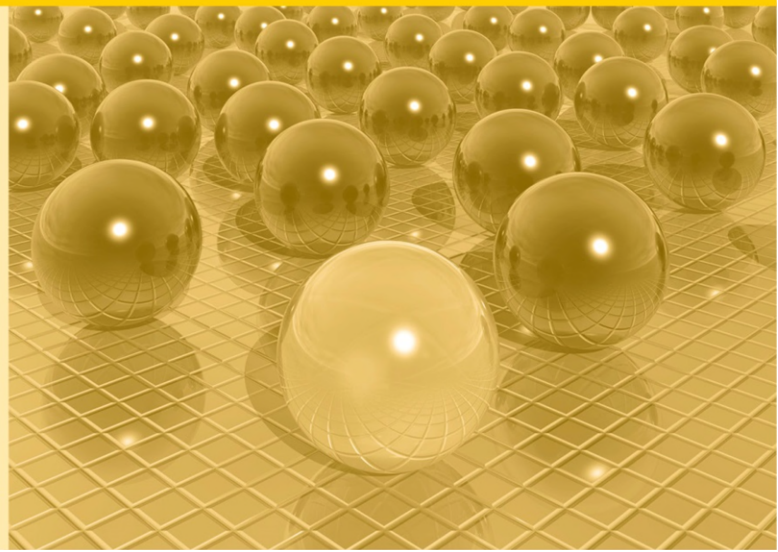


Metadatenreport



Teil I: Allgemeine und methodische Informationen zur Verdiensterhebung -
Beschäftigtenpanel (EVAS-Nummer: 62361)

Version

Impressum

Herausgeber: Statistische Ämter des Bundes und der Länder
Herstellung: Information und Technik Nordrhein-Westfalen
Telefon 0211 9449-01 • Telefax 0211 9449-8000
Internet: www.forschungsdatenzentrum.de
E-Mail: forschungsdatenzentrum@it.nrw.de

Fachliche Informationen zu dieser Veröffentlichung:

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Standort Hessen –
Tel.: 0611 3802-822
Fax: 0611 3802-890
forschungsdatenzentrum@statistik.hessen.de

Informationen zum Datenangebot:

Statistisches Bundesamt
Forschungsdatenzentrum

Tel.: 0611 75-2420
Fax: 0611 72-3915
forschungsdatenzentrum@destatis.de

Forschungsdatenzentrum der
Statistischen Ämter der Länder
– Geschäftsstelle –
Tel.: 0211 9449-2883
Fax: 0211 9449-8087
forschungsdatenzentrum@it.nrw.de

Erscheinungsfolge: unregelmäßig
Erschienen im Oktober 2025

Diese Publikation wird kostenlos als PDF-Datei zum Download unter www.forschungsdatenzentrum.de angeboten.

© Information und Technik Nordrhein-Westfalen, Düsseldorf, 2025
(im Auftrag der Herausbergemeinschaft)

Vervielfältigung und Verbreitung, nur auszugsweise, mit Quellenangabe gestattet. Alle übrigen Rechte bleiben vorbehalten.

Fotorechte Umschlag: ©artSILENCEcom –Fotolia.com

Empfohlene Zitierung:

Forschungsdatenzentren der Statistischen Ämter des Bundes und der Länder: Metadatenreport. Teil I: Allgemeine und methodische Informationen zur Verdiensterhebung - Beschäftigtenpanel (EVAS-Nummer: 62361). Version 1. Wiesbaden 2025.

Metadatenreport

Teil I: Allgemeine und methodische Informationen zur Verdiensterhebung - Beschäftigtenpanel (EVAS-Nummer: 62361)

