

Schmidt, Marcus; Stein, Jana Lucia

**Article**

## Die Hochrechnung im Mikrozensus ab 2020

WISTA - Wirtschaft und Statistik

**Provided in Cooperation with:**

Statistisches Bundesamt (Destatis), Wiesbaden

*Suggested Citation:* Schmidt, Marcus; Stein, Jana Lucia (2021) : Die Hochrechnung im Mikrozensus ab 2020, WISTA - Wirtschaft und Statistik, ISSN 1619-2907, Statistisches Bundesamt (Destatis), Wiesbaden, Vol. 73, Iss. 6, pp. 54-66

This Version is available at:

<https://hdl.handle.net/10419/248260>

**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.*



**Marcus Schmidt**

ist Diplom-Mathematiker und Referent im Referat „Mathematisch-statistische Verfahren für Bevölkerungs- und Sozialstatistiken“ des Statistischen Bundesamtes. Er ist zuständig für methodische Untersuchungen und die mathematische Entwicklung in Haushaltsstatistiken und arbeitet an der Neukonzeption und Umsetzung der Hochrechnung im Mikrozensus ab 2020.



**Jana Lucia Stein**

ist Master of Science Mathematik und Referentin im Referat „Mikrozensus – Methodik und Aufbereitung SILC und IKT“ des Statistischen Bundesamtes. Sie ist zuständig für Hochrechnung und Imputation im Mikrozensus und arbeitet an der Neukonzeption und Umsetzung der Hochrechnung im Mikrozensus ab 2020.

# DIE HOCHRECHNUNG IM MIKROZENSUS AB 2020

Marcus Schmidt, Jana Lucia Stein

↘ **Schlüsselwörter:** Mikrozensus – Haushaltserhebung – Gewichtung – Kalibrierung – Antwortausfälle

## ZUSAMMENFASSUNG

Mit dem Ziel, die amtlichen Haushaltsstatistiken zu harmonisieren, wird der Mikrozensus seit 2020 als integriertes System durchgeführt. Neben der Arbeitskräfteerhebung der Europäischen Union sind auch die europäische Haushaltsbefragung über Einkommen und Lebensbedingungen sowie ab 2021 die Statistik zur Nutzung von Informations- und Kommunikationstechnologien modular in den Mikrozensus integriert. Durch ein sogenanntes Kernprogramm, welches weiterhin über eine 1%-Stichprobe erhoben wird, sind die Erhebungsteile thematisch miteinander verknüpft. Dieser Artikel erläutert die neu konzipierte Methodik der Hochrechnung im neuen System ab 2020. Die Neukonzeption soll die Vielfalt an nationalen und europäischen Anforderungen erfüllen.

↘ **Keywords:** *microcensus – household survey – weighting – calibration – unit nonresponse*

## ABSTRACT

*The microcensus has been implemented as an integrated system since 2020, with the aim of harmonising the official household statistics. In addition to the labour force survey (LFS) of the European Union, the European survey of income and living conditions (SILC) is integrated as a module into the microcensus and the same applies to the statistics on the use of information and communication technologies (ICT) as of 2021. The survey components are thematically linked through the core programme of variables which continue to be collected based on a 1% sample. This article explains the redesigned methodology of extrapolation that has been used in the new system since 2020. The redesigned system is aimed to meet the variety of national and European requirements.*

## 1

---

### Einleitung

---

Auf Basis neu gestalteter Auswahl-<sup>1</sup> und Rechtsgrundlagen<sup>2</sup> von 2016 wurde die Durchführung des Erhebungsgeschäfts und die Aufbereitung der Stichproben für den Mikrozensus (MZ) 2020 umgestellt. Ein zentraler Aspekt dieser Umstellung ist die Integration der bisher freiwilligen und eigenständig durchgeführten europäischen Erhebungen EU-SILC (European Union Statistics on Income and Living Conditions – Erhebung über Einkommen, Armut und Lebensbedingungen der Bevölkerung) und IKT (Informations- und Kommunikationstechnologien). Wie bisher enthalten bleibt die Arbeitskräfteerhebung LFS (Labour Force Survey).<sup>3</sup> Während die integrierte MZ-SILC-Erhebung in weiten Teilen mit Auskunftspflicht belegt wird, ist für den Erhebungsteil MZ-IKT weiter eine freiwillige Beantwortung vorgesehen. Beide Erhebungen finden nur in einem begrenzten Zeitraum des Jahres statt. Die Erhebung des Kernprogramms und der MZ-LFS-Substichprobe erfolgt weiterhin über das gesamte Jahr. Für alle Erhebungsteile wurde ab 2020 zudem eine feste Berichtswoche eingeführt, das heißt auch bei eventuellen Verzögerungen der Befragung erfolgen die Antworten der Haushalte zu einer bestimmten Kalenderwoche. Detaillierte Informationen zu den Neuerungen im Mikrozensus sind der Literatur zu entnehmen (Bihler/Zimmermann, 2016; Hundenborn/Enderer, 2019). Dieser Artikel behandelt die Hochrechnung und sich darauf auswirkende Neuerungen im Mikrozensus ab 2020.

Dazu erläutert Kapitel 2 zunächst die hochrechnungsrelevanten Neuerungen im Mikrozensus, die sich ab dem Erhebungsjahr 2020 aus der Neugestaltung ergeben. Kapitel 3 beschreibt die zum Einsatz kommende Methodik der gebundenen Hochrechnung, die Ziehungswahrscheinlichkeiten sind Thema in Kapitel 4. Die Berechnung von Antwortwahrscheinlichkeiten wird in Kapitel 5 dargestellt, während sich die Kapitel 6 und 7 mit den verschiedenen Hochrechnungsfaktoren, der Bereitstel-

lung von Eckwerten und den verwendeten Hochrechnungsmodellen befassen. Methodischer Anpassungsbedarf aufgrund von Erhebungseinschränkungen im Jahr 2020 ist in Kapitel 8 aufgeführt. Ein kurzes Fazit beschließt den Beitrag.

## 2

---

### Hochrechnungsrelevante Neuerungen im Mikrozensus ab 2020

---

#### 2.1 Periodizität der Erhebung

---

Neben den neu integrierten Erhebungen wurde auch die Erfassung zur Arbeitsmarktbeteiligung neu konzipiert. Um unterjährige Veränderungen im Arbeitsmarkt besser abzubilden, erfolgt die Erfassung dieser Substichprobe unterjährig nach dem Rotationschema 2-(2)-2. Das bedeutet, dass die Befragungen in den entsprechenden Auswahlbezirken zunächst in zwei aufeinanderfolgenden Quartalen erfolgt, in den zwei darauffolgenden Quartalen pausiert und dann in den zwei anschließend folgenden Quartalen nochmals durchgeführt wird (Hundenborn/Enderer, 2019, hier: Grafik 2). Um die Befragungsbelastung zu reduzieren und da einige der Merkmale als zeitstabil angenommen werden, unterscheiden sich MZ-LFS-Befragungen in der Zusammenstellung der Erhebungsmerkmale. Es gibt sogenannte Strukturmerkmale, die einmal jährlich abzufragen sind. Diese werden fest in den Befragungswellen eins und drei abgefragt, sowie zusätzlich in der vierten Befragungswelle, sofern die Berichtswoche im ersten Quartal liegt.<sup>4</sup> In den weiteren Befragungen des MZ-LFS werden nur die Quartalsmerkmale – das sind saisonale Merkmale – erfragt.

