

## SOEP Survey Papers

Series B - Survey Reports (Methodenberichte)

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

2016

# SOEP 2014 – Methodenbericht zum Befragungsjahr 2014 (Welle 2) des IAB-SOEP-Migrationssamples 2013 (M1)

Axel Glemser, Simon Huber (TNS Infratest Sozialforschung)

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. Martin Kroh, DIW Berlin and Humboldt Universität 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:

Axel Glemser, Simon Huber (TNS Infratest Sozialforschung). 2016. SOEP 2014 – Methodenbericht zum Befragungsjahr 2014 (Welle 2) des IAB-SOEP-Migrationsamples 2013 (M1). SOEP Survey Papers 301: Series B. Berlin: DIW/SOEP



This work is licensed under a [Creative Commons Attribution-ShareAlike 4.0 International License](https://creativecommons.org/licenses/by-sa/4.0/). © 2016 by SOEP

ISSN: 2193-5580 (online)

German Socio-Economic Panel (SOEP) | DIW Berlin  
Mohrenstr. 58  
10117 Berlin, Germany  
Contact: [soeppapers@diw.de](mailto:soeppapers@diw.de)

**AXEL GLEMSER, SIMON HUBER  
(TNS INFRATEST SOZIALFORSCHUNG)**

**SOEP 2014 – METHODENBERICHT ZUM  
BEFRAGUNGSJAHR 2014 (WELLE 2) DES IAB-  
SOEP-MIGRATIONSSAMPLES 2013 (M1)**

<b>Inhalt</b>	<b>Seite</b>
<b>1 Einordnung in den SOEP-Kontext</b>	<b>4</b>
<b>2 Übersicht zur zweiten Erhebungswelle des SOEP-Migrationssamples (M<sub>1</sub>)</b>	<b>6</b>
<b>3 Erhebungsinstrumente</b>	<b>8</b>
3.1 <i>Fragebogen</i>	8
3.2 <i>Einsatz von Übersetzungshilfen</i>	9
3.3 <i>Einverständniserklärung zur Registerdatenverknüpfung (Record Linkage)</i>	10
<b>4 Ergebnisse der Feldarbeit</b>	<b>11</b>
4.1 <i>Haushaltsebene</i>	11
4.2 <i>Teilnahme auf Personenebene</i>	13
4.3 <i>Partial Unit Nonresponse</i>	14
4.4 <i>Verteilung der Ankerpersonen und sonstigen Befragungspersonen nach Staatsangehörigkeit</i>	16
<b>5 Übergebener Datenbestand</b>	<b>17</b>

## 1 Einordnung in den SOEP-Kontext

Im Rahmen der seit 1984 durchgeführten Längsschnitterhebung SOEP (Sozio-ökonomisches Panel)<sup>1</sup> wurden im Laufe der Jahre sukzessive neue Teilstichproben implementiert.<sup>2</sup> Neue Teilstichproben können für das SOEP unterschiedliche Funktionen haben, im Wesentlichen aber sind es drei – sich häufig ergänzende – Zielsetzungen: Sie bilden erstens neu hinzukommende oder spezielle Grundgesamtheiten ab,<sup>3</sup> dienen zweitens als bevölkerungsrepräsentative Stichproben der quantitativen Aufstockung des Stichprobenumfangs<sup>4</sup> oder sind drittens eigenständiges Teilprojekt zur Erprobung methodischer Innovationen und innovativer Fragebogeninhalte.<sup>5</sup>

Im Erhebungsjahr 2013 wurde mit dem Migrationssample (Stichprobe M<sub>1</sub>) eine besondere Aufwuchsstichprobe implementiert. Im Vergleich zu den jüngsten Aufwüchsen J (2011) und K (2012) zielt Stichprobe M<sub>1</sub> nicht alleine auf einen quantitativen Ausbau des SOEP-Stichprobensystems, sondern vor allem auf dessen qualitative Erweiterung ab. Durch ein Stichprobendesign, das auf Adressen basiert, die vom Institut für Arbeitsmarkt- und Berufsforschung (IAB) zur Verfügung gestellt wurden, konnte die Repräsentation von Menschen mit Migrationshintergrund in Deutschland deutlich verbessert und das analytische Potential für Integrationsforschung und Einwanderungsdynamiken vermehrt werden.

Stichprobe M<sub>1</sub> ist die dritte Teilstichprobe in der Geschichte des SOEP, die ausschließlich aus Migrantenhaushalten besteht. 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 in Deutschland auftretenden Migrationsdynamiken gegründet. Demzufolge stellt die adäquate Repräsentation von Zuwandererhaushalten seit jeher eine zentrale Zielsetzung im SOEP-Stichprobendesign dar.

Das Migrationssample M<sub>1</sub> unterscheidet sich im Hinblick auf den Stichprobenumfang sowie das Design des Samples wesentlich von den früheren Migranten-Auffrischungsstichproben. Mit über 2.700 Haushalten in der Erstwelle ist es doppelt so groß wie Sample B (1984: 1.393 Haushalte) und sechsmal größer als das Sample D (1994/1995: 522 Haushalte). Im Gegensatz zu der Stichprobe aus dem Einwohnermeldeamt von 1984 und den Screening Samples von 1994/1995 wurden das Sampling Design und der Sampling Prozess für Sample M<sub>1</sub> nicht von TNS Sozialforschung

<sup>1</sup> Das Sozio-ökonomische Panel (SOEP) ist eine seit 1984 jährlich laufende repräsentative Langzeitbefragung privater Haushalte in Deutschland. Die Verantwortung für das Gesamtprojekt liegt bei der SOEP-Gruppe des Deutschen Instituts für Wirtschaftsforschung (DIW). Kooperationspartner für die Datenerhebung, Datenaufbereitung und Beratung hinsichtlich Stichprobenanlage, Fragebogendesign und Panelsteuerung ist seit der ersten Erhebungswelle im Jahr 1984 der Geschäftsbereich Sozialforschung des Markt- und Meinungsforschungsinstituts TNS Infratest in München.

