

## SOEP Survey Papers

Series B - Survey Reports (Methodenberichte)

SOEP – The German Socio-Economic Panel study at DIW Berlin

2018

# Methodenbericht zum Befragungsjahr 2016 des IAB-SOEP- Migrationssamples (M1 / M2)

Anne Bohlender, Simon Huber, Axel Glemser (Kantar Public)

Running since 1984, the German Socio-Economic Panel study (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing. The SOEP Survey Papers is comprised of the following series:

**Series A** – Survey Instruments (Erhebungsinstrumente)

**Series B** – Survey Reports (Methodenberichte)

**Series C** – Data Documentation (Datendokumentationen)

**Series D** – Variable Descriptions and Coding

**Series E** – SOEPmonitors

**Series F** – SOEP Newsletters

**Series G** – General Issues and Teaching Materials

The SOEP Survey Papers are available at <http://www.diw.de/soepsurveypapers>

#### **Editors:**

Dr. Jan Goebel, DIW Berlin

Prof. Dr. Stefan Liebig, DIW Berlin and Universität Bielefeld

Dr. David Richter, DIW Berlin

Prof. Dr. Carsten Schröder, DIW Berlin and Freie Universität Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Please cite this paper as follows:

Anne Bohlender, Simon Huber, Axel Glemser (Kantar Public). 2018. Methodenbericht zum Befragungsjahr 2016 des IAB-SOEP-Migrationssamples (M1 /M2). SOEP Survey Papers 578: Series B. Berlin: DIW/SOEP



This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License.

© 2018 by SOEP

ISSN: 2193-5580 (online)

DIW Berlin

German Socio-Economic Panel (SOEP)

Mohrenstr. 58

10117 Berlin

Germany

soeppapers@diw.de

# Methodenbericht zum Befragungsjahr 2016 des IAB-SOEP-Migrationssamples (M1 /M2)

Anne Bohlender, Simon Huber, Axel Glemser (Kantar Public)

# Inhaltsverzeichnis

<b>Einordnung in den SOEP-Kontext</b>	<b>3</b>
<b>1 Erhebungsdesign und Spezifika der Feldarbeit</b>	<b>4</b>
<b>2 Erhebungsinstrumente</b>	<b>5</b>
2.1 Fragebogen	5
2.2 Einsatz von Übersetzungshilfen	6
2.3 Einverständniserklärung zur Registerdatenverknüpfung (Record Linkage)	7
<b>3 Ergebnisse der Feldarbeit</b>	<b>8</b>
3.1 Haushaltsebene	8
3.2 Teilnahme auf Personenebene	11
<b>4 Übergebener Datenbestand</b>	<b>14</b>

# Einordnung in den SOEP-Kontext

Die adäquate Repräsentation von Zuwandererhaushalten stellt seit jeher eine zentrale Zielsetzung im SOEP-Stichprobendesign dar. Bereits bei der ersten Befragungswelle 1984 wurde die Teilstichprobe B etabliert. Diese bildete Personen und Haushalte aus den fünf Herkunftsländern ab, aus denen am häufigsten die sogenannten „Gastarbeiter“ während der 1960er und 1970er Jahre nach Deutschland zuzogen (Türkei, ehemaliges Jugoslawien, Griechenland, Italien, Spanien). Teilstichprobe D (1994/1995) wurde zur Abbildung der nach 1984 innerhalb Deutschlands auftretenden Migrationsdynamiken gegründet.

Im Erhebungsjahr 2013 wurde mit dem Migrationssample  $M_1$  eine weitere Aufwuchsstichprobe implementiert, die ausschließlich aus Haushalten besteht, in denen mindestens eine Person mit Migrationshintergrund lebt. Auf Basis eines vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) bereitgestellten Adressenpools wurden 2.723 Haushalte von Personen mit Migrationshintergrund mit einem modifizierten Erstwellendesign von Leben in Deutschland befragt. Die Zielpopulation bestand aus zwischen 1995 und 2010 hauptsächlich aus EU-Staaten zugewanderten Personen sowie aus Kindern von Einwanderern (sogenannte zweite Generation von Zuwanderern). In 2015 erfolgte die vierte Erhebungswelle in dieser Stichprobe.

Aufgrund der wissenschaftlichen und politischen Relevanz des Themas Migration wurde im Erhebungsjahr 2015 eine erneute Erweiterung des SOEP-Stichprobensystems implementiert. Mit der 1.096 Haushalte umfassenden Migrationsstichprobe  $M_2$  sollen jüngere Zuwanderungsströme aus EU-Staaten von Anfang 2009 bis Ende 2013 abgebildet werden.

Der Ausbau des SOEP-Haushaltsdatenbestandes zur Repräsentation der in Deutschland lebenden Migrationsbevölkerung beruht auf einer wissenschaftlichen und vertraglichen Kooperation zwischen dem IAB Nürnberg und der Abteilung SOEP am DIW. Diese Kooperation ermöglichte es, auf die Integrierte Erwerbsbiografien-Datenbank (IEBS) der Bundesagentur für Arbeit als Sampleframe zurückzugreifen.

Insgesamt nahmen in den beiden Samples 2.153 der 3.262 kontaktierten Haushalte an der Studie teil. Die Nettostichprobe bestand aus 1.990 Teilnehmern des Vorjahres, 109 temporären Ausfällen und 54 neuen Haushalte (sogenannte Split-off Haushalte). Die Panelstabilität stieg in der vierten Erhebungswelle von  $M_1$  noch einmal spürbar von 82,9 % auf 89,7 % an. In  $M_2$  wurde in der zweiten Welle, die den erfahrungsgemäß schwierigen Übergang von der Querschnittsbefragung in den Längsschnitt bildet, eine Panelstabilität von 60,2 % erreicht. Weitere Ergebnisse der Feldarbeit sind in Abschnitt 3 dargestellt.

