

Forschungsdatenzentrum

der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt-
und Berufsforschung

FDZ

FDZ-Methodenreport

02/2013

DE

Methodische Aspekte zu Arbeitsmarktdaten

Bias in einem komplexen Surveydesign

Ausfallprozesse und Selektivität in der
IAB-Befragung ALWA

Corinna Kleinert,
Michael Ruland,
Annette Trahms



Bundesagentur für Arbeit

Bias in einem komplexen Surveydesign

Ausfallprozesse und Selektivität in der IAB-Befragung ALWA

Corinna Kleinert (Institut für Arbeitsmarkt- und Berufsforschung)

Michael Ruland (Wissenschaftszentrum Berlin für Sozialforschung)

Annette Trahms (Institut für Arbeitsmarkt- und Berufsforschung)

Die FDZ-Methodenreporte befassen sich mit den methodischen Aspekten der Daten des FDZ und helfen somit Nutzerinnen und Nutzern bei der Analyse der Daten. Nutzerinnen und Nutzer können hierzu in dieser Reihe zitationsfähig publizieren und stellen sich der öffentlichen Diskussion.

FDZ-Methodenreporte (FDZ method reports) deal with the methodical aspects of FDZ data and thus help users in the analysis of data. In addition, through this series users can publicise their results in a manner which is citable thus presenting them for public discussion.

Inhaltsverzeichnis

Zusammenfassung	4
Abstract	4
1 Einleitung	5
2 Studiendesign und Ausfallprozesse	6
3 Selektivität in der CATI-Befragung	9
3.1 Ausmaß der Selektivität: Vergleich mit externen Referenzdaten	9
3.2 Quellen der Selektivität I: Analyse der EWO-Merkmale	12
3.3 Exkurs: Zur Repräsentativität der Zusatzbefragungen	14
3.4 Quellen der Selektivität II: Ergebnisse der Non-Response-Befragung	16
4 Selektivität im zweistufigen Design: von CATI zu PAPI	18
4.1 Teilnahmebereitschaft	19
4.2 Erreichbarkeit	22
4.3 Kooperationsbereitschaft	23
5 Zusammenfassung und Schlussfolgerungen	26
Literatur	29

Zusammenfassung

Dieser Methodenreport befasst sich mit der Analyse von Selektivität in der IAB-Befragung „Arbeiten und Lernen im Wandel“ (ALWA). Neben den üblichen Ausfallprozessen in standardisierten Befragungen, welche aus Nichterreichbarkeit und Nichtteilnahme resultieren, gab es hier Ausfälle bei der Telefonnummernrecherche aus der Einwohnermeldeamtsstichprobe und zusätzliche Verzerrungen, die durch das komplexe zweistufige Design von ALWA entstanden sind. Die Ergebnisse zeigen, dass die ALWA-Telefonbefragung hinsichtlich zentraler Merkmale ähnliche Verzerrungen aufweist wie andere Bevölkerungsbefragungen. Diese sind zum großen Teil auf die Telefonnummernrecherche zurückzuführen. Positive Erfahrungen wurden in ALWA damit gemacht, Lese- und Mathematikkompetenzen in einer nachgelagerten persönlichen Befragung getrennt von den übrigen Befragungsinhalten zu erheben. Auch wenn nur etwa 40 Prozent der telefonisch Befragten an diesen Tests teilgenommen haben, ergab sich nur geringfügige zusätzliche Selektivität durch das zweistufige Design.

Abstract

This report contains an analysis of nonresponse bias in the IAB survey ALWA (“Working and Learning in a Changing World”). Besides the usual processes of attrition in standardized surveys, which result from persons who cannot be reached and persons who do not cooperate, two additional sources of nonresponse were relevant here: first, persons in the register based address sample for whom no telephone number could be located, and second, additional dropouts due to the complex two-step design of the study. Our results show that the ALWA telephone survey is biased in similar ways as other population surveys. For the largest part, this selectivity can be traced back to non-location because of missing telephone numbers. Positive experiences were made with collecting reading and mathematical competence data in a face-to-face interview separately from the rest of the survey. Even though only 40 percent of the CATI respondents participated in these tests, the two-step design caused only a slight amount of additional selectivity.

Keywords: Personenbefragung, Datenqualität, Selektivität, Nonresponse-Bias, Mixed-Mode-Design, ALWA-Studie

Wir danken Maria Fleischmann, die erste Analysen und Literaturarbeiten zu diesem Thema als Praktikantin im ALWA-Projekt durchgeführt hat.

1 Einleitung

In der Umfrageforschung, ob in der Sozialwissenschaft oder der Marktforschung, lässt sich in den letzten Jahren ein deutlicher Trend rückläufiger Ausschöpfungsraten beobachten (Schnell 2012). Die IAB-Befragung „Arbeiten und Lernen im Wandel“, kurz ALWA, macht da keine Ausnahme: die Realisierungsquote lag hier bei „nur“ 46 Prozent (zur Studie vgl. Klei- nert et al. 2008, Antoni et al. 2010, Gilberg et al. 2011). Allerdings weisen Experten darauf hin, dass eine hohe Ausschöpfungsrate alleine kein hinreichendes Gütekriterium darstellt und sie somit auch keine hohe Datenqualität hinsichtlich der Repräsentativität garantieren kann (Schnell et al. 2008). Auch wenn die Ausschöpfungsrate niedrig ist, können die Ausfälle zufällig verteilt sein, und die Daten der Studienteilnehmer lassen dann ohne Weiteres Rück- schlüsse auf die Grundgesamtheit zu. Wichtiger als Ausschöpfungsraten sind daher Informa- tionen über Unterschiede zwischen Studienteilnehmern und Nichtteilnehmern, denn nur da- mit lässt sich die Selektivität bzw. der Nonresponse-Bias einer Befragung abschätzen. Sol- che Analysen stehen im Mittelpunkt dieses Berichts.

Grundsätzlich unterscheidet man zwischen Ausfällen, die durch das Nichtauffinden von Per- sonen, durch Nichterreichbarkeit und durch Verweigerung entstanden sind (Blom et al. 2008, Lepkowski und Couper 2002, Schnell 2012, Schnell et al. 2008). In ALWA hat die erste Quel- le von Selektivität besondere Bedeutung, denn hier wurde eine Einwohnermeldeamtsstich- probe gezogen, die ausschließlich Adressen enthält, die Teilnehmer wurden jedoch telefo- nisch befragt. Dazu mussten Telefonnummern für alle Adressen recherchiert werden, was nicht immer möglich war. Weitere Ausfallprozesse sind durch das komplexe Studiendesign von ALWA entstanden: Das Ziel der Studie lag darin, sowohl den retrospektiven Lebensver- lauf zu erheben, als auch kognitive Grundkompetenzen zu messen. Aus Kosten- und Belas- tungsründen wurden die retrospektiven Lebensverläufe telefonisch und die Testung kogniti- ver Grundkompetenzen in einem zweiten persönlichen Interview durchgeführt. Für letzteres musste das Einverständnis der Befragten im telefonischen Interview eingeholt werden. Im Fokus dieses Berichts stehen daher nicht nur die üblichen Ausfallprozesse, die aus Nichter- reichbarkeit und Verweigerungen resultieren, sondern auch die Probleme der Telefonnummerrecherche bei Einwohnermeldeamtsstichproben und die zusätzlichen Verzerrun- gen durch das zweistufige Design von ALWA.

Im nächsten Abschnitt werden das Studiendesign von ALWA, die damit verbundenen Aus- fallprozesse und Analysemöglichkeiten zur Beschreibung von Selektivität dargestellt. Die empirischen Ergebnisse zum Nonresponse-Bias von ALWA sind in zwei Abschnitte aufge- teilt, einen zur telefonischen Befragung und einen zur mündlichen Befragung.¹ Im Abschnitt zur CATI-Befragung wird zunächst das Ausmaß der Selektivität anhand eines Vergleichs mit externen Referenzdaten, dem Mikrozensus 2007, beschrieben. Darauf folgt eine Analyse möglicher Quellen des Bias anhand von Merkmalen, die die Einwohnermeldeämter für alle Stichprobenmitglieder geliefert haben, sowie einer zusätzlich durchgeführten Nonresponse-

¹ Selektivität im ALWA-ADIAB Datensatz ist dagegen nicht Gegenstand dieses Methodenreports. Diese wird ausführlich analysiert in Antoni (2013).

Studie. Im Abschnitt zur PAPI-Befragung wird die zusätzliche Selektivität, die durch Nicht-Einwilligung der CATI-Befragten zur Messung der kognitiven Grundkompetenzen, durch Nichterreichbarkeit im persönlichen Interview und durch Verweigerung des persönlichen Interviews entstanden ist, multivariat überprüft. Dabei kann auf detaillierte Informationen zu Charakteristika der Befragten, Merkmale des CATI-Interviews, sowie Merkmale der CATI- und der PAPI-Interviewer zurückgegriffen werden. Der Beitrag schließt mit einem Fazit zum Studiendesign von ALWA und Anregungen zur Gestaltung künftiger Befragungen.