Für alle weiteren Substichproben gilt das bisherige Rotationsschema. Ein Viertel der Stichprobe rotiert jedes Jahr heraus und wird durch ein neues, hereinrotierendes Viertel ersetzt. Es liegt daher keine einheitliche Rotation über die gesamte Stichprobe vor (Hundenborn/Enderer, 2019).

---

1 Erneuerung der Auswahlgrundlage auf Basis des Zensus 2011.

2 Die Rechtsgrundlagen sind am Ende des Beitrags aufgeführt.

3 Um die in den Mikrozensus integrierten Erhebungen von den eigenständigen beziehungsweise den europäischen Erhebungen abzugrenzen, werden im Beitrag die Bezeichnungen MZ-LFS, MZ-SILC und MZ-IKT verwendet, wo sie zutreffen.

---

4 Diese letzte Bedingung kam hinzu, um den gesetzlichen Anforderungen zu entsprechen: Danach muss jeder Haushalt, der Teil der MZ-LFS-Stichprobe ist, in jedem Kalenderjahr seiner Teilnahme an der Erhebung auch Angaben zu den Strukturmerkmalen geben.

## 2.2 Stichprobe

Beim integrierten Mikrozensus handelt es sich weiterhin um eine geschichtete Klumpenstichprobe von 1 % der Bevölkerung (Bihler/Zimmermann, 2016). Eine proportional einheitliche Verteilung der Stichprobe über alle Bundesländer und Regierungsbezirke kann für den integrierten Mikrozensus nicht mehr gewährleistet werden. Präzisions<sup>5</sup>- und Feldphasenanforderungen in Verbindung mit einem Gesamtauswahlsatz von maximal 1 % der Bevölkerung erfordern, dass die Auswahlsätze der Substichproben auf Ebene der Regierungsbezirke (NUTS 2) variieren (Hundenborn/Enderer, 2019).

Um den Anforderungen eines Panels für die EU-SILC-Erhebung zu genügen, werden Haushalte der MZ-SILC-Substichprobe, die aus ihrer Wohnung wegziehen, bevor sie aus der Stichprobe herausrotieren, weiterverfolgt; ihre Befragung zum Kernprogramm und MZ-SILC erfolgt dann auf freiwilliger Basis. Eine Ausnahme von dieser Regelung sind sogenannte Einzugshaushalte<sup>6</sup>, diese werden im Falle eines Auszugs nicht mehr weiterverfolgt.

## 2.3 Ergebnisbereitstellung

Mit dem neuen System sollen die Ergebnisse des Mikrozensus deutlich früher als bisher zur Verfügung stehen. Dennoch ist auch ein ausreichender Zeitraum notwendig zur Berücksichtigung von Auskünften, die gegebenenfalls durch ein Verwaltungsverfahren infolge der Auskunftspflicht gewonnen werden. Aufgrund dieses Spannungsfelds sind ab 2020 zwei Ergebnisarten vorgesehen: Erst- und Endergebnis.

Beide Ergebnisarten beruhen auf vollständig aufbereiteten und hochgerechneten Daten. Sie unterscheiden sich jedoch in der Menge der vorliegenden Beobachtungen und in der Konsistenz der Hochrechnung der Substichproben. Für Letztere werden hochgerechnete Daten zu den jeweiligen Fristen an Eurostat geliefert und zeitnah

<sup>5</sup> Die Präzisionsanforderungen des Statistischen Amtes der Europäischen Union (Eurostat) an die jeweiligen Substichproben LFS, SILC und IKT sind gegeben als obere Grenze an die geschätzten Standardfehler (siehe Anhang II der Verordnung [EU] 2019/17001700). Diese implizieren wiederum einen vom Designeffekt abhängigen Mindeststichprobenumfang.

<sup>6</sup> Das sind Haushalte, die in einen in Vorwellen bereits befragten Auswahlbezirk der Stichprobe einziehen.

veröffentlicht. Das Endergebnis wird für Kernprogramm und alle Substichproben zum gleichen Zeitpunkt erstellt, sodass Ergebniskonsistenz für genau die in der Hochrechnung verwendeten Anpassungsmerkmale erreicht werden kann. Eine Ausnahme bildet die IKT-Erhebung, für die lediglich ein Endergebnis bereitgestellt wird. Grund dafür ist, dass durch deren Freiwilligkeit die Möglichkeiten einer höheren Ausschöpfung begrenzt sind.

## 3

### Gebundene Hochrechnung

Bei einer Hochrechnung im Sinne einer Vorhersage von unbekanntem Größen einer Grundgesamtheit aus einer Zufallsstichprobe wird allgemein zwischen freier und gebundener Hochrechnung unterschieden. Eine freie Hochrechnung bedingt eine Zufallsauswahl, in der jedes Element mit einer Wahrscheinlichkeit größer 0 unabhängig von anderen Elementen ausgewählt wird. Existieren Einflüsse, bei denen man von Auswirkungen auf die Ziehung beziehungsweise von einer hohen Korrelation mit den Ergebnissen ausgehen kann, und sind diese bekannt, so wird auf eine gebundene Hochrechnung zurückgegriffen. Die gebundene Hochrechnung nutzt Eckwerte aus externen Quellen, um die Schätzung in der freien Hochrechnung zu korrigieren (Deville und andere, 1993). Um die Varianz beziehungsweise Standardabweichung von Ergebnissen erwartungstreu zu schätzen, müssen die externen Eckwerte korrekt sein. Das heißt die Schätzung der Varianz für Ergebnisse der Hochrechnung vernachlässigt jede Unsicherheit in den verwendeten externen Hilfsmerkmalen.

Die gebundene Hochrechnung im Mikrozensus hat grundsätzlich zwei Ziele:

- › Sie soll die systematischen Fehler durch unerkannte Antwortausfälle verringern und ebenso
- › die Zufallsfehler derjenigen Erhebungsmerkmale, die mit den Bezugsmerkmalen hoch korreliert sind.

Durch die Anpassung wird auch eine gewisse Konsistenz zu den in der amtlichen Statistik verfügbaren Bevölkerungsdaten erreicht. Ziel der Hochrechnung ist, den Personen beziehungsweise Haushalten in der Stichprobe Hochrechnungsgewichte zuzuordnen, mit deren Hilfe

Ergebnisse (Totalwerte, Anteilswerte) zu gewissen Merkmalen berechnet werden können.

Wie bereits bisher im Mikrozensus wird der verallgemeinerte Regressionschätzer als Kalibrierungsverfahren zur Erzeugung aller Hochrechnungsfaktoren verwendet.

## 3.1 Der verallgemeinerte Regressionschätzer

Der verallgemeinerte Regressionschätzer (GREG – Generalized Regression Estimator) fungiert als modellunterstützte Erweiterung des designbasierten Horvitz-Thompson-Schätzers (HT-Schätzer). Der HT-Schätzer kann allgemein als freie Hochrechnung verstanden werden und der GREG-Schätzer korrigiert diesen anhand eines Modells aus externen Hilfsvariablen und einem Vektor von geschätzten Regressionskoeffizienten. In einer gewichteten Regression wird also der HT-Schätzer um den Korrekturfaktor des GREG-Schätzers erweitert. Genauere Ausführungen sowie die Herleitung der genutzten Formeln finden sich bei Afentakis und Bihler (2005; hier: Seite 1042). Um die Handhabung bei der Ergebniserstellung aus dem Material des Mikrozensus zu vereinfachen, wird der zu berechnende Hochrechnungsfaktor  $w_k$  wie bisher an das Einzelmateriale angefügt.