Version 1

Inhalt

1. Allgemeine Informationen	2
1.1 Ziel/Zweck der Statistik	2
1.2 Rechtsgrundlagen (Verlinkungen).....	2
1.3 Erhebungsart.....	3
1.4 Erhebungseinheit/Auskunftsgebende/Erhebungsgesamtheit.....	3
1.5 Berichtskreis/Berichtsweg	4
1.6 Berichtszeitraum/Berichtszeitpunkt	6
1.7 Periodizität	7
1.8 Regionale Ebene	7
2. Methodik	7
2.1 Erhebungsmethoden	7
2.2 Erhebungsinhalte.....	8
2.3 Auswahlgrundlagen.....	8
2.4 Methoden der Stichprobenziehung in der VE	8
2.5 Aufbereitungsverfahren	9
2.5.1 Datenaufbereitung	10
2.5.2 Methodik der Verknüpfung.....	10
2.5.3 Imputation.....	11
2.6 Hochrechnungen	13
2.7 Methodische Änderungen	15
2.8 Klassifikationen	16
2.9 Fachliche und räumliche Vergleichbarkeit.....	16
3. Qualität.....	16
4. Zentrale Veröffentlichungen.....	16
5. Angebote der FDZ	17

1. Allgemeine Informationen

Die Verdiensterhebung liefert monatlich umfangreiche Daten zu rund 9 Millionen Beschäftigtenverhältnissen und enthält detaillierte Angaben unter anderem zu Verdiensten, beruflicher Tätigkeit und Arbeitsstunden. Das Beschäftigtenpanel, eine Verdienstverlaufsstatistik, deren Grundlage die Verdiensterhebung ist, ermöglicht die Analyse individueller Verdienstentwicklungen. Diese Daten zu Verdienstverläufen sind zentral für wirtschafts- und sozialpolitische Analysen und bilden die Grundlage für die Entwicklung politischer Maßnahmen in der Einkommens- und Arbeitsmarktpolitik. Ein besonderes Interesse liegt in der Betrachtung der Verdienstverläufe im Niedriglohnbereich und den Auswirkungen des gesetzlichen Mindestlohns.

1.1 Ziel/Zweck der Statistik

Ziel der Statistik sind Aussagen über

- o Verdienstverläufe,
- o Bestimmungsfaktoren der individuellen Verdiensthöhe,
- o Bewertung des Einflusses von arbeitsmarktpolitischen Instrumenten und exogenen Schocks,
- o Verteilung und Streuung der Verdienste,
- o Unterschiede zwischen verschiedenen Bundesländern und Branchen.

1.2 Rechtsgrundlagen (Verlinkungen)

Verdienststatistikgesetz (VerdStatG) vom 21. Dezember 2006 (BGBl. I S. 3291) in Verbindung mit dem Bundesstatistikgesetz (BStatG) vom 22. Januar 1987 (BGBl. I S. 462, 565) in der im Erhebungszeitraum gültigen Fassung. Die rechtliche Verpflichtung zur Erstellung einer Verdienstverlaufsstatistik

ergibt sich aus § 6 VerdStatG. Erhoben werden die Angaben zu § 4 Absatz 3 VerdStatG.

<http://www.gesetze-im-internet.de/verdstatg/>

http://www.gesetze-im-internet.de/bstatg_1987/

1.3 Erhebungsart

Primärerhebung von Betrieben mit sozialversicherungspflichtigen Beschäftigten anhand einer jährlich rollierenden geschichteten Stichprobenziehung in den Wirtschaftszweigabschnitten A bis S (WZ 2008).

Eine Ausnahme stellen die WZ-Abschnitte O „Öffentliche Verwaltung, Verteidigung; Sozialversicherung“ und P „Erziehung und Unterricht“ dar. Im Abschnitt O erfolgte komplett, im Abschnitt P zum größten Teil, eine Auswertung der Personalstandstatistik. Eine weitere Ausnahme stellen Betriebe ohne sozialversicherungspflichtige, aber mit geringfügig entlohnnten Beschäftigten dar. Die Daten dieser Betriebe werden im Berichtsmonat April imputiert.

Im Rahmen des Beschäftigtenpanels werden sowohl die Daten der Personalstandstatistik als auch die Daten der Betriebe ohne sozialversicherungspflichtig Beschäftigte aber mit geringfügig entlohnnten Beschäftigten nicht implementiert, da eine Beobachtung individueller Beschäftigungsverhältnisse über mehrere Monate nicht möglich ist.

1.4 Erhebungseinheit/Auskunftsgebende/Erhebungsgesamtheit

Betrieb / Für Verdienstabrechnung im Betrieb zuständige Abteilung, Steuerberatung, zentrale Personalabrechnungsstellen / Abhängige Beschäftigungsverhältnisse in den Abschnitten A-S ohne den Abschnitt O und Teile von P der WZ 2008.

