

Fortschreibungsverfahren in der Statistik der Grundsicherung für Arbeitsuchende

Vorgehen zur Identifikation und Ausgleich von Datenausfällen



Impressum

Produktlinie/Reihe:	Grundlagen: Methodenbericht
Titel:	Fortschreibungsverfahren in der Statistik der Grundsicherung für Arbeitsuchende
Veröffentlichung:	November 2017
Herausgeberin:	Bundesagentur für Arbeit Statistik/Arbeitsmarktberichterstattung
Rückfragen an:	Dr. Bernd Hofmann Claudia Korn Lena Willert Regensburger Straße 104 90478 Nürnberg
E-Mail:	Zentraler-Statistik-Service@arbeitsagentur.de
Telefon:	0911 179-3632
Fax:	0911 179-1131

Weiterführende statistische Informationen:

Internet:	http://statistik.arbeitsagentur.de
Zitierhinweis:	Statistik der Bundesagentur für Arbeit, Grundlagen: Methodenbericht – Fortschreibungsverfahren in der Statistik der Grundsicherung für Arbeitsuchende, Nürnberg, November 2017

Nutzungsbedingungen:

© Statistik der Bundesagentur für Arbeit

Sie können Informationen speichern, (auch auszugsweise) mit Quellenangabe weitergeben, vervielfältigen und verbreiten. Die Inhalte dürfen nicht verändert oder verfälscht werden. Eigene Berechnungen sind erlaubt, jedoch als solche kenntlich zu machen.

Im Falle einer Zugänglichmachung im Internet soll dies in Form einer Verlinkung auf die Homepage der Statistik der Bundesagentur für Arbeit erfolgen.

Die Nutzung der Inhalte für gewerbliche Zwecke, ausgenommen Presse, Rundfunk und Fernsehen und wissenschaftliche Publikationen, bedarf der Genehmigung durch die Statistik der Bundesagentur für Arbeit.

Inhaltsverzeichnis

Kurzfassung.....	4
1 Einleitung.....	5
2 Fortschreibungsmodell auf Ebene der Jobcenter.....	6
2.1 Verfahrensziele.....	6
2.2 Verfahrensbeschreibung.....	6
2.3 Beispiel – SBG II Typ IIe.....	9
3 Landeshochrechnung bei unplausiblen Daten.....	12

Kurzfassung

Die Statistik der Grundsicherung für Arbeitsuchende nach dem SGB II wird monatlich auf Basis der Verwaltungsdaten der gemeinsamen Einrichtungen sowie der zugelassenen kommunalen Träger erstellt. Um die Qualität der Datenlieferungen zu prüfen und bei vereinzelt Datenausfällen trotzdem Werte auf Landes- und Bundesebene herausgeben zu können, wird ein Fortschreibungsverfahren verwendet sowie eine Gewichtung auf Landesebene vorgenommen.

Bei den Datenlieferungen der Jobcenter an die Statistik der Bundesagentur für Arbeit kann es vereinzelt zu Untererfassungen oder Datenausfällen aufgrund technischer oder organisatorischer Probleme kommen. Um die Berichterstattung von Eckwerten der Grundsicherungsstatistik SGB II (Bestand an Bedarfsgemeinschaften, erwerbsfähigen Leistungsberechtigten, nichterwerbsfähigen Leistungsberechtigten, usw.) sicherzustellen, wird ein Fortschreibungsverfahren angewendet und mit dessen Hilfe ein Fortschreibungswert ermittelt. Dieses Fortschreibungsverfahren ermöglicht eine adäquate Einschätzung der Plausibilität sowie darüber hinaus die Ermittlung eines verlässlichen Ersatzwertes bei Datenausfällen bzw. als unplausibel bewerteten Daten.

Neben dem Fortschreibungsverfahren besteht ein Hochrechnungsverfahren auf Ebene der Bundesländer. Über diese Hochrechnung wird für Bundesländer mit unplausiblen Jobcenter- bzw. Kreis-Werten auf Grundlage der plausiblen Jobcenter- bzw. Kreis-Werte des jeweiligen Bundeslandes eine gewichtete Landeshochrechnung vorgenommen, um neben den Eckwerten auch Strukturmerkmale auf Landes- und Bundesebene ausweisen zu können.

