

Mikrozensus Scientific Use File 1976: Dokumentation und Datenaufbereitung

Lechert, Yvonne; Lüttinger, Paul

Veröffentlichungsversion / Published Version

Verzeichnis, Liste, Dokumentation / list

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Lechert, Y., & Lüttinger, P. (2006). *Mikrozensus Scientific Use File 1976: Dokumentation und Datenaufbereitung*. (ZUMA-Methodenbericht, 2006/02). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-262450>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

ZUMA-Methodenbericht 2006/02

Mikrozensus Scientific Use File 1976
Dokumentation und Datenaufbereitung

Yvonne Lechert / Paul Lüttinger

Januar 2006

ISSN 1610-9953

ZUMA
German Microdata Lab (GML)
Quadrat B2,1
Postfach 12 21 55
68072 Mannheim

Telefon: (0621) 1246 - 253
Telefax: (0621) 1246 - 100
E-mail: mikrodaten@zuma-mannheim.de

INHALTSVERZEICHNIS

1. Einleitung	1
2. Mikrozensus 1976	1
2.1 Kurzbeschreibung	1
2.2 Datenerhebung	2
2.3 Die Stichprobe	2
2.4 Das Erhebungsprogramm	2
3. Der faktisch anonymisierte Mikrozensus 1976 – Das Scientific Use File	3
4. Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 1976 und Überprüfung der Plausibilität des erstellten Systemfiles	6
4.1 Einlesen des Rohdatenmaterials und Erstellung des Systemfiles	6
4.2 Gewichtung und Hochrechnung zum Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes	7
4.3 Bevölkerungsbegriffe und Abgrenzungen im Mikrozensus	8
4.4 Plausibilitätsprüfung des aufbereiteten Mikrozensus Scientific Use Files 1976	8
Anhang A: Randverteilung ausgewählter Variablen des Scientific Use Files 1976	10
Anhang B: SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 1976	12
Anhang C: Weiterführende Literatur	14

1. Einleitung

In dem vorliegenden Bericht werden nach einer kurzen Beschreibung des Erhebungsprogramms, der Stichprobe und der Besonderheiten des Mikrozensus 1976 die einzelnen Schritte der Aufbereitung der faktisch anonymisierten Mikrozensusdaten (des Scientific Use Files) dargestellt.

Die Bereitstellung der faktisch anonymisierten Daten als Rohdatenfile in ASCII-Format erfolgt durch das Forschungsdatenzentrum (FDZ) des Landesamtes für Datenverarbeitung und Statistik Nordrhein-Westfalen (Düsseldorf). In Kooperation mit dem FDZ stellt das German Microdata Lab dem Nutzer ausführliche Informationen über das Scientific Use File zur Verfügung und bietet Setups zur Aufarbeitung des Datenmaterials und zur Erstellung von System-Files in Datenformaten gängiger Softwarepakete an. Ausführlichere Informationen über die verfügbaren Mikrozensus Scientific Use Files (vorhandene Variablen, verwendete Systematiken, Fragebogen, Schlüsselverzeichnis, intertemporale Vergleichbarkeit, Mikrodaten Tools etc.) werden darüber hinaus in ZUMA-Methodenberichten zum intertemporalen Vergleich von Mikrozensus und im World Wide Web des German Microdata Lab¹ bereitgestellt und können dort abgerufen werden.

Im Folgenden wird nach einer kurzen Beschreibung des Mikrozensus und des Mikrozensus Scientific Use Files 1976 der Vorgang der Datenaufbereitung beschrieben und die notwendigen Schritte zur Überprüfung der Plausibilität von eingelesenen Daten erläutert.

2. Mikrozensus 1976

2.1 Kurzbeschreibung

Der Mikrozensus ist mit einem Auswahlsatz von 1 Prozent der Bevölkerung die größte jährliche Haushaltsbefragung in Europa. Sie wird seit 1957 in Westdeutschland und seit 1991 in den neuen Bundesländern durchgeführt. Grundlage für die ausführlichen Angaben zu bevölkerungs- und arbeitsmarktstatistischen Strukturdaten des Mikrozensus ist die gesamte wohnberechtigte Bevölkerung, zu der im Mai 1976 rund 61,5 Millionen Menschen gehörten. Der Mikrozensus ist durch eine hohe Kontinuität im Frageprogramm und aufgrund der Auskunftspflicht durch einen niedrigen Unit-Nonresponse gekennzeichnet. Der Mikrozensus er-

¹ Vgl.: <http://www.gesis.org/Dauerbeobachtung/GML/Service/index.htm>

möglicht differenzierte Analysen auf Personen-, Familien- und Haushaltsebene sowohl im Quer- als auch im Längsschnitt und eignet sich für die Beobachtung von Strukturveränderungen sowie zu tiefgegliederten Analysen.

2.2 Datenerhebung

Die Angaben im Mikrozensus 1976 beziehen sich auf die Berichtswoche vom 2. bis 8. Mai 1976. Die Daten des Mikrozensus 1976 wurden durch die Statistischen Landesämter erhoben. Die Befragungen fanden i.d.R. mündlich statt, ein kleiner Teil erfolgte mittels schriftlicher Auskunft der Befragten. Für alle Beteiligten bestand gemäß der gesetzlichen Regelungen Auskunftspflicht. Nur wenige Fragen waren freiwillig zu beantworten. Der Mikrozensus ist eine Haushaltsstichprobe, d.h. alle Personen in einem Haushalt wurden entweder direkt befragt, oder eine erwachsene Person gab Auskunft über die weiteren (alle minderjährigen) Haushaltsmitglieder. Der Mikrozensus 1976 umfasste auch ausländische Haushalte. Zur Erhebung ihrer Angaben wurde der gleiche Fragebogen verwendet wie bei deutschen Haushalten. Die Daten wurden von den Statistischen Landesämtern aufbereitet und beim Statistischen Bundesamt zu einem Datensatz zusammengestellt.