<sup>2</sup> Die Ausgangsstichprobe des Jahres 1984 setzt sich aus Stichprobe A: Deutsche Haushalte, und Stichprobe B: Ausländer-Haushalte zusammen.

<sup>3</sup> Die bisherigen Erweiterungsstichproben des SOEP sind Stichprobe C: Ostdeutsche Haushalte (seit 1990); Stichprobe D: Zuwanderer der Dekade 1984-1994 (seit 1995); Stichprobe G: Einkommensstarke Haushalte (seit 2002)

<sup>4</sup> Die bisherigen Aufstockungsstichproben sind: Stichprobe E (seit 1998), Stichprobe F (seit 2000), Stichprobe H (seit 2006), Stichprobe J (2011) und Stichprobe K (2012).

<sup>5</sup> Beispiele für eigenständige Innovationsstichproben sind die beiden Biomarkerstudien der Jahre 2008 und 2009 (Messung genetischer Grundlagen von Alltagsentscheidungen unter Surveybedingungen), die drei Mutter-Kind-Studien (Kompetenz- und Verhaltenstests mit Kindern im Vorschulalter) der Jahre 2008 und 2009 sowie insbesondere das seit 2009 etablierte Innovationssample.

durchgeführt. Um einen innovativen Sampling Prozess zu ermöglichen, der auch die neueren Migrationsbewegungen abdeckt, wurde eine Kooperation zwischen der SOEP Abteilung am Deutschen Institut für Wirtschaftsforschung (DIW Berlin) und dem Institut für Arbeitsmarktforschung (IAB Nürnberg) ins Leben gerufen. Auf dieser vertraglichen Basis wurde die Stichprobe der integrierten Erwerbsbiographien (IEBS) der Bundesagentur für Arbeit (BA) als Ausgangspunkt für das Sampling verwendet.<sup>6</sup>

<sup>6</sup> Näheres zum Stichprobendesign und den sonstigen Spezifika können dem entsprechenden Methodenbericht zur ersten Erhebungswelle des SOEP-Migrationssamples 2013 entnommen werden.

## 2 Übersicht zur zweiten Erhebungswelle des SOEP-Migrationsamples (M<sub>1</sub>)

Im Jahr 2014 erfolgte mit der Feldarbeit der zweiten Welle der Übergang des M<sub>1</sub>-Samples in den Längsschnitt. Insgesamt 2.012 der 2.723 in 2013 interviewten Haushalte nahmen erneut teil, was eine Panelstabilität von 73,9 % entspricht. Detaillierte Ergebnisse der Feldarbeit sind in Abschnitt 4 dargestellt.

Die Feldzeit erfolgte von April bis Oktober 2014, **Tabelle 1** zeigt den zeitlichen Verlauf der Feldarbeit. Gut drei Viertel der eingesetzten Haushalte konnten in den ersten vier Monaten final bearbeitet werden. Die verbleibenden drei Monate wurden im Sinne einer ausschöpfungsorientierten Bearbeitung fast ausschließlich auf sogenannte „schwierige“ Haushalte verwendet. Dabei handelt es sich beispielsweise um umgezogene 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:

### Zeitlicher Verlauf der Feldarbeit 2014: Final bearbeitete Haushalte in Prozent der Bruttostichprobe<sup>1</sup>

	Sample M <sub>1</sub>
April	11,6
Mai	48,2
Juni	65,1
Juli	77,4
August	88,1
September	99,0
Oktober	100,0

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

Die Befragung erfolgte ausschließlich persönlich-mündlich mit dem Erhebungsmodus CAPI unter Verwendung aller auch in den sonstigen SOEP-Stichproben eingesetzten Erhebungsinstrumente. Wie bereits in Welle 1 wurde für ausgewählte Haushalte das Einverständnis zur Registerdatenverknüpfung erhoben (siehe Abschnitt 3).

Das Sampling der Migrationsbefragung in der ersten Welle war registerbasiert. Das hatte zur Folge, dass das übliche SOEP Konzept, bei dem 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. Konsequenterweise wurde in einem ersten Schritt ein kurzer Screening-Fragebogen erhoben, um den Migrationshintergrund der Ankerperson zu validieren. Bei negativem Screening wurde der gesamte Haushalt aus der Befragung ausgeschlossen, ungeachtet der Frage, ob andere Haushaltsmitglieder einen Migrationshintergrund hatten oder nicht. Bei positivem Screening hingegen und Teilnahme der Ankerperson an der Befragung wurden alle im Haushalt

lebenden Personen, die vor 1996 geboren wurden, zur Befragung zugelassen, ohne Beachtung des jeweiligen Migrationshintergrunds. Eine logische Folge dieser speziellen Vorgehensweise in der Erstwelle war der deutlich erhöhte Kontaktierungsaufwand für die Interviewer im Vergleich zur üblichen SOEP Befragung, bei der jeder Erwachsene im ausgewählten Haushalt interviewt werden kann.

