudem lassen sich die Daten herunterladen. Der Beitrag stellt die Arbeitsschritte in der Datenaufbereitung und im Datenmanagement dar und veranschaulicht die Nutzungsmöglichkeiten der visualisierten Daten anhand von Beispielen.

📌 **Keywords:** migration – integration indicators – integration report – immigration history – data portal

ABSTRACT

“Dashboard Integration” is a new, separate section of the interactive data portal Dashboard Germany. It provides comprehensive information on the situation of people who immigrated to Germany and their descendants. The dashboard uses interactive graphs to visualise all 60 indicators covered by the 14th Integration Report of the Federal Government Commissioner for Migration, Refugees and Integration. The data can also be downloaded. This article describes the individual steps involved in data processing and data management and illustrates possible uses of the visualised data with the help of examples.

Dr. Coşkun Canan

hat an der Humboldt-Universität zu Berlin im Fach Soziologie promoviert. Er ist Wissenschaftlicher Mitarbeiter im Referat „Bevölkerungsspezifische Auswertungen und Analysen aus dem Mikrozensus“ des Statistischen Bundesamtes und befasst sich vor allem mit Themen der Migration und Integration.

Nina Eppers

ist Sozialwissenschaftlerin und Wissenschaftliche Mitarbeiterin im Referat „Bevölkerungsspezifische Auswertungen und Analysen aus dem Mikrozensus“ des Statistischen Bundesamtes.

Frederik Czerniak

hat einen Master in Betriebswirtschaftslehre und ist Wissenschaftlicher Mitarbeiter in der Gruppe „Externe Kommunikation“ des Statistischen Bundesamtes. Er übt die Funktion des Projektmanagers aus und ist stellvertretender Projektleiter des Dashboard Deutschland.

Rolf Mach

ist Diplom-Informatiker und arbeitet als Wissenschaftlicher Mitarbeiter im Referat „IT-Kompetenzzentrum Informationsbereitstellung und -veröffentlichung“ in der Rolle als Projektmanager beim Dashboard Deutschland.

1

Einleitung

Im Dashboard Deutschland bündelt das Statistische Bundesamt seit dem Jahr 2020 das Datenangebot zur aktuellen Wirtschaftssituation Deutschlands. Am 11. Dezember 2024 ging das Dashboard Integration online, welches das Datenangebot des Dashboard Deutschland um umfangreiche sozioökonomische Strukturdaten zum Integrationsgeschehen erweitert.¹ Das Dashboard Integration bietet einen umfassenden Überblick zur Integration in allen gesellschaftlichen Bereichen, unter anderem Arbeitsmarkt, schulische und berufliche Bildung, Sicherheit und Partizipation. Dargestellt werden in ihm die Indikatoren des Integrationsberichts der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration (Integrationsbeauftragte, 2024).

Im Jahr 2023 lebten in Deutschland 16,2 Millionen Eingewanderte und 5 Millionen Nachkommen, bei denen beide Elternteile eingewandert sind ([Dashboard Integration Demografie](#)): Für ein gutes Viertel der Bevölkerung gehört damit ihre Einwanderungsgeschichte zur Lebenswirklichkeit. Die Integration eingewanderter Menschen und ihrer Nachkommen betrifft alle gesellschaftlichen Bereiche und ist Grundvoraussetzung für Teilhabe und Zusammenhalt. Das Dashboard Integration beschreibt die Teilhabe von Menschen mit Einwanderungsgeschichte am gesellschaftlichen Leben in Deutschland anhand der Indikatoren des Integrationsberichts, den die Integrationsbeauftragte gemeinsam mit Wissenschaftlerinnen und Wissenschaftlern des Deutschen Zentrums für Integrations- und Migrationsforschung (DeZIM) entwickelt hat (Integrationsbeauftragte, 2024). Mit der Veröffentlichung des Dashboard Integration bündelt das Statistische Bundesamt das Indikatorenangebot zu Migration und Integration. Es bietet eine wissenschaftsbasierte Informationsplattform zu einem gesellschaftlichen Thema, das sich neben der breiten Öffentlichkeit auch an relevante Stakeholder wie politische Entscheiderinnen und Entscheider richtet.

1 Das Dashboard Integration wurde vom Statistischen Bundesamt in Kooperation mit der Beauftragten der Bundesregierung für Migration, Flüchtlinge und Integration, zugleich Beauftragte der Bundesregierung für Antirassismus, ergänzend zum Integrationsbericht 2024 aufgebaut.

Dieser Beitrag stellt Aufbau und Inhalt des Dashboard Integration dar. Während Kapitel 2 die Themenfelder benennt und die thematische Verbindung zum Integrationsbericht erläutert, gibt Kapitel 3 zunächst einen Überblick über Datenkoordination und -aufbereitung. Anschließend beschreibt Kapitel 4 die erweiterten IT-Strukturen des Dashboard Deutschland zur Umsetzung des Dashboard Integration. Kapitel 5 präsentiert schließlich Inhalte und Darstellungsmöglichkeiten anhand ausgewählter Dashboard-Grafiken. Zuletzt gibt Kapitel 6 einen Ausblick über die anstehenden Aufgaben nach erfolgreicher Veröffentlichung.

2

Das neue Dashboard Integration

Das [Dashboard Integration](#) ist ein neues Datenportal unter dem Dach des Dashboard Deutschland. Es stellt statistische Ergebnisse interaktiv visualisiert dar und ermöglicht das Herunterladen von Daten. Mit dem [Dashboard Konjunktur](#) lag der Schwerpunkt bislang auf der Konjunkturbeobachtung (Stern und andere, 2024). Durch den Aufbau des Dashboard Integration kommen nun sozioökonomische Strukturdaten zum Integrationsgeschehen hinzu.

Das Dashboard Integration stellt Ergebnisse zu 14 Themenfeldern der Migration und Integration bereit, die anhand von 60 Indikatoren abgebildet werden. ➤ **Übersicht 1** Die Indikatoren basieren außer auf Daten der Statistischen Ämter des Bundes und der Länder auch auf Statistiken anderer Behörden sowie auf wissenschaftlichen Studien und Befragungen. Die unterschiedlichen Datenquellen sind im Dashboard Integration entsprechend kenntlich gemacht.

➤ **Übersicht 2** stellt beispielhaft das Themenfeld „Arbeitsmarktintegration“ dar. Dieses Themenfeld wird insgesamt mit fünf Indikatoren abgebildet. Bei einem Indikator stammen die Daten aus der amtlichen Statistik, zwei Indikatoren beruhen auf administrativen Daten anderer Behörden. Jeweils einem Indikator liegen Daten aus der Wissenschaft beziehungsweise sowohl Daten der amtlichen Statistik als auch administrative Daten von Behörden zugrunde. Insgesamt basiert die Hälfte aller Indikatoren auf Daten der Statistischen Ämter des Bundes und der Länder.

Integrationsberichterstattung im Dashboard Deutschland – die neue Plattform

Dashboard Integration

Übersicht 1

Aufbau des Dashboard Integration

Themenfelder	Indikatoren
Demografie	Bevölkerung nach Einwanderungsgeschichte und Staatsangehörigkeit; Aufenthaltsdauer der Eingewanderten; Lebensform; regionale Verteilung; Wanderungen zwischen Deutschland und dem Ausland
Rechtsstatus	Aufenthaltsstatus und befristete Aufenthaltstitel; Einbürgerungen und ausgeschöpftes Einbürgerungspotenzial; Geduldete und Zahl der erteilten Aufenthaltserlaubnisse nach einer Duldung
Frühkindliche und schulische Bildung	Betreuungsquote; schulische Abschlüsse und Abgänge; Risikolagen für Bildungserfolg; schulische Kompetenzen; Übergangstatus nach der allgemeinbildenden Schule
Sprache	Integrationskursteilnahme und Deutsch-Tests für Zuwanderer; subjektiv eingeschätzte mündliche Deutschkenntnisse; Sprache bei der Mediennutzung
Berufliche und akademische Bildung	Höchster beruflicher Abschluss; Abschluss mindestens der Sekundarstufe II oder in Bildung beziehungsweise Ausbildung; Absolvierende beruflicher Schulen; Einmündungsquote; Studierende; Erfolgsquote von Studierenden; berufliche Fort- und Weiterbildung
Arbeitsmarktintegration	Erwerbstätigenquote; Arbeitslosenquote; Wirtschaftszweige; Anerkennung ausländischer Qualifikationen; wahrgenommene Arbeitsplatzsicherheit
Einkommen	Mittlere Verdienste; Armutsgefährdungsquote; überwiegender Lebensunterhalt aus staatlichen Leistungen; Rentenbezug
Soziale und politische Partizipation	Engagementquote; soziale Kontakte; Wahlbeteiligung; Zugehörigkeitsgefühl zur Gesellschaft
Diversitätsorientierte Öffnung	Beschäftigte im öffentlichen Dienst; Zufriedenheit mit behördlichen Dienstleistungen; Abgeordnete im Bundestag und in den Landtagen
Wohnen	Mietbelastung; Wohnfläche je Person; Wohnzufriedenheit; Wohneigentum
Gesundheit	Allgemeiner Gesundheitszustand; psychische Gesundheit; Prävalenz von Übergewicht und Adipositas; Inanspruchnahme von Leistungen der medizinischen Rehabilitation
Leben in Sicherheit	Allgemeines Sicherheitsgefühl; Kriminalitätsoffer; Verurteilte; Tatverdächtige; spezifische Kriminalitätsfurcht; Vertrauen in die Polizei
Hasskriminelle Gewalt	Amtlich registrierte Fälle hasskrimineller Gewaltdelikte und ihre Opfer; Fälle rechter, rassistischer und antisemitischer Gewalt; Selbsteinschätzung zum Tatmotiv körperlicher Angriffe
Rassismus	Subjektive rassistische Diskriminierungserfahrung; Einstellung zum Zuzug unterschiedlicher ethnischer Gruppen; Zustimmung zu Aussagen unterschiedlicher Facetten gruppenbezogener Menschenfeindlichkeit; Wahrnehmung von Rassismus

Die Indikatoren liegen in der Regel als Zeitreihen ab dem Jahr 2005 vor. Das zentrale Differenzierungsmerkmal bildet das Konzept Einwanderungsgeschichte, da es ermöglicht, die Situation der Eingewanderten (erste Generation) und deren Nachkommen (zweite Generation) mit der Situation von Menschen ohne Einwanderungsgeschichte zu vergleichen (Canan/Petschel,

2023).¹² Um eine intersektionale Betrachtung zu ermöglichen, erfolgen zudem meist Untergliederungen nach Geschlecht, Alter oder anderen relevanten Merkmalen. Entsprechend sind mehrere Grafiken je Indikator möglich. Während die veröffentlichte Print- oder PDF-Version des Integrationsberichts aufgrund einer festgelegten

² In einigen Daten liegt dieses Merkmal nicht vor. In diesen Fällen erfolgt die Differenzierung nach der Staatsangehörigkeit oder dem Migrationshintergrund.

Übersicht 2

Indikatoren und Datenquellen des Themenfelds „Arbeitsmarktintegration“

Themenfeld	Indikator	Datenquelle
Arbeit	Erwerbstätigenquote	Mikrozensus, Beschäftigungsstatistik der Bundesagentur für Arbeit
	Arbeitslosenquote	Arbeitslosenstatistik der Bundesagentur für Arbeit
	Wirtschaftszweige	Beschäftigungsstatistik der Bundesagentur für Arbeit
	Anerkennung ausländischer Qualifikationen	Statistik nach den Berufsqualifikationsfeststellungsgesetzen des Bundes und der Länder
	Wahrgenommene Arbeitsplatzsicherheit	Sozio-oekonomisches Panel

Seitenzahl und des statischen Formats in der Visualisierung der statistischen Daten begrenzt ist, bietet das Dashboard Integration ein erweitertes Grafikangebot der Indikatoren – insgesamt stehen rund 500 Grafiken eingebettet in 189 Seiten (Kacheln) zur Verfügung. Mehr als die Hälfte der Grafiken (51 %) basieren derzeit auf Daten der Statistischen Ämter des Bundes und der Länder. Daten aus der Wissenschaft liegen 35 % und administrative Daten von Behörden 14 % der Grafiken zugrunde. Der Mikrozensus ist mit einem Anteil von rund 34 % die häufigste Datenquelle für Grafiken im Dashboard Integration.

Während der Integrationsbericht die statistischen Ergebnisse wissenschafts- und textbasiert einordnet, konzentriert sich das Dashboard Integration auf die Visualisierung der Daten und begrenzt Textelemente auf Metainformationen, wie methodische Erläuterungen und Lesebeispiele. Im Unterschied zum Integrationsbericht ermöglicht das Dashboard Integration eine Reihe von interaktiven Funktionen und das Herunterladen der visualisierten Daten. Letztlich ergänzen sich beide Veröffentlichungsformate auf eine sinnvolle und ergiebige Weise.

Exkurs

Der Integrationsbericht in seiner jetzigen Form als wissenschaftsbasierter und indikatorgestützter Bericht erschien erstmals 2021. Der aktuelle zweite Integrationsbericht mit dem Titel „Teilhabe in der Einwanderungsgesellschaft“ baut darauf auf (Integrationsbeauftragte, 2024).¹³ Beim Integrationsbericht 2024 kam dem Statistischen Bundesamt mit der Entwicklung des Dashboard Integration eine neue und wichtige Rolle zu. Im Rahmen einer Verwaltungsvereinbarung mit der Integrationsbeauftragten der Bundesregierung hat das Statistische Bundesamt das Dashboard Integration als Informationsplattform entwickelt. Es soll den textbasierten Integrationsbericht ergänzen, mit dem die Integrationsbeauftragte das Deutsche Zentrum für Integrations- und Migrationsforschung (DeZIM) beauftragt hatte.

Dashboards eignen sich vor allem dazu, datenbasierte Themeninhalte zu visualisieren und zu kommunizieren (Sarıkaya und andere, 2019; Bach und andere, 2023). Besonders für die indikatorgestützte Integrationsberichterstattung ist daher das Dashboard Integration ein

wertvolles Instrument, um Daten und Informationen zu präsentieren und zu verbreiten.

3

Vorbereitungen: Datenkoordination und -aufbereitung

Das Dashboard Integration und der Integrationsbericht sind nicht nur inhaltlich, sondern auch visuell und methodisch aufeinander abgestimmt. Das heißt beide Veröffentlichungsformate enthalten wiedererkennbare Grafiken und Metainformationen, die auf derselben Datengrundlage beruhen. Um eine einheitliche Datengrundlage für beide Veröffentlichungsformate herzustellen, wurde zwischen dem Statistischen Bundesamt und dem Deutschen Zentrum für Integrations- und Migrationsforschung ein Datenaustausch etabliert. Das Statistische Bundesamt stellte dabei amtliche Daten der Statistischen Ämter des Bundes und der Länder für den Integrationsbericht zur Verfügung. Das Deutsche Zentrum für Integrations- und Migrationsforschung wiederum teilte die im Integrationsbericht verwendeten Daten der übrigen Datenanbieter. Um das Ziel der gemeinsamen Datengrundlage zu erreichen, war die Datenkoordination entscheidend und aufgrund des großen Datenumfangs sehr komplex. In mehreren Abstimmungsrunden wurden Datenbedarfe ermittelt und Fragen zu methodischen und statistischen Konzepten geklärt. Dabei wurde beispielsweise das seit dem Jahr 2023 vom Statistischen Bundesamt eingeführte Konzept „Einwanderungsgeschichte“ (Canan/Petschel, 2023) als wesentliches Differenzierungsmerkmal zur Beschreibung der Bevölkerung nach Migrationsbezug festgelegt.

Darüber hinaus wurde sich über die gemeinsame Übermittlung von Metadaten abgestimmt. Die Metadaten umfassen Informationen zu den statistischen Daten, wie Quellenhinweise und methodische Erläuterungen. Dazu wurde ein einheitliches Datenblatt zur Synchronisierung der beiden Veröffentlichungsformate verwendet, das neben den Metadaten auch Informationen zur Umsetzung der statistischen Daten im Integrationsbericht erfasste (wie Beispielgrafiken).

Die ausgetauschten Daten wurden nach einer ebenfalls abgestimmten Nomenklatur jeweils als Datentabellen im CSV-Format bereitgestellt. Nach dem Austausch

3 Die Integrationsbeauftragte erstattet gemäß § 94 Absatz 2 Aufenthaltsgesetz dem Deutschen Bundestag alle zwei Jahre Bericht. Dieser Lagebericht wurde ab 2024 auf Dauer in den wissenschaftsbasierten und indikatorergestützten Integrationsbericht überführt (Integrationsbeauftragte, 2023, hier: Seite 7).

etablierte das Statistische Bundesamt einen Qualitätssicherungsschritt, sodass nur geprüfte Daten in das Dashboard eingingen. Auffälligkeiten in den Daten wurden wechselseitig gemeldet und die Daten bei Korrekturbedarf wieder neu aufbereitet und geliefert.

