

Mikrozensus Scientific Use File 1982: 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 1982: Dokumentation und Datenaufbereitung*. (ZUMA-Methodenbericht, 2006/03). Mannheim: Zentrum für Umfragen, Methoden und Analysen -ZUMA-. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-262442>

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 2006/03

Mikrozensus Scientific Use File 1982
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 1982	1
2.1 Kurzbeschreibung	1
2.2 Datenerhebung	2
2.3 Die Stichprobe	2
2.4 Das Erhebungsprogramm	2
3. Der faktisch anonymisierte Mikrozensus 1982 – Das Scientific Use File	3
4. Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 1982 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 1982.....	8
Anhang A: Randverteilung ausgewählter Variablen des Scientific Use Files 1982	11
Anhang B: SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 1982.....	13
Anhang C: Weiterführende Literatur	15

1. Einleitung

Im vorliegenden Bericht werden nach einer kurzen Beschreibung des Erhebungsprogramms, der Stichprobe und der Besonderheiten des Mikrozensus 1982 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 1982 der Vorgang der Datenaufbereitung beschrieben und die notwendigen Schritte zur Überprüfung der Plausibilität von eingelesenen Daten erläutert.

2. Mikrozensus 1982

2.1 Kurzbeschreibung

Der Mikrozensus ist mit einem Auswahlatz von ein 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 April 1982 rund 61,6 Millionen Menschen gehörten. Der Mikrozensus ist durch eine hohe Kontinuität im Frageprogramm und aufgrund der Auskunftspflicht durch einen niedrigen Unit-Nonresponse gekennzeichnet. Der Mikrozensus ermöglicht differenzierte Analysen auf Personen-, Familien- und Haushaltsebene sowohl im

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

Quer- als auch im Längsschnitt und eignet sich für die Beobachtung von Strukturveränderungen sowie zu tiefgegliederten Analysen.

2.2 Datenerhebung

Die Angaben im Mikrozensus 1982 beziehen sich auf die Berichtswoche vom 19. bis 25. April 1982. Die Daten des Mikrozensus 1982 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 1982 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 zusammengespielt.

2.3 Die Stichprobe

Für den Mikrozensus 1982 galt der Stichprobenplan von 1972 (Flächenstichprobe auf der Basis der Volkszählung von 1970). Das stichprobenmethodische Grundkonzept des Mikrozensus 1982 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 Arbeitsmarktes ermöglicht. Im Grundprogramm, das sich an alle Haushalte richtet, sind 1982 folgende Bereiche enthalten:

- Fragen zum Haushalt/Wohnung
- Fragen zu den Personen im Haushalt
- Staatsangehörigkeit
- Vertriebenen/Flüchtlingsstatus
- Schulbesuch
- 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
- Urlaubs- und Erholungsreisen

Im Scientific Use File des MZ82 sind die Fragen zur Gesundheit und zu den Urlaubs- und Erholungsreisen nicht enthalten.

3. Der faktisch anonymisierte Mikrozensus 1982 – Das Scientific Use File

Bei den vorliegenden Daten des Mikrozensus Scientific Use File 1982 handelt es sich um eine faktisch anonymisierte 70-Prozent-Substichprobe des Mikrozensus 1982. 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 1982 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: Gemeindegroßenklasse (im Scientific Use File 1982)

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	<i>EF1 = 04</i>
11	unter 5 000 Einwohnern	<i>EF1 = 06, 08 / EF6 = 1,2</i>
12	2 000 bis unter 10 000 Einwohnern	<i>EF1 = 01,05 / EF6 = 1,2,3</i>
13	5 000 bis unter 20 000 Einwohnern	<i>EF1 = 10 / EF6 = 3,4</i>
14	10 000 bis unter 50 000 Einwohnern	<i>EF1 = 07 / EF6 = 4,5</i>
15	20 000 bis unter 100 000 Einwohnern	<i>EF1 = 01 / EF6 = 5,6</i>
16	20 000 bis unter 200 000 Einwohnern	<i>EF1 = 10 / EF6 = 5,6,7</i>
17	100 000 bis unter 500 000 Einwohnern	<i>EF1 = 09 / EF6 = 7,8</i>
18	200 000 und mehr Einwohnern	<i>EF1 = 03, 06 / EF6 = 8,9</i>