1.5 Berichtskreis/Berichtsweg

Der Berichtskreis bezieht Betriebe der Wirtschaftszweigabschnitte A bis S mit mindestens einem oder einer sozialversicherungspflichtig Beschäftigten ein. Im Einzelnen gehören die Wirtschaftszweige der folgenden Abschnitte der WZ 2008 zum Berichtskreis:

A - Land- und Forstwirtschaft, Fischerei

B - Bergbau und Gewinnung von Steinen und Erden

C - Verarbeitendes Gewerbe

D - Energieversorgung

E - Wasserversorgung; Abwasser- und Abfallentsorgung und Beseitigung von Umweltverschmutzungen

F - Baugewerbe

G - Handel; Instandhaltung und Reparatur von Kraftfahrzeugen

H - Verkehr und Lagerei

I - Gastgewerbe

J - Information und Kommunikation

K - Erbringung von Finanz- und Versicherungsdienstleistungen

L - Grundstücks- und Wohnungswesen

M - Erbringung von freiberuflichen, wissenschaftlichen und technischen Dienstleistungen

N - Erbringung von sonstigen wirtschaftlichen Dienstleistungen

P - Erziehung und Unterricht

Q - Gesundheits- und Sozialwesen

R - Kunst, Unterhaltung und Erholung

S - Erbringung von sonstigen Dienstleistungen

Die Betriebe der WZ-Abschnitte O (komplett) und P (größter Teil von 85.1 bis 85.4) werden nicht direkt befragt, sondern mit Daten der Personalstandstatistik erfasst. Somit sind Betriebe dieser WZ-Abschnitte im Beschäftigtenpanel nicht beziehungsweise nicht komplett abgebildet.

Einbezogen werden ausschließlich Beschäftigte, die für den jeweiligen Berichtsmonat vollständig entlohnt werden. Auch sich in Teilzeit befindende oder geringfügig Beschäftigte, die im Rahmen ihrer üblichen Arbeitszeit für den jeweiligen Berichtsmonat bezahlt werden, werden berücksichtigt.

Nicht einbezogen sind Beschäftigte, die im Laufe des jeweiligen Berichtsmonats eingestellt oder entlassen und nicht für den gesamten Berichtsmonat bezahlt werden. Beschäftigte, deren Lohnfortzahlung im Krankheitsfall durch den Arbeitgeber im Berichtsmonat ausläuft oder die im Berichtsmonat unbezahlten Urlaub nehmen, werden nicht in die Erhebung einbezogen.

Zu den Beschäftigten zählen:

- o Sozialversicherungspflichtige Beschäftigte (einschließlich Auszubildende sowie Beschäftigte in Teilzeit oder Altersteilzeit)
- o Leitende Angestellte (auch Geschäftsführerinnen und Geschäftsführer einer GmbH und Vorstände einer AG) mit einem Arbeitsvertrag, die zumindest teilweise feste, d. h. gewinnunabhängige Verdienstbestandteile für die geleistete Arbeit erhalten
- o Beamtinnen und Beamte
- o Geringfügig und kurzfristig Beschäftigte
- o Saison- oder Gelegenheitsarbeiterinnen und -arbeiter, auch wenn sie nicht in der deutschen Sozialversicherung gemeldet sind
- o Aushilfskräfte, Praktikantinnen und Praktikanten, die als abhängig Beschäftigte eine bezahlte Leistung für den Betrieb erbringen

Zu den Beschäftigten zählen nicht:

- o Tätige Inhaberinnen und Inhaber, Mitinhaberinnen und -inhaber und Familienangehörige, sofern ohne Arbeitsvertrag
- o Ausschließlich auf Provisions- oder Honorarbasis bezahlte Personen
- o Personen im Vorruhestand
- o Betreute Personen in Einrichtungen der Jugendhilfe oder in Werkstätten für Behinderte
- o Personen in berufsfördernden Maßnahmen zur Rehabilitation
- o Personen im Bundesfreiwilligendienst
- o Personen im freiwilligen sozialen oder ökologischen Jahr
- o Personen in Arbeitsgelegenheiten mit Mehraufwandsentschädigung (sogenannte Ein-Euro-Jobs)
- o Personen, die keinen Verdienst für ihre Leistungen erhalten (ehrenamtlich Tätige, Volontärinnen und Volontäre u. Ä.)
- o Personen in Elternzeit und Mutterschutz
- o Langzeitkranke

Leih- oder Zeitarbeitnehmerinnen und -arbeitnehmer werden bei den Verleihern bzw. den Zeitarbeitsfirmen nachgewiesen und nicht dort, wo sie ihre Arbeitsleistung erbringen.