# 1 Erhebungsdesign und Spezifika der Feldarbeit

Das Sampling der Migrationsbefragung in den Stichproben  $M_1$  und  $M_2$  in der ersten Welle war registerbasiert. Das hatte zur Folge, dass das übliche SOEP Konzept, bei der der Haushalt die primäre Stichprobeneinheit darstellt, nicht in der Erstwelle angewandt werden konnte. Stattdessen bildeten die Ankerpersonen, die aus der Datenbank der integrierten Erwerbsbiographien (IEBS) ausgewählt wurden, die primäre Stichprobeneinheit. Seit der zweiten Welle wird in beiden Stichproben das originäre SOEP-Haushaltskonzept mit Haushalten als primäre Stichprobeneinheit angewandt. Aus diesem Grund werden Haushalte auch dann in die Bearbeitung aufgenommen, wenn die ursprüngliche Ankerperson aktuell den Haushalt verlassen hatte oder nicht mehr teilnahmebereit war.

Die Befragung wurde ausschließlich persönlich-mündlich mit dem Erhebungsmodus CAPI unter Verwendung fast aller auch in den sonstigen SOEP-Stichproben eingesetzten Erhebungsinstrumente durchgeführt. Die Kernfragebogen sowie Anschreiben, Broschüre und Datenschutzblatt wurden in fünf Sprachen übersetzt: Englisch, Russisch, Türkisch, Rumänisch und Polnisch. Auch im Jahr 2016 wurden in beiden Stichproben Einverständniserklärungen für die Verknüpfung von Registerdaten mit den Befragungsdaten gesammelt. Zielgruppe waren neben neuen Befragungspersonen auch solche, die in der ersten Welle des Samples  $M_2$  der Verknüpfung nicht zugestimmt hatten (siehe Abschnitt 2.3).

Die Feldzeit erstreckte sich von April bis August 2016 und fiel damit etwas kürzer als im Vorjahr aus (**Tabelle 1.1**). Rund drei Viertel der Haushalte konnten in den ersten drei Monaten final bearbeitet werden. 82,5 % der Nettostichprobe wurde in diesem Zeitraum realisiert. Die verbleibenden beiden Monate wurden fast ausschließlich auf sogenannte „schwierige“ Haushalte verwendet. Dabei handelt es sich beispielsweise um verzogene Haushalte, deren Adresse recherchiert werden muss, neu gebildete Haushalte sowie bisher nicht erreichbare Haushalte und Personen, teilweise in Verbindung mit Krankheit, arbeitsbedingter Abwesenheit, Auslandsaufenthalt etc.

Tabelle 1.1:

**Zeitlicher Verlauf der Feldarbeit: Final bearbeitete Haushalte<sup>1</sup>**

	In % der Bruttostichprobe	In % der Nettostichprobe
April	24,1	27,5
Mai	52,0	60,4
Juni	72,7	82,5
Juli	90,3	94,9
August	100,0	100,0

<sup>1</sup> Angegeben sind kumulierte Prozentwerte basierend auf dem Monat des letzten Haushaltskontakts.

# 2 Erhebungsinstrumente

## 2.1 Fragebogen

Zur Datenerhebung wurden folgende Instrumente aus dem sog. SOEP-Hauptsample 2016 (Stichproben A – L) - meist weitestgehend unverändert - verwendet:

- (1) **Adressen- und Kontaktprotokoll**
- (2) **HAUSHALTS-Fragebogen**
- (3) **PERSONEN-Fragebogen** für jedes Haushaltsmitglied, das 1998 oder früher geboren ist.
- (4) **JUGEND-Fragebogen** für alle Personen im Haushalt, die 1999 geboren sind.
- (5) „**Lust auf DJ**“ ist ein Zusatzfragebogen (Kompetenzmessung) für alle Personen, die den Jugendfragebogen ausgefüllt haben.
- (6) „**Frühe Jugend**“ Fragebogen für Personen im Haushalt, die 2002 geboren sind.
- (7) „**Fragebogen für Schülerinnen und Schüler**“ für Personen im Haushalt, die 2004 geboren sind und erstmals persönlich an der Befragung teilnehmen.
- (8) „**Mutter und Kind A**“ für Mütter von Kindern, die in 2016 geboren sind (bzw. in 2015, aber nach der letzten Befragung).
- (9) „**Ihr Kind im Alter von 2 oder 3 Jahren**“ (Mutter und Kind B) für Mütter von Kindern, die in 2013 geboren sind.
- (10) „**Ihr Kind im Alter von 5 bis 6 Jahren**“ (Mutter und Kind C) für Mütter von Kindern, die in 2010 geboren sind.
- (11) „**Ihr Kind im Alter von 7 bis 8 Jahren**“ (Elternfragebogen D) für Mütter und Väter von Kindern, die 2008 geboren sind.
- (12) „**Ihr Kind im Alter von 9 bis 10 Jahren**“ (Mutter und Kind E) für Mütter von Kindern, die in 2006 geboren sind.
- (13) „**Personenbefragung 2015 Nacherhebung Kurzfassung**“ („LÜCKE“-Fragebogen) für temporäre Ausfälle des Vorjahres um die „Lücke“ in den Längsschnittdaten zu minimieren.
- (14) Zusatzfragebogen „**Die verstorbene Person**“ für Panelteilnehmer, die seit der letzten Befragung einen Todesfall in ihrem Haushalt oder ihrer Familie zu beklagen hatten
- (15) **Einverständniserklärung** zur Verknüpfung der Befragungsdaten mit Registerdaten des IAB.

