

Statistisches Bundesamt (Destatis) (Ed.)

**Periodical Part**

WISTA – Wirtschaft und Statistik, Ausgabe 1/2026

WISTA - Wirtschaft und Statistik

**Provided in Cooperation with:**

Statistisches Bundesamt (Destatis), Wiesbaden

*Suggested Citation:* Statistisches Bundesamt (Destatis) (Ed.) (2026) : WISTA – Wirtschaft und Statistik, Ausgabe 1/2026, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 78, Iss. 1

This Version is available at:

<https://hdl.handle.net/10419/337209>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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

---

# WISTA **Wirtschaft und Statistik**

---

R. Lehmann | S. Linz |  
S. Sauer | J. Wilhelm |  
T. Wollmershäuser

J. Baum

M. Bellinghen | S. Felgendreher |  
D. Kanmegne | J. Oehrlein |  
J. Reith | S. Schürz

N. Gillmann

M. Rengers | T. Körner

H. Brenzel |  
Y. Garcia Ritz | D. Wagner

M. Brauchler | M. Köhlmann |  
F. Kraft | M. Köberl

S. Claußen

O. Hauke | A. Kristiansen

**Strukturwandel in der Industrie: Herausforderung für die Unternehmensstatistik**

**Die neue Statistik zu globalen Wertschöpfungsketten**

**Kühlung in Städten, Kohlenstoffspeicherung, Landschaften für Tourismus: Ökosystemrechnungen zeigen Leistungen der Natur**

**Die Reklassifikation von Verkehrsinfrastruktureinheiten in den Staatssektor**

**Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung**

**Partielle Synthetisierung zur Anonymisierung der verknüpften Verdienststrukturerhebung 2018 – Teil 2: Evaluation der Anonymität und Ausblick zur Datenbereitstellung**

**Entwicklung von statistischen Wirtschaftsindikatoren unter Verwendung der Copernicus-Sentinel-Satelliten**

**Metadatenmodell der amtlichen Statistik**

**Effiziente Auswertung des Business-Tax-Panels mithilfe der Statistiksoftware R**

---

**1 | 2026**

#### 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  | Aktuelle Informationsangebote                                                                                                                                                                                                                                                                                                                                        |
| 6  | Kurznachrichten                                                                                                                                                                                                                                                                                                                                                      |
| 11 | Robert Lehmann, Stefan Linz, Stefan Sauer, Jan Wilhelm, Timo Wollmershäuser<br><b>Strukturwandel in der Industrie: Herausforderung für die Unternehmensstatistik</b><br><i>Structural change in industry presents challenges for business statistics</i>                                                                                                             |
| 28 | Joschka Baum<br><b>Die neue Statistik zu globalen Wertschöpfungsketten</b><br><i>New statistics on global value chains</i>                                                                                                                                                                                                                                           |
| 41 | Marius Bellinghen, Simon Felgendreher, Dan Kanmegne, Johannes Oehrlein, Jonathan Reith, Simon Schürz<br><b>Kühlung in Städten, Kohlenstoffspeicherung, Landschaften für Tourismus: Ökosystemrechnungen zeigen Leistungen der Natur</b><br><i>Urban cooling, carbon sequestration, landscapes for tourism: ecosystem accounts reflect services provided by nature</i> |
| 56 | Niels Gillmann<br><b>Die Reklassifikation von Verkehrsinfrastruktureinheiten in den Staatssektor</b><br><i>Reclassification of transport infrastructure units to the general government sector</i>                                                                                                                                                                   |



# INHALT

|     |                                                                                                                                                                                                                                                                                                                                                                                                                      |
|-----|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 68  | Martina Rengers, Thomas Körner<br><b>Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung</b><br><i>Working hours and working-time preferences of men (and women) – Starting points for identifying gender equality opportunities</i>                                                                                                                 |
| 87  | Hanna Brenzel, Yannik Garcia Ritz, David Wagner<br><b>Partielle Synthetisierung zur Anonymisierung der verknüpften Verdienststrukturerhebung 2018</b><br><b>Teil 2: Evaluation der Anonymität und Ausblick zur Datenbereitstellung</b><br><i>Partial data synthesis for anonymising the linked datasets of the 2018 Structure of Earnings Survey – Part 2: evaluation of anonymity and outlook on data provision</i> |
| 100 | Melanie Brauchler, Maren Köhlmann, Franziska Kraft, Manuel Köberl<br><b>Entwicklung von statistischen Wirtschaftsindikatoren unter Verwendung der Copernicus-Sentinel-Satelliten</b><br><i>Development of economic indicators in statistics using Copernicus Sentinel satellites</i>                                                                                                                                 |
| 113 | Sven Claußen<br><b>Metadatenmodell der amtlichen Statistik</b><br><i>Metadata Model of Official Statistics</i>                                                                                                                                                                                                                                                                                                       |
| 127 | Oliver Hauke, Annette Kristiansen<br><b>Effiziente Auswertung des Business-Tax-Panels mithilfe der Statistiksoftware R</b><br><i>Efficient assessment of the business tax panel using the statistical software R</i>                                                                                                                                                                                                 |

---

# EDITORIAL

---

Dr. Ruth Brand

---



## LIEBE LESERIN, LIEBER LESER,

Wertschöpfungsketten haben sich in den letzten Jahrzehnten zunehmend globalisiert, jedoch gab es in den vergangenen Jahren verschiedene Entwicklungen, die den internationalen Handel beeinträchtigten: Lieferengpässe durch die COVID-19-Pandemie, Sanktionen gegen Russland nach dem russischen Angriff auf die Ukraine, oder auch die Zölle, die die Vereinigten Staaten auf bestimmte importierte Waren einführten. Als eine der am stärksten in den Welthandel integrierten Volkswirtschaften ist Deutschland in besonderem Maße auf belastbare, präzise Daten angewiesen. Auf ihrer Grundlage lassen sich politische Entscheidungen sachgerecht treffen und geopolitische Abhängigkeiten realistisch einschätzen. Solche Daten bietet die neue Statistik zu globalen Wertschöpfungsketten, die ein Beitrag dieser Ausgabe von WISTA vorstellt.

Die amtliche Statistik muss sich aktuell der Frage stellen, wie sich der industrielle Strukturwandel in der Entwicklung der Konjunkturindikatoren niederschlägt und welche Herausforderungen daraus entstehen. Gemeinsam haben Mitarbeiter des Statistischen Bundesamtes und des ifo Instituts untersucht, ob die wachsende Divergenz zwischen dem industriellen Produktionsindex und dem Produktionswert der Volkswirtschaftlichen Gesamtrechnungen für das Verarbeitende Gewerbe Ausdruck für einen solchen Strukturwandel ist. In einem Artikel dieser Ausgabe zeigen sie, dass die Unterschiede zwischen Produktionsindex und Produktionswert überwiegend aus sogenannten sonstigen Umsätzen resultieren. Dazu zählen Umsätze aus Aktivitäten, die nicht unmittelbar aus dem Verkauf von in Deutschland hergestellten Waren stammen, sondern beispielsweise aus produktbegleitenden Dienstleistungen.

Daneben bietet diese Ausgabe weitere interessante Aufsätze, die ich Ihnen zur Lektüre empfehle.

*Ruth Brand*

Präsidentin des Statistischen Bundesamtes



### 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](http://www.dashboard-deutschland.de)



### Ergebnisse des Zensus 2022

Wie viele Menschen leben in Deutschland? Wie wohnen und arbeiten sie? Die Ergebnisse des Zensus 2022 liefern Antworten auf Fragen zur Bevölkerung und zu den Wohnverhältnissen in Deutschland. Kompakte Informationen dazu enthält die Zensus-Themenseite im Internetangebot des Statistischen Bundesamtes. Alle Daten sind kostenlos und frei zugänglich in der [Zensusdatenbank](#) verfügbar. Der [Zensusatlas](#) bietet eine anschauliche Darstellung.

➤ [www.destatis.de](http://www.destatis.de)



### Rente

Wie entwickelt sich in Zukunft die Bevölkerung im Rentenalter? Wovon leben Personen im Ruhestand? Von der demografischen Komponente bis zur Einkommenssituation älterer Menschen – das Statistische Bundesamt bündelt alle relevanten Informationen zum Thema Rente und Alterssicherung auf einer eigenen Sonderseite.

➤ [www.destatis.de](http://www.destatis.de)



### Der Durchschnittsmensch in Deutschland

Angenommen, ein Mensch stünde für alle, die hier leben. Wie groß ist dieser Durchschnittsmensch? Wie alt ist und wird er, wie sehen seine Lebensphasen aus? Wie lebt und wohnt er? Wie arbeitet der Durchschnittsmensch und wie verbringt er seine Zeit? Und wie viel Geld verdient er oder hat er zur Verfügung? Antworten zu diesen und vielen weiteren Fragen bietet die Sonderseite unter

➤ [www.destatis.de](http://www.destatis.de)



### 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](http://www.destatis.de/fachkraefte)



### 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](http://www.destatis.de/klima)

# KURZNACHRICHTEN

## IN EIGENER SACHE

### Fuxi: der digitale Guide auf destatis.de



Seit Januar 2026 assistiert Fuxi auf der Website des Statistischen Bundesamtes. Der FAQ-Chatbot beantwortet einfache sowie häufig gestellte Fragen (FAQ) – schnell und rund um die Uhr. Zudem unterstützt er die Besucherinnen und Besucher der Website aktiv bei der Navigation

auf der Website, indem er sie mit seinen Antworten direkt zu den relevanten Inhalten führt.

Fuxi ist ein regelbasierter, nicht-generativer Chatbot. Das heißt, er kann nur Fragen beantworten, die in seiner Wissensdatenbank hinterlegt sind. Fuxi versteht sowohl vollständige Fragen als auch Schlagworte. Kann der FAQ-Chatbot eine Frage noch nicht beantworten, verweist er automatisch auf das Kontaktformular. Komplexe oder individuelle Anfragen werden also wie gewohnt vom Zentralen Auskunftsdienst des Statistischen Bundesamtes bearbeitet.

➤ [www.destatis.de](https://www.destatis.de)

## Im Fokus: Fakten zum Thema Rente

Die demografische Entwicklung stellt das Rentensystem vor komplexe Herausforderungen. Um eine fundierte Datenbasis für die aktuelle Debatte zu schaffen, hat das Statistische Bundesamt Daten zum Thema Rente und Alterssicherung auf einer neuen Sonderseite gebündelt. Gegliedert nach den Bereichen Demografie, Arbeit, Einkommen und Staatshaushalt sind Informationen zu Rentenbezugsdauer, Altenquotienten und die wirtschaftliche Lage im Alter übersichtlich aufbereitet und zentral verfügbar unter:

➤ [www.destatis.de](https://www.destatis.de)

## AUS EUROPA

### 60. Sitzung des AESS

Der Ausschuss für das Europäische Statistische System (AESS) tagte am 23. und 24. Oktober 2025 in Split, Kroatien. Er hat Folgendem zugestimmt:

> dem Entwurf einer Durchführungsverordnung der Rahmenverordnung LMB (labour market statistics on businesses) der Kommission zur Festlegung der Liste und Beschreibung der Variablen und ihrer technischen Spezifikationen, statistischen Klassifikationen, Untergliederungen, Metadaten und Präzisionsziele für die Themen „Stellenangebote“, „Arbeitskostenindex“ und „Geschlechterentgeltgefälle“ gemäß der Verordnung (EU) 2025/941 des Europäischen Parlaments und des Rates;

- > dem Entwurf einer Durchführungsverordnung (EU) der Kommission über die Maßnahmen zur Vollständigkeit im Sinne der Verordnung (EU) 2019/516 des Europäischen Parlaments und des Rates zur Harmonisierung des Bruttonationaleinkommens (BNE) zu Marktpreisen und zur Aufhebung des Beschlusses 94/168/EG;
- > dem Entwurf einer Durchführungsverordnung (EU) der Kommission zur Festlegung technischer Elemente der Datensätze für das Ad-hoc-Thema 2027 „Psychische Gesundheit und Wohlbefinden“ im Bereich Einkommen und Lebensbedingungen gemäß der Verordnung (EU) 2019/1700 des Europäischen Parlaments und des Rates.

Das Europäische System der Volkswirtschaftlichen Gesamtrechnungen (ESA) soll revidiert werden, um die internationale Vergleichbarkeit mit dem System of National Accounts (SNA) der Vereinten Nationen zu wahren. Der AESS diskutierte über den Zeitpunkt der Implementierung sowie die Bereitstellung von Zeitreihen. Deutschland wies darauf hin, dass nicht alle neuen ESA-Lieferverpflichtungen rückwirkend ab 1999 realisiert werden können.

Hinsichtlich einer geplanten Reform der Unternehmens- und Handelsstatistiken steht die Verordnung (EU) 2019/2152 über die europäischen Unternehmensstatistiken (EBS-Verordnung) im Fokus. Eurostat, das Statistische Amt der Europäischen Union, sieht eine notwendige Anpassung aufgrund gestiegener Anforderungen und Änderungen in vielen Wirtschaftsbereichen. Deutschland und auch andere Mitgliedstaaten betonten, dass gestiegene Datenanforderungen bei gleichzeitiger Entlastung von Unternehmen nicht allein durch neue Datenquellen und Auflösung von Doppelerhebungen zu erfüllen sind. Der AESS schlägt vor, vor allem die Lieferumfänge kritisch zu prüfen und zu reduzieren.

Unter dem Tagesordnungspunkt „Auf dem Weg zu wirksamen und effizienten europäischen Statistiken“ wurde die Überprüfung statistischer Bereiche behandelt, die bei der mehrjährigen Planung für 2021 bis 2027 vereinbart wurde. Mit sogenannten Stresstests soll geprüft werden, ob bestehende statistische Anforderungen weiterhin für den Bedarf von Nutzerinnen und Nutzern relevant sind und ob

die damit verbundenen Kosten und Belastungen verhältnismäßig sind. Der Plan sieht vor, bis Ende 2029 alle EU-Verordnungen im Bereich Statistik zu überprüfen mit dem Ziel, die Meldepflichten für Unternehmen um 25 % zu reduzieren.

Im Bereich der Digitalisierung plant Eurostat, bestehende Datenquellen zu untersuchen, um umfassendere und aktuelle Statistiken über digitale Infrastrukturen und Dienste zu erheben. Ein wichtiger Schritt in diesem Kontext ist, die Altersspanne in der Umfrage zum Einsatz von Informations- und Kommunikationstechnologie (IKT) in Haushalten zu erweitern, um die digitale Teilhabe älterer Menschen besser zu erfassen. Ein weiterer Aspekt ist, neue IKT-Module zur Nutzung von generativer Künstlicher Intelligenz einzuführen. Deutschland begrüßte die vorgestellte EU-Initiative, äußerte jedoch grundsätzliche Bedenken bei der Einführung zusätzlicher Indikatoren zur Digitalisierung, da diese konträr zu den allgemeinen Zielen – den Verwaltungsaufwand verringern und die EU-Gesetzgebung vereinfachen – stehen. Bei den IKT-Haushaltserhebungen wird Deutschland keine freiwilligen Daten zum erweiterten Altersbereich (zusätzliche Personen im Alter von 75 bis 89 Jahren) sowie auf regionaler Ebene (NUTS 2 und NUTS 3) bereitstellen, da die betreffende Rahmenverordnung dies nicht vorsieht. Grundsätzlich ist sicherzustellen, dass freiwillige Datenlieferungen langfristig nicht verpflichtend werden.

### AUS DEM INLAND

---

#### Neues Meldeportal der Bundesregierung

Ab sofort gibt es eine zentrale Anlaufstelle für Beschwerden zur Bürokratie – das EinfachMachen-Portal. Bürgerinnen und Bürger sowie Unternehmen können direkt online melden, in welchen Fällen sie Staat und Verwaltung als unnötig kompliziert erachten. Das Portal zählt zu den Maßnahmen der Modernisierungsagenda Bund, mit der die Bundesregierung Entscheidungen, Prozesse und Strukturen der öffentlichen Verwaltung modernisiert.

Weitere Informationen sind verfügbar unter:

➤ [bmds.bund.de](https://bmds.bund.de)

#### 16. koordinierte Bevölkerungsvorausberechnung vorgestellt

Im Jahr 2035 wird jede vierte Person in Deutschland 67 Jahre oder älter sein. Zu diesem Ergebnis kommen alle Varianten der 16. koordinierten Bevölkerungsvorausberechnung, die das Statistische Bundesamt am 11. Dezember 2025 bei einer Pressekonferenz in Berlin vorgestellt hat. Zum Vergleich: Im Jahr 2024 war nur jede fünfte Person im Rentenalter ab 67 Jahren.

Die Ausgangslage für die insgesamt 27 Berechnungsvarianten war eine andere als bei der vorherigen Vorausberechnung im Jahr 2022: Die zuletzt gesunkene Zuwanderung und die niedrigere Geburtenrate führen zusammen mit der wieder gestiegenen Lebenserwartung dazu, dass die Bevölkerung Deutschlands langfristig stärker altert als nach der 15. Vorausberechnung Ende 2022 angenommen. Schon heute kommen auf 100 Personen im Erwerbsalter 33 Personen im Rentenalter. Im Jahr 2070 werden es im ungünstigsten Fall 61 Personen im Rentenalter sein, also fast doppelt so viele wie heute. Insgesamt wird die Bevölkerung in Deutschland nach der mittleren Variante der Bevölkerungsvorausberechnung bis 2070 auf weniger als 75 Millionen Menschen schrumpfen.

Alle Ergebnisse der Bevölkerungsvorausberechnung für Deutschland und die Bundesländer bietet die Themenseite Bevölkerungsvorausberechnung.

➤ [www.destatis.de](https://www.destatis.de)

### STATISTIK VISUALISIERT

---

#### Neue Kartenanwendung zum Mietenvergleich in der eigenen Nachbarschaft

Deutschland ist das Mieterland Nummer 1 in der Europäischen Union. Über die Hälfte der Bevölkerung lebte hierzulande im Jahr 2024 zur Miete. Dabei hängt die Höhe der Miete maßgeblich ab vom Standort sowie dem Zustand und der Größe der Wohnung. Eine Kartenanwendung auf Basis des Zensus 2022 visualisiert den Mietenvergleich nach Alter des Gebäudes und der Größe der Wohnung. Sie verdeutlicht, wie viel die Miete in vergleichbaren Wohnungen in der Umgebung einer bestimmten Wohnung beträgt. Mit diesem Kartentool lassen sich Bestandsmieten deutschlandweit für gleich große und gleich alte Wohnungen auf der Basis von 100-m-Gitterzellen vergleichen, sofern genügend vergleichbare Mietwohnungen vorhanden sind.

➤ [service.destatis.de](https://service.destatis.de)

### NEUERSCHEINUNGEN

---

#### Hildegard Bartels – Gruppenbild mit Dame

Als erste Frau an der Spitze einer Bundesoberbehörde schrieb Hildegard Bartels deutsche Verwaltungsgeschichte. Von 1972 bis 1980 leitete sie das Statistische Bundesamt – in einer Zeit, in der Führungspositionen in der öffentlichen Verwaltung fest in Männerhand waren. Mit Fachkompetenz, Beharrlichkeit und Führungsstärke setzte sie sich durch und wurde so zur Wegbereiterin für nachfolgende Generationen von Frauen im Staatsdienst. Carsten Stahmer zeichnet in „Hildegard Bartels – Gruppenbild mit Dame. Dokumentation über die erste Präsidentin des Statistischen Bundesamtes und ihrer Zeit“ nicht nur den beruflichen Werdegang einer außergewöhnlichen Persönlichkeit nach, sondern gibt mittels Archivmaterialien und Auszügen aus verschiedenen Veröffentlichungen auch Einblicke in die Entwicklung des Statistischen Bundesamtes und beleuchtet die gesellschaftlichen und politischen Herausforderungen während ihrer Amtszeit.

↘ [reichert-verlag.de](https://reichert-verlag.de)

#### Key figures on European transport 2025

Diese Veröffentlichung des Statistischen Amtes der Europäischen Union (Eurostat) bietet anhand von intuitiven Visualisierungen und kompakten Texten einen umfassenden Überblick über den europäischen Personen- und Güterverkehr. Sie umfasst Indikatoren aus den folgenden Hauptbereichen:

- > Personen- und Warenverkehr
- > Verkehrssicherheit
- > Verkehr, Umwelt und Energie
- > Verkehr und Wirtschaft.

↘ [ec.europa.eu](https://ec.europa.eu)





---

# STRUKTURWANDEL IN DER INDUSTRIE: HERAUSFORDERUNG FÜR DIE UNTERNEHMENSSTATISTIK

---

Robert Lehmann, Stefan Linz, Stefan Sauer, Jan Wilhelm, Timo Wollmershäuser

---

📌 **Schlüsselwörter:** Produktionsindex – Produktionswert – Dienstleistungen der Industriebetriebe – VGR – Konjunkturstatistiken

## ZUSAMMENFASSUNG

Im Zentrum des Beitrags stehen die wachsenden nichtindustriellen Tätigkeiten von Industrieunternehmen und die damit verbundenen Geschäftsmodelle. Industrieunternehmen in Deutschland erzielen zunehmend Umsätze aus Dienstleistungen und anderen nichtindustriellen Tätigkeiten. Deren wachsende Bedeutung trägt wesentlich zur Divergenz zwischen dem industriellen Produktionsindex und dem Produktionswert der Volkswirtschaftlichen Gesamtrechnungen für das Verarbeitende Gewerbe bei. Umsätze aus solchen Aktivitäten gewinnen insbesondere in der Kfz-Industrie an Bedeutung. Eine Sonderbefragung des ifo Instituts bestätigt die Zunahme der nichtindustriellen Umsätze, weist jedoch abweichende absolute Umsatzanteile aus, was auf Abgrenzungsprobleme in der statistischen Erfassung hindeutet.

📌 **Keywords:** *production index – output – services by industrial establishments – national accounts – short-term economic statistics*

## ABSTRACT

*The article focuses on the increasing share of non-industrial activities carried out by industrial enterprises and the associated business models. Industrial enterprises in Germany are generating an ever larger proportion of turnover from services and other non-industrial activities. The latter's growing importance is a major contributing factor to the divergence between the industrial production index and the output recorded in national accounts for the manufacturing sector. Turnover from such activities is playing an increasingly important role particularly in the automotive industry. A special survey conducted by the ifo Institute has confirmed the increase in non-industrial turnover, but revealed diverging turnover shares in absolute figures, indicating problems with regard to statistical delimitation.*

**Dr. Robert Lehmann, Stefan Sauer und Prof. Dr. Timo Wollmershäuser**

sind am ifo Institut für Wirtschaftsforschung in München im Zentrum für Makroökonomik und Befragungen tätig.

**Dr. Stefan Linz und Jan Wilhelm**

sind im Statistischen Bundesamt im Bereich „Konjunkturindizes, Saisonbereinigung“ der Unternehmensstatistiken tätig.

## 1

### Einleitung

Seit Mitte der 2010er-Jahre verläuft der in den Volkswirtschaftlichen Gesamtrechnungen (VGR) ermittelte Produktionswert im Verarbeitenden Gewerbe deutlich dynamischer als der monatlich ermittelte Produktionsindex der Konjunkturstatistik. Die wachsende Diskrepanz interpretieren das ifo Institut – Leibniz-Institut für Wirtschaftsforschung an der Universität München e.V. und das Statistische Bundesamt in einer gemeinsamen Untersuchung als Ausdruck eines industriellen Strukturwandels (Lehmann und andere, 2025).

Während der Produktionsindex die Entwicklung des Realwertes der Warenproduktion abbildet, berücksichtigt der Produktionswert der VGR auch die wachsende Bedeutung nichtindustrieller Tätigkeiten der deutschen Industrieunternehmen. Der Beitrag zeigt, wie sich der industrielle Strukturwandel in der Entwicklung der Konjunkturindikatoren des Statistischen Bundesamtes niederschlägt und welche Herausforderungen daraus für die amtliche Statistik entstehen. Dabei knüpft er an die Ergebnisse einer neuen Umfrage an, in der das ifo Institut die Bedeutung der nichtindustriellen Tätigkeiten von Industrieunternehmen näher analysiert hat.

Der Beitrag ist wie folgt aufgebaut: Kapitel 2 gibt einen Überblick über die langfristige Entwicklung der deutschen Industrie, Kapitel 3 stellt die konzeptionellen Grundlagen und Erhebungsmethoden des Produktionsindex und des Produktionswertes dar und erläutert ihre statistischen Abgrenzungen. Im Kapitel 4 werden die in der Konjunkturstatistik ebenfalls erhobenen Umsätze vorgestellt, weil diese die Entwicklung der nichtindustriellen Umsätze verdeutlichen. Kapitel 5 vertieft die Rolle der nichtindustriellen Umsätze und zeigt deren Bedeutung für einzelne Wirtschaftszweige und Absatzmärkte auf. Kapitel 6 ergänzt die amtlichen Daten um Ergebnisse einer Sonderbefragung des ifo Instituts zu nichtindustriellen Tätigkeiten von Industrieunternehmen. Im Kapitel 7 erfolgt ein Vergleich der Ergebnisse des ifo Instituts mit den nichtindustriellen Umsätzen der Konjunkturstatistik. Kapitel 8 diskutiert die sta-

tistische Abbildung internationaler Produktionsverlagerungen am Beispiel der Automobilindustrie. Nach einer Zusammenfassung der zentralen Ergebnisse in Kapitel 9 werden in Kapitel 10 daraus Handlungsfelder für die Weiterentwicklung der amtlichen Unternehmensstatistik abgeleitet.

## 2

### Langfristige Entwicklung der Industrieaktivität

Die deutsche Industrie steckt in einer tiefgreifenden Transformation. Gemessen am monatlichen Produktionsindex wurden im Jahr 2024 im Verarbeitenden Gewerbe gut 9 % weniger industrielle Erzeugnisse hergestellt als 2015. Ein so deutlicher Rückgang der deutschen Industrieproduktion über viele Jahre hinweg war seit der deutschen Vereinigung und in Westdeutschland seit der Nachkriegszeit nie zu beobachten. [↗ Grafik 1](#) Vielfach wird vor diesem Hintergrund das Ende des deutschen Industriemodells und eine fortschreitende Deindustrialisierung diagnostiziert.

Der alleinige Blick auf den industriellen Produktionsindex zeigt jedoch nur einen Aspekt des Strukturwandels im Verarbeitenden Gewerbe. So hat sich seit Mitte der 2010er-Jahre die wirtschaftliche Aktivität der Industrieunternehmen, gemessen an dem in den Volkswirtschaftlichen Gesamtrechnungen ermittelten preisbereinigten Produktionswert, deutlich weniger negativ entwickelt. Er fiel im Jahr 2024 nur knapp 0,8 % niedriger aus als 2015.<sup>1</sup> [↗ Grafik 2](#)

Beide Indikatoren, Produktionsindex aus der Konjunkturstatistik und Produktionswert der VGR, messen die Entwicklung der Produktion im Verarbeitenden Gewerbe. Dass sich die Indikatoren dennoch unterschiedlich entwickeln, ergibt sich aus der jeweiligen Konstruktion der Statistiken. Im nächsten Kapitel werden die Konstruktionsprinzipien dargestellt, die bei der Berechnung der beiden Indikatoren zur Anwendung kommen.

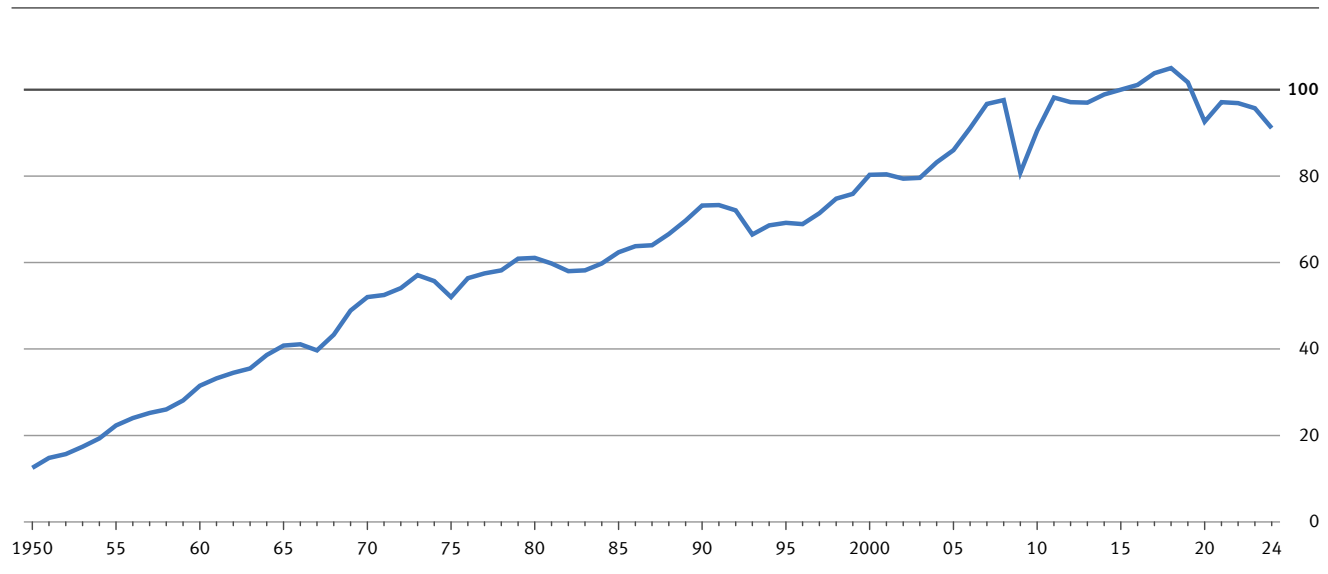
---

1 Der jüngste Wert für das Jahr 2024 zum Produktionswert wurde in den Volkswirtschaftlichen Gesamtrechnungen mittels Fortschreibung des Wertes aus dem Jahr 2023 anhand von verschiedenen unterjährig verfügbaren Indikatoren ermittelt, unter anderem mithilfe des Produktionsindex für das Verarbeitende Gewerbe.

**Grafik 1**

**Produktionsindex für das Verarbeitende Gewerbe**

2015 = 100

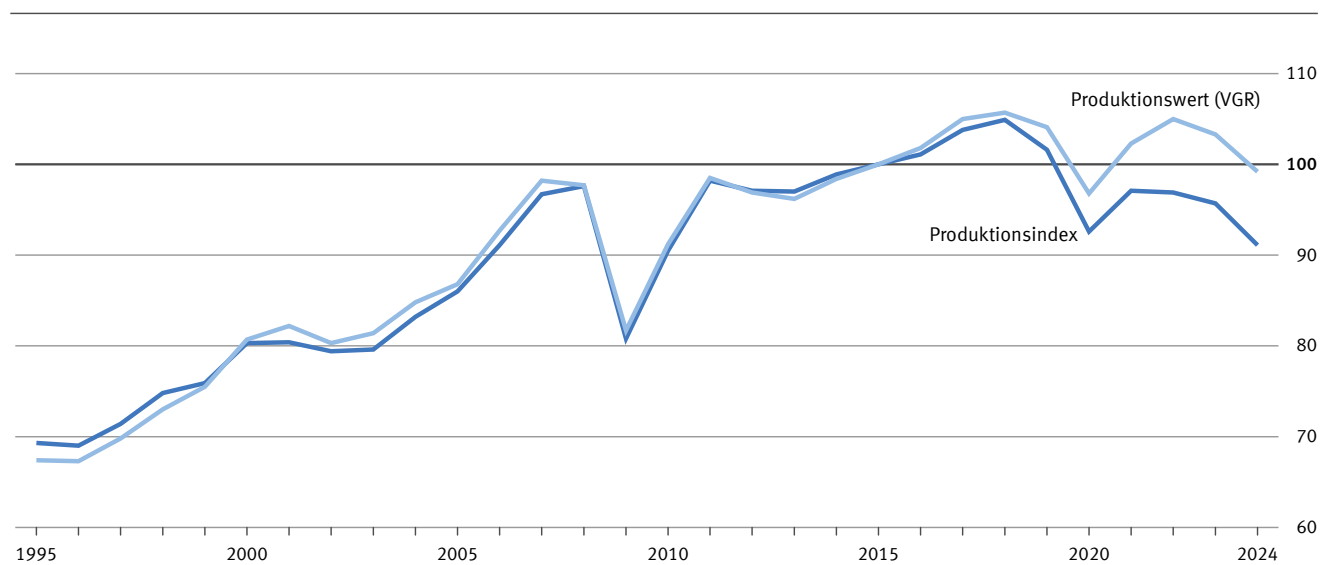


Verarbeitendes Gewerbe einschließlich Bergbau und Gewinnung von Steinen und Erden.  
Ergebnisse vor 1991 beziehen sich auf das frühere Bundesgebiet.  
Die Jahreswerte wurden als Durchschnitte der Monatswerte berechnet.

**Grafik 2**

**Produktionsindex und Produktionswert (VGR) für das Verarbeitende Gewerbe**

2015 = 100



Der Produktionswert der Volkswirtschaftlichen Gesamtrechnungen (VGR) ist ohne Umsatz aus Handelsware, jedoch mit dem Wert der Handelsspanne aus dem Verkauf von Handelsware ausgewiesen.

### 3

## Datenerhebung und Berechnungsmethoden

Der **Produktionsindex** für das Verarbeitende Gewerbe gibt die preisbereinigte Entwicklung des Wertes der Warenproduktion von Industriebetrieben an.<sup>12</sup> Die wichtigste Datengrundlage für den Produktionsindex ist die monatlich durchgeführte Produktionserhebung im Verarbeitenden Gewerbe.<sup>13</sup> In dieser geben die befragten Industriebetriebe unter anderem den Absatzwert der von ihnen produzierten Waren in Euro an. Die Daten werden monatlich bei allen Industriebetrieben ab 50 Beschäftigten und vierteljährlich bei Industriebetrieben ab 20 Beschäftigten erhoben.

Für die Berechnung des Produktionsindex werden die in der Produktionserhebung ermittelten Produktions-Absatzwerte über alle Betriebe eines Wirtschaftszweiges aufsummiert. Der aggregierte Absatzwert der Produktion des aktuellen Monats wird anschließend in Relation gesetzt zum entsprechenden Absatzwert im Basisjahr. Der so berechnete Index wird preisbereinigt und für vierstellige Klassen der Wirtschaftszweigklassifikation<sup>14</sup> ausgewiesen. Der Produktionsindex für das Verarbeitende Gewerbe insgesamt wird schließlich als gewichteter Mittelwert der Wirtschaftszweigindizes berechnet. Bei den Gewichten handelt es sich um nominale Bruttowertschöpfungsanteile der Wirtschaftszweige im entsprechenden Basisjahr, aktuell 2021. Die Wertschöpfungsanteile werden aus den Strukturhebungen im Verarbeitenden Gewerbe ermittelt.<sup>15</sup>

Der jährliche **Produktionswert** in den Volkswirtschaftlichen Gesamtrechnungen – im Folgenden und in den Grafiken: Produktionswert (VGR) – für das Verarbeitende Gewerbe erfasst den Gesamtwert aller im Inland hergestellten Waren und Dienstleistungen der

dem Verarbeitenden Gewerbe zugeordneten rechtlichen Unternehmen.<sup>16</sup> Dabei werden auch der Wert unfertiger eigener Erzeugnisse, Lagerbestandsveränderungen und die Handelsspanne berücksichtigt. Der Produktionswert wird in den Volkswirtschaftlichen Gesamtrechnungen größtenteils aus Umsätzen ermittelt. Die Erhebung der Inputdaten erfolgt im Zuge der jährlichen Strukturhebungen für das Verarbeitende Gewerbe.

Die Ableitung des Produktionswertes des Verarbeitenden Gewerbes nach den Konzepten der Volkswirtschaftlichen Gesamtrechnungen aus den jährlichen Strukturstatistiken erfordert methodische Zu- und Abschläge. Diese Anpassungen dienen dazu, konzeptionelle Unterschiede auszugleichen, insbesondere hinsichtlich der Behandlung von Steuern und Subventionen sowie der Einbeziehung informeller wirtschaftlicher Aktivitäten. Die quantitativ bedeutendste Korrektur stellt die Nettostellung der Handelsware dar, dabei wird der Produktionswert um den Einsatz an Handelsware reduziert. Dadurch werden durchlaufende Handelsaktivitäten ohne eigene Wertschöpfung ausgeblendet. Der Produktionswert wird in den Volkswirtschaftlichen Gesamtrechnungen als absoluter Euro-Betrag – nicht als Index – ausgewiesen und liegt ebenfalls preisbereinigt vor.

Ein wesentlicher Unterschied in der Konstruktion der beiden Indikatoren für die Produktionsentwicklung besteht in der breiteren Definition des Produktionsbegriffs in der VGR. Der Produktionsindex misst die Entwicklung des Realwertes der von den Industriebetrieben selbst produzierten Waren; Umsätze aus nichtindustriellen Tätigkeiten der Industriebetriebe in Deutschland bleiben weitgehend unberücksichtigt.<sup>17</sup> Der Produktionswert (VGR) bezieht sich hingegen auf das rechtliche Unternehmen insgesamt und berücksichtigt neben dem Umsatz aus eigener Erzeugung auch jenen aus sonstigen nichtindustriellen Tätigkeiten in Deutschland. Zum Beispiel umfasst der Produktionswert (VGR) auch Leistungen, die in den Unternehmen nicht unmittelbar zu Umsätzen

2 Informationen zum Produktionsindex und den Datenquellen siehe Linz und andere (2024) und Statistisches Bundesamt (2025a).

3 Informationen zu den Produktionserhebungen siehe Statistisches Bundesamt (2025b).

4 Die aktuell gültige Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008) steht unter [www.destatis.de](http://www.destatis.de) zur Verfügung.

5 Für Informationen zu den Strukturhebungen siehe Statistisches Bundesamt (2018).

6 Als Erhebungseinheit gilt die kleinste rechtliche Organisationseinheit des Unternehmens, die Bücher führt.

7 Eine Ausnahme bilden Umsätze aus der Reparatur und Installation von Maschinen und Ausrüstungen, die auch im industriellen Produktionsindex erfasst werden.

führen, wie selbsterstellte Forschungs- und Entwicklungsleistungen.<sup>18</sup>

Die konzeptionellen Unterschiede zwischen dem Produktionsindex und dem Produktionswert der Volkswirtschaftlichen Gesamtrechnungen ermöglichen unterschiedliche Interpretationen ihrer Ergebnisse. Abhängig vom Untersuchungsziel kann es daher sinnvoll sein, den einen oder den anderen Konjunkturindikator heranzuziehen. Der Produktionsindex eignet sich insbesondere zur Beurteilung der realen Produktionsaktivität. Er ist relevant für die Analyse der Entwicklung von Importen und Exporten, der in der Produktion geleisteten Arbeitsstunden und der dortigen Beschäftigung sowie des Material- und Energieverbrauchs oder der Verkehrsaktivität in Deutschland. Der Produktionswert der Volkswirtschaftlichen Gesamtrechnungen hingegen misst die Entwicklung der tatsächlich erwirtschafteten Leistung. Nach Abzug der Vorleistungen gibt er an, welcher wirtschaftliche Mehrwert in den Unternehmen geschaffen wurde und anschließend beispielsweise in Form von Löhnen, Gehältern oder Unternehmensgewinnen verteilt werden kann.

Im Folgenden werden die Umsätze aus nichtindustriellen Tätigkeiten der Industriebetriebe näher analysiert. Dabei werden sowohl die Ergebnisse der amtlichen Statistik als auch die Erkenntnisse einer Sonderbefragung des ifo Instituts herangezogen. Anschließend werden mögliche Geschäftsmodelle diskutiert, die den nichtindustriellen Umsätzen zugrunde liegen können. Die Verbindung beider Perspektiven soll dazu beitragen, die strukturellen Einflussfaktoren besser zu verstehen, die die unterschiedliche Entwicklung der beiden Konjunkturindikatoren erklären können.

---

<sup>8</sup> Informationen zu den Volkswirtschaftlichen Gesamtrechnungen siehe Statistisches Bundesamt (2025c).

## 4

### Monatliche Erhebung von Umsätzen in der Konjunkturstatistik

---

Der Unterschied zwischen der eigenen Produktions-tätigkeit und den sonstigen Tätigkeiten der Industriebetriebe wird deutlich, wenn man die in der monatlichen Konjunkturstatistik ebenfalls erhobenen Umsätze betrachtet. Die Statistik „Monatsbericht für Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden“ ermittelt monatliche Daten zu den Umsätzen der Industriebetriebe.<sup>19</sup> Die befragten Betriebe müssen die Umsätze differenziert nach Wirtschaftszweigen und getrennt nach Inland und Ausland sowie dem nicht zur Eurozone gehörenden Ausland angeben.

Zunächst geben die Industriebetriebe ihre Umsätze aus eigenen Erzeugnissen an, also die Umsätze aus dem Verkauf von Waren, die sie im eigenen Betrieb produziert haben. Diese bezeichnet die Statistik als „Umsätze der fachlichen Betriebsteile“. Darüber hinaus berichten die Betriebe unter der Bezeichnung „Umsätze der sonstigen Betriebsteile“ auch Umsätze, die sie durch nichtindustrielle Tätigkeiten erzielt haben. Solche sonstigen Umsätze können etwa aus Dienstleistungen oder dem Absatz von zugekaufter Handelsware entstehen. Auch Erlöse aus Lizenzverträgen, Provisionseinnahmen oder Einnahmen aus der Veräußerung von Patenten zählen zu den sonstigen Umsätzen.

Die Umsatzergebnisse der Erhebung werden als „Umsatz der Betriebe“ veröffentlicht. Dieser gliedert sich in den „Umsatz der fachlichen Betriebsteile“ und den „Umsatz der sonstigen Betriebsteile“. Für die weiteren Erläuterungen ist es wichtig klarzustellen, dass der Umsatz der Betriebe sowohl Umsätze aus dem Verkauf von eigenen Erzeugnissen als auch Umsätze aus sonstigen Tätigkeiten umfasst.<sup>10</sup> Die für einen Wirtschaftszweig ausgewiesene Gesamtsumme ent-

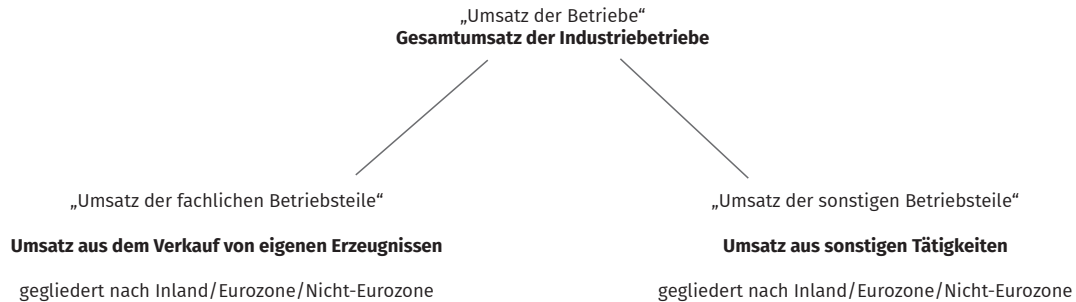
---

<sup>9</sup> Informationen zur Erhebung siehe Statistisches Bundesamt (2025d).

<sup>10</sup> Der monatlich zusammen mit den Auftragseingängen veröffentlichte „Umsatzindex für das Verarbeitende Gewerbe“ der Konjunkturstatistik bezieht sich ausschließlich auf die Umsätze der fachlichen Betriebsteile. Informationen zum Umsatzindex für das Verarbeitende Gewerbe siehe Statistisches Bundesamt (2025e).

### Grafik 3

Umsatz der Betriebe in der Statistik „Monatsbericht für Betriebe des Verarbeitenden Gewerbes“



hält damit sämtliche Umsätze aller diesem Wirtschaftszweig zugeordneten deutschen Industriebetriebe, unabhängig davon, ob sie aus dem Verkauf eigener Erzeugnisse verschiedener Produktbereiche, aus nichtindustriellen Dienstleistungen oder aus dem Handel mit Waren stammen. [➤ Grafik 3](#)

[➤ Grafik 4](#) stellt neben dem Produktionsindex und dem Produktionswert (VGR) auch den Umsatz der Betriebe dar. Weiterhin ist der „Umsatzindex für das Verarbeitende Gewerbe“ angegeben. Der monatlich veröffentlichte Umsatzindex bezieht sich ausschließlich auf die Umsätze der fachlichen Betriebsteile (Statistisches Bundesamt, 2024).

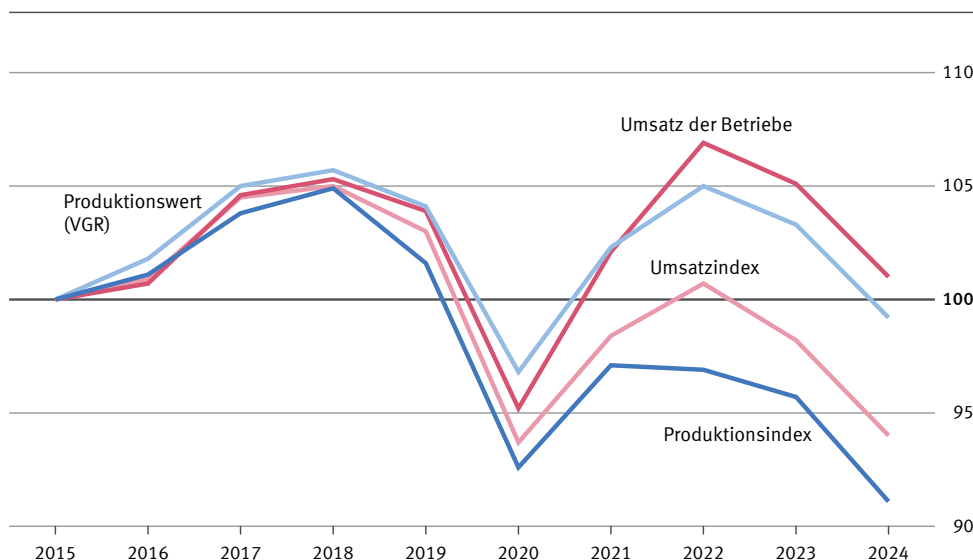
Alle Größen sind preisbereinigt. In den laufenden Veröffentlichungen des Statistischen Bundesamtes werden die Umsätze der Betriebe nicht preisbereinigt ausgewiesen. Daher wurde für diesen Aufsatz eine nachträgliche vereinfachte Preisbereinigung des Gesamttaggregats für das Verarbeitende Gewerbe anhand des impliziten Preisindex des Umsatzindex vorgenommen.<sup>11</sup>

Grafik 4 zeigt, dass der Umsatz der Betriebe insgesamt einem ähnlichen Entwicklungsverlauf folgt wie der Produktionswert (VGR). Sowohl der Produktions-

11 Der Umsatzindex wird als Wert- und Volumenindex veröffentlicht. Das Verhältnis der beiden Indizes kann hilfsweise für die Preisbereinigung des betrieblichen Umsatzes genutzt werden.

### Grafik 4

Preisbereinigte Kennziffern im Verarbeitenden Gewerbe  
2015 = 100



index als auch der Umsatzindex verlaufen hingegen deutlich schwächer. Die wachsende Differenz zwischen Umsatzindex und Produktionsindex auf der einen Seite und Produktionswert (VGR) sowie den gesamten Umsätzen der Betriebe auf der anderen Seite dürfte wesentlich auf einen Anstieg der sonstigen Umsätze der Industriebetriebe zurückzuführen sein. Es folgt daher ein näherer Blick auf die Sonstigen Umsätze. Die in Grafik 4 ebenfalls sichtbaren Unterschiede zwischen Produktions- und Umsatzindex ergeben sich einerseits aus methodischen Abgrenzungen, etwa bei der Behandlung innerbetrieblicher Weiterverarbeitung von selbst erstellten Waren. Andererseits bestehen erhebungstechnisch bedingte Unschärfen, zum Beispiel durch statistisch schwer trennbare Dienstleistungsanteile in den fachlichen Umsätzen.

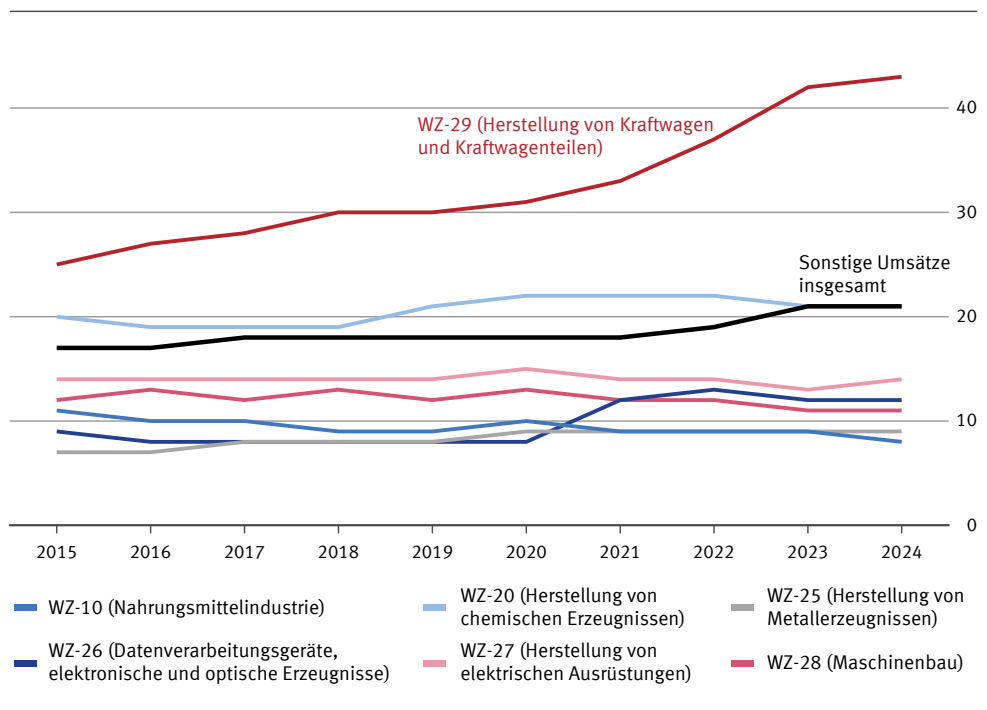
## 5

### Nichtindustrielle Umsätze der Industriebetriebe

➤ Grafik 5 zeigt für ausgewählte umsatzstarke Wirtschaftszweige des Verarbeitenden Gewerbes die Entwicklung des Anteils sonstiger, nichtindustrieller Umsätze am Gesamtumsatz der Betriebe. In vielen Wirtschaftszweigen liegt der Anteil sonstiger Umsätze auf einem eher niedrigen Niveau und variiert nur geringfügig über die Zeit. In den Wirtschaftszweigen WZ-20 (Chemische Industrie), WZ-24 (Metallerzeugung und -bearbeitung) sowie WZ-26 (Herstellung von Datenverarbeitungsgeräten, elektronischen und optischen Erzeugnissen) ist zwar eine leichte Zunahme der Anteile erkennbar, diese bleiben jedoch mit wenigen Prozentpunkten relativ gering. Auffällig hingegen ist der gewichtige Wirtschaftszweig WZ-29 (Herstellung von Kraftwagen

**Grafik 5**

**Anteile des Umsatzes aus sonstigen Tätigkeiten am Gesamtumsatz der Industriebetriebe nach ausgewählten Wirtschaftszweigen<sup>1</sup>**  
in %



<sup>1</sup> Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). - Ausgewählt sind die sieben WZ-2-Steller mit der höchsten Bruttowertschöpfung zu Faktorkosten und damit dem größten Gewicht im Produktionsindex.  
Nominale Umsatzanteile.



und Kraftwagenteilen). Die Automobilindustrie weist den höchsten Anteil sonstiger Umsätze auf und verzeichnet außerdem den stärksten Anstieg dieses Anteils. Die übrigen Branchen haben deutlich geringere Anteile und entwickeln sich meistens weniger dynamisch. Insgesamt steigen die Umsatzanteile, die nicht direkt aus der industriellen Warenherstellung entstehen, kontinuierlich an. Die schwarze Kurve stellt die Entwicklung der Umsätze aus sonstigen Tätigkeiten für das Verarbeitende Gewerbe insgesamt dar. Im Jahr 2024 lag der Anteil bei rund 21 % und damit gut 4 Prozentpunkte höher als im Jahr 2015. Dabei geht das Wachstum maßgeblich von den Betrieben der Kfz-Herstellung aus. Im Bereich der Automobilindustrie lag der Anteil der sonstigen Umsätze im Jahr 2024 bei 43 %.

↳ **Grafik 6** zeigt für die Kfz-Herstellung, zu welchen Anteilen die sonstigen Umsätze der deutschen Industriebetriebe im In- und Ausland erzielt wurden. Demnach ist der gestiegene Anteil der sonstigen Umsätze in der Automobilindustrie überwiegend auf das Wachstum der sonstigen Umsätze im Ausland zurückzuführen, hier sind insbesondere die Umsätze mit der Nicht-Eurozone relevant. Bei den Inlandsumsätzen bleibt der sonstige Umsatz hingegen nahezu

konstant bei etwa 10 % des gesamten betrieblichen Umsatzes.

Die Ergebnisse des Monatsberichts zeigen, dass ein zunehmender Teil der Umsätze von Industriebetrieben in Deutschland nicht mehr ausschließlich durch den Verkauf selbst produzierter physischer Güter erzielt wird. Die Verschiebung der Umsatzstruktur hin zu sonstigen Tätigkeiten geht größtenteils auf die Entwicklung in der Automobilindustrie zurück. Allerdings lassen weder die Daten des Monatsberichts für Betriebe des Verarbeitenden Gewerbes noch die der Strukturerhebung im Verarbeitenden Gewerbe eine Aussage darüber zu, mit welchen konkreten Wertschöpfungsmodellen diese Umsätze erzielt werden.

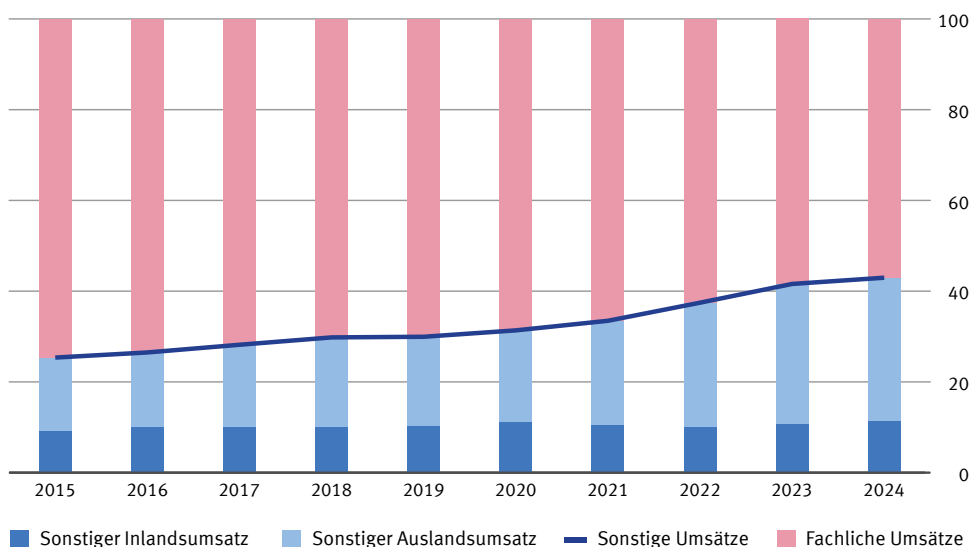
## 6

### Umfrage des ifo Instituts zu nichtindustriellen Umsätzen

Um näher zu analysieren, mit welchen Geschäftsmodellen die sonstigen, nichtindustriellen Umsätze generiert werden, führte das ifo Institut im Juli 2025 im Rahmen seiner Konjunkturumfrage eine Sonder-

**Grafik 6**

Umsätze im Wirtschaftszweig 29 Herstellung von Kraftwagen und Kraftwagenteilen  
in %



erhebung zur Umsatzstruktur im Verarbeitenden Gewerbe durch.<sup>12</sup> Die Fragen orientierten sich an einer früheren Erhebung des Statistischen Bundesamts aus dem Jahr 2002 (Opfermann, 2004) zur Verbreitung von produktnahen Dienstleistungen in Industrieunternehmen. In der Umfrage wurde nach der prozentualen Aufteilung der Umsätze im Jahr 2024 zwischen Umsätzen aus dem Verkauf von Waren aus eigener Herstellung, aus produktbegleitenden Dienstleistungen oder aus sonstigen, produktunabhängigen Tätigkeiten gefragt. Darüber hinaus wurden die Umfrageteilnehmer um eine Schätzung gebeten, wie sich diese Aufteilung vor zehn Jahren für das Jahr 2015 darstellte.

Im Bereich der **produktbegleitenden Dienstleistungen** wurde in der Umfrage nach ergänzenden Serviceleistungen gefragt, die Industrieunternehmen zusätzlich zum Verkauf der von ihnen gefertigten Waren anbieten. Im Fragebogen konnten folgende Kategorien angekreuzt werden

- > Planung, Beratung, Projektierung
- > Logistik, Montage, Inbetriebnahme
- > Maintenance, Wartung, Reparatur
- > Entsorgung, Recycling
- > Softwareprodukte, Datenprodukte
- > Leasing, Vermietung, Finanzierung
- > Schulungen
- > Sonstiges

Bei den Umsätzen aus **produktunabhängigen Tätigkeiten** der Industrieunternehmen wurden in den ankreuzbaren Kategorien unter anderem Umsätze aus dem Verkauf fremd hergestellter Waren, aus Lizenzfertigungen oder Headquarterleistungen als Kategorien angeboten.

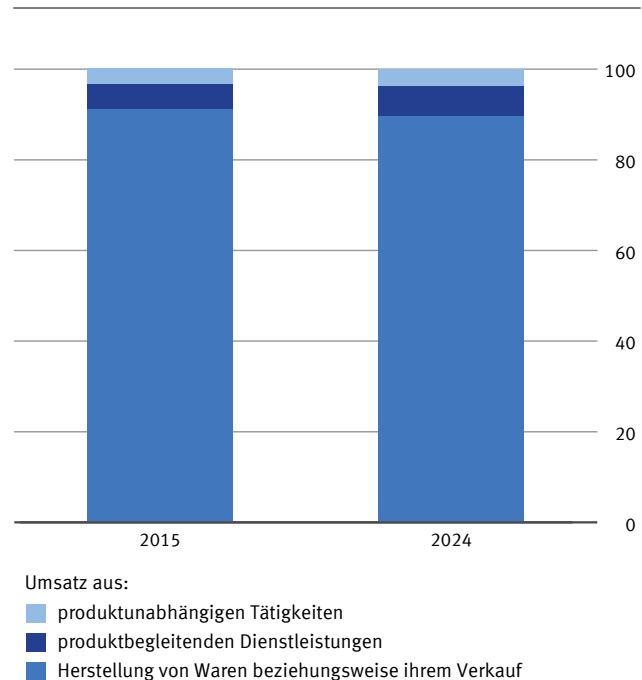
Insgesamt 1 680 Unternehmen gaben sowohl für das Jahr 2024 als auch für 2015 an, wie sie die Verteilung

zwischen Umsätzen aus eigenen Erzeugnissen, produktbegleitenden Dienstleistungen oder produktunabhängigen Tätigkeiten quantitativ einschätzen.

### ➤ Grafik 7

**Grafik 7**

Umsatzverteilung im Verarbeitenden Gewerbe gemäß ifo Konjunkturumfrage  
in %



Quelle: ifo Konjunkturumfragen im Verarbeitenden Gewerbe, Berichtsmonat Juli 2025

Die auf alle Unternehmen des Verarbeitenden Gewerbes hochgerechneten Ergebnisse zeigen, dass die Industrieunternehmen im Jahr 2024 nach wie vor den überwiegenden Anteil ihrer Umsätze mit der Herstellung von Waren und deren Verkauf erzielt haben (89,6 %). Im Vergleich zu 2015 (91,4 %) ist dieser Anteil jedoch etwas geschrumpft. Im selben Zeitraum ist der Anteil der sonstigen Umsätze von 8,6 % auf 10,4 % gestiegen. Betrachtet man die Verteilung innerhalb der sonstigen Umsätze, so sind die Umsätze mit produktbegleitenden Dienstleistungen von 5,3 % auf 6,8 % angewachsen. Der Anteil der Umsätze aus produktunabhängigen Tätigkeiten hat sich dagegen nur geringfügig von 3,4 % auf 3,6 % erhöht.

12 Das Panel der ifo Konjunkturumfragen im Verarbeitenden Gewerbe ist repräsentativ für das Verarbeitende Gewerbe in Deutschland mit einer leichten Überrepräsentation von Großunternehmen (Hiersemenzel und andere, 2022). In der Regel füllen Personen aus der Unternehmensleitung die Fragebogen aus (Sauer/Wohlrabe, 2019; Hennrich und andere, 2023). Für detaillierte Informationen zur Methodik der ifo Konjunkturumfragen siehe Sauer/Wohlrabe (2020) sowie Sauer und andere (2023).

➤ **Grafik 8** zeigt den jeweiligen Anteil der sonstigen Umsätze detaillierter für zusammengefasste Wirtschaftszweige des Verarbeitenden Gewerbes.<sup>13</sup> Auch die Verteilung der sonstigen Umsätze auf produktbegleitende und produktunabhängige Dienstleistungen innerhalb der Branchen ist angegeben.

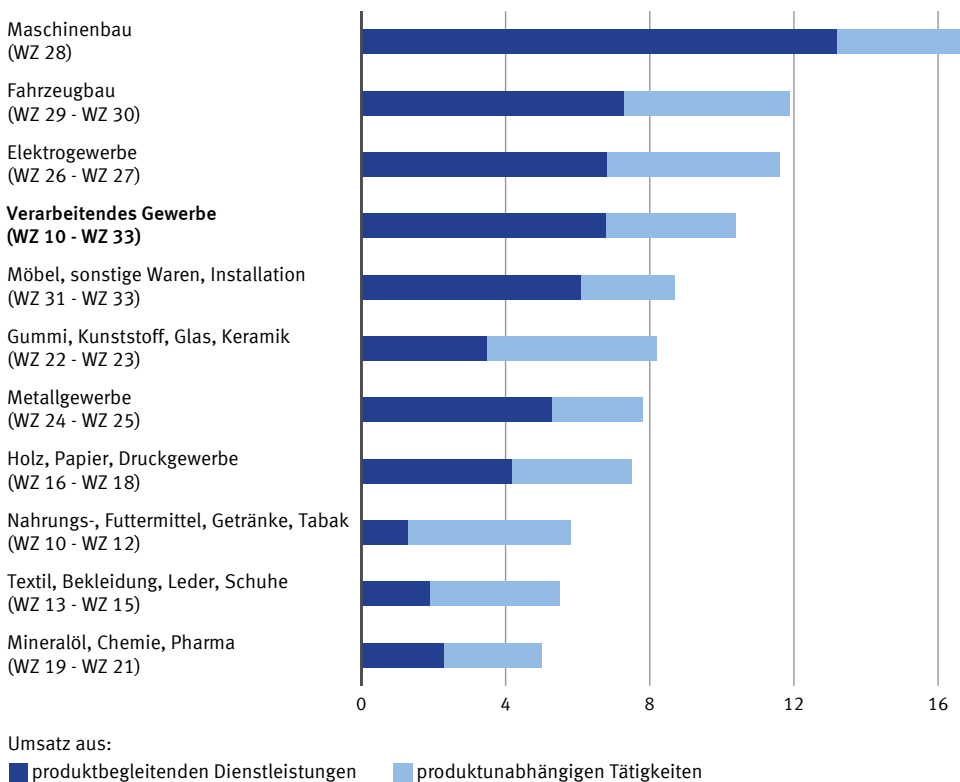
Insgesamt spielen sonstige Umsätze insbesondere im Maschinenbau, im Elektrogewerbe und im Fahrzeugbau eine größere Rolle. Der zusammengefasste Bereich Fahrzeugbau enthält sowohl die Herstellung von Kraftfahrzeugen und -teilen (WZ-29) als auch den sonstigen Fahrzeugbau (WZ-30), zu dem unter anderem Flugzeuge, Schiffe und Schienenfahrzeuge zählen.

13 Wegen teilweise geringer Antwort-Fallzahlen in den einzelnen Wirtschaftszweigen der WZ 2008 werden die Ergebnisse für zusammengefasste Wirtschaftszweige dargestellt.

Bei den produktbegleitenden Dienstleistungen sticht vor allem der Maschinenbau mit einem Umsatzanteil von 13,2 % heraus. Ansonsten liegen bezüglich der produktbegleitenden Dienstleistungen lediglich der Fahrzeugbau (7,3 %) sowie die Elektroindustrie (6,8 %) über dem Durchschnitt des gesamten Verarbeitenden Gewerbes. Im Maschinenbau und der Elektroindustrie wurden in der Umfrage des ifo Instituts Maintenance und Wartung sowie Logistik, Montage und Inbetriebnahme, aber auch Schulungen sowie Softwareprodukte am häufigsten genannt. Die Planung, Beratung und Projektierung kommt am häufigsten im Fahrzeugbau vor. Ebenso sind Entsorgung und Recycling sowie Leasing vorwiegend Geschäftsmodelle des Fahrzeugbaus, wobei Leasing vor allem bei den Kfz-Herstellern zu finden ist und eher seltener bei Zulieferern. In den Bereichen Nahrungs- und Genussmittel oder dem Textil- und Bekleidungs-

#### Grafik 8

Verteilung der sonstigen Umsätze im Verarbeitenden Gewerbe nach Wirtschaftszweigen<sup>1</sup>  
in %



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

Quellen: ifo Konjunkturumfragen im Verarbeitenden Gewerbe, Berichtsmonat Juli 2025

gewerbe spielen produktbegleitende Dienstleistungen dagegen nur eine sehr untergeordnete Rolle.

Bezüglich der Umsätze aus produktunabhängigen Tätigkeiten schwanken die Anteile zwischen den verschiedenen Industriebereichen weniger, mit den durchschnittlich höchsten Werten im Elektrogewerbe (4,8 %) und niedrigsten Anteilen im Metallgewerbe (2,5 %). Anders als bei den produktbegleitenden Dienstleistungen fallen sonstige Umsätze auch im Nahrungs- und Genussmittelbereich (4,5 %) überdurchschnittlich an, da hier über alle Sparten hinweg am häufigsten sonstige fremd hergestellte Handelswaren verkauft werden.<sup>14</sup>

14 Eine ausführliche Darstellung der Ergebnisse findet sich in Sauer/Wollmershäuser (2025).

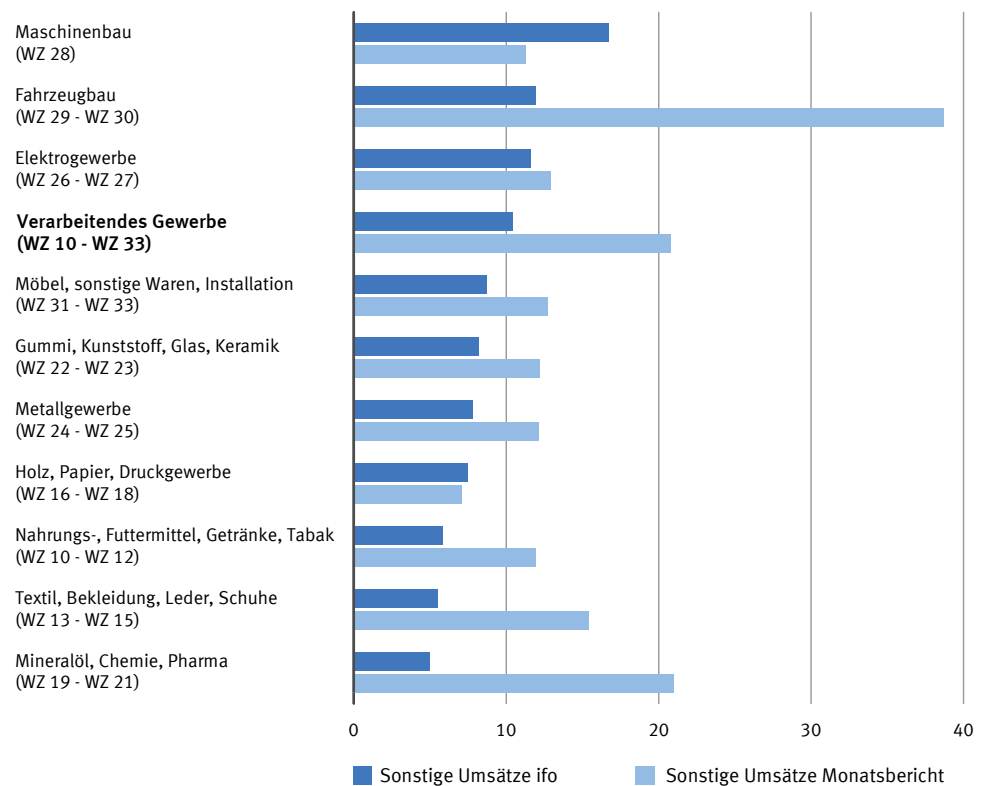
## 7

### Vergleich der Ergebnisse der ifo-Befragung und des Monatsberichts zu nichtindustriellen Umsätzen

Vergleicht man den Anteil der sonstigen Umsätze aus der Befragung des ifo Instituts mit den im Monatsbericht für Betriebe des Verarbeitenden Gewerbes ermittelten Anteilen, so zeigen sich deutliche Unterschiede. Der Anteil der sonstigen Umsätze liegt im Monatsbericht in beiden Jahren 2015 und 2024 nahezu doppelt so hoch wie in der ifo-Umfrage. Dort entfallen 2024 rund 21 % der gemeldeten Umsätze auf sonstige Umsätze, während in der ifo-Umfrage ein Anteil von etwa 10 % ermittelt wurde.

**Grafik 9**

Anteile der sonstigen Umsätze im Verarbeitenden Gewerbe nach Wirtschaftszweigen<sup>1</sup> in %



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

Quellen: Monatsbericht für Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden; ifo Konjunkturumfragen im Verarbeitenden Gewerbe

➤ **Grafik 9** stellt die sonstigen Umsätze aus der ifo Konjunkturumfrage und dem Monatsbericht für Betriebe für das Verarbeitende Gewerbe insgesamt und für zusammengefasste Wirtschaftszweige dar. Sie zeigt, dass sich die Branchen im Hinblick auf die Unterschiedlichkeit der Anteile sonstiger Umsätze in den beiden Ansätzen stark unterscheiden. Vor allem im Maschinenbau wird der Anteil in der ifo Konjunkturumfrage höher eingeschätzt als im Monatsbericht für Betriebe. In fast allen anderen Bereichen ist jedoch der Anteil sonstiger Umsätze im Monatsbericht höher und zum Teil deutlich höher.