War das Ankerpersonenkonzept in der Erstwelle unumgänglich, um die Zielpopulation adäquat abbilden zu können, so verlor dieser Ansatz in der zweiten Welle seine Bedeutung. Es konnte das originäre SOEP-Haushaltskonzept mit Haushalten als primäre Stichprobeneinheit – hier den teilnehmenden Haushalten der ersten Welle der  $M_1$ -Stichprobe des Vorjahres – angewandt werden. Daher wurden Haushalte auch dann in die Bearbeitung aufgenommen, wenn die ursprüngliche Ankerperson aktuell den Haushalt verlassen hatte oder nicht mehr teilnahmebereit war. Insgesamt wurden in 2014 118 Haushalte bearbeitet, in denen die jeweilige ursprüngliche Ankerperson aus dem Erhebungsjahr 2013 nicht mehr anzutreffen war. Exakt 50 % dieser Haushalte konnten realisiert werden. Innerhalb aller teilnehmenden Haushalte war die Teilnahmebereitschaft der Ankerpersonen mit 96,8 % gegenüber 83,3 % bei den abgeleiteten Stichprobenmitgliedern ausgesprochen hoch.



## **3 Erhebungsinstrumente**

### **3.1 Fragebogen**

Zur Datenerhebung wurden alle Fragebogen aus dem aktuellen Hauptsample (Stichproben A – KJ) verwendet. Im Einzelnen waren das im Erhebungsjahr 2014:

- HAUSHALTS-Fragebogen, jährlich pro Haushalt einmal auszufüllen
- PERSONEN-Fragebogen für jedes Haushaltsmitglied, das 1996 oder früher geboren ist.
- LEBENSLAUF-Fragebogen zur einmaligen Erfassung biografischer Daten, für Personen die 1996 oder früher geboren sind und in der Vorwelle nicht teilgenommen haben. Dieser Fragenkomplex war in den Personenfragebogen integriert.
- JUGEND-Fragebogen für alle Personen im Haushalt, die 1997 geboren sind.
- „Lust auf DJ“ ist ein Zusatzfragebogen zur Kompetenzmessung für alle Personen, die den Jugendfragebogen ausgefüllt haben.
- „Fragebogen für Schülerinnen und Schüler“ für Personen im Haushalt, die 2002 geboren sind und erstmals persönlich an der Befragung teilnehmen.
- „Mutter und Kind A“ für Mütter von Kindern, die in 2014 geboren sind (bzw. in 2013, aber nach der letzten Befragung).
- „Ihr Kind im Alter von 2 oder 3 Jahren“ (Mutter und Kind B) für Mütter von Kindern, die in 2011 geboren sind.
- „Ihr Kind im Alter von 5 bis 6 Jahren“ (Mutter und Kind C) für Mütter von Kindern, die in 2008 geboren sind.
- „Ihr Kind im Alter von 7 bis 8 Jahren“ (Elternfragebogen D) für Mütter und Väter von Kindern, die 2006 geborenen sind.
- „Ihr Kind im Alter von 9 bis 10 Jahren“ (Mutter und Kind E) für Mütter von Kindern, die in 2004 geboren sind.
- „Personenbefragung 2013 Nacherhebung Kurzfassung“ („LÜCKE“-Fragebogen) für temporäre Ausfälle des Vorjahres um die „Lücke“ in den Längsschnittdaten zu minimieren.
- Zusatzfragebogen „Die verstorbene Person“ für Panelteilnehmer, die seit der letzten Befragung einen Todesfall in ihrem Haushalt oder ihrer Familie zu beklagen hatten.
- Einverständniserklärung zur Verknüpfung der Befragungsdaten mit Registerdaten des IAB.

Gegenüber den Instrumenten der SOEP-Haupterhebung ergaben sich im Erhebungsprogramm für M<sub>1</sub> einzig im Personenfragebogen nennenswerte Anpassungen, indem die in der ersten Welle etablierte Migrationshistorie für Erstbefragte sowie ergänzende Fragen zum Themenkomplex Migration und (Arbeitsmarkt-) Integration hinzugefügt wurden. Sämtliche Haupt-Instrumente wurden ausschließlich mit CAPI erhoben. Lediglich der Kompetenztest für Jugendliche („Lust auf DJ“) und die Einverständniserklärung zur Registerdatenverknüpfung wurden mit Papierfragebogen erhoben.

### 3.2 Einsatz von Übersetzungshilfen

Da die Zielpopulation (überwiegend) ausländischer Herkunft ist, wurden die Kernfragebogen (Haushalt und Personen) in fünf Sprachen übersetzt: Englisch, Russisch, Türkisch, Rumänisch und Polnisch. Diese Sprachen – mit Ausnahme von Englisch – entsprechen den Nationalitäten, die im Bruttosample der Erstwelle überrepräsentiert waren. Die übersetzten Versionen wurden nicht in CAPI integriert, sondern auf Papier gedruckt und dem Interviewer als Übersetzungshilfen zur Verfügung gestellt, um Sprachbarrieren zu überwinden. **Tabelle 2** zeigt verschiedene Hilfen, die von den Interviewern in sprachlich schwierigen Situationen eingesetzt wurden.