1 Einleitung

Mit der Einführung des SGB II in 2005 wurden Arbeitslosenhilfe und Sozialhilfe zur Grundsicherung für Arbeitsuchende zusammengelegt. Die Statistik der Bundesagentur für Arbeit erstellt und veröffentlicht monatlich die amtliche Statistik der Grundsicherung für Arbeitsuchende nach dem SGB II als Sekundärstatistik aus den operativen Daten, die in den Jobcentern bei der Durchführung der Grundsicherung anfallen.

Die Aufgabenerledigung in den Jobcentern erfolgt in unterschiedlichen Trägerformen: gemeinsame Einrichtungen (gE gem. § 44b SGB II) und zugelassene kommunale Träger (zkT gem. § 6a SGB II). Die Grundsicherungsstatistik SGB II verwendet die Daten aus beiden Jobcenterformen. So erfolgt eine gleichartige, integrierte statistische Berichterstattung über alle leistungsrelevanten Aspekte der Grundsicherung für Arbeitsuchende. Die maßgeblichen Eckwerte in der Berichterstattung der Grundsicherungsstatistik SGB II zeigen den monatlichen Bestand an:

- Bedarfsgemeinschaften (BG),
- Personen in Bedarfsgemeinschaften (PERS)
- Leistungsberechtigte (LB)
- Regelleistungsberechtigten (RLB)
- erwerbsfähigen Leistungsberechtigten (ELB) und
- nichterwerbsfähigen Leistungsberechtigten (NEF).

Im statistischen Verarbeitungsprozess werden die Eckwerte mittels Plausibilitätsüberlegungen hinsichtlich ihrer Vollständigkeit und Qualität geprüft. Für diese Plausibilisierung wird seit Januar 2007 ein Fortschreibungsverfahren angewendet, das auf den Daten der gemeinsamen Einrichtungen und der zugelassenen kommunalen Träger basiert. In der Datenübermittlung kann es bei einzelnen Jobcentern zu Datenausfällen oder Untererfassungen aufgrund technischer bzw. organisatorischer Probleme vor Ort kommen.¹ Auch wenn das Ausmaß der Datenlücken bei den Eckwerten gering scheint, so sind die Konsequenzen doch weitreichend: Die Datenlücken setzen sich ausnahmslos bei allen Grund- und Strukturdaten fort. Wenn für ein Jobcenter zu wenig Personendatensätze geliefert werden und die Zahl der BG bzw. ELB nicht plausibel ausgewiesen werden kann, können auch deren Teilgrößen und weitere abgeleitete Kennzahlen nicht mehr zuverlässig ermittelt werden. In solchen Fällen gewährleistet das im folgenden Kapitel beschriebene Fortschreibungsverfahren, dass auf Ebene der Jobcenter zumindest die Eckwerte sowie auf Landes- und Bundesebene neben den Eckwerten auch Strukturmerkmale und Leistungsdaten vollständig berichtet werden können.

¹ Von Januar 2012 bis Juni 2017 kam es bei den Eckwerten nur bei etwa 0,25 % der Datenlieferungen zu Ausfällen (bezogen auf 66 Berichtsmonate und 407 Jobcenter).

2 Fortschreibungsmodell auf Ebene der Jobcenter

Die Jobcenter übermitteln die Daten, die vor der Erstellung von statistischen Auswertungen, im Rahmen der monatlichen Aufbereitung einer Plausibilitätsprüfung unterzogen werden. Damit können mögliche Fehler bei der Erfassung oder Übermittlung der Daten identifiziert und Verzerrungen in der statistischen Berichterstattung vermieden werden.

Diese Plausibilisierung erfolgt anhand eines Abgleichs der übermittelten Werte mit Erwartungswerten, die auf Basis der Daten des Vormonats ermittelt werden. Das verwendete Fortschreibungsverfahren basiert auf der Annahme, dass sich die Eckwerte in Gebieten mit vergleichbarer Arbeitsmarktstruktur in ähnlicher Weise entwickeln. Auch wenn für ein Jobcenter keine plausiblen Bestandszahlen für Bedarfsgemeinschaften und Personen vorliegen, lassen sich diese Lücken im Vergleich zum Vormonat anhand der Entwicklung in vergleichbaren Regionen abschätzen.