Bei der Interpretation der Abweichungen sind zunächst konzeptionelle Unterschiede zwischen den Erhebungen zu berücksichtigen. Während die ifo Konjunkturumfragen Unternehmen erfassen und auf freiwilliger Teilnahme beruhen, bezieht sich der Monatsbericht für das Verarbeitende Gewerbe auf Betriebe mit Berichtspflicht ab 20 Beschäftigten. Der Berichtskreis umfasst hier rund 22 000 Betriebe, während in der ifo-Befragung die Zahl der einbezogenen Einheiten mit 1 680 Unternehmen deutlich niedriger ist. Zudem unterscheiden sich die Erhebungsmethoden in der Art der Fragestellung – die ifo-Befragungen fragen Umsatz-Anteilswerte ab, während im Monatsbericht absolute Werte zu den verschiedenen Umsatzkategorien anzugeben sind. Trotz dieser konzeptionellen Unterschiede gehen die Autoren aufgrund der Stärke der Abweichungen davon aus, dass im Monatsbericht für Betriebe Umsatzanteile erfasst werden, die sich in der ifo Konjunkturumfrage systematisch nicht widerspiegeln. Besonders auffällig sind die wesentlich höheren Anteile der sonstigen Umsätze des Monatsberichts im Fahrzeugbau, hier dürfte der im Fahrzeugbau dominierende Bereich der Kfz-Herstellung die Unterschiede verursachen. Das folgende Kapitel skizziert für den Bereich der Automobilindustrie Erklärungsansätze für den Anstieg der sonstigen Umsätze im Monatsbericht.

## 8

### Statistische Abbildung von Produktionsverlagerungen in der Automobilindustrie

---

Über die genannten Unterschiede hinaus ist eine zumindest durch bekannte Einzelfälle aus der Automobilindustrie gestützte Erklärung für den Anstieg der sonstigen Auslandsumsätze in den Ergebnissen des Monatsberichts im Wirtschaftszweig WZ-29 die Verlagerung von Produktionsstandorten.<sup>15</sup> Arbeitsintensive und standardisierte Fertigungsschritte werden zunehmend in kostengünstigere Länder verlagert. Gleichzeitig wird von einigen Automobilherstellern die Strategie verfolgt, die Produkte möglichst nahe an den Absatzmärkten zu fertigen, auch als Reaktion auf handelspolitische Rahmenbedingungen. In Deutschland verbleiben hingegen zentrale Funktionen der global aufgestellten Unternehmen, wie Teile der Forschung und Entwicklung, Software- und Datenkompetenzen oder besonders hochwertige und komplexe Produktionsaktivitäten. Ebenso dürften zentrale Unternehmensfunktionen wie Konzernsteuerung, Qualitätssicherung, Einkauf oder Lieferkettenmanagement zum Teil weiter an Standorten in Deutschland ausgeübt werden.

Für die statistische Abbildung dieser Produktionsverlagerungen ist eine Unterscheidung zwischen der Tätigkeit von den in Deutschland ansässigen Konzernteilen und deren Finanzierung relevant. Denn mit der Ausübung der zentralen Unternehmensfunktionen werden oft nicht direkt Umsätze generiert. Die Aktivitäten in Deutschland werden vielmehr zum Teil durch die internationale Wertschöpfung der Konzerne querfinanziert. Die dabei entstehenden Zahlungsströme zwischen den deutschen und ausländischen Konzernteilen können unterschiedlich organisiert sein. Zum Beispiel können zentrale Entwicklungs- und Managementleistungen über interne Verrechnungsströme finanziert werden oder die zen-

---

15 Unternehmen des Verarbeitenden Gewerbes lagern laut einer Sonderauswertung der Statistik zu globalen Wertschöpfungsketten (Baum, 2026) am häufigsten die Unternehmensfunktion „Produktion von Waren“ aus. Als wesentliche Gründe werden vor allem niedrigere Lohnkosten sowie weitere Kostenvorteile genannt.

tral entwickelten Technologien werden über Lizenzgebühren abgerechnet. Weiterhin sind Fälle bekannt, in welchen deutsche Unternehmen die Produkte der Konzerntochter im Ausland ankaufen und selbst wiederverkaufen, ohne dass dabei die Produkte physisch nach Deutschland verbracht würden. Die Handelsmarge fällt dann in Deutschland an und kann für die Finanzierung der inländischen Unternehmensfunktionen genutzt werden.

Die fehlenden direkten Umsätze aus den zentralen Aktivitäten in Deutschland haben möglicherweise dazu geführt, dass in der Befragung des ifo Instituts im Bereich der Automobilindustrie andere Anteile für die sonstigen Umsätze angegeben wurden als sie im Monatsbericht für Betriebe des Verarbeitenden Gewerbes beobachtet werden. Zum Beispiel könnten Zahlungsströme aus der internen Verrechnung unterschiedlich interpretiert worden sein. Die genannten Erklärungsansätze lassen sich anhand der vorliegenden Daten jedoch nicht abschließend und vollständig verifizieren.

## 9

### Zusammenfassung

Ausgangspunkt des Aufsatzes war die wachsende Divergenz zwischen dem industriellen Produktionsindex aus den Konjunkturstatistiken und dem Produktionswert für das Verarbeitende Gewerbe der Volkswirtschaftlichen Gesamtrechnungen. Während der Produktionsindex eine deutliche Abschwächung des Realwertes der industriellen Warenproduktion signalisiert, zeigt der Produktionswert (VGR) eher eine Stagnation seit Mitte der 2010er-Jahre an. Als zentrale Ursache für die wachsende Divergenz wurde bereits in einer vorausgegangenen gemeinsamen Untersuchung des ifo Instituts und des Statistischen Bundesamtes die zunehmende Bedeutung von nichtindustriellen Tätigkeiten der Industrieunternehmen in Deutschland identifiziert (Lehmann und andere, 2025). Die wachsenden Umsätze aus nichtindustriellen Tätigkeiten schlagen sich in den beiden Indikatoren unterschiedlich nieder. Der Produktionswert (VGR) erfasst die gesamte wirtschaftliche Aktivität der Industrieunternehmen in Deutschland

und damit auch Umsätze aus Dienstleistungen oder anderen nichtindustriellen Tätigkeiten. Der Produktionsindex hingegen misst die preisbereinigte Entwicklung des Absatzwertes der von den Industriebetrieben im Inland selbst produzierten Industriewaren und schließt damit (mit Ausnahme von Montage und Reparaturen) Umsätze aus sonstigen Tätigkeiten der Betriebe aus. Abhängig vom Untersuchungsziel kann es daher sinnvoll sein, den einen oder den anderen Konjunkturindikator heranzuziehen. Der Produktionsindex eignet sich insbesondere zur Beurteilung der realen Produktionsaktivität. Der Produktionswert (VGR) hingegen misst die Entwicklung der tatsächlich erwirtschafteten Leistung. Nach Abzug der Vorleistungen gibt er an, welcher wirtschaftliche Mehrwert geschaffen wurde.

Die zunehmende Bedeutung der Umsätze, die nicht aus dem Verkauf von selbst gefertigten Waren stammen, kann mithilfe der betrieblichen Umsätze aus dem Monatsbericht für Betriebe des Verarbeitenden Gewerbes quantifiziert werden, da hier sowohl die Umsätze aus eigenen Erzeugnissen als auch aus sonstigen Umsätzen ausgewiesen werden. Die sonstigen Umsätze sind vor allem bei den Betrieben im Wirtschaftszweig der Kfz-Herstellung angestiegen. Die im Inland ansässigen Kfz-Produktionsbetriebe erzielen diese sonstigen Umsätze vor allem durch Absatz ihrer Leistungen in das Ausland außerhalb der Eurozone.

In einer Sonderbefragung des ifo Instituts haben die Industrieunternehmen in Deutschland angegeben, dass der überwiegende Teil der sonstigen Umsätze aus produktbegleitenden Dienstleistungen stammt, etwa aus Montage, Wartung, Schulungen, Software, Planung oder Leasingangeboten. Der Vergleich zwischen den Befragungsergebnissen des ifo-Instituts und den Ergebnissen des Monatsberichts der amtlichen Statistik zeigt jedoch erhebliche Unterschiede in der Höhe der sonstigen Umsätze, was auf Abgrenzungs- und Interpretationsprobleme bei der Erfassung der sonstigen Umsätze hinweist.

Für die Automobilindustrie könnte der Anstieg der sonstigen Umsätze im Monatsbericht möglicherweise im Kontext internationaler Produktionsverlagerungen erklärt werden. Während standardisierte Fertigungsstufen zunehmend im Ausland erfolgen,



verbleiben zentrale Unternehmensfunktionen wie Forschung und Entwicklung, Software- und Systemkompetenzen, Konzernsteuerung sowie Teile hochwertiger Produktion in Deutschland. Diese Aktivitäten generieren häufig keine unmittelbaren Markterlöse, werden jedoch über die internationale Wertschöpfung der Konzerne finanziert. Die dabei entstehenden konzerninternen Zahlungsströme könnten sich im Monatsbericht für Betriebe des Verarbeitenden Gewerbes und in der Umfrage des ifo Instituts unterschiedlich niederschlagen.

Insgesamt verdeutlicht der Beitrag, dass die beobachteten Unterschiede zwischen Produktionsindex und Produktionswert (VGR) Ausdruck eines Strukturwandels in der Industrie sind, bei dem die Industrieunternehmen in Deutschland zunehmend Umsätze mit Aktivitäten erzielen, die nicht unmittelbar aus dem Verkauf von in Deutschland hergestellten Waren stammen. Die Daten der amtlichen Statistik geben zwar Auskunft über den Gesamtumfang dieser Aktivitäten, sie können jedoch nicht die den Aktivitäten zugrunde liegenden Geschäftsmodelle beschreiben.

## 10

### Handlungsfelder für die amtliche Statistik

Für die amtliche Statistik ergibt sich der Bedarf, die sonstigen Umsätze der Industrieunternehmen genauer nach Geschäftsmodellen zu differenzieren und die Zuordnung zu Industrie, Dienstleistungen und Handel zu schärfen.

Derzeit werden die Unternehmen (Strukturerhebungen) oder Betriebe (Konjunkturerhebungen im Verarbeitenden Gewerbe) für die Datenerhebung zunächst nach ihrem Umsatzschwerpunkt einem Wirtschaftszweig zugeordnet. Die Umsätze werden dann jeweils innerhalb der Wirtschaftszweige erfasst. Dabei werden Umsätze aus branchenfremden Aktivitäten eines Unternehmens entweder dem Wirtschaftszweig zugeschlagen, in dem das Unternehmen seinen Schwerpunkt hat, wie bei den Strukturstatistiken. Oder die branchenfremden Umsätze werden nur nachrichtlich erfasst und gehen nicht in die industriellen Produktions- und Umsatzindizes ein.

Die **unternehmensbezogene Betrachtung**, bei der Unternehmen auf Basis ihrer dominierenden wirtschaftlichen Tätigkeit den Sektoren zugeordnet werden, ist für zahlreiche Fragestellungen sachgerecht und methodisch etabliert. Sie erlaubt jedoch nur eingeschränkt Aussagen darüber, in welchen Tätigkeitsfeldern Umsätze tatsächlich entstehen. Vor dem Hintergrund zunehmender funktionaler Differenzierung innerhalb von Unternehmen sollte daher ergänzend eine **tätigkeitsbezogene Betrachtung** der Umsätze erfolgen. Diese ermöglicht es, industrielle, handelsbezogene und dienstleistungsorientierte Aktivitäten innerhalb der Unternehmen sichtbar zu machen und strukturelle Verschiebungen der Wertschöpfung innerhalb der Unternehmen differenzierter zu analysieren.

Die tätigkeitsbezogene Aufteilung der Umsätze muss gemäß der Klassifikation der Wirtschaftszweige erfolgen, die neben dem Verarbeitenden Gewerbe auch die Bereiche Handel und Dienstleistungen sowie weitere Bereiche umfasst. Im Falle eines Industrieunternehmens könnten zum Beispiel die sonstigen Umsätze eines Kfz-Herstellers dem Bereich Handel zugeordnet werden, wenn sie wie oben beschrieben durch Ankauf und Verkauf von Kraftwagen erzielt wurden. In der WZ 2008 wäre das der Bereich WZ-45.1 Handel mit Kraftwagen. Ebenso ließen sich Umsätze aus anderen Aktivitäten etwa dem Wirtschaftszweig WZ-70.10 Verwaltung und Führung von Unternehmen und Betrieben, dem Wirtschaftszweig WZ-72 Forschung und Entwicklung oder dem Wirtschaftszweig WZ-62.01 Programmierungstätigkeiten zuordnen.

Durch eine solche Erfassung der sonstigen Umsätze kann in den Konjunkturstatistiken erreicht werden, dass die Umsätze eines Industriebetriebs gegebenenfalls zum Teil in den industriellen Produktionsindex und zu einem Teil auch in den Dienstleistungsproduktionsindex eingehen.<sup>16</sup> Ebenso muss in den Strukturstatistiken ergänzend auch eine Aufteilung der branchenfremden sonstigen Umsätze erfolgen. Dies ist die Voraussetzung dafür, dass auch bei der Berechnung der Jahresergebnisse zur Bruttowertschöpfung in den Volkswirtschaftlichen Gesamtrechnungen die Umsätze korrekt zwischen den Wirtschaftsbereichen aufgeteilt werden können. Wenn

---

16 Zum Dienstleistungsproduktionsindex siehe Statistisches Bundesamt (2023).

eine solche Umsatzaufteilung der Unternehmensumsätze auf die Bereiche Industrie, Handel und Dienstleistungen durchgehend sowohl in den Konjunkturstatistiken als auch in den Strukturstatistiken und den Volkswirtschaftlichen Gesamtrechnungen erfolgt, dürften sich die Produktionsindizes der Konjunkturstatistiken, die tätigkeitsbezogenen Jahresergebnisse der Strukturerhebungen und Bruttowertschöpfungsgrößen der Wirtschaftsbereiche der Volkswirtschaftlichen Gesamtrechnungen kohärent entwickeln.

Das Weiterentwicklungsprojekt „System der Unternehmensstatistiken“ (SysdU) strebt eine verbesserte statistische Abbildung der sonstigen Umsätze der Industrieunternehmen an. Der Reformansatz SysdU nimmt das gesamte System der Unternehmensstatistiken in den Blick, um eine bessere Kohärenz zwischen Struktur- und Konjunkturstatistiken sowie eine effizientere Statistikproduktion sicherzustellen (Bundesrat, 2025). [u](#)



## LITERATURVERZEICHNIS

---

Baum, Joschka. [\*Die neue Statistik zu globalen Wertschöpfungsketten\*](#).

In: WISTA 1/2026, Seite 28 ff.

Bundesrat. *Plenarprotokoll 1060. Sitzung vom 19. Dezember 2025*. Anlage 12: Erklärung von Ministerin Dr. Kathrin Wahlmann (Niedersachsen) zu Punkt 43 der Tagesordnung. Berlin 2025, Seite 537 f. [Zugriff am 19. Januar 2026]. Verfügbar unter: [www.bundesrat.de](http://www.bundesrat.de)

Hennrich, Jonas/Sauer, Stefan/Wohlrabe, Klaus. *Wer berichtet von der Stimmung in den deutschen Chefetagen?* In: ifo Schnelldienst. Jahrgang 76. Ausgabe 08/2023, Seite 37 ff.

Hiersemenzel, Magdolna/Sauer, Stefan/Wohlrabe, Klaus. *Zur Repräsentativität der ifo Konjunkturumfragen*. In: ifo Schnelldienst. Jahrgang 75. Ausgabe 07/2022, Seite 26 ff.

Lehmann, Robert/Linz, Stefan/Wollmershäuser, Timo. *Strukturwandel im Verarbeitenden Gewerbe: Hybride Produkte und Produktionsprozesse verändern industrielle Geschäftsmodelle*. ifo Schnelldienst digital. Jahrgang 6. Ausgabe 5/2025, Seite 1 ff.

Linz, Stefan/Flores, Luis Federico/Bolz, Maria/Schächer, Jennifer/Eid, Nicole. [\*Umstellung des Produktionsindex im Produzierenden Gewerbe auf das Basisjahr 2021\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 2/2024, Seite 55 ff.

Opfermann, Rainer. [\*Produktbegleitende Dienstleistungen und ihre statistische Erfassung\*](#). In: Wirtschaft und Statistik. Ausgabe 3/2004, Seite 269 ff.

Sauer, Stefan/Wohlrabe, Klaus. *Chef oder Praktikant – wer beantwortet eigentlich die Fragebögen in den ifo Konjunkturumfragen?* In: ifo Schnelldienst. Jahrgang 72. Ausgabe 03/2019, Seite 30 ff.

Sauer, Stefan/Wohlrabe, Klaus. *ifo Handbuch der Konjunkturumfragen*. ifo Beiträge zur Wirtschaftsforschung, Band 88. München 2020.

Sauer, Stefan/Schasching, Moritz/Wohlrabe, Klaus. *Handbook of ifo Surveys*. ifo Beiträge zur Wirtschaftsforschung, Band 100. München 2023.

Sauer, Stefan/Wollmershäuser, Timo. *Strukturwandel im Verarbeitenden Gewerbe: Produktbegleitende Dienstleistungen gewinnen an Bedeutung*. In: ifo Schnelldienst. Jahrgang 78. Ausgabe 12/2025, Seite 64 ff.

Statistisches Bundesamt. [\*Klassifikation der Wirtschaftszweige 2008\*](#). Wiesbaden 2008.

Statistisches Bundesamt. [\*Produzierendes Gewerbe: Strukturerhebung im Verarbeitenden Gewerbe sowie des Bergbaus und der Gewinnung von Steinen und Erden\*](#). Qualitätsbericht. 2018.

### LITERATURVERZEICHNIS

---

Statistisches Bundesamt. [\*Konjunkturstatistik im Dienstleistungsbereich. Dienstleistung, Gastgewerbe 2023\*](#). Qualitätsbericht. 2023.

Statistisches Bundesamt. [\*Produktionsindex im Produzierenden Gewerbe. Basis 2021 = 100\*](#). Qualitätsbericht. 2025a.

Statistisches Bundesamt. [\*Produktionserhebungen 2025\*](#). Qualitätsbericht. 2025b.

Statistisches Bundesamt. [\*Volkswirtschaftliche Gesamtrechnungen VGR 10/2024-09/2025\*](#). Qualitätsbericht. 2025c.

Statistisches Bundesamt. [\*Monatsbericht für Betriebe des Verarbeitenden Gewerbes sowie des Bergbaus und der Gewinnung von Steinen und Erden 2025\*](#). Qualitätsbericht. 2025d.

Statistisches Bundesamt. [\*Umsatzindex im Verarbeitenden Gewerbe. Basis 2021=100, 2024\*](#). Qualitätsbericht. 2025e.



**Joschka Baum**

ist Soziologe und im Referat „Analyse der Unternehmensstrukturen“ des Statistischen Bundesamtes für die Statistik zu globalen Wertschöpfungsketten verantwortlich.

# DIE NEUE STATISTIK ZU GLOBALEN WERTSCHÖPFUNGSKETTEN

Joschka Baum

➤ **Schlüsselwörter:** Globalisierung – Verlagerung – Beschäftigungsabbau – Unternehmensfunktionen – internationaler Waren- und Dienstleistungsaustausch

## ZUSAMMENFASSUNG

Nach mehreren Piloterhebungen wurde die Statistik zu globalen Wertschöpfungsketten für den Berichtszeitraum 2021 bis 2023 erstmals als Erhebung mit Auskunftspflicht durchgeführt. Sie stellt Informationen über die Einbindung deutscher Unternehmen in globale Wertschöpfungsketten bereit, untergliedert für verschiedene Waren- und Dienstleistungsarten. Zudem untersucht sie die Verlagerung von Unternehmensfunktionen ins Ausland und damit einhergehende Beschäftigungseffekte im Inland. Darüber hinaus berichten Unternehmen über die Motive und Hindernisse der Verlagerung sowie den Effekt globaler Ereignisse auf ihre globalen Wertschöpfungsketten. Der Artikel beschreibt die Inhalte und Methodik der Statistik, stellt ausgewählte Ergebnisse vor und gibt einen Ausblick auf künftige Entwicklungen.

➤ **Keywords:** *globalisation – international sourcing – job losses – business functions – international trade in goods and services*

## ABSTRACT

*Following multiple pilot surveys, statistics on global value chains have been compiled for the first time based on a mandatory response survey and cover the reference period 2021 to 2023. The statistics provide information on the participation of German enterprises in global value chains, broken down by various types of goods and services. In addition, the statistics examine the relocation of business functions abroad and the associated employment effects in Germany. The enterprises also report on motivations for and barriers to international sourcing as well as information on the impact that global events have on their global value chains. The article describes the content and methodology of the statistics, presents selected results and provides an outlook on future developments.*

### 1

## Einleitung

In den letzten Jahrzehnten haben sich die Wertschöpfungsketten zunehmend globalisiert. In jüngerer Zeit gab es jedoch verschiedene Entwicklungen, die den internationalen Handel spürbar beeinträchtigten: Die Corona-Pandemie führte zu Lieferengpässen, die Europäische Union (EU) verhängte im Zusammenhang mit dem Angriff Russlands auf die Ukraine Sanktionen gegen Russland, die Vereinigten Staaten führten Zölle auf bestimmte importierte Waren aus anderen Ländern ein. Deutschland ist überdurchschnittlich stark in den internationalen Handel eingebunden (Ullrich, 2022). Dadurch entsteht ein erhöhter Bedarf an belastbaren Daten als Grundlage für politische Diskussionen und Entscheidungen, zum Beispiel in der Wirtschafts- und Handelspolitik, um geopolitische Abhängigkeiten zu bewerten.

Um diesen Datenbedarf zu decken, hat die EU die Statistik zu globalen Wertschöpfungsketten (Global Value Chains, GVC) eingeführt. Die europäischen Rechtsgrundlagen bilden die Verordnung (EU) 2019/2152 sowie die Durchführungsverordnung (EU) 2022/918. Im Rahmen der Statistik werden Informationen zu globalen Wertschöpfungsketten und Verlagerungen erhoben und Eurostat, dem Statistischen Amt der Europäischen Union, zur Verfügung gestellt. Nationale Rechtsgrundlage in Deutschland ist das Gesetz über die Statistik zu globalen Wertschöpfungsketten vom 17. Januar 2024, das eine Stichprobenerhebung mit Auskunftspflicht vorsieht. Vergangene Piloterhebungen für die Berichtszeiträume 2001 bis 2006 (Statistisches Bundesamt, 2008), 2014 bis 2016 (Statistisches Bundesamt, 2019; Kaus, 2019) sowie 2018 bis 2020 wurden noch als freiwillige Erhebungen nach § 7 Bundesstatistikgesetz durchgeführt.

Der Artikel stellt zunächst in Kapitel 2 die Inhalte der Statistik dar. Kapitel 3 erläutert die Erhebungs- und Auswertungsmethodik. Kapitel 4 erörtert Ergebnisse der aktuellen Erhebung im Hinblick auf Verlagerungen und globale Wertschöpfungsketten und widmet sich speziell der Thematik der fabriklosen Produktion. Das letzte Kapitel gibt einen Ausblick auf die künftige Entwicklung der Statistik.

### 2

## Inhalte der Statistik

➤ **Übersicht 1** auf Seite 30 zeigt die Struktur des Fragebogens, der bei der Erhebung der Statistik zu globalen Wertschöpfungsketten verwendet wird. Die Inhalte der Befragung lassen sich in zwei Themenkomplexe teilen: globale Wertschöpfungsketten und Verlagerung ins Ausland.

### 2.1 Globale Wertschöpfungsketten

Der erste Themenkomplex fragt nach der Einbindung von Unternehmen in globale Wertschöpfungsketten im Berichtsjahr und wird durch vier Frageblöcke erfasst. Die ersten beiden Frageblöcke beziehen sich auf den Bezug von Waren und Dienstleistungen, die beiden anderen auf deren Bereitstellung. Dabei werden sechs Warenarten und sieben Dienstleistungsarten unterschieden. Zusätzlich erfolgt die Unterscheidung danach, ob die Handelspartner für die jeweilige Dienstleistungs- oder Warenart in der EU, im übrigen Europa oder außerhalb Europas ansässig sind. Erfragt wird, ob ein Unternehmen die jeweilige Warenart oder Dienstleistungsart im Wert von über 100 000 Euro für ein Unternehmen in der jeweiligen Region bezogen oder für ein Unternehmen im Ausland zur Verfügung gestellt hat.

Die hier erhobenen Daten überschneiden sich inhaltlich mit der Außenhandelsstatistik (International Trade in Goods Statistics, ITGS) und der Dienstleistungshandelsstatistik der Deutschen Bundesbank (Statistics on International Trade in Services, SITs). Allerdings bestehen auch einige Unterschiede, unter anderem bezüglich der Abschnidegrenzen (Statistisches Bundesamt, 2025, hier: Kapitel 7). In der Statistik zu globalen Wertschöpfungsketten gilt sowohl für Waren als auch für Dienstleistungen eine einheitliche Grenze von 100 000 Euro. In der Außenhandelsstatistik gelten für den Warenverkehr innerhalb der EU deutlich höhere Meldeschwellen, im Berichtsjahr 2023 waren dies 800 000 Euro für den Wareneingang und 500 000 Euro für die Warenversendung. Außerdem weichen die Definitionen der Warenarten in den beiden Statistiken voneinander ab. So

werden in der Statistik zu globalen Wertschöpfungsketten unter anderem Waren danach unterschieden, ob sie im eigenen Unternehmen nur entworfen (factoryless goods production) oder auch hergestellt werden und ob sie im Sinne einer globalen Wertschöpfungskette im empfangenden Unternehmen weiterverarbeitet oder nur verbraucht werden. Letztere werden unter der Kategorie „andere Waren“ erfasst. Aus diesen Gründen ist es nicht möglich, die Merkmale aus den Einzeldaten des Außenhandels zu erstellen. Darüber hinaus unterscheidet sich die Statistik zu globalen Wertschöpfungsketten in ihrer Zielsetzung von der Außenhandelsstatistik dahingehend, dass sie die Anzahl der Unternehmen in globalen Wertschöpfungsketten misst, während die Außenhandelsstatistik den Wert gehandelter Waren erfasst.

Die SITS-Daten decken sich zum Teil mit den Merkmalen der GVC-Statistik, jedoch nur im Bereich des Handels mit Dienstleistungen. In der SITS-Statistik liegt die Meldeschwelle zudem bei 12 500 Euro und damit unter derjenigen der GVC-Statistik.

Zum Themenkomplex globaler Wertschöpfungsketten wird darüber hinaus erfasst, inwieweit Unternehmen im Dreijahreszeitraum die Organisation ihrer globalen Wertschöpfungsketten durch globale Ereignisse beeinflusst sehen und wie sie auf diese Veränderungen reagieren.

## Übersicht 1

### Struktur des Fragebogens zum Inhalt der Statistik zu globalen Wertschöpfungsketten

|     | Merkmal                                                  | Ausprägungen                                           | Art des Merkmals | Zielgruppe der Frage                                   |
|-----|----------------------------------------------------------|--------------------------------------------------------|------------------|--------------------------------------------------------|
| 1   | Tätige Personen                                          | 8 Unternehmensfunktionen                               | metrisch         | alle                                                   |
| 2   | Einbindung in globale Wertschöpfungsketten               | EU/Nicht-EU                                            |                  |                                                        |
| 2.1 | Bezug von Waren                                          | 6 Warenarten                                           | dichotom         | alle                                                   |
| 2.2 | Bezug von Dienstleistungen                               | 7 Dienstleistungsarten                                 | dichotom         | alle                                                   |
| 2.3 | Bereitstellung von Waren                                 | 6 Warenarten                                           | dichotom         | alle                                                   |
| 2.4 | Bereitstellung von Dienstleistungen                      | 7 Dienstleistungsarten                                 | dichotom         | alle                                                   |
| 3   | Verlagerungen                                            |                                                        |                  |                                                        |
| 3.1 | Verlagerung                                              | ja/nein/in Erwägung                                    | kategorial       | alle                                                   |
| 3.2 | Verlagerung                                              | 7 Unternehmensfunktionen, Bezug zur Unternehmensgruppe | dichotom         | Verlagerer                                             |
| 3.3 | Verlagerung                                              | 7 Unternehmensfunktionen, Zielregionen                 | dichotom         | Verlagerer                                             |
| 3.4 | Beschäftigungseffekte                                    | 7 Unternehmensfunktionen, Qualifizierung               | metrisch         | Verlagerer                                             |
| 4   | Motive und Bedenken der Verlagerung                      |                                                        |                  |                                                        |
| 4.1 | Motive                                                   | 13 Motive nach 3 Relevanzstufen                        | kategorial       | Verlagerer                                             |
| 4.2 | Bedenken                                                 | 14 Gründe für Bedenken nach 3 Relevanzstufen           | kategorial       | Verlagerer; Unternehmen, die Verlagerung erwogen haben |
| 5   | Globale Ereignisse                                       |                                                        |                  |                                                        |
| 5.1 | Störung der Organisation globaler Wertschöpfungsketten   | 7 mögliche Störungen nach 3 Relevanzstufen             | kategorial       | alle                                                   |
| 5.2 | Anpassung der Organisation globaler Wertschöpfungsketten | 6 mögliche Anpassungsstrategien nach 3 Relevanzstufen  | kategorial       | alle                                                   |

## 2.2 Verlagerungen ins Ausland

Verlagerungen ins Ausland können als ein Teilaspekt globaler Wertschöpfungsketten betrachtet werden, da sie nur für international tätige Unternehmen relevant sind. In der GVC-Statistik werden unter Verlagerungen ausschließlich Auslandsverlagerungen erfasst. Eine Verlagerung ins Ausland liegt vor, wenn Stellen im Inland abgebaut und gleichzeitig in einem anderen Land neu geschaffen werden. Hiervon abzugrenzen sind zum Beispiel der Aufbau eines Standortes im Ausland ohne gleichzeitigen Abbau von Stellen im Inland oder umgekehrt der Stellenabbau im Inland ohne entsprechenden Stellenzuwachs im Ausland. Bei Auslandsverlagerungen können Arbeitsplätze entweder innerhalb der eigenen Unternehmensgruppe oder an ein Unternehmen außerhalb der Gruppe verlagert werden.

Die Erfassung von Verlagerungen basiert auf dem Konzept der Unternehmensfunktionen (Statistics Denmark und andere, 2008). Erfragt wird hierbei, wie viele tätige Personen in welchen Unternehmensfunktionen beschäftigt sind, wie viele Arbeitsplätze in den vergangenen drei Jahren verlagert wurden und wie viele Arbeitsplätze im Zuge von Verlagerungen im Inland neu entstanden sind.

Darüber hinaus werden Unternehmen gefragt, ob innerhalb der eigenen Unternehmensgruppe (bei multinationalen Unternehmen) verlagert wurde, welche Unternehmensfunktionen betroffen waren sowie in welche Zielregion verlagert wurde. Unternehmen, die Unternehmensfunktionen verlagert haben oder eine Verlagerung zumindest in Erwägung gezogen haben, geben zusätzlich Auskunft über die Motive für die Verlagerung sowie über die Bedenken, die sie bei der Entscheidung berücksichtigt haben.

### 3

## Methodik

### 3.1 Stichprobenziehung und Hochrechnung

Die Statistik wird, beginnend mit dem Berichtszeitraum 2021 bis 2023, alle drei Jahre durchgeführt. Die Grundgesamtheit der Statistik besteht aus Statistischen Unternehmen<sup>1</sup> mit mindestens 50 tätigen Personen in den Abschnitten B bis N der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008). Die Stichprobe wird auf Basis des statistischen Unternehmensregisters gezogen, wobei für jedes Statistische Unternehmen immer die jeweils bestimmende Rechtliche Einheit herangezogen wird. Von 59 127 Unternehmen der Grundgesamtheit wurden 9 813 Unternehmen beziehungsweise bestimmende

Rechtliche Einheiten in die Stichprobe gezogen. Von den 9 813 Einheiten sind 8 607 verwertbare Meldungen eingegangen, was einer Rücklaufquote von 87,7 % entspricht. Die übrigen Einheiten sind in der Zwischenzeit erloschen, liegen durch ihre Angaben zu tätigen Personen außerhalb der Grundgesamtheit oder haben aus sonstigen Gründen nicht geantwortet.

Bei der Stichprobenziehung werden bevorzugt Rechtliche Einheiten gezogen, die im letzten Jahr des Berichtszeitraums wahrscheinlich in globale Wertschöpfungsketten eingebunden waren. Diese Einheiten wurden überproportional gezogen, um diejenigen Unternehmen zu entlasten, die nicht international aktiv sind und für die die meisten Fragen nicht relevant sind. Außerdem können so bessere Schätzer für seltene Merkmale, insbesondere für die Verlagerung, erzielt werden. Die Einbindung in globale Wertschöpfungsketten wird mithilfe eines Random-Forest-Modells geschätzt. Das Random-Forest-Modell trifft mithilfe vieler Entscheidungsbäume (Decision Trees) Vorhersagen. Dabei bezieht jeder einzelne Entscheidungsbaum ein zufälliges Set von Variablen und Beobachtungen ein, um eine Vorhersage zu treffen. Für jede zu schätzende Einheit werden die Vorhersagen der einzelnen Entscheidungsbäume zu einer Vorhersage aggregiert. Wie bei anderen Machine-Learning-Anwendungen wird der Datensatz vorher in einen Trainingsdatensatz zur Berechnung des Random-Forest-Modells und einen Test-Datensatz zur Beurteilung des Random-Forest-Modells unterteilt. Hierfür ist ein Datensatz notwendig, der für ein Set an Beobachtungen sowohl die Zielvariable (das sogenannte Target), also die zu schätzende Variable, als auch die unabhängigen Variablen für das Modell (die sogenannten Features) enthält. Im Fall der GVC-Erhebung wird die Einbindung in globale Wertschöpfungsketten als dichotomes Merkmal aus den Daten der Piloterhebung des Berichtszeitraums 2018 bis 2020 berechnet. Für die Features werden Variablen aus dem statistischen Unternehmensregister und den Außenhandelsdaten gebildet und angespielt. Das Modell führt zu einer etwas besseren Schätzung als eine dichotome Variable, die anzeigt, ob ein Unternehmen importiert oder exportiert. In der Grundgesamtheit der GVC-Erhebung 2021 bis 2023 wurde für 53 % die Einbindung in globale Wertschöpfungsketten geschätzt.

1 Hierbei wird der Unternehmensbegriff gemäß der EU-Einheitenverordnung 696/93 verwendet. Diese definiert als Unternehmen die kleinste Kombination Rechtlicher Einheiten, die eine organisatorische Einheit zur Erzeugung von Waren und Dienstleistungen bildet und insbesondere in Bezug auf die Verwendung der ihr zufließenden laufenden Mittel über eine gewisse Entscheidungsfreiheit verfügt. Ein Statistisches Unternehmen kann gemäß dieser Definition aus einer Rechtlichen Einheit („einfaches Unternehmen“) oder aus mehreren Rechtlichen Einheiten („komplexes Unternehmen“) bestehen.



Um die finalen Ergebnisse zu ermitteln, werden die Stichprobendaten hochgerechnet. Bei der freien Hochrechnung werden unterschiedliche Ziehungswahrscheinlichkeiten und unterschiedliche Rücklaufquoten berücksichtigt. Der freie Hochrechnungsfaktor wird durch Abgleich der Nettostichprobe, also der Menge tatsächlich befragter Einheiten, mit der Grundgesamtheit berechnet und dient dazu, die Zahl der Unternehmen hochzurechnen. Die Berechnung des gebundenen Hochrechnungsfaktors basiert auf einem GREG-Schätzer (Generalized Regression Estimator). Hierbei werden Randverteilungen aus der Unternehmensstrukturstatistik genutzt. Der gebundene Hochrechnungsfaktor dient dazu, die Zahl tätiger Personen beziehungsweise verlagelter Stellen hochzurechnen.

### 3.2 Ermittlung der Ergebnisse für die Darstellungseinheit Statistisches Unternehmen

---

Die Darstellungseinheit in der GVC-Statistik ist das Statistische Unternehmen nach der EU-Einheitenverordnung, die Befragungs- und Berichtseinheit ist hingegen die Rechtliche Einheit (Beck und andere, 2020). Für komplexe Unternehmen, die aus mehreren Rechtlichen Einheiten bestehen, wird von den Erhebungsdaten der bestimmenden Rechtlichen Einheit auf das gesamte Unternehmen geschlossen. Dabei wird bei kategorialen Merkmalen die Antwort der Rechtlichen Einheit als repräsentativ für das Statistische Unternehmen gesehen (sogenannter Repräsentantenansatz). Problematisch ist dies, wenn das jeweilige imputierte Merkmal nicht für die Berichtseinheit, aber für eine andere Rechtliche Einheit des Unternehmens zutrifft. In diesem Fall würde das Merkmal für das Statistische Unternehmen fälschlicherweise als nicht zutreffend klassifiziert. Dieses Problem besteht fort, solange die Daten bei der bestimmenden Rechtlichen Einheit erhoben werden. Allerdings sollen mit der aktuell geplanten Modernisierung der Unternehmensstatistiken die Daten direkt beim Statistischen Unternehmen erhoben werden. Bei den wenigen metrischen Merkmalen der Statistik (Zahl der tätigen Personen und Zahl der Stellen) müssen hingegen die Daten der nicht erhobenen Rechtlichen Einheiten im komplexen Unternehmen imputiert werden.

Die Imputation des Merkmals der tätigen Personen erfolgte folgendermaßen: Für die Zahl tätiger Personen wurden die Daten zu komplexen Unternehmen aus der Strukturstatistik herangezogen. Hierfür wurden die Zahlen tätiger Personen aus den jeweils nicht befragten Rechtlichen Einheiten eines jeden komplexen Unternehmens der Nettostichprobe zu der Zahl tätiger Personen der befragten Rechtlichen Einheit addiert. Da in der GVC-Statistik jedoch auch die tätigen Personen nach Unternehmensfunktion ausgewiesen werden, war ihre Verteilung auf die Unternehmensfunktionen für nicht befragte Rechtliche Einheiten zu schätzen. Mithilfe einer Korrespondenztabelle von Eurostat wurden die Berufe zunächst Unternehmensfunktionen zugeordnet. Auf dieser Grundlage konnte dann die Verteilung der tätigen Personen anhand der Daten der Verdiensterhebung auf WZ-4-Steller-Ebene geschätzt werden. Die so geschätzten Werte wurden den nicht erhobenen Rechtlichen Einheiten der Statistischen Unternehmen in der Stichprobe auf WZ-4-Steller-Ebene zugespielt.

Für die Imputation der Beschäftigungseffekte sind die Zahlen zum Stellenabbau oder -zuwachs auf das komplexe Unternehmen hochgerechnet. Dabei wurden sie mit dem Quotienten aus der Zahl tätiger Personen im Statistischen Unternehmen und der Zahl tätiger Personen in der bestimmenden Rechtlichen Einheit multipliziert.

## 4

---

### Ergebnisse

---

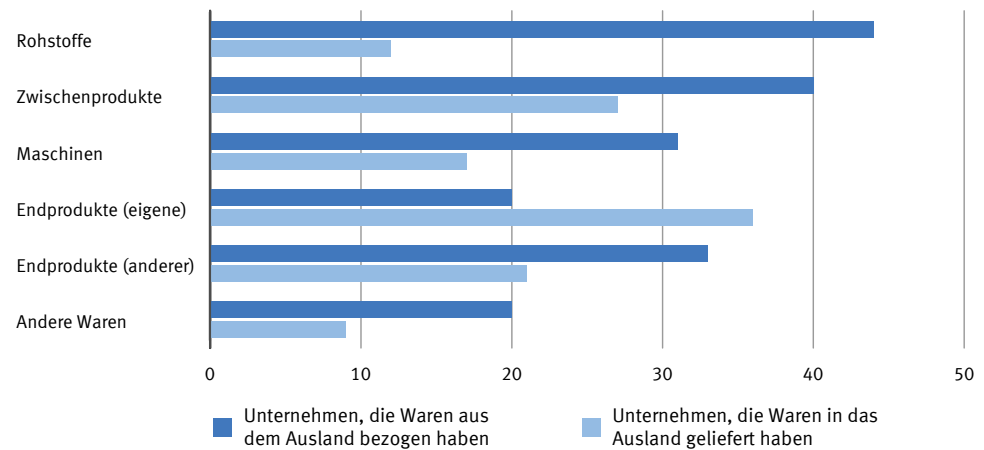
#### 4.1 Einbindung in globale Wertschöpfungsketten

---

Von den 59 127 Unternehmen der Grundgesamtheit der GVC-Erhebung 2021 bis 2023 waren 34 641 Unternehmen (59 %) in globale Wertschöpfungsketten eingebunden. Für diese hat der Intra-EU-Handel eine besondere Bedeutung. So haben 94 % aller im internationalen Warenhandel aktiven Unternehmen Waren mit Unternehmen in anderen EU-Staaten gehandelt. Demgegenüber lag dieser Anteil bei

**Grafik 1**

Unternehmen, die im Jahr 2023 grenzüberschreitend Waren bezogen und/oder geliefert haben in %<sup>1</sup>



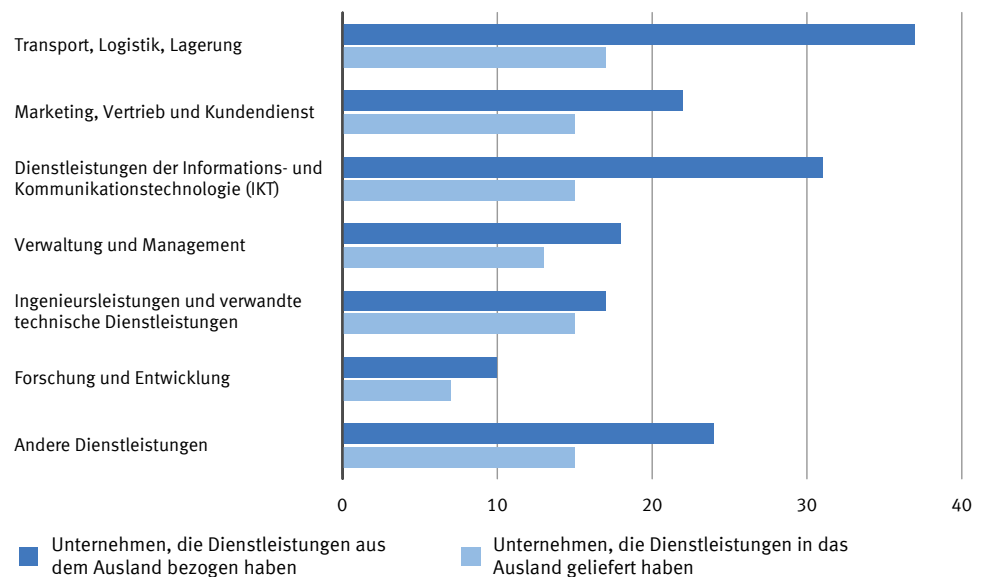
<sup>1</sup> Bezogen auf Unternehmen, die im internationalen Warenaustausch eingebunden sind.

Partnern im übrigen Europa bei 56 %, bei Partnern außerhalb Europas bei lediglich 27 %. Für fast alle Warenarten lässt sich beobachten, dass diese jeweils von mehr Unternehmen aus dem Ausland bezogen

als ins Ausland geliefert wurden. Die Ausnahme bildet die Warenart der vom eigenen Unternehmen entwickelten Endprodukte. Diese wurden von mehr Unternehmen exportiert als importiert. [↗ Grafik 1](#)

**Grafik 2**

Unternehmen, die im Jahr 2023 grenzüberschreitend Dienstleistungen bezogen und/oder erbracht haben in %<sup>1</sup>



<sup>1</sup> Bezogen auf Unternehmen, die im internationalen Dienstleistungshandel eingebunden sind.



Bei den Dienstleistungen zeigt sich in Bezug auf die Regionen der Handelspartner ein ähnliches Muster wie beim Warenverkehr: Von den Unternehmen, die Dienstleistungen international gehandelt haben, haben 88 % mit Unternehmen in anderen Ländern der EU gehandelt, lediglich 47 % mit Unternehmen im übrigen Europa und 44 % mit Unternehmen außerhalb Europas. Auch für die einzelnen Dienstleistungsarten ist dieses Muster zu beobachten. Am häufigsten sind Dienstleistungen aus dem Bereich Transport, Logistik, Lagerung bezogen worden (37 %), am zweithäufigsten Dienstleistungen der Informations- und Kommunikationstechnologie (31 %). Alle sieben erhobenen Dienstleistungsarten sind von mehr Unternehmen aus dem Ausland bezogen als für Unternehmen im Ausland erbracht worden.

#### ↘ Grafik 2

Bei diesen Vergleichen ist allerdings zu berücksichtigen, dass nur der direkte Handel erfasst wird. Beispielsweise geben Unternehmen, die ausschließlich Rohstoffe von einem Zwischenhändler im EU-Ausland beziehen, an, dass sie selbst ausschließlich Rohstoffe aus dem EU-Ausland bezogen haben, auch wenn dieser Zwischenhändler die Rohstoffe von außerhalb der EU bezogen hat. In diesem Fall würde nur der Zwischenhändler angeben, Rohstoffe aus einem Land außerhalb der Europäischen Union bezogen zu haben.

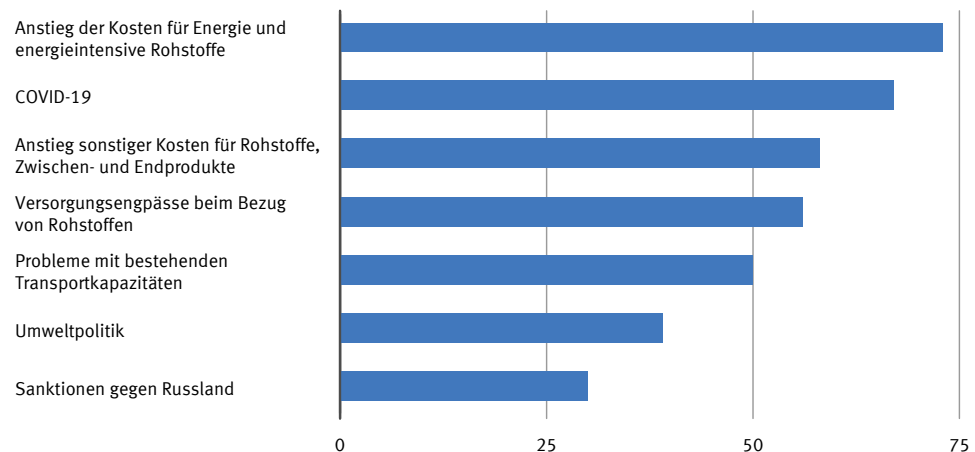
## 4.2 Beeinträchtigung der Organisation globaler Wertschöpfungsketten

↘ Grafik 3 zeigt, dass die Organisation globaler Wertschöpfungsketten insbesondere vom Anstieg der Kosten für Energie (73 % der Unternehmen) und durch die pandemiebedingten Umstände (67 % der Unternehmen) betroffen war. Die hohen Werte sind nur scheinbar ein Widerspruch dazu, dass nur 59 % der Unternehmen berichtet haben, in globale Wertschöpfungsketten eingebunden gewesen zu sein. Weil die Einbindung erst ab einer Schwelle von 100 000 Euro Einkaufs- beziehungsweise Verkaufswert erhoben wird, kann ein Unternehmen unterhalb dieser Schwelle in globale Wertschöpfungsketten eingebunden sein und dennoch eine Beeinträchtigung dieser berichtet haben.<sup>12</sup>

- 2 In einigen Fällen wurden möglicherweise fälschlicherweise Beeinträchtigungen globaler Wertschöpfungsketten angegeben. Beispielsweise berichteten Unternehmen der Gastronomie, dass die Organisation ihrer globalen Wertschöpfungsketten von gestiegenen Energiepreisen betroffen sei. Zwar ist es plausibel, dass die Unternehmen von den gestiegenen Energiepreisen betroffen sind, jedoch ist es eher fragwürdig, dass hier eine Beeinträchtigung einer globalen Wertschöpfungskette vorliegt. In künftigen Erhebungen soll ein zusätzlicher Hinweis an dieser Stelle im Fragebogen mehr Klarheit für die meldenden Unternehmen schaffen.

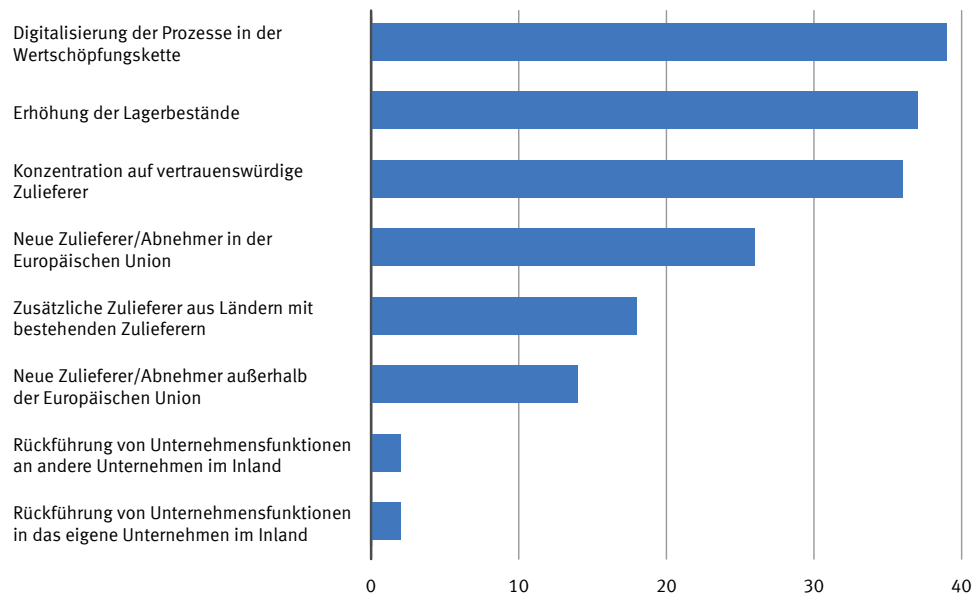
#### Grafik 3

Unternehmen nach dem Grund der Beeinträchtigung der Organisation ihrer globalen Wertschöpfungsketten 2021 bis 2023  
in %



**Grafik 4**

**Unternehmen nach Anpassung der Organisation ihrer globalen Wertschöpfungsketten  
2021 bis 2023  
in %**



Auf Beeinträchtigungen der Organisation ihrer Wertschöpfungsketten reagiert haben Unternehmen, indem sie die Prozesse ihrer globalen Wertschöpfungskette digitalisiert (39 %), ihre Lagerbestände erhöht (37 %) oder sich auf vertrauenswürdige Zulieferer konzentriert haben (36 %). [↗ Grafik 4](#)

### 4.3 Auslandsverlagerung

Im Berichtszeitraum zwischen 2021 und 2023 haben 1297 Unternehmen mit mindestens 50 tätigen Personen Stellen ins Ausland verlagert. Dies entspricht 2,2 % der 59 127 Unternehmen in der Grundgesamtheit.<sup>3</sup>

Die Zahl der im Zeitraum 2021 bis 2023 ins Ausland verlagerten Stellen beläuft sich auf 71 091. Dem-

gegenüber sind im Zuge dieser Auslandsverlagerungen 20270 neue Stellen entstanden, etwa durch Umschichtungen in andere Unternehmensfunktionen oder durch Neueinstellungen aufgrund von Kosteneinsparungen. Daraus ergibt sich ein Nettobeschäftigungsabbau von 50 821 Stellen. Der hier gegengerechnete Beschäftigungszuwachs umfasst aber lediglich den Stellenzuwachs im Zusammenhang mit Auslandsverlagerungen. Verlagerungen von Stellen an ein anderes Unternehmen in Deutschland sind dabei nicht erfasst. Auf der Gegenseite sind Unternehmen nicht abgebildet, die ihre Tätigkeit im Berichtszeitraum vollständig verlagert haben, weil sie zum Zeitpunkt der Stichprobenziehung nicht mehr Teil der Grundgesamtheit waren. Insofern darf der Nettobeschäftigungsabbau nicht als Anzahl der insgesamt durch Auslandsverlagerungen gewonnen beziehungsweise verlorenen Stellen gesehen werden.

[↗ Tabelle 1](#) zeigt, dass 641 Unternehmen im Berichtszeitraum 2021 bis 2023 die Unternehmensfunktion Verwaltung und Management ins Ausland verlagert haben. Während die Zahl verlagernder Unternehmen für die Unternehmensfunktion Ver-

<sup>3</sup> Es ist zusätzlich zu beachten, dass nur die Grundgesamtheit der Unternehmen im Berichtsjahr 2023 betrachtet wird. Ist ein Unternehmen im Zeitraum 2021 bis 2023 vollständig ins Ausland abgewandert, konnte dies nicht in der aktuellen Erhebung erfasst werden. Die berichtete Zahl verlagernder Unternehmen ist daher immer als die Zahl von Unternehmen zu verstehen, die im letzten Jahr des Berichtszeitraums in Deutschland ansässig waren und im Berichtszeitraum Stellen ins Ausland verlagert haben.

**Tabelle 1**

**Unternehmen, die im Zeitraum 2021 bis 2023 Unternehmensfunktionen ins Ausland verlagert haben, nach Art der verlagerten Unternehmensfunktionen**

|                                                                        | Anzahl |
|------------------------------------------------------------------------|--------|
| Herstellung von Gütern                                                 | 437    |
| Transport, Logistik, Lagerung                                          | 193    |
| Marketing, Vertrieb und Kundendienst                                   | 312    |
| Dienstleistungen der Informations- und Kommunikationstechnologie (IKT) | 333    |
| Verwaltung und Management                                              | 641    |
| Ingenieursleistungen und verwandte technische Dienstleistungen         | 252    |
| Forschung und Entwicklung                                              | 154    |
| Andere Dienstleistungen                                                | 137    |

waltung und Management damit am höchsten ist, ist die Zahl abgebauter Stellen in der Unternehmensfunktion Produktion am höchsten. Hier sind für den Berichtszeitraum 26 143 abgebaute Stellen berichtet, in Verwaltung und Management 16 755.

Das mit Abstand häufigste Ziel von Auslandsverlagerungen ist die EU. Von den 1297 Unternehmen, die im Berichtszeitraum Auslandsverlagerungen durchgeführt haben, haben 920 Unternehmen in andere Länder der EU verlagert. In Länder außerhalb der EU haben 736 Unternehmen verlagert. ➤ **Tabelle 2**

**Tabelle 2**

**Unternehmen mit Auslandsverlagerungen im Zeitraum 2021 bis 2023 nach Zielregionen der internationalen Verlagerungen von Arbeitsplätzen**

|                                           | Anzahl |
|-------------------------------------------|--------|
| Innerhalb der Europäischen Union (EU)     | 920    |
| In Staaten außerhalb der EU               | 736    |
| In andere europäische Staaten             | 353    |
| Nach Indien                               | 271    |
| In die Vereinigten Staaten und Kanada     | 128    |
| Nach China                                | 126    |
| In andere asiatische Staaten und Ozeanien | 115    |

Unternehmen, die im Berichtszeitraum Auslandsverlagerungen durchgeführt haben, wurden auch zu ihren Beweggründen befragt. Häufigste Gründe, eine Verlagerung durchzuführen, waren die Lohnkosten (74 %), andere Kostenvorteile (59 %) und eine strategische Entscheidung des Gruppenoberhauptes (62 %). In Bezug auf die Hindernisse, eine Verlagerung durchzuführen, haben Unternehmen vor allem recht-

liche und administrative Hürden (62 %) und steuerliche Probleme (49 %) genannt. Zudem berichteten viele Unternehmen Bedenken, dass die Kosten der Verlagerung den Nutzen übersteigen könnten (43 %).

## 4.4 Fabriklose Produzenten

Wie bereits erwähnt, ermittelt die GVC-Statistik, inwieweit Unternehmen im Ausland produzierte Waren beziehen, die sie zuvor selbst entwickelt haben. Dies ermöglicht damit im Ansatz die Untersuchung der fabriklosen Produktion (factoryless goods production, FGP). Eine fabriklose Produktion liegt vor, wenn ein Unternehmen

- > keine eigene Herstellung durchführt,
- > die Produkte selbst entwickelt,
- > die geistigen Eigentumsrechte am Produkt besitzt,
- > den Produktionsprozess kontrolliert,
- > den Vertrieb selbst durchführt (Allafi und andere, 2017).

Die eigene Entwicklung, die geistigen Eigentumsrechte und die Kontrolle des Produktionsprozesses gehen für die Statistik aus der Formulierung der Frage im Fragebogen und den Erläuterungen hierzu hervor. Das Vorhandensein von eigenem Vertrieb/Marketing und das Fehlen einer eigenen Herstellung kann durch die Daten zu tätigen Personen nach Unternehmensfunktionen festgestellt werden. Problematisch ist dabei der Schritt der Konsolidierung des komplexen Unternehmens, denn es kann nicht ausgeschlossen werden, dass die Produktion in einer anderen Rechtlichen Einheit des Statistischen Unternehmens stattfindet. Aus diesem Grund wird die Besetzung der Unternehmensfunktion Vertrieb/Marketing als Kriterium hier nicht angewandt, weil diese möglicherweise bei einer anderen Rechtlichen Einheit des jeweils betrachteten komplexen Unternehmens liegen kann.

Die fabriklose Produktion ist auf der einen Seite vom Handel abzugrenzen, der nur den Weiterverkauf, nicht aber die Entwicklung des Endprodukts bezeichnet. Auf der anderen Seite ist die fabriklose Produktion von der Herstellung von Waren abzugrenzen, da das Endprodukt lediglich entwickelt wird. Fabriklose

Produktionen werden nicht als solche erkannt, wenn eine andere Rechtliche Einheit des Unternehmens die selbst entwickelten Endprodukte aus dem Ausland bezieht.

Trotzdem sind mit diesen Daten bereits interessante Einblicke in die Struktur von fabriklosen Produzenten möglich. Insgesamt sind 1362 Unternehmen, also 2 % der Unternehmen mit mehr als 50 tätigen Personen, als fabriklose Produzenten tätig. Von diesen sind 7 % dem WZ-Abschnitt C Verarbeiten des Gewerbe und 66 % dem WZ-Abschnitt G Handel zugeordnet. Der größte Anteil der fabriklosen Produktion liegt in absteigender Reihenfolge bei G46 Großhandel (ohne Handel mit Kraftfahrzeugen), G47 Einzelhandel und J62 Erbringung von Dienstleistungen der Informationstechnologie. Nach der WZ-2008-Systematik sind fabriklose Produzenten in WZ-Abschnitt G zu verorten. Unternehmen, die hier als fabriklose Produzenten klassifiziert werden, im statistischen Unternehmensregister aber außerhalb des WZ-Abschnitts G verortet werden, sind entweder hinsichtlich ihrer Tätigkeit nicht korrekt klassifiziert oder werden, beispielsweise durch den Repräsentantenansatz, fälschlicherweise als fabriklose Produzenten klassifiziert. Hinsichtlich ihrer Größe unterscheiden

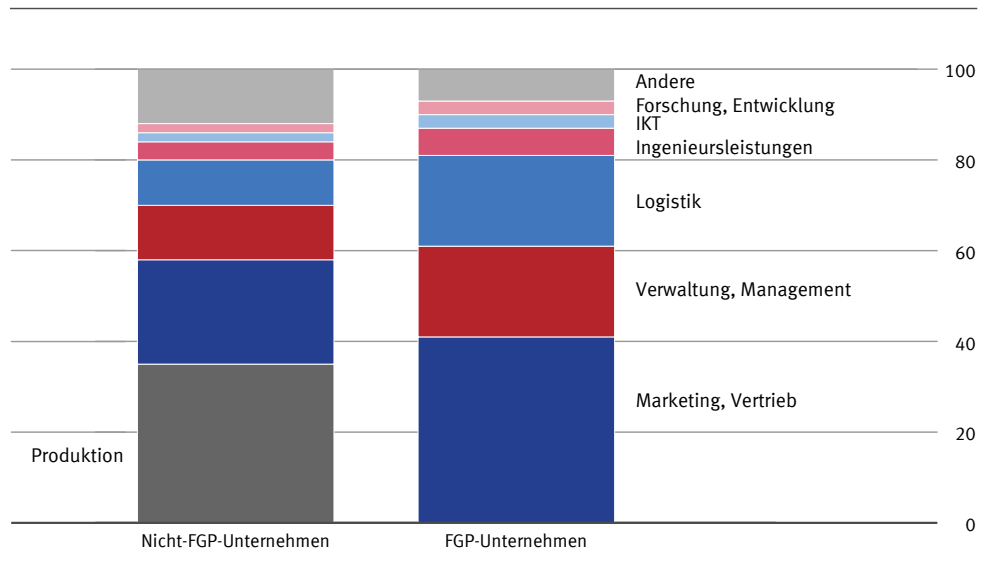
sich fabriklose Produzenten von den übrigen Unternehmen vor allem am rechten Rand der Verteilung: So sind in 8 % der Unternehmen mit fabrikloser Produktion mindestens 1000 Personen tätig, in den übrigen Unternehmen sind es lediglich 4 %.

Allerdings ist zu beachten, dass die meisten der Unternehmen, die FGP-Aktivitäten angegeben haben, zugleich auch andere Warenarten handeln. Im Gegensatz dazu ist in beiden WZ-Abschnitten (C und G) die Zahl der fabriklosen Produzenten, die ausschließlich selbst entwickelte Endprodukte beziehen, sehr gering: In WZ-Abschnitt C sind dies unter 1 % der Unternehmen in diesem WZ-Abschnitt, im Abschnitt G sind es unter 2 %. Nur sehr wenige Unternehmen beschränken sich also auf den Bezug selbst entwickelter Endprodukte.

Vor dem Hintergrund, dass fabriklose Produktionen zwar in der Regel dem Großhandel zugeordnet sind, jedoch innerhalb ihrer Gruppe ein besonderes Tätigkeitsprofil haben, sind die erhobenen Daten zu tätigen Personen nach Unternehmensfunktion an dieser Stelle besonders interessant. Unternehmensfunktionen haben ihre Entsprechungen in Unternehmensaktivitäten.

**Grafik 5**

Verteilung tätiger Personen auf Unternehmensfunktionen in Nicht-FGP- und FGP-Unternehmen 2023  
in %



FGP: factoryless goods production - fabriklose Produktion.

Die Verteilung tätiger Personen auf die Unternehmensfunktionen unterscheidet sich zwischen FGP-Unternehmen und Nicht-FGP-Unternehmen zwangsläufig, da in FGP-Unternehmen per definitionem keine Personen in der Produktion tätig sind. Insbesondere in den Unternehmensfunktionen „Logistik“, „Marketing und Vertrieb“ sowie „Verwaltung und Management“ sind in FGP-Unternehmen anteilmäßig mehr Personen tätig als in Nicht-FGP-Unternehmen. Erwartungsgemäß gilt dies für die Unternehmensfunktion „Forschung und Entwicklung“ allerdings nur geringfügig. [↘ Grafik 5](#)

Insgesamt stellen sich damit fabriklose Produktionen als ein Hybrid aus Unternehmen des Handels und des Verarbeitenden Gewerbes dar. Dabei ist die fabriklose Produktion nur bei sehr wenigen Unternehmen das einzige Geschäftsmodell. Allerdings ist es mit der gegebenen Fallzahl bei vielen Merkmalen nicht möglich, für fabriklose Produktionen verlässliche Schätzungen durchzuführen. Außerdem lässt sich mit den vorhandenen Daten nicht feststellen, welche der identifizierten FGP-Unternehmen tatsächlich keine eigene Produktion im Statistischen Unternehmen durchführen.

## 5

### Weiterentwicklung der Statistik

Vergleicht man die erstmalige Durchführung der GVC-Statistik als Erhebung mit Auskunftspflicht mit den vorangegangenen Piloterhebungen, ist eine große Dynamik festzustellen: So hat sich bislang jede dieser Erhebungen methodisch und inhaltlich von der vorangegangenen stark unterschieden. Im Hinblick auf aktuelle Planungen und Anregungen von Eurostat ist davon auszugehen, dass dies mindestens für die Berichtszeiträume 2024 bis 2026 sowie 2027 bis 2030 ebenfalls gelten wird. Hinzu kommen methodische Verbesserungen der Statistik auf nationaler Ebene.

### 5.1 Neue EU-Anforderungen

Für den Berichtszeitraum 2024 bis 2026 werden aufgrund geänderter EU-Datenanforderungen verschiedene Neuerungen umgesetzt. Erstens wird als neues Merkmal erhoben, ob der Handel mit Waren und Dienstleistungen schwerpunktmäßig innerhalb oder außerhalb der Unternehmensgruppe stattfindet. Dies ist eine Information, die in vorhandenen Datenquellen nicht verfügbar ist und damit eine interessante Ergänzung der GVC-Statistik darstellt. Zweitens werden in den Frageblöcken zu Motiven und Hindernissen bezüglich der Auslandsverlagerung sowie zur Beeinträchtigung der Organisation globaler Wertschöpfungsketten mindestens die auf den jetzigen Berichtszeitraum 2021 bis 2023 zugeschnittenen Fragen mit Bezug zur COVID-19-Pandemie ausgetauscht. Zudem sind Fragen im Zusammenhang mit den Auswirkungen der Zollpolitik zu erwarten.

### 5.2 Weiterentwicklung auf nationaler Ebene

Bei der GVC-Statistik 2021 bis 2023 ist es gelungen, die Zahl tätiger Personen in den Unternehmensfunktionen aus den Daten der Verdiensterhebung zu schätzen. Möglich ist dies, indem von dem ISCO<sup>14</sup>-kodierte Beruf einer jeden tätigen Person auf deren Unternehmensfunktion geschlossen wird. Diese Schätzung wird mithilfe der Daten der Verdiensterhebung auf WZ-4-Steller-Ebene aggregiert und über die Wirtschaftszweige an die GVC-Stichprobe angespielt. Für das Berichtsjahr 2023 ist dies zunächst durchgeführt worden, um die Machbarkeit zu prüfen und die Ergebnisse mit den erhobenen Daten zu vergleichen. Künftig wird diese Frage nicht mehr gestellt, was die befragten Unternehmen entlastet. Unternehmen könnten durch die Aufbereitung vorhandener Daten weiter entlastet werden. Dies wäre jedoch grundsätzlich nur bei den Variablen zur Einbindung von Unternehmen in globale Wertschöpfungsketten möglich. Hierfür könnten zumindest teilweise Daten der Außenhandelsstatistik genutzt werden, allerdings hat Eurostat dazu erst methodische Fragen zu klären.

---

<sup>14</sup> International Standard Classification of Occupations – Internationale Standardklassifikation der Berufe.

Darüber hinaus steht eine Überarbeitung der Umsetzung des Repräsentantenansatzes an. Bis zuletzt wurde hier die Rechtliche Einheit herangezogen, die im statistischen Unternehmensregister als bestimmende Einheit hinterlegt ist. In einigen wenigen Fällen ist die bestimmende Einheit jedoch ungeeignet, um von ihr auf das Statistische Unternehmen zu schließen. Das gilt insbesondere für Holdings, die teilweise nicht einmal tätige Personen führen, sodass es diesen Meldern wegen der im Online-Fragebogen implementierten Muss-Prüfungen gar nicht möglich ist, den Fragebogen auszufüllen. 

## LITERATURVERZEICHNIS

---

Allafi, Sabine/Jung, Sandra/Spies, Veronika. [Globalisierung in der amtlichen Statistik](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 130 ff.

Beck, Martin/Baumgärtner, Luisa/Bürk, Katja-Verena/Redecker, Matthias. [Einführung des EU-Unternehmensbegriffs: Konzept und Umsetzung](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2020, Seite 35 ff.

Kaus, Wolfhard. [Auslandsverlagerung wirtschaftlicher Aktivitäten: Unternehmenscharakteristika und Beschäftigungswirkung](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2019, Seite 11 ff.

Statistics Denmark, Statistics Finland, Statistics Netherlands, Statistics Norway, Statistics Sweden. *International Sourcing. Moving Business Functions Abroad*. Kopenhagen 2008. [Zugriff am 19. Dezember 2025]. Verfügbar unter: [www.dst.dk](http://www.dst.dk)

Statistisches Bundesamt. *Organisation und Verlagerung wirtschaftlicher Aktivitäten – Methodische Erläuterungen und Ergebnisse 2016*. 2019. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Statistisches Bundesamt. *Statistik zu Globalen Wertschöpfungsketten 2021-2023*. Qualitätsbericht. 2025. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

Ullrich, Katrin. *Deutschlands umfangreiche Einbindung in internationalen Handel und globale Wertschöpfungsketten*. KfW Research, Fokus Volkswirtschaft. Nr. 411, 22. Dezember 2022. [Zugriff am 19. Dezember 2025]. Verfügbar unter: [www.kfw.de](http://www.kfw.de)

Zwania, Jacek. [Verlagerung wirtschaftlicher Aktivitäten](#). In: Wirtschaft und Statistik. Ausgabe 6/2008, Seite 477 ff.

## RECHTSGRUNDLAGEN

---

Durchführungsverordnung (EU) 2022/918 der Kommission vom 13. Juni 2022 zur Festlegung der technischen Spezifikationen der Datenanforderungen für das Thema „Globale Wertschöpfungsketten“ gemäß der Verordnung (EU) 2019/2152 des Europäischen Parlaments und des Rates (Amtsblatt der EU Nr. L 159, Seite 43).

Gesetz über die Statistik zu globalen Wertschöpfungsketten (GWStatG) vom 17. Januar 2024 (BGBl. I Nr. 13), zuletzt geändert durch Artikel 10 des Gesetzes vom 22. Dezember 2025 (BGBl. I Nr. 354).