Bevor die Daten im Content Management System umgesetzt werden konnten, mussten sie auf die Datenplattform des Dashboard Deutschland geladen werden. Dies geschah über JSON-Dateien (JavaScript Object Notation).¹⁴ Anschließend erfolgten die Weiterverarbeitung und Visualisierung.

4

Erweiterung der technischen Infrastruktur

Von der Weiterentwicklung der IT-Infrastruktur des Dashboard Deutschland konnte auch das Dashboard Integration deutlich profitieren. Zum einen wurde die Nutzung der Datenplattform flexibilisiert, zum anderen wurde die IT-Anwendung zur Anlage und Verwaltung von Daten zu einem vollständigen Content Management System erweitert. Beide Ausbaustufen heben Synergien bei der Pflege der Inhalte des bereits existierenden Dashboard Konjunktur und ermöglichen künftig einen kosteneffizienten Betrieb der gesamten Dashboard-Deutschland-Technologie.

4.1 Flexibles Datenmodell über Virtual Cubes und JSON-Dateien

Das bisher zugrunde liegende Datenmodell im Dashboard Deutschland war vordefiniert, das heißt jede Grafik hatte eine eigene Datenbasis mit Merkmalen und Merkmalsausprägungen, die zuvor entsprechend definiert und mit Daten geladen werden mussten.

Dieses starre Datenmodell wurde flexibilisiert, sodass fortwährende Änderungen möglich sind – was aufgrund der zahlreichen und komplexen Daten notwendig ist.

Das Dashboard Deutschland speicherte dazu alle darzustellenden Daten in einer Cloud-basierten Datenplattform. Diese Datenplattform konnte multidimensionale Datenstrukturen (Cubes) anbinden und intern nachbilden. So entstanden sogenannte Virtual Cubes.¹⁵ Um diese Form der Datenhaltung auch für das Dashboard Integration zu nutzen, wurden die statistischen Daten für das Dashboard Deutschland über JSON-Dateien angebunden, welche neben den Daten auch die dazugehörige Referenzstruktur enthielten. Die Virtual Cubes ermöglichen eine flexible Nutzung der Datenplattform, da Fachanwenderinnen und -anwender¹⁶ Änderungen in der JSON-Datei selbst vornehmen können und die JSON-Datei ohne großen Aufwand erneut angebunden und zur Nutzung auf der Datenplattform bereitgestellt wird.

Die JSON-Datei entspricht keinem vordefiniertem Datenmodell, sondern definiert vielmehr den gesamten Virtual Cube selbst. Sie ist in den Metadatenblock und den Datenbeziehungsweise Contentblock aufgeteilt (für die schematische Darstellung einer JSON-Datei siehe Grafik 1). Der Metadatenblock enthält sowohl Informationen zur Strukturierung des Virtual Cube als auch Informationen, die sich auf die statistischen Daten beziehen und bei deren Visualisierung mitveröffentlicht werden. Der Datenblock gibt die Datentabelle mit den Merkmalen, Merkmalsausprägungen und den dazugehörigen Werten wider. Die JSON-Datei strukturiert also den Virtual Cube (Metadatenblock) und befüllt ihn anschließend mit Daten (Contentblock). Sofern der über den eindeutigen Identifikator „Topic-ID“ erkennbare Virtual Cube bereits existiert, wird er um neue Daten oder auch Metadaten modifiziert.

Somit werden sämtliche Merkmale und Merkmalsausprägungen dynamisch von Fachanwenderinnen und -anwendern über den JSON-Import gesteuert. Es ist allerdings nur Administratorinnen und Administratoren möglich, Virtual Cubes zu löschen.

Dieser neue Workflow verlangt eine besonders sorgfältige Erstellung der JSON-Dateien durch die Fachanwenderinnen und -anwender. Ansonsten besteht das Risiko, dass bereits vorhandene Datenbezüge (zum Beispiel

4 JSON (JavaScript Object Notation) ist ein textbasiertes Datenaustausch- und Datenverarbeitungsformat, das in webbasierten Anwendungen genutzt wird, um Informationen auszutauschen. Es kann vielfältige Informationen enthalten, die über einen eindeutigen Identifikator abgerufen werden (Bray, 2017; Pezoa und andere, 2016).

5 Im Dashboard Konjunktur entstehen Virtual Cubes mit der automatisierten Anbindung von GENESIS Cube mittels der [GENESIS API](#).

6 Unter Fachanwenderin und Fachanwender werden im Folgenden ganz allgemein die internen Nutzenden einer Anwendung oder eines Systems verstanden (zum Beispiel Mitarbeitende des Statistischen Bundesamtes).

Grafiken, die auf einem Virtual Cube beruhen) verloren gehen. So können beispielsweise Veränderungen im Metadaten- beziehungsweise Contentblock dazu führen, dass bestehende Datenbezüge nicht erkannt und auf deren Basis erzeugte Grafiken unbrauchbar werden (zum Beispiel durch veränderte Bezeichnungen der Merkmalsausprägungen).

Die „Data-Awareness“ der Fachanwenderinnen und Fachanwender für das flexible Datenmodell wurde durch die Administratorinnen und Administratoren entsprechend geschärft, sodass dieses Risiko vermieden wurde.

➤ **Grafik 1** zeigt beispielhaft eine JSON-Datei (einschließlich eines CSV-Anteils) als Grundlage für die Umsetzung der statistischen Daten und der dazugehö-

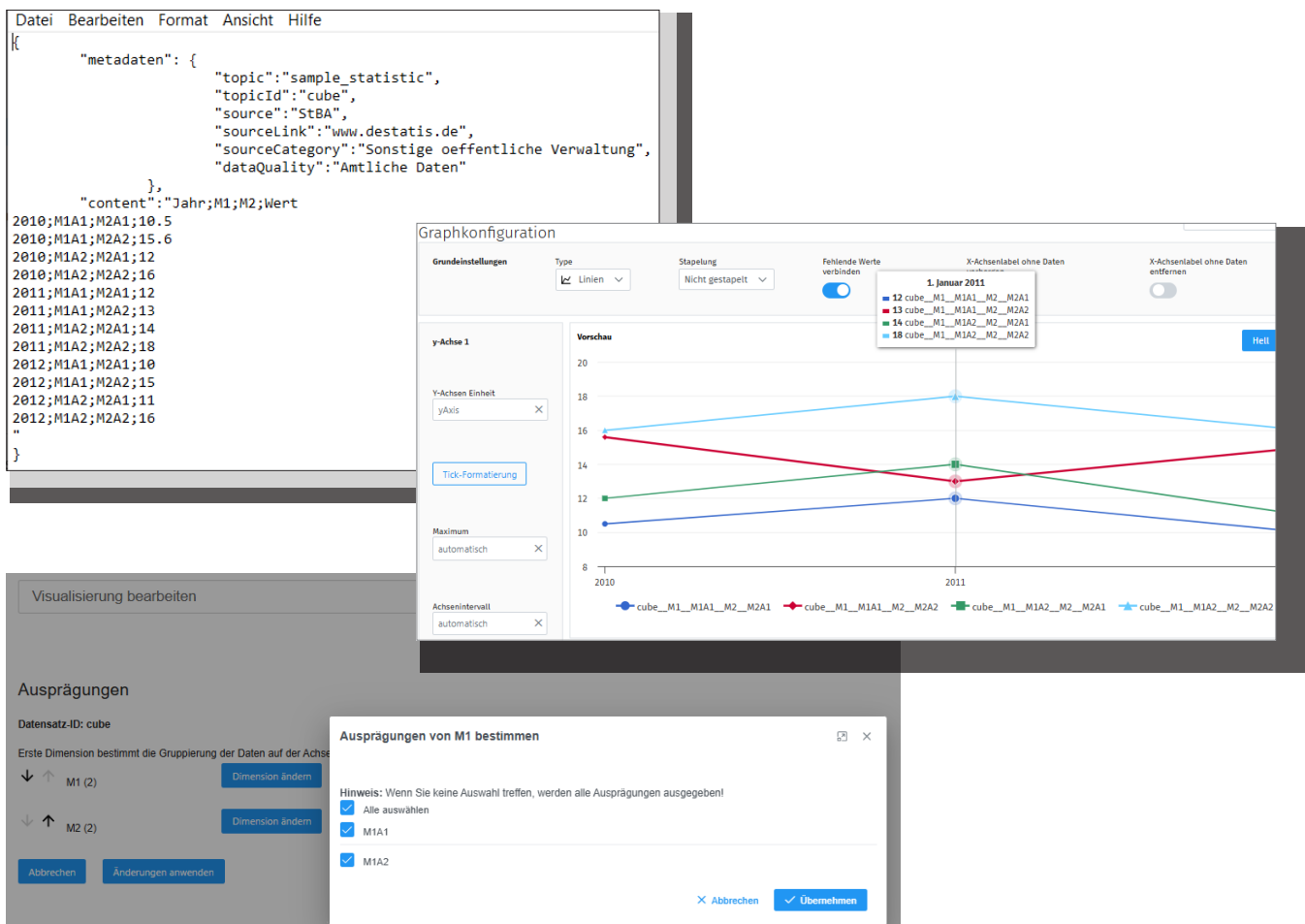
rigen Metadaten im Dashboard Integration. Der entstehende Virtual Cube enthält zusätzlich zu den Dimensionen „Jahr“ und „Wert“ noch zwei Merkmale M1 und M2 mit jeweils den Merkmalsausprägungen A1 und A2. Die Umsetzung umfasst neben der Visualisierung der statistischen Daten auch die Ergänzung der Metainformationen. Die Ergänzung weiterer Metainformationen zu den Daten sowie deren Visualisierung erfolgt über das Content Management System.

4.2 Das Content Management System

Bisher war das für das Dashboard Deutschland verwendete System nur teilweise in der Lage, Inhalte der Website ohne die Unterstützung durch Softwareentwicklerinnen und -entwickler anzupassen. Um die umfangreichen

Grafik 1

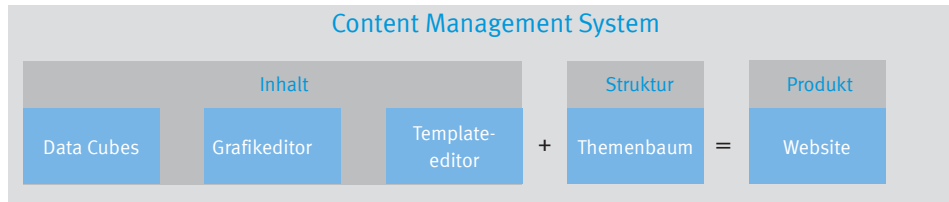
Die JSON-Datei, Auswahl der Merkmalsausprägung und der Grafikeditor



Integrationsberichterstattung im Dashboard Deutschland – die neue Plattform Dashboard Integration

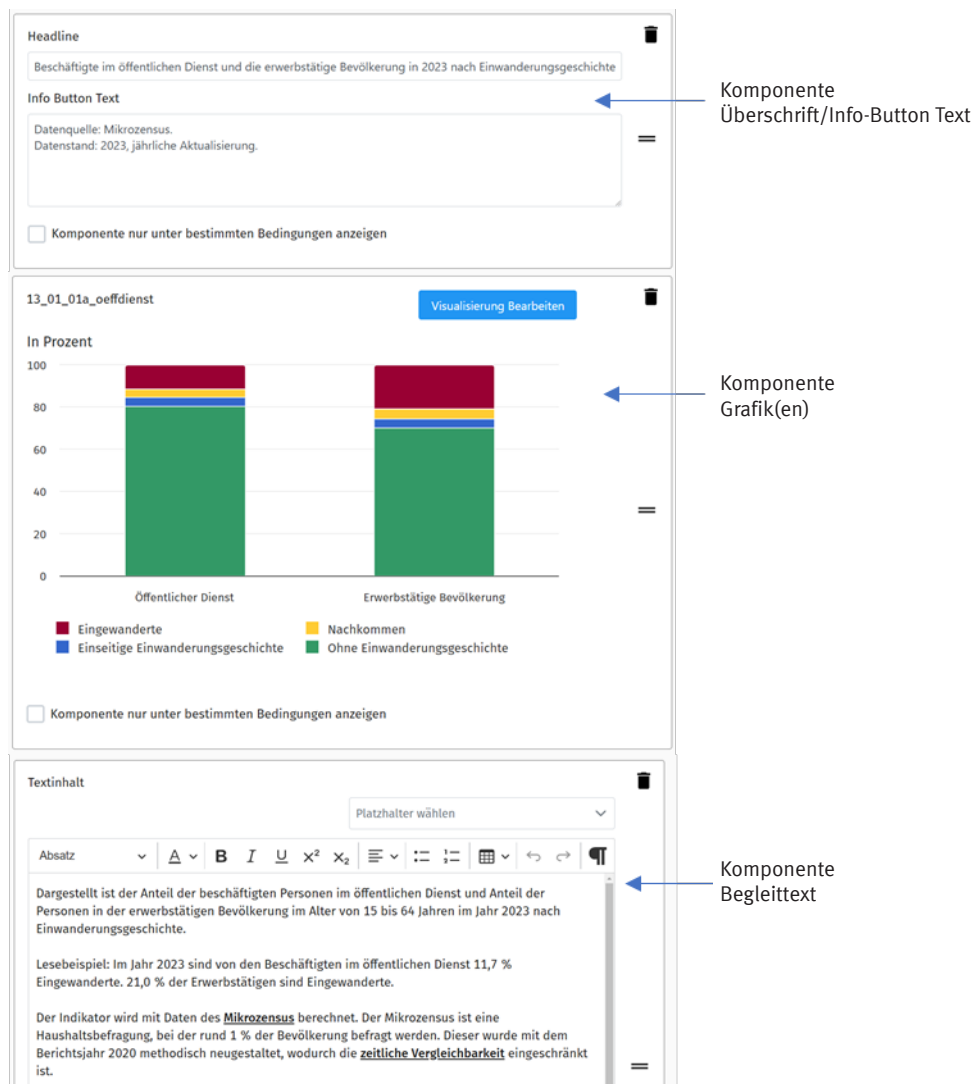
Grafik 2

End-to-End-Prozess des Content Management Systems



Grafik 3

Komponenten eines Indikators



Inhalte des Dashboard Integration zu verwalten sowie eine nachhaltige Lösung für beide Dashboards des Statistischen Bundesamtes zu schaffen, wurde die bisher bestehende Freigabe-Anwendung⁷ des Dashboard Deutschland zu einem vollständigen Content Management System umgebaut. Mit einem Content Management System lassen sich Inhalte erstellen, verwalten und veröffentlichen (Boiko, 2005).

➤ Grafik 2 stellt das Content Management System des Dashboard Integration dar. Es setzt sich aus den drei Bereichen Inhalt, Struktur und Produkt zusammen und befähigt die Fachanwenderinnen und -anwender, selbstständig Inhalte zu erzeugen, hierarchisch anzuordnen und auf einer Website zu publizieren.

Inhalt

Um Inhalte zu entwickeln, werden Daten aus den Virtual Cubes benötigt – wie im vorangegangenen Abschnitt beschrieben. Über den Grafikeditor lassen sich aus den ausgewählten Daten Darstellungsformen wie Linien- oder Balkendiagramme erzeugen und konfigurieren. Unter anderem sind Legenden- und Achsenbeschriftungen sowie weitere Parameter der Grafikeinheit, wie der per Mouseover sichtbare Tooltip mit Wert und Einheit des jeweiligen Datenpunktes, definiert. Hierfür wird auf die in der Freigabeanwendung integrierte und auf der Programmiersprache JavaScript basierende Softwarebibliothek Highcharts zurückgegriffen. Diese eignet sich zur Anwendung im Web-Browser und bietet spezielle Funktionen zur optimierten Bedienung auf mobilen Endgeräten, unter anderem das Zoomen via Touch (Shahzad und andere, 2016).

Um die Anlage der Inhalte im Dashboard Integration möglichst flexibel zu gestalten, sind die Elemente einer Dashboardkachel⁸ in verschiedene Komponenten zerlegt. Der Templateeditor erzeugt durch das Anordnen verschiedener Komponenten eine Schablone dieser Kachel. ➤ Grafik 3

Der Templateeditor ermöglicht es den Fachanwenderinnen und -anwendern, die verschiedenen Komponenten

in beliebiger Anzahl und Reihenfolge zu einer Vorlage zusammzusetzen und mit Inhalten zu befüllen. Das fertige Template repräsentiert eine Dashboardkachel, welche in einer Kompaktansicht sowie einer Detailansicht erzeugt wurde.