Statt des in der SOEP-Hauptwelle eingesetzten LEBENSLAUF-Fragebogens für Neubefragte, erhielten neue Haushaltsmitglieder einen speziell für die Stichproben M<sub>1</sub> und M<sub>2</sub> entwickelten PERSONENBIO-Fragebogen, der neben Fragen zur Biografie und Fragen aus dem Personenfragebogen unter anderem einen Abschnitt zur Migrationshistorie enthält.

Zudem wurde erneut das CAMCES-Modul, ein Instrument zur computergestützten Messung und Kodierung von Bildungsabschlüssen in Umfragen, eingesetzt (CAMCES = Computer-Assisted Measurement and Coding of Educational Qualifications in Surveys). Das von der GESIS entwickelte und dem SOEP zur Verfügung

gestellte Instrument greift auf eine Datenbank internationaler Bildungsabschlüsse zurück und ermöglicht gegenüber einer ungestützten oder kategorialen Abfrage von Bildungsabschlüssen die adäquate Erfassung und Kodierung von (Aus-)Bildungsabschlüssen. In besonderem Maße gilt dies für im Ausland erworbene Qualifikationen, da die Bildungssysteme je nach Land unterschiedlich ausdifferenziert sind, oftmals nur die ausländischen Bezeichnungen bekannt sind oder es zum Teil keine deutschen Entsprechungen für bestimmte Abschlüsse gibt. Die Abfrage erfolgt anhand einer Freitexteingabe in der jeweiligen Landessprache, die einen Such-Algorithmus durchläuft und mit in einer Datenbank hinterlegten Qualifikationsabschlüssen abgeglichen wird. Darauf wird eine Auswahlliste mit den wahrscheinlichsten Abschlüssen angezeigt und die passende Option durch die Zielperson ausgewählt. Die jeweiligen Abschlüsse werden somit automatisch vercodet.

## 2.2 Einsatz von Übersetzungshilfen

Da die Zielpopulation (überwiegend) ausländischer Herkunft ist, wurden die Kernfragebogen (Haushalt, Personen und Personenbio) in fünf Sprachen übersetzt: Englisch, Russisch, Türkisch, Rumänisch und Polnisch. Diese Sprachen – mit Ausnahme von Englisch – entsprechen den Nationalitäten, die im Bruttosample der Erstwelle von M<sub>1</sub> überrepräsentiert waren. Die übersetzten Versionen wurden nicht in CAPI integriert, sondern auf Papier gedruckt und dem Interviewer zur Verfügung gestellt, um Sprachbarrieren zu überwinden. Dieser Rückgriff auf sog. Übersetzungshilfen entspricht dem Vorgehen der SOEP-Hauptwelle. **Tabelle 2.1** zeigt verschiedene Maßnahmen, die von den Interviewern in sprachlich schwierigen Situationen ergriffen wurden.

Tabelle 2.1:

### Sprachprobleme und Verwendung von Übersetzungshilfen

	Anzahl	In % der Nettostichprobe
<b>Nettostichprobe (Personenfragebogen)</b>	3.832	100,0
<b>Keine Sprachprobleme</b>	3.160	82,5
<b>Unterstützung bei Sprachproblemen benötigt</b>	672	17,5
<i>Davon:</i>		
<b>Deutsch sprechende Person im Haushalt</b>	236	6,2
<b>Deutsch sprechende Person außerhalb des Haushalts</b>	75	2,0
<b>Professioneller Dolmetscher</b>	10	0,3
<b>Übersetzungshilfe</b>	351	9,2
<i>Davon:</i>		
<b>Russisch</b>	113	2,9
<b>Türkisch</b>	55	1,4
<b>Rumänisch</b>	63	1,6
<b>Polnisch</b>	51	1,3
<b>Englisch</b>	69	1,8



Bei einem Großteil der Interviews (82,5 %) sind entweder keine Sprachprobleme aufgetreten oder zumindest nicht in dem Maße, dass Unterstützung herangezogen wurde. In 17,5 % der Personeninterviews wurde Unterstützung bei Sprachproblemen benötigt, wobei diese in 6,2 % der Fälle durch deutschsprachige Personen geleistet wurde, die innerhalb oder außerhalb des Haushalts leben. Bei 9,2 % der befragten Personen wurden Übersetzungshilfen eingesetzt. Dabei die russische Übersetzungshilfe mit am häufigsten verwendet (n=113). Die Übersetzungshilfen in türkischer, rumänischer und polnischer Sprache wurden jeweils ca. halb so oft gebraucht.

## 2.3 Einverständniserklärung zur Registerdatenverknüpfung (Record Linkage)

Eine Besonderheit des Migrationssamples ist das Verknüpfen der Befragungsdaten der Teilnehmer mit Registerdaten der Stichprobe für integrierte Erwerbsbiographien (IEBS). Diese Art des Einsatzes von „Big Data“ erweitert das analytische Potential der M-Samples, indem die Bildungs- und Erwerbsbiographien der Befragten mit den zwei unterschiedlichen Datentypen – Befragtenangaben einerseits und Registerdaten andererseits – möglichst vollständig abgebildet werden.