Verordnung (EU) 2019/2152 des europäischen Parlaments und des Rates vom 27. November 2019 über europäische Unternehmensstatistiken, zur Aufhebung von zehn Rechtsakten im Bereich Unternehmensstatistiken (Amtsblatt der EU Nr. L 327, Seite 1).



# KÜHLUNG IN STÄDTEN, KOHLENSTOFFSPEICHERUNG, LANDSCHAFTEN FÜR TOURISMUS: ÖKOSYSTEMRECHNUNGEN ZEIGEN LEISTUNGEN DER NATUR

Marius Belling, Simon Felgendreher, Dan Kanmegne,  
Johannes Oehrlein, Jonathan Reith, Simon Schürz

📌 **Schlüsselwörter:** Umweltökonomische Gesamtrechnungen – Ökosystemleistungen – Feinstaubfilterung – Holzzuwachs – Ernteerträge

## ZUSAMMENFASSUNG

Die Ökosystemrechnungen haben zum Ziel, das räumliche Ausmaß, den Zustand sowie die Leistungen der Ökosysteme für den Menschen systematisch in Konten zu erfassen und im Zeitverlauf darzustellen. Die Leistungsbilanzen der Ökosysteme zeigen eine Auswahl der vielfältigen Leistungen der Natur, beispielsweise Kühlung in Städten, naturbezogener Tourismus oder Kohlenstoffspeicherung. Diese werden auf Basis der Flächen- und Zustandsbilanz in physischen Einheiten gemessen oder modelliert. Die dargestellten Leistungsindikatoren beziehen sich auf bundesweite jährliche Schätzwerte für alle Ökosysteme und erfüllen neue Berichtspflichten der Verordnung über europäische umweltökonomische Gesamtrechnungen.

📌 **Keywords:** *environmental-economic accounting – ecosystem services – absorption of fine particles – wood provision – crop provision*

## ABSTRACT

*The purpose of ecosystem accounting is to record the extent and condition of ecosystems and their services to humans, and present this information over time in the form of accounts. Ecosystem service accounts present some of the manifold services provided by nature. These services are measured or modelled in physical units based on the ecosystem extent and condition accounts. The service indicators shown refer to yearly estimates for all ecosystems in Germany and meet new reporting obligations of the regulation on European environmental economic accounts.*

### Marius Belling

studierte Geographie und Management natürlicher Ressourcen. Im Statistischen Bundesamt ist er in den Bereichen Ökosystemrechnungen sowie Waldgesamtrechnungen tätig.

### Dr. Simon Felgendreher

ist Volkswirt, koordiniert im Statistischen Bundesamt die Ökosystemrechnungen und ist für deren konzeptionelle Weiterentwicklung zuständig.

### Dr. Dan Kanmegne

ist Fernerkundungswissenschaftler und im Statistischen Bundesamt bei den Ökosystemrechnungen für die technische Implementierung der Flächen- und Zustandsbilanz verantwortlich.

### Dr. Johannes Oehrlein

ist Mathematiker im Bereich Geoinformatik. Im Statistischen Bundesamt betreut er die technische Konzeption und Umsetzung der Ökosystemrechnungen.

### Jonathan Reith

ist Geograph und setzt seine ökologische Expertise im Statistischen Bundesamt bei der Berechnung der Ökosystemleistungen ein.

### Dr. Simon Schürz

ist Volkswirt und im Statistischen Bundesamt verantwortlich für methodische Fragen und die monetären Konten der Ökosystemrechnungen.



## 1

### Einleitung

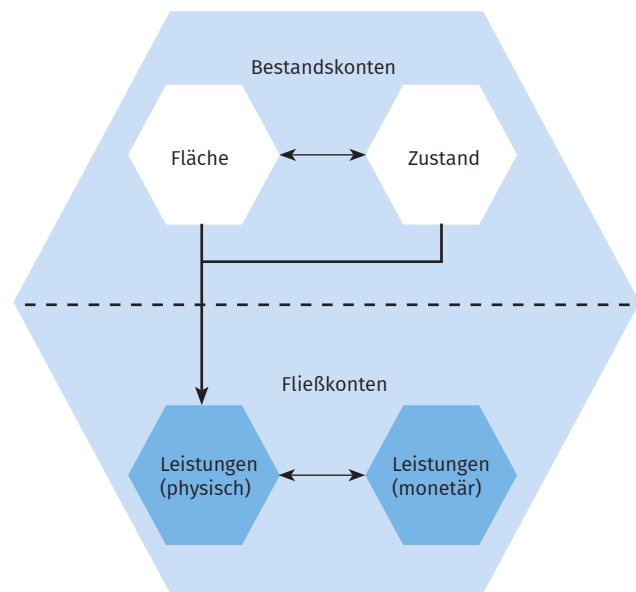
Das im November 2025 verabschiedete Berliner BäumePlus-Gesetz soll die Anzahl der Stadtbäume in der deutschen Hauptstadt verdoppeln und so die Hitzebelastung abmildern und die Luftqualität verbessern.<sup>1</sup> Während der COVID-19-Pandemie stieg die Bedeutung von Tourismus in der Natur.<sup>2</sup> Durch Schädlinge und Trockenheit geben die Wälder in Deutschland inzwischen mehr Kohlenstoff ab, als sie aufnehmen.<sup>3</sup> Diese Schlagzeilen verdeutlichen, wie stark gesellschaftliche Diskussionen mit den Leistungen von Ökosystemen verbunden sind – etwa Kühlung in Städten, Feinstaubfilterung, naturbezogener Tourismus oder Kohlenstoffspeicherung.

Verwunderlich ist das allerdings nicht, denn Ökosystemleistungen schaffen substanziellen gesundheitlichen, sozialen und wirtschaftlichen Nutzen: Urbanes Grün senkte die Temperaturen in deutschen Städten im Sommer 2023 durchschnittlich um 0,9° Celsius. Deutschlands Wälder filterten im Jahr 2023 insgesamt 25 300 Tonnen Feinstaub. Deutschlandweit hingen im Jahr 2023 rund 265 Millionen touristische Übernachtungen von attraktiven Ökosystemen wie Wäldern, Gewässern oder Stränden ab (Statistisches Bundesamt, 2025a). Diese Werte zeigen das breite Spektrum der Ökosystemleistungen und ihren direkten Einfluss auf Menschen und Wirtschaft.

Wie aber können solche Ökosystemleistungen umweltstatistisch bemessen und mit breiterem Blick auf Umwelt und Gesellschaft verstanden werden? Welche Datenangebote gibt es? Die Ökosystemrechnungen des Statistischen Bundesamtes liefern in drei aufeinander aufbauenden Konten zu Ausmaß (Bellinghen und andere, 2021), Zustand (Bellinghen und andere, 2024) und Leistungen räumliche Zeitreihendaten ab dem Berichtsjahr 2015 in Tabellen- und Kartenform. ➤ **Grafik 1** Dieser Artikel stellt mit den Leistungsbilanzen den dritten Teil dieses Berichts-

**Grafik 1**

Aufbau der Ökosystemrechnungen nach dem SEEA<sup>1</sup> der Vereinten Nationen



1 System of Environmental-Economic Accounting (2021).

systems vor und erläutert die dahinterliegenden biophysikalischen Modelle. Er zeigt in Fallbeispielen, wie durch die Verbindung von Informationen zu Flächen (wo gibt es welche Ökosysteme?) und qualitativen Eigenschaften (wie geht es diesen Ökosystemen?) ein ganzheitliches Informationspaket angeboten wird.

Die Ökosystemrechnungen folgen dem internationalen statistischen Standard „System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA EA)“ (Vereinte Nationen, 2024). Im Zuge der Berichtspflicht nach der Verordnung (EU) Nr. 691/2011 über europäische umweltökonomische Gesamtrechnungen wurden die Methoden europaweit abgestimmt und die Ergebnisse somit vergleichbar gemacht. Die Ökosystemrechnungen komplementieren das weitere Datenangebot (unter anderem Umweltbelastungen wie Emissionen oder Energieverbrauch, Maßnahmen zum Umweltschutz oder umweltbezogene Steuern) der Umweltökonomischen Gesamtrechnungen, welche dem internationalen Rahmenwerk „SEEA Central Framework“ folgend die Wechselwirkungen zwischen Umwelt und Wirtschaft quantifizieren. Auch in Bezug auf die Volkswirtschaftlichen Gesamtrechnungen besteht

1 Siehe zum Beispiel [www.rbb24.de](http://www.rbb24.de)

2 Siehe zum Beispiel [www.statistikportal.de](http://www.statistikportal.de). Ländliche Regionen (vor allem in Nord- und Süddeutschland) gewannen relativ an Bedeutung, da sich die verbleibende touristische Nachfrage primär auf diese Räume konzentrierte.

3 Siehe [www.tagesschau.de](http://www.tagesschau.de)

durch einheitliche Accounting-Regeln die Möglichkeit, wirtschaftliche Indikatoren durch die bisher nicht erfassten Ökosystemleistungen zur ergänzen.

Die Ökosystemrechnungen basieren hauptsächlich auf räumlichen, möglichst hochaufgelösten Sekundärdaten. Durch eine automatisierte Verarbeitung werden die einzelnen Zeitschritte der Flächen-, Zustands- und Leistungsbilanzen berechnet. Diese bundesweiten, regelmäßigen, räumlichen Daten erfüllen zum einen die Anforderungen eines Gesamtrechnungssystems, zum anderen eignen sich die Ökosystemrechnungen zusätzlich ideal für Sonderauswertungen, das Ableiten von Indikatoren oder zur Integration in andere Berichtssysteme. Für jede dieser drei Bilanzen (Ökosystemfläche, Zustand und Leistungen) steht ein einheitliches Veröffentlichungsportfolio zur Verfügung:

- > Die **Bilanzen als Tabellenform**: Das Kernprodukt des Accounting-Systems stellt die Resultate im Zeitverlauf für verschiedene Ökosysteme in den jeweiligen administrativen Einheiten dar, beispielsweise die Fläche des Ökosystems Streuobstwiese in Bayern im Zeitverlauf. Im Fall der Leistungsbilanzen sind in einem [Statistischen Bericht](#) sowie in der Datenbank des Statistischen Bundesamtes [GENESIS-Online](#) sowohl Aufkommens- (welches Ökosystem stellt welche Leistung bereit?) als auch Verwendungstabellen (wer nutzt die Leistung?) verfügbar.
- > Der [Ökosystematlas](#) ist eine interaktive Sammlung von Karten zu Fläche, Zustand und Leistungen bis auf Gemeindeebene im Zeitverlauf und stellt diese als Geodaten zum Download bereit.
- > [Steckbriefe](#) zu den einzelnen Ökosystemtypen, Zustandsvariablen und Leistungen beschreiben die Datenquellen, Datenverarbeitung und -modelle sowie weitere Informationen transparent für die Nutzenden. Weitere Informationen zur Methodik der jeweiligen Bilanzen liefern standardisierte [Qualitätsberichte](#) beziehungsweise [Methodenberichte](#).

Die Bilanzen der Ökosystemrechnungen werden regelmäßig aktualisiert; Fläche und Zustand im dreijährigen Turnus und die Leistungsbilanz jährlich. Das Datenangebot reicht momentan von 2015 bis 2023 und wird kontinuierlich fortgeführt.

Der Artikel beschreibt in Kapitel 2 die Methoden und bundesweiten Ergebnisse für drei ausgewählte Ökosystemleistungen und zeigt in Kapitel 3 regionale Fallbeispiele für diese Leistungen. Er endet mit einem Ausblick auf eine potenzielle monetäre Bewertung von Ökosystemleistungen und fasst die Anwendungsfälle der Ökosystemrechnungen zusammen.

## 2

### Methoden und Ergebnisse

Ökosystemleistungen werden als Beiträge definiert, die Ökosysteme zum menschlichen Wohlbefinden leisten. Die Ökosystemrechnungen erfassen finale Ökosystemleistungen, also solche, die Gesellschaft und Wirtschaft direkt nutzen (Vereinte Nationen, 2024). Sie betrachten die Umwelt daher aus einer instrumentellen Perspektive. Ökosystemleistungen werden grundsätzlich in drei Kategorien eingeteilt. [Bereitstellungsleistungen](#) liefern direkt nutzbare Güter wie Holzzuwachs oder Ernteerträge. [Regulierungsleistungen](#), zum Beispiel Feinstaubfilterung, Kohlenstoffspeicherung, Bodenerosionsschutz und städtische Kühlung, stabilisieren natürliche Kreisläufe und mindern Umwelteinflüsse. [Kulturelle Leistungen](#), etwa der naturnahe Tourismus, spiegeln die immateriellen Vorteile wider, die Menschen aus naturnahen Räumen für Erholung, Wohlbefinden und kulturelle Erfahrung ziehen.

Bereitstellungsleistungen werden bereits in den Volkswirtschaftlichen Gesamtrechnungen erfasst. Die Ökosystemrechnungen verorten hier den Beitrag der Natur räumlich und schreiben ihn bestimmten Ökosystemen zu. Regulierungs- und kulturelle Leistungen werden erstmals in einem Berichtssystem der amtlichen Statistik berücksichtigt.

Die sechs bisher veröffentlichten Ökosystemleistungen werden ab 2026 durch drei weitere Leistungen rückwirkend ergänzt. ➤ [Übersicht 1](#)

All diesen Leistungen liegt entweder ein bestehendes statistisches Berichtssystem oder ein eigenes biophysikalisches Modell zugrunde. Ersteres verhindert, dass Doppelstrukturen aufgebaut werden

## Übersicht 1

### Bisher erfasste und mittelfristig geplante Ökosystemleistungen

|                           |                                                  |
|---------------------------|--------------------------------------------------|
| Bereitstellungsleistungen | Bereitstellung von Kulturpflanzen (Ernteerträge) |
|                           | Holzzuwachs                                      |
|                           | Bestäubung <sup>1</sup>                          |
| Regulierungsleistungen    | Luftfilterung (PM <sub>2,5</sub> )               |
|                           | Kohlenstoffspeicherung (Nettosequestrierung)     |
|                           | Kohlenstoffspeicherung (Bestände)                |
|                           | Kühlung in Städten                               |
|                           | Erosionsschutz                                   |
| Kulturelle Leistungen     | Naturnaher Tourismus                             |

■ ab 2026 geplante, rückwirkende Ergänzungen

1 Bestäubung kann als intermediäre, also nicht finale Leistung angesehen werden. Wegen der umweltpolitischen Relevanz gibt es aber auf europäischer Ebene eine Berichtspflicht, um sie als Teilleistung der Bereitstellung von Kulturpflanzen zur Ernte auszuweisen.

und Nutzerinnen und Nutzer mit verschiedenen und möglicherweise widersprüchlichen Daten konfrontiert werden. Beispiele dafür sind die Nutzung der Waldgesamtrechnungen und der Materialflussrechnungen aus den Umweltökonomischen Gesamtrechnungen für die Ökosystemleistungen „Holzzuwachs“ und „Bereitstellung von Kulturpflanzen“ oder die Ableitung der „Kohlenstoffspeicherung“ aus Daten der Treibhausgasberichterstattung (siehe Abschnitt 2.1). Sofern möglich, werden solche bereits bestehenden Werte anhand der Daten aus den Flächen- und Zustandsbilanzen den einzelnen Ökosystemtypen zugewiesen. Wenn kein bereits etabliertes Berichtssystem genutzt werden kann, wird ein speziell für die jeweilige Ökosystemleistung entwickeltes biophysikalisches Modell verwendet, sofern geeignete georeferenzierte Daten zur Berechnung verfügbar sind. Zwei Beispiele sind die Ableitung von Feinstaubfilterung aus Feinstaubkonzentration und Fernerkundungsdaten zu Blattflächen oder von Kühlung in Städten aus Temperatur- und Vegetationsindizes (siehe Abschnitt 2.3). Diese Modelle sind wissenschaftlich fundiert und international abgestimmt, erfordern aber eine Balance zwischen Komplexität und dem Gesamtrechnungscharakter der Ökosystemrechnungen (bundesweit, regelmäßig und vergleichbar).

## 2.1 Kohlenstoffspeicherung (Nettosequestrierung)

Für die Berechnung der Kohlenstoffspeicherung (Nettosequestrierung) werden Daten zu Treibhausgasemissionen aus der Landnutzung, Landnutzungsänderung und Forstwirtschaft (Land Use, Land Use-Change and Forestry – LULUCF) verwendet. Diese stellt das Johann Heinrich von Thünen-Institut, das Bundesforschungsinstitut für Ländliche Räume, Wald und Fischerei, als Teil der Klimaberichterstattung an die Vereinten Nationen zusammen. Berechnungsgrundlage sind die von der Europäischen Union (EU) empfohlenen Methoden zur Erfassung von CO<sub>2</sub>-Nettoemissionen in Böden, Biomasse, Streu- beziehungsweise Totholz sowie weitere Kohlenstoffpools. Holz- und Torfprodukte berücksichtigt die Ökosystemleistung nicht.

Die Landnutzungsdaten der Treibhausgasberichterstattung sind nicht direkt mit den Ökosystemen der Flächenbilanz vergleichbar. Beide Datensätze liegen zwar in einer 100-m-Auflösung vor, die LULUCF-Daten jedoch als Punkte-Grid mit punktspezifischer Landnutzungskategorie. Hingegen sind die Ökosystemflächen als Polygone mit Information zum jeweiligen Ökosystemtyp charakterisiert. Mithilfe von Emissionsfaktoren (Tonnen CO<sub>2</sub> je Hektar je Landnutzungskategorie) und einer Überführungsmatrix werden diese Kategorien anschließend der nationalen [Ökosystemklassifikation](#) zugeordnet und die Emissionsfaktoren auf die Ökosystemtypen der Flächenbilanz übertragen. Die Werte werden abschließend in Kohlenstoff umgerechnet. Dieses Modell macht die Daten der Treibhausgasberichterstattung für die Ökosystemrechnungen nutzbar.

Unterschiede zwischen den beiden Berichtssystemen werden transparent dargestellt und kommuniziert. Somit können widersprüchliche Ergebnisse vermieden werden.

Positive Werte kennzeichnen eine Nettosequestrierung, negative Werte Nettoemissionen. Auch bei negativen Werten kann ein Ökosystem jedoch eine wichtige regulierende Funktion erfüllen, indem es höhere Emissionen – etwa aus organischen Böden – verhindert. Wichtig ist hier zu beachten, dass neben den jährlichen Veränderungen (also den Flussgrößen

der Sequestrierung oder Emission) viele Ökosysteme über sehr große Kohlenstoffspeicher als Bestände verfügen. Dadurch kann ein Ökosystem bei mittel- und langfristiger Betrachtung eine wichtige regulierende Funktion erfüllen, indem diese Bestände langfristig geschützt werden – selbst wenn in Einzeljahren die Emissionen höher sind als die Einspeicherung von Kohlenstoff. Dies wird künftig in der separaten Ökosystemleistung Kohlenstoffspeicherung (Bestände) erfasst.

➤ **Grafik 2** zeigt ab dem Dürrejahr 2018 einen deutlichen Umschlag zu negativen Werten bei der Kohlenstoffspeicherung. Nach einer Kohlenstoffaufnahme der deutschen Wälder mit etwa 1,8 Millionen Tonnen kippt der Trend 2018 deutlich: Mit –20,1 Millionen Tonnen erreicht die Nettosequestrierung ihr bisheriges Minimum. In den Folgejahren bleiben die Werte negativ und schwanken zwischen –15 Millionen und –19 Millionen Tonnen. Damit zeigt sich ein nachhaltiger Verlust der Kohlenstoffsinken-Funktion seit 2018.

## 2.2 Umweltbezogener Tourismus

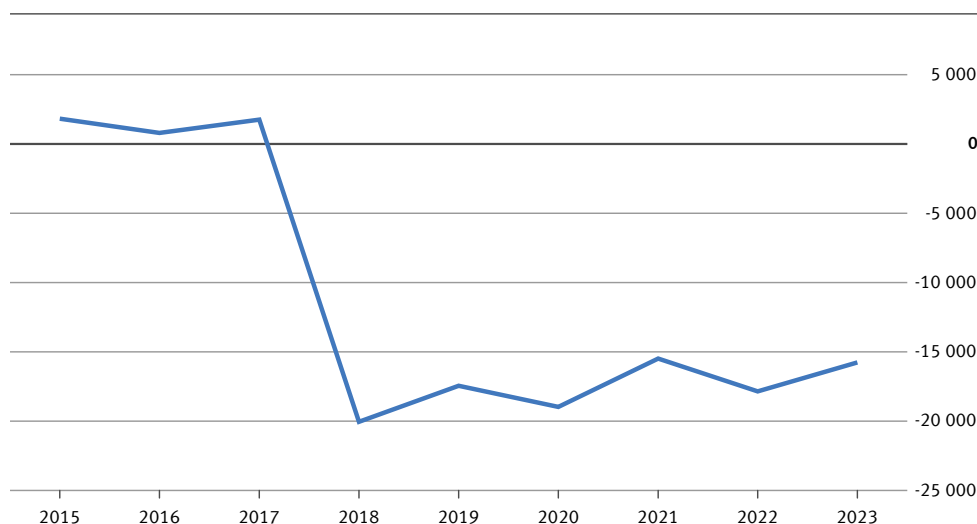
Der naturnahe Tourismus wird auf Basis der touristischen Übernachtungen je Gemeinde berechnet. Entsprechend wird Tagestourismus hierbei nicht berücksichtigt. Übernachtungen in Kurzzeitunterkünften werden über Plattformdaten ergänzt und Geschäftsreisen abgezogen. Alle Übernachtungen werden zunächst nach Herkunft der Gäste differenziert.

Im zweiten Schritt wird der Anteil der Übernachtungen mit Naturbezug je Gemeinde ermittelt, basierend auf bundesweiten Befragungen (GfK Destinationsmonitor), wie wichtig Gästen Naturerleben ist. Die Verteilung der Übernachtungen mit Naturbezug auf Gemeinden berücksichtigt die landschaftliche Attraktivität und die Einwohnerzahl. Landschaftlich attraktive Gebiete verzeichnen mehr naturnahe Übernachtungen, da dort ein überdurchschnittlich hoher Nutzungsanteil für Naturerholung angenommen wird. Dicht besiedelten Gebieten werden weniger naturnahe Übernachtungen zugeordnet, da dort städtische Freizeitaktivitäten eine wichtigere Rolle spielen.

Die ästhetische Landschaftsqualität wird mithilfe einer bundesweiten Karte modelliert, die auf der Vielfalt, Natürlichkeit und Eigenart der Landschaft

**Grafik 2**

**Entwicklung der Ökosystemleistung Kohlenstoffspeicherung (Nettosequestrierung) in Deutschland**  
in 1 000 t



basiert. Das Statistische Bundesamt erstellt die Karte der ästhetischen Landschaftsqualität dreijährlich. Dafür werden unter anderem Ökosystemverteilung, Reliefmerkmale, Strukturelemente wie Hecken, wahrgenommene Natürlichkeit (einschließlich Abwesenheit störender Bauwerke) sowie die Seltenheit von Landschaftstypen einbezogen. Alle Parameter werden mit einer Auflösung von 100 m berechnet, normiert und entsprechend ihrer Bedeutung gewichtet. Die Kombination dieser Komponenten ergibt einen Gesamtwert der landschaftlichen Attraktivität (Hermes und andere, 2023).

Die Zahl der Übernachtungen im naturnahen Tourismus ist ab 2018 verfügbar, 2019 gab es 247 Millionen Übernachtungen. Aufgrund der pandemiebedingten Einschränkungen im Folgejahr sanken die Übernachtungen deutlich auf rund 197 Millionen und fielen 2021 weiter auf rund 186 Millionen. Ab 2022 erholte sich der naturnahe Tourismus stark auf 243 Millionen Übernachtungen und 2023 wurde mit rund 265 Millionen der höchste Wert im betrachteten Zeitraum erreicht. ➔ [Grafik 3](#)

Gleichzeitig ging der gesamte Tourismus während der Pandemie noch deutlich stärker zurück: Die Zahl der Übernachtungen sank von 496 Millionen auf 302 Millionen während der COVID-19-Pandemie, was einem Rückgang von rund 39 % entspricht. Der naturbezogene Tourismus war davon weniger stark betroffen

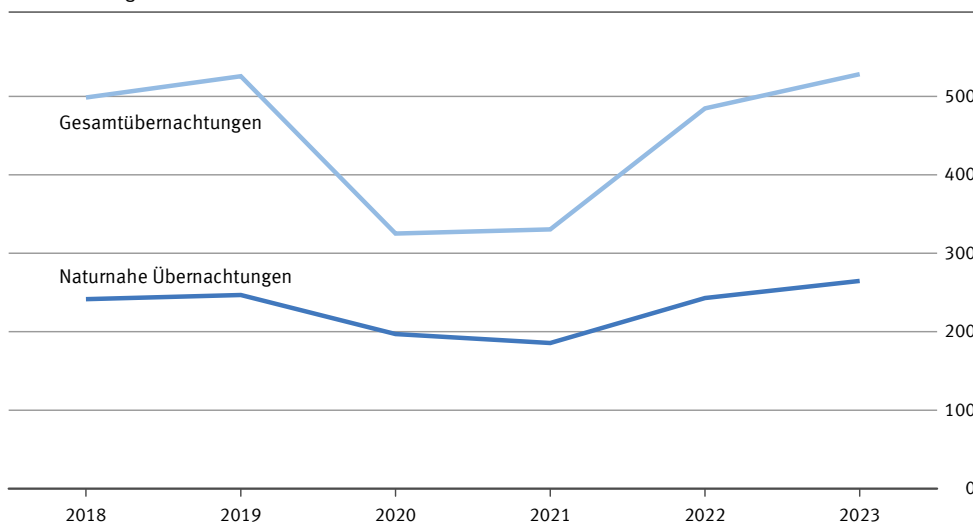
und verringerte sich um etwa 20 %, also nur etwa halb so stark wie der Gesamttourismus. Somit lässt sich feststellen, dass während der Pandemie, die mit Einschränkungen im Sozialleben und des kulturellen Angebots einherging, die relative Bedeutung des naturnahen Tourismus zugenommen hat.

## 2.3 Kühlung in Städten

Vegetation in Städten, wie Baumreihen, Parks oder Dachbegrünung, trägt durch Schattenwurf und Verdunstung dazu bei, die Temperatur in Städten im Sommer zu regulieren. Für die Berechnung dieser Kühlleistung der Stadtvegetation wurde eine regressionsbasierte Methode (Venter und andere, 2020; Marando und andere, 2022) gewählt, die verschiedene räumliche Daten kombiniert: Flächenbilanz der Ökosysteme, Oberflächentemperaturen, Evapotranspiration, Lufttemperatur- und Winddaten sowie Vegetationsinformationen wie Kronendichte und der Vegetationsindex NDVI (Normalized Difference Vegetation Index). Auch Bodenversiegelung und Gebäude-dichte fließen ein (Statistisches Bundesamt, 2025b). Die Analyse bezieht sich auf Städte nach der europäischen Klassifikation des [Verstädterungsgrads](#). Aufgrund der beschränkten zeitlichen Auflösung der Daten zur Oberflächentemperatur wird die Leistung als Mittelwert jeweils für die Sommermonate Juli und August berechnet. Zunächst wird für jede Stadt

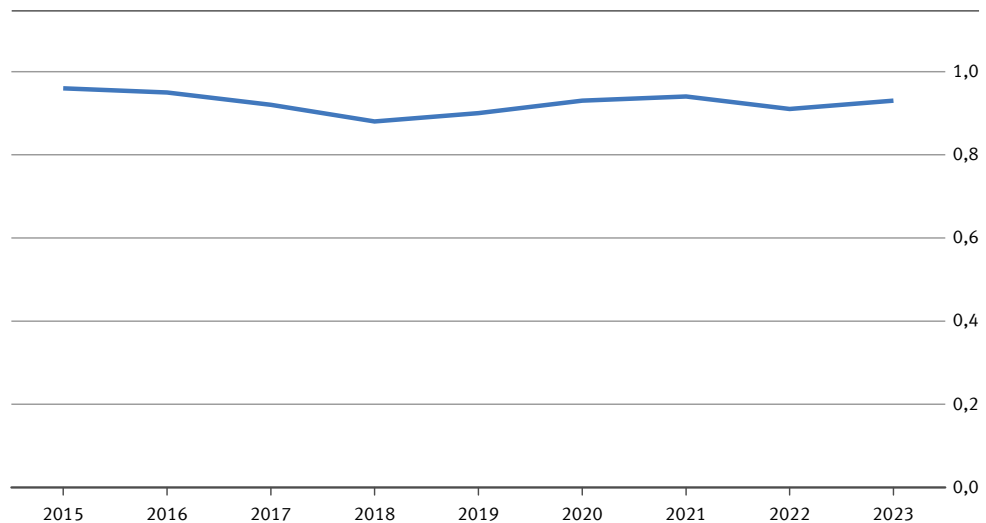
**Grafik 3**

**Entwicklung der Ökosystemleistung Naturnaher Tourismus in Deutschland**  
Übernachtungen in Mill.



**Grafik 4**

**Entwicklung der Ökosystemleistung Kühlung in Städten in Deutschland**  
in °C



auf Rasterebene (100 m) ein Regressionsmodell geschätzt, das zeigt, wie Vegetation die Oberflächentemperatur beeinflusst. Als Faktoren werden Kronendichte, Vegetationsindex NDVI und Verdunstung berücksichtigt – nicht nur der Standort selbst, sondern auch das Umfeld in einem 300-m-Radius. Kontrollvariablen wie Bebauungsdichte und klimatische Parameter ergänzen das Modell. Dieses Modell zeigt die negative Korrelation von Oberflächentemperatur und Vegetation, ausgelöst durch Beschattung und Verdunstung.

Die geschätzten Koeffizienten werden verwendet, um ein Szenario ohne Vegetation zu simulieren: Die Vegetationswerte in allen Rasterzellen werden auf null gesetzt, um im Modell eine fiktive Oberflächentemperatur ohne Kühlwirkung durch Vegetation zu schätzen. Die Differenz zwischen tatsächlicher (gemessener) und modellierter (Szenario ohne Vegetation) Temperatur stellt die Kühlleistung der Vegetation an der Oberfläche dar.

In einem zweiten Schritt wird ein weiteres Modell berechnet, das den Zusammenhang zwischen Oberflächen- und Lufttemperatur beschreibt. Dieser Schritt ist notwendig, da nur die Daten zur Oberflächentemperatur hochaufgelöst (100-m-Auflösung) vorhanden sind, letztlich aber eine Aussage über die Kühlungswirkung auf die Lufttemperatur getroffen

werden soll. Dieses Modell wird mit Daten aus Rasterzellen im Umfeld von Wetterstationen trainiert und dann auf alle Städte angewandt. Die Kühlleistung an der Oberfläche wird so in eine entsprechende Kühlung der Lufttemperatur übersetzt.

➤ **Grafik 4** zeigt, dass die Kühlleistung in Städten im Beobachtungszeitraum leicht abgenommen hat. Die durchschnittliche Temperaturreduktion lag im Jahr 2015 noch bei 0,96 °C und sank bis 2018 auf ihr Minimum von 0,88 °C. Danach steigen die Werte wieder moderat an und bewegen sich ab 2019 zwischen 0,9 °C und 0,94 °C. Unterschiede in der Kühlleistung hängen sowohl vom Ausmaß und Zustand der Vegetation (Leistungspotenzial) als auch von den herrschenden Temperaturen ab.

## 3

### Fallbeispiele

#### 3.1 Wälder im Bergischen Land

Die ausbleibenden Niederschläge im Jahr 2018 hatten in den Wäldern Deutschlands nachhaltige Folgen. Besonders betroffen waren die Mittelgebirge,



wo die Böden über weite Teile des Jahres ungewöhnlich trocken waren. Dieser Trockenstress schwächte vor allem Fichten- und Kiefernbestände und begünstigte massive Borkenkäferbefälle sowie ein erhöhtes Absterben von Bäumen.

➤ **Grafik 5** zeigt die Flächenbilanz der Ökosysteme in der Region um Gummersbach im Oberbergischen Kreis (Nordrhein-Westfalen), die beispielhaft für viele Mittelgebirgslagen Deutschlands steht. Dort veränderte sich die Waldstruktur zwischen 2015 und 2021 deutlich: Nadelwälder starben ab und wurden durch Lichtungen und Naturverjüngungsflächen ersetzt. Diese Flächen stiegen in Nordrhein-Westfalen innerhalb von sechs Jahren um das etwa 336-Fache von 238 ha auf 80 115 ha, während die

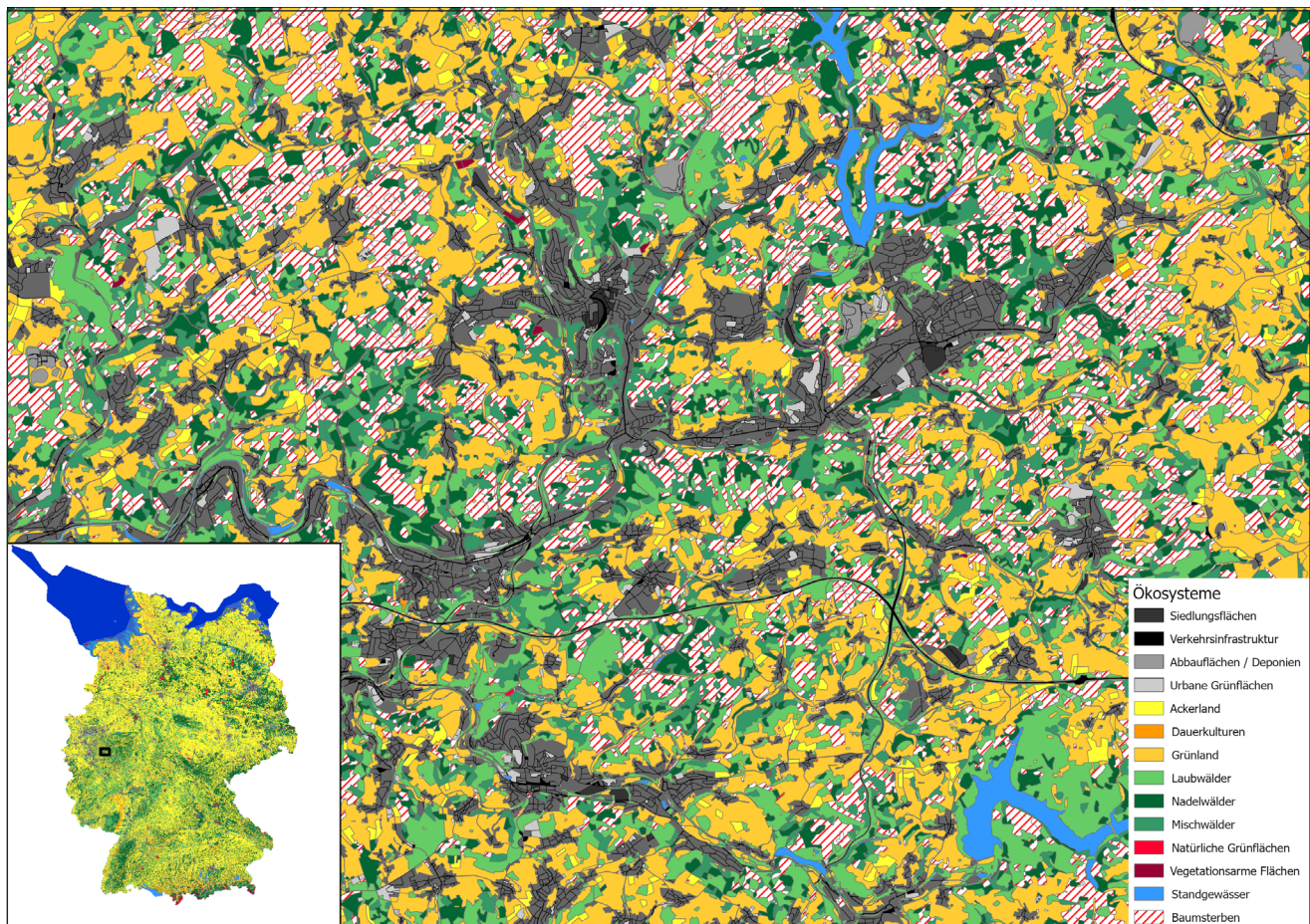
Fläche der Nadelwälder im gleichen Zeitraum von 332 675 ha auf 261 323 ha zurückging. Flächen, auf denen eine solche Änderung stattfand, sind in Grafik 5 rot schraffiert.

Das Absterben der Altbestände führte zu einem drastischen Rückgang der Ökosystemleistung Kohlenstoffspeicherung. Zwar kompensierten die deutschen Wälder 2017 durch Nettoaufnahme noch etwa den jährlichen CO<sub>2</sub>-Ausstoß von rund 120 457 Haushalten<sup>4</sup>, im Jahr 2018 jedoch kehrte sich die Bilanz um: Die freigesetzte Kohlenstoffmenge entsprach nun in etwa dem Ausstoß von 1,4 Millionen Haushalten, da abgestorbene Bäume und Zersetzungspro-

4 Zum CO<sub>2</sub>-Fußabdruck der privaten Haushalte siehe [www.destatis.de](http://www.destatis.de)

#### Grafik 5

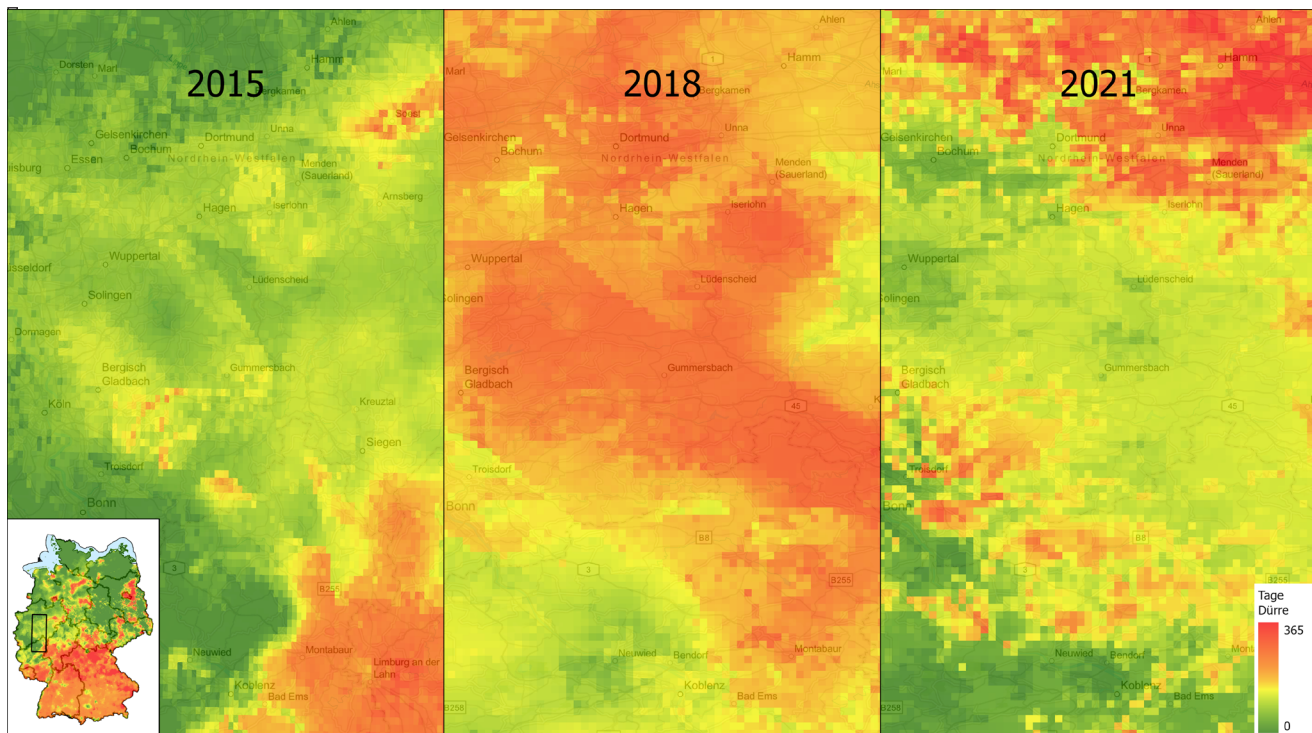
Baumsterben (rot schraffiert) von 2015 auf 2021 im Bergischen Land (Region um Gummersbach im Oberbergischen Kreis)





## Grafik 6

Entwicklung der jährlichen Dürretage im Bergischen Land (Region um Gummersbach im Oberbergischen Kreis)



zesse große Mengen CO<sub>2</sub> freisetzten. Damit zeigten die Mittelgebirgswälder exemplarisch die nachhaltige Schwächung der Kohlenstoffsinken-Funktion seit der Dürre.

Die Kreisstadt Gummersbach war, wie viele weitere Mittelgebirgsgemeinden, besonders von der Dürre betroffen. Wie aus den Daten der Zustandsbilanz der Ökosysteme hervorgeht, litten die Wälder dort an 226 Tagen des Jahres an einer besonderen Trockenheit. Damit gehörte sie zu dem Drittel aller deutschen Gemeinden, die 2018 mehr als 200 Dürretage verzeichneten. Die Dürre 2018 war nicht nur besonders intensiv, sondern betraf zudem weite Teile Deutschlands. Auch im Vergleich zu den anderen Berichtsjahren ist das Ausnahmejahr deutlich zu erkennen. [↗ Grafik 6](#) zeigt die zugrunde liegende Rasterkarte mit einer Auflösung von 1 km in der Region um Gummersbach.

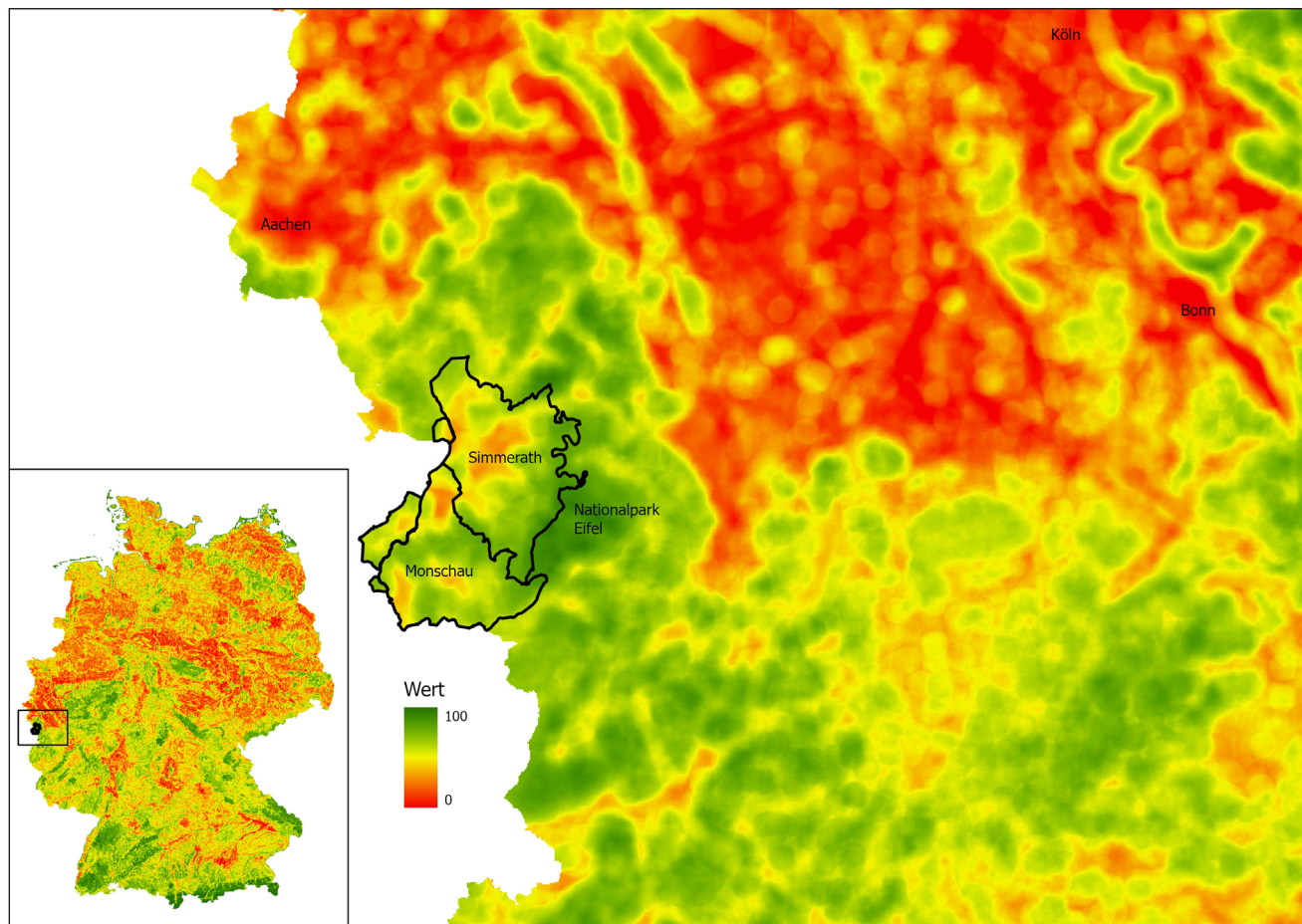
## 3.2 Tourismus in der Eifel

Die Eifel ist mit ihrer abwechslungsreichen Landschaft ein wichtiges Ziel für naturnahe Erholung, insbesondere für Besucherinnen und Besucher aus den nahe gelegenen Ballungsräumen des Rheinlands. Orte wie Simmerath oder Monschau verzeichnen jährlich über 200 000 naturnahe Übernachtungen und zeigen damit die hohe Nachfrage nach Landschaftserleben in der Region.

[↗ Grafik 7](#) macht mit der Karte im oberen Teil die besondere landschaftliche Qualität der Eifel sichtbar. Der Nationalpark Eifel sticht dabei als Gebiet hervor, in dem natürliche Prozesse Vorrang haben und Natur intensiver erfahrbar wird. Im Modell weist die Eifel dementsprechend in den Subindikatoren zum Anteil natürlicher Ökosysteme (zum Beispiel Feuchtgebiete und Laubwälder), zur Abwesenheit störender Elemente (unter anderem Straßen oder Stromleitungen), zur Diversität der Ökosysteme sowie zur Heterogenität der Landschaft besonders hohe Werte auf. Die angrenzenden Gemeinden Simmerath und Monschau profitieren von dieser hohen landschaftlichen

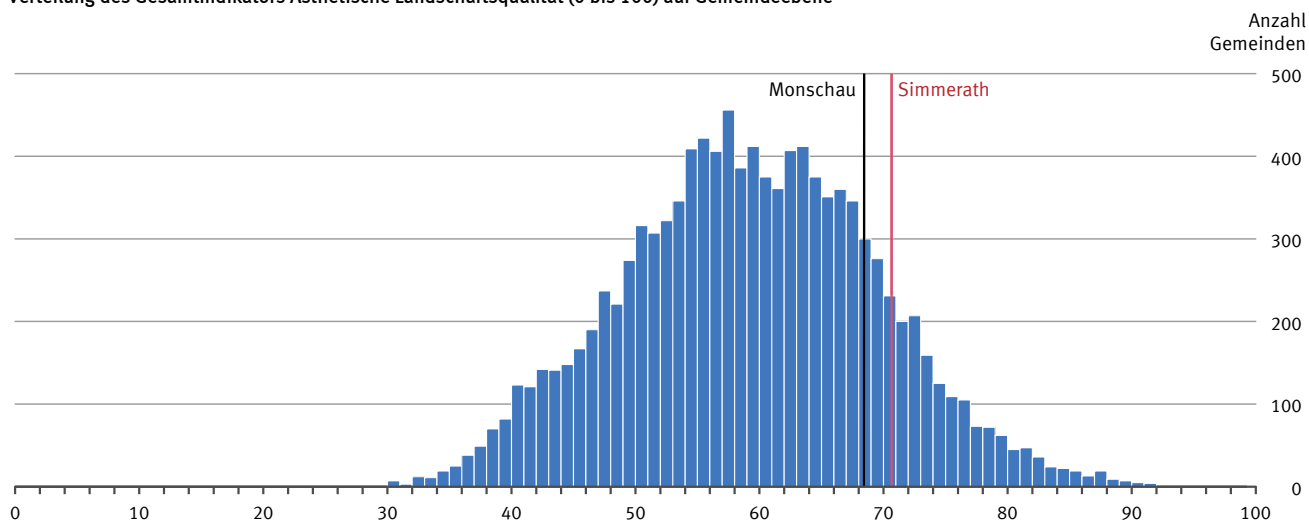
## Grafik 7

Karte der ästhetischen Qualität der Landschaft für die Eifelregion 2023



Grün entspricht einer hohen Landschaftsqualität.

Verteilung des Gesamtindikators Ästhetische Landschaftsqualität (0 bis 100) auf Gemeindeebene



Qualität des Nationalparks und schneiden im Ranking aller Gemeinden überdurchschnittlich ab (siehe den unteren Teil der Grafik 7).

Die Nutzung dieser Landschaft für Naherholung schafft nicht nur wirtschaftliche Impulse, sondern stellt zugleich eine zentrale kulturelle Ökosystemleistung dar: Sie trägt zum Wohlbefinden der Menschen bei und stärkt die Wertschätzung intakter Naturräume. Die Verwendungstabelle der Ökosystemleistungsbilanz zeigt auch die transnationale Bedeutung des umweltbezogenen Tourismus: Rund 38 % dieser Übernachtungen entfallen auf ausländische Gäste.

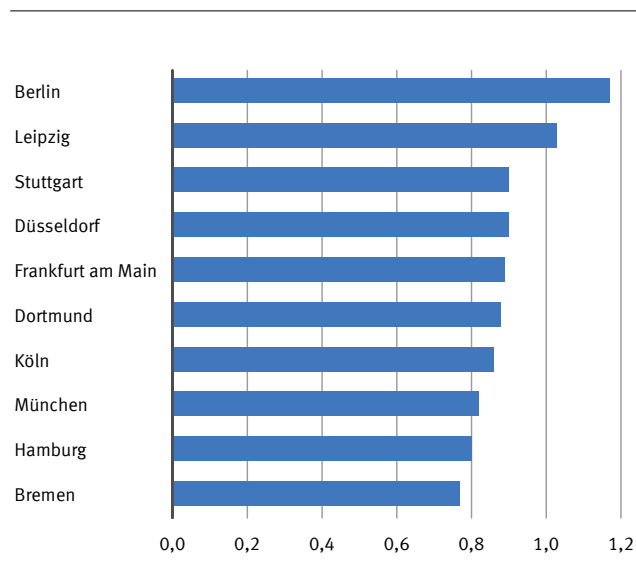
Die direkte Zuordnung von naturnahen Übernachtungszahlen zur landschaftlichen Qualität einzelner Gemeinden ist nur eingeschränkt möglich. Aufenthaltsorte dienen häufig als Ausgangspunkt für Ausflüge in benachbarte attraktive Gebiete, sodass sich das tatsächliche Nutzungsverhalten räumlich über Gemeindegrenzen hinweg verteilt. Für die Interpretation bedeutet dies, dass die Werte eher regionale Tendenzen und räumliche Schwerpunkte der Erholungsnutzung abbilden, nicht jedoch eine präzise gemeindeschärfe Wirkung. Die Ergebnisse liefern damit wertvolle Hinweise, sollten aber stets im Kontext des größeren Landschaftsraums betrachtet werden.

### 3.3 Kühlung durch Vegetation in Berlin

Auch vor der geplanten Pflanzung von hunderttausenden neuen Bäumen bis 2040 durch das Berliner BäumePlus-Gesetz leisten die städtischen Grünflächen und Bäume bereits einen substanziellen Beitrag zur Hitzeregulierung in der Hauptstadt. Tatsächlich stand Berlin 2023 unter den zehn größten Städten Deutschlands mit einer Kühlleistung von 1,17 °C ganz vorne. Die Ergebnisse beziehen sich auf die Hitzereduktion durch städtische Vegetation an einem durchschnittlichen Sommertag (Juli, August). Dies bedeutet, dass die Kühlleistung an besonders heißen Tagen noch höher ausfällt. Eine tägliche Modellierung der Ökosystemleistung ist derzeit noch nicht möglich, da räumlich hochaufgelöste Satellitendaten zur Oberflächentemperatur nicht dementsprechend frequent vorliegen. ➤ Grafik 8

**Grafik 8**

Kühlung durch Vegetation in den zehn bevölkerungsreichsten Städten Deutschlands 2023  
in °C



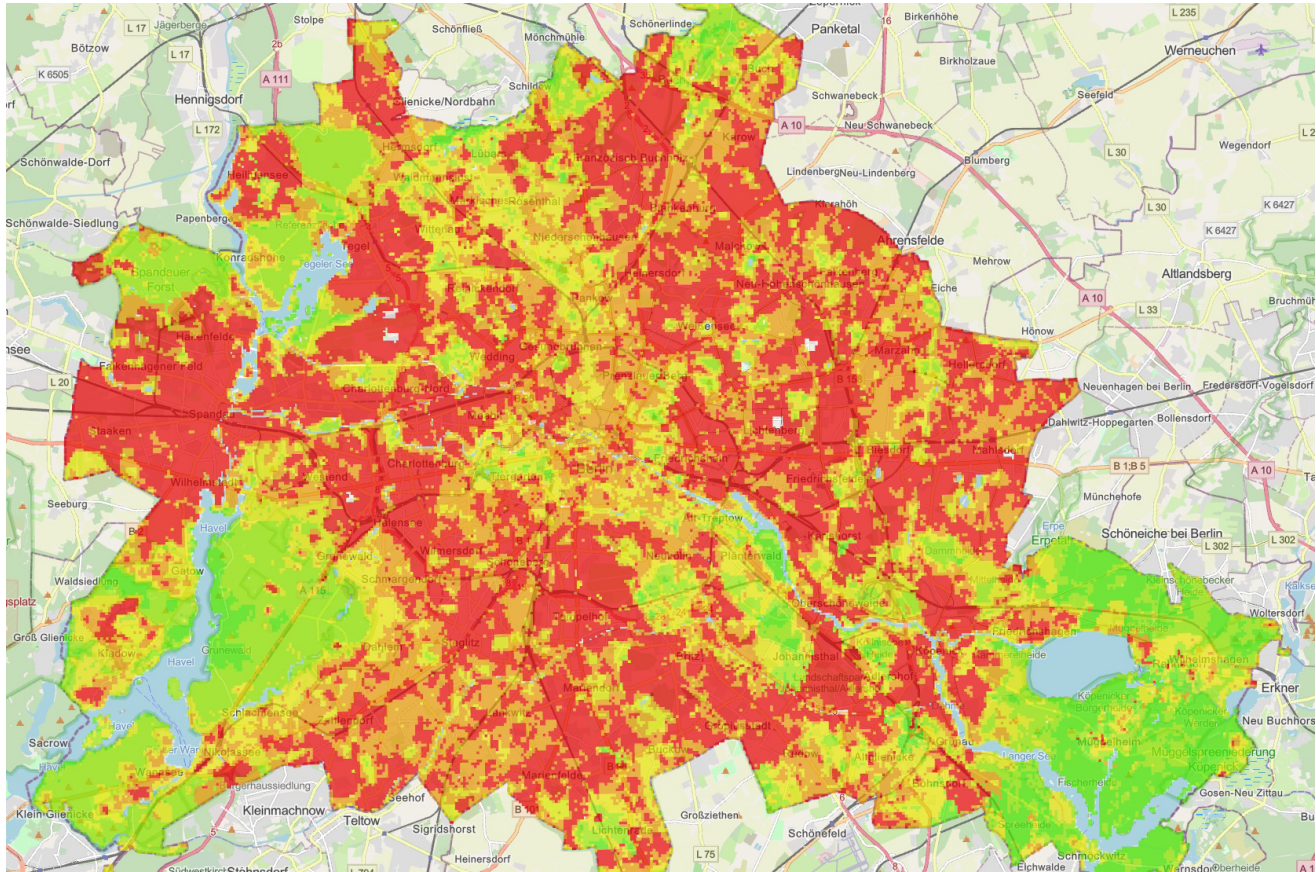
Quelle: Ökosystematlas

➤ Grafik 9 auf Seite 52 zeigt, dass Berlin im Sommer vor allem von großen kühlenden städtischen Grünflächen wie dem Tiergarten oder dem Volkspark Rehberge sowie Stadtwäldern wie dem Spandauer Forst oder dem Grunewald profitiert. Die Verteilung der Ökosystemleistung zeigt aber auch, dass vor allem in den Wohngebieten Unterschiede in der Baumkronenbedeckung zu hoher (orange) oder niedriger (rot) Kühlleistung führen. Die Aufkommenstabelle der Leistungsbilanzen zeigt dies mit durchschnittlichen Kühlleistungen von ganzen 1,68 °C für Berliner Wälder und nur 0,73 °C für Vegetation auf Siedlungsflächen und Verkehrsinfrastruktur.



## Grafik 9

### Verteilung der Ökosystemleistung „Kühlung in Städten“ in Berlin 2022



Durchschnittliche Temperaturreduktion durch Vegetation im Sommer; grün: hohe Kühlleistung, rot: keine bis niedrige Kühlleistung.  
Quelle Hintergrundkarte: OpenStreetMap

## 4

### Ausblick und Anwendungen

Die Ökosystemrechnungen des Statistischen Bundesamtes stellen Daten zu Flächen, Zustand und Leistungen in einem in sich konsistenten und einheitlichen Kontensystem dar. Diese Konten werden regelmäßig und durch eine zentrale und semi-automatisierte Berechnung auch ressourceneffizient aktualisiert. Das auf Python-Skripten basierte Produktionssystem kann flexibel auf neue Datenquellen, methodische Änderungen oder Erweiterungen angepasst werden.

**Erweiterung und Weiterentwicklung:** Die Zustandsbilanz der Ökosysteme gibt einen weitgehend ganz-

heitlichen Überblick über qualitative Merkmale einzelner Ökosystemtypen. Eine Integration zusätzlicher Zustandsvariablen, etwa im Bereich Biodiversität oder Konnektivität der Landschaft, ist im Falle verbesserter Datenverfügbarkeit geplant. Die Leistungsbilanz weist eine spezifische Auswahl der zahlreichen Leistungen der Natur für den Menschen aus. Weitere Leistungen, wie Bereitstellung von Fisch, Überflutungsschutz, Erosionsschutz oder Naherholung (Tagesausflüge), werden auf EU-Ebene bereits methodisch vorbereitet und sollen die Ökosystemrechnungen perspektivisch erweitern.

Ähnlich wie bei neuen Datenquellen verhält es sich bei methodischen Weiterentwicklungen, entweder bei den zugrunde liegenden Inputdaten oder den biophysikalischen Modellen der Ökosystemrechnungen. Methodische Verbesserungen bei der Treib-

hausgasberichterstattung werden in Zukunft zu einer höheren räumlichen Auflösung der Ergebnisse führen. Neue nationale Parameter für Depositionsgeschwindigkeiten erhöhen die Qualität des Modells für Luftfilterung. Die Ökosystemrechnungen können solche Innovationen schnell und ressourceneffizient als Revision der gesamten Zeitreihe integrieren.

**Monetäre Ökosystemleistungsbilanz:** Konzeptionell sieht das Kontensystem der Ökosystemrechnungen nach internationalem Rahmenwerk auch eine monetäre Leistungsbilanz vor. Diese soll monetäre Werte für die einzelnen Ökosystemleistungen ausweisen, um diese leistungsübergreifend und mit anderen, in den Volkswirtschaftlichen Gesamtrechnungen erfassten Gütern und Dienstleistungen vergleichbar zu machen. Monetäre Werte können, verglichen mit den verschiedenen physischen Einheiten der Ökosystemleistungen, für Nutzende und politische Entscheidungsträger leichter verständlich und kommunizierbar sein. Um mit den Volkswirtschaftlichen Gesamtrechnungen konsistent zu sein, wird angestrebt, die Ökosystemleistungen über ihren Tauschwert/Marktpreis zu bewerten. Da Ökosystemleistungen allerdings kaum auf Märkten gehandelt werden, müssten ihr Preise durch teils komplexe Methoden zugeschätzt oder approximiert werden (beispielsweise durch Marktpreise für Substitute oder kostenbasierte Ansätze). Diese Methodik wird derzeit international abgestimmt und unter anderem in einer Machbarkeitsstudie der EU-Kommission untersucht, die bis zum Sommer 2026 vorliegen wird. So wird unter anderem auch geklärt, wie mit multiplen Werten und verschiedenen Nutzendenanforderungen verfahren werden soll und wie diese monetären Werte möglichst transparent kommuniziert werden können.

**Anwendungen:** Die Daten der Flächen-, Zustands- und Leistungsbilanzen haben eine breite Nutzenden-schaft; von der breiten Öffentlichkeit zur Bewusstseinsbildung, über politische Entscheidungsträger als Informationsgrundlage für Maßnahmen oder Evaluierung bis hin zu Forschenden, die Daten der Ökosystemrechnungen weiterverarbeiten. Die Anwendungsfälle können grob in drei Gruppen mit folgenden Beispielen unterteilt werden:

1. Datenquelle zur Erfüllung verschiedener Gesetzgebungen: Mit der Verabschiedung der europä-

ischen Gesetzgebung zur Wiederherstellung der Natur, der Umsetzung durch das Aktionsprogramm Natürlicher Klimaschutz und der EU-Verordnung über europäische umweltökonomische Gesamtrechnungen entstehen neue Anforderungen an ein Umweltmonitoring in Deutschland. Die Ökosystemrechnungen bilden über die einzelnen Bilanzen ein konsistentes System für ein effektives Umweltmonitoring über Raum und Zeit hinweg. Damit können Datenbedarfe unterschiedlicher Interessensgruppen bedient werden. Die hohe räumliche Auflösung ermöglicht es auch, lokale Veränderungen der Ökosysteme hinsichtlich Fläche, Zustand und Leistungen der Ökosysteme auszuweisen.

2. Indikatoren und Sonderauswertungen: Das Produktionssystem Ökosystemrechnungen eignet sich ideal dazu, maßgeschneiderte Indikatoren für verschiedene Zwecke abzuleiten, wie Indikatoren der Sustainable Development Goals der Vereinten Nationen, Indikatoren der Deutschen Nachhaltigkeitsstrategie, des Global Biodiversity Frameworks, der Nachhaltigkeitsberichterstattung von Unternehmen oder des Monitorings für die Wiederherstellungsverordnung (EU-Verordnung 2024/1991). Auch thematische oder regionale Auswertungen, etwa für Schutzgebiete oder Flussgebietseinheiten, sind rasch umsetzbar.
3. Im Falle einer monetären Bewertung: Integration in makroökonomische Modelle und Risikoanalysen, um Ökosystemleistungen erweiterte Input-Output-Tabellen der Volkswirtschaftlichen Gesamtrechnungen. [u](#)

## LITERATURVERZEICHNIS

---

Bellinghen, Marius/Felgendreher, Simon/Ohrlein, Johannes/Schürz, Simon/Arnold, Stephan. *Ökosystemgesamtrechnungen – Flächenbilanzierung der Ökosysteme (Extent Account)*. In: WISTA Wirtschaft und Statistik. Ausgabe 6/2021, Seite 31 ff.

Bellinghen, Marius/Felgendreher, Simon/Reith, Jonathan/Schürz, Simon. *Wie steht es um die Ökosysteme in Deutschland? Einführung einer regelmäßigen Zustandsbilanzierung*. In: WISTA Wirtschaft und Statistik. Ausgabe 1/2024, Seite 80 ff.

Hermes, Johannes/Albert, Christian/Schmücker, Dirk/Bredemeier, Birte/Barkmann, Jan/von Haaren, Christina. *Erfassung und Bewertung kultureller Ökosystemleistungen in Deutschland. Die Qualität der Landschaft für Freizeit- und Wochenenderholung in Deutschland: Potenzial, Dargebot, Präferenzen, Nutzung*. In: BfN-Schriften 659/2023. DOI: [10.19217/skr659](https://doi.org/10.19217/skr659)

Marando, Federica/Heris, Mehdi P./Zulian, Grazia/Udías, Angel/Mentaschi, Lorenzo/Chrysoulakis, Nektarios/Parastatidis, David/Maes, Joachim. *Urban heat island mitigation by green infrastructure in European Functional Urban Areas*. In: Sustainable Cities and Society. Band 77, Februar 2022, Nr. 103564. DOI: [10.1016/j.scs.2021.103564](https://doi.org/10.1016/j.scs.2021.103564)

Statistisches Bundesamt. *Leistungsbilanz der Ökosysteme. Berichtszeitraum 2015-2023*. Statistischer Bericht. 2025a. [Zugriff am 3. Februar 2025]. Verfügbar unter: [www.destatis.de](https://www.destatis.de)

Statistisches Bundesamt. *Umweltökonomische Gesamtrechnungen: Leistungen der Ökosysteme. Kühlung in Städten*. Qualitätsbericht. 2025b. [Zugriff am 9. Februar 2026]. Verfügbar unter: [www.destatis.de](https://www.destatis.de)

Vereinte Nationen. *System of Environmental-Economic Accounting – Ecosystem Accounting (SEEA-EA)*. 2024. [Zugriff am 28. November 2025]. Verfügbar unter: [seea.un.org](https://seea.un.org)

Venter, Zander S./Hjertager Krog, Norun/Barton, David N. *Linking green infrastructure to urban heat and human health risk mitigation in Oslo, Norway*. In: Science of The Total Environment. Band 709, März 2020, Nr. 136193. DOI: [10.1016/j.scitotenv.2019.136193](https://doi.org/10.1016/j.scitotenv.2019.136193)

## Rechtsgrundlagen

---

Berliner Klimaanpassungsgesetz (KAnGBln) vom 7. November 2025 (GVBl. Seite 542).

Verordnung (EU) Nr. 691/2011 des Europäischen Parlaments und des Rates vom 6. Juli 2011 über europäische umweltökonomische Gesamtrechnungen (Amtsblatt der EU Nr. L 192, Seite 1), zuletzt geändert durch die Delegierte Verordnung (EU) 2025/1131 der Kommission vom 26. März 2025 (Amtsblatt der EU Nr. L 1131, Seite 1 ff.).

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).





**Dr. Niels Gillmann**

studierte Volkswirtschaftslehre in Halle (Saale), Saint-Denis (Réunion) und Uppsala (Schweden). Er promovierte in Zeitreihenökonomie an der Technischen Universität Dresden. Seit 2024 ist er im Referat „Staatssektor, EU-Stabilitätspakt“ des Statistischen Bundesamtes beschäftigt. Dort befasst er sich mit dem öffentlichen Personennahverkehr und Fragen der Sektorzuordnung. Zuvor war er als Ökonom am ifo Institut tätig.

# DIE REKLASSIFIKATION VON VERKEHRSINFRASTRUKTUREINHEITEN IN DEN STAATSEKTOR

Niels Gillmann

➤ **Schlüsselwörter:** Volkswirtschaftliche Gesamtrechnungen – Staat – ÖPNV – Abgrenzung des Staatssektors – Eurostat Advice

## ZUSAMMENFASSUNG

Aufgrund einer methodischen Empfehlung von Eurostat, dem Statistischen Amt der Europäischen Union, waren die Einheiten des öffentlichen Personennahverkehrs und der Bahninfrastruktur in den Staatssektor der Volkswirtschaftlichen Gesamtrechnungen zu integrieren. Der Aufsatz beschreibt die Ausgangslage vor der Reklassifikation, die methodischen und datenbedingten Herausforderungen, das Vorgehen und die Ergebnisse der Reklassifikation.

➤ **Keywords:** national accounts – general government – short-distance public passenger transport – delimitation of general government sector – Eurostat advice

## ABSTRACT

*Following methodological advice provided by Eurostat, the Statistical Office of the European Union, it was necessary to integrate local public passenger transport and railway infrastructure units into the general government sector of national accounts. This article describes the situation before the reclassification, the methodological and data-related challenges, the procedure and the results of the reclassification.*

## 1

### Einleitung

Damit die Daten der Mitgliedstaaten der Europäischen Union (EU) im Zuge des Verfahrens bei einem übermäßigen Defizit vergleichbar sind, übermittelt das Statistische Bundesamt kontinuierlich Daten zum Schuldenstand und Defizit an die Europäische Kommission. Die Grundlage für die Berechnung der Daten bildet dabei das Europäische System Volkswirtschaftlicher Gesamtrechnungen (ESVG) 2010 (Eurostat, 2014). Demnach definiert sich der Staatssektor primär über Einheiten, die keine Marktinteressen verfolgen, sondern öffentliche Güter bereitstellen. Entscheidend für die Zuordnung ist dabei die Finanzierungsstruktur: Der Sektor umfasst jene Institutionen, die sich maßgeblich aus Steuern, Sozialabgaben oder staatlichen Zuweisungen finanzieren.

Das Statistische Amt der Europäischen Union (Eurostat) überwacht die Einhaltung der ESVG-Vorgaben und gibt bei Bedarf bilaterale Empfehlungen, sogenannte *Advices*. Ein solcher *Advice* im November 2022 führte in Deutschland zu einer grundlegenden statistischen Neubewertung des Verkehrssektors in den Volkswirtschaftlichen Gesamtrechnungen. Eurostat präzisierte darin, dass Regionalisierungsmittel im öffentlichen Personennahverkehr (ÖPNV) nicht als Markterlöse angerechnet werden dürfen. Da diese Mittel nun als staatliche Transfers eingestuft werden, erfüllen die meisten betroffenen Verkehrs- und Infrastrukturunternehmen nicht mehr die Kriterien für Marktproduzenten. Dies hatte zur Folge, dass nahezu alle Einheiten des ÖPNV und der Bahninfrastruktur in den Staatssektor umklassifiziert werden mussten.

Dieser Artikel beschreibt die Reklassifikation der öffentlichen Personennahverkehrsunternehmen und der Bahninfrastruktur in den Staatssektor aufgrund des *Advice* durch Eurostat.<sup>1</sup> Nach einer ausführlichen Einführung in das Thema in Kapitel 2 beschreibt das Kapitel 3 die Herausforderungen sowie das Vorgehen bei der Sektorneuzuordnung der relevanten Einheiten. Das vierte Kapitel präsentiert die

---

1 Die Arbeiten wurden von der EU finanziell unterstützt durch den Grant 2023-DE-NA-GNI.

Ergebnisse für die Aggregate Staatskonsum, Investitionen und Defizit. Kapitel 5 enthält ein Fazit.

