

Kurzinformation

Fortschreibungsmodell der Daten zur Grundsicherung für Arbeitsuchende

Stand: 20.08.2008

Herausgeber:	Ansprechpartner:
Bundesagentur für Arbeit	
Statistik	Sylvie Breuer, Tel.: 0911 / 179 - 1478
Regensburger Straße 104	Joachim Fritz, Tel.: 0911 / 179 - 1905
90478 Nürnberg	

© Statistik der Bundesagentur für Arbeit, Nürnberg 2008.

Nachdruck, auch auszugsweise, nur mit Quellenangabe gestattet. Die Verbreitung, auch auszugsweise, über elektronische Systeme/Datenträger bedarf der vorherigen Zustimmung. Alle übrigen Rechte vorbehalten.

Zitierhinweis: Kurzinformation: Fortschreibung der Daten zur Grundsicherung für Arbeitsuchende, Nürnberg 20.08.2008

1. Einleitung

Seit der Einführung des SGB II und der Berichterstattung über Eckwerte zu Bestandsfällen von Bedarfsgemeinschaften (BG), erwerbsfähigen Hilfebedürftigen (eHb) und nicht erwerbsfähigen Hilfebedürftigen (nEf) im Rahmen der Grundsicherung für Arbeitsuchende wurde zur Einschätzung der Plausibilität der Daten und zum Ausgleich von Datenausfällen ein Regressionsmodell zur Ermittlung von Schätzwerten dieser Bestandsfälle genutzt. Dieses Modell basierte auf den Arbeitslosenhilfe- und Sozialhilfedaten von Dezember 2004 und den im BA-Verfahren A2LL vollständig erfassten Daten aus Kreisen mit der Aufgabenwahrnehmung in Form von Arbeitsgemeinschaften.

Im Laufe der Jahre 2005 und 2006 hat sich die Datenqualität im Bereich zur Grundsicherung für Arbeitsuchende erheblich verbessert, so dass neben dem Regressionsmodell auch die Zeitreihen der einzelnen Kreise eine gute Einschätzung der Bestandsdaten zulassen. Dennoch kann es speziell bei den zugelassenen kommunalen Trägern, die den Übermittlungsstandard XSozial-BA-SGB II nutzen, zu Datenausfällen aufgrund technischer und / oder organisatorischer Probleme vor Ort kommen. In diesen Fällen muss bei der statistischen Berichterstattung für diese Kreise bzw. Träger entweder auf den Vormonatswert oder einen Schätzwert zurückgegriffen werden, um trotz der Ausfälle Aussagen über die Ergebnisse auf höher aggregierten Ebenen (Bund, Ost / West / Länder) treffen zu können. Von der Realität vor Ort abweichende Darstellungen für die einzelnen Berichtszeiträume und Verzerrungen des tatsächlichen Verlaufs der Zeitreihe für die betroffenen regionalen Einheiten (zkT) ist in diesen Fällen häufig nicht vermeidbar.

Mittlerweile bietet sich durch die mehr und mehr auch im Zeitverlauf als vollständig zu bewertenden Datenlieferungen über XSozial-BA-SGB II die Möglichkeit einer Optimierung der bei Datenausfällen ersatzweise verwendeten Schätzwerte. Daher wurde das bisherige Regressionsverfahren durch ein Fortschreibungsverfahren abgelöst.

2. Fortschreibungsmodell

2.1 Ziele

Ziele des neuen Fortschreibungsverfahrens waren eine adäquatere Einschätzung der Plausibilität der Ergebnisse, bessere Ersatzwerte bei Datenausfällen auch für kleinere regionale Einheiten (Kreise, Träger) und somit lückenfreie und konsistente Zeitreihen.

Es ist sinnvoll, auch die Einschätzung der Plausibilität der ausgewerteten Ergebnisse auf Grundlage der Fortschreibung durchzuführen. Anstatt der Schätzwerte werden die errechneten Fortschreibungswerte mit den gemeldeten Werten verglichen. Liegt ein gemeldeter Wert außerhalb eines Signifikanzintervalls, wird er als unplausibel eingestuft und durch den Fortschreibungswert ersetzt.