2.3 Die Stichprobe

Für den Mikrozensus 1976 galt der Stichprobenplan von 1972 (Flächenstichprobe auf der Basis der Volkszählung von 1970). Das stichprobenmethodische Grundkonzept des Mikrozensus 1976 ist eine einstufige Klumpenstichprobe. Die bebaute Fläche der Bundesrepublik wurde auf der Grundlage der Wohnungszählung von 1970, die im Rahmen der Volkszählung stattfand, in etwa gleich große Flächen (Auswahlbezirke bzw. Segmente) aufgeteilt. Die Neubautätigkeit wurde dabei berücksichtigt. Ein Prozent dieser Auswahlbezirke gelangen in die Stichprobe. Innerhalb eines Auswahlbezirkes wurden alle Haushalte befragt. Grundlage für die Auswahlbezirke sind ganze Gebäude oder bei größeren Gebäuden Gebäudeteile. Die Gebäude der Auswahlbezirke wurden dabei nach der Zahl ihrer Wohnung in drei Größenklassen bzw. Schichten eingeteilt, zusätzlich wurden Gebäude mit vermuteten Gemeinschaftsunterkünften einer Sonderschicht zugeordnet. Ein Auswahlbezirk umfasste durchschnittlich 23 Wohnungen.

2.4 Das Erhebungsprogramm

Das Fragenprogramm des Mikrozensus besteht aus einem gleich bleibenden Grundprogramm, das sozioökonomische Grundinformationen liefert und die laufende Beobachtung des Ar-

beitsmarktes ermöglicht. Im Grundprogramm, das sich an alle Haushalte richtet, sind 1976 folgende Bereiche enthalten:

- Fragen zum Haushalt/Wohnung
- Fragen zu den Personen im Haushalt
- Staatsangehörigkeit
- Fragen zu Ausländern
- Schulbesuch
- allgemeine und berufliche Bildung
- Weiterbildung/Umschulung
- Erwerbsbeteiligung
- Jahr der Eheschließung
- Zahl und Alter der Kinder im HH und Familie
- Erste gegenwärtige Erwerbstätigkeit
- Zweite Erwerbstätigkeit
- Arbeitsuche/Arbeitsplatzwechsel
- Unterhalt, Einkommen
- Altersvorsorge
- Krankenversicherung
- Gesundheit

Im Scientific Use File des MZ76 sind die Fragen zur Gesundheit nicht enthalten.

3. Der faktisch anonymisierte Mikrozensus 1976 – Das Scientific Use File

Bei den vorliegenden Daten des Mikrozensus Scientific Use File 1976 handelt es sich um eine faktisch anonymisierte 70-Prozent-Substichprobe des Mikrozensus 1976. Für die faktisch anonymisierte Substichprobe gilt, dass sie als systematische Zufallsauswahl aus dem Originalmaterial des Statistischen Bundesamtes gezogen wurde. Das Datenmaterial wurde dazu nach Bundesland, Regierungsbezirk, Gemeindegrößenklasse, Anzahl der Personen in Privathaushalten, Auswahlbezirksnummer und laufende Nummer des Haushaltes im Auswahlbezirk sortiert. Durch diese Anordnung ist gewährleistet, dass die Substichprobe hinsichtlich dieser Merkmale nur geringe zufallsbedingte Abweichungen aufweist. Die Haushalte werden bundesweit aufsteigend neu durchnummeriert, dabei werden Anstaltspersonen wie Einpersonenhaushalte behandelt. Zur Stichprobenziehung diente die letzte (Einer-)Stelle der Haushaltsnummer. Alle Haushalte, deren Einerstelle einer von drei zufällig gezogenen einstelligen Zahl entsprach (3,5,9), gelangten nicht in die Stichprobe.

Weiterhin unterscheidet sich das Scientific Use File vom Originalfile des Mikrozensus dadurch, dass bestimmte Variablen im Scientific Use File, bedingt durch die Anonymisierung, in klassierter und vergrößerter Form verfügbar sind.

So sind im Scientific Use File die Regionalangaben Bundesland und eine vergrößerte Gemeindegrößenklasse (EF6) enthalten. Die folgende Abbildung zeigt die Variablen Bundesland und Gemeindegrößenklasse in der Gliederung, wie sie im Originaldatensatz enthalten ist.