Antwortausfälle können nicht vernachlässigt werden (siehe Abschnitt 2.4 in Afentakis/Bihler, 2005). Es wird somit weiterhin auf das zweistufige Hochrechnungsverfahren zurückgegriffen. In der ersten Stufe soll möglichen Verzerrungen vor der Hochrechnung durch Berechnung von Antwortwahrscheinlichkeiten begegnet werden. Die Wahrscheinlichkeit für Person  $k$ , in die Stichprobe zu gelangen, wird daher zusammengesetzt aus der Ziehungswahrscheinlichkeit  $\pi_k$  und der geschätzten Antwortwahrscheinlichkeit  $\hat{\theta}_k$  (siehe Kapitel 5). Für deren Kehrwert  $d_k$  gilt:

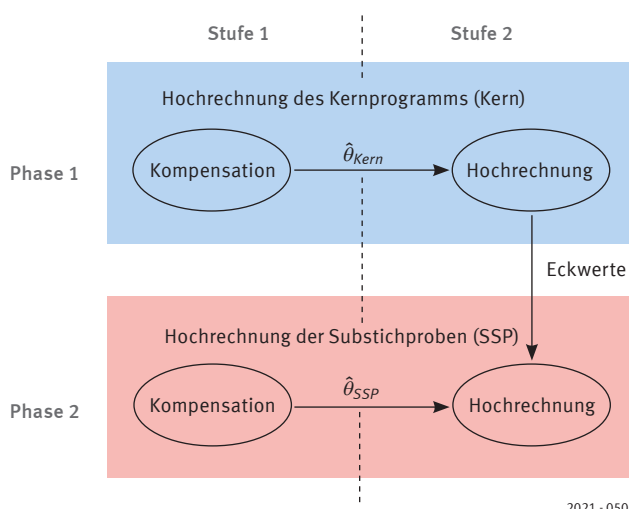
$$(1) \quad d_k = \frac{1}{\pi_k \hat{\theta}_k}$$

Alle Personen eines Haushalts erhalten wie bisher den gleichen Hochrechnungsfaktor (Afentakis/Bihler, 2005).

## 3.2 Zwei-Phasen-Hochrechnung

Zum Zeitpunkt der Substichprobenhochrechnung (Substichproben: MZ-LFS, MZ-SILC und MZ-IKT) steht der bis dahin jeweils erhobene Teil des Kernprogramms zur Verfügung. Dieser enthält auch einige der für die Substichproben relevanten Merkmale. Für diese Merkmale liegen also zum Zeitpunkt der Hochrechnung validere, für eine größere Stichprobe erfasste Informationen vor. Mit dem Ziel der Kohärenz zwischen Kernprogramm und Substichproben (und da einige Merkmale als hochkorreliert angenommen werden können) wird für den Mikrozensus ab 2020 eine Hochrechnung der Substichproben in zwei Phasen angestrebt. Die erste Phase ist die Hochrechnung des kompletten Kernprogramms in der geforderten zeitlichen Untergliederung, beispielsweise Quartal, Halbjahr oder Jahr. In der zweiten Phase findet eine gebundene Hochrechnung der Substichproben an die aus der ersten Phase gewonnenen Eckwerte statt. Dabei wird für jede Phase das zweistufige Verfahren des bisherigen Mikrozensus angewendet. Als erste Stufe erfolgt die Kompensation von Antwortausfällen und in einer zweiten Stufe die eigentliche Hochrechnung an die Grundgesamtheit. Für die Substichproben wird eine getrennte Kompensation mit jeweils eigenen Modellen implementiert. [↪ Grafik 1](#)

**Grafik 1**  
Zwei-Phasen-Hochrechnung des Mikrozensus



2021 - 0508

## 4

### Ziehungswahrscheinlichkeit

Die Einheit für die Hochrechnungen ist die Person beziehungsweise der Haushalt. Im Mikrozensus sind die Auswahlseinheiten jedoch nicht Personen oder Haushalte, sondern Auswahlbezirke. Die Ziehungswahrscheinlichkeiten der Auswahlbezirke gelten für alle Personen/Haushalte des betreffenden Auswahlbezirks, da alle auf der Fläche des Auswahlbezirks in der Berichtswoche wohnenden Personen in die Erhebung einbezogen werden. Bei der Auswahl steht zusätzlich eine zeitliche Komponente<sup>7</sup> zur Verfügung, die Auswahlseinheit ist also im Grunde die Kombination von Auswahlbezirk und Kalenderwoche. Bei den Substichproben MZ-SILC und MZ-IKT sind die Merkmale zeitlich eher stabil, daher kann dieser Aspekt zur einfacheren Darstellbarkeit entfallen. Die zeitliche Komponente ist aber zwingend bei allen Hochrechnungen des MZ-LFS und für die Jahreshochrechnung des Kernprogramms zu berücksichtigen.

<sup>7</sup> Ausgenommen sind diejenigen Merkmale von EU-SILC, die sich nicht wie die demografischen Merkmale oder Haushaltstypen auf die Berichtswoche, sondern auf einen zurückliegenden Stichtag/Zeitraum beziehen.

### 4.1 Kernhochrechnung

Im Jahresmaterial können die Einheiten der Hochrechnung (Personen beziehungsweise Haushalte) mehrfach vertreten sein, da für die MZ-LFS-Substichprobe in Privathaushalten die Kernmerkmale in sieben der neun am Jahr beteiligten Rotationsgruppen für zwei Wochen erhoben wurden. [↪ Grafik 2](#)

Weil das Jahresergebnis als Durchschnitt über die 52 Kalenderwochen angelegt ist, sind alle Befragungen – egal, ob ein oder zwei Befragungszeitpunkte im Jahr vorliegen – gleichwertig. Die normierte Ziehungswahrscheinlichkeit  $\pi_{ij,k}$  einer Einheit  $k$ , die zum Regierungsbezirk  $j$  des Bundeslandes  $i$  gehört und keine Gemeinschaftsunterkunft ist, berechnet sich wie folgt:

$$(2) \pi_{ij,k} = f_{ij}^{gesamt} - f_{ij}^{lfs} + \frac{16}{9} f_{ij}^{lfs} = f_{ij}^{gesamt} + \frac{7}{9} f_{ij}^{lfs}$$

Dabei entsprechen  $f_{ij}^{gesamt}$  und  $f_{ij}^{lfs}$  jeweils dem Auswahlatz der gesamten beziehungsweise der LFS-Stichprobe im Regierungsbezirk  $j$  des Bundeslandes  $i$ .

Für Gemeinschaftsunterkünfte ergeben sich andere Ziehungswahrscheinlichkeiten. Für die Gemeinschaftsunterkünfte der Substichprobe MZ-LFS wird nur die erste

**Grafik 2**

Rotationschema der Auswahlbezirke für die Unterstichprobe zur Arbeitsmarktbeteiligung ab 2020

|                   | I. Quartal   | II. Quartal  | III. Quartal | IV. Quartal  |
|-------------------|--------------|--------------|--------------|--------------|
| Rotationsgruppe 1 | 4. Befragung |              |              |              |
| Rotationsgruppe 2 | 3. Befragung | 4. Befragung |              |              |
| Rotationsgruppe 3 |              | 3. Befragung | 4. Befragung |              |
| Rotationsgruppe 4 |              |              | 3. Befragung | 4. Befragung |
| Rotationsgruppe 5 | 2. Befragung |              |              | 3. Befragung |
| Rotationsgruppe 6 | 1. Befragung | 2. Befragung |              |              |
| Rotationsgruppe 7 |              | 1. Befragung | 2. Befragung |              |
| Rotationsgruppe 8 |              |              | 1. Befragung | 2. Befragung |
| Rotationsgruppe 9 |              |              |              | 1. Befragung |

Hellblau signalisiert Rotationsgruppen mit je zwei Befragungen innerhalb eines Jahres. Dunkelblau sind entsprechend Befragungen mit Strukturmerkmalen (sowie Befragungen von Gemeinschaftsunterkünften).