2.2 Verfahren

Grundlage für die folgenden Berechnungen bildet die vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) entwickelte SGB II-Typisierung, eine regionale Einteilung aller SGB II-Träger. Das IAB hat mit Hilfe einer Regressionsanalyse relevante Einflussfaktoren auf die Integrationsquote von erwerbsfähigen Hilfebedürftigen (eHb) ermittelt. Die Integrationsquote gibt das Verhältnis zwischen den eHb, die in eine reguläre Beschäftigung integriert wurden, zu allen eHb wieder. Die auf diese Weise ermittelten Einflussfaktoren (u.a. Arbeitslosenquote, Bevölkerungsdichte, Verhältnis von SGB II-Kunden zu SGB III-Kunden) wurden in einer nachfolgenden Clusteranalyse zur Typisierung der Träger genutzt, indem Träger mit ähnlichen Ausprägungen der Einflussfaktoren jeweils zu einer Gruppe (einem SGB II-Typ) zusammengefasst wurden.

Die Berechnung des Fortschreibungsfaktors erfolgt ausschließlich auf Basis der Daten der Arbeitsgemeinschaften und der Arbeitsagenturen mit getrennter Aufgabewahrnehmung (AAgtAw) aufgegliedert nach den vom IAB ermittelten SGB II-Typen. Aufgrund von vereinzelt immer wieder auftretenden Inkonsistenzen in den Zeitreihen der zugelassenen kommunalen Träger (zkT) werden diese Daten bei der Berech-

nung des Fortschreibungsfaktors nicht berücksichtigt. Auf diese Weise soll verhindert werden, dass die tatsächliche Entwicklung innerhalb eines SGB II-Trägertyps verzerrt wird.

Für die Bestimmung des Fortschreibungsfaktors ist die durchschnittliche Entwicklung zwischen dem Bestand des Vormonats und des aktuellen Monats für jeden SGB II-Typ ausschlaggebend. Durch Multiplikation des Bestandswerts aus dem Vormonat (Basiswert) mit dem Fortschreibungsfaktor des jeweiligen SGB II-Typs wird dann der Fortschreibungswert berechnet.

Das Vorgehen wird im Folgenden detailliert erläutert:

Im Ergebnis der Clusterbildung ist jeder Träger einem von m regionalen SGB II-Typen zugeordnet. Für alle Träger wird die relative Veränderung d_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 $d_{ik,t}$ von Träger k vom SGB II-Typ i im Monatsmonat t :

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

Um die durchschnittliche relative Veränderung $\bar{d}_{i,t}$ innerhalb eines SGB II-Typs i zwischen $t - 1$ und t zu ermitteln, wird ein gewichtetes Mittel der relativen Veränderungen aller n Träger dieses Typs gebildet. Durch die Gewichtung wird vermieden, dass sich die unterschiedlichen Größen der einzelnen Träger verschieden stark auf die durchschnittliche relative Veränderung auswirken.

$$\bar{d}_{i,t} = \sum_{k=1}^n \frac{w_{ik,t} - w_{ik,t-1}}{w_{ik,t-1}} \cdot \frac{w_{ik,t-1}}{w_{i,t-1}} = \sum_{k=1}^n d_{ik,t} \cdot g_{ik,t-1} \text{ mit } g_{ik,t-1} = \frac{w_{ik,t-1}}{w_{i,t-1}}$$

Mit Hilfe von $\bar{d}_{i,t}$ können anschließend die Fortschreibungswerte für alle Träger im SGB II-Typ i ermittelt werden. Man erhält den Fortschreibungswert $\hat{w}_{ik,t}$ für Träger k

vom Typ i im Berichtsmonat t , indem der Vormonatswert (Basiswert) jedes Trägers mit dem Fortschreibungsfaktor $(1 + \bar{d}_{i,t})$ multipliziert wird, also

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

und analog für alle Träger $1 \dots n$ vom Typ i :

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

Dieses Verfahren wird für alle SGB II-Typen $1 \dots m$ durchgeführt.

Die Fortschreibungswerte dienen sowohl zur Einschätzung der Plausibilität der Daten als auch als Ersatz für den zu meldenden Wert im Fall von Datenausfällen bzw. unplausiblen Lieferungen im jeweiligen Berichtsmonat. Gegebenenfalls können Fortschreibungswerte auch als Basiswert für die Fortschreibung verwendet werden, z.B. bei zwei oder mehr Datenausfällen bzw. unplausiblen Meldungen in Folge.