In 2016 wurden im Sample M<sub>1</sub> nur noch diejenigen Personen um ihr schriftliches Einverständnis zum Record Linkage gebeten, die erstmals an der Befragung teilgenommen haben. Im Sample M<sub>2</sub> kamen zur Gruppe der neuen Befragungspersonen noch diejenigen Teilnehmer hinzu, die in der ersten Welle 2015 eine Einverständniserklärung abgelehnt hatten. Die für das Record Linkage vorgesehene Zielpopulation bestand 2016 über beide Samples hinweg aus 620 Befragungspersonen. Von diesen gaben 48,5 % ihre Einwilligung für die Verknüpfung der Daten (**Tabelle 2.2**).

Seit 2013 wurden in den Samples M<sub>1</sub> und M<sub>2</sub> insgesamt 6.096 Teilnehmer maximal zweimal um ihr Einverständnis zur Registerdatenverknüpfung gebeten. 3.798 Personen waren bereit, diese zu erteilen. So liegt nach vier Erhebungswellen in M<sub>1</sub> und zwei Erhebungswellen in M<sub>2</sub> für 62,3 % der Personen, die für das Record Linkage vorgesehen waren, eine Zustimmungserklärung vor.

Tabelle 2.2:

**Einverständnis zum Record Linkage: Response Rates**

	2013	2014	2015	2015	2016	<b>Gesamt</b>	
	Ausgewählt für RL in W1	M <sub>1</sub>	Nicht Zugestimmt in W1 oder 2 + neue Personen	M <sub>2</sub>	M <sub>1</sub> +M <sub>2</sub>		
		Ausgewählt für RL in W2 + nicht zugestimmt in W1	Alle Befragungspersonen in W1	Nicht Zugestimmt in W 1 (M <sub>2</sub> ) + neue Personen			
	In %	In %	In %	In %	In %	Abs.	In %
Zugestimmt	48,9	44,2	48,8	51,8	48,5	3.798	62,3
Abgelehnt	51,1	55,8	51,2	48,2	51,5	2.298	37,7
<b>Gesamt</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>100,0</b>	<b>6.096</b>	<b>100,0</b>

# 3 Ergebnisse der Feldarbeit

## 3.1 Haushaltsebene

Die **Tabelle 3.1** gibt die Resultate der Feldarbeit in der vierten Welle von  $M_1$  bzw. der zweiten Welle von  $M_2$  wieder. Alles in allem wurden 3.262 Haushalte bearbeitet. Das Bruttosample bestand zu 84,7 % aus Haushalten, die im Vorjahr teilgenommen hatten. Temporäre Ausfälle in 2015 und Neu-Haushalte bildeten einen Anteil von 11,3 % bzw. 4,0 % der Bruttostichprobe. Da in der zweiten Welle üblicherweise keine temporären Ausfälle der Vorwelle in die Befragung aufgenommen werden, war der Anteil der Vorjahresteilnehmer am Gesamtbrutto im Sample  $M_2$  mit 98,0 % deutlich höher als im Sample  $M_1$  (77,8 %).

Insgesamt wurden 2.153 Haushalte realisiert, davon 1.493 in Sample  $M_1$  und 660 in Sample  $M_2$ . Die Response Rate unter den Teilnehmern des Vorjahres lag im Sample  $M_1$ , dessen Haushalte z.T. bereits zum vierten Mal am SOEP teilnahmen, bei 80,3 %. Im Sample  $M_2$  konnten 59,4 % der Haushalte, die im Vorjahr teilgenommen hatten, realisiert werden. Dieser Unterschied spiegelt eine typische Entwicklung bei der Etablierung einer neuen Panelstichprobe wider, deren Teilnahmequoten zunächst jährlich steigen, bis sie sich nach einigen Wellen bei einem mehr oder weniger stabilen Niveau einpendeln. Auch im Sample  $M_1$  scheint dieser Prozess noch nicht gänzlich abgeschlossen zu sein, weshalb die Panelstabilität im noch einmal spürbar von 82,9 % im Jahr 2015 auf 89,7 % im Jahr 2016 gesteigert werden konnte (**Abbildung 3.1**). Die Panelstabilität im Sample  $M_2$  lag in der zweiten Welle bei 60,2 %. Vorläufige Ergebnisse des Jahres 2017 weisen jedoch darauf hin, dass in diesem Sample in der dritten Welle eine deutliche Steigerung der Panelstabilität auf ca. 85 % erreicht werden konnte.

Für die Entwicklung der Fallzahlen im SOEP ist nicht nur der Blick auf die Vorjahresteilnehmer, sondern auch auf die temporären Ausfälle des Vorjahrs und die Neu-Haushalte wichtig, da beide Gruppen einen Beitrag zur Stabilisierung des Panels leisten. Im Sample  $M_1$  konnten 29,5 % der Ausfälle aus dem Vorjahr, die für eine erneute Bearbeitung in 2016 vorgesehen waren, realisiert werden. Sie bilden einen Anteil von 7,3 % der Nettostichprobe. In der Gruppe der Neu-Haushalte lag die Response Rate bei 42,1 % (Anteil an der Nettostichprobe 2016: 3,0 %).