2.1 Verfahrensziele

Die Ermittlung von Fortschreibungswerten für die Eckwerte der Grundsicherungsstatistik SGB II auf Ebene der Jobcenter und Kreise dient zwei Zielen:

- Zum einen ermöglichen die Fortschreibungswerte eine adäquate Einschätzung der Plausibilität der übermittelten Daten.
- Zum anderen gewährleistet die Fortschreibung bei Datenausfällen die Ermittlung verlässlicher Ersatzwerte für die Eckwerte der Grundsicherungsstatistik SGB II auch für kleinere regionale Einheiten, wie Kreise oder Jobcenter. Somit ist eine lückenlose Berichterstattung von konsistenten Zeitreihen möglich.

2.2 Verfahrensbeschreibung

Der Fortschreibungswert für ein Jobcenter ergibt sich durch die Multiplikation des jeweiligen Vormonatswerts mit einem errechneten vergleichstypspezifischen Fortschreibungsfaktor. Der Berechnung der Fortschreibungsfaktoren liegt die vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) entwickelte SGB II-Typisierung zugrunde. Diese ordnet alle SGB II-Träger nach ausgewählten Eigenschaften in Gruppen mit vergleichbarer Arbeitsmarktstruktur ein.² Diese Typisierung bildet die Grundlage für das Fortschreibungsverfahren und die daraus folgenden Berechnungen.

Das IAB hat mit Hilfe einer Regressionsanalyse relevante Einflussfaktoren für die Ziele im SGB II „Verbesserung der Integration in Erwerbstätigkeit“, „Verringerung der Hilfebedürftigkeit“ und „Vermeidung von langfristigem Leistungsbezug“ ermittelt. Diese Einflussfaktoren, wie beispielsweise die Arbeitslosenquote, die Bevölkerungsdichte oder das Verhältnis von SGB II- zu SGB III-Kunden, wurden in einer nachfolgenden Clusteranalyse zur Typisierung der Träger genutzt, indem Träger mit ähnlichen Ausprä-

² Vgl. IAB-Forschungsbericht 11/2013: Neukonzeption der Typisierung im SGB-II-Bereich <http://doku.iab.de/forschungsbericht/2013/fb1113.pdf>

gungen der Einflussfaktoren jeweils zu einer Gruppe, einem SGB II-Typ, zusammengefasst wurden. Innerhalb eines Typs besteht in Bezug auf die Einflussfaktoren eine möglichst hohe Homogenität, während er sich zu anderen Typen deutlich abgrenzt. Im Ergebnis werden die aktuell 407 Jobcenter 15 SGB II-Typen zugeordnet (Stand 2017).

Das Fortschreibungsverfahren basiert auf der Annahme, dass sich innerhalb eines SGB II-Typs aufgrund ähnlicher Einflussfaktoren die Eckwerte ähnlich entwickeln. Für jeden SGB II-Typ wird daher ein Fortschreibungsfaktor für jeden Eckwert ermittelt, der sich aus der durchschnittlichen Veränderung einer Kennzahl aller Jobcenter dieses Typs vom Vormonat zum aktuellen Monat ergibt. Das Fortschreibungsverfahren wird für die folgenden Eckwerte angewendet: Bedarfsgemeinschaften (BG), Regelleistungsbearbeitungsgemeinschaft (RL-BG), Personen in Bedarfsgemeinschaften (PERS), erwerbsfähige Leistungsberechtigte (ELB), nichterwerbsfähige Leistungsberechtigte (NEF) sowie ein Gesamtwert für sonstige Leistungsberechtigte (SLB), vom Leistungsanspruch ausgeschlossene Personen (AUS) und Kinder ohne Leistungsanspruch (KOL).

Vor der Berechnung der Fortschreibungsfaktoren werden diejenigen Jobcenter identifiziert, die eine besonders extreme Veränderung der jeweiligen Eckwerte aufweisen (Ausreißer). Diese werden bei der Ermittlung der Fortschreibungsfaktoren nicht berücksichtigt, um diese nicht zu verzerren.