## 2

### Überblick

#### 2.1 ÖPNV- und Bahninfrastruktureinheiten

Nach § 2 Regionalisierungsgesetz ist der ÖPNV definiert als „die allgemein zugängliche Beförderung von Personen mit Verkehrsmitteln im Linienverkehr, die überwiegend dazu bestimmt sind, die Verkehrsnachfrage im Stadt-, Vorort- oder Regionalverkehr zu befriedigen“. Meistens sind die Reisen im Nahverkehr kürzer als entweder 50 Kilometer oder eine Stunde Reisezeit.

Dieser Artikel betrachtet allerdings nur staatlich kontrollierte Nahverkehrsunternehmen; Einheiten des Fernverkehrs, wie die DB Fernverkehr AG, zählen nicht dazu. Ebenfalls einbezogen sind die staatlich kontrollierten Bahninfrastruktureinheiten wie die DB InfraGO AG.

Die Bundesländer sind zuständig für die Organisation, Planung und Kontrolle der vergebenen Verkehrsleistungen im ÖPNV, der Bund stellt finanzielle Mittel zur Verfügung.

In der Praxis ernennen die meisten Bundesländer Besteller-Organisationen (in der Regel Verkehrsverbünde, kommunale Zweckverbände), die die Organisation des ÖPNV übernehmen. Diese Organisationen definieren das erforderliche Nahverkehrsangebot, legen die gewünschten Verbindungen fest und formulieren, wie häufig bestimmte Strecken bedient werden. Ebenfalls legen sie Mindestvorgaben fest, etwa die Qualität der Fahrzeuge oder die Fahrgeldtarife.

Basierend auf diesen Nahverkehrsplänen schreiben die Besteller-Organisationen konkrete Strecken für die Durchführung der Fahrleistungen aus. Verkehrsunternehmen können sich auf die Ausschreibungen bewerben. Bei den Gewinnern der Ausschreibungen

kaufen die Besteller-Organisationen dann mit den Mitteln der Bundesländer beziehungsweise des Bundes Fahrleistungen.<sup>12</sup>

Der ÖPNV finanziert sich hauptsächlich durch Fahrgeldeinnahmen, Bestellerentgelte und sonstige öffentliche Zuschüsse. Innerhalb eines Verkehrsverbundes kann dieser selbst die Fahrkartenpreise bestimmen. Bei Fahrten über Verbundgrenzen hinweg entscheidet die Deutschlandtarifverbund-GmbH über die Fahrkartenpreise. Seit Mai 2023 gibt es zusätzlich das Deutschlandticket<sup>13</sup>, das mittlerweile mehr als die Hälfte der Fahrgeldeinnahmen im ÖPNV ausmacht.

Mit der Übertragung der Zuständigkeit für den Schienenpersonennahverkehr vom Bund auf die Bundesländer stehen den Bundesländern Mittel für diese Aufgabe zu. Diese sogenannten Regionalisierungsmittel werden seit 1996 auf der Grundlage des Regionalisierungsgesetzes auf die Bundesländer verteilt. Im Jahr 2025 bezifferte sich der relevante Haushalts-titel auf knapp 13 Milliarden Euro. Seit 2023 zahlt der Bund jährlich 1,5 Milliarden Euro an zusätzlichen Mitteln, um die Mindereinnahmen durch das Deutschlandticket auszugleichen.

Zusätzlich zum Schienenpersonennahverkehr wird aus den Regionalisierungsmitteln auch die Bahninfrastruktur finanziert. Dies geschieht indirekt über sogenannte Trassen- und Stationsentgelte, die die Eisenbahnverkehrsunternehmen für die Nutzung von Schienen und Bahnhöfen zahlen. Die Einnahmen aus Trassen- und Stationsentgelten können die Bahninfrastrukturunternehmen für Infrastrukturmaßnahmen wie Bahn-hofsrevitalisierungen, die Sanierung des Schienennetzes sowie den Strecken-ausbau nutzen.

Der größte Teil der Eisenbahninfrastruktur in Deutschland wird von der DB InfraGO AG<sup>14</sup> betrieben. Zur Finanzierung des Schienennetzes stellt der Bund über eine Leistungs- und Finanzierungsverein-

barung jährlich Mittel zur Verfügung.<sup>15</sup> Die Leistungs- und Finanzierungsvereinbarung wurde zwischen der Bundesrepublik Deutschland, den Eisenbahninfrastrukturunternehmen des Bundes und der Deutschen Bahn AG abgeschlossen. Ein weiterer wichtiger Pfeiler in der Finanzierung der Schienenwege sind die Trassenentgelte. Die Trassenentgelte werden als eine Art Schienenmaut von allen Personen- und Güterverkehrsunternehmen erhoben, die auf dem Streckennetz unterwegs sind. Deswegen werden auch die Regionalisierungsmittel indirekt durch die ÖPNV-Einheiten für die Nutzung der Schienenwege verwendet.<sup>16</sup>

## 2.2 Abgrenzung des Sektors Staat

---

Im Zuge des Verfahrens bei einem übermäßigen Defizit nach dem europäischen Stabilitäts- und Wachstumspakt liefert das Statistische Bundesamt regelmäßig Daten zum Finanzierungsdefizit des Staates an die Europäische Kommission. Die Abgrenzung des Staatssektors gegenüber den anderen volkswirtschaftlichen Sektoren ist dabei entscheidend für die Vergleichbarkeit der Daten zwischen den Mitgliedstaaten der EU. Den rechtlichen Rahmen zur Staatsabgrenzung bildet das Europäische System Volkswirtschaftlicher Gesamtrechnungen (ESVG) 2010 (Eurostat, 2014).

Nach den ESVG-Vorgaben zählen zum Staatsektor alle institutionellen Einheiten, die keine Marktproduzenten sind, deren Produktionswert dem Individual- und Kollektivkonsum dient und die sich überwiegend durch Zwangsabgaben finanzieren (Eurostat, 2014, hier: Ziffer 2.111). Einheiten, die sich überwiegend durch staatliche Transfers finanzieren, gehören ebenfalls zum Staatssektor (Eurostat, 2014, hier: Ziffer 20.10).

Zum Staat gehören neben den Kernhaushalten der vier Teilsektoren des Staates (Bund, Länder, Gemeinden und Sozialversicherungen) auch sogenannte Extrahaushalte. Zu diesen zählen alle öffentlichen Fonds, Einrichtungen und Unternehmen, die nach den

---

2 Ein kurzer Überblick findet sich unter [www.dbrégio.de](http://www.dbrégio.de)

3 Das Deutschlandticket erlaubt für einen fixen Preis eine Nutzung des kompletten ÖPNV innerhalb Deutschlands. Die Einnahmen aus dem Verkauf des Deutschlandtickets werden anhand der Postleitzahlen der Kunden dann auf die ÖPNV-Einheiten verteilt.

4 Eine hundertprozentige Tochter der DB AG, welche im Jahr 2023 durch die Fusion der DB Netz AG mit der DB Station&Service AG entstanden ist.

---

5 Ein Überblick über alle Leistungs- und Finanzierungsvereinbarungen findet sich unter [www.eba.bund.de](http://www.eba.bund.de)

6 Ein guter Überblick über die Finanzierung des Schienenpersonennahverkehrs findet sich unter [www.schienen-nahverkehr.de](http://www.schienen-nahverkehr.de)

ESVG-Kriterien dem Sektor Staat zuzurechnen sind, zum Beispiel öffentliche Hochschulen oder staatlich kontrollierte Unternehmen, die ihre Leistungen nicht zu Marktpreisen anbieten (Eurostat, 2014, Ziffer 20.10). Konkrete Beispiele aus dem ÖPNV sind die DB Regio AG, die Hessische Landesbahn GmbH oder auch die ESWE Verkehrsgesellschaft GmbH.<sup>17</sup>

Für die Entscheidung über die Sektorzuordnung zum Sektor Staat gemäß ESVG 2010 sowie ergänzend nach dem Eurostat-Handbuch „Manual on Government Debt and Deficit“ (Version 2022) werden die folgenden drei zentralen Fragen berücksichtigt:

1. Ist die Einheit eine eigenständige „institutionelle Einheit“?
2. Ist die Einheit vom Staat kontrolliert?
3. Handelt es sich um einen Marktproduzenten oder um einen Nichtmarktproduzenten?

Bei der ersten Frage geht es um den Status und die Entscheidungsautonomie der Einheit. Einzubeziehende Aspekte sind unter anderem, ob die Einheit selbstständig Verträge unterschreiben oder Schulden aufnehmen kann. Die zweite Frage zielt darauf ab, welche Einheit die Kontrolle ausübt beziehungsweise die allgemeine (Geschäfts-)Politik bestimmt. Hier wird geprüft, ob der Staat die Mehrheit der Kapital- oder Stimmrechte hält oder zum Beispiel die strategische Ausrichtung bestimmt oder die Leitung einsetzt. Bei der dritten Frage geht es um die Unterscheidung zwischen Markt- und Nichtmarktproduzenten, was über den sogenannten Markt-/Nichtmarkttest geprüft wird: Marktproduzenten sind so definiert, dass sie über einen aussagefähigen Mehrjahreszeitraum<sup>18</sup> mindestens 50 % ihrer Kosten aus den Umsatzerlösen decken. Institutionelle Einheiten, welche staatlich kontrolliert sind und bei denen es sich um Nichtmarktproduzenten handelt, werden dem Sektor Staat zugeordnet. Daneben gibt es noch weitere Zuordnungsvorschriften für spezielle Einheiten wie staatliche Hilfsbetriebe oder staatliche Holdinggesellschaften.<sup>19</sup>

Die Abgrenzung des Sektors Staat wird in Deutschland vom Statistischen Bundesamt jährlich überprüft, da öffentliche Einheiten neu gegründet, ausgegliedert, abgewickelt, privatisiert oder fusioniert werden können. Eine regelmäßige Überprüfung gewährleistet, dass die Abgrenzung des Sektors Staat über die Zeit korrekt erfolgt und in den europäischen Staaten vergleichbar bleibt.

### 2.3 Eurostat-Advice zur Sektorklassifizierung des ÖPNV und der Bahninfrastruktur in Deutschland

---

Eurostat prüft regelmäßig die Einhaltung der ESVG-Vorgaben bei der Berechnung des staatlichen Defizits und des Schuldenstands.

Dabei kann Eurostat den EU-Mitgliedstaaten bei Bedarf bilaterale Empfehlungen (Advices) bezüglich der methodischen Bewertung konkreter Sachverhalte aussprechen.<sup>10</sup> Die Empfehlung erfolgt in Form eines Schreibens an die nationalen statistischen Stellen und wird durch Eurostat veröffentlicht. Im November 2022 hat Eurostat Deutschland einen solchen Advice bezüglich der Klassifizierung bestimmter Transaktionen im Zusammenhang mit der Finanzierung weiter Teile des ÖPNV sowie der Bahninfrastruktur ausgesprochen.

So stellte Eurostat klar, dass staatliche (Pauschal-)Zahlungen zur Deckung von Kosten im Bereich des ÖPNV nicht als Umsatzerlöse im Markt-/Nichtmarkttest zu berücksichtigen sind. Dies betrifft vor allem Bestellerentgelte im Schienenpersonennahverkehr beziehungsweise Trassen-/Stationsentgelte an die Betreiber der Bahninfrastruktur, aber auch sonstige Zuschüsse und Ausgleichszahlungen an kommunale Bus- und Straßenbahnunternehmen des ÖPNV.

Diese geänderte Einordnung der Einnahmen, unter anderem aus den Regionalisierungsmitteln, hatte entscheidende Auswirkungen auf die Ergebnisse des Markt-/Nichtmarkttests für nahezu alle ÖPNV- und

---

7 Die Liste aller Extrahaushalte findet sich auf folgender Seite: [www.destatis.de](http://www.destatis.de)

8 In Deutschland werden in der Regel drei aufeinanderfolgende Jahre betrachtet.

9 Eine detaillierte Zusammenfassung findet sich in Schmidt und andere (2017).

---

10 Ein „Advice“ ist ein Werkzeug von Eurostat um sicherzustellen, dass Defizit und Schuldenstand in allen EU-Ländern korrekt berechnet werden. Konkret gibt Eurostat auf Anfrage eines Mitgliedstaats dabei eine Empfehlung, wie eine bestimmte Transaktion zu buchen ist.

Bahninfrastruktureinheiten, da sie eine Klassifikation dieser Einheiten in den Staatssektor zur Folge hatte.

### 3

## Herausforderungen und Vorgehen

Die Empfehlungen des Eurostat-Advice sind sowohl für die laufende Rechnung, als auch rückwirkend für die vergangenen Jahre zu berücksichtigen. Ursprünglich war daher geplant, die Sektorneuzuordnung der ÖPNV- und Bahninfrastruktureinheiten im Zuge der nächsten Generalrevision der Volkswirtschaftlichen Gesamtrechnungen (VGR) durchzuführen. Dies hätte eine Anpassung der gesamten Zeitreihe ermöglicht, wodurch zu jedem Zeitpunkt Brüche in den Zeitreihen der Volkswirtschaftlichen Gesamtrechnungen vermieden worden wären. Die letzte europaweit harmonisierte Generalrevision der Volkswirtschaftlichen Gesamtrechnung fand im Jahr 2024 statt.

Die Einführung des Deutschlandtickets im Jahr 2023 führte jedoch parallel zu Eurostats Empfehlungen vom November 2022 zu einer Änderung der ökonomischen Realität der ÖPNV-Einheiten: Das Deutschlandticket veränderte die Finanzierungsstruktur der ÖPNV-Einheiten und damit die Ergebnisse der Bewertung ihrer Marktaktivität im Markt-/Nichtmarkttest

erheblich. Dadurch wurde es notwendig, die Einheiten bereits ab dem Frühjahr 2023 im Sektor Staat zu klassifizieren. Die rückwirkende Reklassifikation vor dem zweiten Quartal 2023 zurück bis 1991 aufgrund des Eurostat-Advice erfolgte anschließend wie geplant mit der VGR-Generalrevision 2024.

### ➤ Grafik 1

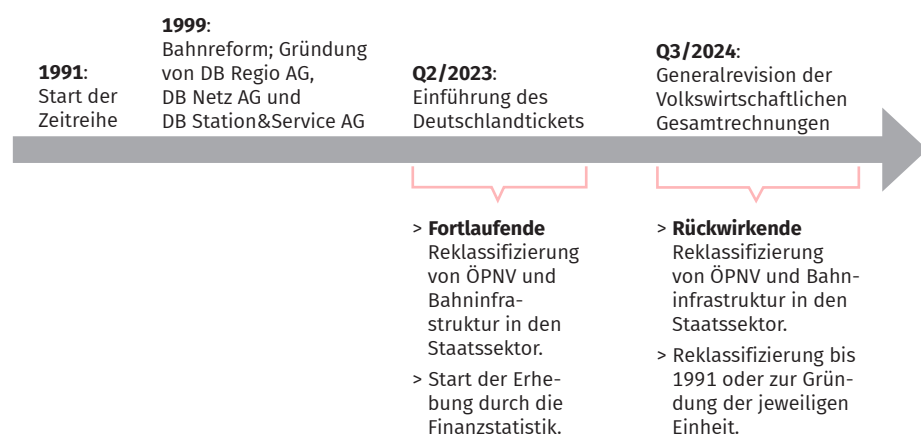
Grundsätzlich wurden alle Einheiten bis 1991 rückwirkend reklassifiziert. Allerdings existierten viele Einheiten vor 1999 nicht, da sie erst im Zuge der Bahnreform aus der DB AG ausgegründet wurden. Diese Einheiten wurden folglich nur bis zu ihrem Gründungsjahr zurück reklassifiziert.

## 3.1 Datenquellen für die Identifizierung der relevanten Einheiten

Die Datengrundlage für die Reklassifikation bildete das Berichtskreismanagement, eine Datenbank der Finanzstatistiken, die gemäß dem Finanz- und Personalstatistikgesetz alle berichtspflichtigen Einheiten enthält. Diese Datenbank umfasst neben den Kernhaushalten auch alle anderen Einheiten, an denen der Staat mit mindestens 50 % des Nennkapitals oder der Stimmrechte beteiligt ist. Hierzu zählen vor allem Zweckverbände sowie Fonds, Einrichtungen und Unternehmen, die in öffentlicher oder privater Rechts- oder Organisationsform existieren.

**Grafik 1**

**Zeitstrahl zur Reklassifikation der ÖPNV- und Bahninfrastruktureinheiten**



ÖPNV: öffentlicher Personennahverkehr

## 3.2 Identifizierte Einheiten

ÖPNV-Einheiten sind in den meisten Fällen institutionelle Einheiten, die über Entscheidungsfreiheit in der Ausübung ihrer Hauptfunktionen verfügen. Da der Staat in der Regel mit mehr als 50 % an diesen Einheiten beteiligt ist, unterliegen sie der Kontrolle des Staates.

Damit sind grundsätzlich zwei von drei notwendigen Kriterien für eine Zuordnung zum Sektor Staat erfüllt. Vor Einführung des Deutschlandtickets und vor dem Advice von Eurostat handelte es sich in der Regel um Marktproduzenten.

Nach den Vorgaben im Advice waren allerdings die vorhandenen Angaben zu den Umsätzen um staatliche Zahlungen zu bereinigen, bevor der Markt-/Nichtmarkttest erneut durchgeführt werden konnte. So wurden von den Umsatzerlösen für ÖPNV-Einheiten die Einnahmen aus Verkehrsverträgen (Bestellerentgelte und sonstige staatliche Zahlungen) abgezogen. Gleichzeitig wurden bei Einheiten des Schienenpersonennahverkehrs auf der Kosten- und der Einnahmenseite die Ausgaben für Trassen- und Stationsentgelte abgezogen. Für die Bahninfrastruktur wurden die Umsatzerlöse um die eingenommenen Trassen- und Stationsentgelte bereinigt. Diese Bereinigung führte dazu, dass sehr viele Einheiten nach Maßgabe des Markt-/Nichtmarkttests nicht mehr als Marktproduzenten angesehen werden konnten.

Alle identifizierten ÖPNV- und Bahninfrastruktureinheiten, die gemäß dem Markt-/Nichtmarkttest keine Marktproduzenten mehr waren, wurden in den Staatssektor überführt.

Zum Stand des Jahres 2023 wurden letztlich 444 Einheiten identifiziert, die nach der methodischen Neubetrachtung der Umsätze und gegebenenfalls Kostenpositionen im Rahmen des Markt-/Nichtmarkttests nun als Nichtmarktproduzenten zu betrachten und in den Staatssektor umzusetzen waren. Die meisten Einheiten sind dabei auf der Gemeindeebene im ÖPNV verortet. Zusätzlich wurden auch größere ÖPNV-Einheiten auf der Landesebene identifiziert und auf der Bundesebene die DB Regio AG und ihre Tochterunternehmen. Für die Bahninfrastruktur gibt es praktisch nur eine

große Einheit: die DB InfraGO AG. Vor dem Jahr 2023 bestand diese Einheit aus den beiden Einheiten DB Netz AG und DB Station&Service AG. Die anderen zehn Einheiten der Bahninfrastruktur fallen im Vergleich zur großen Bundeseinheit kaum ins Gewicht.

➔ **Tabelle 1**

**Tabelle 1**

**Identifizierte ÖPNV- und Bahninfrastruktureinheiten nach staatlichen Teilsektoren**

|                   | Bund | Länder | Gemeinden | Insgesamt |
|-------------------|------|--------|-----------|-----------|
| ÖPNV              | 38   | 25     | 369       | 432       |
| Bahninfrastruktur | 2    | 2      | 8         | 12        |
| Insgesamt         | 40   | 27     | 377       | 444       |

ÖPNV: öffentlicher Personennahverkehr

## 3.3 Datenquellen für die Rückrechnung

Mit der Umsetzung der Einheiten in den Staatssektor begann das Einbeziehen der Einheiten in die vierteljährlichen Erhebungen der Finanzstatistiken. Diese stellen die wichtigste Datengrundlage für den Staatsbereich in den Volkswirtschaftlichen Gesamtrechnungen dar. Ab dem zweiten Quartal des Jahres 2023 sind daher vierteljährliche Erhebungsdaten zu den reklassifizierten Einheiten verfügbar.

Vor der Reklassifikation der Einheiten standen nur jährliche Angaben aus der Statistik der Jahresabschlüsse öffentlicher Fonds, Einrichtungen und Unternehmen für den Zeitraum von 1999 bis 2021 zur Verfügung. Folglich waren für die Jahre 2022 und 2023 Schätzungen vorzunehmen, die später durch die endgültigen Jahresabschlüsse ersetzt werden. Für das Berichtsjahr 2022 ist das schon erfolgt. Für das Berichtsjahr 2023 werden die Jahresabschlüsse im Frühjahr 2026 eingearbeitet. Für den Zeitraum vor 1999 wurde unter anderem auf die Daten des Verbands Deutscher Verkehrsunternehmen zurückgegriffen. Diese Daten und Rückrechnungen ermöglichen eine Imputation bis ins Jahr 1991 zurück.



## 4

## Ergebnisse

## 4.1 Staatskonsum der ÖPNV- und Bahninfrastruktureinheiten

Nach der Reklassifikation fließen die Ausgaben der ÖPNV- und Bahninfrastruktureinheiten auf der Verwendungsseite der Berechnung des Bruttoinlandsprodukts in den Staatskonsum ein, zum Beispiel durch gezahlte Arbeitnehmerentgelte und eingekaufte Vorleistungen.

Im Vergleich der Teilsektoren trägt der Bund für die Bahninfrastruktur, bedingt durch die DB InfraGO AG, mit knapp 99 % den größten Teil zu diesem Teil des Staatskonsums bei. Die Einheit ist verantwortlich für die Eisenbahninfrastruktur in Deutschland und als Tochter der DB AG dem Teilsektor Bund zugeordnet. [↘ Grafik 2](#)

Für den ÖPNV ist die Verteilung ausgeglichener. Der Bund stellt 25,5 % von diesem Teil des Staatskonsums, die Bundesländer 19,2 % und die Gemeinden mit 55,4 % den größten Anteil. Zwar gehört die größte Einheit, die DB Regio AG, dem Bund, gleichzeitig gibt es auf Gemeindeebene aber viele kleine Verkehrsunternehmen.

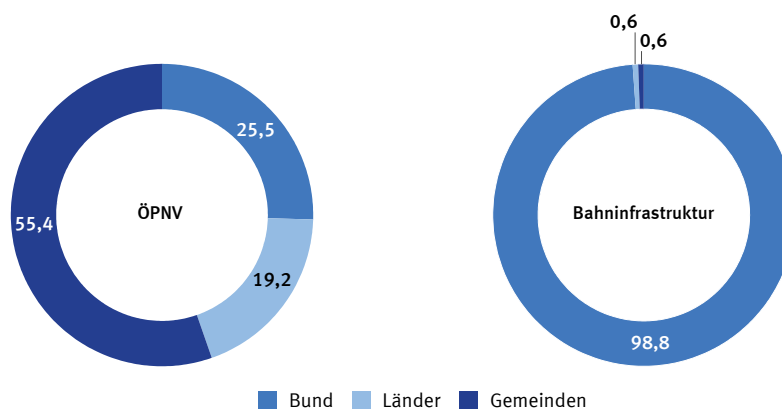
Für die Bahninfrastruktur lag der Staatskonsum vor 1999 bei null, da die Bahninfrastruktur noch Bestandteil der DB AG war, die weiterhin als öffentliche Hauptverwaltung außerhalb des Staates klassifiziert bleibt. Mit der Ausgründung der Bahninfrastruktureinheiten im Jahr 1999 stieg er folglich sprunghaft an. Von 2000 bis 2009 blieb der Staatskonsum im Bereich der Bahninfrastruktur weitgehend konstant. Seit 2010 steigt der Staatskonsum kontinuierlich an, mit einem deutlich beschleunigten Trend. [↘ Grafik 3](#)

Die Trendwende führte die erste Leistungs- und Finanzierungsvereinbarung herbei, die im Jahr 2009 eingeführt wurde. Dadurch stieg die Planungssicherheit für die Unternehmen, sodass die Bahninfrastruktureinheiten expandierten und mehr Mitarbeitende einstellten. Mit jeder neuen Vereinbarung (2014 bis 2019 und 2020 bis 2029) wurden die jährlichen Zuschüsse des Bundes deutlich gesteigert. Zusätzlich erhielt die DB InfraGO AG in den letzten Jahren weitere Investitionszuschüsse für die Sanierung des Schienennetzes in Form von Kapitalzuführungen durch die DB AG.

Für den ÖPNV ist der Staatskonsum auch schon vor 1999 größer null, da es bereits vor 1999 (wenige und kleinere) ÖPNV-Einheiten gab; er stieg jedoch mit der Ausgründung der DB Regio AG deutlich an. Zwischen 2000 und 2010 sank er wieder leicht, verursacht durch Sparmaßnahmen des Bundes, nied-

Grafik 2

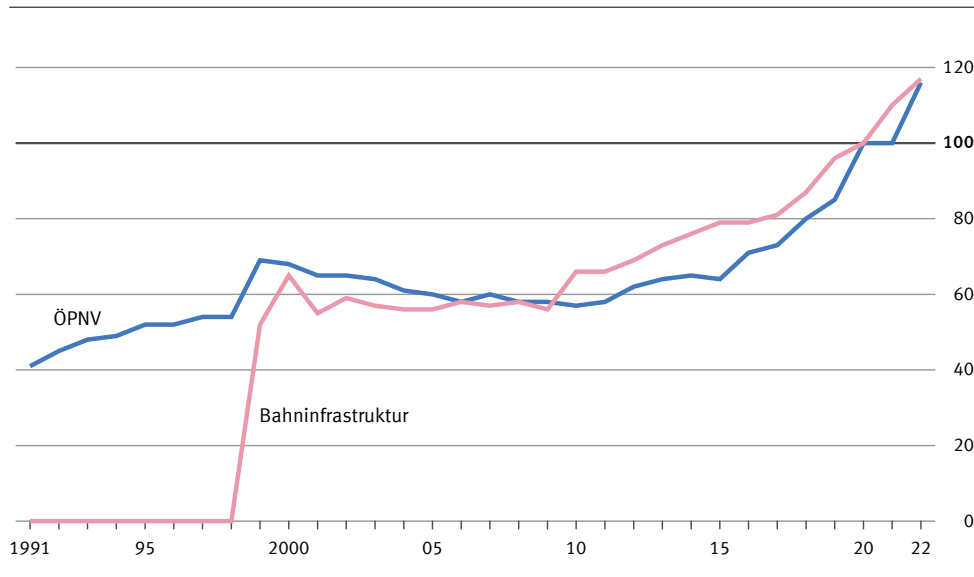
Staatskonsum der ÖPNV- und Bahninfrastruktureinheiten nach staatlichen Teilsektoren in %



ÖPNV: öffentlicher Personennahverkehr  
Durchschnitt von 1991 bis 2022; Stand: Oktober 2025.

**Grafik 3**

**Zeitliche Entwicklung des Staatskonsums der ÖPNV- und Bahninfrastruktureinheiten**  
nominal, 2020 = 100



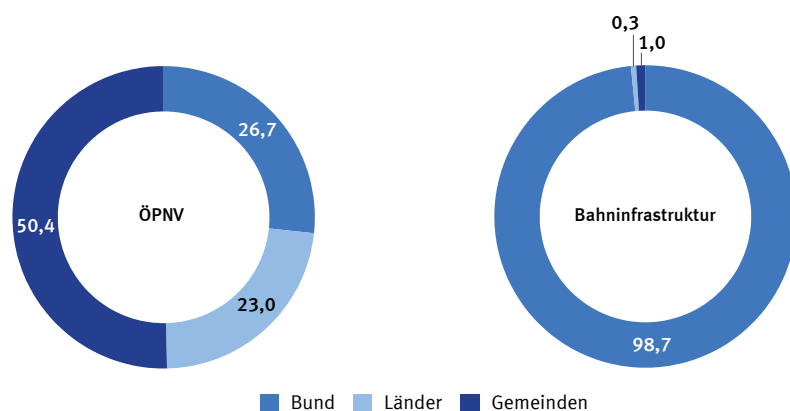
ÖPNV: öffentlicher Personennahverkehr  
Stand: Oktober 2025.

rigere Ticketverkäufe und die Liberalisierung des Schienenverkehrs. Dies führte dazu, dass staatliche Nahverkehrsunternehmen Marktanteile verloren. Ab 2011 stieg der Staatskonsum im Bereich des ÖPNV wieder, begründet unter anderem durch die erfolgreiche Kostenreduktion bei der DB Regio AG, die zu

mehr gewonnenen Ausschreibungen führte. Ab 2016 wurden die Regionalisierungsmittel dynamisiert, wodurch der Staatskonsum in diesem Teilbereich wieder spürbar anstieg.

**Grafik 4**

**Staatliche Bruttoanlageinvestitionen der ÖPNV- und Bahninfrastruktureinheiten nach Teilsektoren**  
in %



ÖPNV: öffentlicher Personennahverkehr  
Durchschnitt von 1991 bis 2022; Stand: Oktober 2025.

## 4.2 Staatliche Investitionen der ÖPNV- und Bahninfrastruktureinheiten

Die Investitionen der ÖPNV-Einheiten bestehen überwiegend aus Ausrüstungsgegenständen wie Zügen und Bussen. Bei den Infrastruktureinheiten sind es hauptsächlich Investitionen in Bauten, da der Bau und die Reparatur von Schienen und Bahnhöfen in diese Kategorie fallen. Aufgrund der Art der Investitionen wenden die Infrastruktureinheiten deutlich mehr Geld auf als die ÖPNV-Einheiten. ➤ **Grafik 4**

Die Verteilung der Investitionsausgaben ähnelt jener bei den Staatskonsumausgaben. Für die Bahninfrastruktur entfallen auch hier knapp 99 % aller Investitionen auf den Bund. Für den ÖPNV weisen auch bei den Investitionen die Gemeinden mit gut 50 % den höchsten Anteil auf. Auf dem zweiten Platz folgt der Bund mit knapp 27 %. Der Anteil der Bundesländer an den Investitionen ist mit 23 % etwas geringer.

➤ **Grafik 5** zeigt die Entwicklung der nominalen Bruttoanlageinvestitionen für die ÖPNV- und Bahninfrastruktureinheiten im Zeitverlauf von 1991 bis 2022. Im Gegensatz zu den relativ stabilen Staatskon-

sumausgaben verlaufen die Investitionen unregelmäßiger. Während Konsumausgaben wie Löhne und Gehälter relativ stabil sind und sich nur langsam verändern, sind Investitionsausgaben deutlich volatiler. Sie steigen zum Beispiel sprunghaft an, wenn in ein neues Großprojekt investiert wird, fallen in den Folgejahren aber oft wieder ab.

Für den ÖPNV ist die Entwicklung der Bruttoanlageinvestitionen von 1991 bis 2011 relativ konstant, mit kleinen Spitzen in den Jahren 1999, 2002 und 2007, verursacht durch Modernisierungen des Fahrzeugparks der DB Regio AG (1999 und 2002) und auf der Gemeindeebene (2007). Ab 2012 steigen die Bruttoanlageinvestitionen in diesem Teilbereich deutlich an, mit zwei großen Spitzen in den Jahren 2014 (Kauf elektrischer Triebzüge durch die DB Regio AG) und 2018 (Fahrzeugbeschaffung der Gemeinden). Die Entwicklung zwischen den beiden Perioden spiegelt den politischen Kurswechsel wider: weg von Sparmaßnahmen hin zu gezielten Investitionen in den ÖPNV.

Für die Bahninfrastruktur gab es aufgrund fehlender staatlicher Einheiten analog zum Staatskonsum vor 1999 keine Bruttoanlageinvestitionen. Erst mit der

**Grafik 5**

**Zeitliche Entwicklung der staatlichen Bruttoanlageinvestitionen der ÖPNV- und Bahninfrastruktureinheiten**

nominal, 2020 = 100



ÖPNV: öffentlicher Personennahverkehr  
Stand: Oktober 2025.

Ausgründung der Bahninfrastruktureinheiten im Jahr 1999 stiegen die Investitionen deutlich an, gefolgt von einem weiteren Anstieg 2002. Bis einschließlich 2009 stagnierten die Investitionen in diesem Teilbereich dann, bevor ab 2009 ein steigender Trend einsetzte, zunächst langsam und dann beschleunigt. Dieser Verlauf korreliert mit den Leistungs- und Finanzierungsvereinbarungen zwischen Bund und Schieneninfrastruktureinheiten.

Insgesamt weist die Entwicklung große Ähnlichkeiten mit der Entwicklung der Staatskonsumausgaben der Bahninfrastruktur aus Grafik 3 auf.

Während der ÖPNV nur 2 % der gesamten staatlichen Bruttoanlageinvestitionen ausmacht, liegt der Anteil der Bahninfrastruktur bei 10 %. Grund dafür ist, dass die Schieneninfrastruktur vergleichsweise hohe Investitionen erfordert. Dementsprechend groß ist der Anteil der Bahninfrastruktur an den gesamten Bruttoanlageinvestitionen des Staates und somit auch der Reklassifikationseffekt, obwohl hier nur wenige Einheiten betroffen sind.

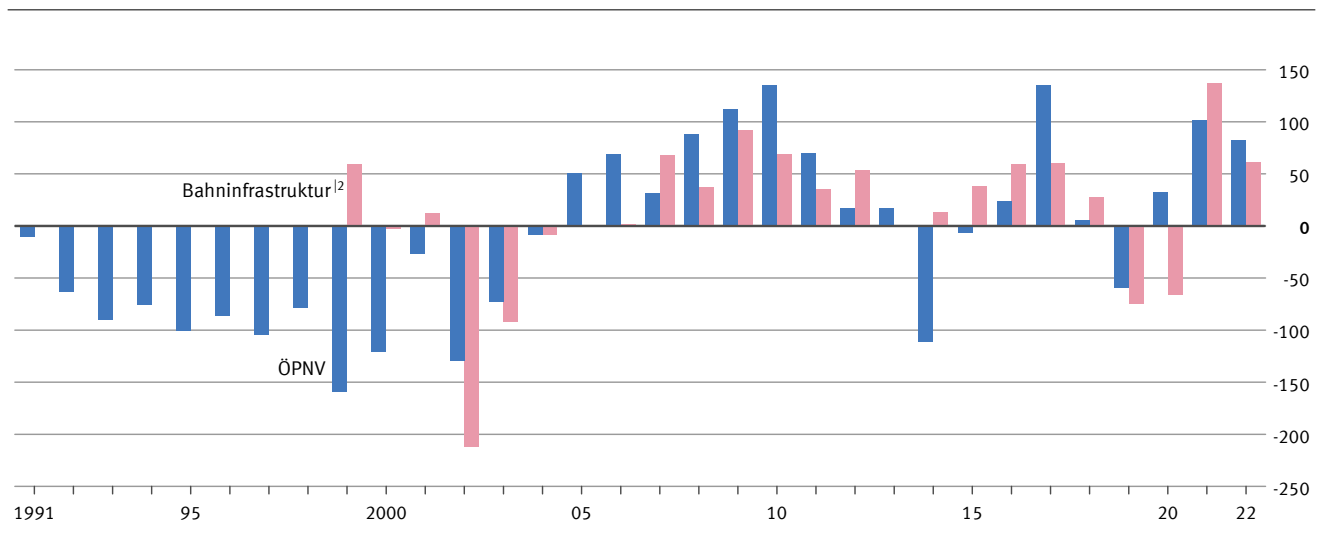
### 4.3 Staatliches Defizit der ÖPNV- und Bahninfrastruktureinheiten

Neben Staatskonsum und staatlichen Investitionen ist das Defizit eine wichtige Kenngröße mit Blick auf den Stabilitäts- und Wachstumspakt, allerdings mit starken jährlichen Schwankungen und sowohl positiven als auch negativen Werten. Aus Gründen der Übersichtlichkeit wird auf eine Darstellung der Defizite nach Teilsektoren verzichtet. Stattdessen zeigt [Grafik 6](#) den zeitlichen Verlauf der gesamtstaatlichen Defizite und Überschüsse für die ÖPNV- und Bahninfrastruktureinheiten.

Die ÖPNV-Einheiten wiesen von 1991 bis einschließlich 2004 durchgehend deutliche Defizite auf. Ab 2005 stabilisieren sich die Bilanzen und weisen in den meisten nachfolgenden Jahren Überschüsse aus, mit großen negativen Ausnahmen in den Jahren 2014 und 2019. Hohe Überschüsse in diesem Teilbereich gab es in den Jahren 2010, 2017 und 2021. Die Defizite in den frühen 2000er-Jahren spiegeln die damalige Sparpolitik wider. Die besonders hohen Defizite der ÖPNV-Einheiten in den Jahren 2002 und 2014 entstanden durch hohe Investitionen in die Modernisierung des Fuhrparks der DB Regio AG. Die

**Grafik 6**

Entwicklung des staatlichen Defizits und Überschusses der ÖPNV- und Bahninfrastruktureinheiten  
Normalisiert auf einen Mittelwert von Null<sup>1</sup>



ÖPNV: öffentlicher Personennahverkehr

<sup>1</sup> Die Normalisierung erfolgt durch die Subtraktion des Mittelwertes. Das Ergebnis wird danach durch die Standardabweichung geteilt. Stand: Oktober 2025.

<sup>2</sup> Für die Bahninfrastruktur können vor 1999 aufgrund fehlender staatlicher Einheiten keine Defizite und Überschüsse nachgewiesen werden.

Reduktion der deutlichen Defizite ab dem Jahr 2005 hat mehrere Gründe: Zum einen stiegen die staatlichen Zuschüsse für den ÖPNV in Form der Regionalisierungsmittel kontinuierlich. Zum anderen gelang der DB Regio AG nach ihrer Ausgründung eine erfolgreiche Sanierung.

Die Bahninfrastruktureinheiten wiesen nur in vier Jahren größere Defizite auf: in den Jahren 2002, 2003, 2019 und 2020. Besonders hohe Überschüsse gab es in diesem Teilbereich in den Jahren 2009 und 2021. Die Jahre 2020 und 2021 sind hierbei gesondert zu betrachten, da die Bahninfrastruktureinheiten im Jahr 2020 einen deutlichen Einbruch der Einnahmen aus Trassen- und Stationsentgelten verzeichneten. Dieser Einbruch war auf die Maßnahmen im Zuge der COVID-19-Pandemie zurückzuführen. Der Überschuss im Jahr 2021 resultierte dann aus Unterstützungszahlungen des Bundes zum Ausgleich der pandemiebedingten Einschränkungen. Grundsätzlich ist anzunehmen, dass es selten zu einem Defizit in diesem Teilbereich kommen sollte, da die Investitionen der Bahninfrastruktureinheiten durch den Staat finanziert werden. Sollte trotzdem ein Defizit eintreten, so liegt dies in der Regel daran, dass die Investitionszuschüsse teilweise zeitverzögert an die Bahninfrastruktureinheiten ausgezahlt werden.

Aggregate in den Volkswirtschaftlichen Gesamtrechnungen umgesetzt wird. Beispielsweise ist der Zugang einer Einheit im Sektor Staat konsistent durch einen Abgang in einem anderen Sektor, hier dem Sektor der Nichtfinanziellen Unternehmensgesellschaften, zu berücksichtigen. Gleichermaßen führt die Berücksichtigung einer Einheit bei der Berechnung der Investitionen des Staates zu einer Nichtberücksichtigung bei der Berechnung der Investitionen der Nichtfinanziellen Unternehmensgesellschaften.

Eine weitere Schwierigkeit bei der Reklassifikation war die Einführung des Deutschlandtickets – eine politische Entscheidung, die Einfluss auf die ökonomische Realität hatte. Die Volkswirtschaftlichen Gesamtrechnungen müssen solche Entwicklungen stets im Blick haben, um eine korrekte Darstellung der Sektorzuordnung sicherstellen zu können. So ist beispielsweise die seit Beginn des Jahres 2025 schrittweise umgesetzte Krankenhausreform ebenfalls eine politische Entscheidung, die Einfluss auf die ökonomische Realität haben könnte. Sie könnte somit Grund für eine erneute Prüfung der Sektorzuordnung der Krankenhäuser sein. [u](#)

## 5

---

### Fazit

---

Dieser Artikel beleuchtet die Auswirkungen eines Eurostat-Advice vom November 2022 für die Sektorabgrenzung des Staates nach dem ESVG 2010, nämlich die Reklassifikation der ÖPNV- und Bahninfrastruktureinheiten in den Staatssektor.

Insgesamt war die Reklassifikation für die deutschen Volkswirtschaftlichen Gesamtrechnungen – und somit auch für die Unternehmensstatistiken und die Finanzstatistiken als Datengrundlage der Volkswirtschaftlichen Gesamtrechnungen – eine große Herausforderung: Eine hohe Anzahl an Einheiten musste zeitgleich und rückwirkend für die gesamte Zeitreihe reklassifiziert werden. Es war zu gewährleisten, dass die Reklassifikation konsistent hinsichtlich der Darstellung der Sektoren sowie der sektorbezogenen

## LITERATURVERZEICHNIS

---

Eurostat. *Europäisches System Volkswirtschaftlicher Gesamtrechnungen – ESVG 2010*. Luxemburg 2014. [Zugriff am 14. Januar 2026]. Verfügbar unter: [ec.europa.eu](https://ec.europa.eu)

Eurostat. *Guidance on the treatment of ÖPNV subsidies*. Advice to Member States. 2022.

Eurostat. *Manual on Government Deficit and Debt – Implementation of ESA 2010. 2022 edition*. 2023. [Zugriff am 14. Januar 2026]. Verfügbar unter: [ec.europa.eu](https://ec.europa.eu)

Schmidt, Pascal/Heil, Nora/Schmidt, Daniel/Kaiser, Julia. [Die Abgrenzung des Staatssektors in den Volkswirtschaftlichen Gesamtrechnungen](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2017, Seite 35 ff.

## RECHTSGRUNDLAGEN

---

Gesetz zur Regionalisierung des öffentlichen Personennahverkehrs (Regionalisierungsgesetz – RegG) vom 27. Dezember 1993 (BGBl. I Seite 2378, 2395), das zuletzt durch Artikel 1 des Gesetzes vom 25. November 2025 (BGBl. I Nr. 287) geändert worden ist.





**Dr. Martina Rengers**

ist Volkswirtin und im Referat „Arbeitsmarkt“ des Statistischen Bundesamtes tätig. Ihre Schwerpunkte liegen in der Analyse ungenutzten Arbeitskräftepotenzials bei Erwerbstätigen, Erwerbslosen und in der Stillen Reserve sowie der konzeptionellen und methodischen Weiterentwicklung der international vergleichbaren Arbeitsmarktstatistik. Zusätzlich befasst sie sich mit dem Thema „Big Data und Arbeitsmarktstatistik“. Im Fokus sind dabei insbesondere Daten zu Online-Stellenanzeigen, ihre Qualitätsbewertung und Eignung für die Arbeitsmarktberichterstattung.



**Thomas Körner**

hat in Mainz und Caen Soziologie studiert und leitet im Statistischen Bundesamt das Referat „Bevölkerungsstatistische Auswertungen und Analysen aus dem Mikrozensus“. Seine Arbeitsschwerpunkte sind die Themen Haushalte und Lebensformen, Migration und Integration, Gleichstellung sowie ältere Menschen.

## ARBEITSZEITEN UND ARBEITSZEITWÜNSCHE VON MÄNNERN (UND FRAUEN) – ANHALTSPUNKTE ZU POTENZIALEN FÜR DIE GLEICHSTELLUNG

Martina Rengers, Thomas Körner

➤ **Schlüsselwörter:** Gleichberechtigung – Geschlechterrollen – Arbeitsmarkt – bezahlte Arbeit – Mikrozensus

### ZUSAMMENFASSUNG

Auf dem Arbeitsmarkt bestehen Unterschiede zwischen Männern und Frauen, die in zahlreichen Aspekten zulasten von Frauen gehen: Indikatoren wie der Gender Pay Gap, der Gender Earnings Gap, der Gender Employment Gap oder der Gender Pension Gap weisen etwa darauf hin, dass Verdienste zuungunsten von Frauen ungleich verteilt sind. Dass Ungleichheiten auf dem Arbeitsmarkt auch Männer bei der Verwirklichung ihrer Lebensentwürfe behindern können, wird jedoch selten thematisiert. Aus diesem Grund untersucht der Beitrag, wie sich die Arbeitszeiten von Männern und Frauen unterscheiden und welche Präferenzen und Wünsche Männer und Väter in Befragungen hinsichtlich ihrer Arbeitszeit angeben.

➤ **Keywords:** gender equality – gender roles – labour market – paid work – microcensus

### ABSTRACT

*The labour market is characterised by gender disparities that, in numerous respects, are disadvantageous to women. Indicators such as the gender pay gap, the gender earnings gap, the gender employment gap or the gender pension gap suggest that earnings are distributed unequally to the disadvantage of women. However, it is seldom discussed that labour market inequalities can also limit men in implementing their own life plans. This article therefore analyses the differences between the working hours of men and women and studies the working-time preferences reported in surveys by men, and more specifically by fathers.*

## 1

### Einleitung<sup>1</sup>

Dass Frauen auf dem Arbeitsmarkt in vielfacher Hinsicht benachteiligt sind, ist seit langem bekannt und wird in zahlreichen statistischen Berichtssystemen national und international thematisiert. Genannt seien etwa der [Gender Equality Index \(Gleichstellungsindex\)](#) des European Institute for Gender Equality (EIGE), der in Zusammenarbeit mit dem Statistischen Bundesamt erstellte [Gleichstellungsatlas](#) des Bundesministeriums für Bildung, Familie, Senioren, Frauen und Jugend (BMBFSFJ) sowie nicht zuletzt die [Themenseite Gleichstellungsindikatoren](#) des Statistischen Bundesamtes. Diese Statistiken geben einen guten Überblick über den Stand der Gleichstellung in verschiedenen Themenbereichen.

Da jeweils die Ergebnisse für Frauen und Männer gegenübergestellt werden, liegt der Fokus dabei meist auf den zahlreichen Benachteiligungen, denen Frauen ausgesetzt sind. Frauen sind noch immer in Entscheidungspositionen in Politik, Wirtschaft und Verwaltung unterrepräsentiert, leisten mehr unbezahlte Arbeit in der Familie und sind durch ein geringeres Volumen an bezahlten Arbeitsstunden sozial weniger abgesichert und auch in ihrer wirtschaftlichen Unabhängigkeit eingeschränkt.

Diese Berichterstattung ist sinnvoll und notwendig, um bestehende Ungerechtigkeiten für Frauen transparent zu machen und zu reduzieren. Seltener wird dabei die Frage gestellt, inwieweit die Ungleichheiten zwischen den Geschlechtern in Teilbereichen auch die Verwirklichungschancen von Männern beeinträchtigen können. Geschlechterrollen, tradierte Normen und Werte sowie die über lange Zeit gelebte Praxis etwa in Familien, Bildungseinrichtungen und Betrieben führen dazu, dass Männer männlich geprägte Berufe und Bildungsgänge ergreifen und seltener unbezahlte Sorgearbeit in der Familie leisten.

Zwischen Männern und Frauen bestehen weiter ausgeprägte Unterschiede in der Berufswahl (Bächmann

und andere, 2024), die sich auch bei den heutigen Absolventinnen und Absolventen von Studium und Berufsausbildung fortsetzen (Statistisches Bundesamt, 2025a; BMBFSFJ, 2025). Diese berufliche Segregation geht nicht nur mit unterschiedlichen Verdienstmöglichkeiten einher (siehe zum Beispiel Mischler, 2021; Kleinert und andere, 2023), sondern führt auch dazu, dass Männer beispielsweise um ein Vielfaches häufiger von Arbeitsunfällen betroffen und ganz allgemein häufiger körperlichen Gefahren bei der Arbeit ausgesetzt sind (Eurofound, 2025; Statistisches Bundesamt, 2025b).

In der aktuellen Gleichstellungsliteratur wird darauf hingewiesen, dass gerade die Frage, wie Männer dabei unterstützt werden können, auch andere als klassisch männliche Lebensentwürfe zu wählen, Grundvoraussetzung gelingender Gleichstellung ist: „For women to achieve career, family, and equity, fathers will have to make the same demands at work that women make, and they will have to take charge at home so women can take charge at work“ (Goldin, 2021, hier: Seite 234; siehe auch Gärtner, 2024; Gers-terkamp, 2014; Heesen, 2022; Volz, 2012). So zeigt der Väterreport 2023 des BMBFSFJ, dass sich heutige Väter viel mehr als früher eine partnerschaftlich organisierte Aufgabenteilung wünschen, zugleich aber die tatsächlichen Ungleichheiten bei Arbeitszeiten und Carearbeit nur geringe Änderungen erfahren haben (Juncke und andere, 2023).

Der Mikrozensus (Statistisches Bundesamt, 2025c) bietet neben umfassenden Informationen zur Erwerbsbeteiligung und zur Arbeitszeit auch differenzierte Angaben zur gewünschten Arbeitszeit. Der vorliegende Beitrag nutzt diese Datengrundlage, um zu untersuchen, wie sich Männer und Frauen sowie Väter und Mütter im Jahr 2024 in ihren Arbeitszeiten und Arbeitszeitwünschen voneinander unterscheiden haben. Kapitel 2 gibt zunächst einen kurzen Überblick über geschlechtsspezifische Ungleichheiten am Arbeitsmarkt, auch über die Merkmale zur Arbeitszeit hinaus. Kapitel 3 vergleicht Arbeitszeiten und Arbeitszeitwünsche. Kapitel 4 stellt Unterschiede zwischen Männern und Frauen zu weiteren Fragen mit Bezug zur Arbeitszeit dar, beispielsweise den Zeitaufwand für den Weg zum Arbeitsplatz sowie Wochenend-, Nacht- und Schichtarbeit.

<sup>1</sup> Wir danken Matthias Keller für die Unterstützung bei der Erstellung von Auswertungen aus dem Mikrozensus.

## 2

## Situation von Männern und Frauen auf dem Arbeitsmarkt

Männer sind in Deutschland häufiger als Frauen erwerbstätig und weisen höhere Verdienste auf, was mit besseren Chancen auf wirtschaftliche Unabhängigkeit einhergeht.

Im Jahr 2024 lag der Anteil der Erwerbstätigen an der Bevölkerung in allen Altersgruppen bei Männern höher als bei Frauen.<sup>12</sup> Dieser Beitrag verwendet das Konzept der realisierten Erwerbstätigkeit, das Personen, die ihre Tätigkeit wegen Mutterschutz oder Elternzeit unterbrochen haben, nicht als (realisiert)

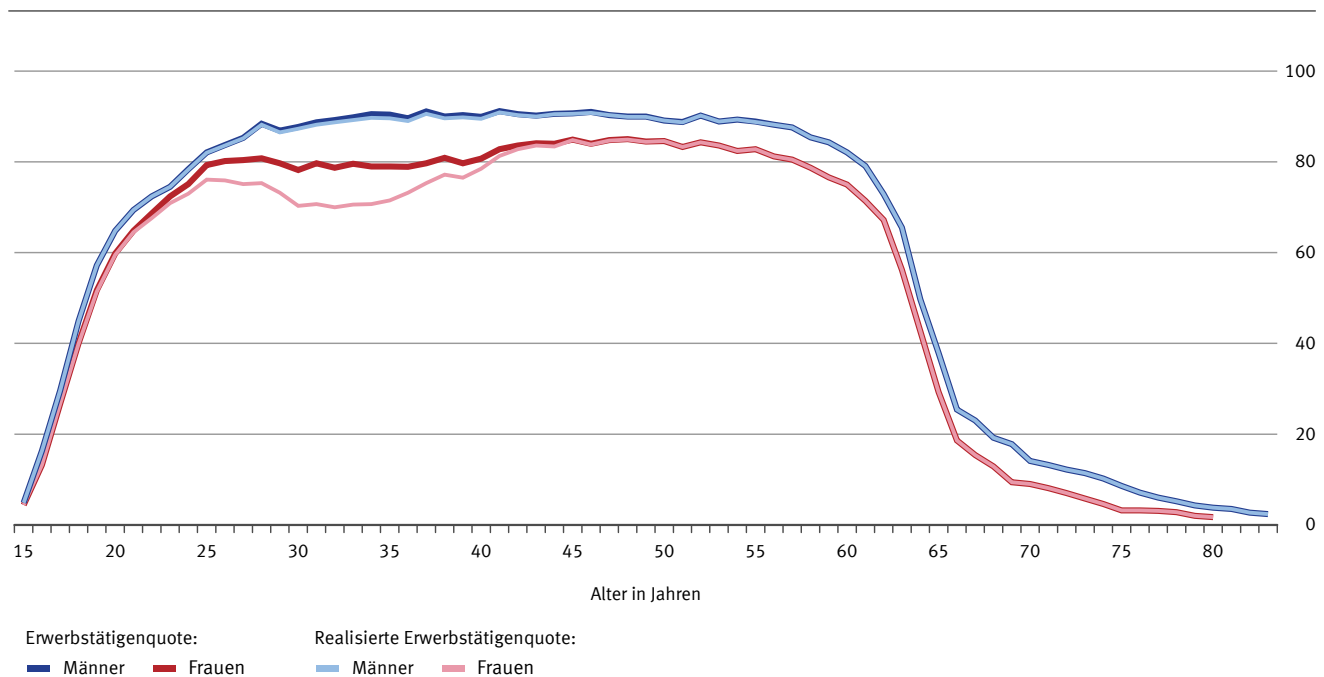
- 2 Um die geschlechtsspezifischen Unterschiede in verschiedenen Altersgruppen differenziert abbilden zu können, werden im Folgenden die Altersgruppen 15 bis 24 Jahre (Ausbildung/Berufseinstieg), 25 bis 54 Jahre (Haupterwerbs- und Familienphase), 55 bis 64 Jahre (reduzierte Erwerbstätigkeit vor dem Ruhestand) und 65 bis 79 Jahre (im Ruhestandsalter) betrachtet.

erwerbstätig zählt.<sup>13</sup> Wegen der unterschiedlichen Inanspruchnahme dieser Leistungen vermeidet dieses Konzept im Gleichstellungskontext Verzerrungen beim Vergleich der Geschlechter. Betrachtet man die realisierten Erwerbstätigenquoten, so waren die Unterschiede zwischen Männern und Frauen 2024 mit 4 Prozentpunkten bei den 15- bis 24-Jährigen am geringsten und in der Haupteerwerbs- und Familienphase mit 10,5 Prozentpunkten am höchsten. Die Abweichung bei den 55- bis 64-Jährigen (7,3 Prozentpunkte) und den 65- bis 79-Jährigen (6,3 Prozentpunkte) lagen dazwischen.<sup>14</sup> Unterschiede zwischen der realisierten Erwerbstätigenquote und der Erwerbstätigenquote fallen dabei vor allem in der Gruppe der 25- bis 54-Jährigen ins Gewicht und zei-

- 3 Realisiert erwerbstätig sind demnach alle Personen, die nach den Vorgaben der Internationalen Arbeitsorganisation (ILO, 2023) als erwerbstätig gezählt werden, sofern diese nicht ihre Tätigkeit in der Berichtswoche wegen Mutterschutz oder Elternzeit unterbrochen haben (Hochgürtel, 2018; Keller/Körner, 2023).
- 4 Hierbei handelt es sich um eine Periodenbetrachtung für das Jahr 2024, daher kann daraus nicht geschlossen werden, wie etwa die Unterschiede bei den heute 15- bis 24-Jährigen ausfallen werden, wenn diese ein höheres Alter erreicht haben.

### Grafik 1

(Realisierte) Erwerbstätigenquoten von Männern und Frauen in %



Ersterggebnisse des Mikrozensus 2024.

## Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung

gen, dass Männer ihre bezahlte Tätigkeit seltener wegen Elternzeit unterbrechen als Frauen. [↗ Grafik 1](#)

Erwerbstätige Männer arbeiten deutlich häufiger Vollzeit als Frauen. Die Vollzeitquote (Anteil der Erwerbstätigen in Vollzeit an allen Erwerbstätigen) fiel im Jahr 2024 bei den Männern in allen Altersgruppen rund eineinhalbmals so hoch aus wie bei den Frauen. Lediglich bei den 15- bis 24-Jährigen war der Abstand etwas geringer. Während etwa gut 90 % der 25- bis 54-jährigen Männer Vollzeit arbeiteten, traf dies nur auf knapp 52 % der Frauen zu. Bei den 55- bis 64-Jährigen und den 65- bis 79-Jährigen waren Männer sogar fast doppelt so häufig in Vollzeit tätig wie Frauen.

Männer waren etwas häufiger erwerbslos, Frauen sind öfter in der Stillen Reserve<sup>15</sup> und in der Gruppe der Sonstigen Nichterwerbspersonen ohne Arbeitswunsch als Männer. [↗ Tabelle 1](#)

Die höhere Erwerbstätigkeit ([Gender Employment Gap](#)) und die höhere Vollzeitquote ([Gender Hours Gap](#)) verbessern die Möglichkeiten von Männern, wirtschaftlich unabhängig zu leben. Darüber hinaus fielen die Bruttostundenverdienste von Frauen

([Gender Pay Gap](#)) 2024 um 16 % niedriger aus als bei Männern. Fasst man die Unterschiede bei Erwerbstätigkeit, bezahlten Arbeitsstunden und Stundenlöhnen zusammen, so kommt man für das Lebenserwerbseinkommen bezogen auf das Jahr 2022 auf einen [Gender Lifetime Earnings Gap](#) von 43 % (OECD, 2025). Die höheren Verdienstmöglichkeiten von Männern gehen damit einher, dass diese weniger Zeit mit unbezahlter Arbeit im Haushalt verbringen ([Gender Care Gap](#); Statistisches Bundesamt, 2025d).

Arbeitszeiten wirken dabei durchaus ambivalent: Hohe Arbeitszeiten ermöglichen einerseits höhere Verdienste, können aber andererseits die Verwirklichungschancen in anderen Lebensbereichen und die Übernahme unbezahlter Arbeit im Haushalt oder die Gesundheit einschränken. Daher werden im Folgenden die geschlechtsspezifische Verteilung von Arbeitszeiten und die Arbeitszeitwünsche von Männern und Frauen betrachtet. Der Beitrag beschränkt sich auf Ergebnisse für Gesamtdeutschland. Jedoch weichen die geschlechtsspezifischen Unterschiede bei Beschäftigung und Arbeitszeit in Ost und West nach wie vor voneinander ab.<sup>16</sup>

5 Unter Stiller Reserve versteht man Nichterwerbspersonen, die zwar Arbeit suchen, jedoch kurzfristig, das heißt innerhalb von zwei Wochen, nicht zur Verfügung stehen, sowie Nichterwerbspersonen, die aus anderen Gründen keine Arbeit suchen, aber grundsätzlich arbeiten würden und für diese auch verfügbar sind. Darüber hinaus werden Sonstige Nichterwerbspersonen mit Arbeitswunsch einbezogen (Rengers/Fuchs, 2022).

6 So sind die geschlechtsspezifischen Unterschiede im Osten weniger ausgeprägt als im Westen. In der Haupteinwerbungs- und Familienphase (25- bis 54-Jährige) betrug die Differenz der realisierten Erwerbstätigenquote zwischen Männern und Frauen im Jahr 2024 in den neuen Ländern (einschließlich Berlins) 6,9 Prozentpunkte, während es im früheren Bundesgebiet 11,3 Prozentpunkte waren. Dies betrifft auch die Arbeitszeit: Während erwerbstätige Männer im Osten im Schnitt gewöhnlich 4,4 Stunden länger gegen Bezahlung gearbeitet haben als Frauen, betrug die Differenz im Westen 8,1 Stunden (Kümmerling, 2024).

**Tabelle 1**

**Ausgewählte Kennzahlen für Männer und Frauen am Arbeitsmarkt nach Altersgruppen**

|                                                               | Männer          |                 |                 |                 |                 | Frauen          |                 |                 |                 |                 |
|---------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                                               | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
|                                                               | %               |                 |                 |                 |                 |                 |                 |                 |                 |                 |
| Erwerbstätigenquote                                           | 69,1            | 52,3            | 89,2            | 78,8            | 16,4            | 61,0            | 48,9            | 81,7            | 71,5            | 10,1            |
| Realisierte Erwerbstätigenquote                               | 69,0            | 52,3            | 88,9            | 78,8            | 16,4            | 59,4            | 48,3            | 78,4            | 71,5            | 10,1            |
| Vollzeitquote                                                 | 86,2            | 74,7            | 90,2            | 88,9            | 38,6            | 50,3            | 60,7            | 51,9            | 46,3            | 19,7            |
| Realisierte Vollzeitquote                                     | 86,2            | 74,7            | 90,3            | 88,9            | 38,6            | 50,0            | 60,5            | 51,4            | 46,3            | 19,7            |
| Erwerbslose <sup>1</sup>                                      | 2,6             | 4,2             | 3,2             | 2,0             | 0,2             | 2,0             | 3,3             | 2,6             | 1,6             | (0,1)           |
| Stille Reserve <sup>1</sup>                                   | 4,2             | 10,4            | 3,8             | 3,2             | 1,9             | 5,3             | 10,7            | 6,6             | 3,5             | 1,1             |
| Sonstige Nichterwerbspersonen ohne Arbeitswunsch <sup>1</sup> | 24,1            | 33,1            | 3,8             | 15,9            | 81,5            | 31,7            | 37,1            | 9,2             | 23,4            | 88,7            |

<sup>1</sup> Anteil an der Bevölkerung.  
Erstergebnisse des Mikrozensus 2024.

## 3

### Arbeitszeiten und Arbeitszeitwünsche

#### 3.1 Operationalisierung

Die Messung von Arbeitszeiten, Arbeitszeitwünschen und damit einhergehend von Unter- und Überbeschäftigung ist mit vielen Herausforderungen verbunden.

Zunächst einmal zu den Begrifflichkeiten und Definitionen: Arbeitszeiten sind nicht gleich Arbeitszeiten. Im Jahr 2008 verabschiedete die 18. International Conference of Labour Statisticians (ICLS) unter anderem eine Resolution zur Messung der Arbeitszeit, die insgesamt sieben Arbeitszeitkonzepte unterscheidet und im Detail definiert (ILO, 2008; siehe dazu auch Rengers und andere, 2017). Dies sind tatsächlich geleistete Arbeitsstunden (hours actually worked), bezahlte Arbeitsstunden (hours paid for), Normalarbeitsstunden (normal hours of work), vertraglich vereinbarte Arbeitsstunden (contractual hours of work), gewöhnlich geleistete Arbeitsstunden (hours usually worked) sowie Überstunden (overtime hours of work) und Stunden in Abwesenheit vom Arbeitsplatz (absence from work hours).

Im Folgenden werden die gewöhnlich geleisteten Arbeitsstunden betrachtet, teilweise auch als durchschnittlich geleistete Arbeitsstunden bezeichnet. Im Mikrozensus werden sie über die Frage „Wie viele Stunden arbeiten Sie normalerweise pro Woche, einschließlich regelmäßiger Mehrstunden und Bereitschaftszeiten?“ erfasst. Ergänzt wird diese Frage mit dem Hinweis beziehungsweise der Bitte, bei stark variierenden Arbeitszeiten die durchschnittlichen Wochenarbeitsstunden auf Basis der letzten 4 bis 12 Wochen zu schätzen.

Auch die Erfassung von Arbeitszeitwünschen ist ein Thema von hohem Interesse, dem sich in Deutschland mehrere Institutionen widmen. Neben dem Mikrozensus dienen auch das Sozio-oekonomische Panel SOEP (DIW, 2025) und die BAuA-Arbeitszeitbefragung (BAuA, 2023; Häring und andere, 2024)

der Bundesanstalt für Arbeitsschutz und Arbeitsmedizin dazu, Arbeitszeiten und Arbeitszeitwünsche zu untersuchen.

Frühere Untersuchungen (Rengers und andere, 2017) zu verschiedenen Umfragen zu Arbeitszeitwünschen haben gezeigt, dass befragte Personen auf Frageformulierungen, Fragereihenfolge und Abfragekontext bei der Ermittlung von Arbeitszeitwünschen sensibel reagieren, obwohl alle Erhebungen einheitlich den Hinweis enthalten, dass mit einer Änderung der Arbeitszeit eine entsprechende Verdienstanpassung verbunden wäre.

Mit diesen Erkenntnissen wurde die Erfassung im Mikrozensus weiter optimiert. Seit 2021 werden Arbeitszeitwünsche im Mikrozensus in einem zweistufigen Abfrageverfahren über die nachfolgenden Formulierungen erfasst:

- > „Würden Sie gerne Ihre normale Wochenarbeitszeit beibehalten oder mit entsprechender Anpassung des Verdienstes verändern?“
- > „Wie viele Stunden pro Woche möchten Sie insgesamt arbeiten?“

Bei beiden Fragen wird darauf hingewiesen, dass zur Wochenarbeitszeit sowohl Haupt- als auch Nebentätigkeiten zählen.

#### Exkurs: Über- und Unterbeschäftigte

Überbeschäftigte sind Erwerbstätige, die den Wunsch haben, ihre Arbeitsstunden zu reduzieren, und dafür ein verringertes Einkommen hinnehmen. Unterbeschäftigte im weiteren Sinne (Unterbeschäftigung i.w.S.) sind Erwerbstätige, die ihre Arbeitszeit mit entsprechend höherem Verdienst erhöhen wollen. Davon zu unterscheiden sind Unterbeschäftigte im engeren Sinne (Unterbeschäftigung i.e.S.), die ihre Arbeitszeit erhöhen möchten und für zusätzliche Arbeitsstunden innerhalb von zwei Wochen zur Verfügung stehen könnten.<sup>17</sup>

<sup>17</sup> Unterbeschäftigte i.e.S. sind in Anlehnung an die Resolution über Arbeitsstatistiken, Erwerbstätigkeit und die Unterauslastung des Arbeitskräfteangebots definiert (ILO, 2023). In der ILO-Resolution wird nur das ungenutzte Arbeitskräftepotenzial von Erwerbstätigen mit Wunsch nach Erhöhung der Arbeitszeit betrachtet, Erwerbstätige mit Wunsch nach Reduzierung der Arbeitszeit nicht.

# Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung

## Exkurs: Einfluss von Befragungsmodes

Die Sensibilität der Befragungsergebnisse scheint auch vom verwendeten Befragungsmode abhängig zu sein. So wurde beim Mikrozensus ab 2020 ein Multi-Mode-Erhebungsdesign eingeführt, das den Befragten die Beantwortung auch online (CAWI) erlaubt.<sup>18</sup> Bei den Ergebnissen zu Arbeitszeitwünschen fielen dabei deutliche Unterschiede zwischen den Ergebnissen von CAWI- und Nicht-CAWI-Antworten auf, die in allen Altersklassen und bei beiden Geschlechtern zu beobachten waren. Im Vergleich zu den Berichtsjahren vor 2020 haben sich die Ergebnisse zu den Arbeitszeitwünschen der Erwerbstätigen erheblich verändert und sogar umgekehrt. Von den Erwerbstätigen mit Wunsch nach einer Arbeitszeitveränderung wollten vor 2020 ungefähr zwei Drittel die Arbeitszeit erhöhen, entsprechend wollte

umgekehrt ein Drittel von ihnen die Arbeitszeit verringern. Nach 2020 wollte nur noch knapp ein Viertel der Erwerbstätigen mit Arbeitszeitdiskrepanzen die Arbeitszeit erhöhen, drei Viertel dagegen wollten reduzieren. Es kann dabei nicht eindeutig festgestellt werden, in welchem Ausmaß die veränderten Arbeitszeitwünsche Abbildungen der Realität oder Effekte der Befragungsmethode sind. Nach Coronapandemie, Work-Life-Balance-Diskussionen und zunehmender Arbeitsmarktbeteiligung der Generation Z<sup>19</sup> ist eine stärkere Fokussierung auf mehr Freizeit denkbar.

8 Neben CAWI (Computer Assisted Web Interview) stehen weiterhin das persönliche Interview CAPI (Computer Assisted Personal Interview), das telefonische Interview CATI (Computer Assisted Telephone Interview) sowie der Papierfragebogen PAPI (Paper and Pencil Interview) als Befragungsmethoden zur Verfügung. Die Modes CAWI und PAPI werden dabei ohne den Einsatz von Interviewenden durchgeführt.

9 Mit Generation Z werden soziologisch verallgemeinernd meist die Geburtsjahrgänge von 1995 bis 2010 bezeichnet, teilweise auch von 1997 bis 2012.

**Tabelle 2**  
**Realisierte Erwerbstätigkeit nach Altersklassen**

|                                     | Männer               |                 |                 |                 |                 | Frauen          |                 |                 |                 |                 |
|-------------------------------------|----------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                     | 15 bis 79 Jahre      | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
|                                     | 1 000   <sup>1</sup> |                 |                 |                 |                 |                 |                 |                 |                 |                 |
| Realisierte Erwerbstätigkeit        | 22 578               | 2 257           | 14 257          | 5 096           | 968             | 19 508          | 1 930           | 12 208          | 4 687           | 683             |
| Vollzeittätige                      | 19 457               | 1 686           | 12 867          | 4 530           | 374             | 9 752           | 1 168           | 6 280           | 2 169           | 134             |
| Teilzeittätige <sup>2</sup>         | 3 121                | 571             | 1 390           | 566             | 594             | 9 757           | 763             | 5 928           | 2 518           | 549             |
| Unterbeschäftigte im weiteren Sinne | 606                  | 77              | 446             | 70              | 12              | 780             | 64              | 598             | 110             | 8               |
| Vollzeittätige                      | 291                  | 25              | 229             | 34              | /               | 93              | 9               | 71              | 12              | /               |
| Teilzeittätige <sup>2</sup>         | 315                  | 52              | 217             | 36              | 9               | 686             | 55              | 527             | 98              | 7               |
| Unterbeschäftigte nicht verfügbar   | 258                  | 38              | 188             | 26              | (5)             | 409             | 33              | 328             | 45              | /               |
| Vollzeittätige                      | 115                  | 11              | 88              | 13              | /               | 43              | (5)             | 33              | 4               | /               |
| Teilzeittätige <sup>2</sup>         | 143                  | 27              | 100             | 13              | /               | 366             | 28              | 294             | 41              | /               |
| Unterbeschäftigte verfügbar         | 348                  | 39              | 258             | 44              | 7               | 370             | 31              | 270             | 65              | (5)             |
| Vollzeittätige                      | 177                  | 13              | 141             | 21              | /               | 50              | (4)             | 38              | 8               | /               |
| Teilzeittätige <sup>2</sup>         | 171                  | 25              | 117             | 23              | (6)             | 320             | 27              | 232             | 57              | (4)             |
| Überbeschäftigte                    | 2 372                | 113             | 1 651           | 522             | 85              | 1 757           | 125             | 1 211           | 374             | 48              |
| Vollzeittätige                      | 2 231                | 101             | 1 587           | 496             | 47              | 1 369           | 105             | 969             | 276             | 19              |
| Teilzeittätige <sup>2</sup>         | 141                  | 12              | 64              | 26              | 38              | 388             | 20              | 242             | 98              | 29              |

1 Personen in Privathaushalten.

2 Die Angabe zur Teilzeittätigkeit bezieht sich auf die Selbsteinschätzung der Befragten. Erstergebnisse des Mikrozensus 2024.