EF1 Land der Bundesrepublik		EF6 Gemeindegrößenklasse (Original)	
Schleswig-Holstein	1	unter 2 000 Einwohner	1
Hamburg	2	2 000 bis unter 5 000 Einwohner	2
Niedersachsen	3	5 000 bis unter 10 000 Einwohner	3
Bremen	4	10 000 bis unter 20 000 Einwohner	4
Nordrhein-Westfalen	5	20 000 bis unter 50 000 Einwohner	5
Hessen	6	50 000 bis unter 100 000 Einwohner	6
Rheinland-Pfalz	7	100 000 bis unter 200 000 Einwohner	7
Baden-Württemberg	8	200 000 bis unter 500 000 Einwohner	8
Bayern	9	500 000 und mehr	9
Saarland	10		
Berlin (West)	11		

Aus Anonymisierungsgründen muss bei jedem Bundesland jede Größenklasse mindestens 400.000 Einwohner (auf die Gesamtbevölkerung hochgerechnet) beinhalten. Folglich waren entsprechende Zusammenfassungen notwendig, die durch eine Kombination der Variablen Bundesland (EF1) und Gemeindegrößenklasse (EF6) erreicht wurde. Die neu gebildete Variable EF6 ist der folgenden Übersicht zu entnehmen. Die Zusammenfassungen zur Anonymisierung sind in den Kategorien 10-18 der Variablen EF6 beschrieben. Tabelle 1 zeigt die Verteilung der Befragten in den einzelnen Bundesländern auf die verschiedenen Gemeindegrößenklassen der neu gebildeten Variablen EF6.

Auch bei den weiteren Variablen des Scientific Use Files werden je nach Bedarf Merkmalsvergrößerungen vorgenommen, so dass jede Ausprägung in den univariaten Verteilungen des Originalmaterials des Mikrozensus 1976 mindestens **5.000** Fälle in der Grundgesamtheit aufweist.

Die Ausprägungen der Variable Nationalität wurden so aggregiert, dass jede Nationalitätengruppe in der Grundgesamtheit des Originalmaterials wenigstens **50.000** Einwohner umfasst.

EF6: Gemeindegrößenklasse (im Scientific Use File 1976)

1	unter 2 000 Einwohner	
2	2 000 bis unter 5 000 Einwohner	
3	5 000 bis unter 10 000 Einwohner	
4	10 000 bis unter 20 000 Einwohner	
5	20 000 bis unter 50 000 Einwohner	
6	50 000 bis unter 100 000 Einwohner	
7	100 000 bis unter 200 000 Einwohner	
8	200 000 bis unter 500 000 Einwohner	
9	500 000 und mehr Einwohner	
10	ohne Aufteilung	$EF1 = 04$
11	unter 5 000 Einwohnern	$EF1 = 08 / EF6 = 1,2$
12	2 000 bis unter 10 000 Einwohnern	$EF1 = 01,05 / EF6 = 2,3$
13	5 000 bis unter 20 000 Einwohnern	$EF1 = 10 / EF6 = 3,4$
14	10 000 bis unter 100 000 Einwohnern	$EF1 = 07 / EF6 = 4,5,6$
15	20 000 bis unter 100 000 Einwohnern	$EF1 = 01 / EF6 = 5,6$
16	20 000 bis unter 200 000 Einwohnern	$EF1 = 06 / EF6 = 5,6,7$
17	20 000 bis unter 500 000 Einwohnern	$EF1 = 10 / EF6 = 5,6,8$
18	200 000 und mehr Einwohnern	$EF1 = 03 / EF6 = 8,9$

Tabelle 1: Bundesland x Gemeindegrößenklasse

		GEMEINDEGRÖßENKLASSEN																		
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	Gesam.
B U N D E S L A N D	1	3999	0	0	2877	0	0	0	3527	0	0	0	3646	0	0	4377	0	0	0	18426
	2	0	0	0	0	0	0	0	0	12264	0	0	0	0	0	0	0	0	0	12264
	3	4186	4295	7357	8503	11599	4156	5995	0	0	0	0	0	0	0	0	0	0	5833	51924
	4	0	0	0	0	0	0	0	0	0	5314	0	0	0	0	0	0	0	0	5314
	5	0	0	0	12379	26483	19029	13052	22391	24385	0	0	4874	0	0	0	0	0	0	122593
	6	10380	6636	5332	4138	0	0	0	3030	4330	0	0	0	0	0	0	5793	0	0	39639
	7	8369	4249	2815	0	0	0	4800	0	0	0	0	0	0	6260	0	0	0	0	26493
	8	0	0	11206	10661	11819	7904	3482	4156	4138	0	12109	0	0	0	0	0	0	0	65475
	9	16680	13212	10176	10782	5320	4706	3272	5370	9377	0	0	0	0	0	0	0	0	0	78895
	10	0	0	0	0	0	0	0	0	0	0	0	0	3291	0	0	0	4525	0	7816
	11	0	0	0	0	0	0	0	0	0	13952	0	0	0	0	0	0	0	0	13952
Ges	43614	28392	36886	49340	55221	35795	30601	38474	68446	5314	12109	8520	3291	6260	4377	5793	4525	5833	442791	

4. Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 1976 und Überprüfung der Plausibilität des erstellten Systemfiles

4.1 Einlesen des Rohdatenmaterials und Erstellung des Systemfiles

Die vom German Microdata Lab bereitgestellten Setups für das Mikrozensus Scientific Use File 1976 dienen zum Einlesen des Rohdatenmaterials und zum Erstellen von Systemfiles für die Programmpakete SPSS, SAS und STATA. Sie beinhalten Programmanweisungen zum Umwandeln von nicht-numerischen Variablen in numerische Variablen, zur Definition von fehlenden Werten sowie zum Versehen der Variablen und ihrer Ausprägungen mit entsprechenden Labeln.