Als Berichtsweg stehen für die Betriebe der WZ-Abschnitte A bis N und P bis S die beiden Online-Meldeverfahren eSTATISTIK.core und ein Online-Formular (IDEV) zur Verfügung. Bei eSTATISTIK.core werden Statistikdaten automatisiert aus dem Rechnungswesen gewonnen und online an eine zentrale Annahmestelle übermittelt. Gemäß Bundesstatistikgesetz sind die Meldungen online zu übermitteln.

Die Möglichkeit, die erfragten Daten auf Papier zu übermitteln, besteht nur auf Antrag und in Ausnahmefällen (vgl. § 11a Bundesstatistikgesetz (BStatG)).

1.6 Berichtszeitraum/Berichtszeitpunkt

Berichtszeitraum ist der jeweilige Kalendermonat im Berichtsjahr.

1.7 Periodizität

Es handelt sich um eine monatliche Erhebung. Die Veröffentlichung der absoluten Werte findet allerdings immer für das gesamte Kalenderjahr statt, sodass die Mikrodaten im FDZ eine jährliche Periodizität aufweisen.

1.8 Regionale Ebene

Die Stichprobenziehung erfolgt auf Ebene der Bundesländer. In den Datensätzen wird zwar der amtliche Gemeindeschlüssel des Betriebssitzes ausgewiesen, jedoch sind Ergebnisse unterhalb der Ebene der Bundesländer nicht per se repräsentativ. Bei Auswertungen unterhalb der Länderebene muss die Zuverlässigkeit deshalb im Einzelfall geprüft und beurteilt werden.

Da die regionale Ebene für die Stichprobenziehung das Bundesland ist und in kleineren Ländern anteilmäßig mehr Betriebe befragt werden als in größeren, sind im allgemeinen statistisch valide Analysen auf Ebene der Bundesländer möglich.

2. Methodik

2.1 Erhebungsmethoden

Die Erhebung von Daten zur Verdiensterhebung (VE) wird nahezu vollständig in elektronischer Form durchgeführt. Hierfür stehen für die Betriebe die in Abschnitt 1.5 genannten Übermittlungswege zur Verfügung.

2.2 Erhebungsinhalte

Daten über die Verdienstsituation (bspw. Bruttomonatsverdienst, Entgeltumwandlung, bezahlte Überstunden) aller Beschäftigtengruppen differenziert nach persönlichen und betrieblichen Merkmalen.

2.3 Auswahlgrundlagen

Für die Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten der WZ-Abschnitte A bis N und P bis S ist die Erhebung als geschichtete Stichprobe konzipiert. Es werden alle jene Betriebe betrachtet, die in der Bundeskopie des statistischen Unternehmensregisters aus dem vorhergehenden Jahr des Berichtsjahres der VE mit Stand August verzeichnet sind, wirtschaftlich aktiv waren und mindestens eine oder einen sozialversicherungspflichtigen Beschäftigten aufweisen. Die ausgewählten Betriebe berichten zu allen Beschäftigten, die zur Grundgesamtheit gehören (siehe Abschnitt 1.5 „Berichtskreis/Berichtsweg“).

2.4 Methoden der Stichprobenziehung in der VE

Bei der Stichprobenziehung der Betriebe der WZ-Abschnitte A bis N und P bis S wird eine geschichtete Zufallsstichprobe durchgeführt. Die Schichtung erfolgt nach den 16 Bundesländern, 85 Wirtschaftsabteilungen (2-Steller der WZ 2008) und sieben Beschäftigtengrößenklassen.