Tabelle 1: Bundesland x Gemeindegroßenklasse

		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	3872	0	0	2972	0	0	0	3089	0	0	0	4209	0	0	4674	0	0	0	18816
	2	0	0	0	0	0	0	0	0	11653	0	0	0	0	0	0	0	0	0	11653
	3	4171	4850	6982	8858	11471	4195	5814	0	0	0	0	0	0	0	0	0	0	5734	52075
	4	0	0	0	0	0	0	0	0	0	4855	0	0	0	0	0	0	0	0	4855
	5	0	0	0	12404	27955	18030	12525	22155	23919	0	0	4551	0	0	0	0	0	0	121539
	6	0	0	6311	10276	6714	3484	3365	0	0	0	3524	0	0	0	0	0	0	6545	40219
	7	7932	4556	3161	0	0	2875	3274	0	0	0	0	0	0	4812	0	0	0	0	26610
	8	0	0	10814	12132	12631	5574	4568	4062	4293	0	12524	0	0	0	0	0	0	0	66598
	9	8411	14153	11369	12887	8064	6831	0	0	9284	0	0	0	0	0	0	0	8602	0	79601
	10	0	0	0	0	0	0	0	0	0	0	0	0	3391	0	0	4218	0	0	7609
	11	0	0	0	0	0	0	0	0	13579	0	0	0	0	0	0	0	0	0	13579
Ges		24386	23559	38637	59529	66835	40989	29546	29306	62728	4855	16048	8760	3391	4812	4674	4218	8602	12279	443154

4. Aufbereitung des Rohdatenfiles zum Mikrozensus Scientific Use File 1982 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 1982 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 1982 (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 1982 ist dies die Fachserie 1: Bevölkerung und Erwerbstätigkeit, Reihe 4.1.1 (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 1982 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 1982 liegen keine Gewichte zur Kompensation der systematischen Ausfälle und des zufallsbedingten Stichprobenfehlers vor.

Zur Anpassung der Daten an die Bevölkerungsfortschreibung wurde beim MZ82 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 1982 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 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 zusätz-

lich 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 1982

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)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) (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) (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) (EF65=8 or EF65=9)

4.4 Plausibilitätsprüfung des aufbereiteten Mikrozensus Scientific Use Files 1982

Die Plausibilitätsprüfung des Mikrozensus Scientific Use Files 1982 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 he-

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

rangezogenen Variablen, Angaben zu notwendigen 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 1982 (siehe Anhang B)

Fachserie 1 / Reihe 4.1.1 'Stand und Entwicklung der Erwerbstätigkeit'
<p>Wohnbevölkerung: (EF58=1 or EF58=2 or EF58=4) and (EF76=1 or EF76=3 or EF76=0) and (EF38=1 or EF38=0) nach Beteiligung am Erwerbsleben, Familienstand, überwiegendem Lebensunterhalt und Altersgruppe</p> <p>Tabelle 1 (S.35) ef66 Altersgruppen ef21 Familienstand ef48 überwiegender Lebensunterhalt</p> <p>Tabelle 1 (S.20) EF18 Geschlecht EF65 Bevölkerungstyp</p> <p>Tabelle 4.1 (S.46) EF55 Nettoeinkommen (Person)</p> <p>Tabelle 25 (S. 78) EF1 Bundesland EF17 Staatsangehörigkeit (dichotom)</p>
<p>Wohnbevölkerung, nur Erwerbstätige 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). (siehe Tabellenprogramm: Tabelle E 2.1; Vorbedingung) FS 1, Reihe 4.1.1 'Stand und Entwicklung der Erwerbstätigkeit' Erwerbstätige nach Wirtschaftsbereichen, Stellung im Beruf und tatsächlich geleisteter Arbeitszeit</p> <p>Tabelle 15 (S. 70) 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 Prozent-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 7 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 1982