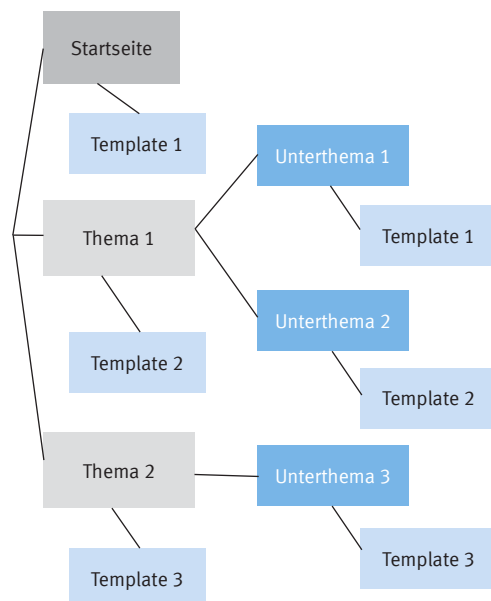
Struktur

Um die erzeugten Inhalte auf einer Website publizieren und beliebig anpassen zu können, wird eine Struktur benötigt. Die Inhalte des Dashboard Integration sind einfach hierarchisch über eine Startseite mit Themen und Unterthemen angeordnet.

➤ Grafik 4 zeigt eine exemplarische Gliederungsstruktur: Startseite, Themen und Unterthemen repräsentieren später die Webseiten, auf denen die Inhalte in Form von Templates erscheinen. Jedes Template stellt eine Kachel des Dashboards dar und wird immer in der Kompaktansicht und der Detailansicht abgebildet. Eine Kachel in den zwei Ansichten veranschaulicht ➤ Grafik 5.

Grafik 4

Ausschnitt aus dem Themenbaum des Dashboard Integration



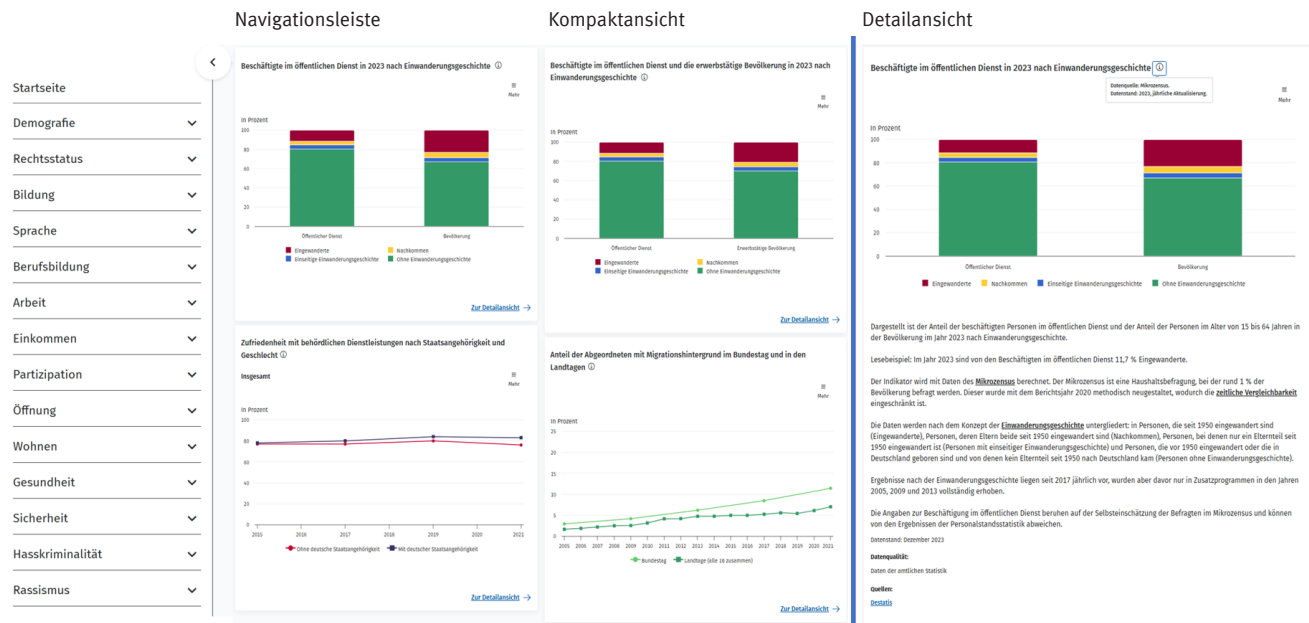
Auf der linken Themenseite sind vier Kacheln in der Kompaktansicht zum Themenfeld „Diversitätsorientierte Öffnung“ in der vorher definierten Reihenfolge abgebildet. Dabei wird der Titel mit dazugehörigem Infobutton und

7 Die Freigabe-Anwendung – auch Release-App genannt – ist eine eigens für das Dashboard Deutschland entwickelte IT-Anwendung zur Anlage und Verwaltung der Inhalte.

8 Eine Kachel bildet den Rahmen der Inhalte im Dashboard und besteht hauptsächlich aus einer Grafikeinheit und dazugehörigem Text. Im Dashboard Deutschland wird hier meist von „Indikator“ gesprochen.

Grafik 5

Ansichten im Dashboard Integration



einer Grafik je Kachel angezeigt. Die detaillierte Kachel auf der rechten Seite ist eine separate Website und über die Verlinkung „Zur Detailsicht“ erreichbar. Hier finden sich zusätzliche Inhalte in Form eines Begleittextes und die Möglichkeit, weitere Grafiken abzubilden ([Dashboard Integration Öffnung](#)). Am linken Rand ist außerdem die zuvor über den Themenbaum erzeugte Navigationsleiste zu finden.

5

Das Ergebnis: visualisierte Integrationsdaten

Das Dashboard Integration stellt statistische Daten zum Thema Migration und Integration visualisiert als Grafiken bereit. Textbasierte Metadaten rahmen diese Grafiken ein und helfen bei der Interpretation der Daten. Die interaktiven Elemente ermöglichen es den Nutzenden, die Grafiken stärker nach eigenen Bedürfnissen zu gestalten und sie in verschiedenen Formaten (svg, png, jpg) herunterzuladen. Die den Grafiken zugrunde liegenden Daten können ebenfalls heruntergeladen und im CSV-Format weiterverarbeitet werden. Besonders in der Integrationsberichterstattung, in der es um die

Teilhabe von Menschen mit Einwanderungsgeschichte in verschiedenen Bereichen des gesellschaftlichen Lebens geht, helfen die interaktiven Elemente, themenspezifisch und individuell ausgewählte Vergleiche vorzunehmen. So können Fragestellungen untersucht und mögliche Erklärungen explorativ geprüft werden. Einige Nutzungsmöglichkeiten der visualisierten Daten im Dashboard Integration werden anhand einer inhaltlichen Fragestellung beispielhaft dargestellt.

5.1 Anwendungsbeispiel: Vergleich von erster und zweiter Generation im Zeitverlauf

Im Kontext von Migration und Integration werden die erste und die zweite Generation von Menschen mit Einwanderungsgeschichte verglichen, um Bedingungen der Integration in der Einwanderungsgesellschaft zu analysieren. Dabei werden Eingewanderte (erste Generation) und ihre Nachkommen (zweite Generation) in ihrer Teilhabe in verschiedenen gesellschaftlichen Bereichen betrachtet. Ein weiterer Vergleich erfolgt schließlich mit der Bevölkerung ohne Einwanderungsgeschichte, um festgestellte Entwicklungen im Integrationsgeschehen im größeren Zusammenhang zu sehen. Eine einfache

Fragestellung wäre beispielsweise: Wie hat sich die Erwerbstätigkeit von Eingewanderten und ihren Nachkommen entwickelt? Die Erwerbstätigenquote, die im Dashboard Integration unter dem Menüpunkt „Arbeit“ zu finden ist, ist ein möglicher Indikator zur Messung der Erwerbstätigkeit ([Dashboard Integration Arbeit](#)).

➤ Grafik 6 stellt die Erwerbstätigenquote der Personen im Alter von 15 bis 64 Jahren nach Einwanderungsgeschichte in einem Zeitreihendiagramm dar.⁹ Die Farbgebung der Linien sowie die Auswahl der Symbole im Dashboard Integration ist für wiederkehrende Kategorien wie die der Einwanderungsgeschichte in Dashboard und Integrationsbericht einheitlich festgelegt, was die Übersichtlichkeit erhöht und die Navigation erleichtert.

Hilfreich in der Analyse der visualisierten statistischen Daten ist ferner die interaktive Ein- und Ausblendefunktion, durch die sich Kategorien auswählen und somit individuell vergleichen lassen. In Grafik 6 sind die Kategorien „Eingewanderte“, „Nachkommen“ sowie „Ohne Einwanderungsgeschichte“ ausgewählt. Es zeigt sich, dass die Erwerbstätigkeit von Eingewanderten und deren Nachkommen seit 2005 zunimmt. Dabei steigt sie

⁹ Zusätzlich stehen in diesem Fall die zusammenfassenden Kategorien „Personen mit Einwanderungsgeschichte“ und „Insgesamt“ zur Verfügung. Die erstgenannte betrachtet Eingewanderte und Nachkommen gemeinsam, die letztgenannte die Gesamtbevölkerung. Erwerbstätigkeit wird nach der Definition der Internationalen Arbeitsorganisation (ILO) erfasst. Erwerbstätige sind Personen ab 15 Jahren, die gegen Entgelt oder als Selbstständige arbeiten oder unbezahlt in einem Familienbetrieb mithelfen, unabhängig vom zeitlichen Umfang der Tätigkeit. Wer sich in einem Beschäftigungsverhältnis befindet, aber im Berichtszeitraum vorübergehend nicht gearbeitet hat, gilt ebenfalls als erwerbstätig.

bei den Nachkommen stärker als bei den Eingewanderten selbst. Anfangs bestehende Unterschiede zwischen beiden Bevölkerungsgruppen werden bis zum Jahr 2023 ausgeglichen. Unterschiede zur Bevölkerung ohne Einwanderungsgeschichte bleiben jedoch bestehen.

5.2 Anwendungsbeispiel: differenzierte Analyse nach Geschlecht

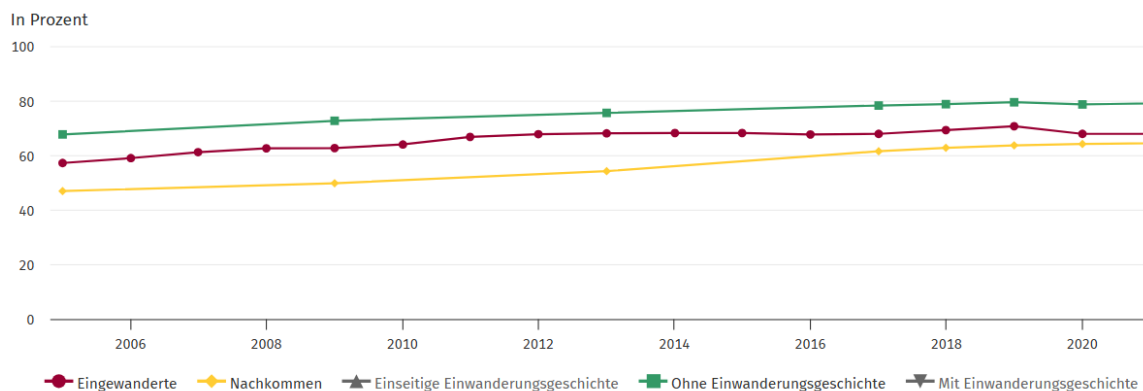
Das Dashboard Integration differenziert die dargestellten Indikatoren neben der Einwanderungsgeschichte zusätzlich nach ausgewählten weiteren relevanten Merkmalen. Das Geschlecht ist dabei ein zentrales Differenzierungsmerkmal. Differenzierungen nach dem Geschlecht liegen in der Regel für jeden Indikator visualisiert vor und erlauben so intersektionale Betrachtungen. Andere häufig verwendete Differenzierungsmerkmale sind Alter oder Aufenthaltsdauer.

Unterschiede in der Erwerbstätigkeit zwischen Eingewanderten und Personen ohne Einwanderungsgeschichte können zu einem großen Teil auf Geschlechterdifferenzen zurückgeführt werden. ➤ Grafik 7 stellt die Erwerbstätigenquoten nach Geschlecht dar. Die zugehörigen Zahlenwerte können aus dem Tooltip abgelesen werden ([Dashboard Integration Arbeit](#)).

Es fällt auf, dass die Erwerbstätigenquote eingewanderter Frauen im Jahr 2023 mit ungefähr 61 % besonders niedrig ist. Der entsprechende Wert liegt für Frauen ohne Einwanderungsgeschichte rund 18 Prozentpunkte höher bei 79 %. Die Unterschiede bei den Männern sind deut-

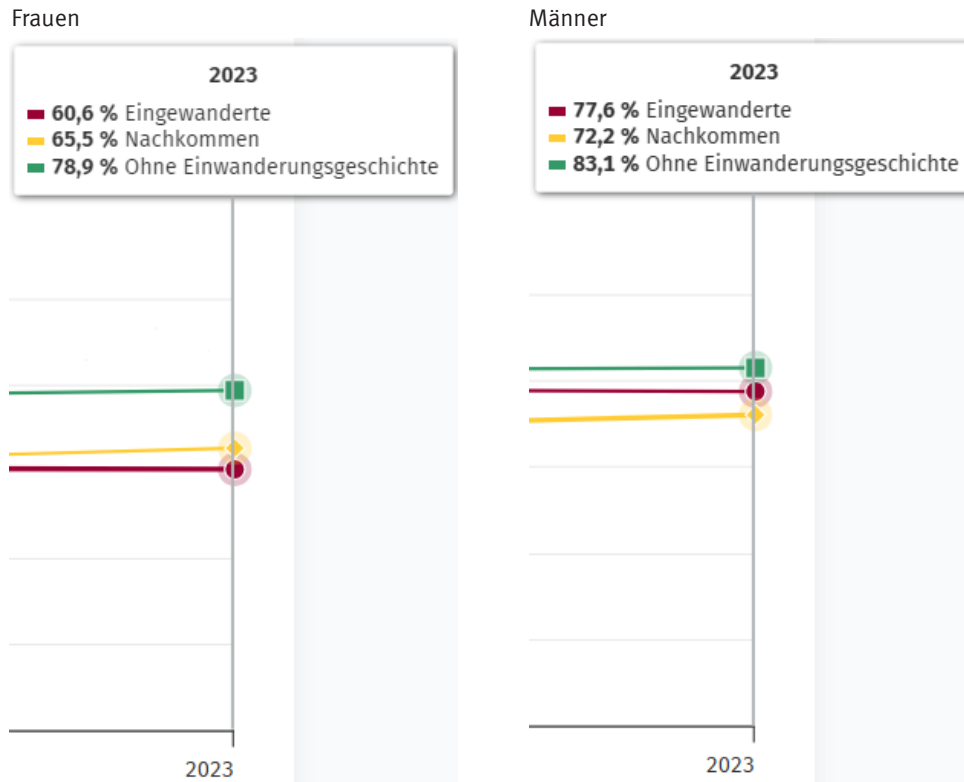
Grafik 6

Erwerbstätigenquote nach Einwanderungsgeschichte



Grafik 7

Erwerbstätigenquote nach Einwanderungsgeschichte und Geschlecht



lich geringer (rund 5 Prozentpunkte). Die Erklärungsansätze für das niedrige Niveau der Erwerbstätigkeit eingewanderter Frauen sind dabei vielfältig und reichen etwa von Hürden beim Arbeitsmarktzugang bis zu fehlender Kinderbetreuung (Grigoriev/Körner, 2024; Keller/Körner, 2024).

Der Geschlechterunterschied ist bei den Nachkommen der Eingewanderten auf den ersten Blick deutlich weniger stark ausgeprägt als bei den Eingewanderten selbst. Allerdings wird die Erwerbstätigkeit der Nachkommen stark von der Altersstruktur geprägt.

Denn im Durchschnitt sind die Nachkommen Eingewanderter noch sehr jung. Ihr Altersdurchschnitt steigt im Zeitverlauf¹⁰ zwar an, aber sie sind weiterhin häufiger als die Eingewanderten selbst und Personen ohne Einwanderungsgeschichte in einem Alter, in dem sie dem Arbeitsmarkt seltener zur Verfügung stehen, weil

sie sich oft noch in Bildung oder Ausbildung befinden ([Dashboard Integration Demografie](#)). [↗ Grafik 8](#) Während rund 31 % der Nachkommen von Eingewanderten zwischen 15 und 29 Jahren alt sind, liegen die entsprechenden Werte für Eingewanderte selbst bei ungefähr 18 % und für Personen ohne Einwanderungsgeschichte bei rund 14 %.