2 Studiendesign und Ausfallprozesse

Die Grundgesamtheit von ALWA umfasst alle in Deutschland lebenden Personen in Privathaushalten, die zwischen 1956 und 1988 geboren sind. Die Stichprobenziehung erfolgte zweistufig (vgl. dazu genauer Gilberg et al. 2011): Zunächst wurde aus den von den Statistischen Landesämtern zur Verfügung gestellten gemeindestatistischen Daten eine Gemeindestichprobe gezogen. Dabei erfolgte die Ziehung der Gemeinden (Primary Sampling Units) mittels pps (= probability proportional to size), d.h. die Auswahlwahrscheinlichkeit der Gemeinden war proportional zur fortgeschriebenen Wohnbevölkerung in der Zielgruppe. Es wurden 281 Sample Points in 250 Gemeinden ermittelt.² Die Ziehung der Personenadressen der Zielgruppe (Secondary Sampling Units) in den ausgewählten Gemeinden erfolgte über eine systematische Zufallsauswahl (Intervallziehung) aus den Einwohnermeldeamtsregistern. Ausgehend von einer zufälligen Startadresse wurden die übrigen Adressen über eine feste Schrittweite ausgewählt. Alle gemeldeten Einwohner der Geburtsjahrgänge 1956 bis 1988, deren Hauptwohnsitz zum Zeitpunkt der Ziehung in einer der ausgewählten Gemeinden lag, wurden in die Auswahl einbezogen. Insgesamt umfasste die Bruttostichprobe knapp 133.500 Adressen.

Nach Übermittlung der Adressen durch die Gemeinden erfolgte eine Telefonnummernrecherche über die gesamte Adressstichprobe mittels aktueller und älterer digitaler Telefonnummernverzeichnisse. Dabei konnten nur für rund 53 Prozent der Adressen Telefonnummern ermittelt werden, deutlich weniger, als bei Festlegung des Studiendesigns zwei Jahre zuvor noch anvisiert worden war. In diesem Zeitraum stieg die Verbreitung von Mobiltelefonen rasant an, und immer mehr Menschen in Deutschland verzichteten ganz auf ein Festnetztelefon (Busse und Fuchs 2012, Schnell 2012). Für Mobiltelefonnummern gab (und gibt es bis heute) allerdings keine vollständigen Verzeichnisse. Von den insgesamt knapp 70.200 Adressen mit ermittelter Telefonnummer wurde nur etwa ein Drittel im Feld eingesetzt, so dass sich die Brutto-Einsatzstichprobe auf 22.656 Personen beläuft (Tabelle 1).

² Aufgrund der vergleichsweise großen Zahl an PSUs ist der erwartete Designeffekt infolge der Clusterauswahl sehr niedrig, d.h. es kann von durchschnittlich kleinen Intraclusterkorrelationskoeffizienten (ICC), mithin einem durchschnittlich kleinen Klumpungsfaktor, ausgegangen werden.

Tabelle 1: Stichprobenrealisierung

	N		Prozent	
Adressstichprobe	133.451	100,0		
Adressen ohne Telefonnummer	63.268	47,4		
Telefonstichprobe	70.183	52,6	100,0	
Brutto-Einsatzstichprobe	22,656		32,3	100,0
CATI: Neutrale Ausfälle	2.797			12,3
CATI: Nicht Erreichte	1.094			4,8
CATI: Verweigerungen	8.361			36,9
CATI: abgebrochene Interviews	64			0,3
CATI: realisierte Interviews	10.404			45,9
CATI: deutschsprachige Interviews	10.177			44,9
PAPI: Teilnahmebereite	5.889		57,9	100,0
PAPI: Neutrale Ausfälle	110			1,9
PAPI: Nicht Erreichte	503			8,5
PAPI: Verweigerungen	1.239			21,0
PAPI: abgebrochene Interviews	47			0,8
PAPI: realisierte Interviews	3.990			67,8
PAPI: korrekt durchgeführte Interviews	3.980			67,6

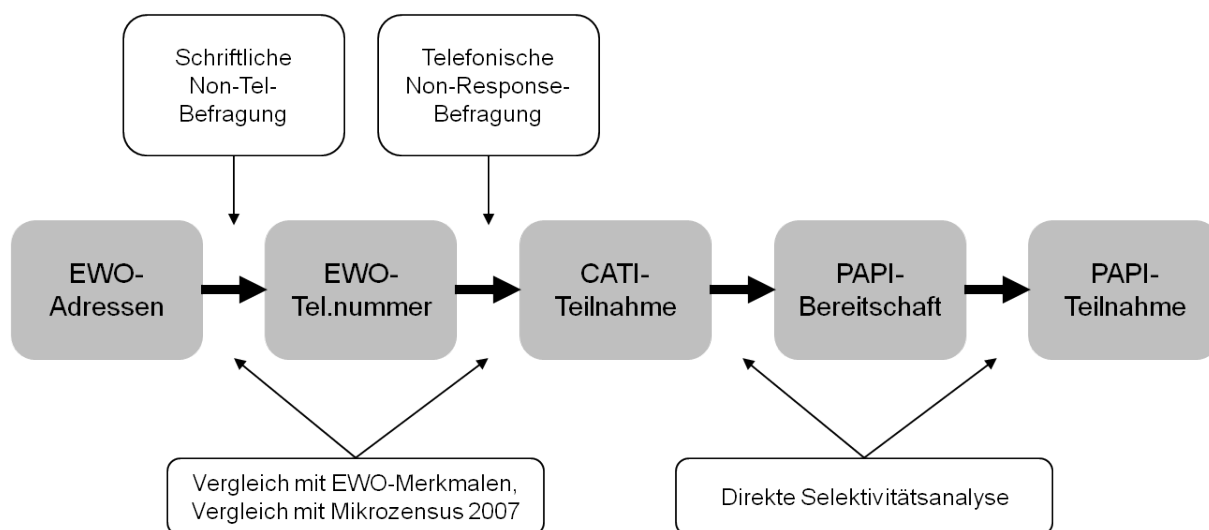
Quelle: Gilberg et al. (2011)

Zwischen Adress- und Telefonnummernstichprobe erfolgten also die ersten Ausfallprozesse der Studie (Abbildung 1). Um die Zufälligkeit dieser Ausfälle zu überprüfen, können erstens die Daten, die von den Einwohnermeldeämtern zur Verfügung gestellt wurden, in beiden Gruppen – den Personen ohne und mit ermittelten Nummern – verglichen werden. Zum zweiten wurde 4.205 zufällig ausgewählten Personen ohne ermittelte Telefonnummer ein kurzer schriftlicher Fragebogen geschickt, um deren Selektivität anhand weiterer Merkmale überprüfen zu können. Leider war der Rücklauf dieser Befragung sehr gering: Nur 375 Personen, knapp neun Prozent, schickten den Fragebogen ausgefüllt zurück, so dass kaum Rückschlüsse aus dieser Befragung gezogen werden können (vgl. Abschnitt 3.3).

Der zweite Ausfallprozess der Studie entstand bei der Realisierung der computergestützten telefonischen Befragung. Um eine hohe Ausschöpfung zu sichern, wurde versucht alle Personen, die zunächst nicht an der Studie teilnehmen wollten, aber nicht grundsätzlich verweigerten, durch besonders geschulte Interviewer zu konvertieren. Außerdem wurden attraktive Sachpreise unter allen Teilnehmern verlost; diese Verlosung wurde im Anschreiben angekündigt. Insgesamt nahmen 10.404 Personen an den CATI-Interviews teil; gemessen an der Einsatzstichprobe entspricht das einer Ausschöpfung von knapp 46 Prozent. Die neutralen Ausfälle (ungültige Telefonnummern, nicht befragbare Personen, Verstorbene, Personen außerhalb der Grundgesamtheit) beliefen sich auf rund 12 Prozent; der größte Teil davon waren ungültige Telefonnummern, also beispielsweise Nummern ohne Anschluss, Faxanschlüsse oder Nummern, unter denen andere Personen zu erreichen waren. Knapp fünf Prozent der Personen in der Einsatzstichprobe wurden in der Feldzeit nicht erreicht und ca. 37 Prozent verweigerten die Teilnahme. Interviewabbrüche kamen mit 0,3 Prozent nur sehr selten vor. Um den Nonresponse-Bias genauer zu untersuchen, wurde eine Zufallsauswahl

von 5.946 Verweigerern im Telefoninterview darum gebeten, einige wenige Nonresponse-Fragen zu beantworten. An dieser Kurzbefragung nahmen 1.767 Personen teil, eine Ausschöpfung von knapp 30 Prozent (vgl. Abschnitt 3.3; zu deskriptiven Ergebnissen dieser Befragung vgl. auch Gilberg et al. 2011).

Abbildung 1: Ausfallprozesse und Analysemöglichkeiten



Die Lese- und Rechenkompetenzen der ALWA-Teilnehmer wurden mit einem insgesamt etwa einstündigen Test ermittelt, der vom Educational Testing Service, USA (ETS) als Papier- und Bleistift-Fragebögen (PAPI) speziell für Erwachsene entwickelt und vom IAB für die deutsche Sprache adaptiert worden war. Dieser Test wurde im Rahmen persönlicher Interviews durchgeführt (vgl. dazu genauer Kleinert et al. 2012). Um die Kosten der Studie im Rahmen zu halten und die Befragten nicht über Gebühr zu belasten, wurden alle übrigen Bestandteile der Befragung (demographische und persönliche Angaben, Lebensverlauf) in das vorangehende telefonische Interview gepackt. Die Einwilligung in ein persönliches Folgeinterview musste aus Datenschutzgründen explizit eingeholt werden. Am Ende des CATI-Interviews wurden daher alle deutschsprachigen Teilnehmer (n=10.177) nach ihrer Bereitschaft gefragt, an einem persönlichen Interview teilzunehmen, das im Abstand von maximal vier Wochen bei ihnen zuhause stattfinden würde. Für die Teilnahme an diesem Interview wurde ein fixes Incentive von 10 EUR angekündigt. 5.889 Personen bzw. knapp 58 Prozent aller CATI-Teilnehmer erklärten sich zur Teilnahme am persönlichen Interview bereit. Hier liegt der dritte Ausfallprozess.