Bei der Aufbereitung der Rohdaten wurden fehlende Werte möglichst in 0 umgewandelt. War dieser Wert schon vergeben, dann wurden die fehlenden Werte in 9 oder 8 (bzw. 99, 999, 98) umkodiert.

Um eine reibungsfreie Aufbereitung der Daten zu gewährleisten, empfiehlt es sich, die in den Setups für das jeweilige Programmpaket vorgesehenen Voreinstellungen zu berücksichtigen und je nach Leistung des verarbeitenden Systems die Setups schrittweise durchführen zu lassen. Sind die Daten eingelesen und ist das entsprechende Systemfile erstellt, können die Verteilungen ausgewählter Variablen des Scientific Use Files 1976 (vgl. Anhang A) zur Überprüfung des fehlerfreien Ablaufs der Datenaufbereitung herangezogen werden.

Als Referenz zur Plausibilitätsprüfung der Scientific Use Files dienen die in den Fachserien des Statistischen Bundesamtes veröffentlichten Ergebnisse des jeweiligen Mikrozensus. Für den Mikrozensus 1976 ist dies die Fachserie 1: Bevölkerung und Erwerbstätigkeit, Reihe 3.1 (Bevölkerung nach Alter und Familienstand), Reihe 4.1 (Stand und Entwicklung der Erwerbstätigkeit). Beim Vergleich der Ergebnisse des Scientific Use Files mit den entsprechenden Veröffentlichungen des Statistischen Bundesamtes ist darauf zu achten, dass die in diesen Reihen veröffentlichten Zahlen gewichtete, auf die Gesamtbevölkerung hochgerechnete Ergebnisse des Mikrozensus sind. Weiterhin beziehen sich die Zahlen auf nach bevölkerungs- oder erwerbstatistischen Konzepten definierte (Teil-) Populationen, die in der jeweiligen Reihe der Fachserien des Statistischen Bundesamtes erläutert werden.

Im Folgenden werden die Modalitäten der Gewichtung und Hochrechnung der Scientific Use Files sowie die für den Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes relevanten Bevölkerungsbegriffe und Abgrenzungen erläutert. Die konkreten Schritte und

Ergebnisse der Plausibilitätsprüfung der aufbereiteten Mikrozensus Scientific Use Files 1976 beim German Microdata Lab werden in Abschnitt 4.4 dargestellt.

4.2 Gewichtung und Hochrechnung zum Vergleich mit den Veröffentlichungen des Statistischen Bundesamtes

Im Scientific Use File 1976 liegen keine Gewichte zur Kompensation der systematischen Ausfälle und des zufallsbedingten Stichprobenfehlers vor.

Zur Anpassung der Daten an die Bevölkerungsfortschreibung wurde beim MZ76 das Verfahren des zufälligen Doppeln und Streichens von Personen vorgenommen. Sollen die Ergebnisse von Häufigkeiten des Scientific Use Files mit Veröffentlichungen des Statistischen Bundesamtes zum Mikrozensus 1976 verglichen werden, muss diese Variable der gedoppelten/gestrichenen Fälle (EF76) bei Auswertungen verwendet werden. Dies geschieht bei der Abgrenzung der verschiedenen Bevölkerungskonzepte, wie z.B. Wohnbevölkerung (siehe Abschnitt 4.3).

Zum Verfahren des Doppeln und Streichens: Den Ist-Zahlen der Stichprobe werden die Soll-Zahlen gegenübergestellt, die sich aus der Bevölkerungsfortschreibung des betreffenden Jahres - gegliedert nach den Merkmalen Regierungsbezirk, Deutsche-Ausländer, Geschlecht - ergeben. Durch das (zufällige) Doppeln und Streichen von Personen (bzw. aller Personen eines Haushaltes) des Mikrozensus wird eine Übereinstimmung mit den Soll-Zahlen der Bevölkerungsfortschreibung herbeigeführt. Dadurch wird zwar strenggenommen nur eine Anpassung an die Randverteilungen der verwendeten Variablen erreicht. Man geht dabei jedoch von der Annahme aus, dass auch für andere Untergliederungen Genauigkeitsverbesserungen erreicht werden.²

Um auf ein Prozent der Bevölkerung, also auf den vollen Auswahlsatz des Mikrozensus hochzurechnen, wird das Gewicht mit dem Kehrwert der Ziehungswahrscheinlichkeit für das Scientific Use File (100/70) multipliziert:

```
compute gewicht = 100/70.  
weight by gewicht.
```

Wie erwähnt, sind die in den Fachserien des Statistischen Bundesamtes veröffentlichten absoluten Zahlen gewichtete, auf die Gesamtbevölkerung der Bundesrepublik hochgerechnete Daten des Mikrozensus. Zur Herstellung der Vergleichbarkeit zwischen den absoluten Häufigkeiten des Scientific Use Files mit den Veröffentlichungen des Statistischen Bundesamtes

müssen somit die Ergebnisse eigener Berechnungen mit dem Scientific Use File noch zusätzlich mit 100 (Kehrwert der Ziehungswahrscheinlichkeit für den vollen Mikrozensus) multipliziert werden.

compute gewicht = 100*100/70.

weight by gewicht.