Der Gesamtumfang der berücksichtigten Stichprobe beträgt 58 000 Betriebe mit Sitz in Deutschland. Bei dem Merkmal (Richtmerkmal), welches die Aufteilung der Stichprobe auf die Schichten (Allokation) bestimmt, handelt es sich um einen Verhältnisschätzwert aus Summen (Summe von Verdiensten / Summe von Arbeitsstunden). Für solch einen Verhältnisschätzwert gibt es keinen unmittelbaren analytischen Ausdruck für die Varianz; diese lässt sich jedoch durch Linearisierung abschätzen. Die Allokation auf die Schichtgruppen

(Bundesland x Wirtschaftsabteilung (WZ-2-Steller)) erfolgt nach abgestufter Präzision. Die abgestufte Präzision schafft einen Ausgleich zwischen der mathematisch optimalen Allokation für das Gesamtergebnis und der Zielsetzung, in allen Schichtgruppen qualitativ ähnliche Ergebnisse zu erreichen. Die Aufteilung auf die Schichten der Beschäftigtengrößenklassen 1 bis 7 innerhalb der Schichtgruppen erfolgt gemäß der Neyman-Tschuprow-Allokation.¹

Die Schichten der Größenklasse 7 mit 1000 und mehr Beschäftigten sind Totalschichten. Hier werden in der Folge alle Betriebe befragt.

2.5 Aufbereitungsverfahren

Zur Sicherstellung der Datenqualität wird ein mehrstufiges Plausibilisierungsverfahren durchgeführt. Bei der Dialogplausibilisierung prüft ein Programm die Angaben und klassifiziert die Fehler. Als Muss-Fehler klassifizierte Unplausibilitäten müssen korrigiert werden. Diese werden manuell bearbeitet oder mit der kanadischen Software CANCEIS (CANadian Census Edit and Imputation System) durch imputierte Werte nach dem Nearest-Neighbour-Prinzip ersetzt. Kann-Fehler können manuell bearbeitet oder bestätigt werden und werden dann nicht von der o.g. Software angepasst, sofern ein Beschäftigtendatensatz komplett fehlerfrei ist.

Angaben über Beruf, Bildungsstand, Befristung und Beschäftigungsumfang (Vollzeit, Teilzeit) der Beschäftigten wurden nicht direkt erfragt, sondern über den Tätigkeitsschlüssel 2010 erhoben, den die Betriebe für die Meldungen zur gesetzlichen Sozialversicherung vorhalten.

¹ Die Aufteilung gemäß der Neyman-Tschuprow-Allokation erfolgt unter einigen Nebenbedingungen.

2.5.1 Datenaufbereitung

Bei der Erstellung des Beschäftigtenpanels wurden einige Merkmale ausgeschlossen, die ausschließlich Betriebe ohne sozialversicherungspflichtig Beschäftigte, aber mit geringfügig Beschäftigten betreffen, da diese im Beschäftigtenpanel nicht betrachtet werden. Die Variable zu Bundesland lt. Stichprobe (F37) wurde im Beschäftigtenpanel nicht übernommen, da sie inhaltlich identisch mit der bereits enthaltenen Variable Bundesland und somit redundant ist. Auch die Variable Referenzdatensatz aus dem Vormonat (F31) aus der Verdiensterhebung wurde nicht ins Beschäftigtenpanel übernommen. Im Gegensatz zur Verdiensterhebung, bei der in der Regel nur der Monatsmonat April veröffentlicht wird und diese Variable notwendig ist, um Längsschnittinformationen herzustellen, enthält das Beschäftigtenpanel Daten über mehrere Monate hinweg. Dadurch lässt sich unmittelbar erkennen, ob für einen Beschäftigten eine Meldung im Vormonat vorliegt. Das Beschäftigtenpanel enthält keine monatlichen Hochrechnungsfaktoren für den Querschnitt der Verdiensterhebung. Stattdessen stehen eigene Längsschnitthochrechnungsfaktoren zur Verfügung (siehe Kapitel 2.6).

Um Beschäftigte über mehrere Monate hinweg zu identifizieren wurde eine Beschäftigten-ID in Form der Variable (EF3) ergänzt. Es wurde eine neue fortlaufende Datensatzbeschreibung für das Beschäftigtenpanel erstellt (siehe 2.1.1 Datensatzbeschreibung im Metadatenreport Teil II). Diese unterscheidet sich sowohl hinsichtlich der Reihenfolge als auch der Auswahl der Variablen von der Datensatzbeschreibung der Verdiensterhebung.