Der vierte und letzte Ausfallprozess fand schließlich bei der Realisierung der PAPI-Befragung statt. 3.990 Personen nahmen an diesem Teil der Studie teil; bezogen auf die Einsatzstichprobe aller teilnahmebereiten Personen entspricht das einer Ausschöpfungsrate von rund 68 Prozent. Neutrale Ausfälle gab es aufgrund des zweistufigen Designs in nur sehr geringem Ausmaß (ca. 2 Prozent der Einsatzstichprobe) und auch die Verweigerungen

waren deutlich seltener als in der CATI-Befragung (21 Prozent). Allerdings waren mehr teilnahmebereite Personen als bei der telefonischen Befragung in der Feldzeit nicht erreichbar, nämlich etwa 9 Prozent der Einsatzstichprobe. Die Rate der nicht Erreichbaren ist bei persönlichen Befragungen grundsätzlich höher als bei telefonischen, weil hier weniger Kontakte realisiert werden können. Bei ALWA lag die Kontakthäufigkeit im CATI-Interview im Schnitt bei 13,3 Anrufen. Im PAPI-Interview wurden dagegen im Schnitt nur 3,5 Kontaktversuche pro Zielperson realisiert.

Zu klären gilt nun, wie stark die unterschiedlichen Ausfallprozesse der ALWA-Studie zu Verzerrungen in der Verteilung der Studienteilnehmer gegenüber der Grundgesamtheit geführt haben. Dies wird in den beiden folgenden Abschnitten für die CATI- und die PAPI-Befragung getrennt dargestellt.

3 Selektivität in der CATI-Befragung

3.1 Ausmaß der Selektivität: Vergleich mit externen Referenzdaten

Erste Anhaltspunkte darüber, wie stark und in welcher Hinsicht die Teilnehmer der ALWA-Studie verzerrt sein könnten, erhält man, wenn man die ALWA-Daten mit externen Referenzdaten vergleicht, die möglichst unverzerrt sein sollten. Dafür bietet sich für die Grundgesamtheit von ALWA nur eine Datenquelle an, der jährlich erhobene Mikrozensus des Statistischen Bundesamtes, denn nur bei dieser Befragung besteht für die Beteiligten eine Auskunftspflicht. Alle hier dargestellten Analysen wurden mit dem Scientific Use File des Mikrozensus 2007 durchgeführt. Die Stichprobe wurde auf die Grundgesamtheit von ALWA beschränkt, d.h. auf Personen der Jahrgänge 1956 bis 1988 in Privathaushalten am Hauptwohnsitz; alle Analysen wurden gewichtet und auf die Bevölkerung hochgerechnet.

Einschränkend muss bei diesem Vergleich bemerkt werden, dass beide Befragungen in weiten Teilen unterschiedliche Philosophien der Datenerhebung aufweisen: Erstens beruhen die Mikrozensus-Angaben auf Querschnittangaben, während ALWA eine modularisierte Lebensverlauserhebung ist (Antoni et al. 2010). Unterschiede in den Verteilungen zur derzeitigen Wohnregion, zu den höchsten Bildungsabschlüssen und zum aktuellen Erwerbsstatus können daher auch dadurch entstanden sein, dass diese Informationen auf völlig anderem Wege erhoben wurden. Zum zweiten ist der Mikrozensus eine Haushalts-, ALWA hingegen eine Individualbefragung. Im Mikrozensus ist es daher relativ schwierig, die Haushaltszusammensetzung für eine einzelne Person zu ermitteln und entsprechende Angaben im SUF beruhen auf einem ausgefeilten Familienkonzept (Nöthen 2005, Lengerer et al. 2005), während Informationen dazu in ALWA direkt erfasst wurden. So sind beispielsweise in den ALWA-Daten auch nicht-ledige Kinder in Haushalten der Befragten enthalten, während diese in den Mikrozensus-Daten ausgeschlossen sind, weil sie eigene Haushalte bilden. Auch dadurch können Unterschiede in den Verteilungen entstanden sein.

In Tabelle 2 sind die Randverteilungen von Grundmerkmalen beider Populationen in Form von Mittelwerten und Standardabweichungen aufgeführt. Außerdem wurde mittels T-Tests

geprüft, ob sich die Mittelwerte beider Gruppen signifikant voneinander unterscheiden. Diese Werte zeigen, dass ALWA in drei Merkmalen deutlich von der Referenzstatistik abweicht:

- Erstens sind Einwanderer und noch stärker Personen ohne deutsche Staatsangehörigkeit in ALWA stark unterrepräsentiert.

Tabelle 2: Randverteilungen von Mikrozensus 2007 und ALWA

Dimension	Merkmal	Mikrozensus		ALWA		T-Test
		MW	SD	MW	SD	
Demographie	Im Ausland geboren	0.172	0.378	0.104	0.305	21.98 ***
	Nichtdeutsche/r	0.114	0.318	0.040	0.195	35.76 ***
	Wohnhaft in neuen Ländern	0.205	0.404	0.184	0.387	5.48 ***
	Weiblich	0.494	0.500	0.507	0.500	-2.53 *
Kohorte	1956-1965	0.352	0.478	0.431	0.495	-15.83 ***
	1966-1975	0.309	0.462	0.267	0.442	9.48 ***
	1976-1980	0.129	0.335	0.087	0.282	14.64 ***
	1981-1985	0.131	0.337	0.108	0.310	7.32 ***
	1986-1988	0.080	0.271	0.108	0.311	-9.21 ***
Bildung	Max. Mittl. Reife, o. Ausbildung	0.185	0.388	0.097	0.296	28.36 ***
	Max. Hauptschule, Ausbildung	0.188	0.390	0.136	0.343	14.29 ***
	Mittlere Reife, Ausbildung	0.303	0.460	0.310	0.462	-1.35
	Abitur	0.188	0.391	0.231	0.421	-9.86 ***
	Fachhochschule	0.047	0.211	0.071	0.257	-9.38 ***
	Universität	0.089	0.285	0.155	0.362	-17.65 ***
Erwerbsstatus	Erwerbstätig	0.734	0.442	0.760	0.427	-5.92 ***
	Arbeitssuchend	0.085	0.279	0.055	0.227	13.02 ***
	In Bildung	0.090	0.286	0.107	0.309	-5.52 ***
	Sonstiges	0.091	0.288	0.079	0.269	4.57 ***
Arbeitszeit	Bis 12 Wo.std.	0.075	0.263	0.083	0.276	-2.60 **
	13-20 Wo.std.	0.094	0.291	0.099	0.299	-1.54
	21-34 Wo.std.	0.098	0.298	0.095	0.293	1.00
	35-42 Wo.std.	0.586	0.493	0.597	0.491	-1.98 *
	43 Wo.std. und mehr	0.148	0.355	0.126	0.332	5.56 ***
Haushalt und Familie	Haushaltsgröße	2.836	1.341	3.088	1.322	-18.92 ***
	Alleinlebend	0.193	0.395	0.137	0.344	15.90 ***
	Mit Partner, ohne Kinder	0.175	0.380	0.157	0.363	5.11 ***
	Ohne Partner, mit Kindern	0.050	0.218	0.044	0.205	2.92 **
	Mit Partner und Kindern	0.407	0.491	0.470	0.499	-12.47 ***
	Bei Eltern	0.150	0.358	0.172	0.378	-5.77 ***
	Mit anderen Personen	0.024	0.152	0.020	0.139	2.87 **
Alter des jüngsten Kindes im HH	9.113	6.184	10.074	5.891	-11.46 ***	

Mittelwerte, Standardabweichungen und T-Tests; Signifikanzniveaus: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Quellen: ALWA (SUF); Mikrozensus 2007 (SUF), hochgerechnet auf Bevölkerung in Tsd.

- Zweitens weist ALWA einen starken Bildungsbias auf: Personen mit niedriger Bildung sind unterrepräsentiert, Akademiker hingegen überrepräsentiert. Beide Untererfassungen wurden auch in vielen anderen telefonischen Befragungen beobachtet (für Bildung vgl. Fitzgerald et al. 1998, Koch 1998, Lepkowski und Couper 2002, Schnell 1997, für Migranten vgl. Koch 1997). Sie werden meist mit Verständnisproblemen, Misstrauen und einem geringen Interesse an Befragungen erklärt. Das geringe Interesse an der Befragung kann

sich in den unteren Bildungsgruppen insbesondere dadurch verstärkt haben, dass im Studentitel von ALWA das Wort „Lernen“ erscheint.

- Drittens sind die jüngsten und ältesten Geburtskohorten in ALWA überrepräsentiert, die mittleren dagegen unterrepräsentiert. Am stärksten trifft das die Geburtsjahrgänge 1976-1980, die zum Zeitpunkt der Befragung zwischen 27 und 32 Jahre alt waren. Dieser Bias geht vermutlich vor allem auf Altersunterschiede in punkto Mobilität und das Vorhandensein von Festnetztelefonen zurück. In dieser Altersspanne leben die meisten jungen Erwachsenen nicht mehr im Elternhaushalt und haben oftmals selbst noch keine eigene Familie gegründet. Aufgrund der frühen Phase ihrer Berufslaufbahn sind sie räumlich mobil und haben oft keinen gelisteten Festnetzanschluss (Busse und Fuchs 2012, Lepkowski und Couper 2002).