4.3 Bevölkerungsbegriffe und Abgrenzungen im Mikrozensus

Zum Vergleich der Ergebnisse des Scientific Use Files (SUF) mit den vom Statistischen Bundesamt veröffentlichten Zahlen sind die jeweiligen Bevölkerungs- und Erwerbskonzepte zu berücksichtigen. Um solche Vergleiche zu erleichtern, gibt die folgende Tabelle einen Überblick über die Bevölkerungsbegriffe und deren Abgrenzung. Die Variable EF76 (Doppelung/Streichung) muss dabei verwendet werden, um eine Anpassung an die Bevölkerungsfortschreibung zu erzielen.

Tabelle 2: Bevölkerungsbegriffe, Abgrenzungen und Gewichtungen im Mikrozensus SUF 1976

Bevölkerungsbegriff Subpopulation	Filter zur Abgrenzung der Subpopulation
Wohnbevölkerung	(EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0)
Bevölkerung in Privathaushalten	(EF58=1 or EF58=2 or EF58=3 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and EF60=1
Bevölkerung am Familienwohnsitz	(EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and EF61=1
Erwerbstätige	(EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and (EF65<=4 or EF65=10 or EF65=11)
Erwerbslose	(EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and (EF65=5 or EF65=6 or EF65=7)
Nichterwerbspersonen	(EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and (EF65=8 or EF65=9)

4.4 Plausibilitätsprüfung des aufbereiteten Mikrozensus Scientific Use Files 1976

Die Plausibilitätsprüfung des Mikrozensus Scientific Use Files 1976 erfolgte im German Microdata Lab durch den Vergleich des aufbereiteten Datensatzes mit den vom Statistischen Bundesamt veröffentlichten Tabellen in den entsprechenden Reihen der Fachserie A. Dabei wurde nach den vom Statistischen Bundesamt verwendeten Bevölkerungs- und Erwerbskonzepten gewichtet und selektiert. Tabelle 2 enthält die Liste der zur Plausibilitätsprüfung herangezogenen Variablen, Angaben zu notwen-

² Erläuterungen zu den Fehlerrechnungen zum Mikrozensus finden sich in den einzelnen Fachserien des Statistischen Bundesamtes.

digen Gewichtungen und Abgrenzungen sowie die entsprechenden Seitenzahlen der jeweiligen Datenreihe des Statistischen Bundesamtes.

Tabelle 3: Gewichtung, Abgrenzung und Vergleich mit den Fachserien des Statistischen Bundesamtes für ausgewählte Variablen des Mikrozensus Scientific Use File 1976 (siehe Anhang B)

Fachserie 1 / Reihe 4.1. 'Stand und Entwicklung der Erwerbstätigkeit'
<p>Wohnbevölkerung (siehe Tabellenprogramm: Tabelle V2 ; Vorbedingung) (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0)</p> <p>Tabelle 1 (S. 15) ef18 Geschlecht ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept</p> <p>Tabelle 20.1 (s. 65) ef55 Nettoeinkommen der Person (monatl.)</p> <p>Tabelle 25 (S. 78) ef1 Bundesland ef17 Staatsangehörigkeit (Dichotom)</p> <p>Tabelle 10 (S. 64) ef66 Altersgruppen ef21 Familienstand ef48 überwiegender Lebensunterhalt</p>
<p>Wohnbevölkerung, nur Erwerbstätige (siehe Tabellenprogramm: Tabelle E 2.1; Vorbedingung) (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and (Erwerb=1).</p> <p>Tabelle 1 (S. 30) ef39 Stellung im Beruf (recodiert) ef44 Tats. gel. Arbeitszeit (recodiert) ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept (Teilpopulation der Erwerbstätigen)</p>

Die Programmanweisungen für die durchgeführten Plausibilitätsprüfungen sind als SPSS-Syntax im Anhang B aufgeführt. Bedingt durch die Ziehung der 70%-Stichprobe ergeben sich Abweichungen zwischen den Häufigkeiten der Variablen des Scientific Use Files und den in Fachserien des Statistischen Bundesamtes veröffentlichten Zahlen. Die Mehrzahl der oben aufgeführten Variablen weichen i.d.R. nur in geringem Maße (meist 0 bis 1 Prozent, max. 5 Prozent) von den veröffentlichten Daten ab. Abweichungen über 5 Prozent wurden bei Variablen festgestellt, deren Merkmalsausprägungen mit sehr kleinen Fallzahlen besetzt sind (z.B. Staatsangehörigkeit und Wirtschaftsbereiche).

Die Überprüfung der Angaben auf Personenebene mit den Angaben in den Bandsatzergänzungen in Bezug auf die Variablen Wirtschaftszweig des Haushaltsvorstands und Wirtschaftszweig der Familienbezugsperson ergab, dass bei insgesamt zehn Fällen die Angaben zum Wirtschaftszweig (ef41) und die Angaben in der Bandsatzergänzung Wirtschaftszweig des Haushaltsvorstandes (ef100) sowie Wirtschaftszweig des Familienvorstandes (ef122) voneinander abweichen.