Tabelle 2:

#### Sprachprobleme und Verwendung von Übersetzungshilfen

	Anzahl	In % der Nettostichprobe
Keine Sprachprobleme	3.212	85,6
Unterstützung bei Sprachproblemen benötigt	540	14,4
<i>Davon:<sup>1</sup></i>		
Deutsch sprechende Person im Haushalt	356	9,5
Deutsch sprechende Person außerhalb des Haushalts	64	1,7
Professioneller Dolmetscher	2	0,1
Übersetzungshilfe	142	3,8
<i>Davon:</i>		
Russisch	64	1,7
Türkisch	31	0,8
Rumänisch	18	0,5
Polnisch	16	0,4
Englisch	13	0,3

<sup>1</sup> In 24 Fällen wurde mehr als eine Art von Unterstützung benötigt.

Wie bereits in der Erstwellenerhebung zeigt sich auch in Welle 2, dass bei einem Großteil der Interviews (85,6 %) entweder keine Sprachprobleme aufgetreten sind oder zumindest nicht in dem Maße, dass Unterstützung in irgendeiner Form herangezogen wurde. In 14,4 % der Personeninterviews wurde Unterstützung bei Sprachproblemen benötigt, wobei diese in den meisten Fällen durch im Haushalt lebende Personen geleistet wurde. Lediglich bei 3,8 % der befragten Personen wurden Übersetzungshilfen eingesetzt. Dabei wurde mehrheitlich die Russische Übersetzungshilfe verwendet.

### 3.3 Einverständniserklärung zur Registerdatenverknüpfung (Record Linkage)

Eine weitere 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 des Samples M<sub>1</sub>, indem die Bildungs- und Erwerbsbiographien der Befragten mit den zwei unterschiedlichen Datentypen – Befragtenangaben einerseits und Registerdaten andererseits – möglichst vollständig abgebildet werden.

Wie schon in der Erstwelle wurde 2014 erneut ein Teil der Zielpersonen am Ende des Personenfragebogens um ihr schriftliches Einverständnis zum sogenannten „Record Linkage“ gebeten. In der Erstwelle wurden 70 % des Bruttosamples (auf Haushaltsebene) nach ihrem Einverständnis befragt; 15 % des Bruttos sollten erst in Welle 2 um ihr Einverständnis gebeten werden und die verbleibenden 15% waren nicht für das „Record Linkage“ vorgesehen (Kontrollgruppe). Diese drei Gruppen wurden ex-ante randomisiert eingeteilt. Die Idee hinter diesem Ansatz ist, dass die Kooperationsraten bei einer derartigen Anfrage innerhalb verschiedener Wellen variieren könnten. Dabei wird vermutet, dass die Kooperationsrate in der zweiten Welle höher sein müsste als in der Erstwelle.

In 2014 wurde der Gruppe, die für die zweite Welle vorgesehen war, die Einverständniserklärung vorgelegt. Zusätzlich wurden alle Personen ein zweites Mal mit einbezogen, die in der Erstwelle ihr Einverständnis nicht gegeben haben. Insgesamt wurden somit 1.379 Personen um ihr Einverständnis zur Registerdatenverknüpfung gebeten, von denen 610 Personen ihre Zustimmung erteilten. Wie in **Tabelle 3** dargestellt, lag die Gesamt-Response-Rate in 2014 bei 44,2 % und damit um 4,7 Prozentpunkte niedriger als noch im Vorjahr. Aussagekräftiger ist jedoch die Rate für die Teilgruppe, die explizit für das „Record Linkage“ in Welle 2 vorgesehen war. Hier ist die Kooperationsrate mit 49,4 % nur marginal höher als im Vorjahr. Es konnte daher nicht statistisch signifikant nachgewiesen werden, dass die Kooperationsrate zur Registerdatenverknüpfung in der zweiten Welle höher ist als in der Erstwelle. Die wiederholte Abfrage zum Record Linkage bei den Personen, die im Vorjahr keine Zustimmung erteilt hatten, erzielte etwas mehr als 40 % Zustimmung. Die erneute Abfrage ist zwar mit einem nicht unerheblichen Aufwand verbunden, jedoch konnte die Rate der Einverständniserklärungen insgesamt nochmals deutlich gesteigert werden.

Tabelle 3:  
**Einverständnis zum „record linkage“: Response Rates**

	2013		2014		
	Zugestimmt	Abgelehnt	Ausgewählt für Record Linkage in Welle 2	Nicht zugestimmt in Welle 1	Gesamt
Zugestimmt	48,9%	51,1%	49,4%	41,5%	44,2%
Abgelehnt	51,1%	48,9%	50,6%	58,5%	55,8%

## 4 Ergebnisse der Feldarbeit

### 4.1 Haushaltsebene

Die **Tabellen 4 und 5** geben die Resultate der Feldarbeit in der zweiten Welle wieder. Alles in allem wurden 2.841 Adressen bearbeitet. Bezüglich der Zusammensetzung des Bruttosamples waren 95,8 % aller Haushalte Teilnehmer des Vorjahres und 4,2 % Neu-Haushalte (Split-off-Haushalte). Insgesamt wurden 2.012 Haushalte interviewt, was zu einer Gesamt-Response-Rate von 70,8 % führt. In der Subgruppe der Vorjahresteilnehmer lag die Response Rate bei 71,9 %.