Hinsichtlich weiterer Merkmale unterscheidet sich ALWA nur geringfügig von der Referenzstatistik. So sind Frauen in ALWA geringfügig überrepräsentiert. Gleiches gilt für Befragte in den alten Bundesländern, was meist damit erklärt wird, dass dort die Telefonverzeichnisse weitaus mehr Nummern enthalten als in den neuen Ländern (Schnell 2012). Unterschiede zeigen sich auch im Erwerbsstatus. Hier sind vor allem Arbeitslose, aber auch Inaktive, die sich nicht im Bildungssystem befinden, unterrepräsentiert. Diese Abweichung ist allerdings mit Vorsicht zu interpretieren. Einerseits könnte sie auf die Bildungsverzerrung von ALWA zurückgehen, denn niedrig Gebildete sind häufiger arbeitslos als höher Gebildete. Andererseits kann sie durch die völlig unterschiedliche Erfassung von Arbeitslosigkeit in beiden Datenquellen zu erklären sein.

Eine weitere Verzerrung von Befragungspopulationen entsteht in der Regel dadurch, dass Befragte, die zeitlich sehr eingespannt sind, seltener bereit sind an langen Befragungen teilzunehmen bzw. nicht erreicht werden (Koch 1998, Schnell 1997, Watson und Wooden 2009). Indikatoren dafür sind neben dem Lebensalter die Arbeitszeit, die Haushaltszusammensetzung und das Alter des jüngsten Kindes. Bei ALWA zeigt sich, dass Personen mit einer Wochenarbeitszeit von mehr als 42 Stunden im Vergleich zum Mikrozensus untererfasst sind. Gleiches gilt für Alleinstehende, Paare ohne Kinder und Alleinerziehende – alles Gruppen, die oft Vollzeit arbeiten und zeitlich sehr stark eingespannt sind. Sie sind daher schwerer erreichbar als Paare mit Kindern. Auch das Alter des jüngsten Kindes (bei Personen mit Kindern im Haushalt) ist in ALWA im Schnitt höher als im Mikrozensus. Diese Abweichung kann allerdings auch damit erklärt werden, dass im Mikrozensus nur ledige Kinder im Haushalt erfasst sind.

Der Vergleich mit dem Mikrozensus lässt zwar Rückschlüsse über die Stärke und Ausprägung des Bias von ALWA zu, sagt aber nichts darüber aus, durch welche Ausfallprozesse diese erzeugt wurden: durch das Nicht-Auffinden von Telefonnummern, durch Nichterreichbarkeit oder durch Nonresponse. Dazu werden im Folgenden diese drei Gruppen von Nicht-Teilnehmern von ALWA mit ihren jeweiligen Kontrastgruppen (Personen mit aufgefundener Telefonnummer, Erreichte, Teilnehmer) anhand der Merkmale verglichen, die in den Einwohnermeldedaten enthalten sind.

3.2 Quellen der Selektivität I: Analyse der EWO-Merkmale

Ein wesentlicher Vorteil von Einwohnermeldeamts- gegenüber Telefonstichproben besteht darin, dass einige Merkmale der Personen in der Bruttostichprobe bekannt sind, so dass ein Vergleich von unterschiedlichen Ausfallgruppen mit Nicht-Ausfallgruppen anhand dieser Daten vorgenommen werden kann (Lynn & Clarke 2002, Heerwegh et al. 2007, Kreuter et al. 2007).³ Über die Grundgesamtheit in Telefonstichproben ist dagegen nichts bekannt. Auch hier sind systematische Verzerrungen aufgrund fehlender Festnetzanschlüsse oder Geheimnummern naheliegend, es gibt jedoch keine Möglichkeit zu überprüfen, wie stark diese Einschränkungen in welchen Bevölkerungsgruppen verbreitet sind.

Für die Analyse der Selektivität mittels der Einwohnermeldedaten sind folgende Merkmale in der Bruttostichprobe vorhanden: deutsche Staatsangehörigkeit, Geschlecht, Altersgruppe, Bundesland, politische Ortsgröße und die BIK-Region, eine Kennziffer für die Ortsgröße unter Einbezug der Verstädterung des Umlandes.⁴ Mit diesen Informationen kann mittels logistischer Regressionen der Effekt dieser Merkmale auf unterschiedliche Wahrscheinlichkeiten geschätzt werden: erstens die Wahrscheinlichkeit einer recherchierten Telefonnummer in der Adressstichprobe, zweitens die Wahrscheinlichkeit, eine Person in der bereinigten Einsatzstichprobe (ohne neutrale Ausfälle) zu erreichen, und drittens die Wahrscheinlichkeit der Interviewteilnahme im Vergleich zu einer Verweigerung (Tabelle 3).

Nichtdeutsche haben eine weitaus geringere Chance als Deutsche, eine recherchierbare Telefonnummer zu ihrer Adresse zu besitzen (Granato 1999) und sie nehmen mit geringerer Wahrscheinlichkeit an der Befragung teil. Die Erreichbarkeit unterscheidet sich dagegen nicht nach Nationalität. Hauptquelle für die starke Unterrepräsentanz von Migranten und Ausländern in ALWA ist somit die Telefonnummernrecherche.

Auch bei Frauen ist die Wahrscheinlichkeit einer recherchierten Telefonnummer etwas geringer als bei Männern. Umgekehrt sind Frauen allerdings besser erreichbar und nehmen mit einer höheren Wahrscheinlichkeit an der Befragung teil (Brehm 1993, Goyder 1987, Lepkowski und Cooper 2002, Schnell 1997, Watson und Wooden 2009). Diese Effekte heben in der Summe die niedrigere Auffindungswahrscheinlichkeit von Frauen auf: nach dem Mikrozensus-Vergleich sind sie in ALWA geringfügig überrepräsentiert.

Das Alter weist einen kurvilinearen Effekt auf die Wahrscheinlichkeit einer Telefonnummer und auf die Erreichbarkeit der ausgewählten Personen auf. Telefonnummern sind bei den 26- bis 30jährigen am wenigsten verbreitet. Bei den Altersgruppen darunter und darüber steigt die Wahrscheinlichkeit einer aufgefundenen Telefonnummer dann wieder an; bei der Referenzgruppe der 46- bis 51jährigen liegt sie am höchsten. Dieser Effekt deutet darauf hin,

³ Für eine deutlich weitergehende Anwendung von administrativen Daten, nämlich die gleichzeitige Verwendung mit Kontaktverlaufsdaten der Befragung, vgl. Kreuter et al. (2010).

⁴ Die Merkmale Geschlecht, Altersgruppe und Nationalität liegen allerdings nicht für alle Teilnehmer der Bruttostichprobe vor. Das Merkmal Geschlecht fehlt bei etwa 0,5% der Personen in der Stichprobe, die Altersgruppe bei 5% und die Nationalität bei 13%.

dass Jüngere häufiger als Ältere ausschließlich Mobiltelefone benutzen oder zumindest keine Festnetznummer in einem Verzeichnis eingetragen haben (vgl. auch Busse und Fuchs 2012). Dies gilt besonders dann, wenn sie bei ihren Eltern ausgezogen sind. Nichterreichbarkeit stellt bei den 26- bis 35jährigen das größte Problem dar (vgl. dazu auch Watson und Wooden 2009). Auch hier liegt die Wahrscheinlichkeit in der Referenzgruppe am höchsten, die Alterseffekte auf die Erreichbarkeit sind allerdings in Stärke und Gewicht nicht ganz so ausgeprägt wie die auf das Auffinden einer Telefonnummer.

Tabelle 3: Wahrscheinlichkeit einer aufgefundenen Telefonnummer, der Erreichbarkeit und der CATI-Teilnahme anhand von EWO-Merkmalen

	Telefonnummer (vs. nicht gefunden)		Erreicht (vs. nicht erreicht)		Interviewteilnahme (vs. Verweigerung)	
	OR	Z	OR	Z	OR	Z
<i>deutsch (Ref.gruppe)</i>	1.000		1.000		1.000	
nicht deutsch	0.341***	(-50.31)	0.902	(-0.70)	0.805**	(-2.84)
keine Angabe	0.819***	(-11.57)	0.862	(-1.59)	1.036	(0.78)
<i>männlich (Ref.gruppe)</i>	1.000		1.000		1.000	
weiblich	0.884***	(-10.81)	1.262***	(3.68)	1.180***	(5.57)
keine Angabe	0.801*	(-2.41)	1.014	(0.02)	1.708	(1.90)
17-20 Jahre	0.688***	(-13.41)	0.714*	(-2.39)	1.933***	(9.57)
21-25 Jahre	0.494***	(-33.51)	0.594***	(-4.55)	1.526***	(7.50)
26-30 Jahre	0.417***	(-41.47)	0.531***	(-5.37)	1.134*	(2.12)
31-35 Jahre	0.504***	(-32.45)	0.529***	(-5.56)	0.920	(-1.46)
36-40 Jahre	0.670***	(-20.58)	0.753**	(-2.63)	0.998	(-0.05)
41-45 Jahre	0.821***	(-10.36)	0.855	(-1.51)	1.062	(1.37)
<i>46-51 Jahre (Ref.gruppe)</i>	1.000		1.000		1.000	
keine Angabe	0.892***	(-3.87)	0.861	(-0.87)	0.759***	(-3.74)
neue Bundesländer	0.566***	(-38.60)	1.091	(1.04)	0.924*	(-2.02)
< 2.000 EW	2.008***	(15.58)	1.433	(1.48)	1.041	(0.38)
2.000 bis < 5.000 EW	2.149***	(20.22)	1.207	(1.01)	0.918	(-0.96)
5.000 bis < 20.000 EW	2.261***	(34.39)	1.626***	(3.56)	0.818***	(-3.43)
20.000 bis < 50.000 EW	1.851***	(31.05)	1.351**	(2.79)	0.934	(-1.36)
50.000 bis < 100.000 EW	1.536***	(21.15)	1.165	(1.38)	0.895*	(-2.05)
100.000 bis < 500.000 EW	1.317***	(19.50)	1.373***	(4.09)	0.986	(-0.39)
<i>ab 500.000 EW (Ref.gruppe)</i>	1.000		1.000		1.000	
N	133451		19859		18765	
Chi ²	10315***		100***		254***	
Pseudo R ²	0.056		0.012		0.010	

Logistische Regressionen, Odds Ratios und Z-Statistik; Signifikanzniveaus: * p<0.05, ** p<0.01, *** p<0.001.