Tabelle 3.1:

**Zusammensetzung der Brutto- und Nettostichprobe und Response Rates**

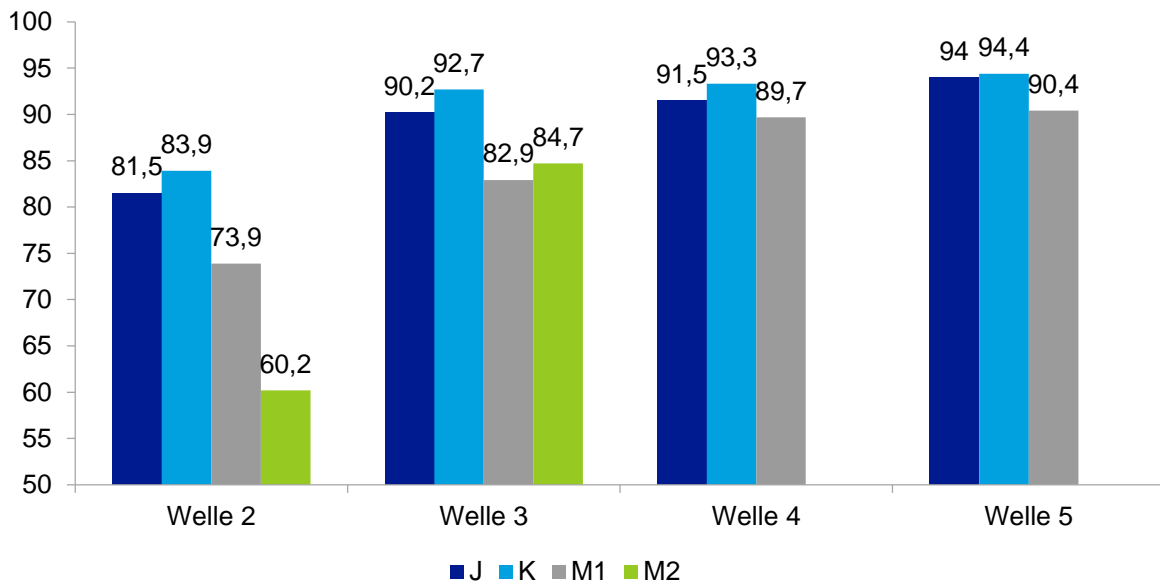
	Sample M <sub>1</sub>		Sample M <sub>2</sub>		Total	
	Abs.	In %	Abs.	In %	Abs.	In %
<b>(1) Zusammensetzung der Bruttostichprobe nach HH-Typ</b>	<b>2.144</b>	<b>100,0</b>	<b>1.118</b>	<b>100,0</b>	<b>3.262</b>	<b>100,0</b>
- Vorjahresteilnehmer	1.668	77,8	1.096	98,0	2.764	84,7
- Temporäre Ausfälle der Vorwelle	369	17,2	0	0,0	369	11,3
- Neue Haushalte (split-off Haushalte)	107	5,0	22	2,0	129	4,0
<b>(2) Zusammensetzung der Nettostichprobe nach HH-Typ</b>	<b>1.493</b>	<b>100,0</b>	<b>660</b>	<b>100,0</b>	<b>2.153</b>	<b>100,0</b>
- Vorjahresteilnehmer	1.339	89,7	651	98,6	1.990	92,4
- Temporäre Ausfälle der Vorwelle	109	7,3	-	-	109	5,1
- Neue Haushalte (split-off Haushalte)	45	3,0	9	1,4	54	2,5
<b>(3) Response Rate nach HH-Typ</b>		<b>69,6</b>		<b>59,0</b>		<b>66,0</b>
- Vorjahresteilnehmer		80,3		59,4		72,0
- Temporäre Ausfälle der Vorwelle		29,5		-		29,5
- Neue Haushalte		42,1		40,9		41,9
<b>(4) Panelstabilität<sup>1</sup></b>		<b>89,7</b>		<b>60,2</b>		<b>7</b>
						<b>7,9</b>
<b>(5) Partial Unit non-response<sup>2</sup></b>		<b>28,8</b>		<b>30,3</b>		<b>29,3</b>

<sup>1</sup> Panelstabilität: Zahl der diesjährigen Teilnehmerhaushalte in Relation zur entsprechenden Vorjahreszahl

<sup>2</sup> Partial Unit non-response: Anzahl teilweise realisierter Haushalte inklusive harter Verweigerer aus Vorwällen geteilt durch die Nettohaushalte mit mehr als einem Haushaltsmitglied im Befragungsalter.

Abbildung 3.1:

**Panelstabilitätsquoten in ausgewählten SOEP-Samples**



M<sub>1</sub> Welle 5 und M<sub>2</sub> Welle 3: Vorläufige Ergebnisse des Jahres 2017

Tabelle 3.2:

**Verteilung der Ausfälle**

	M <sub>1</sub>			M <sub>2</sub>		
	Anzahl	In % Brutto	In % Ausfälle	Anzahl	In % Brutto	In % Ausfälle
Bruttostichprobe	2.144	100,0	-	1.118	100,0	-
Interviews (Nettostichprobe)	1.493	69,6	-	660	59,0	-
<b>Kein Interview</b>	<b>651</b>	<b>30,4</b>	<b>100,0</b>	<b>458</b>	<b>41,0</b>	<b>100,0</b>
- Nicht angetroffen	118	5,5	18,1	121	10,8	26,4
- Ins Ausland verzogen	18	0,8	2,8	27	2,4	5,9
- Weiche Verweigerer (derzeit nicht bereit/nicht in der Lage)	162	7,6	24,9	169	15,1	36,9
- Endgültige Verweigerer	332	15,5	51,0	130	11,6	28,4
- Sprachprobleme	5	0,2	0,8	10	0,9	2,2
- Sonstiges (z.B. Pflegefall; Verstorben)	16	0,7	2,5	1	0,1	0,2