Bei Personen in der Haupterwerbsphase im Alter von 30 bis 59 Jahren reduziert sich der Unterschied in der Erwerbstätigkeit zwischen den Nachkommen von Eingewanderten und Personen ohne Einwanderungsgeschichte deutlich. Dies geht allerdings im Wesentlichen auf die Männer zurück. Unter den Männern liegt in dieser Altersklasse der Unterschied in der Erwerbstätigenquote zwischen den betrachteten Bevölkerungsgruppen nur noch bei rund 3 Prozentpunkten. Unter den Frauen bleibt er hingegen weiter bestehen (rund 11 Prozentpunkte), da nicht nur die Erwerbstätigenquote der Nachkommen in dieser Altersklasse stark steigt, sondern auch die der Personen ohne Einwanderungsgeschichte ([Dashboard Integration Arbeit](#)).

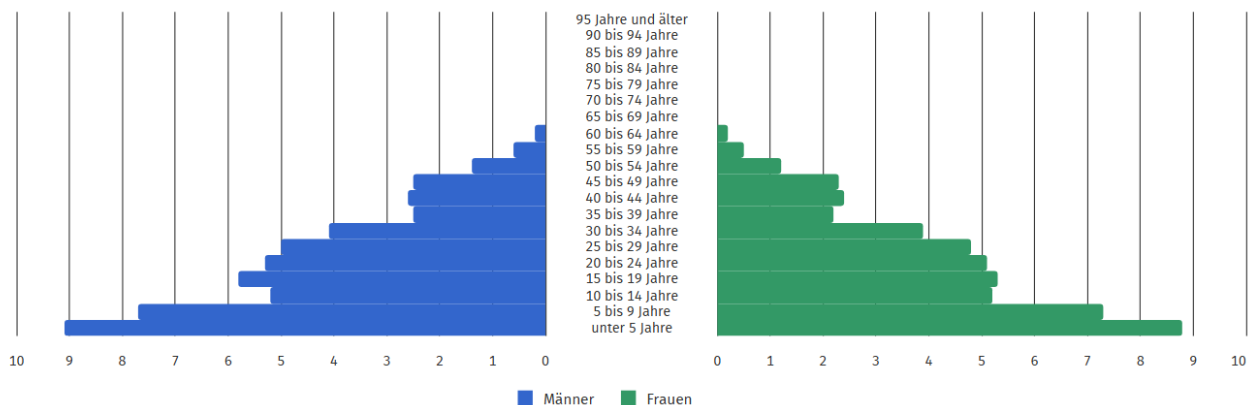
¹⁰ Die Abbildung zum Durchschnittsalter im Zeitverlauf ist im Dashboard Integration online abrufbar unter: [Dashboard Integration Demografie](#).

Grafik 8

Alterspyramiden nach Einwanderungsgeschichte und Geschlecht

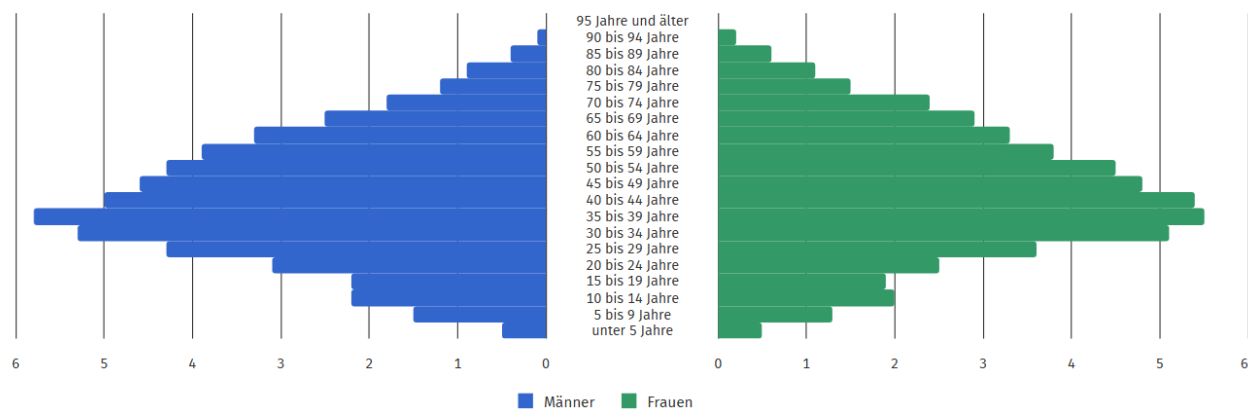
Nachkommen

☰
Mehr



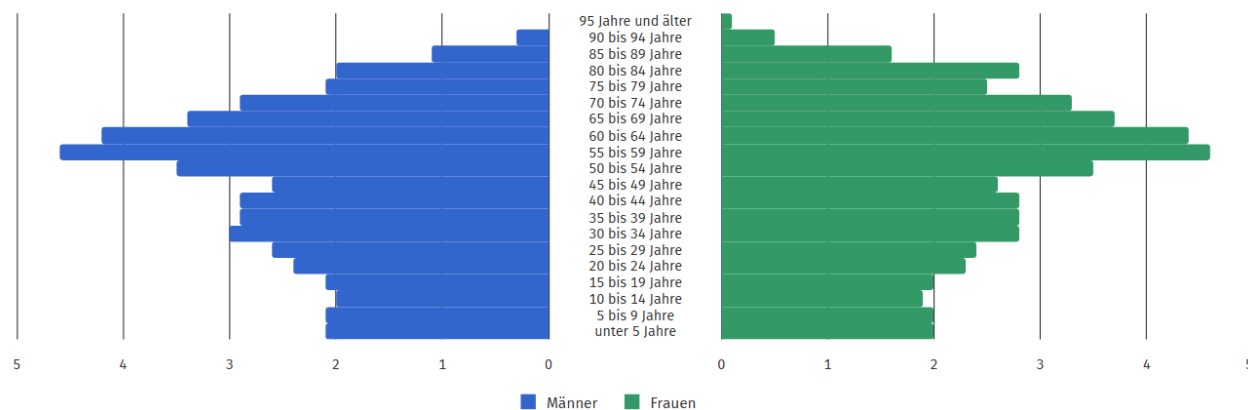
Eingewanderte

☰
Mehr




Ohne Einwanderungsgeschichte

☰
Mehr



Diese Beispiele stehen für die vielfältigen Auswertungsmöglichkeiten, die das Dashboard Integration bietet. Zusammenfassend zeigt sich, dass das Dashboard Integration verschiedene Nutzungsmöglichkeiten visualisierter statistischer Daten zum Thema Migration und Integration eröffnet. Verschiedene Darstellungstypen machen die statistischen Informationen in den Grafiken leicht zugänglich. Interaktive Grafikelemente erlauben es, die Nutzung der visualisierten Informationen stärker an den eigenen Bedürfnissen auszurichten und die den Grafiken zugrunde gelegten Daten herunterzuladen. Darüber hinaus helfen Metadaten bei der Interpretation statistischer Ergebnisse.

zelne Berichtsjahre einer Zeitreihe flexibel auswählbar. Die technische Umsetzung solcher Anpassungen, von denen alle Inhalte des Dashboard Deutschland profitieren, wird fortlaufend geprüft.

Alles in allem steht das neue Dashboard Integration für eine Informationsplattform mit einem breiten Spektrum an inhaltlichen und technischen Möglichkeiten, die statistische Ergebnisse zum Thema Migration und Integration leicht zugänglich macht. 

6

Ausblick

Das Dashboard Integration bietet eine umfassende Informationsplattform zum Integrationsgeschehen in Deutschland anhand von rund 500 Grafiken, die auf dem 14. Integrationsbericht der Integrationsbeauftragten basieren. Die Daten werden fortlaufend aktualisiert.

Interaktive Navigationselemente bieten vielfältige Nutzungsmöglichkeiten der visualisierten statistischen Daten. Eine Download-Funktion ermöglicht es zudem, die Inhalte individuell weiterzuverarbeiten. Die Virtual Cubes auf der Datenplattform sowie das neue Content Management System schaffen zudem eine nachhaltige IT-Struktur mit der Möglichkeit, Inhalte im Dashboard Integration ohne tiefergehende IT-Kenntnisse zu erstellen und zu verwalten. Die vorhandene IT-Struktur wird auch genutzt, um den im Jahr 2026 erscheinenden 15. Integrationsbericht zu erstellen und darauf aufbauend die Daten im Dashboard Integration zu erweitern.

Gleichzeitig bestehen Optimierungspotenziale an den vorhandenen IT-Strukturen. So ist das Dashboard Integration derzeit nicht besonders gut geeignet, Strukturdaten mit langen Zeitreihen anzuzeigen, da die zugrunde liegende Technologie aus dem eher hochfrequenten Dashboard Konjunktur stammt. Auch könnte die Bedienfreundlichkeit auf den verschiedenen Endgeräten weiter verbessert werden. Sinnvoll wäre es außerdem, die Grafikerstellung zu flexibilisieren. So ist beispielsweise die manuelle Einstellung der Skalenrange bei Bevölkerungspyramiden nicht möglich, ebenso wenig sind ein-

LITERATURVERZEICHNIS

Bach, Benjamin/Freeman, Euan/Abdul-Rahman, Alfie/Turkay, Cagatay/Khan, Saiful/Fan, Yulei/Chen, Min. *Dashboard Design Patterns*. In: IEEE Transactions on Visualization and Computer Graphics. Band 29. Ausgabe 1/2023, Seite 342 ff. DOI: [10.1109/TVCG.2022.3209448](https://doi.org/10.1109/TVCG.2022.3209448)

Boiko, Bob. *Content Management Bible, Second Edition*. 2005.

Bray, Tim. *The JavaScript Object Notation (JSON) Data Interchange Format*. 2017. DOI: [10.17487/RFC8259](https://doi.org/10.17487/RFC8259)

Canan, Coşkun/Petschel, Anja. *Die Umsetzung des Konzepts „Einwanderungsschichte“ im Mikrozensus 2022*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2023, Seite 61 ff.

Shahzad, Farrukh/Sheltami, Tarek R./Shakshuki, Elhadi M./Shaikh, Omar. *A Review of Latest Web Tools and Libraries for State-of-the-art Visualization*. In: Procedia Computer Science. Ausgabe 98/2016, Seite 100 ff. DOI: [10.1016/j.procs.2016.09.017](https://doi.org/10.1016/j.procs.2016.09.017)

Grigoriev, Olga/Körner, Thomas. *Eingewanderte und ihre Nachkommen*. In: Sozialbericht 2024. Ein Datenreport für Deutschland. 2024, Seite 30 ff. [Zugriff am 11. März 2025]. Verfügbar unter: www.bpb.de

Integrationsbeauftragte. *Lagebericht: Rassismus in Deutschland. Ausgangslage, Handlungsfelder, Maßnahmen*. 2023. [Zugriff am 11. März 2025]. Verfügbar unter: www.integrationsbeauftragte.de

Integrationsbeauftragte. *Lagebericht: Teilhabe in der Einwanderungsgesellschaft. Ein wissenschaftsbasierter und indikatorengestützter Lagebericht zum Stand der Integration in Deutschland*. 2024. [Zugriff am 4. März 2025]. Verfügbar unter: www.integrationsbeauftragte.de

Keller, Matthias/Körner, Thomas. *Haushalte, Familien und Lebensformen mit Einwanderungsgeschichte – Auswirkungen auf die Erwerbsbeteiligung von Eltern*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2024, Seite 39 ff.

Pezoa, Felipe/Reutter, Juan L./Suarez, Fernando/Ugarte, Martín /Vrgoc, Domagoj. *Foundations of JSON Schema*. In: Proceedings of the 25th International Conference on World Wide Web. 2016, Seite 263 ff. DOI: [10.1145/2872427.2883029](https://doi.org/10.1145/2872427.2883029)

Sarikaya, Alper/Correll, Michael/Bartram, Lyn/Tory, Melanie/Fisher, Danyel. *What Do We Talk About When We Talk About Dashboards?* In: IEEE Transactions on Visualization and Computer Graphics. Band 25. Ausgabe 1/2019, Seite 682 ff. DOI: [10.1109/TVCG.2018.2864903](https://doi.org/10.1109/TVCG.2018.2864903)

Stern, Markus/Gerhard, Marcel/Czerniak, Frederik. *Konjunkturbeobachtung mithilfe innovativer Veröffentlichungsformate: das Dashboard Deutschland und der Pulsmesser für die Wirtschaft*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2024, Seite 15 ff.

RECHTSGRUNDLAGEN

Gesetz über den Aufenthalt, die Erwerbstätigkeit und die Integration von Ausländern im Bundesgebiet (Aufenthaltsgesetz – AufenthG) in der Fassung der Bekanntmachung vom 25. Februar 2008 (BGBl. I Seite 162), das zuletzt durch Artikel 3 des Gesetzes vom 25. Oktober 2024 (BGBl. I Nr. 332) geändert worden ist.



Elena Stäger

hat Statistik an der Technischen Universität Dortmund studiert und arbeitet seit 2021 bei IT.NRW – Statistisches Landesamt Nordrhein-Westfalen im Service „Mathematisch-statistische Methoden und experimentelle Statistik“. Für ihre Masterarbeit „Neuronale Netze in der Baustatistik: Automatisiertes Erkennen von Baustellen anhand von Luftbildern“ wurde sie im Jahr 2022 mit dem Wissenschaftlichen Nachwuchspreis „Statistical Science for the Society“ des Statistischen Bundesamtes in der Kategorie „Master/Bachelorarbeiten“ geehrt. Sie stellt diese von Prof. Dr. Markus Pauly (Technische Universität Dortmund) und Dr. Johannes Rohde (IT.NRW) betreute Arbeit in diesem Beitrag vor.

NEURONALE NETZE IN DER BAU-STATISTIK: AUTOMATISIERTES ERKENNEN VON BAUSTELLEN ANHAND VON LUFTBILDERN

Elena Stäger

🔗 **Schlüsselwörter:** Bilderkennung – Maschinelles Lernen – Orthophotos – räumliche Strukturen – Testregionen

ZUSAMMENFASSUNG

Können Methoden der Künstlichen Intelligenz helfen, Baustellenaktivitäten in Nordrhein-Westfalen zu erfassen und somit mehr Informationen über das dortige Bauge-schehen zu erhalten? Der Artikel untersucht ob es möglich ist, einen Datensatz zu erstellen, um einen Algorithmus zu trainieren, der Baustellen in unbekannten Bildern erkennen und die Vollständigkeit der gemeldeten Bauvorhaben überprüfen kann. Nach einer kurzen Einordnung der Baustatistik beschreibt der Beitrag die Erfassung und Verarbeitung von Bilddaten sowie die Implementierung von Convolutional Neural Networks zur Bildanalyse.

🔗 **Keywords:** image recognition – machine learning – orthophotos – spatial structures – test regions

ABSTRACT

Can artificial intelligence methods be used to capture construction site activities in Nordrhein-Westfalen (NRW) and help gather more information on building activities in the state? This article explores whether it is possible to create a dataset to train an algorithm that can identify construction sites in unknown images and check whether all construction projects have been reported. After a brief explanation of construction statistics in NRW, the article goes on to describe the steps involved in capturing and processing image data and the implementation of convolutional neural networks for image analysis.

1

Einleitung

In Deutschland gibt es in jedem der 16 Bundesländer eine Landesbauordnung; diese basieren auf der Musterbauordnung¹, unterscheiden sich jedoch in Details. In Nordrhein-Westfalen müssen nach der Bauordnung für das Land Nordrhein-Westfalen alle Bauvorhaben an Gebäuden gemeldet werden, bestimmte Kenngrößen daraus werden in der Baugenehmigungsstatistik veröffentlicht. Dies ist wichtig, um Aussagen über die aktuelle Wirtschaftsentwicklung zu treffen, die in der Bautätigkeitsstatistik sichtbar werden, und dient beispielsweise zur Entscheidungsfindung in der Städteplanung. Es ist allerdings unklar, ob wirklich alle Bauvorhaben gemeldet werden.

Ziel der diesem Beitrag zugrunde liegenden Masterarbeit war herauszufinden, ob Methoden der Künstlichen Intelligenz angewendet werden können, um mehr Informationen über das Baugeschehen in Nordrhein-Westfalen zu erhalten. Satellitenbilder sind flächendeckend und in regelmäßigen Zeitabständen für die ganze Welt verfügbar. Diese Bilder haben allerdings eine geringe Auflösung. Für Nordrhein-Westfalen liefert die Bezirksregierung Köln von Flugzeugen aus aufgenommene Luftbilder mit einer höheren Auflösung, die in geringeren Zeitabständen als die Satellitenbilder verfügbar sind. Nun soll mithilfe dieser Daten ein Datensatz erstellt werden, der benutzt wird, um einen Algorithmus zu trainieren. Dieser soll an unbekannten Bildern erkennen können, wo sich Baustellen befinden. Diese Erkenntnisse können mit den gemeldeten Bauvorhaben verglichen werden.