Quelle: ALWA-Methodendaten

Der Nonresponse bezüglich des Alters ist ebenfalls nicht neutral: In den jüngeren Altersgruppen ist die Wahrscheinlichkeit höher als in der Referenzgruppe, am Interview teilzunehmen statt die Teilnahme zu verweigern. Dies gilt besonders stark für die Gruppe der 17- bis 20jährigen. Ab der Altersgruppe der 31- bis 35jährigen unterscheidet sich die Teilnahme-wahrscheinlichkeit nicht mehr nach dem Alter. Diese Verzerrung kann mit zwei Altersdiffe-

renzen erklärt werden: Erstens hat das Thema Bildung und Lernen für Jüngere eine größere Relevanz als für Ältere, weil für sie der Lebensabschnitt in Bildung und Ausbildung entweder noch aktuell ist oder noch nicht weit zurückliegt. Und zweitens hatten sie zum Zeitpunkt der Befragung noch einen kürzeren Lebensverlauf als Ältere und waren vielleicht auch aus diesem Grund eher bereit, an der Befragung teilzunehmen, weil sie von ihnen als weniger zeitintensiv eingeschätzt wurde.

ALWA ist nicht nur nach Merkmalen der Personen in der Stichprobe, sondern auch nach regionalen Merkmalen verzerrt: In den neuen Bundesländern sind Festnetzanschlüsse, wie angenommen, deutlich weniger verbreitet als in den alten und die Wahrscheinlichkeit einer aufgefundenen Telefonnummer damit niedriger. Dazu kommt, dass die Teilnahmewahrscheinlichkeit am Interview hier ebenfalls etwas geringer ist. Und schließlich sind Festnetznummern umso stärker verbreitet, je ländlicher eine Region ist. Personen aus Großstädten hatten damit eine deutlich niedrigere Chance, in die Telefonstichprobe zu gelangen, als Personen aus anderen Regionen. Erreichbarkeit und Teilnahmebereitschaft am Interview unterscheiden sich im Gegensatz dazu nicht systematisch nach dem Urbanisierungsgrad. Dem entspricht, dass sich in der Literatur dazu widersprüchliche Ergebnisse finden. Grooves et al. (1992) und Schnell (1997) berichten ebenfalls keinen Zusammenhang; Schräpler (2000) findet eine Abnahme der Teilnahmebereitschaft mit zunehmender Ortsgröße.

Fasst man die Analysen auf Basis der Einwohnermeldedaten zusammen, dann zeigt sich, dass Verzerrungen vor allem durch die Telefonnummernrecherche entstanden sind. Besonders einflussreich sind hier Nationalität, Alter, alte und neue Länder sowie der Grad der Verstädterung. Die Pseudo-R-Quadrat-Maße der Modelle zur Erreichbarkeit und Teilnahmewahrscheinlichkeit sind im Vergleich dazu deutlich niedriger und die Effektstärken der einbezogenen Merkmale sind gering. Die Erreichbarkeit der Teilnehmer der ALWA-Studie wird hauptsächlich vom Geschlecht und vom Alter beeinflusst, ihre Teilnahmewahrscheinlichkeit von der Nationalität, vom Geschlecht und vom Alter.

3.3 Exkurs: Zur Repräsentativität der Zusatzbefragungen

Aufschluss über weitere Verzerrungen der Studie, insbesondere hinsichtlich des Bildungsstandes der Befragten, sollten zwei Zusatzbefragungen von Nicht-Teilnehmern geben: erstens eine schriftliche Kurzbefragung von Personen ohne ermittelte Telefonnummer (Non-Tel-Befragung) und zweitens eine telefonische Kurzbefragung von Verweigerern (Nonresponse-Befragung).

Will man aus einer Non-Response-Befragung auf Verzerrungen einer Studie schließen, dann muss sie eine grundsätzliche Voraussetzung erfüllen: Die Personen, die an dieser Befragung teilnehmen, müssen der Gesamtheit von Verweigerern gleichen – sie müssen also eine Zufallsauswahl aus dieser Gruppe bilden (Lynn et al. 2002, Schräpler 2000). Ist das nicht der Fall, dann sind die Ergebnisse einer solchen Befragung wertlos. Gleiches gilt natürlich auch für Non-Tel-Befragungen. Diese Voraussetzung kann zumindest hinsichtlich der Merkmale geprüft werden, die in den Einwohnermeldedaten vorliegen. Dazu wurden wieder logistische

Regressionen geschätzt, und zwar einmal die Wahrscheinlichkeit für eine Teilnahme an der Non-Tel-Befragung auf Basis der gesamten Stichprobe ohne ermittelte Telefonnummer und einmal die Wahrscheinlichkeit für eine Teilnahme an der Non-Response-Befragung auf Basis aller Verweigerer (Tabelle 4). Ergeben sich signifikante Abweichungen, dann ist die Nullhypothese einer zufälligen Auswahl abzulehnen, und die Studien sollten nicht inhaltlich ausgewertet werden.

Tabelle 4: Wahrscheinlichkeit für die Teilnahme an den Zusatzbefragungen anhand von EWO-Merkmalen

	Non-Tel		Non-Response	
	OR	Z	OR	Z
<i>deutsch (Ref.gruppe)</i>				
nicht deutsch	0.403***	(-4.16)	1.227	(1.62)
keine Angabe	0.876	(-0.84)	0.882	(-1.50)
<i>männlich (Ref.gruppe)</i>				
weiblich	1.577***	(4.26)	0.979	(-0.40)
keine Angabe	0.793	(-0.23)	1.213	(0.37)
17-20 Jahre	0.983	(-0.07)	1.117	(0.80)
21-25 Jahre	0.652*	(-2.14)	1.061	(0.54)
26-30 Jahre	0.807	(-1.17)	1.132	(1.13)
31-35 Jahre	0.824	(-1.02)	1.228*	(2.06)
36-40 Jahre	0.756	(-1.51)	1.169	(1.82)
41-45 Jahre	1.001	(0.01)	1.245**	(2.76)
<i>46-51 Jahre (Ref.gruppe)</i>				
keine Angabe	0.911	(-0.33)	1.059	(0.45)
<i>neue Bundesländer</i>				
< 2.000 EW	2.593**	(2.96)	1.393	(1.78)
2.000 bis < 5.000 EW	1.335	(0.79)	1.512**	(2.81)
5.000 bis < 20.000 EW	1.721**	(2.58)	1.117	(1.08)
20.000 bis < 50.000 EW	1.401	(1.82)	0.963	(-0.41)
50.000 bis < 100.000 EW	1.304	(1.40)	1.058	(0.58)
100.000 bis < 500.000 EW	1.348*	(2.32)	1.028	(0.40)
<i>ab 500.000 EW (Ref.gruppe)</i>				
N	63268		8361	
Chi ²	74***		39**	
Pseudo R ²	0.016		0.005	

Logistische Regressionen, Odds ratios und Z-Statistik; Signifikanzniveaus: * p<0.05, ** p<0.01, *** p<0.001.

Quelle: ALWA-Methodendaten

Die Ergebnisse der Schätzung zum Non-Tel-Rücklauf zeigen, dass die Teilnehmer im Vergleich zu allen Personen ohne aufgefundene Telefonnummer mehrfach verzerrt sind: Insbesondere Nichtdeutsche und Männer, aber auch Personen im Alter zwischen 21 und 25 Jahren, Personen aus den neuen Ländern und Personen in großstädtischen Regionen sind hier unterrepräsentiert. Bis auf den Geschlechtereffekt gleichen diese Effekte in ihrer Richtung denjenigen der Verzerrung der Telefonstichprobe im Vergleich zur Gesamtstichprobe. Bedenkt man die geringe Fallzahl des Rücklaufs (n=375), verwundern diese Abweichungen

kaum. Damit können aus der Non-Tel-Befragung keinerlei Rückschlüsse auf weitere Verzerrungen gezogen werden; sie wird nicht weiter ausgewertet.

Die Ergebnisse zur Non-Response-Befragung sehen vergleichsweise positiver aus. Hier zeigt sich nur ein signifikanter negativer Effekt: Verweigerer aus den neuen Ländern nehmen mit geringerer Wahrscheinlichkeit an der Zusatzbefragung teil als Verweigerer aus den alten Ländern. Daneben sind die Effekte einiger Alters- und Einwohnerzahl-Gruppen signifikant; systematische Verzerrungsmuster sind bei beiden Merkmalen allerdings nicht erkennbar. Zusätzlich wurde geprüft, ob sich harte und weiche Verweigerer gleichmäßig über Teilnehmer und Nichtteilnehmer an der Non-Response-Studie verteilen; auch dies ist der Fall.

3.4 Quellen der Selektivität II: Ergebnisse der Non-Response-Befragung

Damit können im nächsten Schritt die Daten der Non-Response-Befragten mit den Daten der CATI-Teilnehmer verglichen werden (Tabelle 5). Da die Non-Response-Befragung sehr kurz sein musste, lässt sich der Vergleich nur für wenige zentrale Merkmale der Personen und ihrer Umgebung durchführen.