Für jedes einzelne Jobcenter wird der Vormonatswert eines Eckwertes mit dem Fortschreibungsfaktor, der für den SGB II-Typ dieses Jobcenters ermittelt wurde, multipliziert. Der sich ergebende Fortschreibungswert wird dem aus den übermittelten Daten errechneten Wert gegenübergestellt. Größere Differenzen zwischen dem übermittelten Wert und dem Fortschreibungswert, der die durchschnittliche Entwicklung im Vergleichstyp repräsentiert, können auf ein mögliches Problem bei der Datenqualität, z. B. aufgrund von Fehlern in der Datenübermittlung, hinweisen. In solchen Fällen erfolgt eine weitere Prüfung der Daten und eine Rücksprache mit dem Jobcenter, um zu klären, ob es sich um eine plausible Entwicklung handelt oder tatsächlich fehlerhafte Daten vorliegen.

Das Vorgehen wird im Folgenden detailliert erläutert:

Im ersten Schritt werden über alle Träger hinweg die Ausreißer identifiziert, die für die nachfolgende Ermittlung der Fortschreibungsfaktoren nicht weiter berücksichtigt werden. Zum Ausschluss von Ausreißern werden der bundesweite Mittelwert \bar{x} und die Standardabweichung s der relativen Veränderung $d_{k,t}$ über alle Träger hinweg berechnet. Hierbei werden alle liefernden Jobcenter unabhängig von ihrer Größe bzw. der Größe der Eckwerte gleichwertig behandelt.

Für sämtliche Träger wird die **relative Veränderung** $d_{k,t}$ zwischen dem Wert w_t des Monats t und dem Wert w_{t-1} des Vormonats $t - 1$ (Basiswert) errechnet. Formal dargestellt ist die relative Veränderung $d_{k,t}$ von Träger k im Monatsmonat t :

$$d_{k,t} = \frac{w_{k,t} - w_{k,t-1}}{w_{k,t-1}}$$

Durch Addition bzw. Subtraktion der eineinhalbfachen Standardabweichung zum bzw. vom Mittelwert ergeben sich eine bundesweite Ober- und Untergrenze u für den Ausschluss der Ausreißer:

$$u_{d_{k,t}} = \bar{x} \pm (1,5 * s)$$

Nach Ausschluss der Ausreißer wird der eigentliche Fortschreibungsfaktor $(1 + \bar{d}_{i,t})$ für jeden SGB II-Typ berechnet. Im Ergebnis der Clusterbildung des IAB ist jeder Träger einem regionalen von m SGB II-Typen (aktuell sind es 15 Typen) zugeordnet. Dabei wird für jeden SGB II-Typ i zwischen $t - 1$ und t die **durchschnittliche relative Veränderung** $\bar{d}_{i,t}$ der Jobcenter ohne Ausreißer berechnet:

$$\bar{d}_{i,t} = \frac{\sum_{k=1}^n (w_{ik,t} - w_{ik,t-1})}{\sum_{k=1}^n w_{ik,t-1}}$$

Mit Hilfe des vergleichstypspezifischen Fortschreibungsfaktors $\bar{d}_{i,t}$ können anschließend die **Fortschreibungswerte für die einzelnen Jobcenter im zugehörigen SGB II-Typ i** ermittelt werden. Man erhält den Fortschreibungswert $\hat{w}_{ik,t}$ für das Jobcenter k vom Typ i im Monatsmonat t , indem der Vormonatswert (Basiswert³) jedes Jobcenters mit dem Fortschreibungsfaktor $(1 + \bar{d}_{i,t})$ multipliziert wird:

$$\hat{w}_{ik,t} = w_{ik,t-1} (1 + \bar{d}_{i,t})$$

und analog für alle Jobcenter $1 \dots n$ vom Typ i (in Vektorenschreibweise):

$$\begin{bmatrix} \hat{w}_{i1,t} \\ \hat{w}_{i2,t} \\ \vdots \\ \hat{w}_{in,t} \end{bmatrix} = \begin{bmatrix} w_{i1,t-1} \\ w_{i2,t-1} \\ \vdots \\ w_{in,t-1} \end{bmatrix} \times \begin{bmatrix} 1 + \bar{d}_{i,t} \\ 1 + \bar{d}_{i,t} \\ \vdots \\ 1 + \bar{d}_{i,t} \end{bmatrix}$$

Dieses Verfahren wird für alle SGB II-Typen 1 bis m durchgeführt.