Zur Bildverarbeitung stehen unterschiedliche Algorithmen zur Verfügung, die in der Literatur etabliert sind. Aufgrund ihrer Architektur, die räumliche Strukturen berücksichtigt, haben Neuronale Netze in diesem Zusammenhang besondere Bedeutung. Insbesondere zur Lokalisierung von Gebäuden und zur Schadenserken- nung an ihnen haben in den vergangenen Jahren Con-

volutional Neural Networks (Castelluccio und andere, 2015; Gardner/Nichols, 2017) Anwendung gefunden und vielversprechende Ergebnisse erzielt.

Nach einer kurzen Einführung in die Baustatistik in Kapitel 2 erläutert Kapitel 3, welche Daten verfügbar sind. Kapitel 4 enthält Beschreibungen der verwendeten Methoden, während sich Kapitel 5 der Analyse der erstellten Datensätze widmet. Diese Ergebnisse werden in Kapitel 6 diskutiert und abschließend wird ein Ausblick auf weitere Entwicklungsmöglichkeiten gegeben.

2

Hintergrund Baustatistik

Die Statistiken der Baugenehmigungen und -fertigstellungen sowie der Bauabgänge werden erhoben, um Einblick in die Entwicklung der Bautätigkeit zu bekommen. Aus ihren Ergebnissen wird berechnet, wie sich der Wohnungsbestand entwickelt; diese Daten stehen für Wirtschaft, Forschung und Städteplanung bereit. Zudem dienen diese Erhebungen als Indikator, um die Wirtschaftsentwicklung im Bausektor zu bewerten, und damit der kurzfristigen Konjunkturbeobachtung in Deutschland. Die Statistik der Baugenehmigungen ist eine monatliche Gesamterhebung der Statistischen Ämter des Bundes und der Länder, die mithilfe der Bauaufsichtsbehörden und der Bauherinnen sowie Bauherren durchgeführt wird. Die Statistiken der Baufertigstellungen und der Bauabgänge haben eine jährliche Periodizität.

Rechtsgrundlagen für die Erhebungen sind das Hochbaustatistikgesetz (HBauStatG) in Verbindung mit dem Bundesstatistikgesetz (BStatG).

Es besteht Auskunftspflicht (§ 6 Absatz 1 Satz 1 HBauStatG), deren Einhaltung den Bauherinnen und Bauherren sowie den Bauaufsichtsbehörden obliegt (§ 6 Absatz 2 HBauStatG). Zwar sind die Einzelangaben geheim zu halten (§ 16 BStatG), jedoch dürfen faktisch anonymisierte Daten an Hochschulen und sonstige Einrichtungen mit der Aufgabe unabhängiger wissenschaftlicher Forschung weitergegeben werden (§ 16 Absatz 6 BStatG).

¹ Die von der Bauministerkonferenz verabschiedete Musterbauordnung dient als Grundlage für die Umsetzung in spezifisches Landesrecht. Sie entfaltet somit keine unmittelbare Rechtswirkung. Jedes Land entscheidet, in welchem Umfang die Landesregelung dem Muster folgt. Die aktuelle, zuletzt durch Beschluss der Bauministerkonferenz vom 23./24. November 2023 geänderte Musterbauordnung steht auf deren [Website](#) zur Verfügung.

3

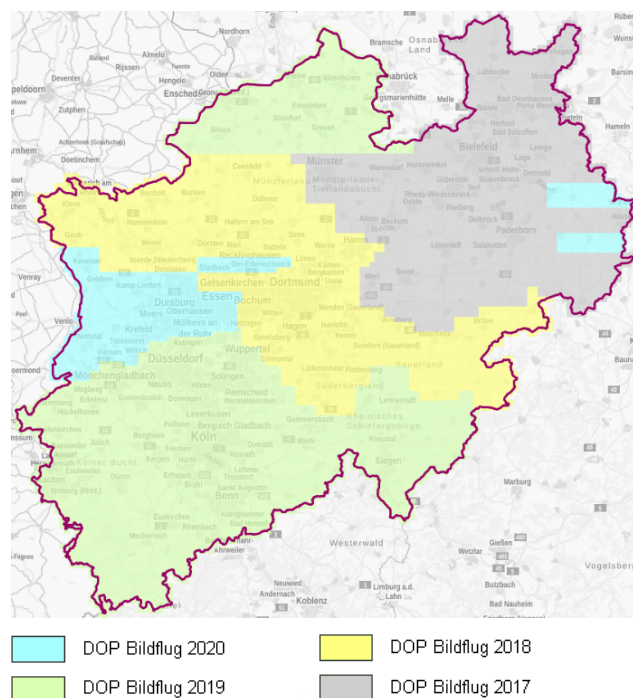
Daten

Die verwendeten Daten zur Analyse sind digitale Orthophotos, die die Bezirksregierung Köln öffentlich zur Verfügung stellt. Die Orthophotos dienen als Planungsgrundlage im Straßenbau, im Umweltschutz und in der Land- und Forstwirtschaft. Zusätzlich können sie bei Schadensfällen, zum Beispiel bei Naturkatastrophen, herangezogen werden.

„Orthophotos sind hochauflösende, verzerrungsfreie, maßstabsgetreue, photographische Abbildungen der Erdoberfläche.“ (Bezirksregierung Köln, 2020a) Die Grundlage bilden Luftbilder, die von Luftfahrzeugen senkrecht zum Boden aufgenommen werden. Diese Bilder werden anschließend digital mithilfe eines Höhenmodells verarbeitet, um die digitalen Orthophotos zu produzieren. Diese Bilder werden für Nordrhein-Westfalen hergestellt und zusätzlich georeferenziert, das heißt die Bilder werden mit den richtigen Koordi-

Grafik 1

Bildflugprogramm Nordrhein-Westfalen nach Jahren



naten versehen. Die Bodenauflösung beträgt zehn Zentimeter je Pixel und die Bilder werden mit vier Kanälen aufgenommen: dem Rot-, dem Grün- und dem Blau-Kanal für die Farbdarstellung sowie in Infrarot. In der diesem Beitrag zugrunde liegenden Arbeit werden nur die drei Farbkanäle verwendet.

Insgesamt wird so das Land Nordrhein-Westfalen alle drei Jahre fotografiert, in jedem Jahr ungefähr ein Drittel.² Welche Gebiete in welchen Jahren überflogen und veröffentlicht wurden, zeigt [Grafik 1](#) zum Stand 10. November 2020. In den folgenden Monaten wurden weitere Daten für 2020 ergänzt, nachdem sie fertig verarbeitet und der Öffentlichkeit zur Verfügung gestellt wurden.

4

Methoden

4.1 Neuronale Netze

(Künstliche) Neuronale Netze basieren auf dem Konzept der synaptischen Verbindungen im Nervensystem. Ein Neuronales Netz besteht aus sogenannten Knoten und Gewichten, die zusammen Neuronen simulieren. Ein solches Netz wird in Schichten organisiert. Es gibt eine Eingangsschicht (Input Layer), eine Ausgangsschicht (Output Layer) und Zwischenschichten (Hidden Layers). Jede Schicht besteht aus mehreren Knoten. Für zwei aufeinanderfolgende Schichten gilt, dass jeder Knoten der Schicht mit jedem Knoten der folgenden Schicht durch ein Gewicht verbunden ist. Knoten einer Schicht sowie Knoten nicht aufeinanderfolgender Schichten sind nicht miteinander verbunden (Goodfellow und andere, 2016).

4.2 Convolutional Neural Network

Ein Convolutional Neural Network (CNN) ist eine besondere Form des Neuronalen Netzes. Es ist besonders für Daten mit räumlichen Strukturen geeignet. Anders als bei Neuronalen Netzen mit vollständig verbundenen aufeinanderfolgenden Schichten benutzen CNNs

² Im Jahr 2020 wurde auf einen Zwei-Jahres-Rhythmus umgestellt.

ein kleines gleitendes Raster, um lokale Eigenschaften zu extrahieren (Convolution) und dann diese Eigenschaften mit dem sogenannten Pooling zu aggregieren. Damit wird die Zahl der Parameter stark reduziert, es können dennoch effizient räumliche Eigenschaften in unterschiedlichen Regionen des Bildes gefunden werden. Außerdem sind die gleitenden Raster translations-invariant, das heißt Merkmale können unabhängig von ihrer Position im Bild gefunden werden (Goodfellow und andere, 2016).

4.3 Vortrainierte CNNs

Es hat sich etabliert, für die Klassifikation von Bilddaten auf CNNs zuzugreifen, die schon auf anderen Datensätzen trainiert und optimiert wurden (Castelluccio und andere, 2015; Gardner/Nichols, 2017; Papadomanolaki und andere, 2016; Perez und andere, 2019).

Beispielsweise nutzten Gardner und Nichols (2017) ein CNN-Modell zur Klassifizierung von Satellitenbildern des Amazonas, um die Abholzung nachzuverfolgen. Es lagen 40 000 Bilder aufgeteilt in 17 Klassen vor. Drei vortrainierte CNNs wurden als Basis benutzt: VGG16, Inception-v3 und ResNet50. ResNet50 erbrachte die besten Ergebnisse mit einer Genauigkeit von 90,8%. Inception-v3 und VGG-16 erreichten eine Accuracy von 90,1% beziehungsweise 89,7%.

Aufgrund der guten Ergebnisse mit Daten basierend auf Satellitenbildern wurde entschieden, die CNNs VGG, ResNet50 und Inception für die Analyse zu benutzen. Alle drei liefern gute Ergebnisse in ähnlichen Anwendungsgebieten und sind in Python Tensorflow verfügbar.

4.4 Stand der Forschung

Convolutional Neural Networks (CNNs) wurden bei der Klassifizierung von Satelliten- und Luftbildern von Regionen nach Naturkatastrophen, zum Beispiel Erdbeben, benutzt, um beschädigte Gebäude zu erkennen und Hilfsaktionen lenken zu können (Duarte und andere, 2018; Ji und andere, 2018). In Duarte und andere (2018) konnte basierend auf Daten aus Italien, Ecuador und Haiti schon eine Genauigkeit von über 90% erreicht werden. Insbesondere führte hier die Kombination von Satelliten- und Luftbildern mit unterschiedlichen Auf-

lösungen zu verbesserten Ergebnissen. Auch in Ji und andere (2018) wurden Luftbilder von Haiti verwendet, und es konnte mit SqueezeNet eine Genauigkeit von 80% erreicht werden.

Außerdem wurden CNNs verwendet, um Satellitenbilder pixelweise verschiedenen Klassen wie Vegetation, Gebäuden und Straßen zuzuordnen (Albert und andere, 2017; Maltezos und andere, 2017; Postadjiana und andere, 2017).

In Maltezos und andere (2017) wurde die pixelweise Klassifikation von CNNs auch mit zwei Arten von Support Vector Machines (SVMs) verglichen. Die SVMs konnten 73% beziehungsweise 76% der Pixel richtig vorhersagen und das CNN 82%.

Des Weiteren wurden CNNs verwendet, um Straßenkreuzungen in Satellitenbildern zu erkennen (Barsia/Heipke, 2003) und um verschiedene Arten von Bautätigkeiten auf Baustellen zu klassifizieren (Luo und andere, 2018).

Aufgrund der vielversprechenden Ergebnisse mit Convolutional Neural Networks, insbesondere mit Daten, die auf Luftbildern basieren und sich mit Gebäuden beschäftigen, wurde beschlossen, für die Entwicklung des Algorithmus zur Erkennung von Luftbildern CNNs zu verwenden.

5

Analyse

5.1 Datensatz

Die diesem Beitrag zugrunde liegende Arbeit ist ein erster Versuch, Baustellen anhand von Luftbildern zu erkennen. Daher lag noch kein Datensatz mit klassifizierten Bildern vor und Baustellen mussten „per Hand“ gefunden werden. Das Programm Arc-GIS wurde verwendet, um die aktuellen Luftbilddaten der Bezirksregierung Köln zu laden und somit Baustellen ausfindig zu machen. Bei der Durchsicht der Daten wurden 2 912 Baustellen gefunden und deren Koordinaten gemeldet. Es wurde darauf geachtet, möglichst alle Regionen Nordrhein-Westfalens und alle Jahre von 2015 bis 2020 abzubilden sowie möglichst viele unterschiedliche Arten

Grafik 2

Beispielbilder mit Baustellen



von Baustellen (zum Beispiel einzelne Wohnhäuser oder gesamte Neubaugebiete) und die verschiedenen Stadien des Bauvorhabens.

Von den 2912 Baustellen wurden etwa 800 auf Grundlage der Baugenehmigungsdaten der Jahre 2015 und 2016 gefunden.

Gut 2000 Areale wurden von der Autorin aufgrund der Informationen, die bestätigten Baustellenbildern entnommen wurden, als wahrscheinliche Baustellen eingestuft.

Da die Luftbilder bis 2020 nur alle drei Jahre erstellt wurden, ist der Anteil der Gebäude in der Baugenehmigungsstatistik, die sich genau am Aufnahmetag in einem Zustand befinden, der sie eindeutig als Baustelle ausweist, eher gering. Weil die zur Verfügung gestellten Adressen aus den Jahren 2015 und 2016 stammen, waren viele dieser Baumaßnahmen vermutlich bereits abgeschlossen, als die Luftbilder in den Jahren 2018 bis 2020 erstellt wurden. Es ist daher zu vermuten, dass die Baustellen, die unabhängig davon gefunden wurden,

in den Genehmigungen der folgenden Jahre auftauchen werden. Zudem ist anzumerken, dass es teilweise schwierig war, Baustellen und ländliche Betriebe auseinanderzuhalten. Baustellen weisen häufig Kräne, Bagger, herumliegende Baumaterialien und einen sandigen Boden auf. Diese Merkmale können leicht mit Traktoren und landwirtschaftlichen Materialien sowie ähnlich sandigem Boden verwechselt werden. Es sollte daher darauf geachtet werden, wie gut der trainierte Algorithmus mit diesen Ähnlichkeiten umgeht, da diese für Menschen schwer zu beurteilen sind.

Die Koordinaten für die Bilder ohne Baustellen wurden zufällig generiert, sodass das Bundesland Nordrhein-Westfalen gleichmäßig abdeckt ist. Insgesamt gibt es 6880 Bilder ohne Baustellen. Mit den 2912 Bildern, bei denen Baustellen zu erkennen sind, besteht der Datensatz für das Training der CNNs aus 9792 Bildern – 30% mit Baustellen und 70% ohne Baustellen. ➤ Grafik 2 zeigt drei Bilder mit Baustellen in unterschiedlichen Stadien des Bauvorgangs, ➤ Grafik 3 drei Bilder ohne Baustellen, nämlich Wald, Häuser in Städten und Straßen.

Grafik 3

Beispielbilder ohne Baustellen



5.2 Training von vortrainierten CNNs

Im Folgenden werden alle Berechnungen mit Python 3.7.7 (van Rossum/Drake, 2009) durchgeführt. Zusätzlich wurde das Paket TensorFlow 2.3.0 verwendet.

Aufgrund der guten Ergebnisse, die vortrainierte CNNs auf ähnlichen Datensätzen erreichen konnten, wurde entschieden, dies auch für diese Anwendung zu verwenden. Ausgesucht wurden die drei folgenden: VGG-16, Inception-v3 und ResNet50, da diese auch in dem Python-Paket Tensorflow implementiert sind. Diese drei Modelle wurden komplett übernommen einschließlich der trainierten Gewichte. Zusätzlich wurden zwei Fully-Connected-Schichten mit 4096 Neuronen und ReLu-Aktivierungsfunktion angehängt sowie eine Schicht mit einer Softmax-Funktion für die binäre Klassifizierung. Außerdem wurde zwischen diesen Schichten jeweils eine Dropout-Schicht mit einem Dropout von 50 % hinzugefügt. Diese Architektur erwies sich nach ersten Versuchen als am besten geeignet. Beim Trainieren der verschiedenen Modelle wurden die Parameter der vortrainierten CNNs fest angenommen und nur die Gewichte der später hinzugefügten Layer wurden berechnet.

Für alle Modelle werden die Standardeinstellungen von TensorFlow übernommen, das heißt die Hyperparameter sind:

- › $\alpha = 0,001$
- › $\beta_1 = 0,9$
- › $\beta_2 = 0,999$
- › $B = 32$

Zusätzlich wurde die Epochenzahl auf 20 und das Early Stopping auf 10 Epochen ohne Verbesserung gesetzt. Der Datensatz wurde aufgeteilt, sodass 80 % zum Trainieren benutzt wurden und die restlichen 20 % zum Validieren. Insgesamt wurde jedes Modell mit dem gleichen Datensatz dreimal berechnet.

Bezogen auf die durchschnittliche Genauigkeit der drei Berechnungen für die Modelle ergab sich, dass VGG und ResNet mit festen Parametern die besten Modelle waren mit Genauigkeiten von 96,46 % beziehungsweise 98,54 %. Inception blieb mit 89,33 % unter diesen Genauigkeiten.