Tabelle 4:

**Zusammensetzung der Bruttostichprobe und Response Rates**

	Sample M <sub>1</sub>	
	Abs.	In %
<b>(1) Zusammensetzung der Bruttostichprobe nach Haushaltstyp</b>	<b>2.841</b>	<b>100,0</b>
Vorjahresteilnehmer	2.723	95,8
Neue Haushalte (split-off Haushalte)	118	4,2
<b>(2) Art der Feldarbeit</b>		
<b>Keine Feldarbeit</b> (Finale Ausfälle, Verstorbene und ins Ausland verzogene)	34	1,2
<b>Interviewer-basiert (CAPI)</b>	2.807	98,8
<b>(3) Nettostichprobe/Response Rate gesamt</b>	<b>2.012</b>	<b>70,8</b>
- Nettostichprobe/RR Vorjahresteilnehmer	1.957	71,9
- Nettostichprobe/RR Neue Haushalte	55	46,6

Bei der Verteilung der Ausfälle war die mit Abstand größte Gruppe die der weichen und harten Ausfälle mit 21,5 % des Bruttosamples, was 73,7% aller Ausfälle entspricht. Der Anteil der Haushalte in Sample M<sub>1</sub>, die während der Feldzeit nicht erreicht werden konnten (6,4 % des Bruttosamples und 21,8 % der Ausfälle) ist auffällig hoch. Verglichen mit der jeweils zweiten Welle der Samples J und K ist dieser Anteil um 8,5 beziehungsweise 6,7 Prozentpunkte höher bezogen auf den Anteil an den Ausfällen insgesamt. Dies bestätigt die Erfahrungen aus der Erstwelle von Sample M<sub>1</sub>, dass die betreffende Zielpopulation der Migrantenhaushalte schwerer zu kontaktieren ist, als dies bei den SOEP-Aufwuchsstichproben zum Bevölkerungsquerschnitt der Fall ist.

Tabelle 5:

**Ergebnisse der Feldarbeit**

	Anzahl	In % des Bruttosample
<b>Bruttostichprobe</b>	<b>2.841</b>	<b>100,0</b>
<b>Kein Interview</b>	<b>829</b>	<b>29,2</b>
- Nicht erreicht während der Feldzeit	181	6,4
- Ankerpersonen mit permanentem Auslandsaufenthalt	13	0,5
- Physisch oder mental nicht in der Lage	18	0,6
- Sprachprobleme	6	0,2
- Weiche Verweigerer (derzeit nicht bereit/nicht in der Lage)	180	6,3
- Endgültige Verweigerer	431	15,2
<b>Interviews (Nettostichprobe)</b>	<b>2.012</b>	<b>70,8</b>
- Haushalt vollständig realisiert	1.677	59,0
- Haushalt teilweise realisiert	335	11,8

**Tabelle 6** vergleicht die 2.-Wellen-Response-Rates und Panelstabilitätsquoten von  $M_1$  mit denen der letzten Auffrischungstichproben J und K. Beide Indikatoren der Feldarbeit sind rund 10 Prozentpunkte niedriger als in den Samples J und K. Dies spiegelt die bekannten Probleme in der Erhebung von Migrantenhaushalten und auch deren Strukturen wider: In einem Migrationssample sind die Anstrengungen, die ein Interviewer unternehmen muss, um einen Haushalt zu kontaktieren und alle im Haushalt lebenden Zielpersonen zu einem Interview zu motivieren erstens größer als in Befragungen, die sich an den Bevölkerungsquerschnitt richten. Ebenso sind die Kontaktaufnahme und die Interviewsituation um einiges komplizierter (z.B. Sprachprobleme, kulturelle Besonderheiten, niedrigeres Bildungsniveau, etc.). Zweitens ist auch die durchschnittliche Anzahl an Personen, die in einem Haushalt leben, erheblich größer, als in der allgemeinen Bevölkerung.

Tabelle 6:

**Response Rates und Panelstabilitätsquoten in Welle 2 in ausgewählten SOEP-Samples**

	J 2012	K 2013	M 2014
Response Rate (Vorjahresteilnehmer)	80,0%	82,0%	71,9%
Panelstabilität	81,5%	83,9%	73,9%

## 4.2 Teilnahme auf Personenebene

**Tabelle 7** zeigt die Volumina sowie die Teilnahmequoten für die einzelnen Fragebogen auf Personenebenen. Die Bezugswerte zur Berechnung der Response Rates sind jeweils die entsprechenden Zielpopulationen für die einzelnen Fragebogen in teilnehmenden Haushalten. Die Teilnahmequoten auf Personenebene sind erfreulich hoch und entsprechen im Gegensatz zur deutlich geringeren Response Rate auf Haushaltsebene in etwa dem Niveau der etablierten Stichproben A – KH der SOEP-Kernbefragung.

Tabelle 7:

### Volumina und Response Rates der eingesetzten Fragebogen in teilnehmenden Haushalten

	Bruttosample/ Referenzwert <sup>1</sup>	Anzahl an Interviews	Response Rate/ Deckungsrate
Personenfragebogen <sup>2</sup>	4.172	3.739	89,6
Jugendfragebogen	93	83	89,3
Kognitiven Kompetenztest <sup>3</sup>	83	75	90,4
Mutter und Kind Fragebogen A	156	140	89,7
Mutter und Kind Fragebogen B	130	121	93,1
Mutter und Kind Fragebogen C	118	110	93,2
Elternfragebogen D <sup>4</sup>	145/290	136/222	93,8/76,6
Mutter und Kind Fragebogen E	114	107	93,9
Schüler	81	72	88,9

<sup>1</sup> Für die kindbezogenen Fragebogen ist der Referenzwert die jeweilige Alterskohorte aller Kinder in teilnehmenden Haushalten. Deswegen gibt die Response Rate für diese Fragebogen an, für wie viele Kinder ein Fragebogen von einem Elternteil vollständig beantwortet wurde (in den meisten Fällen die Mutter).