Die Einschätzung der Plausibilität erfolgt durch den Vergleich von gelieferten Eckwerten mit den Fortschreibungswerten. Hierfür wird für alle k Jobcenter der aktuelle Monatswert w mit dem errechneten Fortschreibungswert \hat{w} verglichen und die relative Abweichung $b_{k,t}$ ermittelt:

$$b_{k,t} = \frac{w_{k,t}}{\hat{w}_{k,t}}$$

³ Dabei ist es möglich, dass es sich auch bei dem Basiswert um einen Fortschreibungswert handelt, z. B. bei zwei oder mehr Datenausfällen bzw. nicht plausiblen Meldungen in Folge. Es wird in diesen Fällen also mit einem Basiswert gerechnet, der im Vormonat durch Anwendung des Fortschreibungsverfahrens ermittelt wurde.

Anschließend werden wiederum über alle Vergleichstypen hinweg eine bundesweite Ober- und eine Untergrenze berechnet. Die Obergrenze ergibt sich aus der Addition des 2-fachen Interquartilsabstands zum 75 %-Quantil, die Untergrenze analog durch Subtraktion des 2-fachen Interquartilsabstands vom 25 %-Quantil. Liegt die relative Abweichung außerhalb dieser Grenzen, wird das Jobcenter als auffällig angezeigt und die Daten sind vertieft zu prüfen.

Die Einschätzung zur Plausibilität der Eckwerte bildet die Grundlage für die weitere Plausibilisierung der Strukturinformationen, wie Bewegungen oder monetäre Größen. Für alle Wartezeitstände gilt, dass diese Strukturinformationen als unplausibel gewertet werden, wenn bereits die Eckwerte als unplausibel eingestuft wurden. Werden die Eckwerte als plausibel eingestuft, findet eine gesonderte Plausibilisierung nach analogen Verfahren für die Strukturinformationen Bewegungen, Einkommen, Bestand und Bewegung von Sanktionen sowie Leistungs- und Zahlungsansprüche statt.⁴ Bei den Strukturinformationen dient der Fortschreibungswert lediglich der Plausibilisierung und wird bei unplausibler Einschätzung nicht veröffentlicht.

2.3 Beispiel – SGB II Typ IIe

Das beschriebene Fortschreibungsverfahren findet monatlich Anwendung und schließt dabei alle SGB II-Typen mit ein. Am Beispiel des Typs IIe wird das Verfahren für den Eckwert der erwerbsfähigen Leistungsberechtigten (ELB) noch einmal für einen beliebigen Monatsmonat exemplarisch dargestellt (siehe Tabelle 1). Der Vergleichstyp IIe beinhaltet Städte und verdichtete Landkreise mit durchschnittlicher ELB-Quote, sehr hohen Wohnkosten, sehr hohem Migrationsanteil und einem Arbeitsmarkt der durch Großbetriebe und einen ausgeprägten Niedriglohnbereich gekennzeichnet ist. Dem SGB II-Typ sind zwölf Jobcenter zugeordnet, davon sechs gemeinsame Einrichtungen und sechs zugelassene kommunale Träger.

Für jedes einzelne Jobcenter wird zunächst die relative Veränderung der erwerbsfähigen Leistungsberechtigten zum Vormonat errechnet. Um Ausreißer zu ermitteln, werden aus diesen Ergebnissen der bundesweite Durchschnitt (-0,57 %) und die bundesweite Standardabweichung (0,82 %) der relativen Veränderung berechnet. Durch Addition bzw. Subtraktion der eineinhalbfachen Standardabweichung zum bzw. vom Mittelwert werden die Obergrenze von 0,67 % sowie die Untergrenze von -1,81 % für den Ausschluss der Ausreißer festgelegt. Im Folgenden werden deshalb die Jobcenter A, H und L aufgrund der Überschreitung der Obergrenze sowie das unterhalb abweichende Jobcenter J nicht in die weiteren Berechnungen zum Fortschreibungsfaktor für den Typ IIe einfließen.