Insgesamt zeigen diese ersten Berechnungen, dass CNNs für die Erkennung von Baustellen ein sinnvolles Verfahren sind. Insbesondere VGG und ResNet mit festen Parametern liefern die besten Ergebnisse mit einer Accuracy von über 95 %, die schon innerhalb der ersten Epochen erreicht werden kann.

Deswegen wurde entschieden, diese beiden Algorithmen weiter zu verwenden, um die Hyperparameter zu optimieren mit dem Ziel, Verbesserungsmöglichkeiten in der Berechnung der Algorithmen zu finden und die Genauigkeit noch weiter zu erhöhen. Mit den Hyperparametern $\alpha = 0,00001$, $\beta_1 = 0,7$, $\beta_2 = 0,999$ und Batchgröße $B = 64$, sowie einer Reduzierung der Epochen auf 10 und Early Stopping auf 5 konnte somit die Accuracy für ResNet auf 99,74 % und für VGG auf 99,48 % verbessert werden.

5.3 Testen auf Testregionen

Um zu prüfen, ob die berechneten Algorithmen für unbekannte Regionen ähnlich gute Ergebnisse liefern, wurde ein neuer Datensatz erstellt. Dazu wurden die Stadt Gelsenkirchen und der Hochsauerlandkreis ausgewählt. Beide haben eine Einwohnerzahl von ungefähr 260 000 (Landesdatenbank NRW, 2020). Der Hochsauerlandkreis als ländliche Gegend sollte mit einer städtischen, hier: Gelsenkirchen, verglichen werden. Es wurden zufällige neue Bilder für die beiden Regionen generiert und von der Autorin jeweils rund 50 Baustellen gesucht, sodass ein Datensatz für Gelsenkirchen mit 986 Nicht-Baustellen und 58 Baustellen entstand sowie ein Datensatz für den Hochsauerlandkreis mit 1 034 Nicht-Baustellen und 50 Baustellen. Diese beiden Datensätze wurden dann jeweils mit dem Algorithmus 5 des VGG und des ResNet des vorherigen Abschnitts klassifiziert.

➤ **Tabelle 1** zeigt die Accuracy des Testdatensatzes für VGG und ResNet, dargestellt für die Stadt Gelsenkirchen und für den Hochsauerlandkreis. ResNet liefert hier für beide Regionen bessere Ergebnisse (93 % beziehungs-

Tabelle 1
Accuracy (Genauigkeit) der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	88,89	97,14
ResNet	92,91	98,15

weise 98 %) als VGG (89 % beziehungsweise 97 %). Die Bilder von Gelsenkirchen werden für beide Algorithmen mit nur 89 % beziehungsweise 93 % Genauigkeit klassifiziert. Der Hochsauerlandkreis hingegen weist eine Genauigkeit von 97 % beziehungsweise 98 % vor. Das beste Ergebnis liegt bei ResNet auf der ländlichen Region vor.

Zu beachten ist hier, dass es jeweils nur rund 50 Bilder mit Baustellen gab und somit der Vergleich der Accuracy zu Bildern mit Baustellen mit um die 1 000 Bildern schwierig ist. Deswegen ist es wichtig, die Sensitivität und die Präzision zu betrachten. [Tabelle 2](#) zeigt die Sensitivität der zwei Datensätze jeweils für das VGG- und das ResNet-Modell. Der Datensatz für Gelsenkirchen weist hier eine geringere Sensitivität (10 beziehungsweise 14 Prozentpunkte) als der Hochsauerlandkreis auf. ResNet liefert für den Hochsauerlandkreis das beste Ergebnis mit 98 %, das heißt 98 % der tatsächlichen Baustellen werden vom Algorithmus auch erkannt. In Gelsenkirchen werden nur 84 % der Baustellen erkannt.

Tabelle 2
Sensitivität der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	82,76	92,00
ResNet	84,48	98,00

[Tabelle 3](#) stellt die Präzision für das VGG- und das ResNet-Modell auf Basis der Daten von Gelsenkirchen und dem Hochsauerlandkreis dar. Für das ResNet-Modell liegt die Präzision für beide Regionen um 10 Prozentpunkte über der des VGG-Modells. Beim VGG liegt die Präzision für Gelsenkirchen bei 31 % und für den

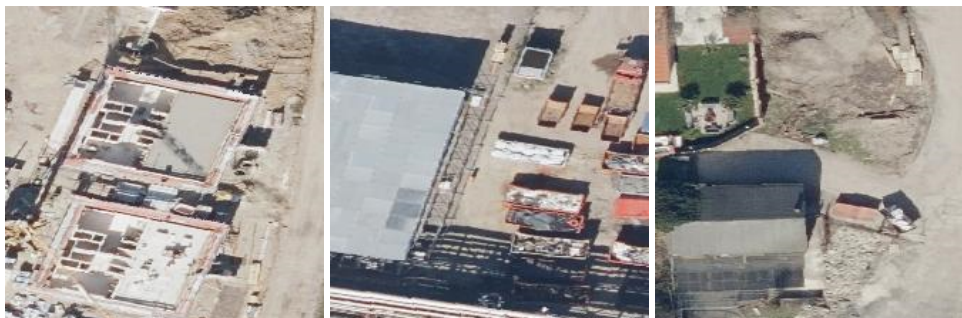
Hochsauerlandkreis bei 63 %. Das heißt in Gelsenkirchen sind nur 31 % der vom Modell als Baustelle klassifizierten Bilder auch wirklich Baustellen. Für ResNet sind dies in Gelsenkirchen 43 %. Für den Hochsauerlandkreis liefert ResNet die größte Präzision mit 72 %. Diese niedrigen Werte sind darauf zurückzuführen, dass in den zwei benutzten Datensätzen 1 000 Bilder ohne Baustellen vorliegen. Selbst wenn diese mit einer Wahrscheinlichkeit von über 90 % richtig klassifiziert werden, bleibt eine hohe Anzahl von Bildern ohne Baustelle, die fälschlicherweise als Bild mit Baustelle klassifiziert werden. Bei dem Verhältnis Bilder von tatsächlichen Baustellen zu Bildern ohne Baustellen von 1 zu 20 führt selbst eine geringe Fehlerquote von 5 % dazu, dass genauso viele Baustellen fälschlicherweise identifiziert werden wie tatsächlich im Datensatz vorhanden sind. Um die Präzision zu erhöhen, könnte der Trainingsdatensatz um mehr Bilder aus städtischen Regionen erweitert werden.

Tabelle 3
Präzision der Algorithmen VGG und ResNet

	Gelsenkirchen	Hochsauerlandkreis
	%	
VGG	31,17	63,01
ResNet	42,98	72,06

[Grafik 4](#) zeigt eine Auswahl von Baustellen in Gelsenkirchen, die von beiden Algorithmen fälschlicherweise als Nicht-Baustelle klassifiziert wurden. Das linke Bild ist für das menschliche Auge recht leicht als Baustelle identifizierbar. Die anderen beiden Bilder könnten auch eine Mülldeponie beziehungsweise Boden ohne Gras darstellen. Da das Urteil des Menschen hier unsicherer als bei anderen Bildern ist, ist es möglich, dass ähnli-

Grafik 4
Fälschlicherweise nicht als Baustellen klassifizierte Baustellen in Gelsenkirchen



Grafik 5

Fälschlicherweise nicht als Baustelle klassifizierte Baustelle im Hochsauerlandkreis



che Bilder seltener im Trainingsdatensatz vorkommen und somit der Algorithmus deren Merkmale schlechter erlernen kann.

Auch die falsch klassifizierte Baustelle aus dem Datensatz für den Hochsauerlandkreis stellt eine Baustelle dar, die für das menschliche Auge schwer als Baustelle zu erkennen ist. Hier hat die Autorin die Umgebung genauer betrachtet, was zur Einstufung als Baustelle führte. ➤ Grafik 5

➤ Grafik 6 zeigt Bilder ohne Baustellen in Gelsenkirchen, die von beiden Algorithmen als Baustellen klassifiziert wurden. Für menschliche Beobachtende ist schnell offensichtlich, dass es sich hier nicht um Baustellen handelt. Eventuell führten Straßenmarkierungen oder die Schatten von Bäumen zu einer Falschklassifizierung. Das rechte Bild zeigt, dass ein Wohngebiet als Baustelle klassifiziert werden kann.

In ➤ Grafik 7 sind drei falsch klassifizierte Nicht-Baustellen des Hochsauerlandkreises abgebildet. Im linken sowie im rechten Bild ist nur Wald beziehungsweise eine Straße zu sehen, wohingegen im mittleren Bild eventuell der graue Boden als Zeichen für eine Baustelle interpretiert wurde.

Grafik 6

Fälschlicherweise als Baustellen klassifizierte Nicht-Baustellen in Gelsenkirchen



Grafik 7

Fälschlicherweise als Baustellen klassifizierte Nicht-Baustellen im Hochsauerlandkreis



6

Diskussion und Ausblick

Das Ziel der diesem Beitrag zugrunde liegenden Arbeit war zu prüfen, ob Convolutional Neural Networks benutzt werden können, um auf Orthophotos Baustellen zu erkennen. Die Ergebnisse der Auswertung verdeutlichen, dass CNNs sehr gut dafür geeignet sind, und es wurden dafür mehrere Algorithmen trainiert und verglichen.

Zu dieser speziellen Anwendung mit Orthophotos in Nordrhein-Westfalen existierte zuvor noch kein Datensatz, daher musste dieser selbstständig erstellt werden. Es konnten zwar die georeferenzierten Daten der Baugenehmigungsstatistik zu Hilfe genommen werden, aber es stand nur der Datensatz für die Jahre 2015 und 2016 zur Verfügung. Die aktuellsten Orthophotos stammen aus den Jahren 2018 bis 2020 und so war es schwierig, aufgrund dieses Datensatzes Bauprojekte zu finden, da die meisten zu diesem Zeitpunkt schon abgeschlossen waren. Deshalb ist mehr als die Hälfte der Bilder nicht mit Eintragungen in der Baugenehmigungsstatistik verbunden. Außerdem gelten rund 40 % der Baugenehmigungen nicht für Neubauten, das heißt viele Bauvorhaben (innerhalb von Gebäuden) sind nicht auf Orthophotos zu sehen, selbst wenn das Foto zum Zeitpunkt des Bauens aufgenommen wird. Zusätzlich ist zu beachten, dass die Aufnahme der Orthophotos in einem Rhythmus von drei beziehungsweise seit 2020 zwei Jahren erfolgt. Dadurch kann ein Großteil der Bauvorhaben, die zwischen diesen Terminen stattfinden, nicht auf den Orthophotos erfasst werden. Hinzu kommt, dass der Datensatz zum Trainieren der CNNs von einer einzigen Person erstellt wurde, wodurch womöglich eine Verzerrung in den Datensatz eingeflossen ist. Beispielsweise wurden vermehrt Baustellen in den Datensatz aufgenommen, die für das menschliche Auge einfach zu erkennen sind, und bei einer Unsicherheit, ob es sich wirklich um eine Baustelle handelt, oblag die Entscheidung der persönlichen Einschätzung der Autorin. Somit sind im Datensatz der Bilder mit Baustellen möglicherweise Aufnahmen vorhanden, die auf den Orthophotos wie eine Baustelle aussehen, aber in Wirklichkeit nicht genehmigungspflichtig sind. Beispiele finden sich bei landwirtschaftlichen Betrieben, Gartenarbeiten oder unaufgeräumten Abstellplätzen. Anzumerken ist außerdem, dass Bauar-

beiten an Straßen oder Windrädern nicht in den Datensatz mit aufgenommen wurden. Es ist aber zu vermuten, dass diese trotzdem als Baustelle klassifiziert werden. Es ist auch möglich, dass deshalb mehr Baustellen aus städtischen Gebieten im Datensatz vorhanden sind, weil diese für eine Person schneller auffindbar sind.

Für die Analyse wurden drei vortrainierte CNNs ausprobiert, die auf ähnlichen Datensätzen vielversprechende Ergebnisse lieferten. Es gibt natürlich noch viele andere CNNs und andere Arten von Neuronalen Netzen, die man hier testen könnte. Zusätzlich wurde die Architektur mit Dropout von 0,5 und zwei Fully-Connected-Schichten nur an einem kleinen Teil des Datensatzes getestet. Hier wäre es theoretisch möglich, noch mehr Zeit zu investieren, um eine bessere Architektur zu finden. Allerdings lieferten die ersten Berechnungen insbesondere mit ResNet Genauigkeiten von über 95 %, sodass selbst mit anderen CNNs und Architekturen der Algorithmen nur geringe Verbesserungen zu erwarten wären.

Für die Optimierung der Hyperparameter wurden die zwei Algorithmen von VGG und ResNet verwendet, die bei den vorherigen Experimenten eine Genauigkeit über 96 % erreichen konnten. Danach wurden die Hyperparameter noch optimiert. Die besten Ergebnisse lieferte die Kombination: $\alpha = 0,00001$, $\beta_1 = 0,7$, $\beta_2 = 0,999$ und Batchgröße $B = 64$ sowohl für VGG und ResNet mit Genauigkeiten über 99 %, wobei auch hier ResNet das bessere Ergebnis lieferte. Insgesamt konnte durch die Optimierung der Hyperparameter eine Verbesserung von 3 % für das VGG-Modell und von 1,2 % für das ResNet-Modell erreicht werden.

Der Testdatensatz mit Bildern aus Gelsenkirchen und dem Hochsauerlandkreis lieferte eine Accuracy von über 85 % für Gelsenkirchen und über 90 % für den Hochsauerlandkreis. Es wurde erwartet, dass der Hochsauerlandkreis schlechter abschneiden würde als Gelsenkirchen, da im Trainingsdatensatz viele Baustellen mit Häusern in Neubaugebieten und in Städten vorhanden sind und vermutet wurde, dass im ländlichen Gebiet mehr landwirtschaftliche Betriebe zu finden sind, bei denen die Klassifikation schwierig sein könnte. Insbesondere die Präzision für die Bilder aus Gelsenkirchen ist sehr gering (VGG: 30 %, ResNet: 40 %). Es könnte eine Verbesserung für Gelsenkirchen erreicht werden, wenn der Datensatz um mehr Bilder aus Städten erweitert werden könnte. Eine andere Möglichkeit wäre, den


Bildausschnitt zu vergrößern, damit dem Algorithmus bei schwer zu klassifizierenden Bildern mehr Informationen zur Umgebung zur Verfügung stehen. Zusätzlich könnte ein vergrößerter Bildausschnitt dazu genutzt werden, den Algorithmus das gesamte Land klassifizieren zu lassen und so flächendeckend Baustellen zu erkennen. Dies wäre bei einem kleinen Bildausschnitt sehr viel aufwendiger. Allerdings ist zu beachten, dass dann der gesamte Datensatz, auf dem diese Arbeit aufbaut, geprüft werden müsste, denn es ist möglich, dass bei Bildern ohne Baustellen dann Baustellen im Ausschnitt liegen. Zusätzlich müssten Bilder von Neubaugebieten eventuell aussortiert werden, falls ein Bild mehrere Baustellen zeigt.

Ein weiteres Problem stellt die Verknüpfung mit der offiziellen Baugenehmigungsstatistik dar. Können durch die Ergebnisse dieser Arbeit nun flächendeckend die Baustellen auf den Orthophotos entdeckt werden, ist dennoch zusätzlich zu prüfen, ob diese Baustellen in der Statistik vorhanden sind. Das Bild gibt keine Information dazu, wann genau die Baugenehmigung erfolgt ist, sodass mindestens die Statistiken der drei vorherigen Jahre geprüft werden müssten. Hinzu kommt, dass die Verarbeitung der Baugenehmigungen einige Zeit beansprucht und teilweise die statistischen Erhebungsbogen nachgereicht werden. Das bedeutet, dass bei aktuellen Bildern einige Zeit abzuwarten ist, bis die dazugehörigen Baugenehmigungen vollständig geprüft werden können.

Darüber hinaus sind 22% der Georeferenzierungen nicht gebäudegenau. Es ist zum Beispiel nicht zu unterscheiden, welche Baustelle genehmigt wurde, wenn für eine Straße mehrere Baustellen entdeckt wurden, aber nur eine offiziell gemeldet ist. Es ist dann auch nicht zu erkennen, ob eine Baustelle gemeldet wurde, die eventuell zum Zeitpunkt der Bildaufnahme gar nicht sichtbar war. Zudem wären solche Genehmigungen zuerst auszusortieren, bei denen es sich nicht um eine Neubebauung handelt und die Bautätigkeit innerhalb des Gebäudes stattfinden könnte – was nicht auf Orthophotos sichtbar sein kann.

Es wäre möglich, den Datensatz mit Satellitenbildern zu kombinieren. In Ji und andere (2018) führte die Kombination von Luft- und Satellitenbildern zu verbesserten Ergebnissen. Die Satellitenbilder liefern zwar eine schlechtere Auflösung, sind aber in kleineren zeitlichen Abständen verfügbar, sodass Baustellen in mehreren Stadien des Bauvorgangs erkennbar sein müssten.