Anhang A: Randverteilung ausgewählter Variablen des Scientific Use Files 1976

Bevölkerungskonzept	Wohnbevölkerung	Bevölkerung in Privathaushalten	Bevölkerung am Familienwohnsitz	Erwerbstätige Bevölkerung	Fallzahlen
Merkmal	<i>select if</i> (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0)	<i>select if</i> (EF58=1 or EF58=2 or EF58=3 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and EF60=1	<i>select if</i> (ef61=1)	<i>select if</i> (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) (EF65<=4 or EF65=10 or EF65=11)	<i>ohne Selektion und ohne Anpassung</i>
Insgesamt	615304	610923	603311	257490	442 791
EF17: Staatsangehörigkeit					
Deutscher/Deutschland+Ausland	576814	573829	566393	236924	415594
Algerien/Marokko/Portugal/Tunesien	1754	1593	1580	1004	1235
Belgien/Dänemark/Frankr./Luxemb.	730	727	720	380	517
Griechenland	4067	4044	4037	2300	2854
GB/Nordirl./Irl./Norw./Polen/Schwed	709	691	691	369	511
Italien	5594	5491	5486	2990	3944
Jugoslawien	5901	5566	5537	3846	4177
Niederlande	1057	1057	1043	431	747
Österreich	1536	1520	1510	826	1099
Schweiz/Tschechoslowakei/Ungarn	679	674	664	340	483
Spanien	2136	2066	2061	1196	1513
Türkei	11080	10623	10589	5426	7781
Vereinigte Staaten von Amerika	596	587	581	221	427
Übriges Ausland	1859	1669	1647	827	1334
Staatenlos	793	786	771	410	575
EF18: Geschlecht					
Männlich	293077	291074	286626	161724	213037
Weiblich	322227	319849	316686	95766	229754
EF20: Alter (Gruppirt)					
76 und älter	30156	28329	28320	799	21272
70 - u. 75	27049	26564	26539	1616	19094
65 - u. 70	33617	33259	33226	3303	23741
60 - u. 65	33567	33290	33204	9883	23944
55 - u. 60	27133	26921	26837	15127	19476
50 - u. 55	38006	37761	37600	24903	27412
45 - u. 50	40003	39757	39580	28647	29022
40 - u. 45	40326	40049	39861	29574	29195
35 - u. 40	49829	49437	49167	36676	35973
30 - u. 35	38834	38487	38184	28174	28086
25 - u. 30	39711	39714	38773	28186	29289
20 - u. 25	40419	41671	37989	28607	30699
15 - u. 20	47303	46680	45276	21996	34292
unter 15	129353	129003	128756	-	91296
EF21: Familienstand					
Ledig	238329	235930	230007	63597	172755
Verheiratet	310119	309891	308244	178776	222662
Verwitwet	53131	51710	51699	6616	37516
Geschieden	13726	13391	13361	8501	9858

EF39: Stellung im Beruf					
Selbst. ohne Beschäftigte	11723	11761	11694	11666	10386
Selbst. mit Beschäftigten	11726	11783	11693	11683	8693
Mithelf. Familienang.	11994	12013	11989	11976	9104
Beamter, Richter, Soldat	16771	17109	16681	16730	12349
Angestellter	87930	88116	86790	85646	63832
Arbeiter/Heimarbeiter	109263	108104	107256	103999	77875
Kaufm.-techn. Auszubild.	4791	4661	4433	4757	3540
Gewerbl. Auszubildender	5739	5723	5544	5690	4155
Wehrpflichtiger	2334	2334	2316	2334	1634
Berufssoldat	3010	2079	2061	3010	2119
Entfällt	350023	347240	342854	-	193687
EF48: Überw. Lebensunterhalt					
Erwerbstätigkeit	232046	231384	228553	232046	169969
Arbeitslosengeld/-Hilfe	5820	5824	5797	50	4141
Rente, Pension	99113	96959	96710	3534	70089
Unterhalt durch Eltern etc.	258946	260650	256496	15684	184576
Eigenes Vermögen	3430	3393	3381	474	2452
Sozialhilfe	10606	8300	7997	357	7799
Soldat	5344	4413	4377	5344	3765
EF55: Nettoeinkommen Person					
Angabe fehlt	16290	14667	14370	8117	11847
Unter 150 DM	6997	6686	6584	1334	5014
150 bis unter 300 DM	20894	20516	20151	10294	14966
300 bis unter 450 DM	21114	20956	20424	10740	15219
450 bis unter 600 DM	27004	26634	25787	10756	19455
600 bis unter 800 DM	38090	37457	36977	16970	27235
800 bis unter 1000 DM	44671	43823	43396	25646	32205
1000 bis unter 1200 DM	49851	48940	48477	36434	36360
1200 bis unter 1400 DM	43200	42453	42146	35131	31374
1400 bis unter 1600 DM	29337	29156	28900	24493	21342
1600 bis unter 1800 DM	17626	17576	17399	14926	12810
1800 bis unter 2000 DM	13557	13557	13423	11564	9849
2000 bis unter 2200 DM	9569	9589	9494	8249	6972
2200 bis unter 2500 DM	8496	8550	8439	7544	6217
2500 bis unter 3000 DM	7857	7921	7814	7123	5789
3000 und mehr	11171	11301	11099	10337	8326
Mith. Fam./Selbst. Landwirt	17831	17854	17821	17831	12804
Kein Einkommen	231747	233287	230610	-	165007
EF57: Personentyp					
Bis 7 Jahre ohne Schulbesuch	45717	45680	45639		32235
Schüler bis 14 Jahre	83636	83323	83117		59061
Schüler/Studenten über 14 Jahre	30910	32320	28509		23825
Erwerbstätige	257490	255893	252724	257490	188135
Sonstige	197551	193707	193323	-	139535
EF78: Allgemeiner Schulabschluß					
Angabe fehlt	17109	16007	15826	5621	12214
Volks(Hauptschul)abschluß	333821	330027	328261	180124	239569
Realschulabschluß (Mittlere Reife)	66187	65819	64449	41616	48179
Fachhochschreife	6316	6423	6100	4030	4723
Abitur (Hochschulreife)	41253	42854	39639	23007	31473
EF79: Beruflicher Ausbildungsab.					
Angabe fehlt	39463	39054	38204	13223	28417
Keinen beruflichen Ausbildungsab.	171894	169373	167607	71313	123315
Lehr-/Anlernausbildung	195586	194261	192571	128169	141026
Berufliches Praktikum	8229	8201	7976	5297	6048
Meister-/Techniker/Fachschulab.	21763	21549	21293	16419	15877
Fachhochschulabschluss	8053	8117	7937	6117	5954
Hochschulabschluss	13101	13199	12873	10517	9957