⁴ Eine Ausnahme bilden die Informationen zu Bildung und Teilhabe. Die Besonderheiten der Leistungsgewährung für diese Leistungen führen dazu, dass nicht von einer gleichförmigen und innerhalb von SGB II-Typen gleichartigen Entwicklung ausgegangen werden kann. Für nähere Informationen siehe: <https://statistik.arbeitsagentur.de/Statischer-Content/Grundlagen/Methodenberichte/Grundsicherung-Arbeitsuchende-SGBII/Generische-Publikationen/Methodenbericht-Bildung-und-Teilhabe.pdf>

Aus den Daten der restlichen Jobcenter des Typs IIe wird nun der Fortschreibungsfaktor $(1+\bar{d}_{i,t})$ für die erwerbsfähigen Leistungsberechtigten ermittelt:

$$\text{Fortschreibungsfaktor}_{IIe,Monat} = 1 + \left(\frac{137.134 - 137.200}{137.200} \right) = 0,999518950$$

Anschließend kann für jedes einzelne Jobcenter der Fortschreibungswert $\hat{w}_{ik,t}$ berechnet werden. Dafür wird der Basiswert des Vormonats mit dem Fortschreibungsfaktor des Typs multipliziert:

$$\text{Fortschreibungswert}_{JC A,Monat} = 19.885 \times (0,999518950) = 19.875,434329$$

$$\text{Fortschreibungswert}_{JC B,Monat} = 9.088 \times (0,999518950) = 9.083,628222$$

$$\text{Fortschreibungswert}_{JC C,Monat} = 52.313 \times (0,999518950) = 52.260,847843$$

...

usw.

Anschließend wird für die jeweiligen Jobcenter k der aktuelle Monatswert w mit dem errechneten Fortschreibungswert \hat{w} verglichen und die relative Abweichung $b_{k,t}$ ermittelt. Bei diesem Vergleich kommen wiederum eine bundesweite Ober- und eine Untergrenze zum Tragen: Die Obergrenze liegt in dem Beispiel bei +1,81 % und die Untergrenze bei -1,99 %. Liegt die relative Abweichung außerhalb dieser Grenzen, wird das Jobcenter als auffällig angezeigt und die gelieferten Daten sind vertieft zu prüfen. Bei der Einschätzung, dass es sich um unplausible Daten handelt, wird der auffällige Wert durch den Fortschreibungswert ersetzt.

Im vorliegenden Beispiel befinden sich das Ergebnis -2,27 % des Jobcenters J außerhalb des erwarteten Rahmens der relativen Abweichung vom Fortschreibungswert. Die weitere Prüfung zeigt ebenfalls, dass es sich um unplausible Daten handelt und deswegen wird anstatt des gemeldeten Wertes der Fortschreibungswert veröffentlicht (siehe Tabelle 1). Die Werte der anderen Jobcenter des Typs IIe befinden sich innerhalb dieser Grenzen und können für die weitere Verarbeitung übernommen werden. Für den ausgefallenen Wert des Jobcenters J dient nun der ermittelte Fortschreibungswert als Ersatzwert in der Berichterstattung.

Tabelle 1**Fortschreibungsverfahren am Beispiel des SGB II-Typs IIe, Bestand an ELB (fiktive Daten)**

	SGB II-Typ	Vormonat (VM)	Aktueller Monat (AM)	Prozentuale Veränderung (AM-VM/VM)	Fortschreibungswert (FW)	Prozentuale Abweichung (AM-FW/FW)	Veröffentlichter Wert
Jobcenter A	IIe	(19.885)	(20.058)	(0,87%)	19.875	0,92%	20.058
Jobcenter B	IIe	9.088	9.095	0,08%	9.084	0,13%	9.095
Jobcenter C	IIe	52.286	52.313	0,05%	52.261	0,10%	52.313
Jobcenter D	IIe	14.397	14.320	-0,53%	14.390	-0,49%	14.320
Jobcenter E	IIe	6.577	6.612	0,53%	6.574	0,58%	6.612
Jobcenter F	IIe	11.645	11.683	0,33%	11.639	0,37%	11.683
Jobcenter G	IIe	7.062	7.070	0,11%	7.059	0,16%	7.070
Jobcenter H	IIe	(11.071)	(11.151)	(0,72%)	11.066	0,77%	11.151
Jobcenter I	IIe	14.653	14.635	-0,12%	14.646	-0,07%	14.635
Jobcenter J	IIe	(15.253)	(14.899)	(-2,32%)	15.246	-2,27%	15.246
Jobcenter K	IIe	21.492	21.406	-0,40%	21.482	-0,35%	21.406
Jobcenter L	IIe	(5.246)	(5.294)	(0,91%)	5.243	0,96%	5.294