2021 - 0509



und dritte Befragung durchgeführt, außerdem die vierte Befragung, falls sie in das erste Quartal fällt. Hierdurch ist die zeitliche Gleichverteilung bei den Gemeinschaftsunterkünften des MZ-LFS-Teils nicht mehr gegeben: Im ersten Quartal sind drei Rotationsgruppen beteiligt, in den übrigen Quartalen jeweils zwei Rotationsgruppen. Der Anteil der Ziehungswahrscheinlichkeit aus der Arbeitskräfteerhebung LFS ist somit  $\frac{12}{9} f_{ij}^{lfs}$  für Gemeinschaftsunterkünfte im ersten Quartal beziehungsweise  $\frac{8}{9} f_{ij}^{lfs}$  für Gemeinschaftsunterkünfte in den anderen Quartalen. Die Ziehungswahrscheinlichkeiten der übrigen Teile – MZ-SILC, MZ-IKT und nur Kernprogramm – unterscheiden sich für Gemeinschaftsunterkünfte und Privathaushalte nicht. Die Ziehungswahrscheinlichkeit des übrigen Teils ist also  $f_{ij}^{gesamt} - f_{ij}^{lfs}$ . Für die Jahreshochrechnung des Kernprogramms ist die normierte Ziehungswahrscheinlichkeit einer Einheit  $k$  einer Gemeinschaftsunterkunft, die zum Regierungsbezirk  $j$  des Bundeslandes  $i$  gehört, somit

$$(3) \quad \pi_{ijk} = (f_{ij}^{gesamt} - f_{ij}^{lfs}) + \frac{12}{9} f_{ij}^{lfs} = f_{ij}^{gesamt} + \frac{3}{9} f_{ij}^{lfs},$$

falls die Berichtswoche der Einheit  $k$  im ersten Quartal liegt, und

$$(4) \quad \pi_{ijk} = (f_{ij}^{gesamt} - f_{ij}^{lfs}) + \frac{8}{9} f_{ij}^{lfs} = f_{ij}^{gesamt} - \frac{1}{9} f_{ij}^{lfs}$$

in den sonstigen Fällen.<sup>18</sup>

### Unterjährige Hochrechnung des Kernprogramms

Für die unterjährige Hochrechnung des Kernprogramms werden die beschriebenen Ziehungswahrscheinlichkeiten  $\pi_{ij,k}$  der Jahreshochrechnung entsprechend der berücksichtigten Zeiträume mit einem Faktor multipliziert. Für die Quartalshochrechnung entspricht der Faktor  $\frac{1}{4}$ , für das Halbjahr  $\frac{1}{2}$ .

## 4.2 Substichprobe MZ-LFS

Je Quartal werden vier der neun Rotationsgruppen befragt, das heißt mehr als ein Viertel des Jahresmaterials. Jedoch wird im Quartal jede Person/jeder Auswahlbezirk nur einmal zur Befragung herangezogen. Die zeitlich stabilen Strukturmerkmale werden fest in den Befragungswellen 1 und 3 abgefragt, sowie zusätzlich in der vierten Befragungswelle, sofern die Berichtswoche im ersten Quartal liegt.<sup>19</sup> Auch hier liegt jede Befragungseinheit nur einmal vor, allerdings besteht keine zeitliche Gleichverteilung, da im ersten Quartal drei statt zwei Rotationsgruppen beteiligt sind. Für die Hochrechnung der MZ-LFS-Stichprobe ergeben sich damit folgende Ziehungswahrscheinlichkeiten:

Quartal

$$(5) \quad \pi_{ijk}^{lfsq} = \frac{4}{9} f_{ij}^{lfs},$$

Jahr

$$(6) \quad \pi_{ijk}^{lfsj} = \frac{12}{9} f_{ij}^{lfs}, \text{ falls die Berichtswoche} \\ \text{der Einheit } k \text{ im ersten Quartal liegt,} \\ = \frac{8}{9} f_{ij}^{lfs} \text{ in den sonstigen Fällen.}$$

## 4.3 Substichproben MZ-SILC und MZ-IKT

Die Ziehungswahrscheinlichkeiten entsprechen für die Substichproben MZ-SILC und MZ-IKT dem Auswahlatz je Regierungsbezirk der jeweiligen Stichprobe. Entsprechend dem Vorgehen in der Kernhochrechnung wird einem weggezogenen MZ-SILC-Haushalt die Ziehungswahrscheinlichkeit des Auswahlbezirks zugeordnet, zu dem er im Vorjahr gehörte. Es erfolgt somit keine Modifikation ihrer ursprünglichen Eingangswahrscheinlichkeiten, auch gibt es bei einem Ausfall keine Kompensation (Antwortwahrscheinlichkeit = 1). Aufgrund der geringen Fallzahlen wird angenommen, dass eine etwaige Verzerrung zu vernachlässigen ist.

<sup>18</sup> Anmerkung: Im Durchschnitt über die vier Quartale ergibt sich für die Gemeinschaftsunterkünfte gerade der vorgegebene Jahresauswahlatz.

<sup>19</sup> Letzteres wurde hinzugenommen, um den gesetzlichen Anforderungen zu entsprechen, wonach jeder Haushalt, der Teil der MZ-LFS-Stichprobe ist, in jedem Kalenderjahr seiner Teilnahme an der Erhebung auch Angaben zu den Strukturmerkmalen geben muss.

5

## Antwortwahrscheinlichkeit – Kompensation von Antwortausfällen

Antwortausfälle können nicht als zufällig angenommen werden. Etwaigen Verzerrungen wird daher durch die Berechnung von Antwortwahrscheinlichkeiten begegnet.

Um die Antwortwahrscheinlichkeit  $\hat{\theta}_k$  eines Haushalts  $k$  zu bestimmen, werden Hilfsinformationen aus der Bruttostichprobe identifiziert, denen ein erwartbarer Unterschied im Antwortverhalten unterstellt werden kann. Geplant ist, den antwortenden Haushalten mittels eines Logitschätzers eine Antwortwahrscheinlichkeit zuzuweisen (Ekholm/Laakson, 1991). Die Teilnahme wird als binäre abhängige Variable  $Y$  der Bruttostichprobe angenommen und die Antwortwahrscheinlichkeit durch ein binäres, logistisches Regressionsmodell zwischen Teilnahme und Hilfsvariablen geschätzt (Körner und andere, 2006):

Teilnahme der Einheit  $k$ :

$$(7) \quad Y_k = 1 \quad \text{Prob}(Y_k = 1) = \rho(\beta^t; x_k)$$

Nichtteilnahme der Einheit  $k$ :

$$(8) \quad Y_k = 0 \quad \text{Prob}(Y_k = 0) = 1 - \rho(\beta^t; x_k)$$

Antwortwahrscheinlichkeit der Einheit  $k$ :

$$(9) \quad \hat{\theta}_k = \text{Prob}(Y_k = 1) = \rho(\hat{\beta}^t; x_k) = \frac{e^{\hat{\beta}^t x_k}}{1 + e^{\hat{\beta}^t x_k}}$$

Dabei entspricht  $x_k$  den Ausprägungen der Hilfsvariablen für Einheit  $k$  und  $\hat{\beta}$  den geschätzten Regressionskoeffizienten.