## 3.2 Ergebnisse

### Geschlecht und Altersklassen

#### Realisierte Erwerbstätigkeit und Arbeitszeiten

Von den knapp 22,6 Millionen 15- bis 79-jährigen Männern in realisierter Erwerbstätigkeit waren im Jahr 2024 fast 19,5 Millionen vollzeitbeschäftigt, gut 3,1 Millionen teilzeitbeschäftigt. Die gut 19,5 Millionen gleichaltrigen Frauen verteilten sich etwa hälftig auf Vollzeit- beziehungsweise Teilzeittätige. Zwischen den einzelnen hier betrachteten Altersklassen gab es hinsichtlich des Verhältnisses von Vollzeit zu Teilzeit bei beiden Geschlechtern deutliche Unterschiede. [↗ Tabelle 2](#)

Auch bei den gewöhnlich geleisteten Wochenarbeitsstunden unterschieden sich die Geschlechter deutlich. So gaben vollzeitbeschäftigte Männer insgesamt mit 41,1 Stunden eine höhere Stundenzahl an als vollzeittätige Frauen, die auf 39,6 Stun-

den kamen. Dagegen sieht das Verhältnis zwischen den Geschlechtern bei einer Teilzeittätigkeit genau anders herum aus: 19,6 Stunden (Männer) zu 21,9 Stunden (Frauen). [↗ Tabelle 3](#)

Insgesamt wollten gut 0,6 Millionen Männer und knapp 0,8 Millionen Frauen im Alter von 15 bis 79 Jahren ihre Arbeitszeit erhöhen, fast 2,4 Millionen Männer und 1,8 Millionen Frauen äußerten den Wunsch nach einer Arbeitszeitreduzierung. In Quoten<sup>10</sup> ausgedrückt sind damit 2,7 % der erwerbstätigen Männer und 4,0 % der erwerbstätigen Frauen unterbeschäftigt, 10,5 % der Männer und 9,0 % der Frauen sind überbeschäftigt. Die nachfolgenden Ausführungen zeigen, dass es nicht nur geschlechtsspezifische Unterschiede zwischen diesen Quoten gibt. Die Höhe der Quoten wird auch davon beeinflusst, ob eine

10 Die Unterbeschäftigtenquote gibt den Anteil der Unterbeschäftigten an den Erwerbstätigen an, die Überbeschäftigtenquote den Anteil der Überbeschäftigten an den Erwerbstätigen. Die Diskrepanzenquote zeigt den Anteil aller Erwerbstätigen mit Arbeitszeitdiskrepanzen (Unterbeschäftigte und Überbeschäftigte) an den Erwerbstätigen.

**Tabelle 3**

#### Arbeitszeiten nach Altersklassen

|                                                         | Männer          |                 |                 |                 |                 | Frauen          |                 |                 |                 |                 |
|---------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                                         | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
| Gewöhnlich geleistete Wochenarbeitsstunden <sup>1</sup> |                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |
| Realisierte Erwerbstätigkeit                            | 38,2            | 33,0            | 39,6            | 39,2            | 23,8            | 30,8            | 29,3            | 32,0            | 30,2            | 17,2            |
| Vollzeittätige                                          | 41,1            | 39,6            | 41,2            | 41,4            | 41,3            | 39,6            | 39,3            | 39,8            | 39,3            | 39,4            |
| Teilzeittätige <sup>2</sup>                             | 19,6            | 13,7            | 23,9            | 22,1            | 12,8            | 21,9            | 14,1            | 23,7            | 22,4            | 11,8            |
| Unterbeschäftigte im weiteren Sinne                     | 29,1            | 23,0            | 30,5            | 29,0            | 19,5            | 22,7            | 19,1            | 23,2            | 23,0            | 13,6            |
| Vollzeittätige                                          | 40,2            | 39,9            | 40,2            | 40,4            | /               | 36,8            | 38,0            | 36,7            | 35,9            | /               |
| Teilzeittätige <sup>2</sup>                             | 18,9            | 14,9            | 20,2            | 18,4            | 11,9            | 20,8            | 16,0            | 21,4            | 21,4            | 10,3            |
| Unterbeschäftigte nicht verfügbar                       | 28,2            | 21,6            | 29,5            | 28,8            | (22,8)          | 22,4            | 19,2            | 22,8            | 22,5            | /               |
| Vollzeittätige                                          | 39,5            | 38,6            | 39,8            | 38,3            | /               | 36,4            | (38,3)          | 36,3            | 35,4            | /               |
| Teilzeittätige <sup>2</sup>                             | 19,1            | 14,4            | 20,5            | 18,9            | /               | 20,7            | 15,6            | 21,2            | 21,2            | /               |
| Unterbeschäftigte verfügbar                             | 29,8            | 24,3            | 31,1            | 29,2            | 17,3            | 23,1            | 18,9            | 23,7            | 23,4            | (14,5)          |
| Vollzeittätige                                          | 40,7            | 41,0            | 40,5            | 41,8            | /               | 37,0            | (37,5)          | 37,1            | 36,2            | /               |
| Teilzeittätige <sup>2</sup>                             | 18,7            | 15,5            | 19,9            | 18,1            | (10,9)          | 21,0            | 16,3            | 21,5            | 21,6            | (10,2)          |
| Überbeschäftigte                                        | 42,4            | 38,6            | 42,9            | 43,2            | 31,4            | 38,7            | 37,3            | 39,4            | 38,2            | 25,6            |
| Vollzeittätige                                          | 43,5            | 41,2            | 43,4            | 44,1            | 44,7            | 41,9            | 40,7            | 41,9            | 42,2            | 42,9            |
| Teilzeittätige <sup>2</sup>                             | 24,9            | 16,6            | 31,7            | 26,1            | 15,2            | 27,2            | 19,0            | 29,5            | 27,1            | 14,0            |

1 Gewöhnlich geleistete Wochenarbeitszeit in Haupt- und – falls vorhanden – Nebentätigkeit. Von den 15- bis 79-jährigen in realisierter Erwerbstätigkeit haben 4,8 % (Männer) beziehungsweise 5,4 % (Frauen) eine Nebentätigkeit.

2 Die Angabe zur Teilzeittätigkeit bezieht sich auf die Selbsteinschätzung der Befragten. Erstergebnisse des Mikrozensus 2024.

## Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung

Tätigkeit in Vollzeit (VZ) oder Teilzeit (TZ) vorliegt.  
[Tabelle 4](#)

### Unter- und Überbeschäftigung bei Vollzeit

Männer sind häufiger VZ-Unterbeschäftigte als Frauen. Bei den 15- bis 79-jährigen Männern lag die VZ-Unterbeschäftigtenquote<sup>11</sup> i.w.S. im Jahr 2024 bei 1,5 % und war damit höher als die der gleichaltrigen Frauen (1,0 %). Der Unterschied zeigt sich über alle Altersklassen. Männliche VZ-Unterbeschäftigte i.w.S. hatten zwar eine niedrigere gewöhnliche Wochenarbeitszeit als die gesamte Gruppe der vollzeiterwerbstätigen Männer (40,2 Stunden zu 41,1 Stunden bei den 15- bis 79-jährigen insgesamt), weshalb ihr Wunsch nach Mehrarbeit „plausibel“ erscheint. Ihre Wochenarbeitszeiten waren aber bereits höher als die der vollzeiterwerbstätigen Frauen, die im Durchschnitt eine gewöhnliche

Wochenarbeitszeit von 39,6 Stunden hatten (siehe Tabelle 3).

Bei vollzeittätigen Personen in realisierter Erwerbstätigkeit war die VZ-Unterbeschäftigungsquote i.w.S. jeweils in der Haupterwerbs- und Familienphase (25 bis 54 Jahre) am höchsten mit 1,8 % (Männer) und 1,1 % (Frauen). Das entspricht 229 000 Männern beziehungsweise 71 000 Frauen dieser Altersklasse, die ihre Arbeitszeit von durchschnittlich 40,2 Stunden (Männer) um 7,7 Wochenstunden beziehungsweise von durchschnittlich 36,7 Stunden um 7,5 Stunden (Frauen) erhöhen wollten.

Im Gegensatz dazu sind die geschlechtsspezifischen Unterschiede bei VZ-Überbeschäftigtenquoten umgekehrt. Hier wollten vollzeittätige Frauen in allen Altersklassen ihre Arbeitszeit häufiger verringern als dies bei den Männern der Fall war. Sowohl Männer als auch Frauen gaben dabei eine gewöhnliche Wochenarbeitszeit an, die deutlich über der

11 Alle Quoten, die vorangestellt ein VZ (TZ) haben, beziehen sich nur auf vollzeittätige (teilzeittätige) Personen.

**Tabelle 4**

### Realisierte Erwerbstätigkeit und Arbeitszeitwünsche in Quoten nach Altersklassen

|                                                          | Männer <sup>1</sup> |                 |                 |                 |                 | Frauen <sup>1</sup> |                 |                 |                 |                 |
|----------------------------------------------------------|---------------------|-----------------|-----------------|-----------------|-----------------|---------------------|-----------------|-----------------|-----------------|-----------------|
|                                                          | 15 bis 79 Jahre     | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre     | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
|                                                          | %                   |                 |                 |                 |                 |                     |                 |                 |                 |                 |
| Realisierte Erwerbstätigenquote                          | 69,0                | 52,3            | 88,9            | 78,8            | 16,4            | 59,4                | 48,3            | 78,4            | 71,5            | 10,1            |
| Teilzeitquote                                            | 13,8                | 25,3            | 9,7             | 11,1            | 61,4            | 50,0                | 39,5            | 48,6            | 53,7            | 80,3            |
| Unterbeschäftigtenquote im weiteren Sinne                | 2,7                 | 3,4             | 3,1             | 1,4             | 1,3             | 4,0                 | 3,3             | 4,9             | 2,3             | 1,1             |
| Vollzeittätige                                           | 1,5                 | 1,5             | 1,8             | 0,8             | /               | 1,0                 | 0,8             | 1,1             | 0,5             | /               |
| Teilzeittätige <sup>2</sup>                              | 10,1                | 9,1             | 15,6            | 6,4             | 1,6             | 7,0                 | 7,2             | 8,9             | 3,9             | 1,3             |
| Überbeschäftigtenquote                                   | 10,5                | 5,0             | 11,6            | 10,3            | 8,8             | 9,0                 | 6,5             | 9,9             | 8,0             | 7,0             |
| Vollzeittätige                                           | 11,5                | 6,0             | 12,3            | 11,0            | 12,5            | 14,0                | 9,0             | 15,4            | 12,7            | 14,2            |
| Teilzeittätige <sup>2</sup>                              | 4,5                 | 2,1             | 4,6             | 4,6             | 6,5             | 4,0                 | 2,6             | 4,1             | 3,9             | 5,2             |
| Diskrepanzenquote                                        | 13,2                | 8,4             | 14,7            | 11,6            | 10,1            | 13,0                | 9,8             | 14,8            | 10,3            | 8,1             |
| Anteil der Unterbeschäftigung im weiteren Sinne          | 20,3                | 40,6            | 21,3            | 11,9            | 12,7            | 30,7                | 33,9            | 33,1            | 22,7            | 14,1            |
| Anteil der Überbeschäftigung                             | 79,7                | 59,4            | 78,7            | 88,1            | 87,3            | 69,3                | 66,1            | 66,9            | 77,3            | 85,9            |
| Vollzeit-Diskrepanzenquote                               | 13,0                | 7,5             | 14,1            | 11,7            | /               | 15,0                | 9,8             | 16,6            | 13,3            | /               |
| Anteil der Vollzeit-Unterbeschäftigung im weiteren Sinne | 11,5                | 19,8            | 12,6            | 6,4             | /               | 6,4                 | 8,0             | 6,9             | 4,1             | /               |
| Anteil der Vollzeit-Überbeschäftigung                    | 88,5                | 80,2            | 87,4            | 93,6            | /               | 93,6                | 92,0            | 93,1            | 95,9            | /               |
| Teilzeit-Diskrepanzenquote                               | 14,6                | 11,2            | 20,2            | 11,0            | 8,0             | 11,0                | 9,8             | 13,0            | 7,8             | 6,5             |
| Anteil der Teilzeit-Unterbeschäftigung im weiteren Sinne | 69,1                | 81,5            | 77,1            | 58,1            | 19,6            | 63,9                | 73,5            | 68,6            | 49,9            | 19,6            |
| Anteil der Teilzeit-Überbeschäftigung                    | 30,9                | 18,5            | 22,9            | 41,9            | 80,4            | 36,1                | 26,5            | 31,4            | 50,1            | 80,4            |

1 Personen in Privathaushalten.

2 Die Angabe zur Teilzeittätigkeit bezieht sich auf die Selbsteinschätzung der Befragten. Erstergebnisse des Mikrozensus 2024.

eines durchschnittsvollzeittätigen Mannes beziehungsweise einer durchschnittsvollzeittätigen Frau lagen: 43,5 Stunden zu 41,1 Stunden (Männer) und 41,9 Stunden zu 39,6 Stunden (Frauen). Trotz unterschiedlicher Ausgangsposition wollten überbeschäftigt vollzeittätige Männer und Frauen im Alter von 15 bis 79 Jahren ihre Arbeitszeit in absoluten Stunden fast in gleichem Umfang reduzieren, nämlich um 10,8 Stunden (Männer) beziehungsweise 10,3 Stunden (Frauen). [↘ Tabelle 5](#)

### Unter- und Überbeschäftigung bei Teilzeit

Teilzeittätige Männer waren insgesamt mit einer TZ-Unterbeschäftigtenquote i.w.S. von 10,1 % relativ häufiger von Unterbeschäftigung betroffen als Frauen mit einer Quote von 7,0 %. In absoluten Zah-

len sieht dies aufgrund der deutlich höheren Teilzeitquote bei Frauen etwas anders aus: Hier stehen 315 000 teilzeittätigen Männern mit Wunsch nach einer Arbeitszeiterhöhung 686 000 teilzeittätige Frauen mit gleichem Wunsch gegenüber. Bei ihrer durchschnittlich geleisteten wöchentlichen Arbeitszeit gaben unterbeschäftigt teilzeittätige Männer dabei mit 18,9 Stunden weniger Stunden an als ihr weibliches Gegenstück (20,8 Stunden). Gleichzeitig ist ihre Angabe bei der gewünschten Stundenveränderung mit +15,4 Stunden aber auch entsprechend höher als diejenige der Frauen (+10,4 Stunden).

Die geschlechtsspezifischen Aussagen zu den TZ-Unterbeschäftigtenquoten i.w.S. gelten für alle Altersklassen. Betrachtet man die Angaben zu den gewöhnlichen Wochenarbeitsstunden und der ge-

**Tabelle 5**

### Arbeitszeitwünsche nach Altersklassen

|                                                              | Männer          |                 |                 |                 |                 | Frauen          |                 |                 |                 |                 |
|--------------------------------------------------------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                                                              | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
| Gewünschte Veränderung der Wochenarbeitsstunden <sup>1</sup> |                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |
| Realisierte Erwerbstätigkeit (Saldo im weiteren Sinne)       | - 0,8           | 0,0             | - 0,8           | - 1,0           | - 1,1           | - 0,5           | - 0,2           | - 0,5           | - 0,6           | - 0,8           |
| Vollzeittätige                                               | - 1,1           | - 0,4           | - 1,1           | - 1,2           | - 2,0           | - 1,4           | - 0,7           | - 1,5           | - 1,4           | - 2,2           |
| Teilzeittätige <sup>2</sup>                                  | + 1,2           | + 1,2           | + 2,0           | + 0,5           | - 0,3           | + 0,4           | + 0,7           | + 0,6           | 0,0             | - 0,3           |
| Realisierte Erwerbstätigkeit (Saldo im engeren Sinne)        | - 0,9           | - 0,2           | - 1,0           | - 1,1           | - 1,1           | - 0,7           | - 0,4           | - 0,7           | - 0,7           | - 0,8           |
| Vollzeittätige                                               | - 1,2           | - 0,4           | - 1,2           | - 1,3           | - 2,0           | - 1,4           | - 0,7           | - 1,5           | - 1,4           | - 2,2           |
| Teilzeittätige <sup>2</sup>                                  | + 0,5           | + 0,5           | + 0,9           | + 0,2           | - 0,4           | 0,0             | + 0,2           | + 0,1           | - 0,1           | - 0,3           |
| Unterbeschäftigte im weiteren Sinne                          | + 12,1          | + 12,9          | + 12,1          | + 11,9          | + 11,2          | + 10,1          | + 11,3          | + 10,1          | + 9,4           | + 8,5           |
| Vollzeittätige                                               | + 7,9           | + 7,9           | + 7,7           | + 8,4           | /               | + 7,4           | + 5,7           | + 7,5           | + 7,5           | /               |
| Teilzeittätige <sup>2</sup>                                  | + 15,4          | + 14,9          | + 15,8          | + 14,8          | + 11,5          | + 10,4          | + 11,9          | + 10,4          | + 9,6           | + 8,7           |
| Unterbeschäftigte nicht verfügbar                            | + 12,4          | + 13,6          | + 12,4          | + 11,1          | (+ 11,5)        | + 9,9           | + 11,1          | + 10,0          | + 9,0           | /               |
| Vollzeittätige                                               | + 7,9           | + 8,8           | + 7,7           | + 8,2           | /               | + 7,0           | (+ 5,1)         | + 7,2           | + 7,3           | /               |
| Teilzeittätige <sup>2</sup>                                  | + 15,2          | + 15,2          | + 15,5          | + 13,6          | /               | + 10,2          | + 11,8          | + 10,2          | + 9,2           | /               |
| Unterbeschäftigte verfügbar                                  | + 11,9          | + 12,3          | + 11,8          | + 12,4          | + 11,0          | + 10,2          | + 11,5          | + 10,2          | + 9,7           | (+ 8,8)         |
| Vollzeittätige                                               | + 7,8           | + 7,2           | + 7,8           | + 8,5           | /               | + 7,6           | (+ 6,4)         | + 7,7           | + 7,6           | /               |
| Teilzeittätige <sup>2</sup>                                  | + 15,6          | + 14,6          | + 16,1          | + 15,5          | (+ 11,6)        | + 10,5          | + 12,1          | + 10,5          | + 9,9           | (+ 9,2)         |
| Überbeschäftigte                                             | - 10,7          | - 8,2           | - 10,4          | - 11,8          | - 13,7          | - 9,9           | - 8,3           | - 9,7           | - 10,7          | - 12,2          |
| Vollzeittätige                                               | - 10,8          | - 8,4           | - 10,4          | - 11,9          | - 16,5          | - 10,3          | - 8,5           | - 10,0          | - 11,3          | - 16,0          |
| Teilzeittätige <sup>2</sup>                                  | - 8,7           | - 6,6           | - 9,1           | - 9,2           | - 8,0           | - 8,3           | - 7,2           | - 8,3           | - 8,5           | - 7,9           |

1 Bei den jeweiligen Personengruppen in realisierter Erwerbstätigkeit ergeben sich die gewünschten positiven oder negativen Veränderungen der gewöhnlichen Wochenarbeitszeiten rein rechnerisch per saldo: Die entsprechenden Angaben der Unter- und Überbeschäftigten werden dazu auf alle Erwerbstätigen umgerechnet.

2 Die Angabe zur Teilzeittätigkeit bezieht sich auf die Selbsteinschätzung der Befragten.  
Erstergebnisse des Mikrozensus 2024.

wünschten Arbeitszeitveränderung zusammen mit den Teilzeitquoten, dann fallen innerhalb der Gruppe der Männer die 25- bis 54-Jährigen im Vergleich zu den anderen Altersklassen mit Extremwerten auf: Sie haben mit 9,7 % die niedrigste Teilzeitquote, mit den höchsten Wochenarbeitsstunden von 20,2 Stunden, der höchsten TZ-Unterbeschäftigtenquote von 15,6 % und der höchsten gewünschten Wochenstundenveränderung von + 15,8 Stunden.

Teilzeittätige Männer sind insgesamt mit einer TZ-Überbeschäftigtenquote von 4,5 % auch häufiger von Überbeschäftigung betroffen als Frauen mit 4,0 %. Dennoch gaben sie gleichzeitig mit 24,9 Stunden eine niedrigere gewöhnliche Wochenarbeitszeit an als teilzeittätige Frauen mit 27,2 Stunden.

### *Arbeitszeitdiskrepanzen insgesamt*

Die Altersklasse der 25- bis 54-Jährigen insgesamt war am wenigsten zufrieden mit ihrer Arbeitszeit. Das gilt für beide Geschlechter. Die Diskrepanzenquote von Männern und Frauen dieser Altersklasse war fast gleich (Männer: 14,7 %, Frauen: 14,8 %). Auch bei einer getrennten Betrachtung von vollzeit- und teilzeittätigen Männern und Frauen gelten diese Aussagen.

### **Lebensformen, Familie und Kinder**

Die folgenden Analysen betrachten Männer und Frauen im Alter von 25 bis 54 Jahren, in dem Familie und Kinder eine zentrale Rolle spielen. Dieser Lebensabschnitt kann als Hauptfamilienphase, aber auch als Haupterwerbsphase gekennzeichnet werden. Mit dieser Alterseinschränkung werden gleichzeitig geschlechtsspezifische Effekte verschiedener Lebensabschnitte so gering wie möglich gehalten.

### *Unter- und Überbeschäftigung bei Vollzeit nach Lebensformen*

↳ **Tabelle 6** zeigt die realisierte Erwerbstätigkeit und Arbeitszeitwünsche in Quoten nach Lebensform der 25- bis 54-Jährigen. Vollzeittätige Männer mit Kindern wollten ihre Arbeitszeit etwas seltener erhöhen als Männer ohne Kind. Am größten sind diese Unterschiede zwischen alleinerziehenden Männern (VZ-Unterbeschäftigtenquote 1,2 %) und allein-

stehenden Männern (VZ-Unterbeschäftigtenquote 2,4 %). Gleichzeitig gaben Männer in Vollzeit deutlich häufiger den Wunsch nach Arbeitszeitreduzierung an, der bei Männern ohne Kind sogar etwas häufiger war als bei Männern mit Kindern. Insgesamt ist die VZ-Diskrepanzenquote damit in allen Lebensformen bei Männern ohne Kind höher als bei Männern mit Kindern. In Lebensgemeinschaften ohne Kind waren beispielsweise 17,6 % der Männer mit ihrer Arbeitszeit unzufrieden, bei Männern mit Kindern waren es 13,4 %. Unabhängig von der Lebensform wollten mehr als 85 % der 25- bis 54-jährigen vollzeittätigen Männer mit Arbeitszeitdiskrepanzen ihre Arbeitszeit reduzieren.

Mit VZ-Unterbeschäftigtenquoten zwischen 0,7 % (Ehepaar ohne Kind) und 1,6 % (Alleinerziehende mit Kindern) wollten Frauen im Alter von 25 bis 54 Jahren tendenziell weniger häufig ihre Arbeitszeit erhöhen als vollzeittätige Männer dieser Altersklasse. Genau wie bei den Männern wollten auf der anderen Seite auch hier Frauen ohne Kind deutlich häufiger ihre Arbeitszeit reduzieren als Frauen mit Kindern.

### *Unter- und Überbeschäftigung bei Teilzeit nach Lebensformen*

Während 25- bis 54-jährige Männer insgesamt die oben beschriebene niedrige Teilzeitquote von 9,7 % aufwiesen, zeigten sich je nach Lebensform deutliche Unterschiede. Die TZ-Quote bei Vätern dieser Altersklasse war in allen Lebensformen höher als bei Männern ohne Kind. Väter, die beispielsweise in der Lebensform Ehepaar mit Kindern lebten, hatten eine TZ-Quote von 8,0 %, Männer ohne Kind in dieser Lebensform lagen bei einer Quote von 6,5 %. Ihren höchsten Wert erreichte die Teilzeitquote mit 14,6 % bei alleinerziehenden Männern.

Die Unterschiede zwischen Vätern und Männern ohne Kind bezüglich der Arbeitszeitzufriedenheit sind nicht einheitlich. So zeigten Männer in einer Lebensgemeinschaft mit Kindern eine höhere TZ-Diskrepanzenquote (22,0 %) als diejenigen ohne Kind (20,1 %), während alleinerziehende Väter in Teilzeit mit ihrer Arbeitszeit zufriedener waren als alleinstehende Männer in Teilzeitbeschäftigung (TZ-Diskrepanzenquote 18,1 % zu 21,5 %). Mindestens

Tabelle 6

Realisierte Erwerbstätigkeit und Arbeitszeitwünsche in Quoten nach Lebensformen<sup>1</sup> der 25- bis 54-Jährigen<sup>2</sup>

|                                                             | Ehepaar     |           | Lebensgemeinschaft |           | Alleinerziehend<br>mit Kindern | Alleinstehende |
|-------------------------------------------------------------|-------------|-----------|--------------------|-----------|--------------------------------|----------------|
|                                                             | mit Kindern | ohne Kind | mit Kindern        | ohne Kind |                                |                |
|                                                             | %           |           |                    |           |                                |                |
|                                                             | Männer      |           |                    |           |                                |                |
| Realisierte Erwerbstätigenquote                             | 91,1        | 93,2      | 90,8               | 93,8      | 78,6                           | 84,4           |
| Teilzeitquote                                               | 8,0         | 6,5       | 9,6                | 9,1       | 14,6                           | 12,7           |
| Unterbeschäftigtenquote<br>im weiteren Sinne                | 2,5         | 2,6       | 2,9                | 3,3       | 3,1                            | 4,2            |
| Vollzeittätige                                              | 1,4         | 1,8       | 1,5                | 2,1       | 1,2                            | 2,4            |
| Teilzeittätige <sup>3</sup>                                 | 14,9        | 14,2      | 15,6               | 15,4      | 14,3                           | 17,0           |
| Überbeschäftigtenquote                                      | 10,3        | 12,2      | 11,3               | 14,5      | 9,6                            | 12,6           |
| Vollzeittätige                                              | 10,8        | 12,7      | 11,9               | 15,5      | 10,6                           | 13,8           |
| Teilzeittätige <sup>3</sup>                                 | 4,4         | 5,5       | 6,5                | 4,7       | 3,7                            | 4,5            |
| Diskrepanzenquote                                           | 12,8        | 14,8      | 14,2               | 17,8      | 12,7                           | 16,8           |
| Anteil der Unterbeschäftigung<br>im weiteren Sinne          | 19,6        | 17,7      | 20,2               | 18,3      | 24,4                           | 25,2           |
| Anteil der Überbeschäftigung                                | 80,4        | 82,3      | 79,8               | 81,7      | 75,6                           | 74,8           |
| Vollzeit-Diskrepanzenquote                                  | 12,2        | 14,5      | 13,4               | 17,6      | 11,8                           | 16,1           |
| Anteil der Vollzeit-Unterbeschäftigung<br>im weiteren Sinne | 11,7        | 12,5      | 11,5               | 11,7      | 10,1                           | 14,7           |
| Anteil der Vollzeit-Überbeschäftigung                       | 88,3        | 87,5      | 88,5               | 88,3      | 89,9                           | 85,3           |
| Teilzeit-Diskrepanzenquote                                  | 19,3        | 19,7      | 22,0               | 20,1      | 18,1                           | 21,5           |
| Anteil der Teilzeit-Unterbeschäftigung<br>im weiteren Sinne | 77,2        | 72,3      | 70,6               | 76,6      | 79,3                           | 79,2           |
| Anteil der Teilzeit-Überbeschäftigung                       | 22,8        | 27,7      | 29,4               | 23,4      | 20,7                           | 20,8           |
|                                                             | Frauen      |           |                    |           |                                |                |
| Realisierte Erwerbstätigenquote                             | 73,6        | 83,6      | 74,7               | 92,1      | 73,6                           | 84,8           |
| Teilzeitquote                                               | 67,3        | 33,5      | 59,3               | 21,2      | 52,3                           | 26,3           |
| Unterbeschäftigtenquote<br>im weiteren Sinne                | 5,5         | 2,7       | 5,8                | 3,3       | 7,4                            | 4,2            |
| Vollzeittätige                                              | 1,1         | 0,7       | 1,0                | 1,0       | 1,6                            | 1,3            |
| Teilzeittätige <sup>3</sup>                                 | 7,6         | 6,7       | 9,1                | 11,6      | 12,8                           | 12,4           |
| Überbeschäftigtenquote                                      | 6,2         | 11,6      | 8,4                | 16,2      | 8,2                            | 14,9           |
| Vollzeittätige                                              | 11,4        | 15,5      | 14,1               | 19,0      | 12,9                           | 18,3           |
| Teilzeittätige <sup>3</sup>                                 | 3,7         | 3,9       | 4,5                | 5,8       | 3,9                            | 5,3            |
| Diskrepanzenquote                                           | 11,6        | 14,4      | 14,2               | 19,5      | 15,7                           | 19,1           |
| Anteil der Unterbeschäftigung<br>im weiteren Sinne          | 46,8        | 19,0      | 41,0               | 16,8      | 47,5                           | 22,2           |
| Anteil der Überbeschäftigung                                | 53,2        | 81,0      | 59,0               | 83,2      | 52,5                           | 77,8           |
| Vollzeit-Diskrepanzenquote                                  | 12,5        | 16,3      | 15,1               | 20,0      | 14,5                           | 19,6           |
| Anteil der Vollzeit-Unterbeschäftigung<br>im weiteren Sinne | 8,6         | 4,6       | 6,8                | 5,2       | 11,1                           | 6,7            |
| Anteil der Vollzeit-Überbeschäftigung                       | 91,4        | 95,4      | 93,2               | 94,8      | 88,9                           | 93,3           |
| Teilzeit-Diskrepanzenquote                                  | 0,1         | 0,1       | 0,1                | 0,2       | 0,2                            | 0,2            |
| Anteil der Teilzeit-Unterbeschäftigung<br>im weiteren Sinne | 67,4        | 63,2      | 67,0               | 66,8      | 76,4                           | 69,9           |
| Anteil der Teilzeit-Überbeschäftigung                       | 32,6        | 36,8      | 33,0               | 33,2      | 23,6                           | 30,1           |

1 Personen, die als gleichgeschlechtliches Ehepaar/in einer gleichgeschlechtlichen Lebensgemeinschaft mit Kindern oder ohne Kind leben, werden wegen zu geringer Fallzahlen hier nicht abgebildet.

2 Personen in Privathaushalten.

3 Die Angabe zur Teilzeittätigkeit bezieht sich auf die Selbsteinschätzung der Befragten.

Erstergebnisse des Mikrozensus 2024.

## Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung

70 % der mit ihrer Teilzeitarbeitszeit unzufriedenen Männer wollten ihre Arbeitszeit erhöhen.

Auch die Teilzeitquote in Höhe von 48,6 % der 25- bis 54-jährigen Frauen setzt sich je nach Lebensform unterschiedlich zusammen. Sie war ebenfalls in allen Lebensformen bei Müttern deutlich höher als bei Frauen ohne Kind. In der Lebensform Ehepaar mit Kindern erreichte sie mit 67,3 % ihren Höchststand und nahm bei Frauen in der Lebensgemeinschaft ohne Kind mit 21,2 % ein Minimum an.

Im Unterschied zu den teilzeittätigen 25- bis 54-jährigen Männern zeigten teilzeittätige Frauen dieser Altersklasse nur eine verschwindend geringe TZ-Discrepanzenquote von 0,1 % bis 0,2 %.

### Hauptgrund für Teilzeit

Einen tieferen Einblick hinsichtlich der Unterschiede zwischen Männern und Frauen bei einer Teilzeittätigkeit ermöglicht [Tabelle 7](#), die den Hauptgrund für eine Teilzeittätigkeit nach Geschlecht differenziert darstellt. Weiterhin wird hier nur die Altersklasse der 25- bis 54-Jährigen betrachtet. „Ich möchte Teilzeit arbeiten.“ nannten 20,9 % der teilzeittätigen Männer und 22,8 % der Frauen als Hauptgrund. Bei Männern spielte auch Schule, Ausbildung/Studium mit 20,6 % eine wichtige Rolle als Grund für Teilzeit, nur 5,2 % der Frauen führten diese Gründe an. Ein weiterer deutlicher geschlechtsspezifischer Unterschied zeigt sich beim Teilzeitgrund Betreuung<sup>12</sup>: Bei den Männern gaben hier 13,1 % Betreuung als Hauptgrund für ihre Teilzeittätigkeit an, während dies bei den teilzeittätigen Frauen 43,3 % waren.

Tabelle 7 ist weiter zu entnehmen, warum Betreuung als Teilzeitgrund angegeben wurde. Wurden die Betreuungsaufgaben freiwillig ausgeübt oder waren geeignete Betreuungsangebote nicht vorhanden beziehungsweise nicht bezahlbar? Von den fast 2,6 Millionen teilzeittätigen Frauen mit Hauptgrund Betreuung gaben 66,2 % an, dies freiwillig zu tun. Von den 182 000 teilzeittätigen Männern mit Teilzeitgrund Betreuung traf dies mit 59,9 % auf anteilig

12 Beim Teilzeitgrund „Betreuung“ werden Betreuungsaufgaben von Kindern zusammen mit denjenigen von Menschen mit Behinderung oder pflegebedürftigen Personen erfasst.

**Tabelle 7**

**Realisierte Erwerbstätigkeit nach Hauptgrund für Teilzeit und Grund für Betreuung**

|                                                      | 25- bis 54-jährige |      |        |      |
|------------------------------------------------------|--------------------|------|--------|------|
|                                                      | Männer             |      | Frauen |      |
|                                                      | 1 000              | %    | 1 000  | %    |
| Hauptgrund für Teilzeit                              |                    |      |        |      |
| Schule, Ausbildung/Studium                           | 287                | 20,6 | 310    | 5,2  |
| Krankheit, verminderte Erwerbsfähigkeit, Behinderung | 114                | 8,2  | 190    | 3,2  |
| Betreuung <sup>1</sup>                               | 182                | 13,1 | 2 566  | 43,3 |
| Andere familiäre Gründe                              | 37                 | 2,7  | 323    | 5,4  |
| Andere persönliche Gründe                            | 127                | 9,1  | 420    | 7,1  |
| Vollzeittätigkeit nicht zu finden                    | 140                | 10,1 | 240    | 4,0  |
| Möchte Teilzeit arbeiten                             | 290                | 20,9 | 1 351  | 22,8 |
| Anderer Hauptgrund                                   | 213                | 15,3 | 528    | 8,9  |
| Grund für Betreuung <sup>1</sup>                     |                    |      |        |      |
| Insgesamt                                            | 182                | 100  | 2 566  | 100  |
| darunter:                                            |                    |      |        |      |
| Geeignete Betreuungsangebote nicht vorhanden/nutzbar | 35                 | 19,3 | 369    | 14,4 |
| Geeignetes Betreuungsangebot nicht bezahlbar         | 7                  | 4,0  | 125    | 4,9  |
| Freiwillige Betreuung                                | 109                | 59,9 | 1 698  | 66,2 |
| Andere Gründe                                        | 31                 | 16,8 | 374    | 14,6 |

1 Betreuung von Kindern, Menschen mit Behinderung oder pflegebedürftigen Personen.

Ersterggebnisse des Mikrozensus 2024.

weniger, aber immer noch deutlich mehr als die Hälfte zu.

### Gründe für Nichtverfügbarkeit bei Unterbeschäftigung

Bisher wurde nur die Unterbeschäftigung i.w.S. betrachtet, weil sie als unmittelbares Gegenstück zur Überbeschäftigung zu sehen ist. Diese Gruppe wurde aber bereits in den Tabellen 2, 3 und 5 in die beiden Subgruppen „Unterbeschäftigte nicht verfügbar“ und „Unterbeschäftigte verfügbar“ unterteilt.

Betrachtet man die absoluten Zahlen der jeweiligen Gruppen von Unterbeschäftigten in Tabelle 2, dann zeigt sich, dass zwischen fast 40 % bis hin zu knapp 55 % der Unterbeschäftigten i.w.S. je nach Geschlecht und Altersklasse für zusätzliche Arbeitszeitstunden in den nächsten zwei Wochen nicht zur Verfügung stehen konnten. Warum ist das so? Zunächst wäre zu vermuten, dass die gewöhnliche Wochenarbeitszeit bei nicht verfügbaren Unterbeschäftigten höher



sein könnte als diejenige von verfügbaren Unterbeschäftigten. Das Gegenteil ist jedoch der Fall: Bis auf wenige Ausnahmen (beispielsweise bei den 15- bis 24-jährigen Frauen) war die gewöhnliche Wochenarbeitszeit bei nicht verfügbaren Unterbeschäftigten geringer als bei verfügbaren Unterbeschäftigten. Die Frage, warum Personen mit Mehrarbeitswunsch das Verfügbarkeitskriterium nicht erfüllen, ist damit nicht geklärt.

In der Arbeitskräfteerhebung (Labour Force Survey – LFS), einer Teilstichprobe des Mikrozensus, werden Befragte deshalb explizit nach den Gründen für ihre Nichtverfügbarkeit trotz Mehrarbeitswunsch gefragt. [Tabelle 8](#) zeigt, welche Gründe 25- bis 54-jährige unterbeschäftigte Männer und unterbeschäftigte Frauen angegeben haben. Bei den Männern dieser Altersklasse gaben 51,1% einen „anderen Hauptgrund“ an, gefolgt in der relativen Häufigkeit von 13,1%, die „andere persönliche Verpflichtungen“

**Tabelle 8**  
Realisierte Erwerbstätigkeit: Gründe für Nichtverfügbarkeit bei Unterbeschäftigten und Nichtverfügbarkeitsgrund Betreuung

|                                                      | 25- bis 54-jährige |        |
|------------------------------------------------------|--------------------|--------|
|                                                      | Männer             | Frauen |
| %                                                    |                    |        |
| Hauptgrund für Nichtverfügbarkeit                    |                    |        |
| Schulische oder berufliche Ausbildung/ Studium       | 11,2               | 6,6    |
| Krankheit, verminderte Erwerbsfähigkeit, Behinderung | 6,2                | 5,1    |
| Betreuung <sup>1</sup>                               | 9,2                | 35,3   |
| Anderer familiäre Verpflichtungen                    | 5,0                | 4,0    |
| Anderer persönliche Verpflichtungen                  | 13,1               | 9,2    |
| Anderer Hauptgrund                                   | 51,1               | 37,0   |
| Kündigungsfrist bestehender Job                      | 4,2                | 2,8    |
| Grund für Betreuung <sup>1</sup>                     |                    |        |
| Insgesamt                                            | 100                | 100    |
| Geeignete Betreuungsangebote nicht vorhanden/nutzbar | 37,9               | 44,0   |
| Geeignetes Betreuungsangebot nicht bezahlbar         | 8,4                | 9,1    |
| Freiwillige Betreuung                                | 33,9               | 38,5   |
| Anderer Gründe                                       | 19,8               | 8,3    |

1 Betreuung von Kindern, Menschen mit Behinderung oder pflegebedürftigen Personen.

Erstergebnisse der Arbeitskräfteerhebung 2024.

nannten. „Betreuung“<sup>13</sup> als Grund für Nichtverfügbarkeit gaben 9,2% der nicht verfügbaren unterbeschäftigten Männer an. Bei den 25- bis 54-jährigen Frauen sieht die Verteilung anders aus: Auch hier nannten 37,0% den anderen Hauptgrund, aber 35,3% gaben Betreuungsaufgaben an.

Analog zu den Gründen für Teilzeittätigkeit wurden auch hier diejenigen, die den Grund „Betreuung“ angaben, gefragt, ob es sich um freiwillige oder unfreiwillige Betreuungsaufgaben handelte. Nicht verfügbare Unterbeschäftigte mit Betreuung als Grund für Nichtverfügbarkeit machten dies in 33,9% (Männer) beziehungsweise 38,5% (Frauen) der Fälle freiwillig. Dass ein geeignetes Betreuungsangebot nicht vorhanden oder nutzbar war, gaben 37,9% der Männer und 44,0% der Frauen an.<sup>14</sup>

4

Weitere arbeitszeitbezogene Aspekte

Geschlechtsspezifische Unterschiede gibt es außer bei den gewöhnlich für bezahlte Arbeit aufgewendeten Stunden hinsichtlich weiterer Aspekte mit Bezug zu Arbeitszeit und Erwerbstätigkeit. Hierzu zählen neben den für die Fahrt zur Arbeitsstätte aufgewendeten Zeiten insbesondere die Arbeit zu ungewöhnlichen Zeiten und die Schichtarbeit. Letztere können soziale und gesundheitliche Einschränkungen mit sich bringen, die auch bei gleicher Arbeitszeit berücksichtigt werden sollten.

Männer hatten nicht nur längere Arbeitszeiten als Frauen, sondern nahmen auch längere Wege zur Arbeit in Kauf. Allerdings fallen die geschlechtsspezifischen Unterschiede etwas geringer aus als bei den Arbeitszeiten. Im Jahr 2024 waren 31,2% der Männer im Alter von 15 bis 79 Jahren 30 Minuten oder länger zur Arbeit unterwegs, bei den Frauen lag dieser Wert mit 24,8% knapp 7 Prozentpunkte

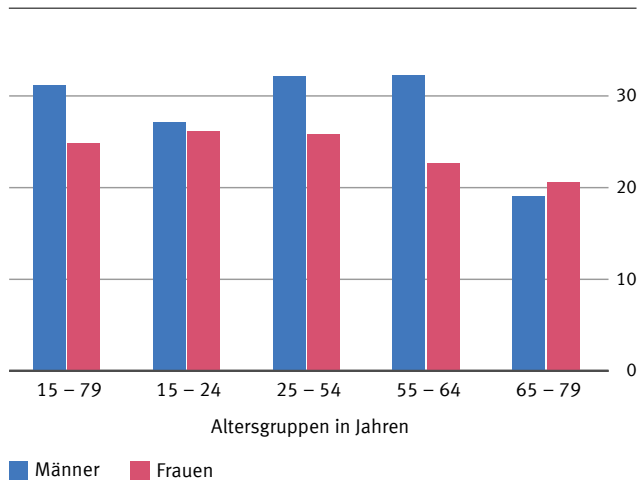
13 Der Grund „Betreuung“ bei den Hauptgründen für Nichtverfügbarkeit umfasst die Betreuung von Kindern, Menschen mit Behinderung oder pflegebedürftigen Personen.

14 Anders als beim Teilzeitgrund Betreuung ist der Anteil der freiwilligen Betreuung hier deutlich geringer. Das liegt auch daran, dass sowohl vollzeit- als auch teilzeittätige Personen mit Mehrarbeitswunsch betrachtet wurden.

## Arbeitszeiten und Arbeitszeitwünsche von Männern (und Frauen) – Anhaltspunkte zu Potenzialen für die Gleichstellung

**Grafik 2**

Erwerbstätige, die 30 Minuten oder länger für den einfachen Weg zur Arbeit benötigen, nach Altersgruppen in %



Erstergebnisse des Mikrozensus 2024. – Erwerbstätige mit Angabe zum Zeitaufwand; ohne Erwerbstätige mit ständig wechselnder Arbeitsstätte.

darunter. [↗ Grafik 2](#) Ähnlich wie bei der Arbeitszeit weichen vor allem bei den 25- bis 54-Jährigen und bei den 55- bis 64-Jährigen die Wegezeiten von Männern und Frauen voneinander ab. Auch die Zeitverwendungserhebung 2022 bestätigt, dass Männer etwas länger zur Arbeitsstätte fahren. In dieser Quelle betrug der Unterschied – bezogen auf die Gruppe der Erwerbstätigen – 8 Minuten je Tag für Hin- und Rückweg (Männer: 56 Minuten; Frauen: 48 Minuten; Statistisches Bundesamt, 2025d).

Der Anteil von Erwerbstätigen, die häufig zu **ungewöhnlichen Zeiten arbeiten**, fällt für Männer und Frauen ebenfalls unterschiedlich aus. Frauen arbeiteten dabei 2023 sogar geringfügig häufiger an

Samstagen und Sonntagen als Männer.<sup>15</sup> Die entsprechenden Anteile lagen – bezogen auf die 15- bis 79-Jährigen – bei Frauen um 0,6 beziehungsweise 0,2 Prozentpunkte höher als bei Männern.<sup>16</sup> Diese Unterschiede verantwortet vor allem die Gruppe der 15- bis 24-jährigen Frauen, dagegen arbeiteten Männer in der Haupterwerbs- und Familienphase etwas häufiger am Wochenende. [↗ Tabelle 9](#)

Abends und nachts arbeiteten Männer dagegen häufiger als Frauen. Der Anteil der Männer, die in den vier Wochen vor dem Erhebungsstichtag an mindestens jedem zweiten Arbeitstag abends oder nachts arbeiteten, lag bei der Abendarbeit um 3,4 Prozentpunkte und bei der Nachtarbeit um 2,6 Prozentpunkte über dem der Frauen. Auch hier fallen die Differenzen in der Haupterwerbs- und Familienphase am größten aus (Abendarbeit: 4,6 Prozentpunkte; Nachtarbeit: 3,1 Prozentpunkte) und waren bei den 15- bis 24-Jährigen am geringsten. Ein sehr ähnliches Bild zeigt sich schließlich auch bei der Verteilung von Schichtarbeit, die häufig mit Abend- oder Nachtarbeit einhergehen dürfte. Diese Unterschiede dürften zum einen das Ergebnis eines unterschiedlichen Verhaltens von Männern und Frauen in Bezug auf die Erwerbstätigkeit in der Familienphase sein, zum anderen aber auch mit der geschlechtsspezifischen Berufswahl zusammenhängen.

15 Da die Variablen zu ungewöhnlichen Arbeitszeiten in der Arbeitskräfteerhebung in zweijährlicher Periodizität erhoben werden, beziehen sich die Ergebnisse hier auf das Jahr 2023.

16 „Häufig“ bedeutet in diesem Zusammenhang in der Arbeitskräfteerhebung an mindestens zwei Samstagen beziehungsweise Sonntagen in den letzten vier Wochen vor dem Berichtsstichtag.

**Tabelle 9**

Erwerbstätige, die an mindestens der Hälfte der Tage zu ungewöhnlichen Zeiten arbeiten, nach Altersgruppen

|                | Männer          |                 |                 |                 |                 | Frauen          |                 |                 |                 |                 |
|----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
|                | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre | 15 bis 79 Jahre | 15 bis 24 Jahre | 25 bis 54 Jahre | 55 bis 64 Jahre | 65 bis 79 Jahre |
| %              |                 |                 |                 |                 |                 |                 |                 |                 |                 |                 |
| Abendarbeit    | 16,7            | 15,4            | 17,6            | 15,2            | 13,2            | 13,3            | 16,8            | 13,0            | 12,9            | 10,8            |
| Nachtarbeit    | 5,8             | 4,6             | 6,4             | 5,3             | (3,2)           | 3,2             | 3,4             | 3,3             | 2,9             | /               |
| Samstagsarbeit | 17,4            | 18,4            | 17,6            | 16,3            | 19,6            | 18,0            | 21,4            | 17,2            | 18,9            | 17,2            |
| Sonntagsarbeit | 9,2             | 8,1             | 9,4             | 8,9             | 9,7             | 9,4             | 10,3            | 9,2             | 9,4             | 9,1             |
| Schichtarbeit  | 16,1            | 16,8            | 17,3            | 13,6            | (4,1)           | 13,0            | 17,5            | 12,8            | 12,3            | (5,9)           |

Endergebnisse 2023 der Arbeitskräfteerhebung.

## 5

### Fazit

Erwerbsbeteiligung und Arbeitszeit von Männern und Frauen unterscheiden sich nach wie vor deutlich. Dabei erweist sich eine Arbeitsteilung entlang traditioneller Geschlechterrollen im Zeitablauf als sehr stabil (Wanger/Weber, 2023). Ein männlicher Familiernährer mit oder ohne weiblichen Zuverdienst ist weiterhin bei den meisten Paaren die Regel, eine gleichmäßige Aufteilung von Erwerbs- und Sorgearbeit die Ausnahme (Gambaro und andere, 2024). Hierzu passt, dass die Unterschiede durchweg bei Lebensformen mit Kindern deutlich größer sind als bei Lebensformen ohne Kind.

Die zentralen Ergebnisse aus den Analysen der Daten des Mikrozensus 2024 bezüglich Arbeitszeit und Arbeitszeitwünschen sind:

#### > Verhältnis Vollzeit zu Teilzeit

Bei der hier betrachteten realisierten Erwerbstätigkeit betrug die Vollzeitquote der Männer 86,2 % im Vergleich zu 50,0 % bei den Frauen. Analog dazu ergab sich ein Verhältnis der Teilzeitquoten von 13,8 % zu 50,0 %. Die Betrachtung nach Altersklassen zeigt, dass in der sogenannten Haupterwerbs- und Familienphase die Teilzeitquote der 25- bis 54-jährigen Männer mit 9,7 % ein Minimum annahm.

#### > Vollzeit

Vollzeittätige Männer hatten mit 41,1 Stunden eine um 1,5 Stunden höhere durchschnittliche Wochenarbeitszeit als vollzeittätige Frauen. Sie hatten relativ häufiger einen Wunsch nach Arbeitszeiterhöhung als Frauen (1,5 % zu 1,0 %) und wollten mit 11,5 % ihre Arbeitszeit seltener reduzieren als Frauen (14,0 %). Der Wunsch nach einer Arbeitszeiterhöhung erreichte bei vollzeittätigen Männern und Frauen in der Haupterwerbs- und Familienphase einerseits mit 1,8 % und 1,1 % ein Maximum, andererseits wies diese Altersgruppe gleichzeitig mit die höchsten Werte beim Wunsch nach Arbeitszeitreduzierung auf. Damit waren die vollzeittätigen 25- bis 54-jährigen Männer (14,1 %) und Frauen (16,6 %) am wenigsten zufrieden mit ihrer Arbeitszeit.

#### > Teilzeit

Männer in Teilzeit hatten im Jahr 2024 mit 19,6 Stunden eine niedrigere gewöhnliche Wochenarbeitszeit als teilzeittätige Frauen (21,9 Stunden). Teilzeittätige Männer äußerten dabei mit 10,1 % häufiger den Wunsch nach einer Arbeitszeiterhöhung als teilzeittätige Frauen (7,0 %). Die gewünschte Stundenerhöhung lag bei durchschnittlich + 15,4 Stunden (Frauen: + 10,4 Stunden). Allerdings gaben auch 4,5 % der teilzeittätigen Männer den Wunsch nach einer Arbeitszeitreduzierung an (Frauen: 4,0 %). Durchschnittlich wollten sie 8,7 Stunden je Woche weniger arbeiten (Frauen: 8,3 Stunden). Insgesamt waren 14,6 % der teilzeittätigen Männer mit ihrer Arbeitszeit unzufrieden (Frauen: 11,0 %). In der Altersklasse der 25- bis 54-Jährigen erreichte die Unzufriedenheit mit der Arbeitszeit bei beiden Geschlechtern ein Maximum: Die TZ-Diskrepanzenquote von Männern dieser Altersklasse lag bei 20,1 % (Frauen: 13,0 %), wobei jeweils wiederum der Großteil von ihnen die Arbeitszeit erhöhen wollte.

#### > Haupterwerbs- und Familienphase

Eine Auswertung nach Lebensformen zeigt, dass in der Altersklasse der 25- bis 54-jährigen die Teilzeitquote bei Vätern generell etwas höher war als bei Männern ohne Kind. Väter, die beispielsweise in der Lebensform Ehepaar mit Kindern lebten, hatten eine Teilzeitquote von 8,0 %, Männer ohne Kind in dieser Lebensform eine von 6,5 %. Auch Mütter dieser Altersklasse hatten in allen Lebensformen eine deutlich höhere Teilzeitquote (67,3 %) als Frauen ohne Kind (33,5 %).

Insgesamt nannten bei den teilzeittätigen Männern dieser Altersklasse 13,1 % „Betreuung“ als Hauptgrund für ihre Teilzeittätigkeit (Frauen: 43,3 %); 59,9 % davon übten die Betreuungsaufgaben freiwillig aus (Frauen: 66,2 %). Bei den unterbeschäftigten teil- oder vollzeittätigen Männern mit fehlender kurzfristiger Verfügbarkeit für Mehrarbeitsstunden gaben 9,2 % „Betreuung“ als Hauptgrund für ihre Nichtverfügbarkeit an (Frauen: 35,3 %). Davon gaben 33,9 % an, dass die Betreuung freiwillig sei (Frauen: 38,5 %).

Vor dem Hintergrund des insgesamt niedrigen Anteils von Erwerbstätigen, die ihre Arbeitszeit ändern wollen, ist der Hinweis wichtig, dass der Mikrozensus vor allem Arbeitszeitwünsche innerhalb der gegebenen ökonomischen und institutionellen Rahmenbedingungen erfasst. Er fragt dagegen nicht nach Idealvorstellungen oder Wünschen, die Änderungen der Rahmenbedingungen voraussetzen. Studien, die beispielsweise keinen Hinweis auf einen entsprechend der Wunscharbeitszeit veränderten Verdienst aufweisen (zum Beispiel DGB, 2025), kommen zu dem Ergebnis, dass sowohl Männer als auch Frauen ihre Arbeitszeit deutlich stärker reduzieren wollen als nach den Ergebnissen des Mikrozensus. Nach einer Erhebung im Kontext des DGB-Index Gute Arbeit würden Männer durchschnittlich 35 Stunden je Woche arbeiten, Frauen 28,7 Stunden. Auch hier bleiben – auf einem niedrigeren Niveau – die traditionellen Rollenbilder erhalten, nach denen Männer länger arbeiten. Einen ähnlichen Befund zeigt das familiendemografische Panel ([FReDA](#)), das nach der idealen Arbeitszeit von Müttern und Vätern fragt. Für beide Gruppen sehen die Befragten eine niedrigere als die gewöhnliche Erwerbsarbeitszeit als ideal an. Väter sollen nach der Einschätzung der Befragten – bei geringen Unterschieden zwischen Männern und Frauen – aber je nach Alter des jüngsten Kindes idealerweise zwischen 2 und 13 Stunden je Woche mehr Erwerbsarbeit leisten als Mütter (Bujard/Kleinschrot, 2024). Hierbei kommt zum Ausdruck, dass etwa die eine Hälfte der Bevölkerung eine gleichberechtigte Aufteilung von Erwerbs- und Sorgearbeit als Ideal ansieht, während die andere Hälfte Varianten des männlichen Ernährermodells befürwortet (Gambaro und andere, 2024; siehe auch Juncke und andere, 2023; Wippermann, 2023; Trübner, 2020). 

## LITERATURVERZEICHNIS

---

Bächmann, Ann-Christin/Kleinert, Corinna/Schels, Brigitte. *Anhaltende berufliche Geschlechtersegregation: In Ost wie West arbeiten Frauen und Männer häufig in unterschiedlichen Berufen*. IAB-Kurzbericht 03/2024. Nürnberg 2024. DOI: [10.48720/IAB.KB.2403](https://doi.org/10.48720/IAB.KB.2403)

Bujard, Martin/Kleinschrot, Leonie. *Wie viel sollten Mütter und Väter arbeiten?* Bevölkerungsforschung aktuell. Ausgabe 1/2024. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.bib.bund.de](http://www.bib.bund.de)

BAuA (Bundesanstalt für Arbeitsschutz und Arbeitsmedizin). *Arbeitszeitreport Deutschland: Ergebnisse der BAuA-Arbeitszeitbefragung 2021*. Dortmund 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.baua.de](http://www.baua.de)

BMBFSFJ (Bundesministerium für Bildung, Familie, Senioren, Frauen und Jugend). *Digitaler Gleichstellungsatlas. Themenfeld Bildung, Berufswahl und wissenschaftliche Qualifizierung*. 2025: Berichtsjahr 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.bmbfsfj.bund.de](http://www.bmbfsfj.bund.de)

DGB (Institut DGB-Index Gute Arbeit). *Wöchentliche Arbeitszeiten. Zwischen Wunsch und Wirklichkeit*. DGB-Index Gute Arbeit Kompakt 05/2025. [Zugriff am 14. Januar 2026]. Verfügbar unter: [index-gute-arbeit.dgb.de](http://index-gute-arbeit.dgb.de)

DIW Berlin. *Auf der Suche nach der idealen Arbeitszeit*. Forschungsgruppe Gender Economics 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.diw.de](http://www.diw.de)

Eurofound (European Foundation for the Improvement of Living and Working Conditions). *European Working Conditions Survey 2024: First findings*. Luxemburg 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.eurofound.europa.eu](http://www.eurofound.europa.eu)

Gambaro, Ludovica/Gehlen, Annica/Spieß, C. Katharina/Wrohlich, Katharina/Ziege, Elena. *Aufteilung von Erwerbs- und Sorgearbeit bei Eltern: Wunsch und Wirklichkeit liegen teils weit auseinander*. In: DIW Wochenbericht 29/2024, Seite 459 ff. [Zugriff am 13. Januar 2026]. Verfügbar unter [www.diw.de](http://www.diw.de)

Gesterkamp, Thomas. *Jenseits von Feminismus und Antifeminismus*. Plädoyer für eine eigenständige Männerpolitik. Wiesbaden 2014.

Gärtner, Marc. *Care-Arbeit, Gleichstellung und der Blick auf Männer*. Berlin 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.bundesstiftung-gleichstellung.de](http://www.bundesstiftung-gleichstellung.de)

Goldin, Claudia. *Career and family. Women's century-long journey toward equity*. Princeton 2021.

Häring, Armando/Schütz, Holger/Thiele, Niels/Kleudgen, Martin/Nold, Johanna/Entgelmeier, Ines/Backhaus, Nils. *Methodenbericht zur BAuA-Arbeitszeitbefragung 2023*. Dortmund/Berlin/Dresden 2024. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.baua.de](http://www.baua.de)

## LITERATURVERZEICHNIS

---

Heesen, Boris von. *Was Männer kosten. Der hohe Preis des Patriarchats*. München 2022.

Hochgürtel, Tim. [Realisierte Erwerbstätigkeit zur Messung des Vereinbarkeitsarrangements von Familie und Beruf](#). In: WISTA Wirtschaft und Statistik. Ausgabe 1/2018, Seite 23 ff.

ILO (International Labour Organization). *Resolution to amend the 19th ICLS resolution concerning statistics of work, employment and labour underutilization*. Genf 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.ilo.org](http://www.ilo.org)

ILO (International Labour Organization). *Resolution concerning the measurement of working time*. Genf 2008. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.ilo.org](http://www.ilo.org)

Juncke, David/Samtleben, Claire/Stoll, Evelyn. *Väterreport 2023. Entwicklungen und Daten zur Vielfalt der Väter in Deutschland*. Bundesministerium für Familien, Senioren, Frauen und Jugend (Herausgeber). Berlin 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.bmbfsfj.bund.de](http://www.bmbfsfj.bund.de)

Keller, Matthias/Körner, Thomas. [Closing the gap? Erwerbstätigkeit und Arbeitszeit von Müttern und Vätern nach 15 Jahren Elterngeld](#). In: WISTA Wirtschaft und Statistik. Ausgabe 4/2023, Seite 88 ff.

Kleinert, Corinna/Leuze, Kathrin/Bächmann, Ann-Christin/Gatermann, Dörthe/Hägglund, Anna Erika/Rompczyk, Kai. *Occupational Sex Segregation and its Consequences for the (Re-)Production of Gender Inequalities in the German Labour Market*. In: Weinert, Sabine/Blossfeld, Gwendolin Josephine/Blossfeld, Hans-Peter (Herausgeber). *Education, Competence Development and Career Trajectories. Methodology of Educational Measurement and Assessment*. Cham 2023, Seite 295 ff. DOI: [10.1007/978-3-031-27007-9\\_13](https://doi.org/10.1007/978-3-031-27007-9_13)

Kümmerling, Angelika. *Arbeitszeiten zwischen Wunsch, Befürchtungen und Wirklichkeit: Ein IAQ-Arbeitszeitmonitor*. IAQ-Report 2024 | 05. DOI: [10.17185/dupublico/81957](https://doi.org/10.17185/dupublico/81957)

Mischler, Frauke. [Verdienstunterschiede zwischen Frauen und Männern. Eine Ursachenanalyse auf Grundlage der Verdienststrukturerhebung 2018](#). In: WISTA Wirtschaft und Statistik. Ausgabe 4/2021, Seite 110 ff.

OECD (Organisation für wirtschaftliche Zusammenarbeit und Entwicklung). *Pensions at a Glance 2025: OECD and G20 Indicators*. Paris 2025. DOI: [10.1787/e40274c1-en](https://doi.org/10.1787/e40274c1-en)

Rengers, Martina/Bringmann, Julia/Holst, Elke. [Arbeitszeiten und Arbeitszeitwünsche: Unterschiede zwischen Mikrozensus und SOEP](#). In WISTA Wirtschaft und Statistik. Ausgabe 4/2017, Seite 11 ff.



## LITERATURVERZEICHNIS

---

Rengers, Martina/Fuchs, Johann. *Stille Reserve in Deutschland: Gemeinsamkeiten und Unterschiede zweier Konzepte. Ergebnisse für das Jahr 2019*. In: AStA Wirtschafts- und Sozialstatistisches Archiv. Jahrgang 16. Ausgabe 3-4/2022, Seite 189 ff. DOI: [10.1007/s11943-022-00309-3](https://doi.org/10.1007/s11943-022-00309-3)

Statistisches Bundesamt. [Gleichstellungsindikatoren](#). 2025a.

Statistisches Bundesamt. [Qualität der Arbeit. Themenfeld Arbeitssicherheit und Gleichstellung](#). 2025b.

Statistisches Bundesamt. [Mikrozensus 2024](#). Qualitätsbericht. 2025c.

Statistisches Bundesamt. [Durchschnittliche Zeitverwendung von Personen ab 10 Jahren für Erwerbstätigkeit, Arbeitssuche nach Geschlecht](#). 2025d.

Trübner, Miriam. *Arbeitsteilung in Paarhaushalten. Eine dyadische Untersuchung partnerschaftlicher Aufgabenverteilung*. Bonner Reihe der Empirischen Sozialforschung, Band 3. Opladen 2020.

Volz, Rainer. *Arbeit ist sein (ganzes) Leben? – Männer und Erwerbsarbeit*.

In: Theunert, Markus (Herausgeber). *Männerpolitik*. Wiesbaden 2012.

DOI: [10.1007/978-3-531-19041-9\\_10](https://doi.org/10.1007/978-3-531-19041-9_10)

Wanger, Susanne/Weber, Enzo. *Arbeitszeit: Trends, Wunsch und Wirklichkeit*. IAB-Forschungsbericht 16|2023. Nürnberg 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [doku.iab.de](https://doku.iab.de)

Wippermann, Carsten. *Männerperspektiven: Einstellungen von Männern zu Gleichstellung und Gleichstellungspolitik*. Sozialwissenschaftliche Repräsentativbefragung im Auftrag des Bundesforum Männer – Interessenverband für Jungen, Männer und Väter e.V. Penzberg 2023. [Zugriff am 13. Januar 2026]. Verfügbar unter: [maennerperspektiven.de](https://maennerperspektiven.de)

# PARTIELLE SYNTHETISIERUNG ZUR ANONYMISIERUNG DER VERKNÜPFTEN VERDIENST- STRUKTURERHEBUNG 2018

## Teil 2: Evaluation der Anonymität und Ausblick zur Datenbereitstellung

Hanna Brenzel, Yannik Garcia Ritz, David Wagner

📌 **Schlüsselwörter:** partiell synthetische Daten – verknüpfte synthetische Daten – Risikokennzahlen – Aufdeckungsszenarios – Anonymitätsverlust – ALC

### ZUSAMMENFASSUNG

Im ersten Teil des Aufsatzes (Brenzel/Garcia Ritz, 2025) wurden die Methodik und Evaluation der Datennützlichkeit dargestellt. Der zweite Teil legt den Fokus auf die Evaluation der Anonymität und einen Ausblick zur Datenbereitstellung. Dafür werden die generierten, partiell synthetischen Daten der Verdienststrukturerhebung 2018 mithilfe von gängigen Risikometriken bezüglich verbleibender Aufdeckungsrisiken bewertet. Abschließend erfolgt eine Einschätzung möglicher Zugangswege im Hinblick auf ihre juristische Dimension.

📌 **Keywords:** *partially synthetic data – linked synthetic data – risk indicators – disclosure scenarios – loss of anonymity – anonymity loss coefficient (ALC)*

### ABSTRACT

*The first part of the article (Brenzel/Garcia Ritz, 2025) described the methodology and evaluation of data utility. This second part focuses on the evaluation of anonymity and provides an outlook on data provision. To this end, common risk metrics are applied to the partially synthetic data generated from the 2018 Structure of Earnings Survey to assess any remaining disclosure risks. The article concludes by evaluating possible forms of access from a legal perspective.*

#### Dr. Hanna Brenzel

studierte internationale Volkswirtschaft mit Ausrichtung auf Mittel- und Osteuropa an der Universität Regensburg. Sie leitet das Referat „Forschungsdatenzentrum, Methoden der Datenanalyse“ des Statistischen Bundesamtes.

#### Yannik Garcia Ritz

studierte nach seiner Ausbildung zum Fachangestellten für Markt- und Sozialforschung Wirtschafts- und Sozialwissenschaften an der Johann Wolfgang Goethe-Universität Frankfurt am Main (Bachelor of Science) sowie der Johannes Gutenberg-Universität Mainz (Master of Science). Er arbeitet aktuell als wissenschaftlicher Mitarbeiter im Referat „Forschungsdatenzentrum, Methoden der Datenanalyse“ des Statistischen Bundesamtes.

#### David Wagner

ist Rechtsanwalt bei Spirit Legal und spezialisiert auf datenintensive Technologien und Europäisches Digitalrecht. Das Projekt AnigeD begleitet er als wissenschaftlicher Mitarbeiter von Prof. Dr. Mario Martini (Lehrstuhl für Recht und Digitalisierung an der Universität der Bundeswehr München).

#### Hinweis zur Förderung

Diese Forschung wurde durch das Cluster „Anonymisierung bei integrierten und georeferenzierten Daten (AnigeD)“ und seinen Träger, das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR), gefördert.

#### Danksagung

Wir möchten uns herzlich bei Paul Francis (Max-Planck-Institut für Softwaresysteme) für seine vielen hilfreichen theoretischen und praktischen Hinweise zu Verständnis und Anwendung des Anonymity Loss Coefficient (ALC) bedanken.

## 1

### Einleitung

Die Wahl der richtigen Anonymisierungsmethode gleicht der Quadratur des Kreises: Daten sollen nutzbar bleiben und gleichzeitig Persönlichkeitsrechte wahren. Meike Kamp, Berliner Beauftragte für Datenschutz und Informationsfreiheit und Vorsitzende der [Datenschutzkonferenz](#) 2025, bringt das Dilemma auf den Punkt: „Für viele Stellen ist es eine Herausforderung, ein geeignetes Verfahren für die Anonymisierung oder Pseudonymisierung personenbezogener Daten auszuwählen“ (Datenschutzkonferenz, 2025).

Genau hier setzt das Forschungscluster „Anonymisierung bei integrierten und georeferenzierten Daten“ (AnigeD) an. Mit Förderung durch das Bundesministerium für Forschung, Technologie und Raumfahrt (BMFTR) im Rahmen des Konjunkturpakets „Next-GenerationEU“ erprobt es eine vielversprechende Alternative: synthetische Daten<sup>1</sup>. Durch die Generierung „künstlicher“ Merkmalsausprägungen im Sinne von Imputationen können die statistischen Eigenschaften der Originale im Datensatz bewahrt werden, bilden damit jedoch nicht mehr reale Personen in ihren Ursprungsausprägungen ab – theoretisch der Königsweg der Anonymisierung.

Doch die Praxis wirft kritische Fragen auf: Ab wann verlassen synthetische Daten tatsächlich den Anwendungsbereich der Datenschutz-Grundverordnung? Wo verläuft die Grenze zwischen zulässiger statistischer Inferenz und unzulässiger Re-Identifikation?

Die Antworten erfordern eine Übersetzungsleistung zwischen zwei Welten. Statistikerinnen und Statistiker messen Risiken in Wahrscheinlichkeiten und Entropien. Juristinnen und Juristen fragen nach Vernünftigkeit und Verhältnismäßigkeit. Beide Perspektiven müssen verschmelzen, um rechtssichere Anonymisierung zu gewährleisten.

Als Testfall dient die Verdienststrukturerhebung 2018. Brenzel/Garcia Ritz (2025) haben für diesen Datensatz multiple, partiell synthetische Versionen erzeugt. Die

---

1 Nicht zu verwechseln mit der Mikrosimulation synthetisch erzeugter Daten.

[Verdienststrukturerhebung](#) eignet sich besonders als Prüfstein: Sie verknüpft Arbeitgeber- mit Arbeitnehmerdaten (ein sogenannter Linked-Employer-Employee-Datensatz), was eine besondere Herausforderung der Anonymisierung bedeutet, weil die Struktur Angreifenden<sup>12</sup> viele Ansatzpunkte bietet.

Wissenschaftliche Nutzbarkeit eines synthetischen Datensatzes allein genügt jedoch nicht für eine Freigabe durch die Forschungsdatenzentren. Der Knackpunkt bleibt die Anonymität des Datensatzes. Zwar existieren diverse Risikokennzahlen, doch ihre juristische Interpretation gleicht oft einem Ratespiel. Welche Metrik misst, was Gerichte unter „vernünftigerweise identifizierbar“ verstehen?