Tabelle 5: Randverteilungen von Non-Response-Befragten und CATI-Teilnehmern

Dimension	Merkmal	Non-Response		CATI		T-Test
		MW	SD	MW	SD	
Demographie	Im Ausland geboren	0.130	0.336	0.104	0.305	3.276 **
	Wohnhaft in neuen Ländern	0.164	0.370	0.182	0.386	-1.834
	Weiblich	0.467	0.499	0.507	0.500	-3.090 **
Kohorte	1956-1965	0.447	0.497	0.431	0.495	1.285
	1966-1975	0.320	0.467	0.267	0.442	4.612 ***
	1976-1980	0.086	0.280	0.087	0.282	-0.212
	1981-1985	0.086	0.281	0.108	0.310	-2.709 **
	1986-1988	0.061	0.240	0.108	0.311	-5.985 ***
Bildung	Max. Realschule ohne Ausbildung	0.106	0.308	0.126	0.332	-2.309 *
	Hauptschule mit Ausbildung	0.216	0.411	0.127	0.333	10.027 ***
	Realschule mit Ausbildung	0.338	0.473	0.296	0.456	3.550 ***
	Abitur	0.161	0.367	0.234	0.424	-6.870 ***
	Fachhochschule	0.072	0.259	0.082	0.274	-1.337
	Universität	0.107	0.309	0.136	0.342	-3.287 **
Erwerbsstatus	Erwerbstätig	0.734	0.442	0.744	0.437	-0.874
	Arbeitssuchend	0.040	0.195	0.052	0.221	-2.127 *
	In Bildung	0.068	0.253	0.105	0.307	-4.791 ***
	Sonstiges	0.158	0.365	0.099	0.299	7.375 ***
BIK-Regionen	unter 2.000 EW	0.024	0.152	0.022	0.146	0.489
	2.000 bis u. 5.000 EW	0.041	0.199	0.029	0.168	2.739 **
	5.000 bis u. 20.000 EW	0.092	0.289	0.074	0.262	2.583 **
	20.000 bis u. 50.000 EW	0.113	0.316	0.118	0.322	-0.619
	50.000 bis u. 100.000 EW	0.104	0.306	0.097	0.297	0.870
	100.000 bis u. 500.000 EW	0.320	0.467	0.331	0.471	-0.878
	500.000 und mehr EW	0.306	0.461	0.329	0.470	-1.879

Mittelwerte, Standardabweichungen und T-Tests; Signifikanzniveaus: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Quelle: ALWA-Methodendaten

Im Ergebnis zeigen sich einige deutliche Unterschiede: So ist der Anteil an Befragten, die im Ausland geboren sind, in der Non-Response-Befragung signifikant höher als in der CATI-Befragung. Umgekehrt nahmen signifikant weniger Frauen und weniger Mitglieder aus den beiden jüngsten Kohorten an der Non-Response-Befragung als an der Hauptbefragung teil. Die Ergebnisse des Mikrozensus-Vergleichs werden durch diese Analyse also bestätigt: Verweigerungen werden systematisch vom Migrationshintergrund, Geschlecht und Alter beeinflusst.

Zusätzlich können nun jedoch auch die Bildung und der Erwerbsstatus betrachtet werden, beides Merkmale, für die in den Einwohnermeldedaten keine Informationen vorliegen. In Bezug auf den Bildungsgrad zeigen sich die größten Unterschiede zwischen beiden Gruppen: Personen mit niedriger Schulbildung sind in der Non-Response-Gruppe deutlich häufiger vertreten als in der Gruppe der CATI-Teilnehmer, und Personen mit Abitur bzw. Hochschulabsolventen sind umgekehrt deutlich geringer vertreten. Damit korrespondiert, dass sich Personen, die zum Zeitpunkt der Befragung in Bildungsgängen waren, wesentlich seltener in der Non-Response-Gruppe als in der Gruppe der CATI-Befragten wiederfinden. Insgesamt finden wir damit auch in ALWA den aus vielen Studien bekannten Bildungsbias zugunsten höher Gebildeter.

Vergleicht man die Randverteilungen der beiden Gruppen in Tabelle 5 mit der Randverteilung der Bevölkerung nach Bildung im Mikrozensus (Tabelle 2), dann zeigt sich, dass die Mikrozensus-Anteile bei Haupt- und Realschulabsolventen mit Ausbildung sowie bei Abiturienten in der Mitte zwischen den ALWA-Verweigerern und den ALWA-Teilnehmern liegen. Inhaltlich kann man dieses Ergebnis so deuten, dass die Verzerrungen der ALWA-Teilnehmer in diesen Bildungsgruppen wohl hauptsächlich auf Unterschiede in der Teilnahmebereitschaft zurückgeführt werden können – zumindest wenn man annimmt, dass die Non-Response-Teilnehmer repräsentativ für alle Verweigerer sind. Umgekehrt können die Abweichungen in der Gruppe der Ungelernten und bei den Akademikern nicht alleine auf die Antwortbereitschaft der Zielpersonen in ALWA zurückgeführt werden, denn die erste Gruppe ist auch bei den ALWA-Verweigerern unterrepräsentiert, die zweite überrepräsentiert. Die Quelle für diese Verzerrungen liegt also nicht alleine in Differenzen der Teilnahmebereitschaft, sondern sie ist teilweise auch woanders zu suchen. Unterschiede in der Erreichbarkeit erscheinen hier unwahrscheinlich; der wichtigere Unterschied liegt vermutlich in der Dichte (eingetragener) Festnetzanschlüsse.

Was schließlich den Erwerbsstatus angeht, zeigt sich in der Gruppe der Erwerbstätigen kein Unterschied zwischen Befragten und Verweigerern, sehr wohl aber bei Arbeitssuchenden und Inaktiven. Beide Gruppen finden sich deutlich häufiger in der Gruppe der CATI-Teilnehmer als bei den Non-Response-Befragten, was auf eine Überrepräsentanz in der Studie schließen lässt. Verantwortlich dafür sind vermutlich die größeren zeitlichen Spielräume von Arbeitssuchenden und Inaktiven, die sich nicht nur auf ihre Erreichbarkeit, sondern auch auf ihre Teilnahmebereitschaft auswirken. Vergleicht man die Randverteilungen der beiden ALWA-Gruppen nach dem Erwerbsstatus mit dem Mikrozensus (Tabelle 2), zeigt

sich, dass in beiden ALWA-Gruppen Arbeitssuchende unterrepräsentiert und Inaktive überrepräsentiert sind. Für ersten Befund könnte ebenfalls eine geringe Festnetzdicke verantwortlich sein. Zudem ist zu bedenken, dass sich Arbeitssuchende überproportional aus Niedriggebildeten rekrutierten; das heißt, hier spiegelt sich die Bildungsverzerrung in ALWA wider. Der zweite Befund ist dagegen eher mit der besonders hohen Erreichbarkeit von Inaktiven zu erklären.

4 Selektivität im zweistufigen Design: von CATI zu PAPI

Wie in Abschnitt 2 erläutert, fand im Anschluss an die telefonische Befragung im Rahmen der ALWA-Studie ein mündliches Interview statt – allerdings nur, wenn die Befragten im CATI-Interview ihre Bereitschaft dazu erklärt hatten. ALWA ist somit eine Mixed-Mode-Studie, allerdings eine mit einem sehr speziellen Design, das in einschlägigen Literatursurveys bislang nicht beschrieben ist (Dillman 2009). Weitaus häufiger werden unterschiedliche Modi eingesetzt, um den Befragungsrücklauf zu erhöhen, d.h. jeder Befragte bekommt das gleiche Befragungsprogramm entweder in dem einen oder dem anderen Modus. Wenn der Modus bei gleichen Befragten zu unterschiedlichen Zeitpunkten gewechselt wird, dann werden die Teilnehmer meist zuerst persönlich interviewt und in den Folgewellen nur noch telefonisch, um Geld zu sparen (de Leeuw 2005).

Durch das zweistufige Design der ALWA-Studie kam es zu zusätzlichen Ausfallprozessen bei der PAPI-Befragung: innerhalb der Ausgangsstichprobe dieser Studie, 10.177 realisierten deutschsprachigen CATI-Interviews, wurden 3.990 mündliche Interviews realisiert. Da dieses zweistufige Design keine Vorbilder hat, fehlt auch empirische Forschung zur Selektivität dieser Ausfallprozesse, zumindest in der Mixed-Mode-Forschung. Aus einer anderen Perspektive ist das zweistufige Befragungsdesign von ALWA jedoch nur ein Sonderfall einer „normalen“ Panelbefragung: Auch hier muss die Erlaubnis für eine weitere Kontaktierung in der ersten Welle explizit eingeholt werden, und damit treten die gleichen Ausfallprozesse auf. Die einzige Besonderheit ist die dichte zeitliche Abfolge der „zweiten Befragungswelle“, wodurch sich die Selektivität stark erhöhen dürfte. Grundsätzlich gilt damit auch die Grundannahme aller Untersuchungen zur Panelmortalität: Nonresponse-Prozesse tendieren dazu, mit zunehmender Zeitdauer zu kumulieren (Watson und Wooden 2009).