Neu für den Mikrozensus ab 2020 ist, dass nicht nur finale Ausfälle, sondern auch verzögerte Dateneingänge zu berücksichtigen sind. Aufgrund der festen Berichtswoche sind Befragungen einem festen Berichtszeitraum zugeordnet und können nicht wie bisher dem folgenden Berichtszeitraum zugeschlagen werden. Insbesondere für die frühe Hochrechnung der Erstergebnisse ist dies relevant. Um den zeitlichen Verzug zu berücksichtigen, kann auf Basis der Daten aus vorherigen Erhebungsjahren die Antwortwahrscheinlichkeit bestimmt werden. Hierzu können Materialstände des Erstergebnisses mit den Ständen zum Endergebnis abgeglichen werden.

Unter Verwendung des logistischen Regressionsmodells werden Antwortwahrscheinlichkeiten je Merkmalskombination bestimmt. Vorteilhaft hierbei ist, dass für alle berücksichtigten Datensätze der volle Merkmalskranz vorliegt und somit eine große Auswahl an erklärenden Merkmalen zur Verfügung steht. Berechnete Antwortwahrscheinlichkeiten berücksichtigen allerdings lediglich die Verzögerungen, nicht die Totalausfälle. Aus diesem Grund wird eine weitere Antwortwahrscheinlichkeit, welche die aktuelle Bruttostichprobe berücksichtigt, berechnet. Dies geschieht unter Verwendung von Informationen, die für die gesamte Bruttostichprobe vorliegen. Auch für die Berechnung dieser Wahrscheinlichkeiten ist ein logistisches Regressionsmodell vorgesehen. Der Merkmalskranz hierfür ist begrenzt. Vollständig und qualitativ gut liegen alle Merkmale zur Verwendung in der Stichprobenziehung vor. Allerdings besteht für diese Merkmale eine geringere Korrelation zum Antwortverhalten. Aus Beobachtungen der Interviewenden liegen Daten für weitere soziodemografische Merkmale vor. Umfang und Qualität zu den frühen Hochrechnungszeitpunkten müssen allerdings künftig in einem funktionierenden System untersucht werden.

Die Antwortwahrscheinlichkeit  $\hat{\theta}_k$  für eine Einheit  $k$  wird ermittelt durch

$$(10) \quad \hat{\theta}_k = \alpha \hat{\theta}_{k1} + (1 - \alpha) \hat{\theta}_{k2}$$

mit

$\hat{\theta}_{k1}$ : Antwortwahrscheinlichkeit der Einheit  $k$  berechnet auf vorherigem Datenmaterial,

$\hat{\theta}_{k2}$ : Antwortwahrscheinlichkeit der Einheit  $k$  berechnet aus der aktuellen Stichprobe und

$\alpha \in [0,1]$ . Durch die Wahl des Parameters  $\alpha$  kann der Fokus stärker auf die aktuelle Stichprobe oder auf das Antwortverhalten der Vorjahre gelegt werden.

### Sonderfall Substichprobe MZ-IKT

Für die freiwillige Substichprobe MZ-IKT ergibt sich die Antwortwahrscheinlichkeit als Produkt aus der Antwortwahrscheinlichkeit des Mikrozensus und Teilnahme-wahrscheinlichkeit, das heißt der Wahrscheinlichkeit, an der MZ-IKT-Erhebung teilzunehmen. Auch die Teilnahmewahrscheinlichkeit wird über das logistische Regressionsverfahren bestimmt. Hierfür werden die an der IKT teilnehmenden Haushalte den restlichen Haushalten, die



auskunftspflichtig zum Kernteil geantwortet haben und Teil der IKT-Bruttostichprobe sind, gegenübergestellt.

Für die gesonderte Hochrechnung der Personen wird analog eine Teilnahmewahrscheinlichkeit auf Personenebene erstellt. Die Antwortwahrscheinlichkeit des Haushalts wird für alle Personen übernommen.

## 6

### Hochrechnungsfaktoren

Für jeden Erhebungsteil findet eine getrennte Hochrechnung statt, die die jeweils zugrunde liegende Stichprobe berücksichtigt. Die wichtigsten Hochrechnungsfaktoren des Mikrozensus einschließlich ihrer entsprechenden Bezeichnung bei europäischen Datenlieferungen zeigt [Übersicht 1](#).

Für Auswertungen von Merkmalskombinationen muss immer auf die gemeinsame Stichprobe zurückgegriffen werden. Sollen beispielsweise Kernmerkmale zusammen mit MZ-SILC-Merkmalen ausgewertet werden, ist hierfür die MZ-SILC-Stichprobe mit dem zugehörigen Hochrechnungsfaktor zu verwenden.

#### Übersicht 1

##### Hochrechnungsfaktoren im Mikrozensus

| Faktoren  | Stichprobe          | Nationale Bezeichnung     | Bezeichnung Eurostat       |
|---|---------------------|---------------------------|----------------------------|
| Monatliche Arbeitsmark-Schnellmeldung                 | MZ-Kern             | HR000MM                   | –                          |
| Nationales Quartalsergebnis                           | MZ-Kern             | HR000QQ                   | –                          |
| Nationales Jahresergebnis                             | MZ-Kern             | HR000JJ                   | –                          |
| Tief regionalisierte Auswertung des Jahresergebnisses | MZ-Kern             | HR000RG                   | –                          |
| MZ-LFS-Quartalsergebnis                               | MZ-LFS              | HR100QQ                   | COEFFQ                     |
| MZ-LFS-Quartalsdurchschnitt zur Jahresauswertung      | MZ-LFS              | HR100JQ                   | COEFFQ/4                   |
| MZ-LFS-Jahresergebnis                                 | MZ-LFS-Struktur     | HR100JJ                   | COEFFY = COEFF2Y = COEFFHH |
| Ad-Hoc-Modul LFS                                      | MZ-LFS-Ad-Hoc-Modul | HR100MO                   | COEFFMOD                   |
| MZ-SILC-Querschnitt                                   | MZ-SILC             | HR200JJ                   | DB090 = RB050 = PB040      |
| MZ-SILC-Faktor für Kinder bis 12 Jahre                | MZ-SILC             | HR200KI                   | RL070                      |
| 2-, 3-, 4-jähriger Längsschnitt                       | MZ-SILC             | HR200L2, HR200L3, HR200L4 | RB062, RB063, RB064        |
| MZ-IKT Personen                                       | MZ-IKT              | HR300PN                   | IND_WGHT                   |
| MZ-IKT Haushalte                                      | MZ-IKT              | HR300HH                   | HH_WGHT                    |

## 7

### Bereitstellung der Eckwerte

Für die Substichproben steht zum Zeitpunkt der Hochrechnung der bis dahin jeweils erhobene Teil des Kernprogramms zur Verfügung. Hieraus können Aggregate für die zweite Phase der gebundenen Hochrechnung generiert werden. Für die erste Phase der gebundenen Hochrechnung werden Informationen über die Größe und Struktur der Gesamtbevölkerung verwendet, sogenannte externe Eckwerte. Die Datenquelle hierfür ist die Fortschreibung des Bevölkerungsbestands (Bevölkerungsfortschreibung). Die jeweils aktuellen Bevölkerungszahlen werden durch die Fortschreibung der Ergebnisse des jeweiligen letzten Zensus ermittelt und sind drei bis vier Monate nach dem Ablauf des Berichtsmonats beziehungsweise sechs Monate nach dem Berichtsjahr verfügbar. Zum Zeitpunkt der ersten Hochrechnung liegen noch keine aktuellen Werte vor, daher werden die Bevölkerungseckzahlen quartalsweise für die drei zurückliegenden und die vier kommenden Monate vorausgeschätzt. Da die Mikrozensuserhebung gleichmäßig über alle Kalenderwochen eines Berichtszeitraums durchgeführt wird, stellen Quartalseckwerte den Mittelwert aus den Monatswerten, Jahreseckwerte den Durchschnitt aus den Quartalen dar. Monatliche Eckwerte werden als Durchschnitt aus dem Vormonat und dem Berichtsmonat gebildet.