Dieser Artikel schlägt die Brücke. Er gibt in Kapitel 2 einen Überblick über etablierte und innovative Kennzahlen der Risikoanalyse und evaluiert die Verdienststrukturerhebung 2018 anhand dieser Kennzahlen in Kapitel 3. Kapitel 4 übersetzt technische Risiken in juristische Kategorien und zeigt: Mit modernen Synthetisierungsverfahren können Daten so verändert werden, dass sie sowohl nützlich als auch rechtssicher anonym sind.<sup>13</sup>

## 2

### Risikokennzahlen: von klassischen Kennzahlen zum Non-member-Framework

Die Datensynthetisierung wird seit über drei Jahrzehnten als mögliche Anonymisierungsmaßnahme erforscht und erprobt. Bei partiell synthetischen Daten werden nur ausgewählte Spalten oder Werte synthetisiert, während andere Originalwerte erhal-

---

2 Analog zur einschlägigen (meist englischsprachigen) Literatur zur „Statistical Disclosure Control“ sprechen wir bei allen Datennutzenden von „Angreifern“. Dies bedeutet nicht zwingend, dass diese einen Re-Identifizierungsversuch unternehmen. Es ist beispielsweise denkbar, dass eine Re-Identifizierung auch versehentlich erfolgen kann.

3 Der Artikel untersucht den Schutz personenbezogener Daten innerhalb des Linked-Employer-Employee-Datensatzes. Die Bewertung der Aufdeckungsrisiken für einzelne Unternehmen in diesem Datensatz ist nicht Gegenstand der hier dargelegten wissenschaftlichen Betrachtung und erfordert eine separate Analyse.

ten bleiben. Dies ermöglicht einen Kompromiss zwischen Datennutzen und Schutz personenbezogener Daten, birgt aber inhärente Risiken, da Originaldaten teilweise erhalten bleiben.

Ihr Erfolg hängt davon ab, wie präzise sich verbleibende Aufdeckungsrisiken in voll oder partiell synthetischen Daten messen lassen. Die entwickelten Kennzahlen müssen zwei Risikoarten erfassen: die Re-Identifizierung einer Person (Identity Disclosure) und die Offenlegung ihrer Verhältnisse (Attribute Disclosure) (Raab, 2024). Nicht alle Metriken eignen sich zur Evaluation dieser Risiken sowohl bei voll als auch bei partiell synthetischen Daten (Drechsler/Haensch, 2023).

## 2.1 Grundlegende Metriken nach Reiter/Mitra

Reiter und Mitra (2009) etablierten drei grundlegende Metriken, um Aufdeckungsrisiken für partiell synthetische Daten zu messen.

Der Expected Match Risk (EMR) beantwortet die Grundfrage jedes Angreifers: Wie wahrscheinlich treffe ich bei mehreren Kandidaten den richtigen? Die Autoren modellieren einen rationalen Angreifer, der stets die wahrscheinlichste Übereinstimmung wählt. Bei mehreren gleichwertigen Optionen würfelt er aus dem Kandidatenpool  $c_i$ . Die Trefferwahrscheinlichkeit hängt davon ab, ob das wahre Ziel überhaupt Teil des Pools ist ( $I_i = 1$ ) oder nicht ( $I_i = 0$ ):

$$(1) \text{ EMR} = \sum \frac{I_i}{c_i}$$

True Match Rate (TMR) und False Match Rate (FMR) verfeinern die Analyse bei eindeutigen Zuordnungen. Die True Match Rate zeigt den Anteil korrekt identifizierter Treffer ( $K_j$ ) an allen Zielbeobachtungen ( $N$ ) – die Erfolgsquote des Angreifers:

$$(2) \text{ TMR} = \sum \frac{K_i}{N}$$

Die False Match Rate misst fälschlich als korrekt klassifizierte Matches innerhalb der einzigartigen Zuordnungen ( $s$ ):

$$(3) \text{ FMR} = \sum \frac{F_j}{s}$$

Die Metriken von Reiter und Mitra etablieren eine probabilistische Betrachtung von Anonymität.

## 2.2 Das Synthpop-Framework von Raab

Das R-Paket synthpop ermöglicht nicht nur die Datensynthetisierung, sondern bietet auch Algorithmen, um die Nützlichkeit und verbleibende Aufdeckungsrisiken zu evaluieren (Raab und andere, 2016). Die aktuelle Version implementiert etablierte Metriken wie Differential Correct Attribution Probability (DCAP) und Target Correct Attribution Probability (TCAP) sowie eigene Kennzahlen für Identity und Attribute Disclosure. Die Attribute-Disclosure-Metriken bauen dabei auf den Prinzipien von Differential Correct Attribution Probability und Target Correct Attribution Probability auf (Raab, 2024).

Die Identity-Disclosure-Kennzahlen erfassen strukturelle Vulnerabilitäten. Unique in Original (UiO) identifiziert einzigartige Merkmalskombinationen in Originaldaten – bei 5 % Unique in Original sind 50 von 1000 Beobachtungen einzigartig und damit potenzielle Angriffspunkte. Unique in Original in Synthetic (UiOiS) misst, welcher Anteil dieser vulnerablen Beobachtungen in den synthetischen Daten wiederkehrt:

$$(4) \text{ UiOiS} = 100 \sum \frac{(d_{\cdot q} = 1 | s_{\cdot q} = 1 \cap d_{\cdot q} > 0)}{N_d}$$

Replicated Uniques (repU) markiert die kritischste Konstellation: Beobachtungen, die ihre Einzigartigkeit über die Transformation bewahren:

$$(5) \text{ repU} = 100 \sum \frac{(s_{\cdot q} | d_{\cdot q} = 1 \cap s_{\cdot q} = 1)}{N_d}$$

Das Risiko der Attribute Disclosure bewerten drei Vergleichskennzahlen (Raab und andere, 2024): Disclosive in Original (DiO) misst den Anteil offengelegter Zielvariablenwerte in Originaldaten, Disclosive in Synthetic (DiS) erfasst denselben Anteil in synthetischen Daten und Disclosive in Synthetic, Correct in Original (DiSCO) bestimmt den Anteil synthetischer Werte, die mit den Originalwerten übereinstimmen.

$$(6) \text{ DiSCO} = 100 \sum_q \sum_t \frac{(d_{tq} | p_{stq} = 1)}{N_d}$$

Disclosive in Synthetic, Disclosive in Original (DiS-DiO) erfasst Werte, die sowohl in synthetischen als auch in Originaldaten offengelegt werden und dabei korrekt übereinstimmen.

$$(7) \text{DiSDiO} = 100 \sum_{t=1}^q \sum_{s=1}^t \frac{(d_{tq} | p_{s tq} = 1 \cap p_{d tq} = 1)}{N_d}$$

Alle beschriebenen Kennzahlen setzen voraus, dass Angreifer über bestimmtes Vorwissen verfügen, das Re-Identifizierungen oder die Offenlegung persönlicher Verhältnisse ermöglicht.

Unser Basisszenario unterstellt folgendes Angreiferwissen über Angestellte in der Verdienststruktur-erhebung 2018: Geschlecht, Alter, Dauer der Unternehmenszugehörigkeit sowie Betriebsstandort. Den Standort testen wir auf drei Präzisionsebenen: Gemeinde, Kreis und Regierungsbezirk.

### 2.3 Der Non-member-Ansatz von Francis/Wagner

Francis und Wagner (2025) messen Privatsphäreverletzungen am tatsächlichen Wissenszugewinn: durch die Differenz zwischen Angriffsvorhersagen und privacy-neutralen Baseline-Vorhersagen.

Der Anonymity Loss Coefficient (ALC – Anonymitätsverlustkoeffizient) wird folglich von Francis und Wagner zur Quantifizierung des Risikos eingeführt. Um diesen zu bestimmen, extrahiert das Verfahren einzelne Beobachtungen (Non-member) aus dem Originaldatensatz. Der Angriff wird auf den anonymisierten Datensatz simuliert, während für den Non-member-Datensatz die beste Vorhersage mittels Machine-Learning-Techniken getroffen wird. Analog zu klassischen Metriken wie Expected Match Risk oder True Match Rate nutzt das Verfahren vorausgesetzte Schlüsselvariablen, die es in unterschiedlichen Angriffsszenarien abbildet, die jeweils gesondert betrachtet werden.

Die Methodik misst zwei Kerngrößen: Precision (P) erfasst den Anteil korrekter Vorhersagen, Recall (R) den Anteil der Personen, über die der Angreifer überhaupt Vorhersagen treffen kann. Der Recall spiegelt die ökonomische Rationalität wider: Kein Angreifer investiert hohen Aufwand, wenn er nur über wenige Betroffene etwas erfährt – zumal er vorab nicht weiß, welche Personen betroffen sind. Dabei birgt gleiche Precision bei höherem Recall größere Risiken (Francis/Wagner, 2025). Für die praktische Anwendung

werden kontinuierliche Variablen in Bins unterteilt, um kategoriale Vorhersagen zu ermöglichen und die Vergleichbarkeit zwischen verschiedenen Variablentypen sicherzustellen. Diese Gruppierung ist notwendig, da die Methodik mit kategorialen Werten arbeitet.

Der Precision Recall Coefficient (PRC) kombiniert beide Maße durch Multiplikation der Precision mit einem gewichteten Recall. Die Gewichtung erfolgt durch Logarithmierung und die Konstanten  $\alpha$  und  $R_{min}$  (0,0001), wobei  $\alpha$  den tendenziell kleineren Einfluss des Recalls steuert: Je größer  $\alpha$ , desto schneller erreicht der Precision Recall Coefficient seinen Minimalwert:

$$(8) \text{PRC} = \begin{cases} \left(1 - \left(\frac{\log_{10}(R)}{\log_{10}(R_{min})}\right)^\alpha\right) * P & \text{if } R > R_{min} \\ R & \text{otherwise} \end{cases}$$

Der finale Anonymity Loss Coefficient verwendet ausschließlich den relativen Anonymitätsverlust (Formel 9). Er kann negative Werte annehmen (verbesserte Anonymität) und maximal 1 erreichen (kompletter Anonymitätsverlust). Francis und Wagner empfehlen 0,75 als kritischen Grenzwert.

$$(9) \text{ALC}_{rel} = \frac{\text{PRC}_{atk} - \text{PRC}_{base}}{1 - \text{PCC}_{base}}$$

## 3

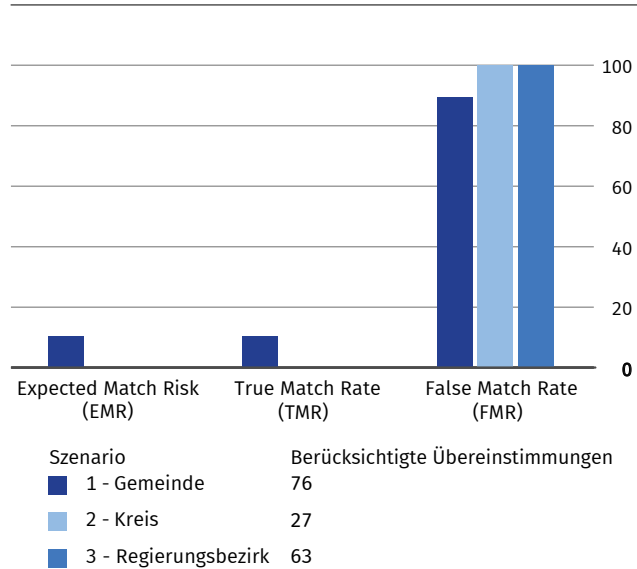
### Ergebnisse

#### 3.1 Expected Match Risk, True/False Match Rate

Im Basisszenario führt eine Zufallsauswahl aus allen potenziellen Matches in etwa 10 % der Fälle zur korrekten Re-Identifizierung (Expected Match Risk). Marginal höher liegt die Trefferquote, wenn nur die wahrscheinlichsten Matches als solche klassifiziert werden (True Match Rate). Allerdings erweisen sich 90 % dieser vermeintlichen Treffer als Fehlzuschreibungen (False Match Rate). Reduziert man die geografische Präzision von Gemeinde auf Kreis oder Regierungsbezirk, sinken Expected Match Risk und True Match Rate auf null – bei gleichzeitig 100 % Fehlzuschreibungen. [➤ Grafik 1](#)

**Grafik 1**

Bewertung der Anonymität der multiplen, partiell synthetischen Verdienststrukturerhebung 2018 anhand der Risikometriken in %



Berechnungen auf Basis der Verdienststrukturerhebung 2018.

### 3.2 Unique in Original in Synthetic, replicated Uniques

Die Betriebsgemeinde einer Zielperson lässt sich nahezu sicher ermitteln, weshalb dieses Wissen als vorausgesetzt eingestuft wird. Bei Preisgabe dieser Information zeigt das Basisszenario: 17 % der Berichtseinheiten in den Originaldaten sind einzigartig identifizierbar, in den partiell synthetischen Daten sogar 18 %. Nur die Hälfte der einzigartigen Originaleinheiten taucht in den synthetischen Daten

**Tabelle 1**

Vergleich der Re-Identifizierungen zwischen Originaldaten (Unique in Original – UiO) sowie den multiplen, partiell synthetischen Daten (replicated Uniques – repU)

|                    | Gemeinde | Kreis  | Regierungsbezirk |
|--------------------|----------|--------|------------------|
|                    | %        |        |                  |
| Originaldaten UiO  | 16,8     | 7,9    | 0,5              |
| Durchschnitt UiS   | 18,1     | 9,9    | 1,0              |
| Durchschnitt UiOiS | 7,5      | 4,4    | 0,3              |
| Durchschnitt repU  | 5,1      | 2,1    | 0,1              |
| (repU/UiO) – 1     | – 69,5   | – 73,5 | – 75,8           |

Berechnungen auf Basis der Verdienststrukturerhebung 2018.

wieder auf. Lediglich 5 % der Berichtseinheiten bleiben in beiden Datenformaten einzigartig – eine Risikoreduktion von 70 % durch die Synthetisierung.

#### ➤ Tabelle 1

Da die Verdienststrukturerhebung 2018 für kleinräumige Analysen ungeeignet ist (Forschungsdatenzentren, 2020a), reduziert größere geografische Aggregation die Risiken erheblich, ohne das Analysepotenzial wesentlich einzuschränken. Der Wechsel von Gemeinde- auf Kreisebene halbiert das Risiko sowohl in Original- als auch in synthetischen Daten. Im direkten Vergleich sinkt das Risiko (replicated Uniques zu Unique in Original) um 73 %.

Bei Aggregation auf Regierungsbezirksebene sind weniger als 0,5 % der Originalbeobachtungen einzigartig. Die Synthetisierung reduziert das Risiko zusätzlich um 76 % – selbst bei diesem groben geografischen Raster.

### 3.3 Disclosive in Synthetic, Correct in Original; Disclosive in Synthetic, Disclosive in Original

Das Bruttojahreseinkommen als zentrale Variable der Verdienststrukturerhebung 2018 wird in den faktisch anonymen Daten durch ein Top-Coding<sup>4</sup> bei 750 000 Euro geschützt. Alle einkommensbezogenen Variablen oberhalb dieser Schwelle werden unter einer Einkommenskategorie zusammengefasst (Forschungsdatenzentren, 2020b). Da Einkommensaufdeckungen besonders kritisch betrachtet werden, wird diese Grenze ebenso für die Analyse des Attributaufdeckungsrisikos in den formal anonymen Daten und den multiplen, partiell synthetischen Daten verwendet. Den Anonymisierungsmaßnahmen in den faktisch anonymen Daten entsprechend wird die Information, ob eine Person 750 000 Euro oder mehr verdient, als potenzielles Angriffsinteresse betrachtet. Zur Berechnung der Metriken für das Attributaufdeckungsrisiko wird die Variable Bruttojahreseinkommen folglich sowohl im originalen, for-

<sup>4</sup> Top-Coding ist eine Anonymisierungsmethode zur Wahrung des Datenschutzes für Extremwerte (in diesem Fall hoher Bruttoeinkommen). Dabei wird ein Maximalwert festgelegt, ab welchem alle Werte dieser Höhe und größer unter eine Kategorie fallen.



mal anonymen Datenmaterial der Verdienststrukturerhebung 2018 als auch in den generierten partiell synthetischen Datensätzen in eine Dummy-Variable transformiert. Die Dummy-Variable unterscheidet, ob die Berichtseinheiten 750 000 Euro und mehr (Dummy = 1) oder weniger (Dummy = 0) verdienen. Diese Dummy-Variable wird unter dem Basisszenario auf Gemeindeebene bezüglich des Attributaufdeckungsrisikos evaluiert.

Bei 97 % der Berichtseinheiten in den Originaldaten kann aus der eindeutigen Merkmalskombination auf das Bruttojahreseinkommen geschlossen werden, unabhängig davon, ob dieses ober- oder unterhalb von 750 000 Euro liegt. In den multiplen, partiell synthetischen Datensätzen sind es durchschnittlich nur 84 %. Bei 82 % stimmen die aufgedeckten Werte in beiden Datenformaten (Original und Synthetisch) überein und sind auch jeweils eindeutig. Die Synthetisierung reduziert das Aufdeckungsrisiko im Vergleich zu den verfügbaren formal anonymen Daten somit um rund 15 %. [↪ Tabelle 2](#)

**Tabelle 2**

Verringerung der Aufdeckungen in multiplen, partiell synthetischen Daten (Disclosive in Synthetic, Disclosive in Original – DiSDiO) im Vergleich zu Originaldaten

|                | Original-<br>daten | DiS    | DiSCO  | DiSDiO | Verringe-<br>rung des Auf-<br>deckungs-<br>risikos |
|----------------|--------------------|--------|--------|--------|----------------------------------------------------|
|                | %                  |        |        |        |                                                    |
| m <sub>1</sub> | 97,25              | 83,73  | 83,72  | 81,48  | 16,22                                              |
| m <sub>2</sub> | 97,25              | 84,28  | 84,27  | 82,13  | 15,55                                              |
| m <sub>3</sub> | 97,25              | 83,97  | 83,96  | 82,34  | 15,33                                              |
| m <sub>4</sub> | 97,25              | 84,33  | 84,32  | 82,33  | 15,34                                              |
| m <sub>5</sub> | 97,25              | 84,06  | 84,04  | 82,02  | 15,66                                              |
| Durchschnitt   | 97,25              | 84,074 | 84,062 | 82,06  | 15,62                                              |

Berechnungen auf Basis der Verdienststrukturerhebung 2018.

Die Aufdeckungsrate von 84 % erscheint dennoch zunächst hoch. Es gibt jedoch zwei Aspekte, die diesen Eindruck relativieren: Erstens, die Schnittmenge der Aufdeckungen zwischen synthetischen Berichtseinheiten und den Berichtseinheiten in den Originaldaten beträgt nicht 100 %. Das heißt, eindeutige Attributaufdeckungen in den synthetischen Daten müssen nicht in den Originaldaten vorliegen. Ebenso sind nicht alle Berichtseinheiten mit eindeutiger

Attributaufdeckung in den Originaldaten auch in den partiell synthetischen Daten zu finden.

Zweitens, beim angenommenen Basisszenario auf Gemeindeebene sind fast ausschließlich Berichtseinheiten mit einem Bruttojahreseinkommen unter 750 000 Euro von einer Aufdeckung dieses Attributs betroffen (über 99 %). Spannender ist es folglich, die bei der faktischen Anonymisierung als besonders schutzwürdig erachteten Berichtseinheiten mit einem Bruttojahreseinkommen von 750 000 Euro oder höher näher zu betrachten.

In den Originaldaten gibt es 111 Berichtseinheiten mit eindeutiger Merkmalskombination, welche über 750 000 Euro verdienen (Dummy-Variable = 1). Nur für 53 bis 61 dieser Hochverdiener kann diese identisch zu den in den Originaldaten eindeutige Merkmalskombination auch in den synthetischen Datensätzen beobachtet werden. Damit können etwa 50 % der ursprünglichen Hochverdiener aus den Originaldaten durch die Synthetisierung erfolgreich verschleiert werden. [↪ Tabelle 3](#)

**Tabelle 3**

Beobachtungen mit einer eindeutigen Merkmalskombination und einem jährlichen Bruttoeinkommen über 750 000 Euro

|                | Originaldaten | Identische<br>Merkmalskombi-<br>nation in synthe-<br>tischen Daten | Verringerung der<br>Aufdeckungen |
|----------------|---------------|--------------------------------------------------------------------|----------------------------------|
|                | Anzahl        |                                                                    | %                                |
| m <sub>1</sub> | 111           | 53                                                                 | 52,25                            |
| m <sub>2</sub> | 111           | 56                                                                 | 49,55                            |
| m <sub>3</sub> | 111           | 53                                                                 | 52,25                            |
| m <sub>4</sub> | 111           | 61                                                                 | 45,05                            |
| m <sub>5</sub> | 111           | 55                                                                 | 50,45                            |
| Durchschnitt   | 111           | 55,6                                                               | 49,91                            |

Berechnungen auf Basis der Verdienststrukturerhebung 2018.

### 3.4 Non-member und basic framework

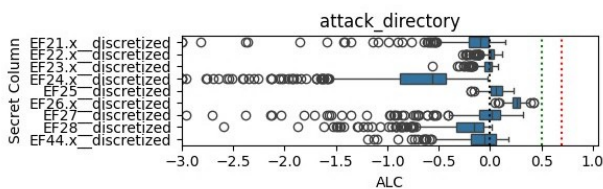
Für die ALC-Berechnung dienten alle neun einkommensbezogenen Variablen als potenzielle Zielvariablen. Als mögliches Angreiferwissen wurden fünf Schlüsselvariablen angenommen: Geschlecht, Geburtsjahr, Unternehmenseintritt, Bildungsabschlüsse (schulisch und beruflich) sowie Befris-

tungsstatus. Verschiedene Kombinationen dieser Variablen bildeten unterschiedliche Angriffsszenarien, womit im Vergleich zum Basisszenario noch mehr mögliche Wissensszenarien in unterschiedlicher Stärke getestet werden können.

Das Ergebnis: Kein einziger De-Anonymisierungsversuch erreicht einen Anonymity Loss Ccoefficient über 0,5. Die meisten Werte liegen sogar deutlich darunter. Die synthetischen Daten weisen damit durchgehend starke Anonymität auf. [➤ Grafik 2](#)

**Grafik 2**

ALC-Werte für Zielvariablen

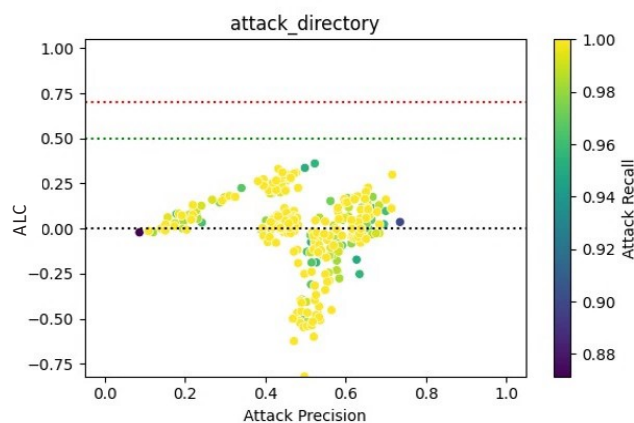


Berechnungen auf Basis der Verdienststrukturerhebung 2018.  
ALC: Anonymity Loss Coefficient – Anonymitätsverlustkoeffizient

[➤ Grafik 3](#) zeigt einen klaren Zusammenhang zwischen Precision eines Aufdeckungsszenarios und dem Anonymitätsverlust (ALC). Der Recall hingegen korreliert bei der Verdienststrukturerhebung 2018 und den getesteten Szenarien nicht erkennbar mit dem Anonymity Loss Coefficient. Dies deutet darauf hin, dass der zugrunde gelegte Anteil, über den ein Angreifer Vorhersagen treffen kann, für die aktuellen

**Grafik 3**

Beste ALC-Werte, zusätzliche kombinierte Darstellung mit Precision und Recall



Berechnungen auf Basis der Verdienststrukturerhebung 2018.  
ALC: Anonymity Loss Coefficient – Anonymitätsverlustkoeffizient

Daten keine Bedeutung hat. Somit hätte ein möglicherweise unterschätztes Angriffspotenzial in Bezug auf die Menge an vorhersagbaren Beobachtungen wohl keine signifikante Unterschätzung des Risikos zur Folge.

Die 351 untersuchten Kombinationen aus Schlüssel- und Zielvariablen erreichen durchweg hohe Anonymitätswerte: 40,7 % haben einen Anonymity Loss Coefficient von 0,5 oder weniger, 59,3 % von unter 0,75. Keine einzige Kombination überschreitet den Schwellenwert von 0,75 (Francis/Wagner, 2025).

## 4

### Die juristische Dimension

#### 4.1 Der normative Rahmen

Artikel 4 Nr. 1 Datenschutz-Grundverordnung definiert personenbezogene Daten als „alle Informationen, die sich auf eine identifizierte oder identifizierbare natürliche Person [...] beziehen“. Als identifizierbar gilt eine Person, wenn sie „direkt oder indirekt [...] identifiziert werden kann“. Erwägungsgrund 26 Datenschutz-Grundverordnung operationalisiert diese Definition: Die Identifizierbarkeit einer natürlichen Person beurteilt sich nach sämtlichen Mitteln, „die von dem Verantwortlichen oder einer anderen Person nach allgemeinem Ermessen wahrscheinlich genutzt werden, um die natürliche Person direkt oder indirekt zu identifizieren [...]“. Hierzu zählt etwa der Abgleich mit anderen Daten (Zusatzwissen). Die Wahrscheinlichkeit der Nutzung eines Mittels ist anhand objektiver Faktoren wie „Kosten der Identifizierung und der dafür erforderliche Zeitaufwand“ unter Berücksichtigung verfügbarer und absehbarer Technologien zu bestimmen.

Satz 5 des Erwägungsgrunds 26 Datenschutz-Grundverordnung stellt klar: „Die Grundsätze des Datenschutzes sollten daher nicht für anonyme Informationen gelten [...] oder personenbezogene Daten, die in einer Weise anonymisiert worden sind, dass die betroffene Person nicht oder nicht mehr identifiziert werden kann.“

## 4.2 Die gefestigte Rechtsprechungslinie

Der Europäische Gerichtshof (EuGH) konkretisierte in der Rechtssache Breyer<sup>5</sup> den Maßstab für die Identifizierbarkeit natürlicher Personen: Relevant sind nur Mittel, die „vernünftigerweise“ zur Identifizierung eingesetzt werden können. Eine Identifizierung ist nicht vernünftigerweise möglich, wenn sie gesetzlich verboten oder praktisch undurchführbar ist, weil sie „einen unverhältnismäßigen Aufwand an Zeit, Kosten und Arbeitskräften erfordern würde, so dass das Risiko einer Identifizierung de facto vernachlässigbar erschiene“ (C-582/14, hier: Randnummer 46). Zusatzinformationen eines Dritten sind zu berücksichtigen, wenn dem Verantwortlichen rechtliche Mittel zur Verfügung stehen, um diese Zusatzinformationen zu erlangen (C-582/14, hier: Randnummern 42–44, 47–49; vergleiche auch [IAB Europe, C-604/22](#) Randnummern 43 und 48).

Diese Grundsätze entwickelte der Europäische Gerichtshof in der Rechtssache Gesamtverband Autoteile-Handel<sup>6</sup> fort: An sich nicht personenbezogene Daten werden zu personenbezogenen Daten, wenn der Verantwortliche sie weitergibt an Dritte, die über Mittel verfügen, mit denen eine Identifizierung der betroffenen Person vernünftigerweise möglich ist (C-319/22, hier: Randnummern 46, 49). Die Daten gelten dann sowohl für den Verantwortlichen als auch für den Dritten als personenbezogen.

Die Rechtssache OC v Commission<sup>7</sup> reiht sich nicht nur nahtlos in die etablierte Rechtsprechungslinie ein, sondern erweitert diese um eine verhaltensökonomische Betrachtungsebene. Erstmals berücksichtigte der Europäische Gerichtshof auch die konkreten Interessen, die Dritte an einer Re-Identifikation haben (C-479/22 P, hier: Randnummer 62 f.). Der Europäische Gerichtshof kann hier so verstanden werden, dass nicht allein die technische Machbarkeit einer Re-Identifikation maßgeblich sein soll, sondern auch die Motivationslage potenzieller Angrei-

fer. Dieses Verständnis überzeugt, weil sich Kosten und Zeitaufwand einer Identifizierung zwar objektiv quantifizieren lassen. Aussagekräftig sind diese Indikatoren jedoch erst durch die Gegenüberstellung mit dem erwarteten Nutzen für den potenziellen Angreifer. Erst aus diesem Verhältnis lässt sich beurteilen, ob eine Identifizierung vernünftigerweise zu erwarten ist.

## 4.3 Relative Anonymität

Der Europäische Gerichtshof hat in der Rechtssache EDSB/SRB<sup>8</sup> eine Position bestätigt, die die Datenschutzaufsicht bislang ablehnte: Daten können für einen Verantwortlichen pseudonymisiert und für einen Dritten zugleich anonymisiert sein. Der Gerichtshof führt damit seine bisherige Rechtsprechung konsequent fort und konkretisiert sie. Die [Artikel-29-Datenschutzgruppe](#) und auch der [Europäische Datenschutzausschuss](#) (EDSA) hatten eine solche relative Betrachtungsweise pauschal verworfen (Artikel-29-Datenschutzgruppe, 2014, hier: Seiten 3, 12, 24 ff.; EDSA, Guidelines 1/2025, Rn. 22).

Der Gerichtshof knüpft die Anonymität bei Weitergabe an eine konkrete Bedingung: Es muss ausgeschlossen werden können, dass Dritte nach allgemeinem Ermessen in der Lage sind, die pseudonymisierten Daten einzelnen betroffenen Personen zuzuordnen, beispielsweise durch den Abgleich mit anderen ihnen verfügbaren Daten.

## 4.4 Grenzen klassischer Metriken aus rechtlicher Perspektive

Die etablierten Risikokennzahlen wie Expected Match Risk, True Match Rate und False Match Rate (Reiter/Mitra, 2009) sowie die Synthpop-Metriken (Raab, 2024) messen primär statistische Wahrscheinlichkeiten und strukturelle Eigenschaften der Daten. Sie quantifizieren, wie viele Beobachtungen einzigartig sind (Unique in Original, replicated Uniques), oder wie wahrscheinlich eine korrekte Zuordnung ist (Expected Match Risk, True Match Rate). Sie adressieren eine technische Möglichkeit der Re-Identifi-

---

5 [Urteil des Europäischen Gerichtshofs vom 19. Oktober 2016, Breyer, C582/14, EU:C:2016:779.](#)

6 [Urteil des Europäischen Gerichtshofs vom 9. November 2023, Gesamtverband Autoteile-Handel, C-319/22, EU:C:2023:837.](#)

7 [Urteil des Europäischen Gerichtshofs vom 7. März 2024, OC v Commission, C-479/22 P, EU:C:2024:215.](#)

8 [Urteil des Europäischen Gerichtshofs vom 4. September 2025, EDSB/SRB, C-413/23 P, EU:C:2025:645.](#)

zierung, geben aber nicht ohne Weiteres Auskunft darüber, ob Re-Identifikationsversuche „vernünftig“ erscheinen.

Nach der Rechtsprechung des Europäischen Gerichtshofs steht eine denkbare technische Möglichkeit einer Re-Identifikation aber nicht per se einer Anonymisierung entgegen – entscheidend ist, ob die Identifizierung „vernünftigerweise“ zu erwarten ist. Die klassischen Metriken zeigen zwar, dass eine Re-Identifikation mit 10 % Wahrscheinlichkeit möglich wäre (Expected Match Risk), sagen aber wenig darüber aus, ob ein rationaler Akteur auch bei einer zehnpromzentigen Chance einen relevanten Erkenntnisgewinn erzielen kann. Sie messen das abstrakte Können, ohne dieses Können in die juristisch maßgebliche Kategorie des „Vernünftigen“ einzuordnen. Gleichwohl bleiben diese Metriken als Risikoindikatoren relevant: Sehr hohe Werte würden unabhängig von der Motivation ein inakzeptables Risiko indizieren.

#### 4.5 Der Anonymity Loss Coefficient als partielle Operationalisierung der EuGH-Rechtsprechung

---

Der Anonymity Loss Coefficient nach Francis/Wagner (2025) misst, wie viel besser ein Angreifer mit Zugang zu den synthetischen Daten eine Zielvariable vorhersagen kann im Vergleich zu jemandem, der nur die statistischen Muster des Datensatzes kennt (Non-member-Baseline). Der Anonymity Loss Coefficient quantifiziert diesen relativen Informationsgewinn. Er weist also primär den Nutzen eines Angriffs aus, nicht direkt den Aufwand. Die Metrik basiert auf der Annahme, dass bei geringem Informationsgewinn kein rationaler Akteur die nötigen Ressourcen investieren würde. Für die rechtliche Bewertung liefert der Anonymity Loss Coefficient gleichwohl ein Indiz: Ein niedriger Anonymity Loss Coefficient deutet auf fehlendes rationales Interesse hin. Dabei gilt: Je geringer der Informationsgewinn, desto höher der relative Aufwand je gewonnene Information. Ein Anonymity Loss Coefficient nahe null deutet darauf hin, dass der Anreiz für Akteure, einen potenziell hohen Aufwand zu betreiben, vernachlässigbar ist. Dies stützt die Annahme, dass eine Re-Identifikation nicht „vernünftigerweise“ zu erwarten ist.

#### 4.6 Die Verdienststrukturhebung im Lichte des Anonymity Loss Coefficient

---

Für die Verdienststrukturhebung 2018 zeigen die empirischen Ergebnisse durchgehend niedrige Risikoindikatoren: Von 351 untersuchten Szenarien erreicht keines den kritischen Schwellenwert von 0,75. Mit 40,7 % der Szenarien im Bereich  $ALC \leq 0$  liefern die synthetischen Daten in diesen Fällen keinen verwertbaren Informationsgewinn oder erschweren sogar korrekte Inferenzen. Diese Werte deuten auf eine substantielle Reduktion der Re-Identifikationsrisiken hin.

Die niedrigen ALC-Werte legen nahe, dass rationale Akteure wenig Anreiz für Re-Identifikationsversuche hätten. Wenn der Informationsgewinn marginal ist, erscheint die Investition erheblicher Ressourcen in einen Angriff (verhaltens-)ökonomisch fragwürdig. Dies gilt besonders für Szenarien mit  $ALC \leq 0$ , bei denen die Synthetisierung die Inferenzmöglichkeiten verschlechtert. Allerdings basiert diese Einschätzung auf der Annahme rationalen Verhaltens – nicht ökonomische Motive könnten trotz geringer Erfolgsaussichten zu Angriffsversuchen führen. Der damit erzielte Erkenntnisgewinn bleibt aber gering.

#### 4.7 Kontextuelle Schutzebenen im Forschungsumfeld

---

Die technischen Kennzahlen sind im spezifischen Bereitstellungskontext zu bewerten. Drei Faktoren sind relevant:

1. Empfängerkreis: Die Datenbereitstellung erfolgt ausschließlich an wissenschaftliche Einrichtungen mit legitimem Forschungsinteresse. Es liegen keine Anhaltspunkte für Re-Identifikationsabsichten in diesem Nutzendenkreis vor.
2. Rechtliche Bindung: Die Empfänger der Daten werden auf das Statistikgeheimnis nach § 16 Bundesstatistikgesetz verpflichtet. Re-Identifikationsversuche wären nicht nur technisch aufwendig und ökonomisch unattraktiv, sondern auch strafbewehrt. Nach der Breyer-Rechtsprechung sind „gesetzlich verbotene“ Mittel aus der

Betrachtung auszuschließen (C-582/14, hier: Randnummer 46).

3. Technische und organisatorische Maßnahmen:  
Die Forschungsdatenzentren implementieren je nach Zugangsweg technische und organisatorische Schutzmaßnahmen. Beispielsweise macht die Protokollierung der Datennutzung (Remote Scientific Use File) Re-Identifikationsversuche nachvollziehbar; Zugangskontrollen im Rahmen der Onsite-Nutzung schränken die verfügbaren Mittel ein.

Diese Faktoren unterstreichen, dass keine vernünftigen Mittel für eine Re-Identifikation bestehen. Wissenschaftlich Nutzende, die auf das Statistikgeheimnis verpflichtet sind und unter kontrollierten Bedingungen arbeiten, haben weder Anreiz noch rechtliche Möglichkeit zur Re-Identifikation. Die durch den niedrigen Anonymity Loss Coefficient indizierte technische Anonymität wird durch diese rechtlichen und organisatorischen Schutzschichten zusätzlich abgesichert.

#### 4.8 Anwendung der rechtlichen Kriterien auf die ALC-Ergebnisse

---

Die ALC-Werte lassen sich zu den etablierten Kriterien des Europäischen Gerichtshofs in Beziehung setzen:

Nach dem Breyer-Urteil (C-582/14) ist eine Identifizierung nicht vernünftigerweise möglich, wenn sie einen unverhältnismäßigen Aufwand erfordert. Ein  $ALC < 0,5$  zeigt, dass der Informationsgewinn weniger als die Hälfte des theoretisch Möglichen beträgt und ein Angreifer nur in geringem Maße Informationen erlangen kann, die über ein privacy-neutrales Maß hinausgehen. Dies indiziert geringen Anreiz für rationale Akteure, wobei der Anonymity Loss Coefficient den tatsächlichen Aufwand nicht direkt misst.

Die Entscheidung OC v Commission (C-479/22 P) berücksichtigt das Interesse Dritter an einer Re-Identifikation. Der Anonymity Loss Coefficient operationalisiert dieses Interesse als messbaren Informationsgewinn. Bei 40,7 % der Szenarien mit  $ALC \leq 0$  besteht kein positiver Informationsgewinn, mithin kein rationales Interesse an einer Re-Identifikation.

Die EDSB/SRB-Rechtsprechung (C-413/23 P) bestätigt, dass Daten für verschiedene Empfänger unterschiedliche Anonymitätsgrade aufweisen können. Im Forschungskontext mit seinen spezifischen Schutzmaßnahmen erreichen die synthetischen Daten ein höheres Anonymitätsniveau als bei unkontrollierter Weitergabe.

Schon die durchgehend niedrigen ALC-Werte der Verdienststrukturerhebung 2018 (alle  $< 0,5$ ; keiner  $> 0,75$ ) sprechen dafür, dass die Voraussetzungen vernünftiger Identifizierbarkeit nicht erfüllt sind. Jedenfalls in Verbindung mit dem Statistikgeheimnis und den organisatorischen Schutzmaßnahmen erfüllen die partiell synthetischen Daten der Verdienststrukturerhebung 2018 die Anforderungen an anonymisierte Daten im Sinne von Erwägungsgrund 26 Datenschutz-Grundverordnung für den spezifischen Forschungskontext.

## 5

---

### Praktische Implikationen und Ausblick

---

Um synthetische Daten zu bewerten, ist es essenziell, die erreichte Anonymität zu evaluieren. Die generierten multiplen, partiell synthetischen Daten der Verdienststrukturerhebung 2018 wurden mit klassischen Risikokennzahlen bis hin zum Non-member-Framework untersucht.

Die etablierten Metriken Expected Match Risk, True Match Rate und False Match Rate (Reiter/Mitra, 2009) zeigen vernachlässigbare Aufdeckungsrisiken. Selbst bei Bereitstellung auf Gemeindeebene liegen Expected Match Risk und True Match Rate bei nur 10,5 %. Da die Verdienststrukturerhebung ohnehin nicht für kleinräumige Analysen konzipiert ist, würde eine Aggregation auf Kreisebene die Risiken vollständig eliminieren.

Die von Raab (2024) entwickelten Kennzahlen belegen eine deutliche Reduktion einzigartiger Beobachtungen (replicated Uniques) in den synthetischen gegenüber den Originaldaten. Auch die Aufdeckungen der Zielvariable Bruttoeinkommen (Disclosive in



Synthetic, Disclosive in Original) fallen in den synthetischen Datensätzen geringer aus.

Die Ergebnisse des Anonymity Loss Coefficient nach Francis/Wagner (2025) bestätigen starke Anonymität für alle Szenarien ( $ALC < 0,5$ ), bei 41 % sogar perfekte Anonymität.

Der Ansatz von Francis und Wagner (2025) fügt sich in die bestehende Rechtsprechung ein, wenn „vernünftigerweise“ ökonomisch interpretiert wird: Technische Machbarkeit allein begründet keine vernünftige Erwartung. Re-Identifikationsversuche ohne Informationsgewinn erscheinen unvernünftig (Breyer) und für rationale Angreifer unattraktiv (OC v Commission). Über den tatsächlichen Aufwand eines Angriffs gibt der Anonymity Loss Coefficient allerdings keine Auskunft.

Alle Kennzahlen deuten darauf hin, dass die multiple partielle Synthetisierung der Verdienststrukturerhebung 2018 eine gesteigerte Anonymität im Vergleich zu formal anonymen Daten ermöglicht. Das Verhältnis zu faktisch anonymen Daten (Scientific Use Files und Remote Scientific Use Files) lässt sich jedoch nicht präzise bestimmen. Wünschenswert wäre eine universelle, juristisch interpretierbare Kennzahl, die sowohl für synthetisierte als auch klassisch anonymisierte Daten Aussagekraft bietet.

Zusammenfassend lässt sich festhalten, dass die Ergebnisse der beiden Teile dieses Artikels grundsätzlich vielversprechend sind: Zum einen konnte ein synthetischer Datensatz mit einer hohen Nützlichkeitsbewertung unabhängig von der Forschungsfrage generiert werden. Hier zeigten die Ergebnisse der globalen Nützlichkeit in Teil 1 (Brenzel/Garcia Ritz, 2025) zufriedenstellende Messwerte. Jedoch zeigte sich auch, dass eine allgemeingültige Aussage für alle wissenschaftlichen Fragestellungen bezüglich der Qualität nicht möglich ist und schlussendlich die Bewertung der Nützlichkeit einzig abhängig von der konkreten Forschungsfrage stattfinden kann. Zum anderen konnte im vorliegenden Teil 2 aufgezeigt werden, dass geringe Aufdeckungsrisiken im Allgemeinen und explizit eine Reduktion des Aufdeckungsrisikos für die angenommenen Wissensszenarien im Vergleich zu den verfügbaren formal anonymen Daten durch die Synthetisierung erzielt

werden konnte. Jedoch muss auch hier konstatiert werden, dass eine eindeutige juristische Einordnung auf dem Kontinuum von formaler zu absoluter Anonymität herausfordernd bleibt und final nicht erfolgen kann.

Damit lässt sich grundsätzlich bestätigen, dass unabhängig von der künftigen Forschungsfrage qualitativ hochwertige nutzbare synthetische Daten erstellt werden können, welche grundsätzlich geringere Aufdeckungsrisiken bergen als Originaldaten. Der Zielkonflikt zwischen Anonymität und Analysepotenzial kann jedoch auch bei der Datensynthetisierung nicht aufgelöst werden. Zudem ist festzuhalten, dass die Einordnung des Anonymitätsgrads letztlich eine mathematisch-methodische Fragestellung bleibt. Die Datensynthetisierung ist somit eine Möglichkeit, das Aufdeckungsrisiko zu verringern; allerdings fordert sie hohe Aufwände und geht mit einem Verlust des Analysepotenzials einher. Aus Sicht der Autoren hält die Synthetisierung daher den hohen Erwartungen in den laufenden Debatten nicht stand. Final bleibt der Trade-off zwischen Anonymität und Analysepotenzial bestehen. [📄](#)



## LITERATURVERZEICHNIS

---

Artikel-29-Datenschutzgruppe. *Stellungnahme 5/2014 zu Anonymisierungstechniken*. 0829/14/DE, WP216. 2014. [Zugriff am 6. November 2025]. Verfügbar unter: [ec.europa.eu](http://ec.europa.eu)

Brenzel, Hanna/Garcia Ritz, Yannik. *Partielle Synthetisierung zur Anonymisierung der verknüpften Verdienststrukturerhebung 2018. Teil 1: Methodik und Evaluation der Datennützlichkeit*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2025, Seite 83 ff.

Datenschutzkonferenz. *DSK plant Hilfestellungen zur Anonymisierung und Pseudonymisierung von personenbezogenen Daten*. Pressemitteilung der DSK vom 30. Januar 2025. [Zugriff am 4. November 2025]. Verfügbar unter: [www.datenschutz-berlin.de](http://www.datenschutz-berlin.de)

Drechsler, Jörg/Haensch, Anna-Carolina. *30 years of synthetic data*. 2023. DOI: [10.48550/arXiv.2304.02107](https://doi.org/10.48550/arXiv.2304.02107)

Europäischer Datenschutzausschuss (European Data Protection Board). *Guidelines 1/2025 on Pseudonymisation*. Version 1.0. 2025. [Zugriff am 11. Februar 2026]. Verfügbar unter: [www.edpb.europa.eu](http://www.edpb.europa.eu)

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. *Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung der Verdienststrukturerhebung 2018 per On-Site-Nutzung*. 2020a. DOI: [10.21242/62111.2018.00.00.1.1.0](https://doi.org/10.21242/62111.2018.00.00.1.1.0)

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder. *Metadatenreport. Teil II: Produktspezifische Informationen zur Nutzung der Verdienststrukturerhebung 2018 per Scientific-Use-File*. 2020b. DOI: [10.21242/62111.2018.00.00.3.1.0](https://doi.org/10.21242/62111.2018.00.00.3.1.0)

Francis, Paul/Wagner, David. *Towards Better Attribute Inference Vulnerability Measures*. 2025. DOI: [10.48550/arXiv.2507.01710](https://doi.org/10.48550/arXiv.2507.01710)

Raab, Gillian M. *Privacy Risk from Synthetic Data: Practical Proposals*. In: Domingo-Ferrer, Josep/Önen, Melek (Herausgeber). *Privacy in Statistical Databases*. 2024, Seite 254 ff. DOI: [10.1007/978-3-031-69651-0\\_17](https://doi.org/10.1007/978-3-031-69651-0_17)

Raab, Gillian M./Nowok, Beata/Dibben, Chris. *Practical Data Synthesis for Large Samples*. In: *Journal of Privacy and Confidentiality*. Jahrgang 7. Ausgabe 3/2016, Seite 67 ff. DOI: [10.29012/jpc.v7i3.407](https://doi.org/10.29012/jpc.v7i3.407)

Reiter, Jerome P./Mittra, Robin. *Estimating Risks of Identification Disclosure in Partially Synthetic Data*. In: *Journal of Privacy and Confidentiality*. Ausgabe 1/2009, Seite 99 ff. DOI: [10.29012/jpc.v1i1.567](https://doi.org/10.29012/jpc.v1i1.567)

## RECHTSGRUNDLAGEN

---

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 – DSGVO) (Amtsblatt der EU Nr. L 119 vom 4. Mai 2016).

#### Melanie Brauchler

ist Umweltwissenschaftlerin mit einem Schwerpunkt auf Geoinformatik und Fernerkundung.

Seit November 2024 ist sie wissenschaftliche Mitarbeiterin im Referat „Erforschung neuer digitaler Daten“ des Statistischen Bundesamtes im Bereich Fernerkundungsdaten.

#### Maren Köhlmann

ist Volkswirtin und Soziologin.

Der thematische Schwerpunkt der Referentin im Referat „Erforschung neuer digitaler Daten“ des Statistischen Bundesamtes liegt im Bereich Fernerkundungsdaten, in welchem sie in verschiedenen nationalen sowie internationalen Forschungsprojekten arbeitet.

#### Franziska Kraft

ist Geografin und seit 2024 wissenschaftliche Mitarbeiterin in der Abteilung für Georisiken und zivile Sicherheit des Deutschen Zentrums für Luft- und Raumfahrt (DLR) in Oberpfaffenhofen. Ihr Forschungsinteresse gilt der Entwicklung datengestützter Lösungen zur Verbesserung der Klimaresilienz und -anpassung.

#### Manuel Köberl

ist Geoinformatiker, M.Sc. mit Fokus auf Urbanisierung und Fernerkundung. Seit Oktober 2022 ist er in der Abteilung für Georisiken und zivile Sicherheit im Team „Stadt und Gesellschaft“ des Deutschen Zentrums für Luft- und Raumfahrt (DLR) in Oberpfaffenhofen tätig.

#### Danksagung

Unser besonderer Dank gilt Klara Marie Schönenberger, die das Projekt organisatorisch und methodisch maßgeblich entwickelt und begleitet hat. Nur aufgrund ihres tatkräftigen Engagements konnte das Projekt zu diesem Entwicklungsstand gebracht werden.

# ENTWICKLUNG VON STATISTISCHEN WIRTSCHAFTSINDIKATOREN UNTER VERWENDUNG DER COPERNICUS-SENTINEL-SATELLITEN

Melanie Brauchler, Maren Köhlmann, Franziska Kraft, Manuel Köberl

➤ **Schlüsselwörter:** experimentelle Statistik – Satellitendaten – neue digitale Daten – Konjunkturschnellschätzung – Automobilproduktion – Bautätigkeit

## ZUSAMMENFASSUNG

Das Statistische Bundesamt hat gemeinsam mit dem Deutschen Zentrum für Luft- und Raumfahrt und unter Beratung der Europäischen Raumfahrtagentur das von Eurostat geförderte Forschungsprojekt „Satellite-based economic flash estimation“ (Sat4EC) durchgeführt. Darin wurde der Einsatz von Fernerkundungsdaten, speziell Satellitendaten, zur Berechnung von Wirtschaftsindikatoren für die ökonomischen Schnellschätzungen untersucht. Ziel war, unter Berücksichtigung der speziellen Anforderungen der Statistik, zwei satellitenbasierte Indikatoren zu erarbeiten, die ein Abbild der ökonomischen Entwicklungen in der Automobil- sowie der Baubranche zulassen.

➤ **Keywords:** *experimental statistics – satellite data – new digital data – economic flash estimate – automotive production – building activity*

## ABSTRACT

*In collaboration with the German Aerospace Center and in consultation with the European Space Agency, the Federal Statistical Office conducted a research project on “Satellite-based economic flash estimation” (Sat4EC) which was funded by Eurostat. The project investigated the use of remote sensing data, specifically satellite data, to calculate economic indicators for econometric flash estimation. The aim was to develop two satellite-based indicators that reflect the economic developments in the automotive and construction industries while considering the specific requirements of statistics.*

## 1

### Einleitung

Aufgrund der hohen Dynamik der Weltwirtschaft besteht eine große Nachfrage nach Wirtschaftsdaten am aktuellen Rand. Insbesondere in Krisenzeiten können traditionelle makroökonomische Indikatoren die kurzfristigen wirtschaftlichen Entwicklungen nur schwer abbilden, denn diese stehen aufgrund der aufwendigen Datenaufbereitung meist nur mit einer zeitlichen Verzögerung zur Verfügung. Fernerkundungsbasierte Ansätze eröffnen Möglichkeiten, diese Lücken zu schließen. Wirtschaftliche Aktivitäten können optische Spuren auf der Erdoberfläche hinterlassen, welche wiederum von Sensoren erfasst und quantifiziert werden können. Im Kontext der ökonometrischen Schnellschätzung könnten solche Daten das zeitnahe Beobachten von Wirtschaftstrends ermöglichen und somit Behörden, Unternehmen und Politik helfen, schneller auf Krisen oder Veränderungen zu reagieren.

Der Einsatz von Satellitendaten für (sozio-)ökonomische Fragestellungen ist weit verbreitet und reicht vom Einsatz in der Land- und Forstwirtschaft bis zum Einsatz im Bereich der Urbanisierung und des Flächenverbrauchs (Wuepper und andere, 2024; Taubenböck und andere, 2024). Weiterhin gibt es eine Fülle an Studien, die mithilfe der nächtlichen Beleuchtung (die von Satelliten erfassbar ist) auf die ökonomische Entwicklung von Ländern oder größeren Regionen zu schließen versuchen (Kondmann und andere, 2021; Gibson und andere, 2020). Diese Studien sind für Volkswirtschaften geeignet, in denen der ökonomische Aufschwung auch mit der Elektrifizierung einhergeht und in denen entsprechende Statistiken nicht verfügbar oder sehr unzuverlässig sind.

Studien, die sich im Speziellen mit Satellitendaten für ökonometrische Schnellschätzungen beschäftigen, sind selten. In ihrer systematischen Literaturauswertung zu Datenquellen in der ökonometrischen Schnellschätzung erwähnen Stundziene und andere (2024) Satellitendaten zwar als eine alternative Quelle für die ökonometrische Schnellschätzung, benennen jedoch keine konkreten Anwendungen. Es existieren allerdings Studien, die sich mit der

Auslastung von Containerhäfen (Spelta und andere, 2025) oder der Auslastung von Parkplätzen im Einzelhandel beschäftigen (Drouyer/de Franchis, 2020).

Das Statistische Amt der Europäischen Union (Eurostat) finanziert das Projekt „Satellite-based economic flash estimates“ (kurz: Sat4EC). In diesem Rahmen haben sich die Projektpartner Statistisches Bundesamt und Deutsches Zentrum für Luft- und Raumfahrt (DLR) zur Aufgabe gemacht, wirtschaftliche Indikatoren aus Satellitendaten zu entwickeln. Diese sollten so ausgestaltet sein, dass sie als Zuschätzer für die ökonometrische Schnellschätzung in den Volkswirtschaftlichen Gesamtrechnungen (VGR) eingesetzt werden können. In dem Projekt wurden ein Prototyp eines wirtschaftlichen Indikators der Europäischen Raumfahrtagentur (European Space Agency, kurz: ESA) zur Automobilindustrie weiterentwickelt und ein Indikator zur Detektion von Bautätigkeiten von Grund auf neu entwickelt.

Nach einer Einführung in die Quellen von Fernerkundungsdaten in Kapitel 2 stellt Kapitel 3 neben dem Copernicus-Programm, über welches die verwendeten Satellitendaten bezogen wurden, auch die bei der Analyse des Automobilindikators genutzte Datenplattform vor. Anschließend beleuchtet Kapitel 4 die Methoden zur Erstellung der satellitenbasierten Indikatoren im Detail. Abschließend folgen in Kapitel 5 eine Bewertung der aus dem Projekt heraus identifizierten Potenziale für den Einsatz von Satellitendaten für die ökonometrische Schnellschätzung und ein Ausblick.

## 2

### Datengrundlage: Fernerkundungsdaten im Projekt Sat4EC

Fernerkundungsdaten werden mithilfe von Sensoren erfasst, die sich auf unterschiedlichen Trägersystemen<sup>1</sup> befinden. Eine bedeutende Quelle von Fernerkundungsdaten sind Luftbilder, die durch Bildflüge erfasst werden. Durch die geringe Flughöhe

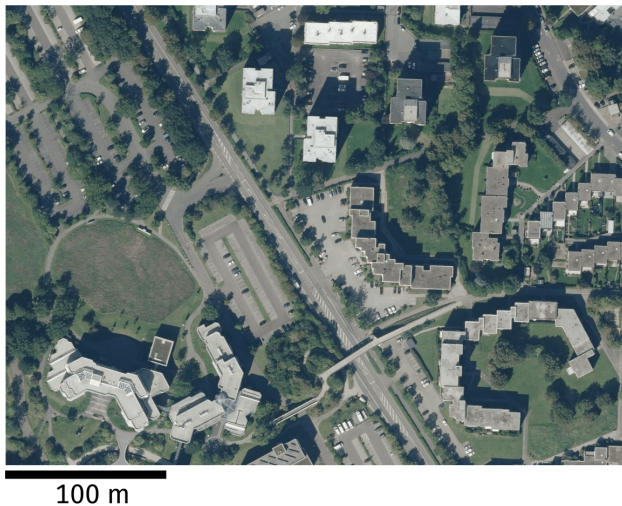
---

1 Unter anderem Drohnen, Flugzeuge und Satelliten.

kann eine hohe Bodenauflösung<sup>12</sup> erreicht werden, die eine detaillierte Betrachtung der Erdoberfläche sicherstellt. Ein Beispiel dafür zeigt [Grafik 1](#) mit einem Luftbild, dessen Bodenauflösung 20 Zentimetern entspricht. Die Interpretation eines Luftbildes erlaubt es, visuell Strukturen (wie Gebäude, Bäume, Straßen) in den Aufnahmen zu erfassen. Aus diesem Grund sind Luftbilder für automatisierte Objekterkennungen grundsätzlich gut geeignet.

#### Grafik 1

Ausschnitt aus einem Digitalen Orthophoto (DOP20)



© GeoBasis-DE / LVermGeoRP (2025), dl-de/by-2-0

Ein Nachteil von Luftbildern besteht jedoch in der zeitlichen Verfügbarkeit der Daten. Aufgrund des hohen organisatorischen und finanziellen Aufwands sowie der großen Flächenausdehnung größerer Bundesländer können die Landesgebiete nicht in einem einzigen Bildflug erfasst werden. Stattdessen erfolgen die Bildflüge abschnittsweise über bestimmte fachliche Zeiträume hinweg. Entsprechend werden vollständige Datensätze in der Regel im Rhythmus von zwei bis drei Jahren produziert. Sie sind daher für die Detektion von Objekten zum Aufnahmezeitpunkt geeignet, wie sich im Projekt „Satelliten- und weitere Fernerkundungsdaten für das Gebäude und Wohnungsregister“ (Sat4GWR) darstellen ließ (Hen-

nig und andere, 2025). Aktuell wird dazu in dem von Eurostat geförderten Projekt „Earth Observation and AI for Construction Statistics“ (EO4ConStat<sup>13</sup>) anhand der KI-basierten Detektion von Baustellen geforscht. Die größeren Zeitabstände zwischen den Bildflügen lassen jedoch keine kontinuierliche Beobachtung der Erdoberfläche zu, was für eine adäquate Abbildung von frühwirtschaftlichen Entwicklungen nötig wäre.

Daher wurde im Projekt eine weitere Quelle von Fernerkundungsdaten untersucht. Fernerkundungssensoren auf Satellitensystemen erlauben es, die Erdoberfläche großflächig und in kurzen zeitlichen Abständen zu betrachten. Neben kommerziellen Anbietern, die vorrangig räumlich hoch- bis sehr hochaufgelöste Daten anbieten, existieren im Rahmen des Copernicus-Programms der Europäischen Kommission Satellitensysteme, welche die Erdoberfläche mit einer Bodenauflösung von bis zu 10 Metern kontinuierlich in geringen Zeitabständen erfassen. Neben der zeitlichen weist diese Datenquelle noch einen weiteren entscheidenden Informationsvorteil auf: die spektrale Dimension ([Grafik 2](#), die Grafik 1 in der geografischen Lage entspricht). Während Luftbilder maximal vier Spektralkanäle (Blau, Grün, Rot und nahes Infrarot) erfassen, werden bei Satellitendaten häufig weitere Spektralkanäle eingesetzt.

Die im linken Teil der Grafik 2 eingefügte Aufnahme des Sentinel-2-Satelliten (in Echtfarbdarstellung, 10-m-Auflösung) scheint keine hohe Informationsdichte zu liefern. Die Aufnahme ist jedoch nur aus den farblich dargestellten Spektralkanälen zusammengesetzt (Blau, Grün und Rot, um die Echtfarbdarstellung zu erhalten). Ein Satellitenbild des Sentinel-2 liefert jedoch aufgrund der verbauten Sensoren Informationen in weiteren Spektralkanälen, die schematisch in der Mitte der Grafik 2 abgebildet sind. Ein Blick auf die Reflektionsspektren<sup>14</sup> im rechten Teil der Grafik zeigt, dass sich Oberflächen

12 Die Begriffe der sehr hohen, hohen oder mittleren Bodenauflösung sind fluide. Es ist jedoch üblich, bei weniger als 1 Meter je Pixel von sehr hochauflösenden Daten, bei unter 5 Metern je Pixel von hochauflösenden Daten zu sprechen.

3 Ein Werkstattbericht des Projektes steht unter [destatis.de](https://destatis.de) --> EXSTAT --> EO4ConStat zur Verfügung.

14 Ein Reflektionsspektrum beziehungsweise eine spektrale Signatur ist die charakteristische Rückstreuung elektromagnetischer Energie, die eine Oberfläche abhängig von der Wellenlänge aufweist. Während Wasser über alle Wellenlängen eine geringe Reflektion hat, zeigt vitale Vegetation ein charakteristisches lokales Maximum im visuellen Grün (~ 600 Nanometer) und einen starken Anstieg zum nahen Infrarot hin (~700 bis 800 Nanometer).

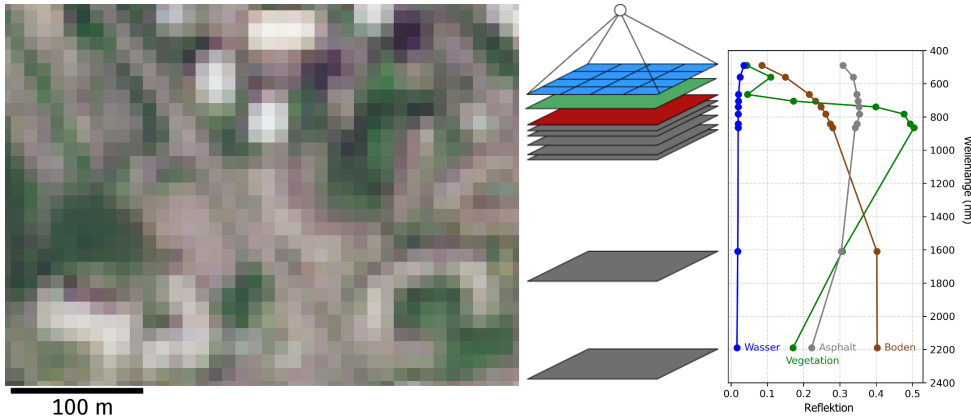
## Grafik 2

### Spektrale Dimension einer Sentinel-2-Satellitenaufnahme

Aufnahme des Sentinel-2-Satelliten<sup>1</sup>

Spektralkanäle

Spektralsignaturen



<sup>1</sup> In Echtfarbdarstellung, 10-m-Auflösung.

insbesondere auch in ihrem weiteren Verlauf und nicht nur im roten, grünen und blauen Kanal unterscheiden. Diese Informationen tragen dazu bei, dass visuell gleichartige Oberflächen durch die spektrale Signatur unterschieden werden können.

## 3

### Das Copernicus-Programm

Das Copernicus-Programm ist ein von der Europäischen Kommission beaufsichtigtes und mit den Mitgliedstaaten der Europäischen Union (EU) koordiniertes Programm zur Erdbeobachtung. Die Ziele des Copernicus-Programms sind eine bedarfsgerechte Infrastruktur, ein kostenfreier und offener Datenzugang, die Datenerfassung und -bereitstellung sowie die Stärkung von Märkten und Innovation (Europäische Union, 2015). Neben der In-Situ-Komponente, in der Daten direkt auf der Erdoberfläche durch die Akteure in den Mitgliedstaaten erfasst werden (zum Beispiel an Wetterstationen), enthält das Programm eine stark ausgeprägte Komponente zur satellitenbasierten Fernerkundung. Diese Komponente ermöglicht es Nutzenden, aktuelle und kostenfreie Satellitendaten zur Beobachtung der Erdoberfläche und der Atmosphäre zu beziehen. Die Satellitendaten werden darüber hinaus in themenspezifische

[Copernicus-Kerndienste](#) gespeist (zum Beispiel in das [Katastrophen- und Krisenmanagement](#)). Auch die Daten und Produkte dieser Dienste sind kostenlos verfügbar.

### 3.1 Sentinel-Satelliten

Unter dem Namen „Sentinel“ starteten seit 2014 zehn eigenständige Satelliten in die Erdumlaufbahn – unter dem Schirm des Copernicus-Programms und mithilfe der technischen Koordination der Europäischen Raumfahrtagentur. Diese beobachten kontinuierlich die Erdoberfläche und können so die Dynamik von Veränderungen an der Erdoberfläche abbilden. Mit den Konstellationen Sentinel-1 (Radardaten) und Sentinel-2 (Multispektraldaten, wie in Grafik 2 dargestellt) befinden sich Satelliten in der Erdumlaufbahn, die eine Bodenauflösung von bis zu 10 Metern liefern. Die Sentinel-Satelliten eröffnen so neue Möglichkeiten der Informationsbeschaffung und der Analyse, die für Forschungszwecke genutzt werden können. Betrachtet man einzelne Satellitenaufnahmen von Sentinel-1 und Sentinel-2, so stehen die erfassten Daten bereits innerhalb einer Stunde nach der Erfassung zum Download bereit. Ein einzelner Sentinel-1-Satellit ist in der Lage, über eine Zeit von zwölf Tagen die gesamte Erdoberfläche zu erfassen. Die Konstellation aus aktuell zwei Satelli-



ten erreicht einen Zyklus von sechs Tagen. Aufgrund der sich überlappenden Umlaufbahnen in den höheren Breiten kann etwa alle drei Tage eine Abdeckung des gesamten deutschen Staatsgebiets erreicht werden. Die langfristigen Investitionen in das Copernicus-Programm und in die verschiedenen Satelliten erfüllen eine Forderung nach einer konsistenten und unabhängigen Quelle von hochqualitativen Daten (Europäische Union, 2015).

### 3.2 Das Copernicus Data Space Ecosystem

Das [Copernicus Data Space Ecosystem](#) dient seit 2023 als zentraler Zugang zu Copernicus-Satellitendaten. Es wird kontinuierlich um Produkte aus den themenspezifischen Kerndiensten erweitert. Mit einem kostenfreien Nutzerkonto können grundlegende Anwendungen und Analysemöglichkeiten genutzt werden. Das Copernicus Data Space Ecosystem umfasst verschiedene Plattformen und Services, darunter den Copernicus-Browser, der die Visualisierung von Satellitendaten in vordefinierten oder benutzerdefinierten Darstellungen ermöglicht. Weiterhin werden auch die Datenrecherche sowie ein direkter Download von Sentinel-Daten unterstützt. Eine weitere Komponente des Copernicus Data Space Ecosystem sind Arbeitsumgebungen in JupyterLab und ein eigener Speicherbereich<sup>5</sup>, ergänzt durch Schnittstellen zu verschiedenen Services. Zu diesen Services zählt auch der SentinelHub, ein ursprünglich kostenpflichtiger Service des Unternehmens [Sinergise](#), der über das Copernicus Data Space Ecosystem kostenfrei zugänglich ist. Der [SentinelHub](#) ermöglicht eine serverseitige Datenprozessierung der Sentinel-Daten, sodass nicht die Rohdaten, sondern bereits aufbereitete Analyseprodukte für die Nutzenden bereitgestellt werden. Bei dem in Abschnitt 4.1 erläuterten Automobilindikator kommt eine solche serverseitige Prozessierung über die statistische Schnittstelle des SentinelHub zum Einsatz. Gerade im europäischen Kontext und in Fragen der Übertragbarkeit von Prozessierungsschritten auf andere Länder bietet das Copernicus Data Space Ecosystem große Vorteile, da

alle Services und Daten europaweit einheitlich zur Verfügung stehen.

## 4

### Indikatoren aus Satellitendaten

Aufgrund der hohen Aktualität der Schnellschätzungen des Bruttoinlandsprodukts in den Volkswirtschaftlichen Gesamtrechnungen werden Zuschätzungen zur Berechnung benötigt. Um die Genauigkeit dieser Zuschätzungen zu verbessern, sollen neue Datenquellen erschlossen werden. Da sich die ökonomische Schnellschätzung im Spannungsfeld zwischen Aktualität und Genauigkeit befindet, können hierfür nach einer gründlichen Prüfung auch nicht amtliche Indikatoren eingesetzt werden (Ackermann und andere, 2021). Diese müssen zeitnah nach Ende eines Quartals, mindestens monatlich und im Optimalfall für alle Monate des zu berechnenden Quartals verfügbar sein und Aktivitäten abbilden, die einem Wirtschaftszweig zugeordnet werden können. Die Verarbeitung von Satellitendaten ist jedoch mit besonderen Herausforderungen verbunden.

Die Ableitung von statistischen Indikatoren aus Satellitendaten erfordert Methoden, welche die komplexen Bildinformationen in eine nutzbare statistische Form überführen. Dazu gehören zeitliche, räumliche und spektrale Filterungen der Daten sowie die Aggregation in diesen Dimensionen. Ziel ist, die komplexen Daten in eine einfache, zweidimensionale Struktur zu überführen, die in der Statistik nutzbar ist (zum Beispiel Zeitvariable und Wert).

Satellitendaten werden nicht gezielt für statistische Zwecke erhoben, sondern zu festen Umlaufzeiten aufgenommen, sodass sie letztendlich Momentaufnahmen darstellen. Das bedeutet, dass der Zeitpunkt des Überflugs für die statistische Fragestellung ungeeignet sein kann (wenn zum Beispiel die Aufnahme nicht dem Zeitpunkt der ökonomischen Aktivität entspricht)<sup>6</sup>. Gleichzeitig eröffnen sie neue Möglichkeiten, wirtschaftliche Entwicklungen schnell

---

5 Dieser Speicher wird nach 15 Tagen ohne Zugriff gelöscht, was jedoch durch eine Systemnachricht per E-Mail rechtzeitig angekündigt wird und durch Anmeldung verhindert werden kann.

6 Sentinel-1-Aufnahmen außerhalb von Ladenöffnungszeiten wären beispielsweise nicht geeignet, um die Auslastung von Parkplätzen im Einzelhandel zu betrachten.

und unabhängig von traditionellen Datenquellen abzubilden.

## 4.1 Automobilindikator

Die Automobilindustrie ist ein bedeutender Treiber der konjunkturellen Entwicklung in Deutschland. Ein von der Europäischen Raumfahrtagentur entwickelter Prototyp eines [Indikators](#) zur Auslastung von Produktionsparkplätzen nutzt Sentinel-1-Rückstreuwerte als Schätzer für die Automobilproduktion. Die zugrunde liegende Annahme ist, dass sich Fahrzeuge zeitnah nach der Produktion auf den Parkplätzen der jeweiligen Produktionsstandorte befinden und die Auslastung dieser Parkplätze über Satelliten erfasst werden kann. Die Weiterentwicklung dieses Prototyps wurde maßgeblich von der Europäischen Raumfahrtagentur unterstützt, indem sie den Quellcode für den Prototyp bereitstellte und das Projekt beratend betreute.