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	616 704	614123	606 379	268 290	443 154
EF17: Staatsangehörigkeit					
Deutscher/Deutschland+Ausland	570404	568507	561029	247219	410439
Algerien/Marokko/Portugal/Tunesien	1823	1771	1766	849	1282
Belgien/Dänemark/Frankr./Luxemb.	986	989	973	450	708
Griechenland	3087	3054	3046	1556	2180
GB/Nordirl./Irl./Norw./Polen/Schwed	1380	1357	1347	643	981
Italien	6181	6117	6097	3223	4350
Jugoslawien	6369	6259	6200	3636	4525
Niederlande	1054	1053	1047	554	748
Österreich	1379	1374	1353	821	985
Schweiz/Tschechoslowakei/Ungarn	721	714	704	356	514
Spanien	1801	1779	1767	1014	1268
Türkei	17220	17043	17000	6299	12106
Vereinigte Staaten von Amerika	584	579	576	274	414
Übriges Ausland	3057	2914	2866	1151	2188
Staatenlos	657	613	609	246	466
EF18: Geschlecht					
Männlich	295144	294097	289623	166357	213814
Weiblich	321560	320026	316756	101933	229340
EF20: Alter (Gruppiert)					
76 und älter	37486	35404	35384	763	26440
70 - u. 75	29133	28754	28714	1176	20554
65 - u. 70	28147	27926	27897	1801	19881
60 - u. 65	29881	29690	29604	7291	21264
55 - u. 60	36979	36799	36674	20461	26520
50 - u. 55	38781	38599	38420	26549	28011
45 - u. 50	43073	42800	42611	31463	31053
40 - u. 45	50384	50133	49899	37581	36287
35 - u. 40	37376	37219	36971	28117	26829
30 - u. 35	41073	40907	40593	30244	29676
25 - u. 30	40564	40643	39509	28956	29964
20 - u. 25	46634	48467	44586	32870	35177
15 - u. 20	53980	53784	52727	21017	38751
unter 15	103213	102999	102789	-	72747
EF21: Familienstand					
Ledig	236494	235801	229727	75469	171560
Verheiratet	309203	309167	307547	176570	221328
Verwitwet	54229	52694	52679	5833	38252
Geschieden	16779	16460	16426	10419	12014

EF39: Stellung im Beruf					
Selbst. ohne Beschäftigte	10974	11003	10939	10911	9811
Selbst. mit Beschäftigten	12404	12474	12379	12354	9133
Mithelf. Familienang.	8243	8240	8230	8233	6205
Beamter, Richter, Soldat	18054	18311	17959	17970	13210
Angestellter	99217	99447	98226	96640	71640
Arbeiter/Heimarbeiter	113667	113183	112417	102813	80850
Kaufm.-techn. Auszubild.	5941	5970	5731	5914	4344
Gewerbl. Auszubildender	8206	8307	8043	8130	5950
Wehrpflichtiger	2243	2243	2234	2243	1570
Berufssoldat	3081	2243	2236	3081	2166
Entfällt	334673	332701	327986	-	238275
EF48: Überw. Lebensunterhalt					
Erwerbstätigkeit	242834	243124	240566	242834	176961
Arbeitslosengeld/-Hilfe	8809	8776	8730	73	6226
Rente, Pension	105593	103404	103230	2526	74557
Unterhalt durch Eltern etc.	238587	240541	236251	16729	170027
Eigenes Vermögen	2953	2870	2856	284	2096
Sozialhilfe	12604	10921	10276	520	9535
Soldat	5324	4486	4470	5324	3752
EF55: Nettoeinkommen Person					
Angabe fehlt	16014	14060	13864	8764	11608
Unter 300 DM	19034	18244	17844	5493	13676
300 bis unter 450 DM	21184	21263	20764	12364	15264
450 bis unter 600 DM	20281	20494	19680	10401	14727
600 bis unter 800 DM	26281	26184	25446	11040	18886
800 bis unter 1000 DM	29951	29776	29479	13156	21385
1000 bis unter 1200 DM	34174	34023	33729	17669	24500
1200 bis unter 1400 DM	34940	34757	34466	22190	25198
1400 bis unter 1600 DM	38764	38499	38221	27807	28059
1600 bis unter 1800 DM	34124	33749	33581	26317	24615
1800 bis unter 2000 DM	30137	29914	29726	23673	21813
2000 bis unter 2200 DM	23424	23291	23130	18644	16891
2200 bis unter 2500 DM	20751	20710	20537	17110	15036
2500 bis unter 3000 DM	17977	17923	17777	15023	13055
3000 bis unter 3500 DM	11574	11431	11303	9790	8415
3500 bis unter 4000 DM	6059	6004	5914	5240	4440
4000 bis unter 4500 DM	3726	3763	3691	3313	2767
4500 bis unter 5000 DM	2171	2186	2143	1989	1613
5000 und mehr	5661	5746	5606	5244	4275
Mith. Fam./Selbst. Landwirt	13063	13061	13049	13063	9371
Kein Einkommen	207410	209044	206429	-	147560
EF57: Personentyp					
Entfällt	3	3	3	1	2
Bis 7 Jahre ohne Schulbesuch	39230	39217	39180	-	27605
Schüler bis 14 Jahre	63983	63781	63609	-	45142
Schüler/Studenten über 14 Jahre	41767	43763	39609	-	31698
Erwerbstätige	268289	267726	264810	268289	195152
Sonstige	203433	199633	199169	-	143555
EF78: Allgemeiner Schulabschluß					
Angabe fehlt	8201	7116	7060	2516	5831
Volks(Hauptschul)abschluß	346387	342780	341309	180251	247429
Realschulabschluß (Mittlere Reife)	81471	81737	80467	52479	58910
Fachhochschreife	9450	9583	9196	6369	7028
Abitur (Hochschulreife)	41490	43589	39673	25501	32237
EF79: Beruflicher Ausbildungsab.					
Angabe fehlt	11850	10827	10713	3609	8459
Keinen beruflichen Ausbildungsab.	170861	168901	166901	66063	122340
Lehr-/Anlernausbildung	222966	222654	221040	142179	160005
Berufliches Praktikum	7361	7287	7127	4539	5376
Meister-/Techniker/Fachschulab.	23723	23803	23524	17620	17238
Fachhochschulabschluss	9547	9620	9434	7194	7017
Hochschulabschluss	19703	19824	19364	15873	14852