Eine weitere Neuerung ist, dass die Eckwerte zur Staatsangehörigkeit künftig aus der Bevölkerungsfortschreibung und nicht wie bisher aus den Zahlen des Ausländerzentralregisters generiert werden.<sup>10</sup> Seit dem Zensus 2011 erfolgt die Fortschreibung nach einzelnen Staatsangehörigkeiten und bietet somit eine konsistente Datenquelle für alle Eckwerte zur Hochrechnung der Mikrozensusstichprobe.

Zur Vorausschätzung ab 2020 wurde das bisher verwendete Verfahren weiterentwickelt: Zunächst wird die „Bevölkerung insgesamt“ unter Verwendung des sogenannten Holt-Winters-Prognoseverfahrens je Bundesland vorausgeschätzt. Basis hierfür sind Zeitreihen über die monatliche Entwicklung der Geburten, Sterbefälle und Zu- und Fortzüge über die Grenzen des Bundeslands, die mit der letzten vorhandenen Bevölkerungszahl verrechnet werden. Im nächsten Schritt werden hieraus die untergliederten Eckwerte auf Landesebene abgeleitet. Dafür werden zunächst die Anteile dieser Eckwerte an der Gesamtbevölkerung des Bundeslands in der Vergangenheit nach Monaten berechnet. Die monatlichen Zeitreihen werden anschließend extrapoliert und mit der projizierten Bevölkerung multipliziert. Die verwendeten Zeitreihen basieren auf der monatlichen Entwicklung ab Januar 2016. Lediglich für die Prognose der Anteile nach den drei nichtdeutschen Staatsangehörigkeitsgruppen<sup>11</sup> haben die Testrechnungen ergeben, dass ein kür-

zerer Zeitraum von 24 Werten eine bessere Anpassung liefert.

Im letzten Schritt werden die Eckwerte nach Kreisen berechnet. Hierfür wird das bisherige Vorgehen übernommen. Die vorausgeschätzten großen Aggregate auf Landesebene werden nach Kreisen verteilt. Da die Entwicklung der einzelnen Eckwerte nach Kreisen stark volatil ist, wird auf eine Zeitreihenextrapolation verzichtet. Die relativen Verteilungen nach Kreis-Eckwerten werden deshalb im Prognosezeitraum konstant gehalten.

Zum Zeitpunkt der zweiten Hochrechnung für Endergebnisse liegen bereits endgültige Zahlen der Bevölkerungsfortschreibung vor. Diese werden in gleicher Untergliederung verwendet.

### Geschlechtsangabe divers/kein Eintrag

Da aus dem Zensus 2011 keine Angaben zu den Merkmalsausprägungen divers/ohne Angabe vorliegen, stehen aktuell für die Hochrechnung keine Eckwerte hierzu aus der Bevölkerungsfortschreibung zur Verfügung. Derzeit werden die Personen ohne Angabe zum Geschlecht in der Bevölkerungsfortschreibung und im Mikrozensus noch zufällig den Ausprägungen männlich oder weiblich zugeordnet.

### Hilfsvariablen für die gebundene Hochrechnung

↳ **Übersicht 2** stellt den Anpassungsrahmen für die gebundene Hochrechnung des Kernmaterials an die Bevölkerung dar. Grundlage zur Bestimmung des Anpas-

10 Es gab aufgrund unterschiedlicher Berechnungsmethodik regelmäßig Abweichungen zwischen den Zahlen der Bevölkerungsfortschreibung und denen des Ausländerzentralregisters.

11 Folgende Staatsangehörigkeitsgruppen werden gesondert berücksichtigt: türkisch, EU-25, nicht EU-25.

## Übersicht 2

### Anpassungsrahmen Kernprogramm

| Merkmal   | Ausprägung  | Regionale Ebene              | Berichtszeitraum |
|---|---|------------------------------|------------------|
| Alter (zum 31.12. des Jahres) *Geschlecht                         | (<15, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+)<br>*(männlich, weiblich) | NUTS- 1                      | Jahr/ Quartal    |
| Alter (zum 31.12. des Jahres)<br>*Geschlecht *Staatsangehörigkeit | (15-24, 25-54, 55-74)<br>*(männlich, weiblich)<br>*(deutsch, nichtdeutsch)    |                              | Jahr/ Quartal    |
| Staatsangehörigkeit *Geschlecht                                   | (deutsch, türkisch, EU, Rest) *(männlich, weiblich)                           |                              | Jahr/ Quartal    |
| Monatsanpassung   | Bevölkerung je Monat  |                              | Jahr/ Quartal    |
| Staatsangehörigkeit<br>*Geschlecht                                | (deutsch, nichtdeutsch)<br>*(männlich, weiblich)                              | NUTS-2                       | Jahr/ Quartal    |
| Geschlecht  | (männlich, weiblich)  | Regionale Anpassungsschicht  | Quartal          |
| Staatsangehörigkeit<br>*Geschlecht                                | (deutsch, nichtdeutsch)<br>*(männlich, weiblich)                              |                              | Jahr             |
| Geschlecht  | (männlich, weiblich)  | Regionale Schichtuntergruppe | Jahr             |

## Übersicht 3

### Anpassungsrahmen des MZ-LFS für die Quartalshochrechnung

| Haushalt/Personen  | Merkmal                            | Ausprägung   | Regionale Ebene             |
|--|------------------------------------|--|-----------------------------|
| Personen<br>„Bevölkerung in Privathaushalten am Hauptwohnsitz“ | Monatsanpassung                    | Bevölkerung je Monat   | NUTS-0                      |
|  | Staatsangehörigkeit<br>*Geschlecht | (deutsch, türkisch, EU, Rest)<br>*(männlich, weiblich)   | Ost/ West                   |
|  | Erwerbsstatus*Alter<br>*Geschlecht | (erwerbstätig, erwerbslos,<br>Nichterwerbsperson)<br>*(15-24, 25-54, 55-74)<br>*(männlich, weiblich) | NUTS-1                      |
|  | Alter*Geschlecht                   | (15-24, 25-34, 35-44, 45-54,<br>55-64, 65-74)<br>*(männlich, weiblich)                               | NUTS-2                      |
|  | Bevölkerung insgesamt              |  | Regionale Anpassungsschicht |

sungsrahmens waren fachliche Analysen sowie die erwarteten Besetzungszahlen der Anpassungsklassen. Die Untergliederung mit einer gesonderten Ausweisung türkischer Staatsangehöriger wurde beibehalten (Afen-takis/Bihler, 2005). Grund dafür ist, dass auch auf Basis aktueller Daten die Stichprobenverteilung nach freier Hochrechnung für Personen mit türkischer Staatsangehörigkeit weniger von den Eckwerten abweicht als bei der übrigen nicht deutschen Bevölkerung.