Die Satelliten der Sentinel-1-Konstellation senden Radarsignale (C-Band SAR<sup>7</sup>) zur Erde und erfassen die Rückstreuung. Anders als optische Satelliten (siehe Grafik 2) liefern sie Informationen über Form und Materialeigenschaften von Oberflächen. Ein Vorteil von Radarsatelliten ist, dass sie unabhängig von Wolkenbedeckung sind. Sie werden mitunter für die Schnellerkennung von Überflutungen genutzt, wenn Wolkenbedeckung die Analyse der Aufnahmen optischer Systeme erschwert. Da Radarsysteme das Radarsignal aktiv erzeugen, können sie im Gegensatz zu optischen Satelliten auch unabhängig von der Tageszeit die Erdoberfläche erfassen. Glatte Flächen wie Wasser, Straßen oder unbewachsener Boden wirken als spiegelnde Flächen. Da das vom Satellitensystem ausgesendete Radarsignal nicht senkrecht, sondern in einem bestimmten Winkel auf die Oberfläche trifft, reflektiert das gespiegelte Signal vom Sensor weg und ein entsprechend niedriges Signal wird am Sensor aufgezeichnet. Komplex strukturierte Oberflächen, wie Bäume, oder hohe Objekte, wie Gebäude, werden mit einem stärkeren Signal

am Sensor erfasst, da das Radarsignal an diesen Flächen streut oder mehrfach reflektiert wird. Diese Charakteristik von Radardaten ermöglicht es, die Auslastung von Parkplätzen anhand der Stärke der Rückstreuung zu erkennen. Je mehr Fahrzeuge auf der Fläche stehen, desto stärker ist das am Sensor erfasste Radarsignal.

Der weiterentwickelte Indikator basiert auf der Auslastung von Produktionsparkplätzen an 18 Standorten in Deutschland (Kraft und andere, 2025). Hierfür mussten die Produktionsparkplätze manuell identifiziert und digitalisiert werden. Bei der Auswahl der Parkplatzflächen wurden einige Voraussetzungen geprüft. Die Produktionsstätte sollte möglichst nur über Produktionsparkplätze im Freien verfügen. Diese sollten auch weiterhin nicht durch andere Strukturen (zum Beispiel Kräne oder Vegetation) verdeckt sein. Die Parkplätze sollten nur für die Lagerung produzierter Fahrzeuge dienen und keine Parkplätze für Mitarbeitende einschließen. Letztlich sollten die Parkplätze auch über die komplette Beobachtungszeit unverändert bleiben.

Der von dem Projektpartner Deutsches Zentrum für Luft- und Raumfahrt entwickelte Code<sup>8</sup> konnte auf das Copernicus Data Space Ecosystem (CDSE) übertragen werden, sodass eine zunächst geplante kostenpflichtige Option nicht mehr benötigt wurde. Die bereits in Kapitel 3 erwähnte statistische Schnittstelle zum SentinelHub erlaubt es, einen vordefinierten Flächenausschnitt an den Server zu übermitteln und statistische Auswertungen auf diese Fläche zu begrenzen. Die Produktionsparkplätze werden in ihrer geografischen Ausdehnung an den Server übermittelt. Als Resultat werden nicht alle Satellitendaten zurückgeliefert, die in den Parkplatzbereich fallen, sondern die Daten serverseitig aggregiert und statistische Größen (beispielsweise Mittelwert, Standardabweichung) geliefert. Diese räumlichen Statistiken können letztlich als einfache Tabellen im eigenen Arbeitsbereich des CDSE gespeichert oder heruntergeladen werden.

Die für die Konjunkturstatistik nötige zeitliche Dimension wurde durch unterschiedliche Konfigurationen adressiert. Die Indikatorenberechnung des Deutschen Zentrums für Luft- und Raumfahrt erlaubt

7 Der Sensor operiert in einer Wellenlänge von 6 Zentimetern. SAR steht für „Synthetic Aperture Radar“; es handelt sich um bildgebende Radarsysteme. Die gelieferten Satellitendaten sind somit ebenso räumlich darstellbar (2D) wie die multispektralen Satellitendaten des Sentinel-2.

8 <https://github.com/dlr-eoc/sat4ec>

es, Einstellungen bezüglich der zeitlichen Auflösung (semi-täglich und monatlich) vorzunehmen. Weiterhin wurde die Detektion von Anomalien eingeführt, die einen Hinweis auf mögliche strukturelle Veränderungen der Produktionsparkplätze geben soll. **► Grafik 3** stellt die mittlere Parkplatzauslastung aller betrachteten Parkplätze den gemeldeten produzierten Waren gegenüber. Hier ist eine gewisse Übereinstimmung sichtbar. In der Validierung mit Einzeldaten aus der Produktion, aber auch mit anderen Datenanbietern, zeigten sich jedoch große Unterschiede der einzelnen Produktionsstandorte und Hersteller (Kraft und andere, 2025).

Fahrzeuge, die länger auf den Parkplätzen gelagert werden, sowie Fahrzeuge, die lediglich umgeparkt werden, lassen sich mit dieser Methode nicht identifizieren, da nicht die einzelnen Fahrzeuge kontinuierlich beobachtet werden, sondern lediglich die Auslastung der Produktionsparkplätze. Die Annahme, dass eine Lagerung von Fahrzeugen auf eine hohe Produktion hinweist, lässt sich anhand der aktuellen Datenlage nicht plausibilisieren. Des Weiteren könnten auch hier je nach Standort und Hersteller die Gegebenheiten variieren. Trotz der Kriterien bei der

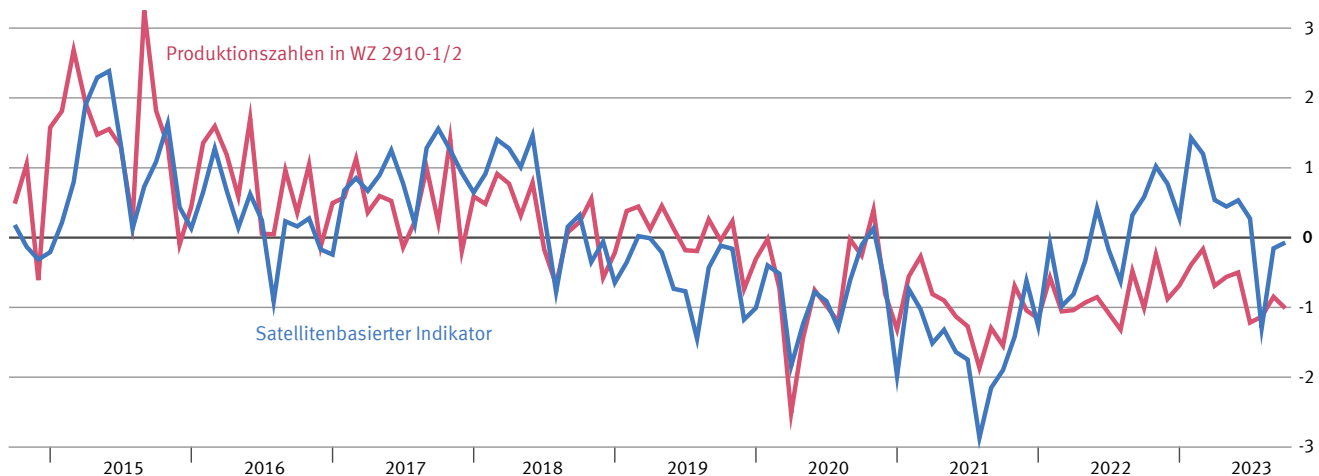
Auswahl der Parkplätze könnten weitere überdachte Produktionsparkplätze existieren, die sich in der Analyse nicht abbilden lassen und die nicht von anderen zum Standort gehörigen Gebäuden unterscheidbar sind. Auch die Komplexität des Radarsignals und fehlende Validierungsdaten erschweren die Auswertung. Die monatlich und räumlich aggregierten und normierten Werte erleichtern die Auswertung, jedoch wird durch sie auch die potenzielle Fluktuation der Fahrzeuge auf den Parkplätzen geglättet.

Der Indikator wurde aufgrund der großen Unterschiede zwischen den Produktionsstandorten in zwei Versionen in der ökonometrischen Schnellschätzung getestet. Neben der Version, die alle Standorte enthält und in Grafik 3 abgebildet ist, wurde eine zweite Version lediglich der Standorte erstellt, die in der standortspezifischen Validierung mit den Einzeldaten aus der Produktion einen Korrelationskoeffizienten nach Bravais und Pearson von über 0,6 aufweisen.<sup>9</sup> In der ökonometrischen Schnellschät-

9 Bei diesem Korrelationskoeffizienten wird die Kovarianz der Merkmale durch das Produkt der Standardabweichungen der Randverteilungen beider Merkmale dividiert. Er misst den linearen Zusammenhang zweier metrischer normalverteilter Merkmale.

### Grafik 3

**Monatliche Zeitreihen des satellitenbasierten Indikators und der gemeldeten Produktionszahlen<sup>1</sup>**  
z-transformierte Werte



Aufgrund der unterschiedlichen Skalen der Indikatoren wurden die Zeitreihen zur Darstellung z-transformiert.

1 Unterklassen 29.10.1 Herstellung von Personenkraftwagen und Personenkraftwagenmotoren sowie 29.10.2 Herstellung von Nutzkraftwagen und Nutzkraftwagenmotoren der Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).

zung können neue Datenquellen schnell und zielgerichtet getestet werden.<sup>10</sup> Die Ergebnisse zeigten, dass der Indikator für die Version der ausgewählten Standorte einen Effekt auf die Schnellschätzung der Bruttowertschöpfung im Kraftfahrzeug-Handel (Wirtschaftsabteilung G45 der Klassifikation der Wirtschaftszweige, Ausgabe 2008) hatte. Die Version des Indikators aller Standorte hatte jedoch keinen Effekt auf die Schätzung der Wertschöpfung. In Anbetracht der Berechnungsmethode und der verfügbaren Frequenz der Daten ist grundsätzlich problematisch, dass a priori nicht eindeutig ist, ob ein hoher monatlicher Mittelwert der Parkplatzbelegung auch tatsächlich einer hohen Produktion beziehungsweise einem hohen Umsatz mit Kraftfahrzeugen entspricht.

### 4.2 Bautätigkeitsindikator

Die Baubranche und Bautätigkeiten zählen zu den zentralen Triebkräften der deutschen Wirtschaft, da sie nicht nur erheblich zur Wertschöpfung und Beschäftigung beitragen, sondern auch als Indikator für Konjunktorentwicklung und gesellschaftlichen Wohlstand gelten. Um Entwicklungen in der Baubranche abzubilden, stehen in der Bautätigkeitsstatistik monatsweise die neu erteilten Baugenehmigungen bereit. Der tatsächliche Baubeginn wird jedoch aktuell nicht erfasst. Ebenso werden Baufertigstellungen oftmals erst zum Jahresende übermittelt (Schumann und andere, 2025). Hierdurch entsteht eine Informationslücke in der Abbildung der konjunkturellen Entwicklung, die durch die Informationen aus Satellitendaten gefüllt werden könnte.

Unter Verwendung maschineller Lernverfahren wurde im Projekt ein auf multispektralen Sentinel-2-Daten basierter, satellitengestützter Indikator entwickelt, welcher der räumlichen Erfassung und Quantifizierung von Bautätigkeit dient. Die Grundidee des Indikators ist, monatlich Landbedeckungsklassifikationen zu berechnen, aus denen dann die für die Bautätigkeit relevanten Informationen extrahiert werden können. Der Indikator wurde für fünf Gemeinden in Deutschland (München, Freiburg, Würzburg, Frankfurt am Main, Hamburg), in denen

nachweislich hohe Bautätigkeitsaktivitäten zu verzeichnen waren, von 2016 bis 2023 pilotiert.

Da die Satellitenkonstellation Sentinel-2 aus aktuell zwei Satelliten besteht, kann in den europäischen Breiten durch die Überlappung der Umlaufbahnen etwa alle zwei bis drei Tage mit einer Aufnahme gerechnet werden. Die Gebiete, die durch Sentinel-2 erfasst werden, können jedoch komplett oder teilweise von Wolken bedeckt sein, womit eine Erfassung der Erdoberfläche nicht möglich ist. Um Wolkenbedeckung aus den Daten herauszurechnen, werden monatliche Sentinel-2-Komposite, die als Basis für den Indikator dienen, im Deutschen Zentrum für Luft- und Raumfahrt vorprozessiert.<sup>11</sup> Da sich Wolken, Schnee und andere Witterungsbedingungen auch negativ auf die Datenqualität der Komposite auswirken und somit die komplette Prozessierung beeinflussen, blieben die Wintermonate (November bis Februar) in der Analyse unberücksichtigt.

Bis zur finalen Berechnung der Statistiken wird die geometrische Auflösung erhalten und die Analyse pixelbasiert durchgeführt. In der spektralen Dimension werden die Daten durch die Berechnung von Indizes ergänzt, die pixelweise aus den Spektralwerten der Kanäle berechnet werden. Da sich Oberflächen in ihrer spektralen Signatur unterscheiden (siehe Grafik 2), werden Indizes genutzt, um Unterschiede zwischen den Oberflächen hervorzuheben. Indizes fließen neben den Werten der spektralen Kanäle als ergänzende Variablen in den Klassifikationsprozess ein.

Die monatlichen Landbedeckungsklassifikationen werden mithilfe eines Machine-Learning-Ansatzes erstellt. Dabei wurden vier Zielklassen (Urban, Vegetation, offener Boden, Wasser) definiert und weitere Landbedeckungen herausgefiltert. Bei der Klasse „Urban“ handelt es sich vor allem um Gebäude und versiegelte Flächen. Die Klassen werden in den Sentinel-2-Kompositen mit einem Random-Forest-Klassifizierungsalgorithmus bestimmt. Da es sich bei Random-Forest um eine überwachte Klassifikation handelt, sind Trainingsdaten erforderlich. Monat-

10 Für die ökonomische Schnellschätzung wird in den Volkswirtschaftlichen Gesamtrechnungen eine R-Shiny-App eingesetzt, die es erlaubt, experimentelle Indikatoren zu testen.

11 Die Daten sind für Deutschland über den [EOC Geoservice](#) des Deutschen Zentrums für Luft- und Raumfahrt beziehbar (DOI: [10.15489/4hcq6dgkj648](#)). Die Prozessierung ist in Hagolle und andere (2018) dokumentiert.

liche Referenzdaten wurden automatisiert aus [Open-StreetMap](#)- und [Dynamic World](#)-Daten extrahiert, von denen jeweils 80 % für das Training des Modells und 20 % für die Validierung des Modells genutzt wurden. Die monatlich trainierten Modelle werden dann auf die monatlich aggregierten Satellitendaten angewendet.

Als einer der bedeutendsten Indizes im Klassifikationsprozess erwies sich der Normalized Difference Vegetation Index (NDVI), der aus der normierten Differenz zwischen dem sichtbaren roten und dem nahinfraroten Kanal berechnet wird. Dieser Index unterstützte das Lernmodell insbesondere bei der Erkennung von Vegetationsflächen und entsprechend dem Ausschluss versiegelter Bereiche:

$$NDVI = \frac{NIR - RED}{NIR + RED}$$

Die räumliche monatliche Landbedeckungsklassifikation wurde anschließend in einem nachgelagerten regelbasierten Bestätigungsschritt zur finalen Ausweisung der Landbedeckung überführt. Da in den monatlichen pixelbasierten Klassifikationen eine fehlerhafte Klasse nicht auszuschließen ist, werden nur diejenigen Flächen (zum Zeitpunkt  $t_0$ ) in die Klasse „Urban“ aufgenommen, die sich auch über die

folgenden drei Monate ( $t_{1-3}$ ) in der Klasse „Urban“ verstetigt haben.

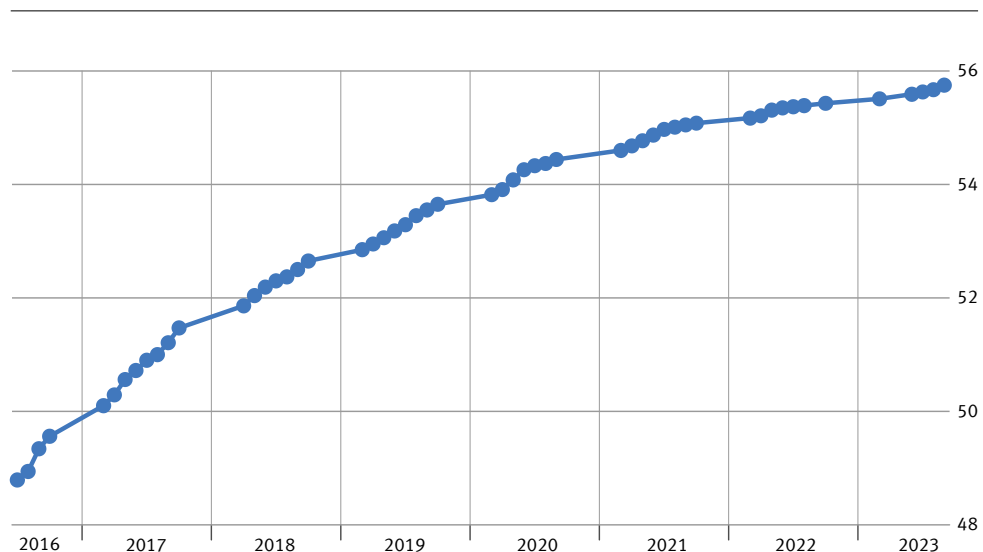
Um die Klassifikationen mit einer weiteren unabhängigen Datenquelle zu verifizieren, wurden Daten der „Land Use/Cover Area Frame Survey“ ([LUCAS](#)) genutzt, die für die Europäische Union einen harmonisierten Datensatz an In-situ-Daten liefern. Für die Klasse „Urban“ ergab sich hierbei eine Genauigkeit von 92 % im August 2018 und eine Genauigkeit von 90 % im August 2022.

Die final ausgewiesenen monatlichen Landbedeckungsklassifikationen können mithilfe des Untersuchungsgebietes in eine Tabellenstruktur überführt werden, indem die Pixel innerhalb einer Klasse gezählt werden. Da die Pixelgröße bekannt ist, kann so die Fläche ausgewiesen werden, die im Untersuchungsgebiet in eine bestimmte Klasse fällt. Daraus kann die monatliche Änderung der Fläche erfasst werden oder die Fläche ins Verhältnis zu der Gesamtfläche aller Pixel im Untersuchungsgebiet gesetzt werden. ➔ [Grafik 4](#)

Da die Klasse „Urban“ die Zielklasse für die Identifikation für Bautätigkeiten ist, wird im Falle der Verstetigung der Klassenzuordnung das entsprechende Pixel mit dem Datum der Änderung versehen. Das

**Grafik 4**

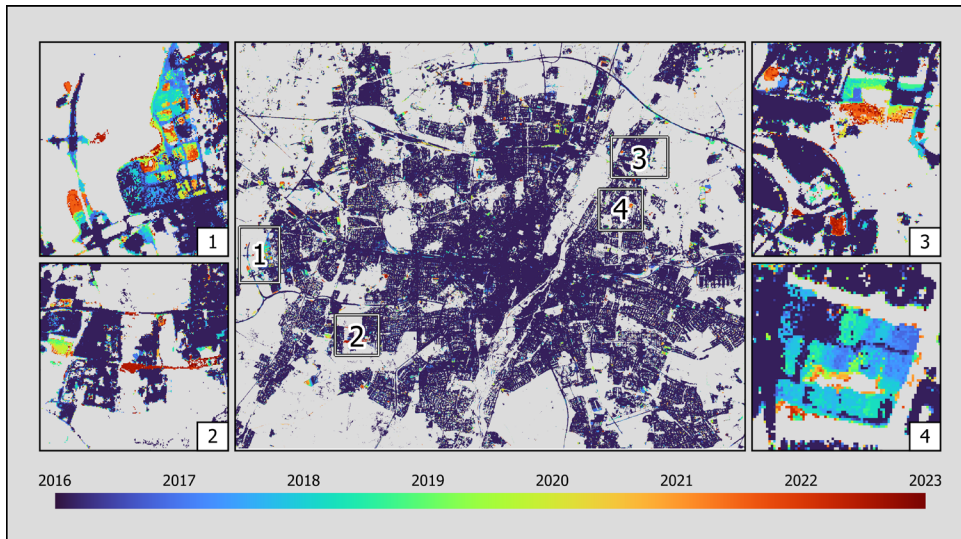
**Räumliche Statistik: Anteil der Bedeckungsklasse „Urban“ in der Stadt München im Zeitverlauf in %**





## Grafik 5

Zeitpunkt der Landbedeckungsänderung in die Klasse „Urban“ für das Stadtgebiet München



[1] bis [4] sind Vergrößerungen einzelner Abschnitte der Karte.

Ergebnis ist in [Grafik 5](#) dargestellt. Die Farbkodierung stellt dabei den Zeitpunkt der Veränderung zu städtischen Flächen dar, wobei die urbanen Flächen, die bereits im Mai 2016 bestanden, dunkelblau dargestellt sind. Bei den Kartenausschnitten 1 und 4 lassen sich anhand dieser Farbcodierung der Beginn einzelner Bauabschnitte erkennen.

Der entwickelte Prototyp des Indikators zeigt in diesen Auswertungen sein Potenzial und wäre bei Weiterentwicklung neben dem Einsatz in der ökonomischen Schnellschätzung auch für interne Validierungsarbeiten in der Bautätigkeitsstatistik interessant (Wurm und andere, 2025).

Hinsichtlich der eindeutigen Zuordnung der detektierten Veränderungen zu genehmigten Bautätigkeiten sind die Möglichkeiten durch die fehlende Geolokalisierung jedoch begrenzt. Ein Einsatz der Methode wäre jedoch direkt in den Bauämtern möglich, in denen die Baugenehmigungen mit geographischen Merkmalen erfasst werden.

## 5

### Fazit und Zukunftsperspektiven

Die Entwicklung der vorgestellten Indikatoren zeigt Potenziale sowie bestehende Hürden auf. Im Zuge des Projektes zeigte sich, dass der Zusammenhang zwischen der Automobilproduktion oder den Umsätzen in diesem Bereich und der Parkplatzbelegung deutlich komplexer ist, als es der Indikator darstellen kann. Zudem existieren alternative Indikatoren, die eine höhere Prognosequalität aufweisen und als Zuschätzer in der ökonomischen Schnellschätzung des Bruttoinlandsprodukts genutzt werden (zum Beispiel Kraftfahrzeug-Produktion oder -Zulassungen des Kraftfahrt-Bundesamtes). Der Indikator könnte jedoch in Ländern zum Einsatz kommen, in denen alternative Quellen nicht zur Verfügung stehen.

Gleichwohl eröffnet die Nutzung von Satellitendaten in anderen Sektoren – etwa im Bereich des untersuchten Bauwesens, aber auch in möglichen anderen Wirtschaftssektoren, beispielsweise in der Landwirtschaft oder im Bereich von Energieinfrastrukturen (Zhai und andere, 2024) – Perspektiven für die Entwicklung zusätzlicher Wirtschaftsindikatoren.



Die monatliche satellitenbasierte Detektion von potenzieller Bautätigkeit liefert eine neue Informationsebene zum Start von Bautätigkeiten, die aktuell in der Bautätigkeitsstatistik nicht erfasst werden. Der Bautätigkeitsindikator ist jedoch in seiner aktuellen Form in der ökonometrischen Schnellschätzung des Bruttoinlandsprodukts nicht einsetzbar. Zum einen bildet er mit den sechs Pilotgebieten nur einen geringen Teil der Gesamtfläche Deutschlands ab. Zum anderen verringert die nachgeschaltete regelbasierte Ausweisung der urbanen Flächen ( $t_{1-3}$ ) die Aktualität. Eine Weiterentwicklung des Indikators sollte an diesen Punkten ansetzen.

Die erstmals in dem Projekt genutzte Infrastruktur des Copernicus Data Space Ecosystem zeigt, dass hier ein enormes Potenzial für die automatisierte und dezentralisierte Verarbeitung von Satellitendaten liegt. Besonders wenn, wie am Beispiel des Automobilindicators, eine überschaubare Anzahl an zu beobachtenden Standorten vorliegt, kann dies die Verarbeitung erheblich erleichtern. Die lokale Speicherung der Satellitendaten entfällt. Die Überführung der Informationen aus den Satellitendaten in einfache Tabellenstrukturen erlaubt es dann, die Informationen in gängiger Statistiksoftware und in geschützten Umgebungen weiterzuverarbeiten.

Für die künftige Forschung ergibt sich daraus die Aufgabe, die methodische Integration von Fernerkundungsdaten in bestehende Indikatorensysteme systematisch zu prüfen. Von besonderer Bedeutung ist hierbei die Validierung der Aussagekraft im Vergleich zu etablierten konjunkturellen Kennzahlen sowie die Untersuchung, in welchen Teilbereichen diese Daten einen substanziellen Mehrwert liefern können. Hierfür ist die methodische Weiterentwicklung zentral, insbesondere im Hinblick auf die Verknüpfung von Fernerkundungsdaten mit bestehenden amtlichen und administrativen Datenquellen. Zudem sollte die Validierung der Aussagekraft in unterschiedlichen Sektoren systematisch erfolgen, um Stärken und Grenzen der Datenbasis transparent zu machen.

Schließlich ist auch die Zusammenarbeit mit anderen Institutionen, wie dem Projektpartner Deutsches Zentrum für Luft- und Raumfahrt, und Datenanbietern ein wichtiges Feld, um Synergien zu nutzen und eine nachhaltige Integration in die amtliche Statistik

zu gewährleisten. Aspekte wie die Automatisierung der Datenverarbeitung, die Reproduzierbarkeit der Ergebnisse und eine langfristige Verfügbarkeit der Satellitendaten werden auf europäischer Ebene mit der weiteren Entwicklung der Sentinel-Satellitenmissionen sowie der Bereitstellung des Copernicus Data Space Ecosystem erfolgreich adressiert. Der kontinuierliche Austausch im Europäischen Statistischen System und die umfangreiche Förderung seitens Eurostat von spezifischen nationalen Projekten zur Integration von Fernerkundungsdaten in die amtliche Statistik trägt ebenso zur methodischen Weiterentwicklung bei.

Perspektivisch werden Fernerkundungsdaten eine wertvolle Ergänzung bestehender Indikatorensysteme darstellen, indem sie punktuell dazu beitragen, die Datenverfügbarkeit zu beschleunigen und die Analyse bestimmter Wirtschaftsbereiche zu erweitern. 

## LITERATURVERZEICHNIS

---

Ackermann, Arne/Dickopf, Xaver/Mucha, Tanja. *Flash und Nowcast: Schnellschätzungen des Bruttoinlandsprodukts in der Corona-Pandemie*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2021, Seite 17 ff.

Drouyer, Sébastien/de Franchis, Carlo. *Parking occupancy estimation on Sentinel-1 images*. In: Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Volume V-2-2020, Seite 821 ff.

DOI: [10.5194/isprs-annals-V-2-2020-821-2020](https://doi.org/10.5194/isprs-annals-V-2-2020-821-2020)

Europäische Union. *Copernicus Broschüre – Copernicus – Europas Blick auf die Erde*. DOI: [10.2873/65425](https://doi.org/10.2873/65425)

Hagolle, Olivier/Morin, David/Kadiri, Mohamed. *Detailed Processing Model for the Weighted Average Synthesis Processor for Sentinel-2*. 2018. In: Zenodo.

DOI: [10.5281/zenodo.1401360](https://doi.org/10.5281/zenodo.1401360)

Hennig, Florian/Köhlmann, Maren/Weißmann, Julius/Schepers, Marianne. *Erforschung von Satelliten- und weiteren Fernerkundungsdaten zur Ermittlung von Gebäudeangaben*. In: WISTA Wirtschaft und Statistik. Ausgabe 4/2025, Seite 70 ff.

Kondmann, Lukas/Taubenböck, Hannes/Zhu, Xiao Xiang. *Blinded by the Light: Monitoring Local Economic Development Over Time With Nightlight Emissions*. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS. Seite 5708 ff. DOI: [10.1109/IGARSS47720.2021.9554428](https://doi.org/10.1109/IGARSS47720.2021.9554428)

Kraft, Franziska/Martinis, Sandro/Krullikowski, Christian/Plank, Simon/Angheloa, Anca/Delgado Blasco, José Manuel/Schönenberger, Klara/Köhlmann, Maren/Brauchler, Melanie. *Satellite Data for Economic Insights: Towards Tracking Automotive Production in Germany With Sentinel-1 for Economic Nowcasting*. In: IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing. Ausgabe 18/2025. DOI: [10.1109/JSTARS.2025.3601351](https://doi.org/10.1109/JSTARS.2025.3601351)

Schumann, Carsten/Schepers, Marianne/Weigert, Alexander. *Digitalisierungspotenzial in der Bautätigkeitsstatistik*. In: WISTA Wirtschaft und Statistik. Ausgabe 3/2025, Seite 74 ff.

Spelta, Alessandro/Pagnottoni, Paolo/Pecora, Nicoló. *Big brothering the economy: nowcasting and forecasting with port satellite images*. In: Journal of the Royal Society Series A: Statistics in Society. 2025. DOI: [10.1093/jrssa/qnae157](https://doi.org/10.1093/jrssa/qnae157)

Stundziene, Alina/Pilinkiene, Vaida/Bruneckiene, Jurgita/Grybauskas, Aandrius/Lukauskas, Mantas/Pekarskiene, Irena. *Future directions in nowcasting economic activity: A systematic literature review*. In: Journal of Economic Surveys. Jahrgang 38. Ausgabe 4/2024, Seite 1199 ff. DOI: [10.1111/joes.12579](https://doi.org/10.1111/joes.12579)

## LITERATURVERZEICHNIS

---

Taubenböck, Hannes/Mast, Johannes/Geiß, Christian/Wurm, Michael/Esch, Thomas/Seto, Karen C. *Global differences in urbanization dynamics from 1985 to 2015 and outlook considering IPCC climate scenarios*. 2024. In: *Cities*. Band 151. Artikel 105117. DOI: [10.1016/j.cities.2024.105117](https://doi.org/10.1016/j.cities.2024.105117)

Wuepper, David/Oluoch, Wyclife Agumba/Hadi, Hadi. *Satellite Data in Agricultural and Environmental Economics: Theory and Practice*. In: *Agricultural Economics*. Jahrgang 56. Ausgabe 3/2025, Seite 493 ff. DOI: [10.1111/agec.70006](https://doi.org/10.1111/agec.70006)

Wurm, Michael/Köberl, Manuel/Cevallos, Gabriel/Köhlmann, Maren/Brauchler, Melanie/Schumann, Carsten/Taubenböck, Hannes. *Satellite-based fine-grained spatio-temporal monitoring of urban building activities as an indicator of economic development*. In: 2025 Joint Urban Remote Sensing Event (JURSE). DOI: [10.1109/JURSE60372.2025.11075972](https://doi.org/10.1109/JURSE60372.2025.11075972)

Zhai, Yichen/Chen, Xuehong/Cao, Xin/Cui, Xihong. *Identifying wind turbines from multiresolution and multibackground remote sensing imagery*. In: *International Journal of Applied Earth Observation and Geoinformation*. 2024. Ausgabe 126. Artikel 103613. DOI: [10.1016/j.jag.2023.103613](https://doi.org/10.1016/j.jag.2023.103613)

# METADATENMODELL DER AMTLICHEN STATISTIK

Sven Claußen

➤ **Schlüsselwörter:** Metadaten – Metadatenmanagement – Standardisierung – Digitalisierung – Automatisierung

## ZUSAMMENFASSUNG

Für die Statistikproduktion ist ein gut funktionierendes, übergreifendes Management von Metadaten von großer Bedeutung, um die notwendigen Geschäftsprozesse qualitativ hochwertig und effizient umzusetzen. Mithilfe einer zentralen Systemkomponente werden dazu verschiedene Teilsysteme in einer Gesamtarchitektur für Metadaten miteinander kombiniert. Diese dient in den Produktionsprozessen sowohl als zentrale Bezugsquelle als auch zur Speicherung von Metadaten. Konzeptionelle Grundlage ist hierbei das Metadatenmodell der amtlichen Statistik (MedaS), das auf einem internationalen Standard basiert und 2025 in einer neuen Version im Metadatenmanagementsystem implementiert wurde. MedaS bietet darüber hinaus die Grundlage für eine verbesserte Kommunikation zwischen Fachbereichen, Informationstechnik und Managementebene.

➤ **Keywords:** metadata – metadata management – standardisation – digitalisation – automation

## ABSTRACT

*In statistics production, a well-functioning and comprehensive metadata management system is crucial to ensure that the necessary business processes are implemented efficiently and to a high quality standard. A central system component is used to combine various individual solutions into an overall metadata architecture to this end. In the production processes, it is used both as a central source of reference and for storing metadata. The conceptual framework for this is the “Metadata Model of Official Statistics” (MedaS), which is based on an international standard. A new version of MedaS was implemented in the metadata management system in 2025. MedaS also facilitates communication among specialised statistics units, IT experts and management.*



**Dr. Sven Claußen**

ist Diplom-Informatiker und Referent im Referat „Metadaten, Qualität und Datenmanagement“ des Statistischen Bundesamtes. Zu seinen Arbeitsschwerpunkten zählen die Konzeption und Qualitätssicherung von Methoden und Verfahren zur Verwaltung von Metadaten.

## 1

### Einleitung

Die Statistikproduktion wird durch eine Vielzahl von IT-Werkzeugen unterstützt, die untereinander Daten und Metadaten austauschen. Wichtige Basis für eine erfolgreiche Interoperabilität ist ein standardisiertes Informationsmodell, das ein gemeinsames Verständnis für die notwendigen Informationsobjekte sowie deren Eigenschaften und Beziehungen zueinander vermittelt. Im Statistischen Verbund<sup>1</sup> wird hierzu [das Metadatenmodell der amtlichen Statistik \(MedaS\)](#) eingesetzt, das auf einem internationalen Standard der Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) aufsetzt. Im Zusammenhang mit dem [Geschäftsprozessmodell Amtliche Statistik \(GMAS\)](#) wird mithilfe von MedaS definiert, welche Objekte in jedem Teilprozess benötigt und produziert werden. MedaS hat auch außerhalb der Informationstechnik (IT) eine besondere Bedeutung, da sich durch einen einheitlichen Sprachgebrauch die Kommunikation auf verschiedenen Ebenen verbessert.

Der Beitrag stellt das Metadatenmodell der amtlichen Statistik (MedaS) vor. Adressaten sind beispielsweise diejenigen, die in statistischen Organisationen arbeiten. Er skizziert die Vorteile und den Nutzen des Modells für die Statistischen Ämter des Bundes und der Länder, wenn es sowohl fachlich als auch technisch flächendeckend angewendet wird.

Nach einem Überblick über den Modernisierungsbedarf, den Stand der Standardisierung und das Metadatenmanagement des Statistischen Verbundes in Kapitel 2 erläutert Kapitel 3 ausführlich das Metadatenmodell der amtlichen Statistik. Wie konkret die Anwendung von MedaS als Basis für die Architektur des Metadatenmanagements im Statistischen Verbund erfolgt, wird in Kapitel 4 thematisiert. Ein Fazit in Kapitel 5 rundet den Artikel ab.

---

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

## 2

### Sachstand und Weiterentwicklung

#### 2.1 Modernisierungsbedarf

Die Statistikproduktion steht zunehmend vor der Herausforderung, bei schwindenden Ressourcen weiterhin eine gute Datenqualität zu gewährleisten. Höhere Ansprüche, insbesondere an die Aktualität, sind unter diesen Rahmenbedingungen nicht mehr ohne Weiteres zu erfüllen. Erst durch Investitionen in die Modernisierung der Produktionsprozesse lassen sich langfristig und nachhaltig in größerem Umfang Kosten einsparen, ohne die bestehenden Qualitätsansprüche zu vernachlässigen.

Die Modernisierung umfasst eine übergreifende Standardisierung und Kopplung von Standardwerkzeugen, um das vorhandene Automatisierungspotenzial möglichst effizient zu nutzen (Statistisches Bundesamt, 2022). Anstelle von manuellen Export- und Importvorgängen werden benötigte Daten und Metadaten ohne Zusatzaufwand direkt zwischen den Komponenten ausgetauscht. Dies setzt eine IT-Architektur voraus, die verteilte Systeme auf Basis von Standards miteinander koppelt (Lindenstruth/Claußen, 2017; Blumöhr und andere, 2017). Die inhaltlichen sowie fachlichen Standards sind Voraussetzung für Kompatibilität und Austausch der Informationsobjekte in passenden Formaten, um diese automatisiert verarbeiten zu können.

#### 2.2 Standardisierung

Auf der ganzen Welt erledigen statistische Organisationen ähnliche Aufgaben. Die dabei eingesetzten Prozesse unterscheiden sich zwar im Detail, jede dieser Aktivitäten verwendet und produziert aber ähnliche Informationen (zum Beispiel definieren alle Organisationen Grundgesamtheiten für ihre statistischen Beobachtungen, verwenden statistische Klassifikationen, erstellen Datensätze und verbreiten Informationen). Grundlage der Statistikproduktion sind insbesondere die international abgestimmten

Standards [Generic Statistical Business Process Model \(GSBPM\)](#) und [Generic Statistical Information Model \(GSIM\)](#), die in Arbeitsgruppen der UNECE mit Beteiligung des Autors in den Jahren 2007 bis 2019 entwickelt wurden (UNECE, 2025). Das GSBPM definiert den gesamten Geschäftsprozess der Erstellung einer Statistik von der Planung bis zur Auswertung anhand von acht Phasen und 44 Teilprozessen (Gehle/Lüüs, 2017). Das GSIM standardisiert alle Informationsobjekte, die bei der Gestaltung und Erstellung von Statistiken notwendig sind, durch eine konsistente Beschreibung und die Beziehungen untereinander. Vor einer nationalen Adaptierung und technischen Umsetzung sind aufgrund bewusst vorhandener Freiheitsgrade insbesondere die Objektattribute umfassender zu spezifizieren.

Der Statistische Verbund hat GSBPM und GSIM in eigenen Standards adaptiert. GMAS und MedaS sind nicht nur die deutschen Übersetzungen der internationalen Modelle, sondern passen sie durch Präzisierungen und Ergänzungen auch an die Bedürfnisse und Gepflogenheiten der amtlichen Statistik in Deutschland an. Die Verwendung standardisierter Begrifflichkeiten und zusätzliche Beispiele erleichtern das Verständnis von GMAS und MedaS. Gleichwohl bleiben Bedeutungen der adaptierten Geschäftsprozesse und Metadatenobjekte sowie ihre Beziehungen untereinander bestehen, um auch langfristig die Kompatibilität zu GSBPM und GSIM zu sichern.

### 2.3 Metadatenmanagement der Statistischen Ämter des Bundes und der Länder

---

Das Metadatenmanagement des Statistischen Verbundes basiert auf GMAS und MedaS, die sich sinnvoll ergänzen. Der Fokus dieses Artikels liegt auf der einführenden Beschreibung von MedaS, das 2025 in der Version 2.0 erstmals umfassend alle Informationsobjekte der Statistikproduktion enthält. Von diesem Standard profitieren nicht nur IT-Verantwortliche, sondern alle am Produktionsprozess Beteiligten, da MedaS (in Kombination mit GMAS) statistik- und organisationsübergreifend die gemeinsame Sprache für die Beschreibung der Statistikproduktion im Statistischen Verbund bildet. Trotz eingedeutsch-

ter Begrifflichkeiten sind die integrierten Verweise auf GSIM auch bei der internationalen Zusammenarbeit, zum Beispiel bei Studienbesuchen, Kooperationen sowie Sitzungen des Statistischen Amtes der Europäischen Union (Eurostat) und der UNECE, von großem Nutzen.

Die Aufgabe, geeignete technische Infrastrukturkomponenten für das Management von Metadaten bereitzustellen, obliegt dem Statistischen Bundesamt. Eine der wichtigsten Komponenten ist hierbei das Metadatenmanagementsystem, das als zentrale Drehscheibe die Speicherung, Verteilung und Standardisierung aller notwendigen Metadatenobjekte übernimmt. Langjährig bewährte metadatenhaltende Systeme wie die Statistikdatenbank, die Erhebungsdatenbank oder der Klassifikationsserver werden dabei nicht ersetzt, sondern deren Metadaten (zum Beispiel Klassifikationen) mithilfe sogenannter Adapter übernommen und in das standardisierte MedaS transformiert. Die Zwischenspeicherung und Bereitstellung über zentrale Standardschnittstellen gewährleisten einen einfachen, dauerhaften Zugriff und eine hohe Verfügbarkeit notwendiger Metadaten. Für den weiteren Ausbau – um letztlich alle MedaS-Objekte über moderne Bedienoberflächen verwalten zu können – kann das Metadatenmanagementsystem sukzessive um sogenannte Anwendungsmodule erweitert werden. Diese lassen sich insofern schlank und ressourcengünstig entwickeln, als die Speicherung der Inhalte direkt im Metadatenmanagementsystem erfolgen kann. Eine eigene Datenbank in den Anwendungsmodulen sowie aufwendige Modelltransformationen sind nicht notwendig.



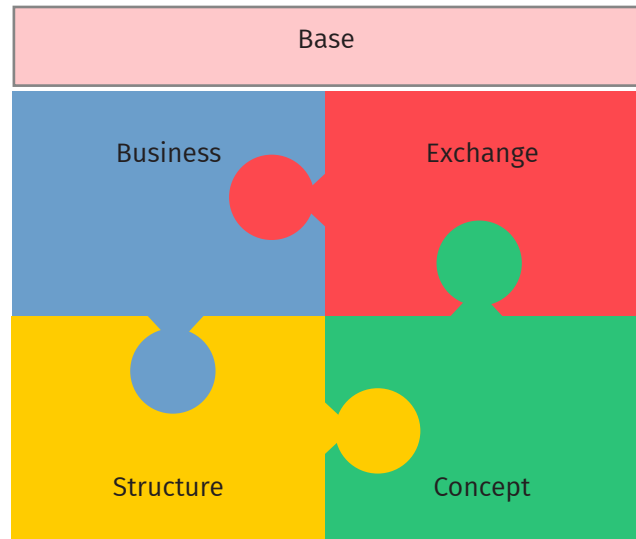
## 3

## Das Metadatenmodell der amtlichen Statistik

Ein Informationsmodell beschreibt alle Objekte, die Informationen über die reale Welt spezifizieren – sogenannte Informationsobjekte – und deren Beziehungen zueinander. Für spezielle Bedürfnisse wie die Statistikproduktion kann es spezialisierte Informationsmodelle geben, wie GSIM oder MedaS, die auf die Verarbeitung und Verwaltung von Daten und Metadaten in statistischen Organisationen fokussieren. MedaS als deutsche Adaption des internationalen Standards GSIM wurde im Wesentlichen an den deutschen Sprachgebrauch angepasst und um einige wenige Objekte sowie zahlreiche Objektattribute ergänzt. Die Einteilung aller Informationsobjekte in

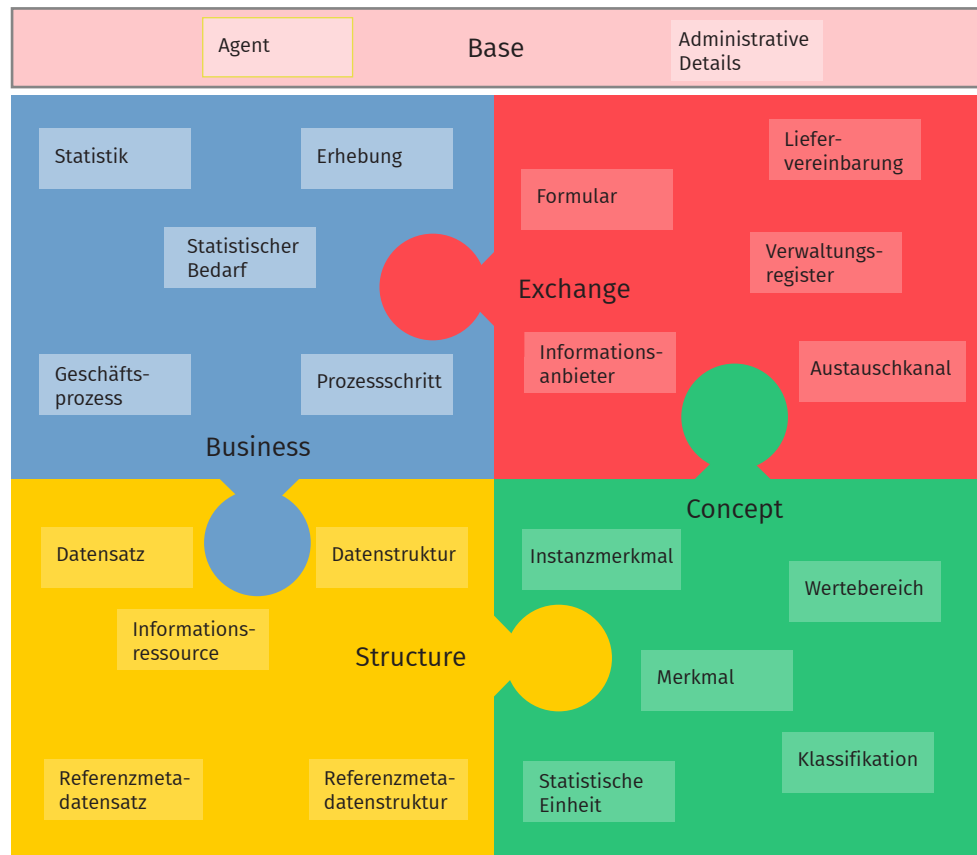
Grafik 1

Die fünf Hauptgruppen von Informationsobjekten im Metadatenmodell der amtlichen Statistik



Grafik 2

Auswahl einiger Informationsobjekte mit Zuordnung zu den fünf Hauptgruppen im Metadatenmodell der amtlichen Statistik



fünf Hauptgruppen wurde einschließlich der Originalbezeichnungen und farblichen Kennzeichnungen beibehalten (UNECE, 2025; Lindenstruth/Claußen, 2017). [↗ Grafik 1](#)

Die aktuelle Version des Metadatenmodells der amtlichen Statistik MedaS 2.0 enthält 131 Informationsobjekte, die den fünf Hauptgruppen zugeordnet sind. [↗ Grafik 2](#)

In einem Informationsmodell stehen die Informationsobjekte miteinander in Beziehung.

[↗ Grafik 3](#) zeigt die folgenden Zusammenhänge zwischen den Informationsobjekten:

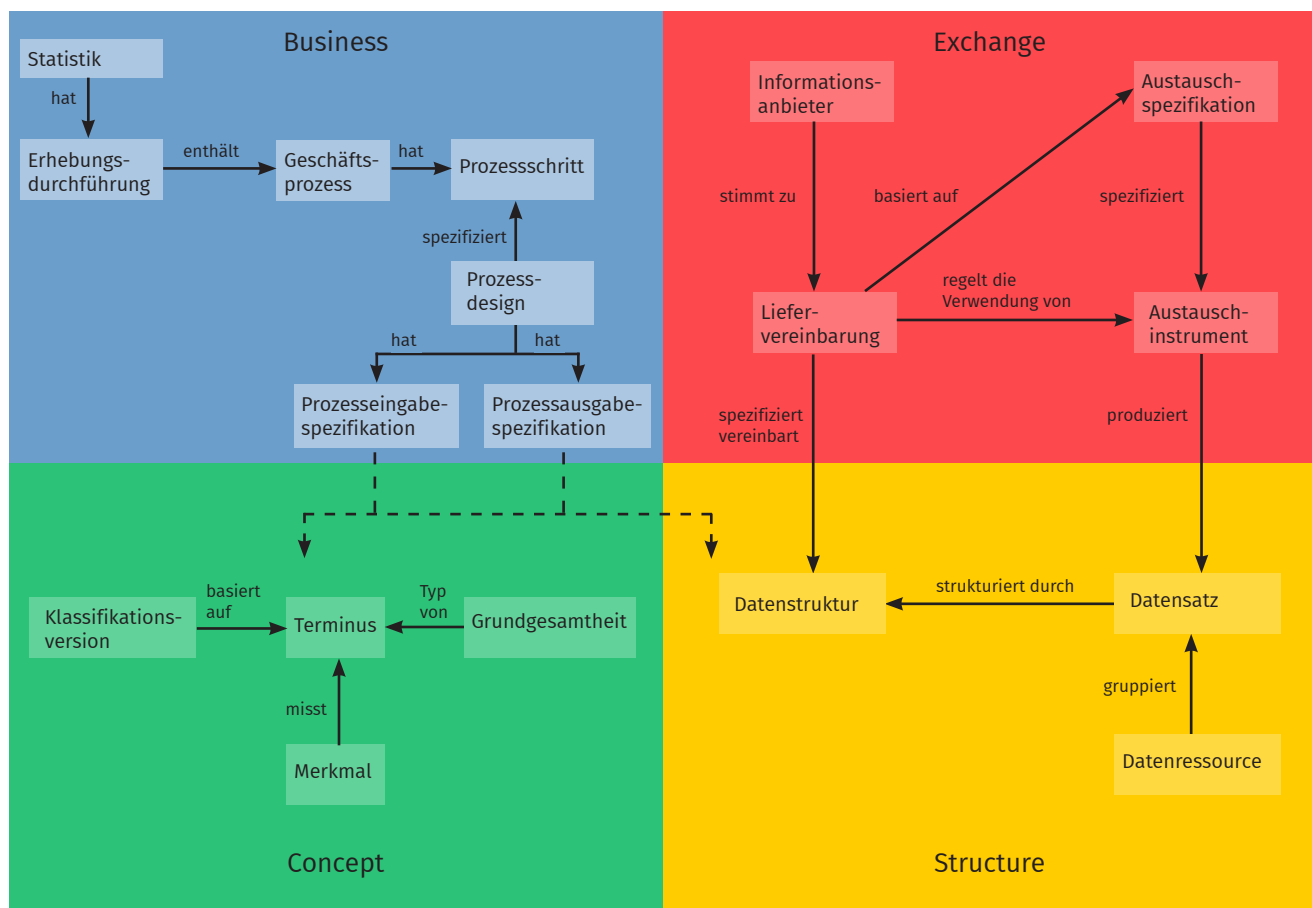
Eine statistische Organisation initiiert eine Statistik. Die Statistik entspricht einer fortlaufenden Aktivität wie einer Erhebung oder einer Ergebnisreihe und

hat eine Erhebungsdurchführung, beispielsweise vierteljährlich oder jährlich. Die Erhebungsdurchführung umfasst eine Reihe von Geschäftsprozessen. Die Geschäftsprozesse bestehen aus mehreren Prozessschritten, die durch ein Prozessdesign spezifiziert werden. Diese Prozessdesigns haben Prozesseingangsspezifikationen und Prozessaussgangsspezifikationen. Bei den Spezifikationen handelt es sich häufig um Informationen, die sich auf Konzepte und Strukturen beziehen (zum Beispiel Klassifikation, Merkmal, Grundgesamtheit, Datenstruktur und Datensatz).

Bezieht sich der Geschäftsprozess beispielsweise auf die Erhebung von Daten, gibt es einen Informationsanbieter, der auf Basis einer Liefervereinbarung der statistischen Organisation Daten zur Verfügung stellt. Diese Liefervereinbarung legt eine vereinbarte

### Grafik 3

Beziehungen zwischen den Informationsobjekten im Metadatenmodell der amtlichen Statistik



Datenstruktur fest und regelt den Einsatz eines Austauschinstrument, das für die eingehenden Informationen verwendet wird. Das Austauschinstrument könnte in diesem Fall ein Fragebogen oder ein Verwaltungsregister sein, das durch die Austauschspezifikation definiert wird.

Der vom Austauschinstrument produzierte Datensatz wird in einer Datenressource (gegebenenfalls mit anderen Datensätzen) gruppiert und durch eine Datenstruktur strukturiert.

Dieses Beispiel zeigt, dass MedaS keine statistik- oder erhebungsspezifischen Bezeichnungen verwendet; dadurch ist es übergreifend einsetzbar. Es abstrahiert von Besonderheiten einer Erhebungsdurchführung und deckt dennoch die Vielfalt der Statistikproduktionen ab. So verbergen sich hinter einem Informationsanbieter alle Auskunftseinheiten, ob mit oder ohne Auskunftspflicht, ob in einer Stichprobe oder einer Vollerhebung, ob über eine Online-Befragung, ein Verwaltungsregister oder über einen klassischen Papierfragebogen. Erst durch die Befüllung der abstrakten Objekte konkretisiert sich eine Statistikproduktion, beispielsweise durch Festlegung eines Merkmalsträgertyps mit „Person“, der Definition der Grundgesamtheit mit „Personen über 18 Jahre mit Wohnsitz in Deutschland“ und der Auflistung der zugehörigen (etwa 70 Millionen) statistischen Einheiten. Als Austauschinstrument ist gegebenenfalls ein zentrales Melderegister vorgesehen. Liegen die notwendigen technischen und juristischen Voraussetzungen vor, können Werkzeuge der Statistikproduktion auf Basis der hinterlegten Datenstruktur entsprechende Datensätze aus diesem Melderegister automatisiert entnehmen und auf Basis der weiteren Metadaten weiterverarbeiten (einschließlich Geheimhaltung, Imputation und so weiter).

An dieser Stelle sei ausdrücklich darauf hingewiesen, dass MedaS kein fachliches Modell der zu verarbeitenden Daten ist. So hat beispielsweise eine Person ein Geschlecht, ein Geburtsdatum und einen Geburtsort. Ein Unternehmen ist beispielsweise durch einen Unternehmenstyp, eine Geschäftsadresse und eine Anzahl Beschäftigter (die wiederum Personen sind) gekennzeichnet. Diese Zusammenhänge obliegen einem semantischen Daten-

modell über die in der Statistik zu betrachtenden Merkmalsträgertypen, das gesondert zu definieren und innerhalb des Statistischen Verbundes aufgrund der vorhandenen Verzahnungen statistikübergreifend abzustimmen wäre.

Zwar ist in diesem Aufsatz keine detaillierte Definition aller MedaS-Objekte möglich, nachfolgend soll allerdings eine vollständige Auflistung mit einer leichtverständlichen Kurzbeschreibung das Verständnis fördern:

### 3.1 Base

---

Die Gruppe „Base“ stellt Informationen bereit, die von anderen Informationsobjekten wiederverwendet werden können, um Funktionen wie Identifikation und Versionierung zu unterstützen:

- > **Administrative Details:** speichert insbesondere technische Angaben (Status, Zeitpunkt der Erstellung und so weiter)
- > **Adresse:** umfasst räumliche Informationen, um einen Akteur lokalisieren zu können
- > **Agent:** stellt einen Akteur dar und wird durch folgende Untertypen konkretisiert:
  - > **Organisation:** eine Ansammlung von Menschen, die gemeinsam organisiert sind
  - > **Person:** Einzelpersonen, die Metadaten bearbeiten und nutzen
  - > **Software:** Akteur, insbesondere für automatisierte Prozessabläufe
- > **Agent in einer Rolle:** Verknüpfung eines Akteurs mit einer bestimmten Rolle bezüglich eines konkreten Informationsobjekts
- > **Änderungsereignis:** bestimmtes Ereignis, das einen konkreten Einfluss auf Informationsobjekte hat
- > **Artefakt:** Basisobjekt für (nahezu) alle Informationsobjekte, um grundlegende Objekteigenschaften zu bündeln
- > **Attribut:** Eigenschaft eines Objekts
- > **i18n-Text:** dient der Verwaltung mehrsprachiger Texte

- > **Informationsobjekt:** dient der Modellierung des eigenen Informationsmodells
- > **Rolle:** Akteure können je nach Kontext verschiedene Rollen einnehmen:
  - > **Besitzer:** verantwortlich für die Definition, Spezifizierung, Erstellung und Pflege eines Objekts
  - > **Betreuer:** kümmert sich um administrative und betriebliche Angelegenheiten, aber ohne Entscheidungsbefugnis
  - > **Kontakt:** verantwortlich für die Bereitstellung zusätzlicher Informationen über ein Objekt

### 3.2 Business

---

Die Gruppe „Business“ wird verwendet, um die Entwürfe und Pläne von Statistiken sowie die Prozesse zu erfassen, die zur Bereitstellung statistischer Ergebnisse durchgeführt werden. Dies umfasst auch die Identifizierung eines statistischen Bedarfs, die Geschäftsprozesse einer Statistik sowie deren Evaluierung.

- > **Änderungsdefinition:** Ergebnis konzeptioneller Arbeiten aufgrund eines statistischen Bedarfs
- > **Bewertung:** Evaluation zur Verbesserung von Produktionsprozessen (und Produkten) im Hinblick auf den statistischen Bedarf
- > **Erhebung:** spezifiziert Prozesse und Objekte für konkrete Erhebungsdurchführungen unter Berücksichtigung der spezifischen Anforderungen
- > **Erhebungsdurchführung:** beschreibt die konkrete Ausführung einer Iteration einer Erhebung für einen bestimmten Berichtszeitraum
- > **Geschäftsdienst:** ermöglicht den Zugriff auf eine bestimmte Geschäftsfunktion
- > **Geschäftsfall:** begründet die Durchführung eines Statistikunterstützungsprogramms, um eine neue Erhebung für eine bestehende oder neue Statistik zu initiieren
- > **Geschäftsfunktion:** definiert den Geschäftszweck, dem ein Geschäftsdienst oder ein Prozessschritt dient
- > **Geschäftsprozess:** strukturierte und wiederholbare Aktivität, die eine oder mehrere Geschäftsfunktionen erfüllt
- > **Prozessausgabe:** neue oder veränderte Objekte, die bei der Ausführung eines Prozessschritts entstehen (nutzbar als Prozesseingabe für andere Prozessschritte):
  - > **Kernausgabe:** Objekt, für das der Prozessschritt maßgeblich dient (typischerweise die Kerneingabe für einen nachfolgenden Prozessschritt oder das Endprodukt der Statistikproduktion)
- > **Prozessausführungsprotokoll:** zeichnet technische Informationen über die Ausführung eines Prozessschritts auf, um Prozessabläufe und etwaige Störungen zu dokumentieren
- > **Prozessmetrik:** zeichnet qualitative und inhaltliche Informationen über die Ausführung eines Prozessschritts auf, um das Qualitätsmaß der Kernausgabe beurteilen zu können
- > **Prozessausgabespezifikation:** definiert die Prozessausgaben für die Ausführung eines Prozessdesigns
- > **Prozessdesign:** definiert die Ausführung eines Prozessschritts als Teil eines bestimmten Geschäftsdienstes
- > **Prozesseingabe:** Objekte, die zur Ausführung eines Prozessschritts benötigt werden:
  - > **Kerneingabe:** Objekt, für das der Prozessschritt maßgeblich dient (das beispielsweise transformiert oder als Grundlage für die Erzeugung neuer Objekte verwendet wird)
  - > **Parametereingabe:** dient der Konfiguration generischer Prozessschritte, um sich an spezifische Bedürfnisse anzupassen
  - > **Prozessunterstützungseingabe:** Objekt, das für den Zweck des Prozessschritts hilfsweise benötigt wird
- > **Prozesseingabespezifikation:** definiert die notwendigen Prozesseingaben für die Ausführung eines Prozessdesigns
- > **Prozessmethode:** beschreibt unabhängig von Technologien und Werkzeugen eine bestimmte

Methodik zum Ausführen eines Prozessschritts (gegebenenfalls Anwendung bestimmter Regeln)

- > **Prozessmuster:** Definition einer Vorlage zur Beschleunigung von Designprozessen und zur gemeinsamen Nutzung und Wiederverwendung von Designmustern, die sich als effektiv erwiesen haben
- > **Prozessschritt:** implementiert das definierte Prozessdesign, um die gewünschten Ergebnisse zu erzeugen
- > **Prozessschrittinstanz:** führt einen Prozessschritt in einem bestimmten Kontext mit tatsächlichen Instanzen der Eingaben (zum Beispiel bestimmte Datensätze) aus, sodass einzigartige Ergebnisse erzielt werden, auch wenn der Prozessschritt konstant bleibt
- > **Prozesssteuerung:** steuert die Ausführung der Prozessschritte gemäß Prozesssteuerungsdesign
- > **Prozesssteuerungsdesign:** definiert den Fluss zwischen den Prozessschritten auf Basis von Entscheidungspunkten
- > **Referenzdokument:** dokumentiert – meist unstrukturiert – Methodik, Rahmenbedingungen, Grundlagen, Best Practices und so weiter in physischen Publikationen (zum Beispiel Büchern) oder elektronischen Dateien (zum Beispiel Qualitätsberichten)
- > **Regel:** ermöglicht die Prüfung oder Transformation beliebiger Objekte durch Anwendung einer hinterlegten Anweisung
- > **Statistik:** sorgt für organisatorischen Umgebungskontext, in dem Aktivitäten zur Erstellung von Statistiken innerhalb einer statistischen Organisation durchgeführt werden
- > **Statistikunterstützungsaktivität:** umfasst Funktionen wie Metadatenmanagement, Datenmanagement, Qualitätsmanagement, methodische Forschung und Designfunktionen
- > **Statistischer Bedarf:** dokumentiert eine vorgeschlagene oder auferlegte Anforderung, Anfrage oder sonstige Mitteilung, die bei einer statistischen Organisation eingegangen ist:

> **Informationsanfrage:** dokumentiert eine Anfrage nach neuen Informationen für einen bestimmten Zweck

> **Umgebungsänderung:** dokumentiert eine Änderung im Tätigkeitsbereich einer statistischen Organisation

### 3.3 Concept

---

Die Gruppe „Concept“ wird verwendet, um die Bedeutung von Daten zu definieren und ein Verständnis dafür zu vermitteln, was die Daten messen:

- > **Ausprägung:** spezifiziert übergreifend eine bestimmte Eigenschaft, die in einem beliebigen Kontext verwendet werden kann
- > **Bezeichnung:** benennt ein Objekt zur (eindeutigen) Identifizierung
- > **Datum:** enthält den tatsächlichen Wert, der einen Datenpunkt füllt, und verknüpft nützliche Informationen, beispielsweise zur Herkunft oder Qualität
- > **Ebene:** verknüpft eine Ebene einer (hierarchischen) Knotenmenge mit einem unverwechselbaren Terminus
- > **Gegenüberstellung:** dokumentiert die Zuordnung von Knotenmengen
- > **Grundgesamtheit:** spezifiziert die Menge aller statistischen Einheiten mit übereinstimmenden Identifikationskriterien (sachlich, räumlich und zeitlich)
- > **Instanzmerkmal:** beschreibt Instanzen von Daten, die gesammelt wurden (erst die Kombination von Standardmerkmal, Kontextmerkmal und Instanzmerkmal definiert das konkrete Merkmal für eine Erhebungsdurchführung)
- > **Klassifikationsfamilie:** fasst Klassifikationsreihen mit einem gemeinsamen Terminus (zum Beispiel die Wirtschaftstätigkeit) zusammen
- > **Klassifikationsreihe:** bündelt Klassifikationsversionen
- > **Knoten:** definiert ein einzelnes Element einer Knotenmenge, speziell je nach Einsatzzweck:

- > **Ausprägungselement:** ordnet eine Ausprägung einer Ausprägungsliste zu
- > **Klassifikationselement:** definiert Inhalt und Grenzen der verknüpften Ausprägung für eine Klassifikationsversion
- > **Verschlüsselungszuordnung:** verknüpft eine Ausprägung mit Schlüsselwert für eine Schlüssel-liste
- > **Knotenmenge:** bündelt Knoten basierend auf einer Menge von Termini
  - > **Ausprägungsliste:** enthält Knoten ohne Schlüsselwerte
  - > **Klassifikationsversion:** definiert Ausprägungen mit standardisierten Schlüsselwerten unter Bezugnahme auf ein- und ausschließende Charakteristiken
  - > **Schlüsselliste:** verwaltet mögliche Schlüsselwerte für aufgezählte Wertebereiche
- > **Kontextmerkmal:** verwendet ein Merkmal in einem konkreten Kontext, das heißt einer konkreten Erhebung
- > **Konzeptioneller Wertebereich:** definiert den Wertebereich des Merkmals auf Ebene des Standardmerkmals „konzeptionell“ im Sinne eines Wertebereichstyps vor, speziell:
  - > **Aufgezählter konzeptioneller Wertebereich:** konzeptioniert einen Wertebereich auf Basis einer Ausprägungsliste
  - > **Beschriebener konzeptioneller Wertebereich:** konzeptioniert einen Wertebereich durch eine Beschreibung oder Spezifikation des Wertebereichs, zum Beispiel eine Regel, eine Prozedur oder ein Intervall, in dem die Werte liegen
- > **Maßeinheit:** definiert eine Größeneinheit, in der ein Wert gemessen wird
- > **Maßeinheitentyp:** gruppiert alle Maßeinheiten, die ineinander umgerechnet werden können
- > **Merkmalsträgertyp:** gruppiert Statistische Einheiten auf Basis einer einzelnen sachlichen Eigenschaft (ohne zeitliche und geografische Angabe)
- > **Sachgebiet:** bündelt gleichartige Statistiken, die auf den gleichen Terminologien basieren
- > **Schlüsselwert:** kodiert Ausprägungen
- > **Standardmerkmal:** kombiniert die Bedeutung eines Terminus mit einem Merkmalsträgertyp und dem Konzeptionellen Wertebereich, um die zu messende Eigenschaft zu definieren
- > **Statistische Einheit:** legt den konkreten Träger der Informationen für die statistische Untersuchung fest
- > **Stichwort:** definiert die Charakteristiken eines Knotens, jedes Stichwort bezieht sich normalerweise eindeutig auf einen Knoten
- > **Stichwortverzeichnis:** bündelt Stichworte und verknüpft damit Text aus statistischen Datenquellen mit Knotenmengen
- > **Terminologie:** bündelt Termini in Abgrenzung zu anderen Terminologien
- > **Terminus:** definiert eine Begrifflichkeit, die als Basis für viele andere Objekte (Grundgesamtheit, Universum, Ausprägung, Merkmal, Merkmals-trägertyp) dient
- > **Universum:** definiert einen spezialisierten Merkmalsträgertyp ohne zeitliche und geografische Angabe
- > **Wertebereich:** definiert eine abstrakte Oberklasse für mögliche Eigenschaften, die durch die Objekte „Beschriebener Wertebereich“ und „Aufgezählter Wertebereich“ hinreichend konkretisiert werden muss:
  - > **Aufgezählter Wertebereich:** verknüpft eine Knotenmenge, die zulässige Ausprägungen und Schlüsselwerte enthält
  - > **Beschriebener Wertebereich:** konkretisiert die erlaubten Werte durch einen maschinell auswertbaren Ausdruck
  - > **Fehlwertebereich:** verknüpft ein Instanzmerkmal mit einem (aufgezählten oder beschriebenen) Wertebereich, der Aussagen zu Fehlwerten dokumentiert
  - > **Inhaltlicher Wertebereich:** verknüpft ein Kontextmerkmal mit einem (aufgezählten oder beschriebenen) Wertebereich, der die spezifischen gültigen Werte dokumentiert



- > **Zuordnung:** ordnet (im Rahmen einer Gegenüberstellung) Knoten einer Quellknotenmenge und einer Zielknotenmenge einander zu

### 3.4 Exchange

---

Die Gruppe „Exchange“ dient zum Katalogisieren der Informationen, die innerhalb einer statistischen Organisation und zwischen statistischen Organisationen über Austauschinstrumente ausgetauscht werden. Sie umfasst Informationsobjekte, die die Sammlung und Verbreitung von Informationen beschreiben:

- > **Ausgabespezifikation:** spezifiziert Produkte und verwendet dazu die Präsentationen, die sie enthalten
- > **Aussage:** enthält Hinweise in Formularen für den Interviewer oder Befragten
- > **Austauschinstrument:** spezifiziert Werkzeuge zum Empfangen oder Senden von Informationen:
  - > **Formular:** Instrument für Datenerhebungen (zum Beispiel persönlich, computergestütztes Interview, Online-Fragebogen)
  - > **Harvesting:** Sammeln von Daten zu einem bestimmten Sachthema über netzbasierte Dienste (zum Beispiel Verwaltungsdaten über Webservices auslesen oder Preise von Webseiten extrahieren)
  - > **Verbreitungsinstrument:** Werkzeug, um Daten oder Referenzmetadaten zur weiteren Nutzung anzubieten
- > **Austauschspezifikation:** legt das Design der Austauschinstrumente fest (unter anderem Informationsstruktur), speziell:
  - > **Formularspezifikation:** enthält das vollständige Formulardesign, mit Beziehung zur Formularkomponente der obersten Ebene
  - > **Formularkomponente:** definiert die Struktur der Formularspezifikation als eine Kombination aus Fragen, Formularblöcken und Aussagen, speziell:
    - > **Formularblock:** ermöglicht Wiederverwendung in mehreren Formularen
    - > **Formularlogik:** steuert Darstellung der Fragen, Aussagen, Formularkomponenten und so weiter auf Basis hinterlegter Regeln
    - > **Frage:** spezifiziert ein Antwortfeld als Teil eines Formulars, üblicherweise in Verbindung mit einem Merkmal und dessen Wertebereich
    - > **Informationsanbieter:** definiert einen Lieferanten für Informationen mit einer bestimmten Informationsstruktur
    - > **Informationsverbraucher:** definiert einen Konsumenten für Informationen mit einer bestimmten Informationsstruktur
    - > **Instanzaussage:** spezifiziert eine Aussage in einer konkreten Formularkomponente
    - > **Instanzformularblock:** spezifiziert einen Formularblock in einer bestimmten Formularkomponente
    - > **Instanzfrage:** definiert die Verwendung einer Frage in einer bestimmten Formularkomponente
    - > **Liefervereinbarung:** regelt die Nutzung von Austauschinstrumenten zwischen Informationsanbietern und -verbrauchern bezüglich Informationsstruktur (und damit die Inhalte) der auszutauschenden Informationen
    - > **Präsentation:** definiert die Darstellung von Daten und Referenzmetadaten aus Informationssätzen, zum Beispiel Tabellen, Grafiken, strukturierte Datendateien
    - > **Produkt:** verpackt Informationssätze für einen Informationsverbraucher
    - > **Register:** speichert Einzeldaten zum regelmäßigen Abruf

### 3.5 Structure

---

Die Gruppe „Structure“ wird verwendet, um Informationen im gesamten statistischen Geschäftsprozess zu strukturieren:

- > **Datenpunkt:** definiert die Zelle als Platzhalter für einen Wert, den die Zelle enthalten kann
- > **Datenstrukturkomponente:** definiert einen Bestandteil einer Datenstruktur, speziell:

- > **Identifikatorkomponente:** identifiziert die statistische Einheit
- > **Messkomponente:** dient einem beobachteten/abgeleiteten Wert
- > **Attributkomponente:** hält eine andere Information als zur Identifikation oder Messung bereit
- > **Datentupel:** enthält bestimmte Eigenschaften zu einer Statistischen Einheit in einer festen Reihenfolge (bezogen auf eine Tabelle entspricht diese einer Tabellenzeile mit einer bestimmten Satzart)
- > **Informationsressource:** bündelt Informationssätze, speziell:
  - > **Datenressource:** bündelt Datensätze
  - > **Referenzmetadatenressource:** bündelt Referenzmetadaten
- > **Informationssatz:** enthält Informationen, die während der Statistikproduktion erhoben, verarbeitet, analysiert und verbreitet werden, speziell:
  - > **Datensatz:** enthält Daten (unter anderem aus Register, Zeitreihen, Längsschnittdaten, Umfragedaten, Ereignisverlaufsdaten, Datentabellen, Datenwürfel, Data Warehouses und so weiter)
  - > **Referenzmetadatenatz:** enthält Referenzmetadaten mit einem bestimmten Referenzmetadatenbezug
- > **Informationsstruktur:** beschreibt das Aussehen einer Informationsstruktur, speziell:
  - > **Datenstruktur:** beschreibt den Aufbau eines Datensatzes durch Kombination entsprechender Datenstrukturkomponenten
    - > **Makrodatenstruktur:** beschreibt speziell die Struktur eines Makrodatensatzes
    - > **Mikrodatenstruktur:** beschreibt speziell die Struktur eines Mikrodatensatzes
  - > **Referenzmetadatenstruktur:** beschreibt den Aufbau eines Referenzmetadatenatzes durch Kombination entsprechender Referenzmetadatenattribute
- > **Referenzmetadatenattribut:** definiert den Bestandteil einer Referenzmetadatenstruktur
- > **Referenzmetadatenbezug:** identifiziert das Thema, für das Referenzmetadaten mit einer bestimmten Referenzmetadatenstruktur grundsätzlich gemeldet werden können (zum Beispiel Datensätze oder Statistiken)
- > **Referenzmetadatengegenstand:** konkretisiert ein bestimmtes Objekt, für den ein Referenzmetadatenatz angelegt wird (zum Beispiel ein bestimmter Datensatz oder eine bestimmte Statistik)
- > **Referenzmetadateninhalt:** enthält den eigentlichen Inhalt für eine bestimmte Eigenschaft eines Referenzmetadaten-themas
- > **Tupelbeziehung:** dokumentiert Abhängigkeiten zwischen Tupelstrukturen
- > **Tupelstruktur:** beschreibt den Aufbau eines Datentupels und kann darüber zusätzliche Beziehungen definieren (zum Beispiel eine Person ist Teil eines Haushalts)

## 4

### Anwendung

---

MedaS ist die Basis für die Architektur des Metadatenmanagements im Statistischen Verbund. Mit dieser ist es möglich, bestehende metadatenhaltende Systeme und Werkzeuge mithilfe eines zentralen Metadatenmanagementsystems miteinander zu koppeln. Die technische Anbindung der externen Systeme, die Übernahme ausgewählter Inhalte und die gegebenenfalls notwendigen Objekttransformationen übernimmt das Metadatenmanagementsystem. Dieses verbindet somit alle Systeme mit einer zentralen Metadaten-speicherung und stellt über standardisierte Schnittstellen übergreifend notwendige Metadatenobjekte im standardisierten Informationsmodell MedaS für eine konsistente Nachnutzung durch Standardwerkzeuge (zum Beispiel das Statistikportal) und Fachanwendungen (zum Beispiel zur Übertragung von Konfigurationsparametern und Qualitätsinformationen) zur Verfügung. Entsteht neuer Bedarf für die Verwaltung weiterer Metadaten, ermöglicht das Metadatenmanagementsystem die kostengünstige Ergänzung schlanker Anwendungs-module ohne eigene Datenbank, die zur persistenten Speicherung

von Inhalten die bereits existierende Datenbank des Metadatenmanagementsystems nutzen.