2.5.2 Methodik der Verknüpfung

Grundlage der Verknüpfung der monatlichen Beschäftigtenmeldungen ist die Kombination der Berichtseinheit-ID, der pseudonymisierten Personalnummer

und des Geburtsdatums. Diese wird im Statistischen Bundesamt der sogenannten Beschäftigten-ID (EF3) zugeordnet. Dabei handelt es sich um einen zufällig generierten, eindeutigen Identifikator, der keine Rückschlüsse mehr auf die ursprüngliche pseudonymisierte Personalnummer zulässt. Gemäß § 6 Abs. 5 VerdStatG ist die pseudonymisierte Personalnummer spätestens drei Jahre nach der letzten Erhebung zu löschen. Liegt für eine Beschäftigte oder einen Beschäftigten innerhalb dieses Zeitraums keine neue Meldung vor, wird die zugehörige intern gespeicherte pseudonymisierte Personalnummer sowie deren Verknüpfung zur Beschäftigten-ID gelöscht. Eine Person, die nach mehr als drei Jahren erneut im Datensatz auftaucht, erhält einen neuen Identifikator. Eine Verbindung zum früheren Datensatz ist dann weder intern noch extern möglich.

2.5.3 Imputation

Es ist zwischen echtem und unechtem Antwortausfall zu unterscheiden. Um echte Antwortausfälle handelt es sich, wenn

- Betriebe zur Auswahlgrundlage gehören, als Stichprobeneinheiten ausgewählt und damit auskunftspflichtig sind und die Auskunft (ggf. auch nach Durchführung von Zwangsmaßnahmen) verweigern.
- Betriebe aus faktischen Gründen (z. B. Konkurs) die gewünschten Daten nicht liefern können.

Unechte Antwortausfälle liegen vor, wenn

- sich nach Aussage des auskunftsgebenden Betriebs die registrierte WZ-Zuordnung nicht mit der tatsächlichen wirtschaftlichen Haupttätigkeit des Betriebs deckt und diese außerhalb des Erfassungsbereichs der Erhebung liegt.
- der Betrieb im Unternehmensregister als Dublette existiert und in der Erhebung doppelt befragt wurde.

- der Betrieb erloschen ist und im Berichtsjahr nicht mehr aktiv war.
- die Erhebungsunterlagen nicht zustellbar waren und die Existenz des Betriebs unklar ist.

Eine eindeutige Unterscheidung zwischen echtem und unechtem Antwortausfall ist nicht in allen Fällen möglich. Daher wird auf Betriebsebene imputiert, wenn ein meldepflichtiger Betrieb im Berichtsmonat nicht gemeldet hat, im Vormonat jedoch vollständige Angaben vorliegen. Grundlage der Imputation ist die Fortschreibung der letzten gültigen Beobachtung. Es gibt zwei Einschränkungen: Bei Imputation des Berichtsmonats Januar anhand der Dezembermeldung des Vorjahres wird sichergestellt, dass die Schichtmerkmale (EF29-EF31) aus dem Berichtsjahr übernommen werden, wenn es in diesen Abweichungen zur Meldung im Dezember gibt. Außerdem werden im Januar des Berichtsjahres nur Betriebe imputiert, die im Verlauf des Berichtsjahres mindestens eine Meldung abgegeben haben.

Zur Reduktion selektiven Antwortausfalls im Beschäftigtenpanel wird eine kombinierte Methode angewendet, bestehend aus:

- LOMF (Last Observation Moving Forward): Anwendung bei kurzfristigem Ausfall mit vollständiger Vormonatsmeldung. Dabei wird der zuletzt beobachtete Wert für maximal eine Periode fortgeschrieben.
- Begrenzung auf einzelne Perioden: Imputiert wird nur, wenn der Ausfall isoliert auf einen einzelnen Berichtsmonat beschränkt ist. Mehrere Imputationen innerhalb eines Jahres sind möglich, solange es sich nicht um aufeinanderfolgende Ausfälle handelt. Damit wird kurzfristige Non-response von dauerhafter Panelmortalität abgegrenzt.
- Reweighting-Verfahren: Zusätzlich werden Gewichtungsfaktoren verwendet, um Verzerrungen auf aggregierter Ebene zu korrigieren. Dies

ermöglicht eine Anpassung der realisierten Stichprobe an die Zielpopulation (siehe 2.6 Hochrechnung).