**Tabelle 3.2** enthält eine Übersicht der Verteilung der Ausfallgründe. Besonders in der zweiten Welle des Samples M<sub>2</sub> fällt auf, dass ein vergleichsweise hoher Anteil von Haushalten während der gesamten Feldzeit nicht erreicht werden konnte. Rund ein Viertel der Ausfälle in M<sub>2</sub> entfallen auf diesen Code (26,4 %). Zum Vergleich: In den zweiten Wellen der Samples J (2012) und K (2013) der SOEP-Hauptwelle lag der Anteil der Ausfälle aufgrund von fehlender Erreichbarkeit an allen Ausfällen bei 13,1 % bzw. 14,8 %. Dies bestätigt erneut die Erfahrungen aus den ersten Erhebungsjahren des Samples M<sub>1</sub>, wonach die Zielpopulation der Migrantenhaushalte schwerer zu kontaktieren ist, als dies bei den SOEP-Aufwuchsstichproben zum Bevölkerungsquerschnitt der Fall ist.

Neben Response Rates und Stabilitätsquoten stellt die Teilnahme auf Personenebene einen weiteren wichtigen Faktor zur Beurteilung der Ergebnisse der Feldarbeit dar. Zum einen wird hier die Teilnahmequote auf Personenebene herangezogen, die im Jahr 2016 über beide Teilstichproben hinweg bei 86,4 % lag (**Tabelle 3.4**). Zum anderen betrachtet man den Anteil von Haushalten, bei denen nicht von allen potentiellen Befragungspersonen ein Interview vorliegt (Partial Unit non-response). Dieser lag in den Samples M<sub>1</sub> und M<sub>2</sub> mit 29,3 % (**Tabelle 3.1**) 2016 auf einem vergleichsweise hohen Niveau im Vergleich zu den Stichproben der SOEP-Hauptwelle (**Tabelle 3.3**). Diese Ergebnisse auf Personenebene zeigen erneut, dass in den Migrationsstichproben nicht nur die Kontaktierung der Haushalte, sondern auch die Kontaktierung und Motivation aller im Haushalt lebenden Zielpersonen aufwändiger ist, als in den bevölkerungsrepräsentativen Stichproben des SOEP.

Tabelle 3.3:

### Anteil teilweise realisierter Haushalte am Gesamtnetto 2016 (in %)

	A-L1	M <sub>1</sub>	M <sub>2</sub>
<b>PUNR 1</b> (inklusive Einpersonenhaushalte und exklusive harter Verweigerer aus Vorwellen)	5,5	13,3	19,2
<b>PUNR 4</b> (Mehrpersonenhaushalte inklusive harter Verweigerer aus Vorwellen)	21,2	28,8	30,3

PUNR: Partial Unit non-response

## 3.2 Teilnahme auf Personenebene

### Kernfragebogen

Im Vergleich zu anderen Repräsentativbefragungen, bei denen in der Regel nur eine Person im Haushalt befragt wird, ist für das als Haushaltspanel angelegte SOEP die möglichst vollständige Teilnahme aller Haushaltsmitglieder wesentlich. Ein Haushalt gilt dann als realisiert, wenn der Haushaltsfragebogen und mindestens ein Personen- oder Jugend-Fragebogen auswertbar vorliegen. Ausgehend von dieser Mindestanforderung besteht die Zielsetzung der Feldarbeit darin, möglichst alle weiteren Haushaltsmitglieder im Befragungsalter zur Teilnahme zu gewinnen.

Mittlerweile gibt es im SOEP vier anhand des Alters bzw. des Geburtsjahres abgrenzbare Personengruppen im Befragungsalter: der Fragebogen für Schülerinnen und Schüler des Geburtsjahrganges 2004 stellt seit dem Jahr 2014 den Einstieg in die SOEP-Befragung dar und hat das Befragungsalter für die erste Befragung im

Zyklus eines SOEP-Teilnehmers auf 11 bis 12 Jahre gesenkt. Hinzu kommt seit 2016 der Fragebogen „Frühe Jugend“, der sich an junge Jugendliche im Alter von 13 bis 14 Jahre richtet (Jahrgang 2002). Die dritte Personengruppe sind im Haushalt lebende Jugendliche im Alter von 16 bis 17 Jahren (2016: Jahrgang 1999). Diese beantworten den Jugendfragebogen und den Kompetenztest „Lust auf DJ“. Jede weitere erwachsene Person im Haushalt beantwortet den Personenfragebogen bzw. den Personenbiofragebogen (2016: Jahrgang 1998 und älter).

Tabelle 3.4:

**Basiszahlen auf Personenebene für 2016: Stichprobe M<sub>1</sub> und M<sub>2</sub>**

	Alle Personen	Personen im Befragungsalter					Ges.	Nicht zu befragende Personen
		Jg. 1998 und älter	Jg. 1999	Jg. 2002	Jg. 2004			
<b>Personen im Bruttoband 2016</b>	9.511	6.650	114	109	164	7.037	2.474	
<b>davon:</b>								
<b>In teilnehmenden Haushalten<sup>1</sup></b>	6.482	4.482	74	74	107	4.737	1.745	
<b>Brutto für RR-Berechnung<sup>2</sup></b>	-	4.432	74	73	107	4.686	1.742	
<b>Fragebogen liegt vor<sup>3</sup></b>	-	3.830	49	64	93	4.036	-	
<b>Response Rate in %</b>	-	86,4	66,2	87,7	86,9	86,1	-	

<sup>1</sup> Haushaltsfragebogen liegt vor.