Anhang B: SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 1976

*+++ Plausibilitätsprüfung: Vergleich mit den Fachserien des Statistischen Bundesamtes

*+++ Mikrozensus 1976

*+++ Plausi_FS_mz76.sps

*+++++

missing value all ().

*** Vorarbeiten: Recodierungen einzelner Variablen

***** Erwerbskonzept**

Recode EF65 (1 thro 4=1) (10 thro 11=1) (5 thro 7=2) (else= 3) into Erwerb.

Var Lab Erwerb 'Erwerbsstatus'.

Val Label Erwerb 1 'Erwerbstätige'

2 'Erwerbslose'

3 'Nichterwerbspersonen'.

***** Tatsächlich geleistete Stunden**

Recode EF44 (0=0) (1 thro 9=1) (10 thro 20=2) (21 thro 30=3) (31 thro 35= 4)

(36 thro 39=5) (40=6) (41=7) (42 thro 44=8) (45 thro 47=9) (48 thro 54=10)

(55 thro 69=11) (else=12) into Stunden.

Var Lab Stunden 'tats. gel. Arbeitsstunden'.

Val Label Stunden 1 '1-9 Stunden'

2 '10-20 Stunden'

3 '21-30 Stunden'

4 '31-35 Stunden'

5 '36-39 Stunden'

6 '40 Stunden'

7 '41 Stunden'

8 '42-44 Stunden'

9 '45-47 Stunden'

10 '48-54 Stunden'

11 '55-69 Stunden'

12 '70 unde mehr Stunden'.

***** Stellung im Beruf**

Recode EF39 (0 thro 1=1) (2 thro 2=2) (3 thro 9=3) into Beruf.

Var Lab Beruf 'Stellung im Beruf'.

Val Label Beruf 1 'Selbständige'

2 'Mithelfende'

3 'Abhängige'.

***** Alter (gruppiert) / Tabelle 19 (S.64)**

Recode EF66 (65 thro 94=1) (60 thro 64=2) (55 thro 59=3) (45 thro 54=4) (35 thro 44=5)

(25 thro 34=6) (20 thro 24=7) (15 thro 19=8) (0 thro 14=9) into Alter.

Var Lab Alter 'Altersgruppen'.

Val Label Alter 1 '65 und mehr'

2 '60 b.u. 65 Jahre'

3 '55 b.u. 60 Jahre'

4 '45 b.u. 55 Jahre'

5 '35 b.u. 45 Jahre'

6 '25 b.u. 35 Jahre'

7 '20 b.u. 25 Jahre'

8 '15 b.u. 20 Jahre'

9 'unter 15 Jahren'.

***** Nettoeinkommen der Person / Tabelle 20.1 (S.65)**

Recode ef55 (0=0) (1 thro 2=1) (3 thro 4=2) (5=3) (6=4)(7=5) (8=6)

(9 thro 10=7) (11 thro 13=8) (14 thro 15=9) into Einko.

Var lab Einko 'Nettoeinkommensgruppen'.

Val lab Einko 0 'keine Angabe'

1 'unter 300 DM'
 2 '300 b.u. 600 DM'
 3 '600 b.u. 800 DM'
 4 '800 b.u. 1000 DM'
 5 '1000 b.u. 1200 DM'
 6 '1200 b.u. 1400 DM'
 7 '14000 b.u. 1800 DM'
 8 '1800 b.u. 2500 DM'
 9 '2500 und mehr'.

*** **Staatsangehörigkeit (dichotom)**

Recode EF17 (1=1) (2 thro 40=2) (50=3) into Nation.

Var Lab Nation 'Staatsangehörigkeit (dichotom)'.

Val Lab Nation 1 'Deutsche'

2 'Ausländer'

3 'Staatenlos'.

*** **Wirtschaftsbereiche**

Recode EF41 (1 thro 5=1) (7 thro 11=2) (12 thro 44=2) (45 thro 47=2) (48 thro 56=3) (57 thro 63=3)
 (64 thro 65=4) (66 thro 85=4) (86 thro 90=4) (91 thro 95=4) (else=5) into Wirt.

Var Lab Wirt 'Wirtschaftsbereiche'.

Val Label Wirt 1 'Landwirtschaft etc.'

2 'Produzierendes Gewerbe'

3 'Verkehr/Handl'

4 'sonstige Dienstl.'

5 'ohne Angabe'.