() Werte in Klammern bezeichnen die Ausreißer

3 Landeshochrechnung bei unplausiblen Daten

Neben den in Kapitel 2 beschriebenen Eckwerten wird in der Grundsicherungsstatistik SGB II auch eine große Zahl an Strukturmerkmalen und Leistungsinformationen differenziert berichtet. So wird beispielsweise über Bedarfsgemeinschaften und Leistungsberechtigte nach soziodemographischen Merkmalen (u. a. Geschlecht, Alter, Staatsangehörigkeit) sowie alle Aspekte der Leistungsgewährung (Bedarfe, Einkommen, Leistungen etc.) berichtet. Diese Strukturmerkmale und Leistungsinformationen werden nur für Jobcenter ausgewiesen, wenn deren Daten als plausibel eingeschätzt wurden.

Eine Schätzung von Strukturmerkmalen und Leistungsinformationen auf Ebene der einzelnen Jobcenter, vergleichbar wie sie bei den Eckwerten mittels Fortschreibungsverfahren vorgenommen wird, ist aufgrund der Fülle an Merkmalen und sich daraus ergebenden Merkmalskombinationen in der Statistik nicht möglich. Es besteht ein hohes Interesse daran, diese Strukturmerkmale und Leistungsinformationen zumindest auf Landes- und Bundesebene ausweisen zu können. Aus diesem Grund wird ein Verfahren zur Landeshochrechnung angewendet, das auf den Fortschreibungswerten der Eckwerte der Jobcenter basiert.

Aufgrund der hierarchischen Gebietsstruktur wird das Hochrechnungsverfahren ausgehend von den Daten der einzelnen Jobcenter bzw. Kreise für jedes Bundeslandes vorgenommen. Es ist davon auszugehen, dass sich die hochzurechnenden Strukturmerkmale und Leistungsinformationen der Jobcenter in einem Bundesland nur in so geringem Maße unterscheiden, dass sich die auf Grundlage der plausiblen Jobcenter hochgerechneten Landeswerte der jeweiligen Merkmale nicht signifikant von den tatsächlichen Werten im jeweiligen Bundesland unterscheiden.

Die Plausibilitätseinstufung anhand des Fortschreibungsverfahrens auf Ebene der SGB II-Typen bildet die Basis für die Berechnung der Gewichtungsfaktoren. Für die Hochrechnung werden für jedes einzelne Bundesland j Gewichtungsfaktoren $g(j)$ für die Eckwerte ermittelt. Diese Berechnung wird für jedes zu plausibilisierende Thema in der Grundsicherungsstatistik SGB II anhand der jeweils vorliegenden Plausibilitätseinstufungen durchgeführt.

Dabei wird der Wert aller Jobcenter (unplausible mit Fortschreibungswert und plausible) \hat{w} durch den Wert der ausschließlich plausiblen Jobcenter \check{w} geteilt:

$$g(j) = \frac{\sum_{k=1}^n \hat{w}_{jk,t}}{\sum_{k=1}^n \check{w}_{jk,t}}$$

Zur Ermittlung der Landeswerte werden nun die Merkmalsträger (BG bzw. Personen) der plausiblen Jobcenter mittels des auf Grundlage der Eckwerte errechneten Hochrechnungsfaktors hochgewichtet. Durch Addition der gewichteten Werte der plausiblen Jobcenter ergeben sich die jeweiligen hochgerechneten Strukturmerkmale und Leistungsinformationen des Bundeslandes.

Die Berechnung von Gewichtungsfaktoren wird am Beispiel eines fiktiven Bundeslandes für den BG-Typ Alleinerziehende-BG dargestellt. Dem Bundesland sind sechs Jobcenter (bzw. Kreise) zugeordnet. Der Datenausfall eines Jobcenters hätte zur Folge, dass kein Landeswert zur Zahl der Alleinerziehenden-BG

veröffentlicht werden könnte. Um dies zu vermeiden wird ein Gewichtungsfaktor auf Grundlage des relevanten Eckwertes (BG Bestand) für das Bundesland ermittelt, der diesen Datenausfall ausgleicht.

