

Im Mikrozensus (MZ) werden zur Anpassung der Stichprobe an ausgewählte Populationsverteilungen Anpassungsvariablen (Hilfsvariablen, X) verwendet, die sowohl für die Stichprobe als auch für die Population bekannt sind, z. B. Alter und Geschlecht. Die Konstruktion der Hochrechnungsfaktoren (Gewichte) erfolgt seit MZ 2005 mittels verallgemeinerter Regressionsschätzung (GREG). Diese wird auch bei der Hochrechnung eingesetzt. Die Varianz des GREG-Schätzers entspricht näherungsweise der Varianz mit Designgewichtung, wenn in der Varianzformel statt der interessierenden Variablen (Y) die Residuen der Regression von X auf Y eingesetzt werden; siehe

Schimpl-Neimanns, B., 2011: Schätzung des Stichprobenfehlers in Mikrozensus Scientific Use Files ab 2005. AStA Wirtschafts- und Sozialstatistisches Archiv 5 (1): 19-38. DOI: [10.1007/s11943-011-0092-4](https://doi.org/10.1007/s11943-011-0092-4). <http://www.ssoar.info/ssoar/handle/document/43866>.

Für jeden Erhebungszeitpunkt (JJ) stehen zur Konstruktion der Hilfsvariablen für die gebundene Hochrechnung (Zell-/Anpassungsgewichtung) SAS-, SPSS- und StataProgramme xvar* bereit. Des Weiteren zeigen Beispielprogramme die praktische Umsetzung für Scientific Use Files des MZ (beispiele*):

1. Designbasierte Schätzung des Gesamtwertes Zahl der Erwerbslosen
2. Schätzung des Gesamtwertes mit gebundener Hochrechnung (Regressionsschätzung)
3. Schätzung der Erwerbslosenquote mit Designgewichtung
4. Schätzung der Erwerbslosenquote mit gebundener Hochrechnung

Die Beispiele beziehen sich auf den Erhebungszeitpunkt 2015. Sie können mit geringfügigen Änderungen auch für die anderen Daten ab 2005 verwendet werden.

Im Vergleich zu früheren Versionen (06/2019) wurde die Programmsyntax vereinfacht und teilweise an neue Programmversionen angepasst.

Bei der Berechnung der Mittelwerte der Indikatoren der Hilfsmerkmale im Haushalt wird nur noch die Bevölkerung am Hauptwohnsitz berücksichtigt. Dadurch ergeben sich geringfügige und i. d. R. unwesentliche Änderungen gegenüber früheren Versionen.

Die Abgrenzung der Schicht mit der Kombination der Merkmale EF1 Bundesland, EF12 Berichtsquartal und EF570 Gebäudegrößenklasse führt dazu, dass einige Schichtgruppen nur einen Auswahlbezirk („singuläre PSUs“) enthalten. In den Beispielprogrammen betrifft dies wenige Schichten der Gebäudeschicht Gemeinschaftsunterkunft (EF570=4). Mit SAS und SPSS bleiben singuläre PSUs bei der Schätzung unberücksichtigt. Mit Stata kann analog die Option „singleunit(certainty)“ gewählt werden.

Im Zeitverlauf von 2005 bis 2015 sind folgende Änderungen zu beachten:

- * Die Codierung der Bundesländer (EF1) liegt im MZ 2005 und 2006 nach alphabetischer, im MZ ab 2007 nach geographischer Sortierung vor.
- * In den SUF des Mikrozensus 2006 und 2007 sind die Identifikatoren (IDs, Ordnungsnummern) für Haushalte nicht eindeutig. Für die Abgrenzung der Hilfsmerkmale und die beispielhaften Varianzschätzungen werden Daten mit rekodierten Identifikatoren verwendet; siehe https://www.gesis.org/fileadmin/upload/dienstleistung/daten/amtl_mikrodaten/mz/syntax/misc/ID_2006.do, *.sas, *.sps und https://www.gesis.org/fileadmin/upload/dienstleistung/daten/amtl_mikrodaten/mz/syntax/misc/ID_2007.do, *.sas, *.sps.
- * In den MZ 2005 bis 2011 basieren die Gewichtungsvariablen (EF951, EF952) auf den Original MZ-Hochrechnungsfaktoren, ab MZ 2006 sind die Gewichtungsvariablen bereits mit dem Faktor 1/0.7 der 70%-Substichprobe multipliziert.
- * In den MZ ab 2012 wurden die Hochrechnungsfaktoren der Originaldaten nachträglich an hochgerechnete Randverteilungen des MZ angepasst nach Geschlecht, Altersgruppen (<15 Jahre, 15-44 Jahre, 45-64 Jahre, >=65 Jahre), Erwerbsstatus (erwerbstätig, nicht erwerbstätig) pro Bundesland. Im MZ 2012 erfolgte dies ohne Berücksichtigung des Haushaltszusammenhangs; die Hochrechnungsfaktoren sind deshalb im Haushalt nicht konstant. In den MZ ab 2013 wurde der Haushaltszusammenhang berücksichtigt.
- * Bei der Substichprobenziehung der SUF wurden bis MZ 2011 Haushalte bzw. Wohnungen als Auswahlinheiten verwendet. Ab MZ 2012 sind es Auswahlbezirke innerhalb eines Rotationsviertels. Ab 2012 sind auch Auswahlbezirke ohne erhobene Personen in der Substichprobe enthalten (siehe EF30=-8, -7).