* +++ **Wohnbevölkerung** (nur Erwerbstätige)
 * +++ (siehe Tabellenprogramm: Tabelle E 2.1; Vorbedingung)
 * +++ FS 1, Reihe 4.1 'Stand und Entwicklung der Erwerbstätigkeit'
 * +++
 * +++ Tabelle 1 (S. 30) ef39 Stellung im Beruf (recodiert)
 * +++ ef44 Tats. gel. Arbeitszeit (recodiert)
 * +++ ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept
 * +++ (Teilpopulation der Erwerbstätigen)
 * +++ ef41 Wirtschaftszweig (recodiert)

Compute gewicht=100/70.

weight by gewicht.

temporary.

select if (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) and (Erwerb=1).

fre Stunden Beruf Wirt.

* +++ **Wohnbevölkerung**
 * +++ (siehe Tabellenprogramm: Tabelle V2 ; Vorbedingung)
 * +++
 * +++ FS 1, Reihe 4.1 'Stand und Entwicklung der Erwerbstätigkeit'
 * +++ Tabelle 1 (S. 15) ef18 Geschlecht
 * +++ ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept
 * +++ Tabelle 20.1 (s. 65) ef55 Nettoeinkommen der Person (monatl.)
 * +++ Tabelle 25 (S. 78) ef1 Bundesland
 * +++ ef17 Staatsangehörigkeit (Dichotom)
 * +++ Tabelle 10 (S. 64) ef66 Altersgruppen
 * +++ ef21 Familienstand
 * +++ ef48 überwiegender Lebensunterhalt

Compute gewicht=100/70.

weight by gewicht.

temporary.

select if (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0).

fre Alter ef21 ef48 Erwerb ef65 ef18 Einko Nation ef1.

Anhang C: Weiterführende Literatur

Informationen zum Mikrozensus und seiner Konzeption:

- Emmerling, Dieter/Riede, Thomas (1997): 40 Jahre Mikrozensus. *Wirtschaft und Statistik* (3):160-174.
- Hartmann, P. (1989): Der Mikrozensus als Datenquelle für die Sozialwissenschaften. ZUMA-Nachrichten Nr. 24, Mai 1989: 6-25.
- Heidenreich, H.-J.(1994): Hochrechnung des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis*. Westdeutscher Verlag, Opladen 1994, S. 112-123.
- Lüttinger, P./Riede, T.(1997): Der Mikrozensus: amtliche Daten für die Sozialforschung, ZUMA-Nachrichten Nr. 41, November 1997: 19-43.
- Meyer, Kurt (1994): Zum Auswahlplan des Mikrozensus ab 1990. In: Gabler, S./Hoffmeyer-Zlotnik, J./Krebs, D. (Hrsg.): *Gewichtung in der Umfragepraxis*. Westdeutscher Verlag, Opladen 1994, S. 106-111.
- Statistisches Bundesamt (Hrsg.) (2004): *Arbeitsunterlagen zum Mikrozensus. Das Erhebungsprogramm des Mikrozensus seit 1957*. Wiesbaden, November 2004.
- Statistisches Bundesamt (Hrsg.)(1989): *Inhaltliche Fragen bevölkerungsstatistischer Stichproben am Beispiel des Mikrozensus. Bericht zur Konferenz vom 21. und 22. Oktober 1988. Heft 10 der Schriftenreihe Ausgewählte Arbeitsunterlagen zur Bundesstatistik*. Wiesbaden 1989.
- Schmidt, S. (2000): *Erwerbstätigkeit im Mikrozensus. Konzept, Definition, Umsetzung*. ZUMA-Arbeitsbericht 2000/01.

Zur Weitergabe faktisch anonymisierter Einzeldaten:

- Müller, W./Blien, U./Knoche, P./Wirth, H. (1991): *Die Faktische Anonymität von Mikrodaten*, (Band 19 der Schriftenreihe Forum der Bundesstatistik). Metzler-Poeschel, Stuttgart 1991.
- Wirth, H. (1992): *Die faktische Anonymität von Mikrodaten: Ergebnisse und Konsequenzen eines Forschungsprojektes*. ZUMA-Nachrichten Nr. 30, Mai 1992.

Zur Fehlerrechnung im Mikrozensus:

- Statistisches Bundesamt (1976): *Fachserie 1, Reihe 4.1, Stand und Entwicklung der Erwerbstätigkeit*. Stuttgart/Mainz: Kohlhammer.
- Krug, W./Nourney, M./Schmidt, J. (1994): *Wirtschafts- und Sozialstatistik. Gewinnung von Daten.*, München: Oldenbourg (3. völlig neubearb. Auflage)

Ergebnisse des Mikrozensus 1976:

- Statistisches Bundesamt (1976): *Fachserie 1, Reihe 3.1, Bevölkerung nach Alter und Familienstand*. Stuttgart/Mainz: Kohlhammer.
- Statistisches Bundesamt (1976): *Fachserie 1, Reihe 4.1, Stand und Entwicklung der Erwerbstätigkeit*. Stuttgart/Mainz: Kohlhammer.

Darüber hinaus sind Veröffentlichungen in der Zeitschrift 'Wirtschaft und Statistik' des Statistischen Bundesamtes zu finden.

Ein umfassendes Informationsangebot bietet das German Microdata Lab, ZUMA an:

<http://www.gesis.org/Dauerbeobachtung/GML/>

Publikationen, die auf den Scientific-Use-Files der Mikrozensen basieren finden sich auf der Webseite

http://www.gesis.org/Dauerbeobachtung/GML/Service/MZ-SUF-Literatur/MZ_Literatur.htm