Beibehalten wurde auch die automatische Reduktion der Hochrechnungsterme, wenn eine Anpassungszelle weniger als zehn befragte Personen enthält. Ziel ist es, eine instabile Schätzung durch zu geringe Besetzungszahlen in den Anpassungsklassen zu vermeiden. Hierbei wird eine mögliche fachliche gegenüber einer geografischen Reduktion bevorzugt, sodass möglichst alle in

der Anpassung vorgesehenen regionalen Ebenen verwendet werden.

Zur Hochrechnung der Substichproben kann auf die Merkmale des Kernprogramms zugegriffen werden. Hierbei sind die ab 2021 geltenden Vorgaben Eurostats und die gewünschte fachliche Konsistenz zwischen Stichproben zu berücksichtigen. Die Lieferungen an Eurostat umfassen lediglich Personen in privaten Haushalten am Hauptwohnsitz (Hochgürtel/Weinmann, 2020), daher bezieht sich die Hochrechnung für Substichproben auf diesen Personenkreis.

Insbesondere die neue Durchführungsverordnung des EU-LFS enthält deutlich stärkere Anforderungen als bisher. Künftig soll die Verteilung der Referenzpopulation nach Geschlecht, Altersgruppen und NUTS-2-Region

## Übersicht 4

### Anpassungsrahmen des MZ-LFS für die Jahreshochrechnung

| Haushalt/Personen  | Merkmal                         | Ausprägung  | Regionale Ebene             |
|--|---------------------------------|---|-----------------------------|
| Haushalt<br>„Privathaushalte am Hauptwohnsitz“                 | Anzahl Haushalte                |   | NUTS-0                      |
|  | Haushaltsgröße                  | (1, 2, 3, 4, 5 und mehr)  |                             |
| Personen<br>„Bevölkerung in Privathaushalten am Hauptwohnsitz“ | Erwerbsstatus*Alter*Geschlecht  | (erwerbstätig, erwerbslos, Nicht-<br>erwerbsperson) *(15-24, 25-34,<br>35-44, 45-54, 55-64,<br>> 64 nM))*(männlich, weiblich) | NUTS-0                      |
|  | Monatsanpassung                 | Bevölkerung je Monat  |                             |
|  | Staatsangehörigkeit* Geschlecht | (deutsch, türkisch, EU, Rest)<br>*(männlich, weiblich)  | Ost/West                    |
|  | Erwerbsstatus*Alter*Geschlecht  | (erwerbstätig, erwerbslos,<br>Nichterwerbsperson) *(15-24, 25-<br>54, 55-74)*(männlich, weiblich)                             | NUTS-1                      |
|  | Alter*Geschlecht                | (15-24, 25-34, 35-44, 45-54,<br>55-64, 65-74)<br>*(männlich, weiblich)  | NUTS-2                      |
|  | Bevölkerung insgesamt           |   | Regionale Anpassungsschicht |

## Übersicht 5

### Der Anpassungsrahmen für MZ-SILC

| Haushalt/Personen  | Merkmal             | Ausprägung   | Regionale Ebene |
|--|---------------------|--|-----------------|
| Haushalt<br>„Privathaushalte am Hauptwohnsitz“                 | Haushaltsgröße      | (1, 2, 3, 4, 5+)   | NUTS-0          |
|  | Haushaltstyp        | (Einpersonenhaushalt, Paare ohne Kind, Alleinerziehende, Paare mit Kindern, Sonstige)              |                 |
|  | Haushaltseinkommen  | (<1 500, 1 500-2 000, 2 000-3 000, 3 000-4 000, >4 000)  |                 |
|  | Haushaltstyp2       | (ohne Kind/mit Kindern)  | NUTS-1          |
|  | Anzahl              |  | NUTS-2          |
| Personen<br>„Bevölkerung in Privathaushalten am Hauptwohnsitz“ | Alter*Geschlecht    | (25-49, 50-64, 65+) *Geschlecht  | NUTS-0          |
|  | Staatsangehörigkeit | (deutsch, EU, nicht EU)  |                 |
|  | Alter               | (0-15, 16-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75+)  |                 |
|  | ISCED               | (niedrig, mittel, hoch)  |                 |
|  | Hauptstatus         | (arbeitslos, angestellt, selbstständig, Schüler/-in/Studierende, im Ruhestand; nicht erwerbstätig) | NUTS-1          |
|  | Geschlecht          | (männlich, weiblich)   |                 |
|  | Alter               | (0-24, 25-64, 65+)   |                 |
|  | Anzahl              |  |                 |

berücksichtigt werden. Eine Kombination der Merkmale ist gewünscht, aber nicht zwingend notwendig. Zusätzlich soll bei der Gewichtung der Strukturvariablen (Jahresergebnis) die Anzahl der Haushalte sowie die Verteilung nach Haushaltsgröße einbezogen werden. Zudem ist Konsistenz zum Quartalsdurchschnitt hinsichtlich des Erwerbsstatus aufgedgliedert nach Alter und Geschlecht sicherzustellen. Somit müssen bei Hochrechnung der Strukturmerkmale die Eckwerte der LFS-Quartalsstichprobe berücksichtigt werden. Aufgrund der monatlichen Schnellmeldung zum Arbeitsmarkt liegen einige der wichtigsten LFS-Merkmale im gesamten Kernprogramm vor, an Eurostat wird jedoch nur die Substichprobe geliefert. Europäische Konsistenzanforderungen zwischen Stichproben werden gegenüber nationalen Konsistenz-

anforderungen zum Kernjahresergebnis priorisiert. Daher wird die Stichprobe der LFS-Strukturmerkmale an den Quartalsdurchschnitt der Kernstichprobe hochgerechnet und nicht an das Jahresmaterial. Gewählte Anpassungsrahmen für die Quartals- und Jahreshochrechnung enthalten die [Übersichten 3 und 4](#).

[Übersicht 5](#) zeigt den geplanten Hochrechnungsrahmen für MZ-SILC. Es wird wie bisher an das Haushaltsnettoeinkommen angepasst. Die Klassenzusammenfassung wurde entsprechend der geschätzten Quintilverteilung im Mikrozensus vor 2020 vorgenommen. Aufgrund der Erhöhung des Stichprobenumfangs im System ab 2020 ist ein gesonderter Faktor für Personen ab 16 Jahren, wie er bisher vorlag, nicht mehr notwendig.

## Übersicht 6

### Anpassungsrahmen MZ-IKT für Haushalte

| Haushalt/Personen  | Merkmal                 | Ausprägung  | Regionale Ebene        |
|--|-------------------------|---|------------------------|
| Haushalt<br>„Privathaushalte am Hauptwohnsitz“<br>mindestens eine Person zwischen 16 und 74 Jahren | Internetzugang          | (ja, nein)  | NUTS-0                 |
|  | Haushaltsnettoeinkommen | (<1 500, 1 500-2 500, 2 500-4 000, >4 000)                      |                        |
|  | Haushaltstyp            | (Einpersonenhaushalt, Haushalt ohne Kind, Haushalt mit Kindern) |                        |
|  | Anzahl Haushalte        |   | NUTS-1 Zusammenfassung |

## Übersicht 7

### Anpassungsrahmen MZ-IKT für Personen

| Haushalt/Personen   | Merkmal           | Ausprägung   | Regionale Ebene        |
|---|-------------------|--|------------------------|
| Personen<br>„Bevölkerung in Privathaushalten am Hauptwohnsitz“<br>zwischen 16 und 74 Jahren | Internetnutzung   | (ja, nein)   | NUTS-0                 |
|   | Alter *Geschlecht | (16-24, 25-34, 35-44, 45-54, 55-64, 65-74) *(männlich, weiblich) |                        |
|   | ISCED             | (niedrig, mittel, hoch)  |                        |
|   | Soziale Stellung  | (erwerbstätig, nicht erwerbstätig)                               |                        |
|   | Gesamtbevölkerung |  | NUTS-1-Zusammenfassung |

Für MZ-IKT werden auch weiterhin zwei Hochrechnungsfaktoren bereitgestellt, um Personen- und Haushaltsauswertungen abzudecken. Die Anpassungsrahmen sind abgebildet in den [Übersichten 6 und 7](#).