Gegenüber dieser statischen Perspektive aus Sicht der Systemarchitektur offenbart ein Blick auf die Geschäftsprozesse der Statistikproduktion eine eher dynamische Sichtweise: Entlang der Prozesskette zur Statistikproduktion werden in bestimmten Teilprozessen (gemäß GMAS) Metadatenobjekte (gemäß MedaS) produziert und in anderen Teilprozessen als Eingabeobjekte benötigt. [↗ Grafik 4](#)

Wichtige Aspekte sind hierbei die konsistente Wiederverwendung bereits vorhandener Metadaten, die Reduzierung manueller Aufwände und die lückenlose Dokumentation der Produktionsprozesse, um die Qualität der Produkte und Prozesse jederzeit bewerten und stetig verbessern zu können (Europäische Kommission, 2019, hier: Seite 19, 28 f., 32, 35, 38, 40, 44, 54 ff.; Statistisches Bundesamt, 2017, hier: Seite 11, 15 f., 18 f.; Statistisches Bundesamt, 2020 hier: Seite 22, 24 f., 27). Hinsichtlich des Ziels einer digitalen, standardisierten und weitgehend automatisierten Statistikproduktion sind die oben aufgeführten MedaS-Objekte unterschiedlich relevant. Der Ausbau und die weitere IT-Unterstützung durch das Metadatenmanagementsystem erfolgen daher sukzessive nach Bedarf unter Berücksichtigung vorhandener Ressourcen, die in den weiteren Ausbau der Infrastruktur des Metadatenmanagements investiert werden können. MedaS dient dabei als Orientierung für das weitere Vorgehen, um perspektivisch eine vollständig digitale und medienbruchfreie Verwaltung aller Metadatenobjekte nutzen zu können. Durch die Kompatibilität mit dem internationalen Standard GSIM können auch Austausch und Nutzung technischer Komponenten anderer statistischer Organisationen gelingen.

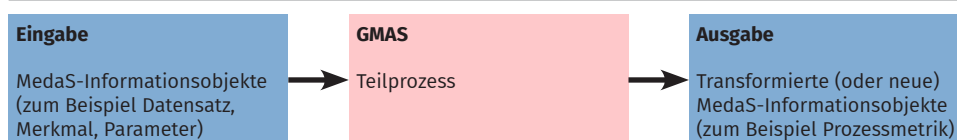
Bei der Modernisierung des Werkzeugportfolios für die Statistikproduktion wird bevorzugt auf Fertigprodukte gesetzt, die in die bestehende System-

landschaft zu integrieren sind. Zwar wird der Austausch von Inhalten über technische Schnittstellen der Komponenten im Prozess grundsätzlich unterstützt, bei voneinander abweichenden Informationsmodellen können aber größere Aufwände für Anbindung und Transformationen entstehen. Muss jedes metadatenhaltende System seine Inhalte mit allen anderen Systemen abgleichen, um einen konsistenten Stand der Informationsobjekte entlang der Produktionskette zu gewährleisten, kann die Integration zu einem exponentiell steigenden Programmieraufwand für entsprechend leistungsfähige Adapter führen. Durch den Einsatz einer zentralen Komponente wie dem Metadatenmanagementsystem, das einen abgestimmten Standard wie MedaS implementiert, reduziert sich der Aufwand spürbar. Für jedes neue System ist dann lediglich ein Adapter zur Anbindung der Webservices und Transformation auszutauschender Informationsobjekte notwendig.

Schließlich übernimmt MedaS nicht nur eine wichtige Rolle als technischer Standard, sondern definiert darüber hinaus auch eine gemeinsame Fachsprache, um ein einheitliches Verständnis für Metadaten zu erreichen. Dies fördert sowohl die organisationsinterne Kommunikation als auch den Austausch auf nationaler und internationaler Ebene über Organisationsgrenzen hinweg. Das spiegelt sich insbesondere in der Spezifikation von Geschäftsprozessen (zum Beispiel im GSBPM beziehungsweise GMAS) oder auch der methodischen Abstimmung einer Statistikproduktion wider.

### Grafik 4

Definition von Informationsobjekten entlang der Prozesskette zur Statistikproduktion



### 5

#### Fazit

International und national sind standardisierte Geschäftsprozessmodelle und Informationsmodelle wichtige Grundlagen für die Modernisierung der amtlichen Statistik. MedaS und GMAS sind komplementäre Modelle, um statistische Informationen zu erstellen und zu verwalten. GMAS bietet Bausteine und eine gemeinsame Kommunikationssprache für die Modellierung des statistischen Geschäftsprozesses und identifiziert die Aktivitäten von Statistik erstellenden Organisationen, die zu Informationsausgaben führen. Diese Aktivitäten sind in Teilprozesse wie „Produkte konzipieren“ und „Aggregate berechnen“ unterteilt. MedaS hilft bei der Beschreibung von GMAS-Teilprozessen, indem es die Informationsobjekte im Prozessfluss von Statistiken definiert.

Ein gutes Metadatenmanagement ist für die effiziente Implementierung statistischer Geschäftsprozesse von wesentlicher Bedeutung. Metadaten sind in jeder Phase von GMAS vorhanden, entweder erstellt, aktualisiert oder unverändert aus einer vorherigen Phase übernommen. Im GMAS liegt der Schwerpunkt des übergreifenden Prozesses des Metadatenmanagements auf dem Erstellen, Aktualisieren, Verwenden und Wiederverwenden von Metadaten. Metadatenmanagement-Strategien und -Systeme sind daher für den Betrieb von GMAS von entscheidender Bedeutung und werden durch MedaS erleichtert.

Eine Referenzimplementierung von MedaS findet sich im Metadatenmanagementsystem des Statistischen Bundes. Dieses dient der zentralen Ablage von standardisierten Metadaten entlang der Prozesskette der Statistikproduktion, um deren Wiederverwendbarkeit zu erleichtern und Inkonsistenzen zu vermeiden. Hierzu stellt das Metadatenmanagementsystem eine Reihe von Schnittstellen zur Verfügung, über die die Standardwerkzeuge und Fachanwendungen auf Metadaten zugreifen können. Zur initialen Eingabe und Verwaltung von Metadaten können über browserbasierte Benutzeroberflächen diverse Anwendungsmodule zum Beispiel zur Dokumentation der Produkt- und Prozessqualität genutzt

werden. So ist es möglich, die Implementierungslücken ergänzend zu den existierenden metadatenhaltenden Produktionswerkzeugen sukzessive zu schließen.

Das Metadatenmodell der amtlichen Statistik, MedaS, ist die nationale Adaption des internationalen Standards GSIM, der die konzeptionelle Grundlage für die Modernisierung der amtlichen Statistik weltweit bildet. In seine Entwicklung sind im Zeitverlauf Wissen und Erfahrung aus verschiedenen statistischen Organisationen eingeflossen. Da immer mehr statistische Organisationen GSIM als Teil ihres Modernisierungsprogramms verwenden oder dies planen, wächst die Benutzerbasis des Modells stetig. Der Austausch von Wissen und gewonnenen Erkenntnissen kommt neuen und erfahrenen Nutzenden gleichermaßen zugute und das Feedback aus diversen Anwendungen ist für die Überarbeitung des Modells von wesentlicher Bedeutung. Somit kann MedaS auch künftig von der Weiterentwicklung von GSIM profitieren. 

## LITERATURVERZEICHNIS

---

Blumöhr, Torsten/Teichmann, Corina/Noack, Anke. [\*Standardisierung der Prozesse: 14 Jahre AG SteP\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 58 ff.

Europäische Kommission. *Qualitätssicherungsrahmen des Europäischen Statistischen Systems*. Version 2.0. 2019. [Zugriff am 13. Januar 2026]. Verfügbar unter: [ec.europa.eu](https://ec.europa.eu)

Gehle, Christian/Lüüs, Hans-Peter. [\*Prozessmanagement im Statistischen Bundesamt\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 46 ff.

Lindenstruth, Thomas/Claußen, Sven. [\*Metadatenmanagement als neue Integrationsarchitektur\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 5/2017, Seite 76 ff.

Statistisches Bundesamt. *Die Digitalstrategie 2025 des Statistischen Bundesamtes*. Wiesbaden 2022. (Internes Dokument)

Statistisches Bundesamt. *Das System der amtlichen Statistik: Organisation und Zusammenarbeit im nationalen, europäischen und internationalen Kontext*. Wiesbaden 2020. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.statistischebibliothek.de](http://www.statistischebibliothek.de)

Statistisches Bundesamt. *Verhaltenskodex für europäische Statistiken*. Version November 2017. [Zugriff am 13. Januar 2026]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

UNECE. *Generic Statistical Business Process Model*. Version 5.2. 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: [statswiki.unece.org](https://statswiki.unece.org)

UNECE. *Generic Statistical Information Model*. Version 2.0. 2025. [Zugriff am 13. Januar 2026]. Verfügbar unter: [statswiki.unece.org](https://statswiki.unece.org)

# EFFIZIENTE AUSWERTUNG DES BUSINESS-TAX-PANELS MITHILFE DER STATISTIKSOFTWARE R

Oliver Hauke, Annette Kristiansen

➤ **Schlüsselwörter:** effiziente Programmierung – Datenverarbeitung – Apache Parquet – Steuerstatistiken – Forschungsdatenzentrum

## ZUSAMMENFASSUNG

Das Statistische Bundesamt steht ständig vor der Herausforderung, deutlich größere Datenvolumen zu verarbeiten; zugleich steigen die Erwartungen der Öffentlichkeit an die Aktualität und Granularität der amtlichen Statistiken. Aktuelle technologische Weiterentwicklungen bieten das Potenzial, diese Herausforderungen zu bewältigen. So können geeignete Werkzeuge Auswertungen, die zuvor mehrere Stunden gedauert haben, in wenigen Sekunden ausführen. Der Artikel evaluiert am Beispiel des Business-Tax-Panels die Performanz und Handhabung entsprechender Werkzeuge in der Statistiksoftware R mit dem Ziel, Handlungsempfehlungen für die Praxis abzuleiten.

➤ **Keywords:** efficient programming – data processing – Apache Parquet – tax statistics – research data centre

## ABSTRACT

*The Federal Statistical Office of Germany faces the constant challenge of having to process significantly increasing data volumes, while public expectations regarding the granularity and timeliness of official statistics continue to increase. Recent technological developments hold the potential to meet this challenge. Analyses that used to take several hours can be conducted within seconds using suitable tools. Taking the business tax panel as an example, the article assesses the performance and user-friendliness of tools provided in the statistical software R, with the aim of deriving recommendations for their use in practice.*



**Oliver Hauke**

ist Mathematiker und Referent im Referat „IT-Kompetenzzentrum Auswertung und Analyse“ des Statistischen Bundesamtes. Er verantwortet die Weiterentwicklung der technischen Infrastruktur des Forschungsdatenzentrums des Statistischen Bundesamtes und berät das Netzwerk empirische Steuerforschung in der effizienten Nutzung der Systeme.



**Annette Kristiansen**

ist Ökonomin und Referentin im Referat „Unternehmenssteuern“ des Statistischen Bundesamtes. Sie beschäftigt sich mit der Zusammenführung der Unternehmenssteuerstatistiken und betreut Arbeiten für das Netzwerk empirische Steuerforschung.



## 1

### Einleitung

Die zunehmende Digitalisierung hat Auswirkungen auf viele Lebensbereiche und somit auf die Gesellschaft. Sie beeinflusst und verändert beispielsweise Kommunikation, Arbeitswelt und Bildung. Als ein Produkt der Digitalisierung entstehen große Datenmengen, die die amtliche Statistik vor die Herausforderung stellen, diese deutlich größeren Datenmengen zu verarbeiten. Zugleich nehmen die Erwartungen der Öffentlichkeit an die Aktualität, Genauigkeit und Gliederungstiefe von amtlichen Statistiken zu. Technologische Weiterentwicklungen bieten das Potenzial, diese Herausforderungen zu bewältigen. Geeignete Werkzeuge können Auswertungen, die bisher Stunden dauerten, in wenigen Sekunden ausführen. Der Artikel evaluiert die Leistung solcher Werkzeuge in der [Statistiksoftware R](#) sowie deren Handhabung durch Nutzende von amtlichen Mikrodaten.

Die Analysebedarfe der Wissenschaft in den Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder (FDZ)<sup>1</sup> verdeutlichen die Herausforderungen besonders. Die On-Site-Zugangswege bieten formal anonymisierte Mikrodaten der amtlichen Statistik in großem Umfang für wissenschaftliche Auswertungen an. Diese umfassen Gastwissenschaftsarbeitsplätze (GWAP), die Forschende direkt in den Liegenschaften der Forschungsdatenzentren nutzen können, oder kontrollierte Datenfernverarbeitung (KDFV), bei der entwickelte Auswertungsprogramme durch die FDZ-Mitarbeitenden ausgeführt werden. Die Infrastruktur der Forschungsdatenzentren muss in der Lage sein, umfangreiche Forschungsprojekte auf Basis großer Datenmengen effizient zu unterstützen.

Auch aus der Politik und der evidenzbasierten Politikberatung haben sich weitere Anforderungen ergeben. So fordert der Wissenschaftliche Beirat beim Bundesministerium der Finanzen (2020) eine Ausweitung des Datenbestands für die empirische Steuerforschung, während der Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwick-

lung (2023) die Forschungsdateninfrastruktur zudem als rückständig bezeichnete.

Seit seiner Gründung im Jahr 2022 ist die Verbesserung der Datenbereitstellungsinfrastruktur für die empirische Steuerforschung ein Fokusthema des Netzwerks empirische Steuerforschung (NeSt). Insgesamt vernetzt das Netzwerk Kompetenzen aus der Finanzverwaltung, Steuerforschung und amtlichen Statistik mit dem Ziel, die evidenzbasierte Steuerforschung zu stärken. Es hat die Erweiterung des Datenbestands an Steuerstatistiken seither maßgeblich vorangetrieben. Seit 2024 steht das Business-Tax-Panel (BTP) mit fast 3 000 Merkmalen und 67 Millionen Beobachtungen über die Forschungsdatenzentren zur Verfügung (Kapitel 2) – ein enormes Analysepotenzial, das zugleich hohe Anforderungen an die Analysesysteme und deren Nutzende stellt.

Um diese Herausforderungen zu bewältigen, wurde ein interner R-Server des Statistischen Bundesamtes im Januar 2025 um erhebliche Rechenkapazitäten erweitert. Vorangetrieben durch das Netzwerk empirische Steuerforschung stehen auch der Wissenschaft seit November 2025 im FDZ Bund die [On-Site-Server](#) mit erhöhten Rechenkapazitäten für Analysen in R und Stata exklusiv für GWAP- und KDFV-Nutzungen zur Verfügung (Kapitel 3) – ein bedeutender Fortschritt für die Infrastruktur des FDZ Bund.<sup>2</sup>

Die freie Statistiksoftware R hat sich zu einem Standardwerkzeug für Statistik und Datenanalyse entwickelt, das zunehmend in der amtlichen Statistik sowie in der Wissenschaft genutzt wird. Eine weltweite Gemeinschaft erweitert die Open-Source-Software R laufend um neue Funktionalitäten in Form von R-Paketen – insbesondere für effiziente Datenverarbeitung (Kapitel 4). Dieser Artikel bewertet die Eignung von R-Paketen für die Verarbeitung großer Datenmengen im Statistischen Bundesamt und FDZ Bund (Kapitel 5) und demonstriert, wie mit R relevante Auswertungen des Business-Tax-Panels sekundenschnell durchgeführt werden können (Kapitel 6). Das Ergebnis zeigt, dass geeignete Technologien die Effizienz erheblich steigern, zugleich jedoch weiterer Handlungsbedarf besteht (Kapitel 7).

---

1 FDZ Bund steht im Folgenden für das Forschungsdatenzentrum des Statistischen Bundesamtes.

2 Demonstriert in einem Vortrag anlässlich eines NeSt-Webinars des Bundesministeriums der Finanzen am 28. März 2025: [www.bundesfinanzministerium.de](http://www.bundesfinanzministerium.de)

## 2

### Anwendungsfall Business-Tax-Panel

#### 2.1 Datengrundlage

Das Business-Tax-Panel kombiniert die Angaben aus verschiedenen Ertrags- sowie Umsatzsteuerstatistiken im Quer- und Längsschnitt. Verfügbar sind Informationen aus den folgenden Statistiken: Gewerbesteuerstatistik, Körperschaftsteuerstatistik, Umsatzsteuer-Voranmeldung und Veranlagung, Statistik über die Personengesellschaften und Gemeinschaften, Einnahmenüberschussrechnung sowie einige Merkmale aus dem statistischen Unternehmensregister. Insgesamt deckt der Forschungsdatensatz eine Vielzahl an Analysemöglichkeiten ab, wie die Analyse von Anpassungsreaktionen oder Verteilungsanalysen, die für evidenzbasierte steuerpolitische Entscheidungen nötig sind (Kristiansen, 2023).

Für die vielfältigen Analysewünsche der Forschenden, die damit einhergehende Dokumentation des Datensatzes in den FDZ-Metadatenreports (FDZ Bund, 2024) und den Auswertungsbedarf des Fachbereichs im Statistischen Bundesamt ist der Datensatz möglichst wenig beschränkt worden und dadurch ressourcenintensiv. Anstelle einer Stichprobe wurden die zugrunde liegenden Vollerhebungen verwendet. Die erste Version des Business-Tax-Panels (2013 bis 2019) besteht aus knapp 67 Millionen Beobachtungen und liegt in komprimierter Form mit insgesamt 314 Gigabyte<sup>13</sup> vor. Mit der Aktualisierung um das Berichtsjahr 2020 wird sich die Zahl der Beobachtungen auf etwa 78 Millionen erhöhen. Ebenso steigt der Merkmalsumfang von fast 3000 Merkmalen durch weitere Berichtsjahre. Bereits jetzt – und insbesondere durch die anstehende Aktualisierung – stehen der Fachbereich im Statistischen Bundesamt sowie das FDZ Bund vor der Herausforderung, den hohen Bedarf an Festplattenspeicher, Arbeitsspeicher (sogenanntes Random-Access-Memory, kurz RAM) und Prozessorleistung zu decken.

<sup>3</sup> Der Datenumfang ist in komprimierter SAS-Datei gemessen. Dies variiert für andere Dateitypen (siehe Kapitel 5).

#### 2.2 Auswertungsbedarfe

Neben den standardisierten Veröffentlichungsprodukten (wie Tabellen in der Datenbank GENESIS-Online, Statistische Berichte, Pressemitteilungen) der zugrunde liegenden Statistiken erstellt das Statistische Bundesamt zusätzlich Sonderauswertungen. Diesen liegen ebenso wie den Standardauswertungen meist spezifische Anforderungen zugrunde:

Ein häufiger Anwendungsfall ist die deskriptive Auswertung in der Statistikproduktion. Dafür sind verschiedene Bearbeitungsschritte in Form von Gruppierung der Beobachtungen, Auswahl von Merkmalen und Bearbeitung der Daten nötig. Auch für den FDZ-Metadatenreport werden deskriptive Auswertungen benötigt. Daher wird als Anwendungsfall die Tabelle zur Häufigkeitsverteilung der Beobachtungen je Jahr und Steuerstatistik (FDZ Bund, 2024) herangezogen.

➤ **Tabelle 1** auf Seite 130

Ein weiterer Anwendungsfall simuliert eine exemplarische Datenanalyse, die den klassischen Bedarf der Wissenschaft repräsentiert. Die verschiedenen Datenbearbeitungsschritte und Analysen zeigen ein potenzielles Forschungsprojekt und sind entsprechend vereinfacht. Konkret wird das Vorliegen eines Verlustvortrages<sup>14</sup> im jeweiligen Berichtsjahr anhand verschiedener Einflussfaktoren untersucht. Die Körperschaftsteuerstatistik zeigt über die Jahre eine Zunahme der Zahl der Steuerpflichtigen mit Verlustvortrag im jeweiligen Berichtsjahr (GENESIS-Tabellen-Code 73211-0002). Anzunehmen ist, dass Personalkosten den Verlustvortrag erhöhen können, daher dienen die Zahl der sozialversicherungspflichtig Beschäftigten sowie die Summe der Löhne und Gehälter als erklärende Variable. Der jeweilige Gesamtbetrag der Einkünfte und die abgezogenen Spenden können ebenfalls den Verlustvortrag beeinflussen. Um den Internationalisierungsgrad des Unternehmens zu erfassen, wird ein Dummy aufgenommen. Die Ausgabe der Regressionsergebnisse ist hier nicht von Relevanz und wird daher nicht erzeugt.

Die beiden beschriebenen Anwendungsfälle (Fallzahl und Regression) werden für den Performanzvergleich genutzt (siehe Kapitel 5).

<sup>4</sup> Verluste aus einem Veranlagungsjahr können mit Gewinnen in späteren Veranlagungsjahren verrechnet werden.

**Tabelle 1**

**Häufigkeitsverteilung der Beobachtungen je Jahr und Steuerstatistik**

|      | Gewerbesteuer | Körperschaftsteuer | Statistik der Personengesellschaften und Gemeinschaften | Umsatzsteuer-Voranmeldung | Umsatzsteuer-Veranlagung | Statistisches Unternehmensregister | Einnahmen-überschussrechnung |
|------|---------------|--------------------|---------------------------------------------------------|---------------------------|--------------------------|------------------------------------|------------------------------|
| 2013 | 3 633 451     | 1 210 838          | 1 205 110                                               | 3 243 538                 | 6 435 280                | 3 474 431                          | 3 703 982                    |
| 2014 | 3 714 176     | 1 236 129          | 1 220 521                                               | 3 240 221                 | 6 446 620                | 3 550 775                          | 3 720 318                    |
| 2015 | 3 819 938     | 1 264 717          | 1 234 537                                               | 3 255 537                 | 6 535 948                | 3 634 643                          | 4 230 871                    |
| 2016 | 3 887 556     | 1 293 431          | 1 247 699                                               | 3 266 429                 | 6 550 960                | 3 674 354                          | 4 581 982                    |
| 2017 | 3 981 853     | 1 316 520          | 1 220 067                                               | 3 266 806                 | 6 664 535                | 3 708 949                          | 5 685 661                    |
| 2018 | 4 050 674     | 1 350 837          | 1 284.121                                               | 3 279 136                 | 6 831 036                | 3 736 550                          | 6 310 028                    |
| 2019 | 4 099 690     | 1 388 679          | 1 294 898                                               | 3.288 306                 | 6 972 252                | 3 723 169                          | 6 501 292                    |

Anmerkung: Das statistische Unternehmensregister ist nicht komplett im Business-Tax-Panel enthalten. Nur Einheiten, die in den Steuerstatistiken enthalten sind, werden um Angaben des statistischen Unternehmensregisters ergänzt.

### 3

## Ausbau der Analyseserver

Mit der Bereitstellung neuer Analyseserver im Jahr 2025 wurden entscheidende Fortschritte beim Ausbau der Infrastruktur des Statistischen Bundesamtes und des FDZ Bund erzielt. Die Einführung der neuen Systeme ist als Ergebnis langjähriger Planung und Umsetzung ein bedeutender Schritt.

Änderungen von Softwareversionen sind in einer abgeschotteten Infrastruktur wie im Statistischen Bundesamt organisatorisch aufwendig. Daher standen für die folgenden Darstellungen nicht immer die neuesten Softwareversionen zur Verfügung – dies spiegelt die realistischen Bedingungen der Statistikproduktion wider. Die neue Infrastruktur wurde daher von Beginn an zukunftsorientiert konzipiert und bietet moderne Software auf deutlich erweiterter Hardwarebasis.

### 3.1 Ausgangslage

Das Hauptwerkzeug der Statistikproduktion des Statistischen Bundesamtes ist die kommerzielle Analysesoftware SAS 9.4, die auf dedizierten Servern bereitgestellt wird.<sup>5</sup> SAS verarbeitet Daten überwiegend zeilenorientiert und belastet den Arbeits-

speicher nur gering – ein entscheidender Vorteil in der amtlichen Statistik, da große Datenmengen mit begrenzter Hardware verarbeitet werden können. Für viele Anwendungsfälle ist jedoch die spaltenbasierte Datenverarbeitung effizienter. In R sind Verfahren verfügbar, die moderne Prozessor- und Arbeitsspeicher gezielt ausnutzen und damit Laufzeiten gegenüber SAS erheblich verkürzen (siehe die Kapitel 4 und 5).

Im FDZ Bund ist die Statistiksoftware [Stata](#) das am häufigsten genutzte Werkzeug. Viele wissenschaftliche Fachrichtungen nutzen es, da weniger tiefgreifende Kenntnisse nötig sind und spezialisierte Verfahren, insbesondere für ökonometrische Analysen, bereitstehen. Bei großen Datenmengen stößt Stata jedoch an Grenzen, da Daten vollständig in den Arbeitsspeicher geladen werden müssen und in Stata nur wenige Operationen auf mehrere Prozessorkerne parallelisiert durchgeführt werden. Nur durch manuelle Variablenauswahl lassen sich größere Datensätze bearbeiten. Bisher waren die Gastwissenschaftsarbeitsplätze des FDZ Bund auf die Single-Core-Version von Stata mit begrenztem Arbeitsspeicher beschränkt, wodurch viele FDZ-Produkte ohne Datenzuschnitt<sup>6</sup> an den Gastwissenschaftsarbeitsplätzen nicht nutzbar waren. Leistungsfähigere Server standen vor der Bereitstellung der On-Site-Server lediglich in der kontrollierten Datenfernverarbeitung zur Verfügung.

<sup>5</sup> Da SAS bereits langjährig eingesetzt wird, ist für die kommenden Performanzvergleiche zu beachten, dass die SAS-Server nicht die neueste Hardware nutzen. Im Jahr 2025 war SAS auf Version 9.4M8 verfügbar.

<sup>6</sup> Ein Datenzuschnitt ist eine Teilmenge der Datengrundlage, die aus projektspezifischen Merkmalen und zufällig ausgewählten Beobachtungen besteht.

## 3.2 Anforderungen an die Analyseinfrastruktur

Damit eine moderne Infrastruktur den technischen Anforderungen des Statistischen Bundesamtes und des FDZ Bund gerecht wird, muss sie ein breites Spektrum an Datenverarbeitungswerkzeugen mit hoher Performanz leicht handhabbar bereitstellen.

Die eingesetzte Infrastruktur muss skalierbar die Verarbeitung wachsender Datenmengen zuverlässig mit den vorhandenen Rechenkapazitäten ermöglichen, damit steigende Datenverarbeitungsbedarfe gedeckt werden. Um eine hohe Reaktionsfähigkeit und eine reibungslose GWAP- beziehungsweise KDFV-Nutzung im FDZ Bund sicherzustellen, sollten die Systeme kurze Rechenzeiten ermöglichen – und idealerweise Auswertungen großer Datenmengen in Echtzeit (das heißt innerhalb weniger Sekunden) ausführen können.

Ein wesentlicher Erfolgsfaktor der Systeme ist eine intuitive Handhabung. Moderne Interfaces und nachvollziehbare Programmiersyntax müssen sicherstellen, dass die Infrastruktur für verschiedene Fachrichtungen leicht zugänglich ist. So kann die simple Nutzung den Gesamtaufwand der Datenauswertung reduzieren, da neben der Rechenzeit auch die Entwicklungszeit berücksichtigt werden muss (Gomolka und andere, 2021).

Ebenso entscheidend sind die Bereitstellungs- und Wartungsaufwände der Systeme innerhalb der abgeschotteten Systemarchitektur des Statistischen Bundesamtes für einen stabilen Einsatz in der Statistikproduktion und die Verfügbarkeit im FDZ Bund.

## 3.3 Erweiterter R- und Python-Server

Die Statistiksoftware R bietet hohe Anwenderorientierung und fokussiert sich auf Datenverarbeitung und -auswertung. Als Open-Source-Software erfreut sich R einer großen internationalen Gemeinschaft, die R um mehr als 20 000 R-Pakete erweitert.<sup>17</sup> Passend zu den Anforderungen aus Abschnitt 3.2 sind

darunter viele R-Pakete, die die Performanz und/oder Handhabung von R verbessern. Auch die amtliche Statistik hat die Vorzüge von R erkannt und die Software erfreut sich nun steigender Beliebtheit in den Statistikämtern.<sup>18</sup> Diese Entwicklungen sind ebenfalls im Statistischen Bundesamt zu beobachten. So steht bereits seit 2020 ein Server für R und Python zur Verfügung, der für die Datenverarbeitung in der Statistikproduktion sowie für die kontrollierte Datenfernverarbeitung im FDZ Bund genutzt wurde.

Im Januar 2025 wurde der interne R- und Python-Server, der exklusiv für die Mitarbeiterinnen und Mitarbeiter des Statistischen Bundesamtes zur Verfügung steht, erheblich erweitert. Dabei wird die kommerziell unterstützte [Posit Workbench](#) verwendet, die kollaboratives Arbeiten und den Betrieb der Systeme vereinfacht. So werden die aktuellsten R-Versionen und R-Pakete in einer modernen Oberfläche bereitgestellt. Zusätzlich wurden die Systemkapazitäten deutlich ausgebaut, sodass Grafikkarten ganz neue Anwendungsfälle ermöglichen. [↪ Tabelle 2](#)

**Tabelle 2**

**Ausbau der R-Server des Statistischen Bundesamtes im Januar 2025**

| Komponente       | Vorher                  | Nachher                                        |
|------------------|-------------------------|------------------------------------------------|
| Anzahl Server    | 1                       | 10                                             |
| Arbeitsspeicher  | 700 Gigabyte            | je 2 Terabyte                                  |
| Kernanzahl       | 72                      | je 96                                          |
| Kerntaktfrequenz | je bis zu 4,0 Gigahertz | je bis zu 3,4 Gigahertz                        |
| Grafikkarte      | keine                   | je zwei Nvidia A6000 (mit je 48 Gigabyte VRAM) |

Mehr als 200 Mitarbeiterinnen und Mitarbeiter setzen den R- und Python-Server inzwischen regelmäßig ein.

## 3.4 On-Site-Server des FDZ Bund

Neben dem internen Server sind seit November 2025 die sogenannten [On-Site-Server](#) im FDZ Bund einsatzfähig. Sie bieten der Wissenschaft dedizierte Stata- und R-Server mit umfangreichen Systemka-

<sup>7</sup> Am 9. Dezember 2025 waren 23 035 Pakete auf [CRAN](#), dem umfassenden R-Archivnetzwerk, verfügbar.

<sup>8</sup> Dies belegt beispielsweise die Entwicklung der Teilnehmezahlen der Konferenz „The Use of R in Official Statistics“: [r-project.ro/conference2025.html](https://r-project.ro/conference2025.html)

pazitäten für die exklusive Nutzung an den Gastwissenschaftsarbeitsplätzen sowie der kontrollierten Datenfernverarbeitung. Alle Analyseserver sind über schnelle Direktverbindungen an einen zentralen Netzwerkspeicher angebunden, sodass R und Stata kombiniert werden können. ➤ **Tabelle 3**

**Tabelle 3**

**On-Site-Server des Forschungsdatenzentrums des Statistischen Bundesamtes: Spezifikationen**

| Komponente       | Anzahl/Größe     |
|------------------|------------------|
| Stata-Server     | 3                |
| R-Server         | 6                |
| Arbeitsspeicher  | je 512 Gigabyte  |
| Kernanzahl       | je 16            |
| Kerntaktfrequenz | je 3,1 Gigahertz |

Stand: November 2025

Stata ist in der effizienten Version „MP“ einheitlich auf leistungsstarken Servern verfügbar. Zudem steht seit 2025 ein R-Server für GWAP-Nutzungen im FDZ Bund bereit. Der Systemaufbau orientiert sich am System aus Abschnitt 3.3. Beide R-Server unterstützen moderne, speicherschonende Werkzeuge, mit denen auch Datenmengen oberhalb des Arbeitsspeichers effizient verarbeitet werden können.

## 4

### Datenverarbeitungswerkzeuge

Im Folgenden wird die Performanz von Datenverarbeitungswerkzeugen, das heißt Softwarelösungen, die Daten einlesen, modifizieren, aggregieren, auswerten oder analysieren,<sup>9</sup> betrachtet.

Ein Datenverarbeitungswerkzeug wird als hochperformant bezeichnet, wenn es Datenmengen über dem verfügbaren Arbeitsspeicher verarbeiten kann und dabei eine flexible und effiziente Nutzung der Ressourcen ermöglicht. Denn für die Skalierung auf grö-

<sup>9</sup> Das Schreiben von Daten wird nicht vertieft behandelt, da die betrachteten Ergebnistabellen in der Regel klein sind. Die genannten Lösungen sind dennoch im Schreiben von Daten ebenfalls effizient. Aufgrund externer Abhängigkeiten werden Datenbanksysteme sowie verteilte Rechenlösungen nicht berücksichtigt. Der Fokus liegt auf den Standard-R-Nutzenden, die mit den üblichen R-Werkzeugen vertraut sind.

ßere Datenmengen ist ein Wechsel von der vollständigen Verarbeitung im Arbeitsspeicher zur flexiblen Auslagerung von Daten nötig. Im Arbeitsspeicher ist zwar eine schnelle Datenverarbeitung möglich, doch das Einlesen großer Datenmengen kann einen Großteil der Gesamtlaufzeit ausmachen, und der verfügbare Arbeitsspeicher ist meist begrenzt. Geeignete Software und moderne Hardware ermöglichen heute sehr schnelle Verarbeitungen von Datenmengen oberhalb des Arbeitsspeichers.

### 4.1 Klassische Werkzeuge in R

Die Kernfunktionen von R (sogenannte Base R) bieten grundlegende Werkzeuge, die für Einsteiger oder in stark abgeschotteten Infrastrukturen zweckdienlich sind. Base R lässt sich in Bezug auf Handhabung und Performanz durch eine Vielzahl von R-Paketen verbessern, wie die folgenden R-Pakete zeigen:

- > **data.table** (Barrett und andere, 2024) gilt als „Goldstandard“ für die schnelle Datenverarbeitung im Arbeitsspeicher sowie das effiziente Einlesen von CSV-Dateien. Dank der Anlehnung an die Base-R-Syntax ist es sehr beliebt. Solange die Daten vollständig in den Arbeitsspeicher passen, ist data.table in den meisten Fällen die schnellste klassische Lösung. Dabei ist einschränkend zu beachten, dass eine CSV-Datei im Arbeitsspeicher ein Vielfaches (etwa das Drei- bis Vierfache) ihrer Größe beansprucht.
- > **tidyverse** (Wickham und andere, 2016) ist ein Ökosystem von R-Paketen, das besonders auf einfache und nachvollziehbare Programmierung setzt. Die eigene dplyr-Syntax ist auch für neue Nutzende gut verständlich. ➤ **Tabelle 4** Diese intuitive Syntax wird von vielen anderen R-Paketen adaptiert. Die exemplarische Syntax gibt, nach ergänzenden Transformationen, die in Tabelle 1 dargestellten Fallzahlen wider.

### 4.2 Hochperformante Werkzeuge in R

Für Datenmengen, die einen Großteil des Arbeitsspeichers belegen, ist die nachfolgend beschriebene Variablenselektion eine oft notwendige Best Practice, die direkt spürbare Performanzgewinne



**Tabelle 4**

Exemplarische Auswertung des Business-Tax-Panels in dplyr-Syntax

| Nr. | Befehl                             | Beschreibung                                                          |
|-----|------------------------------------|-----------------------------------------------------------------------|
| 1   | ergebnis <-                        | Speichert die Fallzahl je Jahr und Statistik des Business-Tax-Panels. |
| 2   | read_csv('daten/btp/daten.csv')  > | Einlesen des Business-Tax-Panels im CSV-Format.                       |
| 3   | group_by(jahr)  >                  | Gruppiert die Daten nach Jahren.                                      |
| 4   | summarize(                         | Aggregiert die Daten auf eine Beobachtung je Gruppe.                  |
| 5   | g = sum(str_detect('g', verk)),    | Zählt in jeder Gruppe die Fälle mit 'g' in verk.                      |
| 6   | k = sum(str_detect('k', verk)),    | Zählt in jeder Gruppe die Fälle mit 'k' in verk.                      |
| 7   | p = sum(str_detect('p', verk)),    | Zählt in jeder Gruppe die Fälle mit 'p' in verk                       |
| 8   | u = sum(str_detect('u', verk)),    | Zählt in jeder Gruppe die Fälle mit 'u' in verk .                     |
| 9   | v = sum(str_detect('v', verk)),    | Zählt in jeder Gruppe die Fälle mit 'v' in verk.                      |
| 10  | r = sum(str_detect('r', verk)),    | Zählt in jeder Gruppe die Fälle mit 'r' in verk.                      |
| 11  | e = sum(str_detect('e', verk)),    | Zählt in jeder Gruppe die Fälle mit 'e' in verk .                     |
| 12  | )                                  |                                                                       |

bringt. Besonders für klassische Werkzeuge kann das Risiko für Arbeitsspeicher-Abbrüche und Wartezeiten deutlich reduziert werden. Zusätzlich ist die Datenaufteilung kombiniert mit sogenannter Lazy-Computation für große Datenmengen oder komplexe Analysen vorteilhaft.

- > **Variablenselektion:** Die Einlesefunktion sollte bei größeren Datenmengen stets auf notwendige Variablen begrenzt werden, um Ladezeit und Arbeitsspeicher-Bedarf massiv zu reduzieren. Das Einlesen zunächst aller Daten (gegebenenfalls mit späterer Variablenselektion) verursacht unnötige Wartezeiten, Ressourcenbelegung und gegebenenfalls Abbrüche. Auch Anwendungsfälle in Stata können von einer Variablenselektion profitieren – zum Beispiel mit einer Beschleunigung um 33,9 % in Anwendungen der Forschungsdatenzentren (Gomolka und andere, 2021).
- > **Datenaufteilung:** Wenn der Arbeitsspeicher nicht ausreicht, um die benötigten Daten einzulesen, kann eine Datenaufteilung auf mehrere Arbeitsspeicher-gerechte Dateien helfen. Hochperformante Werkzeuge können auch direkt mehrere Dateien nutzen, ohne dass sie manuell zusammengeführt werden müssen.
- > **Lazy-Computation:** Ohne Anpassungen in der Programmierung kann die Lazy-Computation geplante Datenverarbeitungsschritte speichern, ohne sie sofort auszuführen. Erst bei expliziter

Abfrage eines Ergebnisses führt das System die Berechnung durch. Dadurch werden nur tatsächlich benötigte Daten verarbeitet und unnötige Schritte vermieden – eine manuelle Variablenauswahl ist nicht erforderlich.

[Apache Parquet](#) ist ein spaltenbasiertes, komprimiertes Open-Source-Datenformat, das speziell für effiziente Speicherung und Verarbeitung großer Datenmengen entwickelt wurde. Im Folgenden wird das Format vereinfacht als Parquet bezeichnet. Durch eine spaltenbasierte Struktur ermöglicht Parquet den schnellen Zugriff auf einzelne Merkmale und reduziert gleichzeitig den Speicherbedarf erheblich. Gomolka und andere (2021) zeigten im Kontext der Forschungsdatenzentren massive Geschwindigkeitsgewinne anhand Parquet, sodass im Forschungsdaten- und Servicezentrum der Deutschen Bundesbank bereits standardmäßig Parquet-Dateien angeboten werden.

Für die hochperformante Verarbeitung von Datenmengen oberhalb des Arbeitsspeichers werden die folgenden Werkzeuge betrachtet, die in Form von R-Paketen verfügbar sind:

- > **arrow** (Richardson und andere, 2024) nutzt speziell für die Arbeit mit großen Datensätzen intelligente Datenaufteilung und Parallelisierung durch C++ im Hintergrund. Es verarbeitet Daten spaltenweise und verwendet Lazy-Computation, was



besonders bei aufwendigen analytischen Abfragen auf großen Datensätzen Vorteile bietet. arrow nutzt standardmäßig dplyr-Syntax (in Tabelle 4 muss lediglich „csv“ durch „parquet“ ersetzt werden). arrow ist oft die beste Wahl in R, um mehrere Parquet-Dateien zu nutzen, die insgesamt ein Vielfaches größer sind als der vorhandene Arbeitsspeicher.

- > **duckdb** (Mühleisen/Raasveldt, 2024) ist ein eingebettetes Datenbanksystem, das ohne separaten Datenbankserver direkt in R ebenfalls effizient Daten spaltenweise verarbeitet. Nativ nutzt duckdb SQL-Syntax. Durch ergänzende R-Pakete kann auch dplyr-Syntax verwendet werden. Eine direkte Schnittstelle zu arrow ermöglicht die nahtlose Kombination von arrow und duckdb.
- > **polars** (Shima und andere, 2024) ist ein jüngeres R-Paket, das anhand der Programmiersprache Rust für hohe Geschwindigkeit und sparsame Ressourcennutzung sorgt. Nativ nutzt es Python-orientierte Syntax und kann auch durch das Paket tidypolars minimal verlangsamt dplyr-Syntax verwenden.

### 4.3 Externe Benchmarks

Öffentlich verfügbare Benchmarks können zu abweichenden Ergebnissen führen.<sup>10</sup> Daher ist die Untersuchung in der eigenen Infrastruktur mit relevanten Anwendungsfällen (insbesondere Datenmengen) und den verfügbaren Software-Versionen wichtig. Die zum Zeitpunkt des Performanzvergleichs (Kapitel 5) betrachteten R-Paket-Versionen können im Statistischen Bundesamt aus organisatorischen Gründen nicht flexibel angepasst werden. Sie sind im Literaturverzeichnis explizit aufgeführt. Entscheidend ist bei der Versionswahl nicht das absolute Performanz-Optimum, sondern eine Version mit langfristig stabiler und verlässlich guter Performanz.

---

10 Benchmark von data.table ([cran.r-project.org](https://cran.r-project.org)), duckdb ([duckdblabs.github.io](https://github.com/duckdblabs/duckdb)) und polars ([ddotta.github.io](https://github.com/pola-rs/polars)).

## 5

### Performanzvergleich

Um zwischen den in Kapitel 4 vorgestellten Datenverarbeitungswerkzeugen abzuwägen, wurde deren Performanz in einem umfangreichen Vergleich auf dem SAS-Server des Statistischen Bundesamtes und den On-Site-R- und Stata-Servern des FDZ Bund exploriert. Dabei wurde schrittweise die Datenmenge erhöht und nur Werkzeuge weiterbetrachtet, die eine hohe Performanz aufzeigten. Weiterführende Performanzmessungen sowie der vollständig reproduzierbare Quellcode sind auf Opencode verfügbar.<sup>11</sup>

#### 5.1 Versuchsaufbau

Für die Untersuchung wurden simulierte Einzeldaten in verschiedenen Größen (1 % sowie 10 % der insgesamt 67 Millionen Beobachtungen des Business-Tax-Panels) betrachtet. Sie wurden aus öffentlich verfügbaren Informationen generiert und bilden die wesentlichen Strukturen des Business-Tax-Panels für die Analyse künstlich nach. Abgrenzend zu den im FDZ Bund bereitgestellten Datenstrukturfiles wurden die Testdaten nicht aus echten Mikrodaten abgeleitet. Der simulierte Datensatz ist ausschließlich für technische Testzwecke der folgenden Anwendungsfälle (siehe Kapitel 2) bestimmt:

1. Fallzahlen (je Statistik und Jahr; siehe Tabelle 1),
2. Regression (Einflussfaktoren auf Verlustvorträge).

Um eine zuverlässige Performanz-Messung zu erhalten, wurde jedes Datenverarbeitungswerkzeug dreifach angewandt. Der Vergleich erfolgt anhand der mittleren Messung der Laufzeit (real verstrichene Zeit).

Daneben ist der beanspruchte Arbeitsspeicher der Auswertung entscheidend, da ein übermäßiger Arbeitsspeicherverbrauch zu Abbrüchen führen kann. Dennoch wurde der Arbeitsspeicher-Verbrauch

---

11 [gitlab.opencode.de](https://gitlab.opencode.de)

nur für ausgewählte Tests anhand eines Systemmonitors beobachtet.<sup>12</sup>

Das Ziel war, Werkzeuge mit einer Mindestperformanz in verschiedenen Szenarien zu finden und nicht in jedem Anwendungsfall ein Werkzeug für die schnellste Verarbeitung zu bestimmen. Denn der Einsatz von möglichst wenigen verschiedenen Werkzeugen reduziert Entwicklungs- und Wartungsaufwände.

Feine Abstufungen der Zielgrößen sind für die Aufgaben nicht gleichermaßen relevant. Sobald eine Mindestperformanz erreicht wurde, rückt die Reduktion des Entwicklungsaufwands in den Vordergrund. Entsprechend sind Aspekte wie einfache Anwendbarkeit und Nachvollziehbarkeit – auch für unerfahrene Nutzende – ausschlaggebend.

## 5.2 Vergleich von R zu SAS und Stata

Zunächst werden 1 % der simulierten BTP-Daten (etwa 700 000 Beobachtungen; 12 Gigabyte als CSV-Datei) geprüft. Um die nachfolgenden Ergebnisse einzuordnen, wurden die Performanzmessungen auch in SAS und Stata als Vergleichsmaßstab durchgeführt. ➔ **Tabelle 5**

SAS zeigt seine Stärke anhand der geringen Beanspruchung des Arbeitsspeichers. Mit etwa 55 Sekunden je Auswertung ist Stata aber etwa doppelt so schnell wie SAS.

Für die Datenmenge von 1 % der simulierten BTP-Daten beschleunigen alle betrachteten R-Datenverarbeitungswerkzeuge die Berechnungen um mindestens 76 % gegenüber SAS und Stata. ➔ **Tabelle 6** Die

schnellsten Ausführungen benötigen in R für Fallzahlen oder Regression etwa eine Sekunde.

Aus weiteren Untersuchungen für 1 % der simulierten Daten lassen sich die folgenden Erkenntnisse ableiten:

1. Das **Parquet-Dateiformat** ermöglicht die schnellsten Ergebnisse und spart gleichzeitig 50 % Festplattenspeicher ein.
2. Die **Datenaufteilung** verlangsamt für geringere Datenmengen die Berechnung der Fallzahlen. In der Berechnung der Regression profitiert arrow von einer Aufteilung.
3. Die **Variablenauswahl** zeigte immer eine Reduktion der Laufzeit und Arbeitsspeicher-Beanspruchung, schränkt aber die Flexibilität der Auswertung ein.
4. Unter den **klassischen Werkzeugen** war data.table deutlich am schnellsten<sup>13</sup>, benötigte aber etwa die dreifache Größe der CSV-Datei im Arbeitsspeicher. data.table wird hier durch Variablenauswahl konkurrenzfähig zu den besten Methoden.
5. In der Berechnung der Fallzahl sind alle hochperformanten Werkzeuge – arrow, duckdb und polars – nahezu gleichwertig und benötigen höchstens 2 Sekunden für das Ergebnis.
6. In der Berechnung der Regression war duckdb mit nur etwa 1 Sekunde deutlich am schnellsten. Hier ist der klassische Ansatz (data.table mit Variablenselektion) sogar schneller als arrow.
7. Bei der Regression wurde polars nicht betrachtet, da es im R-Kontext weniger intuitiv zu nutzen war als die anderen Werkzeuge und die Implementierung aufgrund von Fehlern nicht durchgelaufen ist.

<sup>12</sup> Im Rahmen des Experiments konnte der Arbeitsspeicher-Verbrauch nicht systematisch korrekt in R gemessen werden, da der Verbrauch von C++ oder Rust nicht mit R-Paketen gemessen werden kann. Dies zeigt eine Warnung des R-Pakets bench in dieser Benchmark: [tidypolars.etiennebacher.com](https://tidypolars.etiennebacher.com)

<sup>13</sup> Übersichtshalber wird data.table als repräsentativer Vergleichswert für klassische Werkzeuge betrachtet, da das R-Paket deutlich die beste Performanz zeigte. Der Vergleich zu anderen klassischen Werkzeugen ist zu finden auf [gitlab.opencode.de](https://gitlab.opencode.de)

**Tabelle 5**

Baselines in SAS und Stata (anhand 1 % des simulierten Business-Tax-Panels)

| Software | Anwendung  | Arbeitsspeicher | Laufzeit      | Festplattenspeicher |
|----------|------------|-----------------|---------------|---------------------|
| SAS      | Fallzahlen | 65 Megabyte     | 2,4 Minuten   | 33 Gigabyte         |
|          | Regression | 7 Megabyte      | 2,8 Minuten   | 33 Gigabyte         |
| Stata    | Fallzahlen | 33 Gigabyte     | 55,2 Sekunden | 32 Gigabyte         |
|          | Regression | 33 Gigabyte     | 55,8 Sekunden | 32 Gigabyte         |

**Tabelle 6**

**Benchmarks für die Berechnung der Fallzahlen und der Regression anhand 1 % des simulierten Business-Tax-Panels**

| R-Paket    | Dateiformat | Anpassung                      | Laufzeit       | Arbeitsspeicher | Festplatte  |
|------------|-------------|--------------------------------|----------------|-----------------|-------------|
|            |             |                                | Fallzahlen     |                 |             |
| polars     | Parquet     |                                | 0,97 Sekunden  | –               | 5 Gigabyte  |
| duckdb     | Parquet     |                                | 1,68 Sekunden  | –               | 5 Gigabyte  |
| duckdb     | Parquet     | Datenaufteilung (Berichtsjahr) | 1,94 Sekunden  | –               | 4 Gigabyte  |
| arrow      | Parquet     | Variablenauswahl               | 2,07 Sekunden  | –               | 5 Gigabyte  |
| data.table | CSV         | Variablenauswahl               | 2,38 Sekunden  | 41 Megabyte     | 12 Gigabyte |
| arrow      | Parquet     |                                | 2,72 Sekunden  | –               | 5 Gigabyte  |
| arrow      | Parquet     | Datenaufteilung (gleichmäßig)  | 2,99 Sekunden  | –               | 5 Gigabyte  |
| arrow      | Parquet     | Datenaufteilung (Berichtsjahr) | 3,42 Sekunden  | –               | 4 Gigabyte  |
| arrow      | Parquet     |                                | 4,18 Sekunden  | –               | 5 Gigabyte  |
| data.table | CSV         |                                | 8,76 Sekunden  | 31 Gigabyte     | 12 Gigabyte |
| data.table | CSV         | Datenaufteilung (Berichtsjahr) | 13,32 Sekunden | 47 Gigabyte     | 10 Gigabyte |
|            |             |                                | Regression     |                 |             |
| duckdb     | Parquet     |                                | 1,13 Sekunden  | –               | 5 Gigabyte  |
| data.table | CSV         | Variablenauswahl               | 2,90 Sekunden  | 552 Megabyte    | 12 Gigabyte |
| arrow      | Parquet     | Datenaufteilung (Berichtsjahr) | 4,02 Sekunden  | –               | 4 Gigabyte  |
| data.table | CSV         |                                | 7,51 Sekunden  | 32 Gigabyte     | 12 Gigabyte |
| arrow      | Parquet     |                                | 13,35 Sekunden | –               | 5 Gigabyte  |

Anhand der hier betrachteten Datenmenge von 700 000 Beobachtungen können die schnellsten Werkzeuge kaum differenziert werden. Daher wurde im nächsten Experiment für ausgewählte Werkzeuge die Datenmenge auf 7 000 000 Beobachtungen erhöht (etwa 10 % des Business-Tax-Panels).

### 5.3 Größere Datenmengen in R

Auch für 10 % der simulierten BTP-Daten können hochperformante Werkzeuge in R beide Anwendungsfälle in unter 10 Sekunden beantworten. [Tabelle 7](#) Die Auswertungswerkzeuge mit der höchsten Performanz sind:

- > arrow mit nur 4 Sekunden für die Berechnung der Fallzahl,
- > duckdb mit nur 5,2 Sekunden für die Berechnung der Regression.

Für 10 % der simulierten Daten lassen sich weitere Erkenntnisse ableiten:

1. Das Parquet-Dateiformat ermöglicht die schnellsten Ergebnisse und reduziert die Datenmengen

um 75 Gigabyte gegenüber den CSV-Dateien (etwa 62 %).

2. duckdb verbraucht weniger als 1 Gigabyte Arbeitsspeicher und liefert beide Ergebnisse in etwa 5 Sekunden. Eine Datenaufteilung hat das Ergebnis verschlechtert.
3. Mit der schnellsten Fallzahlberechnung zeigt arrow seine Kernkompetenz, mehrere Parquet-Dateien besonders effizient nutzen zu können. Die spezifische Aufteilung hat einen großen Einfluss auf Arbeitsspeicher-Bedarf und Laufzeit. Unerwarteterweise hat eine zufällige Aufteilung bessere Ergebnisse erzielt als die Aufteilung nach Berichtsjahren.
4. Klassische Werkzeuge kommen bei der Größe der CSV-Datei von 121 Gigabyte bereits an Kapazitätsgrenzen der Server. Somit ist bei der Anwendung klassischer Methoden deutlich mehr Erfahrung notwendig, um zwischen Variablenauswahl und Arbeitsspeicher-Verbrauch abzuwägen.
5. Die Laufzeit von 41 Sekunden von polars ist um den Faktor 10 langsamer als arrow und duckdb. Da polars das jüngste dieser R-Pakete ist, könnte

**Tabelle 7**

**Benchmarks für die Berechnung der Fallzahlen und der Regression anhand 10 % des simulierten Business-Tax-Panels**

| R-Paket    | Dateiformat | Anpassung                      | Laufzeit              | Arbeitsspeicher | Festplatte   |
|------------|-------------|--------------------------------|-----------------------|-----------------|--------------|
| Fallzahlen |             |                                |                       |                 |              |
| arrow      | Parquet     | Datenaufteilung (gleichmäßig)  | 4,00 Sekunden         | 2,7 Gigabyte    | 45 Gigabyte  |
| duckdb     | Parquet     |                                | 5,39 Sekunden         | 400 Megabyte    | 46 Gigabyte  |
| arrow      | Parquet     | Variablenauswahl               | 7,45 Sekunden         | –               | 46 Gigabyte  |
| duckdb     | Parquet     | Datenaufteilung (Berichtsjahr) | 7,57 Sekunden         | –               | 45 Gigabyte  |
| arrow      | Parquet     | Datenaufteilung (Berichtsjahr) | 9,61 Sekunden         | 8,3 Gigabyte    | 45 Gigabyte  |
| arrow      | Parquet     |                                | 14,85 Sekunden        | 163 Gigabyte    | 46 Gigabyte  |
| data.table | CSV         | Variablenauswahl               | 20,80 Sekunden        | 408 Megabyte    | 121 Gigabyte |
| polars     | Parquet     |                                | 41,02 Sekunden        | 186 Gigabyte    | 46 Gigabyte  |
| data.table | CSV         |                                | 53,04 Sekunden        | 313 Gigabyte    | 121 Gigabyte |
| data.table | CSV         | Datenaufteilung (Berichtsjahr) | 4 Minuten 41 Sekunden | 472 Gigabyte    | 97 Gigabyte  |
| Regression |             |                                |                       |                 |              |
| duckdb     | Parquet     |                                | 5,19 Sekunden         | 760 Megabyte    | 46 Gigabyte  |
| arrow      | Parquet     | Datenaufteilung (Berichtsjahr) | 7,21 Sekunden         | 8 Gigabyte      | 45 Gigabyte  |
| data.table | CSV         | Variablenauswahl               | 27,16 Sekunden        | 5 Gigabyte      | 121 Gigabyte |
| data.table | CSV         |                                | 1 Minuten 12 Sekunden | 317 Gigabyte    | 121 Gigabyte |
| arrow      | Parquet     |                                | 2 Minuten 59 Sekunden | 336 Gigabyte    | 46 Gigabyte  |

ein Update der vorliegenden Version 1.0.1<sup>14</sup> dies beheben.

## 6

### Ergebnis für das vollständige Business-Tax-Panel

Hochperformante R-Werkzeuge sind in der Lage, die betrachteten Anwendungsfälle in Echtzeit zu beantworten. Die Verarbeitung von Parquet-Dateien mit arrow und duckdb für die Auswertung großer Datenmengen wird empfohlen. Parquet-Dateien haben in allen Untersuchungen deutlich schnellere

Auswertungen ermöglicht und gleichzeitig den Speicherbedarf reduziert. Arrow und duckdb können ohne Verzögerungen kombiniert werden und bieten die nachvollziehbare dplyr-Syntax. Für besonders große Datenmengen kann arrow aufgeteilte Parquet-Dateien besonders gut nutzen.

Das vollständige Business-Tax-Panel 2013 bis 2019 hat einen Umfang von 610 Gigabyte in Stata-Dateien, sodass selbst der Arbeitsspeicher der On-Site-Stata-Server mit 512 Gigabyte nicht ausreicht, um das Business-Tax-Panel vollständig in Stata einzulesen. Dadurch muss in Stata zwingend mit Variablenselektion gearbeitet werden. Für SAS liegt das Business-Tax-Panel in einzelnen komprimierten Jahresscheiben vor und belegt insgesamt 314 Gigabyte Festplattenspeicher. SAS kann die betrachteten Anwendungsfälle für diese Datenmengen weiterhin durchführen und erreicht die in [Tabelle 8](#) dargestellte Performanz.

<sup>14</sup> Das R-Package polars in der verwendeten Version 1.0.1 war nur der erste Minor Release nach einer kompletten Überarbeitung. Inzwischen sind fünf neuere Major Releases mit diversen Performanz-Verbesserungen verfügbar.

**Tabelle 8**

**Business-Tax-Panel 2013 bis 2019 in SAS als komprimierte Jahresscheiben**

| Software | Anwendung  | Arbeitsspeicher | Laufzeit   | Festplattenspeicher |
|----------|------------|-----------------|------------|---------------------|
| SAS      | Fallzahlen | 1 Gigabyte      | 93 Minuten | 314 Gigabyte        |
| SAS      | Regression | 141 Megabyte    | 92 Minuten | 314 Gigabyte        |

**Tabelle 9**

**Business-Tax-Panel 2013 bis 2019 in R mit arrow und Berichtsjahren in Parquet-Dateien**

| Software | Anwendung  | Arbeitsspeicher  | Laufzeit      | Festplattenspeicher |
|----------|------------|------------------|---------------|---------------------|
| R        | Fallzahlen | etwa 7 Gigabyte  | 5,5 Sekunden  | 17,5 Gigabyte       |
| R        | Regression | etwa 10 Gigabyte | 11,8 Sekunden | 17,5 Gigabyte       |

Anhand einer Aufteilung von Parquet-Dateien je Berichtsjahr zeigt arrow für das gesamte Business-Tax-Panel erst das volle Potenzial von R-Werkzeugen.

↘ **Tabelle 9.**

Gegenüber SAS reduziert R die Laufzeit von über 90 Minuten auf etwa 12 Sekunden mit dennoch geringem Arbeitsspeicher-Bedarf (unter 10 Gigabyte). Die Parquet-Dateien beanspruchen nur noch 17,5 Gigabyte Speicher. Hier lässt sich beobachten, dass arrow und Parquet-Dateien für große Datenmengen höhere Effizienz bieten. Obwohl gegenüber Abschnitt 5.3 die Datenmenge um 90 % erhöht wurde, steigt die Laufzeit der Berechnung der Fallzahlen nur um eine Sekunde (38 %) und der Regression um 4,6 Sekunden (64 %).

## 7

### Fazit

Eine Datenhaltung im Parquet-Format für Datenverarbeitungen beziehungsweise -auswertungen in der Statistikproduktion und in den Forschungsdatenzentren ist gewinnbringend. Es reduziert den notwendigen Festplattenspeicher stark und mehrere R-Pakete (arrow, duckdb und polars) können Daten großen Umfangs im Parquet-Format sehr schnell verarbeiten. Daher wird empfohlen, eine Datenhaltung im Parquet-Format und eine Datenverarbeitung in R für große Datensätze in der amtlichen Statistik und in den Forschungsdatenzentren zu erproben. Denn die Untersuchung demonstriert, dass dadurch relevante analytische Fragestellungen anhand großer Datenmengen in Echtzeit ressourcenschonend beantwortet werden können.

Das R-Paket arrow (kombiniert mit duckdb) zeigte die intuitivste Handhabung, sodass die Entwicklungsaufwände am geringsten waren. Das neuere R-Paket

polars<sup>15</sup> bietet eine vielversprechende Alternative, die sich bei kleineren Datenmengen als sehr schnell erwiesen hat. In der vorliegenden Version zeigte es jedoch bei größeren Datenmengen keine durchgängig stabile Performanz, und die Umsetzung in R war vergleichsweise weniger intuitiv. Dadurch wird deutlich, dass es für einen optimalen Einsatz der Werkzeuge entscheidend ist, die Ergebnisse des Statistischen Bundesamtes mit der Open-Source-Gemeinschaft zu teilen, sodass die Anwendungen reproduziert und verbessert werden können.

In der Statistikproduktion und den Forschungsprojekten der Forschungsdatenzentren bieten die vorgestellten Werkzeuge das Potenzial, Effizienz erheblich zu steigern sowie die Nachvollziehbarkeit der Programmierung zu erhöhen. Dafür ist der Aufbau eines Datenbestands in Parquet notwendig und die schrittweise Ausweitung auf weitere Statistiken. Begleitend sind Schulungsangebote für die Nutzenden und Mitarbeitenden der Forschungsdatenzentren und der amtlichen Statistik wesentlich für den Erfolg der Umsetzung. Nutzende der effizienten Werkzeuge können dann sekundenschnelle Auswertungen großer Datensätze im Vollmaterial durchführen, wodurch sich Analysen in der empirischen Forschung mithilfe von R sowohl beschleunigen als auch ausweiten lassen.

Die Untersuchung hat gezeigt, dass R als zentrales Analysewerkzeug im FDZ Bund und im Statistischen Bundesamt eingesetzt werden kann. Da R vielseitig hohe Performanz und intuitive Handhabung bietet, kann ein Fokus auf R Hardwarekosten senken und die Softwarevielfalt verschlanken. Die Analyse großer Datensätze in der amtlichen Statistik und den Forschungsdatenzentren kann von einer massiven Performanz-Verbesserung profitieren, wenn die vorgestellten Werkzeuge von Infrastrukturbetreibern, amtlicher Statistik, Forschungsdatenzentren und Forschenden adaptiert werden. 

15 Siehe Fußnote 14.

## LITERATURVERZEICHNIS

---

Barrett, Tyson/Dowle, Matt/Srinivasan, Arun/Gorecki, Jan/Chirico, Michael/Hocking, Toby/Schwendinger, Benjamin/Krylov, Ivan. *data.table: Extension of 'data.frame'* (Version 1.17.8). 2024. [Zugriff am 16. Januar 2026]. Verfügbar unter: [CRAN.R-project.org](https://CRAN.R-project.org)

Forschungsdatenzentrum der Statistischen Ämter des Bundes und der Länder. *Metadatenreport. Teil II. Produktspezifische Informationen zum Business-Tax-Panel 2013–2019 für die On-Site-Nutzung. Version I.* Wiesbaden 2024. [Zugriff am 16. Januar 2026]. Verfügbar unter: [www.forschungsdatenzentrum.de](https://www.forschungsdatenzentrum.de)

Gomolka, Matthias/Blaschke, Jannick/Hirsch, Christian. *Working with Large Data at the RDSC*. In: Deutsche Bundesbank (Herausgeber). Technical Report 2021-04. [Zugriff am 16. Januar 2026]. Verfügbar unter: [www.bundesbank.de](https://www.bundesbank.de)

Kristiansen, Annette. [\*Business-Tax-Panel – Zusammenführung von Unternehmenssteuerstatistiken\*](#). In: WISTA Wirtschaft und Statistik. Ausgabe 4/2023, Seite 47 ff.

Mühleisen, Hannes/Raasveldt, Mark. *duckdb: DBI Package for the Duckdb Database Management System (Version 1.3.2)*. 2024. [Zugriff am 16. Januar 2026]. Verfügbar unter: [CRAN.R-project.org](https://CRAN.R-project.org)

Richardson, Neal/Cook, Ian/Crane, Nic/Dunnington, Dewey/François, Romain/Keane, Jonathan/Mecum, Bryce und andere. *Arrow: Integration to 'Apache' 'Arrow' (Version 20.0.0.2)*. 2024. [Zugriff am 16. Januar 2026]. Verfügbar unter: [CRAN.R-project.org](https://CRAN.R-project.org)

R Core Team. *R: A language and environment for statistical computing (Version 4.5.0)*. R Foundation for Statistical Computing, Vienna, Austria. 2021. [Zugriff am 16. Januar 2026]. Verfügbar unter: [www.R-project.org](https://www.R-project.org)

Sachverständigenrat zur Begutachtung der gesamtwirtschaftlichen Entwicklung. *Jahresgutachten 2023/24: Wachstumsschwäche überwinden – In die Zukunft investieren*. 2023. [Zugriff am 16. Januar 2026]. Verfügbar unter: [www.sachverstaendigenrat-wirtschaft.de](https://www.sachverstaendigenrat-wirtschaft.de)

Shima, Tatsuya/Bacher, Etienne und andere. *polars: R Bindings for the 'polars' Rust Library (1.0.1)*. 2024. [Zugriff am 16. Januar 2026]. Verfügbar unter: [pola-rs.github.io](https://pola-rs.github.io)

Wickham, Hadley/Averick, Mara/Bryan, Jennifer/Chang, Winston/ D'Agostino McGowan, Lucy/François, Romain/Grolemund, Garrett und andere. *Welcome to the tidyverse (Version 2.0.0)*. In: Journal of Open Source Software. Ausgabe 4 (43): 1686. 2019. DOI: [10.21105/joss.01686](https://doi.org/10.21105/joss.01686).

Wissenschaftlicher Beirat beim Bundesministerium der Finanzen. *Notwendigkeit, Potenzial und Ansatzpunkte einer Verbesserung der Dateninfrastruktur für die Steuerpolitik*. 2020. Gutachten. [Zugriff am 16. Januar 2026]. Verfügbar unter: [www.bundesfinanzministerium.de](https://www.bundesfinanzministerium.de)



**Herausgeber**

Statistisches Bundesamt (Destatis), Wiesbaden

---

**Schriftleitung**

Dr. Daniel Vorgrimler

Redaktion: Ellen Römer

---

**Ihr Kontakt zu uns**

[www.destatis.de/kontakt](http://www.destatis.de/kontakt)

---

**Erscheinungsfolge**

zweimonatlich, erschienen im Februar 2026

Ältere Ausgaben finden Sie unter [www.destatis.de](http://www.destatis.de) sowie in der [Statistischen Bibliothek](#).

---

Artikelnummer: 1010200-26004-4, ISSN 1619-2907

---

© Statistisches Bundesamt (Destatis), 2026

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.