Anhang B: SPSS-Syntax für die Plausibilitätsprüfungen des Scientific Use Files 1982

```
*+++ Plausibilitätsprüfung: Vergleich mit den Fachserien des Statistischen Bundesamtes  
*+++ Mikrozensus 1982  
*+++ Plausi_FS_MZ82.sps
```

```
missing value all ().
```

```
*** Fachserie 1: Reihe 4.1.1 'Stand und Entwicklung der Erwerbstätigkeit, Mikrozensus 1982  
*** 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 1 (S.35)
```

```
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 4.1 (S.47)
```

```
Recode ef55 (0=0) (1 thro 3=1) (4 thro 5=2) (6=3) (7=4)(8 thro 9=5) (10 thro 11=6)  
(12 thro 13=7) (14 thro 15=8) (16 thro 18=9) into Einko.  
Var lab Einko 'Nettoeinkommensgruppen'.
```



```

Val lab Einko 0 'keine Angabe'
              1 'unter 600 DM'
              2 '600 b.u. 1000 DM'
              3 '1000 b.u. 1200 DM'
              4 '1200 b.u. 1400 DM'
              5 '1400 b.u. 1800 DM'
              6 '1800 b.u. 2200 DM'
              7 '2200 b.u. 3000 DM'
              8 '3000 b.u. 4000 DM'
              9 '4000 und mehr'.

```

*** Staatsangehörigkeit (dichotom)

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

Var Lab Nation 'Staatsangehörigkeit (dichotom)'.
 Val Label 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**

*+++ FS 1 / Reihe 4.1.1 'Stand und Entwicklung der Erwerbstätigkeit (S. 35 / Tabelle 1)

*+++ Wohnbevölkerung nach Beteiligung am Erwerbsleben, Familienstand,
 *+++ überwiegendem Lebenunterhalt und Altersgruppe 1982

*+++

*+++ Tabelle 1 (S. 35) ef66 Altersgruppen

*+++ ef21 Familienstand

*+++ ef48 überwiegender Lebensunterhalt

*+++ Tabelle 1 (S.20) ef18 Geschlecht

*+++ ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept

*+++ Tabelle 4.1 (S. 46) ef55 monatliches Nettoeinkommen (Person) =Einko

*+++ Tabelle 25 (S. 78) ef1 Bundesland

*+++ ef17 Staatsangehörigkeit (dichotom) = Nation

*+++++

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 ef18 ef65 Einko Nation ef1.

*+++++

*+++ **ERWERBSTÄTIGE**

*+++ (siehe Tabellenprogramm: Tabelle E 2.1; Vorbedingung)

*+++ FS 1, Reihe 4.1.1 'Stand und Entwicklung der Erwerbstätigkeit'

*+++ Erwerbstätige nach Wirtschaftsbereichen, Stellung im Beruf und tatsächlich geleisteter Arbeitszeit

*+++

*+++ Tabelle 15 (S. 70) ef39 Stellung im Beruf (recodiert)

*+++ ef44 Tats. gel. Arbeitszeit (recodiert)

*+++ ef 65 Bevölkerungstyp nach Erwerbs- und Unterhaltskonzept (Teilpopulation der Erwerbstätigen)

*+++ ef41 Wirtschaftszweig

*+++++

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.

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 (1982): *Fachserie 1, Bevölkerung und Erwerbstätigkeit. Reihe 4.1.1, 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 1982:

- Statistisches Bundesamt (1982): *Fachserie 1, Bevölkerung und Erwerbstätigkeit. Reihe 4.1.1, 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