Für die Auswertung werden zwei Hochrechnungen erstellt, eine mit allen (inkl. imputierten) Einheiten sowie eine nur mit tatsächlich beobachteten Fällen.

2.6 Hochrechnungen

Längsschnittgewichte sind erforderlich, um Betriebe über mehrere Zeitpunkte hinweg analysieren zu können, ohne systematische Verzerrungen zu erzeugen. Sie gewährleisten, dass die Ergebnisse aus einem Panel repräsentativ für die zugrunde liegende Grundgesamtheit der Betriebe sind.

Für einjährige Panels werden Längsschnittgewichte berechnet, die die Grundgesamtheit der Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten in den Wirtschaftsabschnitten² A–S (ausgenommen O und Teile von P) repräsentieren, die im Januar des Startjahres des Panels aktiv waren.

Für Panels mit einer Laufzeit von zwei oder mehr Jahren werden die Längsschnittgewichte so konstruiert, dass sie die Betriebe mit mindestens einem sozialversicherungspflichtig Beschäftigten in den oben genannten Wirtschaftsabschnitten repräsentieren, die im Januar zu Beginn des Panels aktiv waren und auch zum Ende des Panels aktiv sind.

Berücksichtigt werden daher nur Betriebe, die zu Beginn des Panels im Januar eine Meldung abgegeben haben – und bei Panels mit einer Laufzeit von zwei oder mehr Jahren³ zusätzlich nur jene, die bis zum Ende (d.h. bis Dezember eines zwei- oder mehrjährigen Panels) nicht aus der Stichprobe herausrotiert

² Klassifikation der Wirtschaftszweige 2008.

³ Die Anzahl an Jahren, für die stabile Ergebnisse im Längsschnitt mithilfe von Gewichten abgebildet werden können, befindet sich derzeit in der Prüfung.

und noch aktiv sind. Eine Einbeziehung von Betrieben, die im Januar nicht gemeldet haben, würde die Gewichtungsberechnung erheblich verkomplizieren und ist daher nicht vorgesehen.

Die Januar-Gewichte basieren auf den Designgewichten der Verdiensterhebung (VE) und werden um einen Ergänzungsfaktor angepasst, um Antwortausfälle im Januar (sowie bei mehrjährigen Panels auch Panelrotation) zu kompensieren. Das Produkt aus Designgewicht und Ergänzungsfaktor bildet die Grundlage für ein generalisiertes Kalibrierungsverfahren, das in R mit dem Paket *gencalib* durchgeführt wird. Ziel der Kalibrierung ist es, diese Januar-Gewichte so anzupassen, dass die mit dem Panel hochgerechneten Eckwerte – also die Anzahl der Betriebe, der sozialversicherungspflichtig Beschäftigten und der geringfügig Beschäftigten – die entsprechenden Eckwerte im Verwaltungsdatenspeicher der Statistischen Ämter (möglichst) genau treffen.

Bei der Kalibrierung der geringfügig Beschäftigten wird eine Toleranz von 5 % zugelassen:

Zum einen kommt es im Jahr 2022 zu abweichenden Definitionen zwischen der VE und dem Verwaltungsdatenspeicher. Letzterer erfasst jede Person nur einmal (Personenkonzept) bei der Berechnung der Beschäftigtenanzahl. Nebenjobs, die häufig als geringfügige Beschäftigungen ausgeübt werden, bleiben dort unberücksichtigt. Im Panel kann dieselbe Person dagegen auch mehrfach erhoben werden, wenn sie mehr als nur einer Beschäftigung nachgeht (Jobkonzept). Dadurch ist die geringfügige Beschäftigung im Panel im Vergleich zu den Eckwerten aus dem Verwaltungsdatenspeicher tendenziell überrepräsentiert.