Tabelle 2

Beispiel-Bundesland mit unplausiblen Jobcenter C (fiktive Daten)⁵

	Plausible Werte (BG Bestand)	Bestand BG mit Fort- schreibungswert
Jobcenter A	24.156	24.156
Jobcenter B	6.957	6.957
Jobcenter C	-	2.232
Jobcenter D	5.148	5.148
Jobcenter E	3.038	3.038
Jobcenter F	7.328	7.328
GESAMT Bundesland	46.627	48.859

- unplausibler Wert

Aus den BG-Eckwerten der Jobcenter des Bundeslandes bzw. deren Fortschreibungswerten wird zunächst der Gewichtungsfaktor g ermittelt:

$$\text{Gewichtungsfaktor}(\text{Bundesland}) = \frac{48.859}{46.627} = 1,047869$$

Um zu verhindern, dass für ein Bundesland aufgrund unplausibler oder fehlender Werte einzelner Jobcenter keine Daten veröffentlicht werden, wird zum Ausgleich auf Grundlage der anderen Jobcenter-Werte mittels des Gewichtungsfaktors auf den Landeswert hochgerechnet. Für das Bundesland werden in dem fiktiven Beispiel also alle Bedarfsgemeinschaften des Jobcenter C mit 0 gewichtet, während die Bedarfsgemeinschaften der anderen Jobcenter mit dem errechneten Gewichtungsfaktor 1,047869 multipliziert (gewichtet) werden.

Bei einer Auswertung der Bedarfsgemeinschaften des BG-Typs Alleinerziehend für das Beispiel-Bundesland wird, aufgrund des in diesem Beispiel unplausiblen Wertes des Jobcenters C, ein gewichteter Wert ausgegeben. Die veröffentlichte Anzahl an alleinerziehenden Bedarfsgemeinschaften für das Bundesland wäre demnach nicht 7.371 (Landessumme ohne Jobcenter C) sondern 7.724 (hochgerechnet), siehe Tabelle 3.

⁵ Die Ermittlung der Gewichtungsfaktoren wird analog zum genannten Beispiel für Kreise eines Bundeslandes durchgeführt.

Tabelle 3

**Beispiel-Bundesland mit unplausiblem Jobcenter C (fiktive Daten):
Bedarfsgemeinschaften des BG-Typs Alleinerziehend**

	Plausible Werte (alleinerz. BG)
Jobcenter A	3.715
Jobcenter B	1.126
Jobcenter C	-
Jobcenter D	775
Jobcenter E	516
Jobcenter F	1.239
GESAMT Bundesland (ohne Jobcenter C)	7.371
GESAMT Bundesland (hochgerech- ter, veröffentlichter Wert)	7.724

- unplausibler Wert

Ergebnisse auf Bundesebene sowie Ergebnisse für Ost- und Westdeutschland ergeben sich aus der Summe der hochgerechneten Landesergebnisse, sofern es im Betrachtungszeitraum auf Ebene der Jobcenter zu Datenausfällen gekommen ist oder Werte als unplausibel eingestuft wurden. Mit Hilfe dieser hochgerechneten Ergebnisse sind dennoch Aussagen zu Strukturmerkmalen von Personen und Bedarfsgemeinschaften sowie zu Leistungsinformationen auf Landes- und Bundesebene möglich.

Statistik-Infoseite

Im Internet stehen statistische Informationen unterteilt nach folgenden Themenbereichen zur Verfügung:

- [Arbeitsmarkt im Überblick](#)
- [Arbeitslose, Unterbeschäftigung und Arbeitsstellen](#)
- [Arbeitsmarktpolitische Maßnahmen](#)
- [Ausbildungsstellenmarkt](#)
- [Beschäftigung](#)
- [Grundsicherung für Arbeitsuchende \(SGB II\)](#)
- [Leistungen SGB III](#)
- [Statistik nach Berufen](#)
- [Statistik nach Wirtschaftszweigen](#)
- [Zeitreihen](#)
- [Eingliederungsbilanzen](#)
- [Amtliche Nachrichten der BA](#)
- [Kreisdaten](#)

Glossare sind zu folgenden Fachstatistiken veröffentlicht:

- [Arbeitsmarkt](#)
- [Ausbildungsstellenmarkt](#)
- [Beschäftigung](#)
- [Förderstatistik/Eingliederungsbilanzen](#)
- [Grundsicherung für Arbeitsuchende \(SGB II\)](#)
- [Leistungen SGB III](#)

Die [Methodischen Hinweise der Statistik](#) bieten ergänzende Informationen.