Insgesamt sind die vielversprechendsten Vorschläge, den Trainingsdatensatz um Baustellen aus städtischen Regionen zu erweitern und für eine flächendeckende Analyse des Landes den Bildausschnitt zu vergrößern. Soll der Algorithmus auf anderen Regionen oder Ländern als Nordrhein-Westfalen getestet werden, wäre es sinnvoll, den Datensatz um einen kleinen Datensatz mit Bildern aus der jeweiligen Region zu erweitern. In Xu und andere (2019) konnte damit die Accuracy für neue Regionen um ein paar Prozentpunkte gesteigert werden.

Insgesamt hat die Autorin in ihrer Masterarbeit festgestellt, dass CNNs sehr gut geeignet sind, um Baustellen in Orthophotos zu erkennen. 

LITERATURVERZEICHNIS

Albert, Adrian/Kaur, Jasleen/Gonzalez, Marta. *Using convolutional networks and satellite imagery to identify patterns in urban environments at a large scale*. 2017. [Zugriff am 4. Februar 2025]. Verfügbar unter: arxiv.org

Barsia, Arpad/Heipke, Christian. *Artificial neural networks for the detection of road junctions in aerial images*. ISPRS Archives. 2003. [Zugriff am 5. Februar 2025]. Verfügbar unter: www.pf.bgu.tum.de

Bezirksregierung Köln. *Digitale Orthophotos*. 2020a. [Zugriff am 20. Februar 2025]. Verfügbar unter: www.bezreg-koeln.nrw.de

Bezirksregierung Köln. *TIM-online*. 2020b. [Zugriff am 20. Februar 2025]. Verfügbar unter: www.bezreg-koeln.nrw.de/geobas

Castelluccio, Marco/Poggi, Giovanni/Sansone, Carlo/Verdoliva, Luisa. *Land Use Classification in Remote Sensing Images by Convolutional Neural Networks*. 2015. [Zugriff am 24. Januar 2025]. Verfügbar unter: arxiv.org

Duarte, Diogo/Nex, Francesco/Kerle, Norman/Vosselman, George. *Satellite Image Classification of Building Damages using Airborne and Satellite Image Samples in a Deep Learning Approach*. In: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Ausgabe IV-2/2018, Seite 89 ff.

Gardner, Daniel/Nichols, David. *Multi-Label Classification Of Satellite Images With Deep Learning*. 2017. [Zugriff am 24. Januar 2025]. Verfügbar unter: cs231n.stanford.edu

Goodfellow, Ian/Bengio, Yoshua/Courville, Aaron. *Deep Learning. An MIT Press book*. 2016. [Zugriff am 4. Februar 2025]. Verfügbar unter: www.deeplearningbook.org

Ji, Min/Liu, Lanfa/Buchroithner, Manfred. *Identifying Collapsed Buildings Using Post-Earthquake Satellite Imagery and Convolutional Neural Networks: A Case Study of the 2010 Haiti Earthquake*. In: Remote Sensing. Ausgabe 10(11)2018, Seite 1689. DOI: [10.3390/rs10111689](https://doi.org/10.3390/rs10111689)

Landesdatenbank NRW. *Datenlizenz Deutschland – Namensnennung – Version 2.0*. 2020. [Zugriff am 11. März 2025]. Verfügbar unter: www.govdata.de

Luo, Xiaochun/Li, Heng/Cao, Dongping/Dai, Fei/Seo, JoonOh/Lee, SangHyun. *Recognizing Diverse Construction Activities in Site Images via Relevance Networks of Construction-Related Objects Detected by Convolutional Neural Networks*. In: Journal of Computing in Civil Engineering. Jahrgang 32, Ausgabe 3/2018. DOI: [10.1061/\(ASCE\)CP.1943-5487.0000756](https://doi.org/10.1061/(ASCE)CP.1943-5487.0000756)

Maltezos, Evangelos/Doulamis, Nikolaos/Doulamis, Anastasios/Ioannidis, Charalabos. *Deep convolutional neural networks for building extraction from orthoimages and dense image matching point clouds*. In: Journal of Applied Remote Sensing. Jahrgang 11, Ausgabe 4/2017, Seite 1. DOI: [10.1117/1.JRS.11.042620](https://doi.org/10.1117/1.JRS.11.042620)

LITERATURVERZEICHNIS

Papadomanolaki, Maria/Vakalopoulou, Maria/Zagoruyko, Sergey/Karantzalos, Konstantinos. *Benchmarking deep learning frameworks for the classification of very high-resolution satellite multispectral data*. In: ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences. Ausgabe III-7/2016, Seite 83 ff. [Zugriff am 6. Februar 2025]. Verfügbar unter: noa.gwlb.de

Perez, Husein/Tah, Joseph H. M./Mosavi, Amir. *Deep Learning for Detecting Building Defects Using Convolutional Neural Networks*. In: Sensors. Ausgabe 19(16)/2019, Seite 3556. [Zugriff am 6. Februar 2025]. Verfügbar unter: arxiv.org

Postadjiana, Tristan/Le Bris, Arnaud/Sahbi, Hichem/Mallet, Clémont. *Investigating the potential of deep neural networks for large-scale classification of very high-resolution satellite images*. *ISPRS Annals of the Photogrammetry*. In: Remote Sensing and Spatial Information Sciences. Ausgabe IV-1/W1/2017, Seite 183 ff. [Zugriff am 5. Februar 2025]. Verfügbar unter: isprs-annals.copernicus.org

Van Rossum, Guido/Drake, Fred L. *Python 3 Reference Manual*. Scotts Valley, 2009. [Zugriff am 20. Februar 2025]. Verfügbar unter: dl.acm.org

Xu, Joseph Z./Lu, Wenhan/Li, Zebo/Khaitan, Pranav/Zaytseva, Valeriya. *Building Damage Detection in Satellite Imagery Using Convolutional Neural Networks*. 2019. [Zugriff am 7. Februar 2025]. Verfügbar unter: arxiv.org

RECHTSGRUNDLAGEN

Bauordnung für das Land Nordrhein-Westfalen (BauO NRW) in der Fassung der Bekanntmachung vom 21. Juli 2018, zuletzt geändert durch das Gesetz vom 31. Oktober 2023 (GV. NRW. Ausgabe 2023 Nr. 31 Seite 1167).

Gesetz über die Statistik der Bautätigkeit im Hochbau und die Fortschreibung des Wohnungsbestandes (Hochbaustatistikgesetz – HBauStatG) vom 5. Mai 1998 (BGBl. I Seite 869), das zuletzt durch Artikel 3 des Gesetzes vom 8. August 2020 (BGBl. I Seite 1728) geändert worden ist.

Gesetz über die Statistik für Bundeszwecke (Bundesstatistikgesetz – BStatG) in der Fassung der Bekanntmachung vom 20. Oktober 2016 (BGBl. I Seite 2394), das zuletzt durch Artikel 14 des Gesetzes vom 8. Mai 2024 (BGBl. I Nr. 152) geändert worden ist.



Prof. Dr. Ralf Münnich

ist Inhaber der Professur für Wirtschafts- und Sozialstatistik der Universität Trier und Stellvertretender Vorsitzender der Deutschen Statistischen Gesellschaft. Seit 2022 hat er den Vorsitz des unabhängigen Gremiums von Gutachterinnen und Gutachtern inne, das die wissenschaftlichen Arbeiten zur Auszeichnung mit dem Nachwuchspreis „Statistical Science for the Society“ des Statistischen Bundesamtes empfiehlt.

WISSENSCHAFTLICHER NACHWUCHSPREIS „STATISTICAL SCIENCE FOR THE SOCIETY“ 2024

Ralf Münnich

➤ **Schlüsselwörter:** amtliche Statistik – Nachwuchspreis – Record-Linkage-Methoden – Prävalenzschätzungen – Mikrosimulationen – Klimawandel

ZUSAMMENFASSUNG

Das Statistische Bundesamt vergibt seit 1999 Förderpreise an junge Wissenschaftlerinnen und Wissenschaftler, seit 2022 den wissenschaftlichen Nachwuchspreis „Statistical Science for the Society“. Der Name zeigt die Internationalität des Preises und bringt zum Ausdruck, dass sowohl die akademische als auch die amtliche Statistik dem Gemeinwohl verpflichtet sind. Mit dem Wissenschaftspreis „Statistical Science for the Society“ werden Nachwuchskräfte ausgezeichnet, deren Forschung Methodenprobleme mit einem engen Bezug zum Aufgabenspektrum der amtlichen Statistik behandelt oder empirische Fragestellungen unter intensiver Nutzung von Daten beantwortet.

➤ **Keywords:** official statistics – young scientist award – record linkage methods – prevalence estimates – microsimulations – climate change

ABSTRACT

The Federal Statistical Office has awarded prizes to junior scientists since 1999 and the “Statistical Science for the Society” prize since 2022. The prize’s name reflects its international character and shows that both academic and official statistics are committed to public welfare. The “Statistical Science for the Society” science prize is awarded to young researchers whose research addresses methodological problems that are closely related to the range of tasks of official statistics, or answers empirical questions with intensive use of data.

Einleitung

Auf Empfehlung eines unabhängigen Gremiums von Gutachterinnen und Gutachtern hat das Statistische Bundesamt am 28. November 2024 vier herausragende wissenschaftliche Arbeiten mit dem wissenschaftlichen Nachwuchspreis „Statistical Science for the Society“ ausgezeichnet. Im Mittelpunkt der prämierten Arbeiten stehen Fragestellungen rund um Statistik, Datenanalyse und neue digitale Anwendungen. Mit den Wissenschaftspreisen fördert das Statistische Bundesamt innovative Datennutzungen sowie Methodenforschungen in wissenschaftlichen Abschlussarbeiten mit engem Bezug zur angewandten Statistik.

- › Severin Weiland nahm den Wissenschaftspreis in der Kategorie „Master-/Bachelorarbeiten“ für seine Bachelorarbeit zum Thema „Vergleich von Record-Linkage Methoden anhand der Mikro-Simulation eines bundesweiten Schülerregisters“ entgegen. Hintergrund der Arbeit sind Bestrebungen von Bund und Ländern, ein bundesweites Schülerregister zu führen, das über Record-Linkage anhand von Personenmerkmalen erzeugt werden soll. Herr Weiland hat in seiner Bachelorarbeit verschiedene Record-Linkage-Methoden unter verschiedenen Datenqualitätsannahmen miteinander verglichen. Die Arbeit wurde betreut von Prof. Dr. Rainer Schnell an der Universität Duisburg-Essen.
- › Inken Siems erhielt den Wissenschaftspreis in der Kategorie „Master-/Bachelorarbeiten“ für ihre Masterarbeit zum Thema „Sampling for COVID-19 Statistics. A Comparison of Sampling Designs for Prevalence Estimates“. Am Beispiel der Coronakrise untersucht sie, wie qualitativ hochwertige epidemiologische Statistiken idealerweise gesammelt werden können, wann das etablierte Meldesystem verlässliche Zahlen liefert und an welcher Stelle Stichprobenstrategien die Ausbreitung der Infektion besser abschätzen können. Die Arbeit entstand an der Universität Trier bei Prof. Dr. Ralf Münnich.
- › In der Kategorie „Dissertationen“ bekam Dr. Simon Schmaus den Wissenschaftspreis für die Arbeit zum Thema „Methoden regionalisierter dynamischer Mikrosimulationen“. Im Erstellungsprozess regionalisierter Simulationsmodelle stehen Forschende vor großen Herausforderungen, beginnend mit der Generierung geeigneter Ausgangsdatensätze über

die Erfassung und Umsetzung der dynamischen Prozesse bis hin zur Auswertung der Ergebnisse und Quantifizierung von Unsicherheiten. Die Dissertation beschreibt ausgewählte Komponenten, die für regionalisierte Mikrosimulationen von besonderer Relevanz sind, und analysiert diese systematisch. Die Arbeit wurde betreut von Prof. Dr. Ralf Münnich an der Universität Trier.

- › Einen weiteren Wissenschaftspreis in der Kategorie „Dissertationen“ konnte Dr. Elisa Rottner für die Arbeit mit dem Thema „Electricity prices, emissions and firm performance – the case of German manufacturing“ entgegennehmen. Die kumulative Dissertation zeigt in vier Kapiteln am Beispiel des deutschen Verarbeitenden Gewerbes, dass amtliche Daten relevante Informationen für die Politikevaluation und -gestaltung enthalten. Damit können sie entscheidend dazu beitragen, die gesellschaftliche Herausforderung des Klimawandels effizienter und zielgerichteter zu lösen. Die Arbeit entstand an der Universität Basel bei Prof. Dr. Beat Hintermann.

Die Preisverleihung fand im Rahmen des [33. Wissenschaftlichen Kolloquiums „Energiewende und Energiepreiskrise – zur Rolle der Daten“](#) in Wiesbaden statt. Dr. Ruth Brand, die Präsidentin des Statistischen Bundesamtes, überreichte die Urkunden und hob die Bedeutung neuer Impulse aus der Forschung für die Zukunftsfähigkeit der amtlichen Statistik hervor. Die Laudationes auf die prämierten Arbeiten hielt der Vorsitzende des Gutachtergremiums, Professor Dr. Ralf Münnich (Universität Trier).

Die [Kurzfassungen der prämierten Arbeiten](#) sowie weitere Details zur Verleihung der Preise sind im [Internetangebot des Statistischen Bundesamtes](#) zu finden. Die Preisträgerinnen und Preisträger berichten im Jahr 2025 in dieser Zeitschrift ausführlich über ihre Arbeiten.

Laudationes wissenschaftliche Nachwuchspreise „Statistical Science for the Society“ 2024

Sehr geehrte Präsidentin Brand,
sehr verehrte Damen und Herren,

eine Gesellschaft wie die unsere braucht kluge Köpfe, die ihre Ideen mit besonderem Enthusiasmus und großem Engagement zu etwas Besonderem machen. Und gerade der wissenschaftliche Nachwuchs demonstriert uns immer wieder, zu was er imstande ist. Die großen Preisträger haben ihre besten Ideen auch meist in jungen Jahren gehabt. Insofern begrüße ich Sie auch dieses Jahr wieder mit besonderer Freude zur Preisverleihung des Wissenschaftlichen Nachwuchspreises „Statistical Science for the Society“.

Als Vorsitzender des Gutachterausschusses habe ich das Privileg, die Laudationes vornehmen zu dürfen. Und so ist es mir natürlich eine besondere Freude und Ehre, Sie heute hier im Museum Wiesbaden – aber auch online – zur Verleihung der Preise begrüßen zu dürfen. Und Sie werden sehen, dass auch das Thema des Kolloquiums „Energiewende und Energiepreiskrise – zur Rolle der Daten“ eine besondere Bedeutung bei der Preisverleihung einnimmt.

Auch dieses Jahr findet die Preisverleihung wieder im Rahmen des gemeinsamen Kolloquiums des Statistischen Bundesamtes und der Deutschen Statistischen Gesellschaft statt. Der besondere Dank gilt aber Frau Brand und ihrem Team, die diesen einzigartigen Rahmen wieder so überzeugend geschaffen haben – hier im Museum Wiesbaden tagen zu dürfen.

Viele aktuelle Themen, so auch das heutige Thema des Kolloquiums, zeigen, wie wichtig Daten sind. Aber auch das Wissen ob der Bedeutung der Qualität der Daten und derjenigen Methoden, die sie zu wertvollen Statistiken machen, ist von besonderem Gewicht. Eine Fähigkeit, die leider zu oft vernachlässigt wird.

Umso wichtiger ist es, einen ambitionierten Nachwuchs in der Statistik als Wissenschaft von der Datengewinnung bis zur Datenanalyse auszubilden und natürlich zu ehren, wenn sie bemerkenswerte Erkenntnisse errungen haben. Und so ist auch der Nachwuchspreis „Statistical Science for the Society“ zu verstehen. Seit Langem steht

dieser Preis sowohl für statistische Methoden im weiten Feld der amtlichen Statistik, aber auch für spannende Anwendungen von Daten der amtlichen Statistik. Und den besonderen Reiz macht die Interdisziplinarität aus, die stets mit Daten und Statistik verbunden ist – und so sollte man auch die Datenwissenschaften interpretieren, als Disziplin für die Gesellschaft!

Vielen Dank Frau Präsidentin Brand, Ihnen und Ihrem Statistischen Bundesamt, die Sie diesen Nachwuchspreis schon lange in so vorbildlicher Weise unterstützen – und dieses Jahr war das eine besondere Herausforderung, da wir einige Preise vergeben dürfen.

Ich habe es bereits angedeutet, dieses Jahr wurde das Gutachtergremium in besonderer Weise gefordert.