<sup>2</sup> Ohne Verstorbene und ins Ausland verzogene

<sup>3</sup> Zusätzlich haben 2 Personen einen Personenfragebogen und eine Person einen Jugendfragebogen ausgefüllt, es liegt jedoch kein Haushaltsinterview vor.

**Tabelle 3.4** gibt die Basiszahlen auf Personenebene für das Erhebungsjahr 2016 wieder. Der Gesamtbestand an Personen, der in der sog. Bruttodatei der Stichproben M<sub>1</sub> und M<sub>2</sub> dokumentiert ist, umfasst insgesamt 9.511 Personen. Darin enthalten sind alle Haushaltsmitglieder in allen Haushalten, die im Jahr 2016 für die Befragung vorgesehen waren (Bruttostichprobe). Insgesamt waren das 7.037 Personen. Die restlichen 2.474 Personen sind nicht in einem der oben angegebenen Jahrgänge geboren und sind damit nicht für die Befragung vorgesehen. Als Grundlage zur Berechnung der Teilnahmequoten auf Personenebene werden zunächst die teilnehmenden Haushalte herangezogen. Darüber hinaus sind verstorbene oder ins Ausland verzogene Haushaltsmitglieder vom ursprünglichen Bruttobestand abzuziehen. Es verbleiben 4.686 Personen im Befragungsalter, die die Basis für die Berechnung der Teilnahmequote (Response Rate) bilden.

Insgesamt haben 3.830 Befragte in teilnehmenden Haushalten den Personen- oder Personenbiofragebogen beantwortet. Die sich daraus ergebende Teilnahmequote liegt bei 86,4 % (SOEP-Hauptwelle 2016: 91,4 %). Die Teilnahmequote der Jugendlichen des Jahrgangs 1999 fällt mit 66,2 % deutlich geringer aus. Eine um mehrere Prozentpunkte geringere Response Rate des Jugendfragebogens im Vergleich zum Personenfragebogen kommt allerdings auch in der SOEP-Hauptwelle, wenn auch auf einem insgesamt etwas höherem Niveau, vor. Die gegenüber der SOEP-Kernstichprobe etwas niedrigeren Teilnahmequoten auf Personenebene spiegeln die bereits auf Haushaltsebene beschriebenen generellen Schwierigkeiten hinsichtlich Erreichbarkeit und Teilnahmebereitschaft der Haushalte und Zielpersonen wider.

Die Response Rate des 2016 erstmals eingesetzten Fragebogens „Frühe Jugend“ fällt erfreulicherweise deutlich höher aus als die des Jugendfragebogens. 87,7 % der 13-14-Jährigen gaben ein Interview. Die Teilnahmequote beim Fragebogen für Schülerinnen und Schüler liegt mit 86,9 % auf einem ähnlichen Niveau. Alles in allem liegen 4.036 Fragebogen in teilnehmenden Haushalten vor. Dies entspricht einer Gesamt-Response-Rate auf Personenebene von 86,1 %.

### Zusatzfragebogen

In **Tabelle 3.5** sind die Bearbeitungsergebnisse zu den in 2016 eingesetzten Zusatzfragebogen aufgeführt. Bis auf den Kompetenztest für Jugendliche („Lust auf DJ“), der als Ergänzung zum Jugendfragebogen durch die Jugendlichen selbst bearbeitet wird, erheben die kindbezogenen Instrumente Informationen über Kinder und Jugendliche eines bestimmten Geburtsjahrgangs. Die Fragebogen werden – bis auf den Elternfragebogen, der möglichst von beiden Elternteilen vorliegen sollte – von einem Elternteil, in der Regel der Mutter, bearbeitet. Bei Betrachtung der Deckungsrate, die angibt, für wie viele der in teilnehmenden Haushalten lebenden Kinder der entsprechenden Altersgruppe ein Fragebogen vorliegt, zeigt sich eine erfreulich hohe Teilnahmebereitschaft für die Zusatzfragebogen von 88,0 % (MuKi A) bis 92,0 % (MuKi C).

Tabelle 3.5:

#### Anzahl und Deckungsraten der Zusatz Fragebogen<sup>1</sup>

	Referenzgröße	Anzahl Interviews	Deckungsrate
Kognitiver Kompetenztest („Lust auf DJ“) <sup>2</sup>	49	45	91,8
„Mutter und Kind“ (MuKi A)	125	110	88,0
„Ihr Kind im Alter von 2 oder 3 Jahren“ (MuKi B)	129	116	89,9
„Ihr Kind im Alter von 5 oder 6 Jahren“ (MuKi C)	150	138	92,0
„Ihr Kind im Alter von 7 oder 8 Jahren“ (Eltern D) <sup>3</sup>	117/234	104/166	88,9/70,9
„Ihr Kind im Alter von 9 oder 10 Jahren“ (MuKi E)	119	108	90,8

<sup>1</sup> Die Zahlen beziehen sich auf die jeweilige Zielpopulation in teilnehmenden Haushalten. Für die kindbezogenen Instrumente ist die Referenzgröße die Anzahl der Kinder der entsprechenden Altersgruppe, die in teilnehmenden Haushalten leben. Die Deckungsrate für diese Fragebogen gibt an, für wie viele Kinder ein Fragebogen vorliegt.

<sup>2</sup> Der Test wird nur in Haushalten eingesetzt, die durch Interviewer bearbeitet werden. Zudem muss die Zielperson den Jugendfragebogen ausgefüllt haben.