Im Folgenden wird die Selektivität der Ausfallprozesse an der zweiten Befragungsstufe von ALWA ausführlich untersucht. Wie bei anderen Untersuchungen zur Panelmortalität kann diese Verzerrung multivariat untersucht werden, da für alle potentiellen PAPI-Teilnehmer eine Fülle detaillierter Merkmale bereits im CATI-Interview erhoben wurde (Lipps 2012). Zusätzlich kann der Effekt von Merkmalen des telefonischen Interviews und der Interviewer geprüft werden. Wo die Quellen der zusätzlich entstandenen Selektivität liegen, wird analog zu Abschnitt 3 mit getrennten Modellen für die unterschiedlichen Ausfallprozesse im zweistufigen Studiendesign untersucht: Die ersten Ausfälle treten bereits im telefonischen Interview bei der Frage nach der PAPI-Teilnahmebereitschaft auf, im zweiten Schritt sind nicht alle

Teilnehmer erreichbar und im dritten Schritt verweigern Personen aus der Zielpopulation an der Haustür das mündliche Interview.

4.1 Teilnahmebereitschaft

Die Wahrscheinlichkeit, im CATI-Interview die Frage nach der Bereitschaft zur Teilnahme am mündlichen Interview positiv zu beantworten, wurde anhand von vier logistischen Regressionsmodellen geschätzt (Tabelle 6). Das erste Modell enthält nur Merkmale der Befragten, das zweite enthält zusätzlich Merkmale des telefonischen Interviews. In den Modellen 3a und 3b wurden schließlich noch Merkmale der Telefoninterviewer berücksichtigt, und zwar einmal in absoluter Form und einmal in relativer Form, d.h. in Bezug zu Merkmalen der Befragten.

Die **Merkmale der Befragten** beeinflussen die Teilnahmebereitschaft insgesamt nur geringfügig (Modell 1, Pseudo R^2). Dies heißt auch, dass nur eine geringe zusätzliche Verzerrung im Hinblick auf sozio-demografische Merkmale an diesem Schritt des Untersuchungsdesigns entstanden ist. Die stärksten Effekte weisen das Bildungsniveau der Befragten und ihre Lese- und Schreibkompetenzen auf: Befragte mit Abitur, Fachhochschulabschluss und insbesondere mit Hochschulabschluss haben eine höhere Wahrscheinlichkeit, ihre Bereitschaft zum mündlichen Interview zu geben, als Befragte mit Mittlerer Reife und Ausbildung (die Referenzgruppe) oder Befragte mit niedrigerem Bildungsniveau (vgl. auch Watson und Wooden 2009, Abraham et al. 2006). Und je besser die Befragten ihre Lese- und Schreibkompetenzen einschätzen, desto eher stimmen sie der Teilnahme am PAPI-Interview zu. Gleiches gilt, wenn auch in schwächerem Maße, für die Mathematikkompetenzen. Diese Verzerrungen sind erwartbar, weil im Fragetext zur PAPI-Bereitschaft ein deutlicher Stimulus auf das Wort ‚Lernen‘ gesetzt wurde und die Bearbeitung von Aufgabenheften angekündigt wurde.⁵ Daneben stimmen ältere Befragte, Personen aus den neuen Ländern, Befragte mit sehr vielen Wochenarbeitsstunden, aber auch Arbeitslose und Alleinerziehende dem mündlichen Interview eher zu als jüngere Befragte, Befragte mit weniger Arbeitsstunden oder Singles. Befragte, die mit Partner und Kindern in einem Haushalt leben, haben eine geringere Zustimmungsbereitschaft als die Referenzgruppe der Singles. Auch wenn Merkmale des Interviews und der Interviewer mit einbezogen werden (Modelle 2 bis 3b), verändern sich die Effekte der Teilnehmermerkmale hinsichtlich Richtung und Stärke kaum.

⁵ Der Text lautete: „Wie bereits der Studientitel ‚Arbeiten und Lernen im Wandel‘ andeutet, ist ein Schwerpunkt dieser Studie ja auch das Thema ‚Lernen‘. Lernen heißt dabei für uns nicht nur ‚Lernen in der Schule‘, sondern auch im Beruf, im Alltag und im Leben allgemein. Wir wollen wissen, welche Bedeutung das durch Lernen erworbene Wissen für den Beruf und das Leben allgemein hat. Deshalb interessiert uns besonders, wie die erwachsene Bevölkerung in Deutschland lesen und mit Zahlen umgehen kann. Dazu möchten wir mit Ihnen gerne noch ein persönliches Interview führen, in dem wir Ihnen Aufgabenhefte vorlegen. In den Aufgabenheften werden Sie gebeten, Fragen zu verschiedenen Themen zu beantworten, die alle mit Situationen des täglichen Lebens zu tun haben. Zu diesem Zweck würde Sie in den nächsten Tagen eine Interviewerin oder ein Interviewer von infas bei Ihnen zu Hause besuchen. Wären Sie bereit, an diesem Teil der Studie teilzunehmen? Wenn Sie sich für eine Teilnahme entscheiden, werden wir Ihnen ein kleines Dankeschön in Höhe von 10 Euro überreichen!“.

Tabelle 6: Wahrscheinlichkeit der Bereitschaft zur Teilnahme am PAPI-Interview

	Modell 1		Modell 2		Modell 3a		Modell 3b	
	OR	z	OR	z	OR	z	OR	z
Frau	1.020	(0.47)	1.026	(0.58)	1.019	(0.43)	1.019	(0.44)
Alter (in Jahren)	1.010***	(3.63)	1.014***	(4.71)	1.013***	(4.53)	1.018***	(4.90)
Migrationshintergrund (Ref. kein)								
selbst zugewandert	0.941	(-0.79)	1.019	(0.22)	1.016	(0.19)	1.013	(0.16)
Eltern zugewandert	0.972	(-0.42)	0.978	(-0.32)	0.994	(-0.08)	0.986	(-0.21)
neue Bundesländer	1.224***	(3.55)	1.212***	(3.33)	1.225***	(3.40)	1.216***	(3.29)
ländliche Region (bis 20.000 EW)	1.115	(1.90)	1.092	(1.46)	1.088	(1.39)	1.086	(1.38)
mehrere Wohnorte	1.103	(0.65)	1.088	(0.56)	1.059	(0.38)	1.085	(0.55)
Bildung (Ref. Mittlere Reife + Ausb.)								
ungelehrt / Hauptschule + Ausb.	0.989	(-0.18)	1.044	(0.71)	1.050	(0.82)	1.039	(0.62)
Abitur	1.198**	(2.84)	1.145*	(2.02)	1.144*	(2.02)	1.226*	(2.48)
Fachhochschule	1.200*	(2.38)	1.188*	(2.13)	1.183*	(2.04)	1.304	(1.79)
Universität	1.465***	(5.44)	1.451***	(5.01)	1.461***	(5.09)	1.596***	(3.53)
Lese- und Schreibkompetenzen	1.256***	(7.62)	1.251***	(7.23)	1.250***	(7.21)	1.251***	(7.18)
Mathematikkompetenzen	1.060*	(2.11)	1.060	(1.92)	1.061	(1.96)	1.060	(1.93)
Erwerbsstatus (Ref. 35-49 Std. erw.)								
bis 19 Std. erwerbstätig	1.113	(1.39)	1.132	(1.60)	1.132	(1.64)	1.127	(1.59)
20-34 Std. erwerbstätig	0.982	(-0.24)	1.024	(0.29)	1.016	(0.19)	1.021	(0.25)
50 und mehr Std. erwerbstätig	1.131*	(1.98)	1.152*	(2.19)	1.155*	(2.25)	1.148*	(2.14)
arbeitslos	1.296*	(2.46)	1.412**	(3.14)	1.400**	(3.13)	1.402**	(3.13)
in Bildung / Wehrdienst	1.118	(1.37)	1.076	(0.89)	1.085	(1.00)	1.077	(0.91)
nicht erwerbstätig	1.066	(0.69)	1.127	(1.31)	1.123	(1.28)	1.116	(1.22)
Haushaltsform (Ref. alleine/WG)								
nur mit Partner	0.914	(-1.35)	0.884	(-1.78)	0.888	(-1.70)	0.884	(-1.81)
nur mit Kind(ern)	1.393**	(2.58)	1.367*	(2.23)	1.374*	(2.26)	1.369*	(2.24)
mit Partner und Kind(ern)	0.876*	(-2.27)	0.828**	(-2.97)	0.830**	(-2.95)	0.827**	(-2.97)
bei Eltern	0.872	(-1.63)	0.862	(-1.68)	0.858	(-1.73)	0.867	(-1.63)
Ermüdung Interviewer			1.056*	(2.34)	1.065**	(2.75)	1.057*	(2.45)
Ermüdung Befragter			0.881***	(-4.98)	0.877***	(-5.82)	0.879***	(-5.40)
schlechte Antwortbereitschaft			0.512***	(-7.34)	0.515***	(-6.94)	0.508***	(-7.14)
Unsicherheit (Anzahl weiß nicht)			0.944***	(-3.88)	0.945***	(-3.73)	0.944***	(-3.86)
Anzahl Verweigerungen			0.903***	(-5.26)	0.905***	(-5.01)	0.905***	(-5.03)
kein Einverständnis Zuspiegung			0.408***	(-9.54)	0.417***	(-8.69)	0.415***	(-9.08)
CATI: Int.erfahrung (in Jahren)					1.033	(0.74)	1.017	(0.36)
CATI: Anzahl Interviews pro Int.					1.000	(-0.79)	1.000	(-0.98)
CATI: Anzahl Interview pro Tag					1.007	(0.34)	1.005	(0.28)
CATI: Geschlecht Int.: Frau					0.911	(-1.08)		
CATI: Alter Interviewer					1.011**	(2.59)		
CATI: Bildung Int. (Ref. Ausb.)								
keine Angabe					1.031	(0.18)		
ungelehrt					1.831**	(2.66)		
Abitur					1.372*	(2.16)		
FH / Universität					1.177	(1.22)		
CATI: Int.+ Bef. gl. Geschlecht							0.984	(-0.46)
CATI: rel. Alter Int. (Ref. +/-10 Jahre)								
Int. min. 11 Jahre jünger							0.814*	(-2.06)
Int. min. 11 Jahre älter							1.050	(0.66)
CATI: rel. Bildung Int. (Ref. gleich)								
Int. niedriger gebildet							1.068	(0.72)
Int. höher gebildet							1.196	(1.65)
Pseudo R ²	0.019		0.058		0.063		0.060	
Chi ²	326***		602***		849***		778***	