Ich danke den Kolleginnen Katja Ickstadt, Monika Jungbauer-Ganz und Notburga Ott sowie den Kollegen Roland Döhrn und Rainer Schnell für ihr besonderes Engagement, wieder die richtigen Arbeiten ausgewählt zu haben.

Zunächst darf ich in eigener Sache auch darauf hinweisen, dass genau die Thematik des Preises ausgewiesene Expertise in diesem Feld erfordert. Und so ist es wenig verwunderlich, dass sich aus dem Umfeld der Gutachterinnen und Gutachter gelegentlich herausragende Arbeiten anbieten, eingereicht zu werden. Ich kann Ihnen aber versichern, dass wir die Thematik einer möglichen Befangenheit bei der Begutachtung selbstverständlich korrekt berücksichtigen und betroffene akademische Lehrende auch bei entsprechenden Entscheidungen ausschließen müssen – wie das in diesem Jahr der Fall war.

Wir haben jetzt Nachwuchspreise in den Kategorien Bachelor- sowie Masterarbeiten und auch Dissertationen zu vergeben. Kommen wir zunächst zum Nachwuchspreis im Bereich Bachelorarbeiten.

Nach dem Zensus ist vor dem Zensus. Die Veröffentlichungen zum Zensus 2022 sind noch nicht abgeschlossen, aber die Diskussion um einen Registerzensus ist im vollen Gange. Auch wenn wir wohl noch einen Zwischenschritt zum reinen Registerzensus benötigen, ist es in jedem Falle sinnvoll, wesentliche Forschungsarbeiten zur Verwendung von Registern voranzutreiben. Eine besondere Herausforderung bei einem Registerzensus ist die Zusammenführung unterschiedlicher Regis-

ter, wenn keine eindeutigen Identifikatoren vorhanden sind. Mithilfe von Record-Linkage-Verfahren können auf Basis von Personenmerkmalen verschiedene Datensätze zusammengespült werden. Allerdings gibt es zahlreiche Verfahren mit vielen Parametern, die zu wählen sind, um sogenannte False-positive-Links beziehungsweise False-negative-Links zu reduzieren. Einen besonderen Einfluss auf die Anwendbarkeit und Effizienz der Verfahren hat die Datenqualität der jeweiligen Register. Selbstverständlich kann kein Register als vollständig exakt angenommen werden.

Der Autor der Arbeit bedient sich zunächst des Instrumentariums der Mikrosimulation im Sinne der Datengenerierung, um Register unter verschiedenen Annahmen fortzuführen. Dabei wurden ausgewählte Qualitätsszenarien implementiert und anschließend verschiedene Record-Linkage-Verfahren angewandt und miteinander verglichen.

Alleine die Implementation all dieser Verfahren ist für eine Bachelor-Arbeit exzellent. Beeindruckend ist dabei aber die Präzision der Analysen zur Beurteilung der Linkage-Güte, die in dieser Arbeit durchgeführt wurden.

Zweifelsohne ist die Anwendung im Kontext des Bildungsregisters von besonderer Bedeutung für die amtliche Statistik und sicher auch für zukünftige Zensus.

Der Förderpreis 2024 für Bachelorarbeiten geht an Herrn Severin Vincent Weiland vom Lehrstuhl für Empirische Sozialforschung am Institut für Soziologie der Universität Duisburg-Essen mit seiner Bachelorarbeit zum Thema „Vergleich von Record-Linkage Methoden anhand der Mikro-Simulation eines bundesweiten Schülerregisters“.

Herr Weiland, herzlichen Glückwunsch zum Nachwuchspreis 2024 für Ihre Bachelorarbeit!

Kommen wir zum Nachwuchspreis in der Kategorie Masterarbeiten:

Die Corona-Krise scheint vorbei zu sein. Die Frage lässt sich aber stellen, was wir daraus gelernt haben: Waren die Maßnahmen gerechtfertigt, und vor allem, waren die Daten von geeigneter Qualität für eine evidenzbasierte Politik? Bemerkenswert war, dass für wissenschaftliche Studien in Deutschland vielfach nicht-deutsche Daten herangezogen wurden. Ebenso bemerkenswert war, dass nach einigen Monaten die Politik sich darüber bewusst war, dass individuelle Meldungen von potenziellen oder

tatsächlichen COVID-19-Erkrankungen ein verzerrtes Bild über mögliche Prävalenzen und Inzidenzen geben, aber keinerlei Konsequenzen daraus gezogen wurden. Während in vielen Ländern qualifizierte Stichproben gezogen wurden, wie etwa in Großbritannien, wurde in Deutschland recht kurz über die Größe einer möglichen Stichprobe zur Schätzung regionaler Prävalenzen diskutiert. Zu lesen war, dass das Robert Koch-Institut einen Stichprobenumfang unter 15 000 Personen vorgeschlagen hat. Bei über 400 Kreisen und einer damals noch vorhandenen 7-Tage-Inzidenz von unter 10 durfte man sich aus statistischer Sicht durchaus fragen, ob das zielgerecht wäre. Schließlich hat sich Deutschland entschieden, auf eine echte Zufallsstichprobe zu verzichten.

Die Autorin dieser Arbeit sollte sich genau diesem Thema widmen, wie präzise beziehungsweise wie verzerrt COVID-Prävalenzen unter den Bedingungen, die wir in Deutschland hatten, gewesen sind. Ziel der Untersuchungen war herauszufinden, ob und wie eine probabilistische Stichprobe die Qualität der Schätzung von regionalen Prävalenzen verbessert hätte. Zur Lösung dieser Frage wurde das Instrumentarium der Mikrosimulation herangezogen. Unter Verwendung amtlicher Daten aus Trier und Trier-Saarburg sowie Daten des Robert Koch-Instituts wurde eine geeignete Simulationsgesamtheit erzeugt. Schließlich wurden verschiedene epidemiologische Modelle in die Untersuchungen integriert.

Auf Basis dieser Daten wurden anschließend sowohl unterschiedliche Arten der Selbstmeldungen wie auch zahlreiche echte Stichprobenmodelle analysiert und miteinander verglichen. Bei der Ermittlung der Prävalenzen wurde auch die Qualität unterschiedlicher Testverfahren zum Nachweis von COVID-19 berücksichtigt.

Die Autorin konnte nachweisen, dass in allen Modellen die Selbstmeldung, wie sie in Deutschland durchgeführt wurde, zu überwiegend erheblich verzerrten, teilweise unbrauchbaren Prävalenzen geführt hat. Ebenso konnte sie zeigen, dass echte Stichprobenverfahren zu teilweise deutlichen Verbesserungen der Qualität der COVID-Berichte geführt hätten. Allerdings zeigt sich auch, dass bei niedrigen Infektionszahlen der Stichprobenumfang sehr hoch hätte ausfallen müssen, um Informationen zu regionalen Prävalenzen in hoher Qualität zu erhalten.

Somit liefert die Arbeit eine ausgezeichnete Grundlage, bei aufkommenden Pandemien helfen zu können, in Deutschland Daten guter Qualität über das Infektions-

geschehen zu erhalten und mit dieser Datenbasis politische Maßnahmen zu begründen.

Der Nachwuchspreis 2024 für Masterarbeiten geht an Frau Inken Siems von der Professur für Wirtschafts- und Sozialstatistik an der Universität Trier mit ihrer Masterarbeit zum Thema „Sampling for COVID-19 Statistics – A Comparison of Sampling Designs for Prevalence Estimates“.

Herzlichen Glückwunsch, Frau Siems, zum Nachwuchspreis 2024 für Ihre Masterarbeit!

Kommen wir zum Nachwuchspreis in der Kategorie Dissertationen:

Tatsächlich freue ich mich, dass wir heute zwei herausragende Dissertationen prämiieren dürfen.

Nachdem in beiden Preisen zuvor bereits die Mikrosimulation als Forschungsmethode verwendet wurde, kann man sich natürlich fragen: Was ist das und warum spielt sie für eine amtliche Statistik eine große Rolle?

Mit der Novellierung des § 3 Absatz 1 Nr. 6 Bundesstatistikgesetz im Jahre 2016 erhielt das Statistische Bundesamt eine neue Aufgabe, „Entwicklung und Anwendung von Mikrosimulationsmodellen sowie mikroökonomische Analysen durchzuführen“. Und so ist es wenig verwunderlich, dass das Statistische Bundesamt zusammen mit den Universitäten Trier und Duisburg-Essen eine DFG-Forschungsgruppe zu regionalen Mikrosimulationen aufgesetzt hat. Ziel ist die Erarbeitung einer methodischen Grundlage für regionale dynamische Mikrosimulationen sowie die Erschaffung eines Datenlabors für Deutschland.

Folgt man Li und O'Donoghue (2013)¹, dann bestehen Mikrosimulationen aus zwei essenziellen Teilen: der Schaffung einer geeigneten Datenbasis sowie der eigentlichen Durchführung von Wenn-dann-Analysen zur Lösung von Fragen, die man sonst nicht hätte beantworten können. Ohne Zweifel kann eine Mikrosimulation nur so gut sein wie die für eine Mikrosimulation verwendeten Daten. Verwendet man verfügbare Daten, etwa Register oder Stichprobenerhebungen, so verfügt man meistens nur über eine gewisse Auswahl an Varia-

blen, aber selten über einen umfangreichen Datenkranz, der regional, eventuell sogar geokodiert verfügbar ist. Möchte man zusätzlich die Entwicklung über die nächsten Jahre, vielleicht sogar über einige Jahrzehnte, berücksichtigen, so benötigt man eine Reihe von modernen survey-statistischen Verfahren und auch Prädiktionsmethoden, um derartige Datenbasen von hoher Qualität aufbauen zu können.

Es mag wenig überraschen, dass die meisten Publikationen zu Mikrosimulationen zu Anwendungen, also Wenn-dann-Analysen, vorliegen. In den letzten Jahren befassen sich aber immer mehr Forscherinnen und Forscher mit der synthetischen Generierung beziehungsweise Erweiterung von Daten, die für dynamische, regionale Mikrosimulationen herangezogen werden konnten.

Die vorliegende Dissertation liefert eine äußerst gelungene aktuelle Übersicht und kritische Würdigung über dynamische Mikrosimulationsmethoden und ihre Probleme, insbesondere aus survey-statistischer Sicht. Die ausführliche Diskussion von kleinräumigen Mikrosimulationen sowie die Entwicklung von Binnenmigrationsmodulen liefern wichtige innovative Ansätze zur Weiterentwicklung von Mikrosimulationsmodellen.

Mit dem Nachwuchspreis 2024 in der Kategorie Dissertationen wird Herr Dr. Simon Schmaus von der Professur für Wirtschafts- und Sozialstatistik an der Universität Trier mit seiner Dissertation zum Thema „Methoden regionalisierter dynamischer Mikrosimulationen“ ausgezeichnet.

Ganz herzliche Gratulation an Simon Schmaus zum Nachwuchspreis 2024 in der Kategorie Dissertationen!

Bisher mag der Eindruck erweckt worden sein, methodische Arbeiten stünden im besonderen Fokus des Nachwuchspreises. Ich kann Ihnen versichern, dass der Augenblick des Jahres und seiner zahlreichen hervorragenden Einreichungen dazu geführt haben mag. Aber selbstverständlich betrachten und diskutieren die Gutachter alle unterschiedlichen Arbeiten in ihrer Vielfalt sehr ausführlich, auch wenn das nicht immer ganz einfach ist. Aber es freut mich, dass wir noch den Gegenbeweis antreten und final eine Dissertation aus dem Bereich Anwendung von amtlichen Daten prämiieren. Und ein Drehbuch hätte es kaum besser hinbekommen, diese Arbeit passt auch noch bestens in das Thema des diesjährigen Kolloquiums.

1 Li, Jinjing/O'Donoghue, Chatal. *A survey of dynamic microsimulation models: Uses, model structure and methodology*. In: International Journal of Microsimulation. Jahrgang 6. Ausgabe 2/2013, Seite 3 ff. DOI: [10.34196/ijm.00082](https://doi.org/10.34196/ijm.00082)

Die zu prämierende Dissertation wurde als kumulative Dissertation eingereicht. Ich kann Ihnen versichern, dass wir gemäß den Statuten der Preisverleihung die Eigenleistung der Arbeit bewertet haben, da drei der vier Beiträge in Ko-Autorenschaft entstanden sind.

Empirische Grundlage der vier Beiträge sind die Amtlichen Firmendaten für Deutschland (AFiD) für das Verarbeitende Gewerbe.

Ich erlaube mir, kurz auf die vier Beiträge einzugehen, die sich allesamt politisch hoch relevanten Fragestellungen widmen:

Das erste Papier befasst sich mit den Treibern der CO₂-Emissionen in Industrieunternehmen. Für diese Untersuchung werden die Daten zum Produktionsspektrum der Unternehmen auf tief disaggregierter Ebene genutzt. Die Untersuchungen zeigen, dass eine Verschiebung des Produktionsspektrums hin zu sauberen Produkten stattgefunden hat, was zu einer Reduktion der CO₂-Emissionen in Deutschland geführt hat. Belege für Carbon-leakage-Effekte findet die Analyse nicht.

Im zweiten Beitrag wird der Einfluss von Elektrizitätspreisen auf den Elektrizitätsverbrauch in Unternehmen des Verarbeitenden Gewerbes untersucht. Zudem wird analysiert, ob die Unternehmen aufgrund steigender Preise zur Selbsterzeugung von Strom übergehen. Die Analyse findet negative Effekte von steigenden Netzentgelten auf Energiebeschaffung, Energieverbrauch und Emissionen sowie positive Effekte auf die Eigenerzeugung von Strom. Ein Einfluss auf die Wettbewerbsfähigkeit wird nicht gefunden.

Im dritten Beitrag untersucht die Autorin, ob Klimapolitiken das Outsourcing fördern, was mit Carbon-leakage-Effekten verbunden sein könnte. Dabei macht sie sich zunutze, dass nur ein Teil der Unternehmen von Ausnahmeregelungen bei der EEG-Umlage profitieren. Es zeigt sich mithilfe eines Difference-in-difference-Ansatzes, dass weder die inländische Beschaffung von Vorprodukten noch die Importe von der Variation der EEG-Abgaben beeinflusst werden. Die Autorin analysiert sehr kritisch die eigenen Ergebnisse mithilfe zahlreicher Tests auf Robustheit.

Ausgangspunkt des vierten Beitrags der Dissertation ist, dass die CO₂-Emissionen in der Industrie in Deutschland im Gegensatz zu anderen EU-Ländern zwischen 2005

und 2019 nicht gesunken sind. Als Erklärung fokussiert die Arbeit auf regulatorische Unterschiede zwischen Deutschland und anderen EU-Ländern. Mithilfe eines quantifizierten Handelsmodells wird gezeigt, dass die deutsche Regulierung dazu beiträgt, dass die CO₂-Emissionen in Deutschland relativ zum EU-Ausland höher sind.

Besonders herausragend aus Sicht der Gutachter – und damit auch wesentliche Grundlage für die Preisverleihung – ist die intensive und überzeugende Auseinandersetzung der Autorin mit den AFiD-Daten. Die Autorin hat sich intensiv mit dem Datenangebot der amtlichen Statistik auseinandergesetzt und die Daten eigenständig aufbereitet. Wo AFiD für die jeweilige Fragestellung benötigte Daten nicht enthält, wurden diese geschätzt oder aus anderen Quellen der amtlichen Statistik entnommen. Darüber hinaus hat die Verfasserin der Arbeit die Grenzen der Aussagefähigkeit der Daten ausgelotet und sich mit möglichen, mit der Datenbasis verbundenen Problemen auseinandergesetzt. Ein verantwortungsvoller Umgang mit Daten, ihrer Qualität und Aussagekraft ist mehr denn je eine essenzielle Grundlage für Forschung von höchster Qualität!

Der Nachwuchspreis 2024 für Dissertationen geht an Frau Dr. Elisa Rottner von der Professur Public Economics an der Wirtschaftswissenschaftlichen Fakultät der Universität Basel mit ihrer Dissertation zum Thema „Electricity prices, emissions and firm performance – the case of German manufacturing“.

Herzlichen Glückwunsch, Frau Rottner, zum Nachwuchspreis 2024 für Ihre Dissertation! 🇩🇪

Herausgeber

Statistisches Bundesamt (Destatis), Wiesbaden

Schriftleitung

Dr. Daniel Vorgrimler

Redaktion: Ellen Römer

Ihr Kontakt zu uns

www.destatis.de/kontakt

Erscheinungsfolge

zweimonatlich, erschienen im April 2025, Seite 95 korrigiert am 06.05.2025

Ältere Ausgaben finden Sie unter www.destatis.de sowie in der [Statistischen Bibliothek](#).

Artikelnummer: 1010200-25002-4, ISSN 1619-2907

© Statistisches Bundesamt (Destatis), 2025

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.