<sup>3</sup> Im Gegensatz zu den anderen kindbezogenen Instrumenten soll der Elternfragebogen nicht nur von einem Elternteil (in der Regel die Mutter) sondern (wenn möglich bzw. vorhanden) von beiden Eltern ausgefüllt werden. Für 104 (88,9%) von 117 Kindern, die im Jahr 2008 geboren sind und in Haushalten leben, die 2016 teilgenommen haben, wurde mindestens ein Fragebogen ausgefüllt. Insgesamt waren es 166 Elternfragebogen.

# 4 Übergebener Datenbestand

Folgender Datenbestand der SOEP-Migrationssamples M<sub>1</sub> und M<sub>2</sub> 2016 wurde dem DIW am 5. Dezember 2016 übergeben:

## 1. Bruttodaten

<b>Bezeichnung</b>	<b>Anzahl Fälle</b>	<b>Dateiname</b>
Codebuch Paneldatei Stand Dez. 2016		Codebuch 2016.pdf
Bruttoband Haushalte	3.262	hbru_M1M2_2016.sav
Bruttoband Personen	9.511	pbru_M1M2_2016.sav
Interviewerband	221	Intband_M1M2_2016.sav
SOEP-Personen M <sub>1</sub> +M <sub>2</sub>	12.221	P_M1M2_2016.sav

## 2. Nettodaten

<b>Bezeichnung</b>	<b>Anzahl Fälle</b>	<b>Dateiname</b>
Nettodaten Haushalte geprüft	2.153	H16_M1M2.sav
Nettodaten Haushalte ungeprüft	2.153	H16_M1M2_u.sav
Nettodaten Personen geprüft	3.703	P16_M1M2.sav
Nettodaten Personen ungeprüft	3.703	P16_M1M2_u.sav
Nettodaten Kognitiver Test (Zeichen-Zahlen)	3.703	P16_M1M2_ZZ.sav
Nettodaten Personenbio geprüft	129	PB16_M1M2.sav
Nettodaten Personenbio ungeprüft	129	PB16_M1M2_u.sav
Nettodaten Jugend geprüft	50	J16_M1M2.sav
Nettodaten Jugend ungeprüft	50	J16_M1M2_u.sav
Nettodaten Lust auf DJ	45	DJ16_M1M2.sav
Nettodaten Mutter-und-Kind A	139	MA16_M1M2.sav
Nettodaten Mutter-und-Kind B	116	MB16_M1M2.sav
Nettodaten Mutter-und-Kind C	138	MC16_M12.sav
Nettodaten Elternfragebogen	166	EL16_M1M2.sav
Nettodaten Mutter-Kind E	108	ME16_M1M2.sav
Nettodaten Schüler	93	S16_M1M2.sav
Nettodaten Frühe Jugend	64	FJ16_M1M2.sav
Nettodaten „Die verstorbene Person“	41	VP16_M1M2.sav
Lückedaten aus 2015	192	LUE15_16_M1M2.sav
Einverständniserklärung zur Registerdaten- verknüpfung	3.882	EV16_M1M2_DIW.sav
Angaben zu 7 und mehr Kindern	5	7bis10Kind_2016_M1M2.sav



Berufe, Branchen, Abschlussvercodung  
 Hochschulvercodung  
 Variablenliste  
 Zusatzcodes für Haushalts-, Personen-,  
 Personenbio- und Jugenddaten: Nationen,  
 Sorgen, Bundesland, Sport

Berufe\_Branchen\_Ausb\_2016\_M1M2.sav  
 Hochschul-Abschl\_2016\_M1M2.sav  
 Struktur2016.xlsx  
 Zusatzcodes.xlsx

### 3. CAMCES-Modul 2016

CAMCES-Gesamtfile  
 gezippte Einzeldatensätze vom CAMCES-  
 Modul  
 Erklärung zur Benennung der  
 Einzeldatensätze und Schlüsselvariablen

659 P16\_M1M2\_CAMCES.sav  
 CAMCES\_Einzelfiles  
 READ ME.txt

### 4. Fragebogen

Codierter Haushaltsfragebogen  
 Codierter Personenfragebogen  
 Codierter Personenbiofragebogen  
 Codierter Jugendfragebogen  
 Codierter DJ-Fragebogen  
 Codierter Kind A Fragebogen  
 Codierter Kind B Fragebogen  
 Codierter Kind C Fragebogen  
 Codierter Eltern D Fragebogen  
 Codierter Kind E Fragebogen  
 Codierter Schülerfragebogen  
 Codierter Frühe Jugendfragebogen  
 Codierter Fragebogen zu „Verstorbene  
 Person“  
 Codierter Lücke-Fragebogen

Co\_M1+M2\_Haushalt 2016.pdf  
 Co\_M1+M2\_Personen\_2016.pdf  
 Co\_M1+M2\_PersBio\_2016.pdf  
 Co\_M1+M2\_Jugend 2016.pdf  
 Co\_LiD\_DJ 2016.pdf  
 Co\_LiD\_Kind A 2016.pdf  
 Co\_LiD\_Kind B 2016.pdf  
 Co\_LiD\_Kind C 2016.pdf  
 Co\_LiD\_Eltern D 2016.pdf  
 Co\_LiD\_Kind E 2016.pdf  
 Co\_LiD\_Schüler 2016.pdf  
 Co\_LiD\_Frühe Jugend 2016.pdf  
 Co\_LiD\_VP 2016.pdf  
 Co\_LiD\_Lücke 2015\_2016.pdf