<sup>2</sup> Für 13 Personen liegt ein Personenfragebogen vor, allerdings kein entsprechender Haushaltsfragebogen.

<sup>3</sup> Der Test kann nur erhoben werden, wenn der Jugendfragebogen bearbeitet wurde. Deswegen unterscheidet sich das betreffende Bruttosample der Zielpopulation (n=83) von dem des Jugendfragebogens (n=93).

<sup>4</sup> Im Gegensatz zu den anderen kindbezogenen Fragebogen soll dieser Fragebogen nicht nur von einem, sondern (wenn möglich bzw. vorhanden) von beiden Elternteilen beantwortet werden. Für 136 (93,7%) von 145 Kindern, die im Jahr 2006 geboren sind und in Haushalten leben, die 2014 teilgenommen haben, wurde mindestens ein Fragebogen ausgefüllt. Insgesamt waren es 222 Elternfragebogen. Somit haben in 76,6% der Fälle beide Elternteile den Fragebogen für das betreffende Kind ausgefüllt.

Insgesamt haben 89,6 % der Befragungspersonen (Jahrgang 1996 oder älter) in teilnehmenden Haushalten einen Personenfragebogen beantwortet. Betrachtet man die Response Rates nach Migrationshintergrund, ist auffällig, dass Personen mit ausländischer Nationalität mit 90,5 % und vor allem Migranten der zweiten Generation (mindestens ein Elternteil ist nicht in Deutschland geboren) mit 93,7 % eine signifikant höhere Ausschöpfungsquote aufweisen als Personen ohne jeden Migrationshintergrund (81,4 %) (vgl. **Tabelle 8**). Auf Haushaltsebene scheint es, wie oben erläutert, für die eingesetzten Interviewer bei der Bearbeitung von Haushalten, in denen überwiegend Personen mit Migrationshintergrund leben, schwieriger zu sein diese zur Teilnahme zu gewinnen, als bei der allgemeinen Bevölkerung. Gelingt das jedoch, sind die Teilnahmequoten auf Personenebene aber durchaus mit anderen SOEP-Samples vergleichbar. Aufgrund der thematischen Ausrichtung der Studie gelingt es bei Personen mit Migrationshintergrund sogar deutlich besser, diese zur Teilnahme zu gewinnen als das bei Personen ohne Migrationshintergrund.

Tabelle 8:

**Teilnahme auf Personenebene in teilnehmenden Haushalten nach Migrationshintergrund**

	N Brutto	N Netto	Response Rate
Nationalität nicht Deutsch	2.045	1.851	90,5 %
Migrationshintergrund (2. Generation)	1.271	1.191	93,7 %
Kein Migrationshintergrund	856	697	81,4 %
Gesamt	4.172	3.739	89,6 %

### 4.3 Partial Unit Nonresponse

Ein weiterer wichtiger Indikator zur Bemessung der Qualität der Feldarbeit im Querschnitt aber auch für die Stabilität im Längsschnitt ist der Anteil teilrealisierter Haushalte gemessen an der Gesamtzahl der realisierten Nettostichprobe, also der sogenannte Partial Unit Nonresponse (PUNR). Ein Haushalt gilt im SOEP-Kontext dann als realisiert, wenn ein Haushaltsfragebogen und mindestens ein Personen- oder Jugendfragebogen vorliegen. Ein Haushalt ist dann „vollständig realisiert“, wenn alle zu befragenden Personen auch tatsächlich teilgenommen haben. Hat mindestens eine Befragungsperson nicht teilgenommen gilt der Haushalt als „teilweise realisiert“. Grundsätzlich wird der PUNR also als Anteil teilrealisierter Haushalte am Gesamtnetto berechnet. Je nach Bemessungsgrundlage zur Berechnung des PUNR kann dieser Wert auf vier Arten berechnet werden:

- PUNR 1: Diese Berechnung berücksichtigt nicht die Haushaltsgröße, d. h. Einpersonenhaushalte mit Interview, die zwangsläufig immer „vollständig realisiert“ sind, sind hier mit eingeschlossen. Zudem sind hier unter den „vollständig realisierten“ Haushalten auch Haushalte vertreten, die mindestens eine Person im Befragungsalter beinhalten, die bereits in einer Vorwelle endgültig verweigert hat. Streng genommen sind diese Haushalte ebenfalls nur teilweise realisiert (siehe PUNR 3 und 4), da die betreffenden Personen zur Zielgruppe gehören. Da Personen, die in Vorwellen verweigert hatten, nicht in das Bearbeitungsbrutto der aktuellen Welle mit einfließen und somit auch nicht bearbeitet werden,<sup>7</sup> werden bei der Berechnung von PUNR 1 derartige Haushalte als vollständig definiert, insofern alle als Befragungsperson der laufenden Welle geführten Personen teilgenommen haben. PUNR 1 ist somit ein Indikator für die Bearbeitungsgüte einer aktuellen Welle, da er sich auf die tatsächlich eingesetzte Bruttosichprobe auf Personenebene bezieht. Diese Berechnungsgrundlage wird auch bei der Codierung der Ergebnisvariablen im Bruttoband (Haushaltsebene) herangezogen.
- PUNR 2: Diese Berechnung entspricht PUNR 1, allerdings nur bezogen auf Haushalte, in denen mehr als eine Befragungsperson lebt. Einpersonenhaushalte, die den PUNR-Anteil automatisch reduzieren, werden hier nicht berücksichtigt.