## 8

### Erhebungseinschränkungen im Jahr 2020

Im Jahr 2020 war die Durchführung der Mikrozensus-erhebung eingeschränkt. Dies lag zum einen an der Umstellung des IT-Systems und wurde zum anderen durch die Corona-Pandemie zusätzlich verschärft, da bisher überwiegend persönlich vor Ort stattfindende Befragungen nicht mehr durchgeführt werden konnten. Die durchschnittliche Ausfallquote auf Bundesebene lag 2020 hierdurch bedingt bei etwa 38% und damit deutlich höher als in vorherigen Jahren (nähere Informationen hierzu unter [www.destatis.de/mikrozensus2020](http://www.destatis.de/mikrozensus2020)).

Die beobachteten Antwortausfälle waren fachlich, regional und zeitlich über das Jahr 2020 sehr unterschiedlich verteilt. Zusätzlich war aufgrund der Erhebungseinschränkungen eine vollständige Erfassung der Bruttostichprobe mit Kompensationsmerkmalen auf Haushaltsebene nicht möglich. Um dennoch etwaigen Verzerrungen entgegenzuwirken, wurde die Annahme getroffen, dass für zeitstabile Merkmale die Bruttostichprobenverteilung 2020 der Stichprobenverteilung 2019 entspricht. Unter Berücksichtigung der neuen Auswahl-sätze wurden hierdurch synthetische Bruttofallzahlen für 2020 bestimmt. Die Antwortwahrscheinlichkeiten für den Mikrozensus 2020 wurden unter Verwendung des bisherigen Verfahrens (Anpassung der Netto- an die Bruttofallzahlen [Afentakis/Bihler, 2005]) je Stich-

probe auf Bundeslandebene berechnet. Das in Kapitel 5 beschriebene Verfahren wurde bisher nicht angewandt.

Weiteren Anpassungsbedarf gab es in den Hochrechnungsmodellen. Es konnten weniger Terme als in den beschriebenen Modellen (Kapitel 7) aufgenommen werden. Eine Anpassung für die Kernstichprobe wurde nur bis zur Ebene der Anpassungsschichten durchgeführt, auf eine monatliche Anpassung wurde verzichtet. Auch für die Substichproben waren Reduktionen der Modelle notwendig. Die Qualität der Jahresergebnisse aus dem Mikrozensus 2020 ist auf Bundesebene weiterhin gewährleistet, jedoch sind die Möglichkeiten zur Auswertung in fachlicher und regionaler Tiefe für dieses außergewöhnlich schwierige Erhebungsjahr eingeschränkt.

## 9

### Fazit

Mit der Umstellung auf den Mikrozensus 2020 ist es möglich, eine Vielzahl von Anforderungen an die amtlichen Haushaltsstatistiken aus einer gemeinsamen Quelle konsistent und flexibel zu bedienen. Da die Komplexität des Systems stark gestiegen ist, wurde die Methodik soweit möglich auf moderne und den Anpassungen gerechte Weise aktualisiert und erweitert. Die wichtigsten methodischen Änderungen wurden in diesem Beitrag vorgestellt. Aufgrund des Umfangs und der Komplexität konnte jedoch nur eine Auswahl der erzeugten Hochrechnungsfaktoren des Mikrozensus beleuchtet werden. Darüber hinaus werden derzeit weitere Entwicklungen bearbeitet und analysiert. [!!!](#)

## LITERATURVERZEICHNIS

---

Afentakis, Anja/Bihler, Wolf. [Das Hochrechnungsverfahren beim unterjährigen Mikrozensus ab 2005](#). In: Wirtschaft und Statistik. Ausgabe 10/2005, Seite 1039 ff.

Bihler, Wolf/Zimmermann, Daniel. [Die neue Mikrozensusstichprobe ab 2016](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2016, Seite 20 ff.

Deville, Jean-Claude/Särndal, Carl-Erik/Sautory, Olivier. *Generalized Raking Procedures in Survey Sampling*. In: Journal of the American Statistical Association. Jahrgang 88. Ausgabe 09/1993, Nr. 423, Seite 1013 ff.  
[DOI 10.1080/01621459.1993.10476369](https://doi.org/10.1080/01621459.1993.10476369)

Ekholm, Anders/Laaksonen, Seppo. *Weighting via Response Modeling in the Finnish Household Budget Survey*. In: Journal of Official Statistics. Jahrgang 1991. Ausgabe 7, Nr. 3, Seite 325 ff.

Hochgürtel, Tim/Weinmann, Julia. [Haushalte in der Berichterstattung des Mikrozensus ab 2020](#). In: WISTA Wirtschaft und Statistik. Ausgabe 3/2020, Seite 89 ff.

Hundenborn, Janina/Enderer, Jörg. [Die Neuregelung des Mikrozensus ab 2020](#). In: WISTA Wirtschaft und Statistik. Ausgabe 6/2019, Seite 9 ff.

Körner, Thomas/Nimmergut, Anja/Nökel, Jens/Rohloff, Sandra. [Die Dauerstichprobe befragungsbereiter Haushalte](#). In: Wirtschaft und Statistik. Ausgabe 5/2006, Seite 451 ff.

Statistisches Bundesamt. *Zu den Auswirkungen der methodischen Neugestaltung sowie der Corona-Krise auf die Ergebnisse des Mikrozensus 2020*. 2021. [Zugriff am 8. November 2021]. Verfügbar unter: [www.destatis.de](http://www.destatis.de)

## RECHTSGRUNDLAGEN

---

Verordnung (EU) 2019/17001700 des Europäischen Parlaments und des Rates vom 10. Oktober 2019 zur Schaffung eines gemeinsamen Rahmens für europäische Statistiken über Personen und Haushalte auf der Grundlage von Einzeldaten aus Stichprobenerhebungen, zur Änderung der Verordnungen (EG) Nr. 808/2004, (EG) Nr. 452/2008 und (EG) Nr. 1338/2008 des Europäischen Parlaments und des Rates sowie zur Aufhebung der Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates und der Verordnung (EG) Nr. 577/98 des Rates (Amtsblatt der EU Nr. L 261, Seite I/1).

Gesetz zur Durchführung einer Repräsentativstatistik über die Bevölkerung und die Arbeitsmarktbeteiligung sowie die Wohnsituation der Haushalte (Mikrozensusgesetz – MZG) vom 7. Dezember 2016 (BGBl. I Seite 2826), das durch Artikel 178 der Verordnung vom 19. Juni 2020 (BGBl. I Seite 1328) geändert worden ist.



**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 Dezember 2021  
Ältere Ausgaben finden Sie unter [www.destatis.de](http://www.destatis.de) sowie in der [Statistischen Bibliothek](#).

---

Artikelnummer: 1010200-21006-4, ISSN 1619-2907

---

© Statistisches Bundesamt (Destatis), 2021  
Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.