Zum anderen lässt sich aber auch ab dem Jahr 2023 (ab dem die Eckwerte des Verwaltungsdatenspeichers nach dem Jobkonzept gebildet werden können) anhand von Tests eine höhere Stabilisierung der Gewichte beobachten,

wenn weiterhin eine Toleranz von 5 % bei der Kalibrierung an die Anzahl der geringfügig Beschäftigten erlaubt wird. Aus diesem Grund wird die eingeräumte Toleranz von 5 % dauerhaft, d.h. über das Jahr 2022 hinaus, erlaubt.

Die Kalibrierung erfolgt auf Betriebsebene und differenziert nach Bundesland, Wirtschaftsabschnitt und Größenklasse. Anschließend wird allen Beschäftigten eines Betriebs der jeweilige errechnete betriebsspezifische Hochrechnungsfaktor für den jeweiligen Monat zugeordnet.

Für jeden Folgemonat werden die Gewichte angepasst, um Antwortausfälle in diesem Monat zu korrigieren. Dazu werden logistische Regressionsmodelle verwendet, welche die Wahrscheinlichkeiten schätzen, mit denen ein Betrieb im jeweiligen Monat gemeldet hat. Diese Modelle umfassen sowohl die Fortsetzer als auch die Rückkehrer. Fortsetzer meint dabei diejenigen Betriebe, die im Vormonat gemeldet haben; Rückkehrer meint diejenigen Betriebe, die im Vormonat nicht gemeldet, aber in einem der davorliegenden Monate eine Meldung abgegeben haben. Hierzu wurde ein bewusst sparsames Modell gewählt, das lediglich die Merkmale Bundesland, Gesamtzahl der Beschäftigten und Wirtschaftsabschnitt berücksichtigt, da diese Modellspezifikation stabile Ergebnisse liefert und verhindert, dass Gewichte über die Zeit hinweg sehr stark schwanken. Auch die für jeden Folgemonat auf Ebene der Betriebe berechneten Längsschnittgewichte werden anschließend allen Beschäftigten eines jeweiligen Betriebs zugeordnet.

2.7 Methodische Änderungen

-

2.8 Klassifikationen

Klassifikation der Berufe, Ausgabe 2010 – überarbeitete Fassung 2020 (KldB 2010, V. 2020). <https://www.klassifikationsserver.de/klassService/jsp/common/url.jsf?variant=kldb2010v2020>

Klassifikation der Wirtschaftszweige, Ausgabe 2008 (WZ 2008).
<https://www.klassifikationsserver.de/klassService/index.jsp?variant=wz2008>

2.9 Fachliche und räumliche Vergleichbarkeit

Die Erhebung basiert auf Verordnungen der Europäischen Union und wird in allen Mitgliedstaaten durchgeführt. Ein Beschäftigtenpanel existiert jedoch nicht EU-weit. Die tiefste repräsentative regionale Ebene sind in Deutschland die Bundesländer. Ergebnisse auf Ebene der Kreise und Gemeinden sind nicht Teil der Zielsetzung der Statistik.

3. Qualität

Siehe Qualitätsberichte des Statistischen Bundesamtes zu den Erhebungen der Arbeitsverdienste nach § 4 Verdienststatistikgesetz.

<https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Verdienste/einfuehrung.html?nn=206008>

4. Zentrale Veröffentlichungen

Mit den Daten aus dem Beschäftigtenpanel können die veröffentlichten Ergebnisse der Verdiensterhebung nicht berechnet werden. Das Panel unterscheidet

sich in mehrfacher Hinsicht vom amtlichen Veröffentlichungsstand, insbesondere in Bezug auf die abgegrenzte Grundgesamtheit der Beschäftigten, dem Hochrechnungsverfahren sowie dem Konzept der Imputation.

5. Angebote der FDZ

Für das Beschäftigtenpanel stehen die On-Site Zugangswege (kontrollierte Datenfernverarbeitung und Gastwissenschaftlerarbeitsplatz) zur Verfügung.

Weitere Informationen zu den verfügbaren FDZ-Produkten des Beschäftigtenpanels finden Sie auf:

<https://www.forschungsdatenzentrum.de/de/sonstige-wirtschaftsstatistiken/vse>