<sup>7</sup> Personen, die in einer beliebigen Welle endgültig ihre (weitere) Teilnahme verweigert haben, werden nicht in das Feldeinsatzbrutto einer aktuellen Erhebungswelle mit einbezogen. Bei der Auflistung der Haushaltszusammensetzung für die Interviewer, in der auch angegeben ist, welche Person im Haushalt welche Fragebogen ausfüllen soll, sind diese Personen zwar enthalten, aber entsprechend gekennzeichnet. Diese werden nicht als Zielperson angegangen und somit auch nicht um ein Interview gebeten. Es kommt allerdings in jeder Erhebungswelle vor, dass Personen, die ursprünglich einmal „hart“ verweigert hatten, doch wieder teilnehmen möchten. Diese sind dann in der Folgewelle wieder als zu befragende Person geführt. Im Erhebungsjahr 2014 haben 55 Personen in der Stichprobe M<sub>1</sub>, die aufgrund der Verweigerung in der Vorwelle zum Feldstart nicht als Befragungsperson geführt wurden, doch an der Erhebung teilgenommen.

- PUNR 3: Hier werden ohne Berücksichtigung der Haushaltsgröße zusätzlich auch diejenigen Haushalte als teilrealisiert gewertet, die Personen enthalten, die in Vorwellen verweigert haben.
- PUNR 4: Diese Berechnung entspricht PUNR 3, allerdings bezogen auf Haushalte, in denen mehr als eine Befragungsperson lebt.

Die PUNR-Anteile des Erhebungsjahres 2014 der Stichproben  $M_1$  in allen vier Varianten sind in **Tabelle 9** aufgeführt. Die Zielsetzung bei der jährlichen Bearbeitung der Bruttostichprobe liegt darin, den Anteil teilrealisierter Haushalte so gering wie möglich zu halten. Der hierfür herangezogene Indikator ist PUNR 1. Zur Minimierung dieses Anteils werden während der Feldbearbeitung zahlreiche Maßnahmen und Nachfassaktionen ergriffen. So wenden sich beispielsweise bei fehlenden Fragebogen (das schließt auch die weiteren Zusatzfragebogen mit ein) die Interviewer erneut an die Haushalte, bei Unklarheiten bezüglich der Haushaltzusammensetzung werden die Haushalte direkt telefonisch und/oder postalisch um fehlende Fragebogen gebeten. Sobald Hinweise auf neue oder nicht erfasste Haushaltsmitglieder (auch in den Nettodaten) vorliegen, werden diese Personen ebenfalls in eine Nachbearbeitungswelle einbezogen. Eine Partial Unit Nonresponse, die gegen Null tendiert, ist dennoch im Rahmen einer derartigen Studie ausgeschlossen, da es immer Personen geben wird, die in einer Erhebungswelle nicht erreichbar sind, oder nicht teilnehmen können (z. B. längere Krankheit) oder wollen. Sofern die weitere Teilnahme nicht eindeutig auch für Folgewellen verweigert wurde, werden diese Personen im Folgejahr erneut in die Bruttostichprobe aufgenommen und wenn möglich befragt.

Tabelle 9:

**Anteil teilweise realisierter Haushalte (PUNR) am Gesamtnetto 2014**

Sample	K Welle 2	$M_1$ Welle 2
<b>PUNR 1</b> (inklusive Einpersonenhaushalte und exklusive harter Verweigerer aus Vorwellen)	9,8 %	16,7 %
<b>PUNR 2</b> (Mehrpersonenhaushalte exklusive harter Verweigerer aus Vorwellen)	15,1 %	20,1 %
<b>PUNR 3</b> (inklusive Einpersonenhaushalte und harter Verweigerer aus Vorwellen)	12,1 %	22,0 %
<b>PUNR 4</b> (Mehrpersonenhaushalte inklusive harter Verweigerer aus Vorwellen)	18,7 %	26,6 %

**Tabelle 8** stellt die beschriebenen PUNR-Varianten für  $M_1$  den Werten für die zweite Welle der jüngsten SOEP-Aufwuchsstichprobe K gegenüber. Es zeigt sich in allen Varianten ein deutlich höherer PUNR-Anteil bei  $M_1$  als bei K. Dies scheint auf den ersten Blick die Ergebnisse des vorigen Abschnittes – nämlich eine vergleichsweise hohe Ausschöpfung auf Personenebene – zu widerlegen. Ein wesentlicher Grund für die erhöhten PUNR-Anteile ist allerdings darin zu sehen, dass die durchschnittliche Anzahl an Befragungspersonen pro Haushalt bei allgemeinen Bevölkerungstichproben signifikant niedriger ist als bei Stichproben mit Migrantenhaushalten. So liegt die durchschnittliche Anzahl an Personen im Befragungsalter bei  $M_1$  bei 2,21 Personen, während sie bei Sample K bei 1,79 liegt. Mit steigender Haushaltsgröße ist es für die Interviewer zunehmend schwierig, alle relevanten Personen zu befragen, selbst wenn die Teilnahmebereitschaft auf Personenebene auf einem vergleichbaren Niveau ist wie bei anderen SOEP-Aufwuchsstichproben.



#### 4.4 Verteilung der Ankerpersonen und sonstigen Befragungspersonen nach Staatsangehörigkeit

**Tabelle 10** vergleicht die Verteilung der Ankerpersonen und sonstigen Befragungspersonen nach Staatsangehörigkeit der Erhebungsjahre 2013 und 2014. Die Werte beziehen sich auf Personen in teilnehmenden Haushalten, für die ein Personenfragebogen vorliegt. Es zeigt sich in beiden Subgruppen, dass sich in der zweiten Erhebungswelle gegenüber der Erstwellenbefragung keine nennenswerten Verschiebungen – etwa durch selektive Ausfallprozesse – in den Verteilungen hinsichtlich der Nationalität der Befragungspersonen vollzogen haben.

Tabelle 10:

#### Verteilung der Ankerpersonen und sonstigen Befragungspersonen nach Staatsangehörigkeit (Nettostichprobe)

	2013				2014			
	Ankerpersonen		Sonstige Befragungspersonen		Ankerpersonen		Sonstige Befragungspersonen	
Staatsangehörigkeit	N	%	N	%	N	%	N	%
Deutschland	1.177	43,2	1.265	56,4	851	45,0	1.037	56,1
Türkei	237	8,7	217	9,7	152	8,0	172	9,3
Polen	170	6,2	63	2,8	117	6,2	55	3,0
Italien	138	5,1	81	3,6	94	5,0	65	3,5
Rumänien	136	5,0	66	2,9	90	4,8	47	2,5
Griechenland	109	4,0	78	3,5	67	3,5	68	3,7
Russland	103	3,8	40	1,8	69	3,6	45	2,4
Kosovo-Albaner	68	2,5	79	3,5	45	2,4	69	3,7
Ukraine	67	2,5	30	1,3	55	2,9	23	1,2
Spanien	64	2,4	10	0,4	39	2,1	9	0,5
Serbien	49	1,8	52	2,3	33	1,7	41	2,2
Kroatien	37	1,4	23	1,0	28	1,5	23	1,2
Kasachstan	32	1,2	14	0,6	23	1,2	12	0,6
Irak	19	0,7	15	0,7	12	0,6	12	0,6
Sonstige	318	11,6	208	9,3	216	11,4	170	9,2
<b>Gesamt</b>	<b>2.724</b>	<b>100,0</b>	<b>2.241</b>	<b>100,0</b>	<b>1.891</b>	<b>100,0</b>	<b>1.848</b>	<b>100,0</b>

## 5 Übergebener Datenbestand

Der Datenbestand des SOEP-Migrationssamples wurde zunächst als 1. Vorablieferung am 23. 09. 2013 und als 2. Vorablieferung am 16. 12. 2013 an das DIW übergeben. Der endgültige und vollständige Datenbestand für die Stichprobe M<sub>1</sub> wurde am 20. 01. 2014 übergeben.

### 1. Bruttodaten

<b>Bezeichnung</b>	<b>Anzahl Fälle</b>	<b>Dateiname</b>
Bruttoband Haushalte	2.841	hbru_M_2014.sav
Bruttoband Personen	8.745	Pbru_M_2014.sav
Interviewerband	203	Intband_M_2014.sav

### 2. Nettodaten

<b>Bezeichnung</b>	<b>Anzahl Fälle</b>	<b>Dateiname</b>
Einverständniserklärung zur Registerdaten- verknüpfung	3.835	EV14_M_DIW.sav
Nettodaten Haushalte geprüft	2.012	H14_M.sav
Nettodaten Haushalte ungeprüft	2.012	H14_M_u.sav
Nettodaten Personen geprüft	3.752	PB14_M.sav
Nettodaten Personen ungeprüft	3.752	PB14_M_u.sav
Nettodaten Jugend geprüft	83	J14_m.sav
Nettodaten Jugend ungeprüft	83	J14_m_u.sav
Nettodaten Kompetenztest für Jugendliche	75	DJ14_M.sav
Nettodaten Schüler	72	S14_M.sav
Nettodaten Mutter-und-Kind A	140	MA14_M.sav
Nettodaten Mutter-und-Kind B	121	MB14_M.sav
Nettodaten Mutter-und-Kind C	110	MC14_M.sav
Nettodaten ELTERN-Fragebogen	222	EL14_M.sav
Nettodaten Mutter-und-Kind E	107	ME14_M.sav
Nettodaten „Die verstorbene Person“	59	VP14_M.sav
Angaben zum 7. Kind	6	7Kind_M_2014.sav
Vercodung offener Angaben: Berufe und Branchen (kldb92; isco88; wz2003)	12.074	Berufe_Branchen_Ausb_ M_2014_alteKI.SAV
Vercodung offener Angaben: Berufe und Branchen (kldb2010; isco08; wz2008)	12.074	Berufe_Branchen_Ausb_ M_2014_neueKI.SAV
Vercodung offener Angaben: Hochschulabschlüsse	41	Hochschulabschluss _M_2014.sav