Logistische Regressionen, Odds Ratios und z-Statistik, robuste Standardfehler (geclustert nach CATI-Interviewer); Signifikanzniveaus: * p<0.05, ** p<0.01, *** p<0.001; N = 10.177 Teilnehmer, 115 PAPI-Interviewer. Quelle: ALWA-Methodendaten und SUF, eigene Berechnungen

Größere Effekte auf die Teilnahmebereitschaft als die Merkmale der Befragten haben ihre **Erfahrungen mit dem bisherigen Interview** (Modell 2). Zu deren Erfassung wurden drei Indikatoren aus den Interviewerfragen verwendet: Einschätzungen der Ermüdung des Befragten sowie des Interviewers sowie eine Dummyvariable für Teilnehmer, deren Antwortbereitschaft nicht ‚durchgehend gut‘ eingeschätzt wurde. Zusätzlich wurden drei Indikatoren aus den CATI-Daten generiert: die Anzahl von „weiß nicht“-Angaben im Interview, die den Grad an Unsicherheit bzw. kognitiver Überforderung durch das Interview anzeigen, die Anzahl von Verweigerungen im Interview und eine Dummyvariable für Personen, die der Zustimmung von Prozessdaten nicht zugestimmt haben. Diese wurde mitten im telefonischen Interview erfragt, also mit deutlichem Abstand zur Zustimmung zum mündlichen Interview. Die letzten beiden Variablen erfassen Misstrauen gegenüber dem Interview, insbesondere Datenschutzvorbehalte. Alle verwendeten Indikatoren zu den Befragtererfahrungen – größere Ermüdung, schlechte Antwortbereitschaft, mehr Unsicherheit, größeres Misstrauen – stehen im negativen Zusammenhang mit der Teilnahmebereitschaft am mündlichen Interview. Die Ermüdung des Interviewers steht dagegen in positivem Zusammenhang mit der Teilnahmebereitschaft. Dies kann damit erklärt werden, dass Befragte, die gerne und viel erzählen, den Interviewer stärker ermüden lassen. Diese Befragten sind gleichzeitig eher bereit, an einem zweiten Interview teilzunehmen. Dass die Interviewerfahrung sehr prädiktiv für die Teilnahmebereitschaft der Folgewelle ist, hat eine Fülle von empirischer Forschung gezeigt, dass diesem Faktor eine große Rolle zukommt (vgl. dazu ausführlich Watson und Wooden 2009). In der ALWA-Studie wurde die Teilnahmebereitschaft für das mündliche Interview zudem direkt im Erstinterview erfragt, so dass situationale Faktoren hier überaus prägend für das Antwortverhalten der Befragten gewesen sein dürften.

Interviewermerkmale tragen nur wenig zur Erklärung der Teilnahmebereitschaft der Befragten bei (Modelle 3a und 3b). Die Erfahrung als Telefoninterviewer, die Zahl der Interviews in der ALWA-Studie, die bereits durchgeführt wurde und die Zahl der Interviews, die am jeweiligen Tag durchgeführt wurde, beeinflusst die Teilnahmebereitschaft der Befragten am mündlichen Interview nicht signifikant. Dieses Ergebnis kann vor allem mit der hohen Standardisierung und den vielen Kontrollen im CATI-Feld erklärt werden. Das Alter der CATI-Interviewer hat einen positiven Effekt auf die Teilnahmebereitschaft (vgl. dazu auch Watson und Wooden 2009). Die Bildung zeigt ein uneinheitliches Muster: Interviewer ohne Ausbildungsabschluss und Interviewer mit Abitur (aber ohne Hochschulabschluss) erreichen eine signifikant höhere Teilnahmebereitschaft als Interviewer mit Berufsausbildung und maximal Realschulabschluss.

Während dieses Ergebnis kaum zu interpretieren ist, gestalten sich die Effekte der Interviewermerkmale sinnhafter, wenn man sie in Beziehung zu Merkmalen der Befragten setzt. Hintergrund dafür ist die gut belegte Tatsache, dass Personen offener und gesprächsbereiter gegenüber solchen Personen sind, die ihnen selbst ähnlich sind (Aronson et al. 2008: 315 ff.). Ob Interviewer und Befragte das gleiche Geschlecht haben oder nicht, spielt empirisch keine Rolle. Sind Interviewer deutlich jünger als die Befragten, dann sinkt deren

Teilnahmebereitschaft. Umgekehrt steigt sie, wenn Interviewer höher gebildet sind als sie selbst – dieser Effekt ist allerdings nur auf dem 10%-Niveau signifikant. Dass diese Effekte relativ geringe Bedeutung haben, erscheint vor dem Hintergrund plausibel, dass es sich um telefonische Interviews gehandelt hat, bei denen die Befragten ihr Gegenüber nicht sehen und nur verbal mit ihnen interagieren können.

4.2 Erreichbarkeit

Im nächsten Schritt wird untersucht, welche Merkmale die Erreichbarkeit derjenigen Befragten, die ihre Zustimmung zum persönlichen Interview gegeben haben, im PAPI-Interview beeinflussen. Im ersten Modell werden wieder ausschließlich Befragtenmerkmale einbezogen. Im zweiten Modell werden zusätzlich auch Merkmale der Interviewer – diesmal natürlich der Interviewer im PAPI-Feld – berücksichtigt, allerdings nur in absoluter Form, weil es ja bei den nicht-erreichten Zielpersonen zu keinem persönlichen Kontakt kam.

Wie theoretisch zu erwarten, gibt es nur relativ wenige Merkmale der Befragten, die ihre Erreichbarkeit im persönlichen Interview beeinflussen, und diese stehen allesamt in Zusammenhang mit räumlicher Mobilität oder mit zeitlichen Restriktionen (Watson und Wooden 2009). So sind Ältere besser erreichbar als Jüngere, während Personen mit mehreren Wohnorten und solche mit sehr vielen Wochenarbeitsstunden schlecht erreichbar sind. Leben Zielpersonen mit einem Partner in einem Haushalt, sind sie besser erreichbar als Singles. Besonders gut erreichbar sind Personen, die noch bei ihren Eltern leben.

Merkmale der PAPI-Interviewer üben einen erstaunlich großen Effekt auf die Erreichbarkeit der Befragten aus: Erfahrenen Interviewern gelingt es besser, die Zielpersonen zu erreichen als unerfahrenen – sie haben also vermutlich bessere Strategien des Aufsuchens und lassen sich nicht so leicht von Misserfolgen frustrieren. Umgekehrt beeinflusst die Zahl der Interviews, die am gleichen Tag oder in der gleichen Studie schon durchgeführt wurden, die Erreichbarkeit negativ (zum Forschungsstand vgl. Nicoletti und Buck 2004). Die Motivation der Interviewer scheint also mit zunehmender Felddauer und mit zunehmender Zahl von Interviews am Tag abzunehmen. Dabei ist allerdings zu berücksichtigen, dass sich Personen, die öfter nicht erreicht wurden, mit zunehmender Felddauer häufen, weil sie immer wieder angegangen werden müssen. Interviewer brauchen mit daher immer länger, um ein erfolgreiches Interview zu absolvieren. Wenn diese Mühe nicht entsprechend entlohnt wird, passen Interviewer ihre Strategien an, indem sie eher für andere Studien oder andere Auftraggeber arbeiten. Der Effekt der Anzahl der bereits durchgeführten Interviews am Tag kann ebenfalls mit Strategien der Interviewer erklärt werden: So ist es wahrscheinlich, dass Personen, die öfter nicht zu erreichen waren, erst spät am Abend erneut aufgesucht werden, während neue Adressen zuerst abgearbeitet werden. Sozio-demografische Merkmale der Interviewer haben schließlich – wie theoretisch nicht anders zu erwarten – keine signifikanten Effekte auf die Erreichbarkeit der Studienteilnehmer.

Tabelle 7: Wahrscheinlichkeit der Erreichbarkeit im PAPI-Interview

	Modell 1		Modell 2	
	OR	z	OR	z
Frau	0.895	(-1.23)	0.838	(-1.84)
Alter (in Jahren)	1.034***	(5.70)	1.032***	(5.23)
Migrationshintergrund (Ref. kein)				
selbst zugewandert	1.166	(0.89)	1.057	(0.32)
Eltern zugewandert	0.811	(-1.42)	0.787	(-1.66)
neue Bundesländer	0.888	(-0.64)	0.882	(-0.80)
ländliche Region (bis 20.000 EW)	1.204	(1.09)	1.317	(1.84)
mehrere Wohnorte	0.550*	(-2.16)	0.496*	(-2.41)
Bildung (Ref. Mittlere Reife + Ausbildung)				
ungelernt / Hauptschule + Ausbildung	0.969	(-0.24)	0.938	(-0.44)
Abitur	1.168	(1.06)	1.203	(1.21)
Fachhochschule	1.104	(0.60)	1.044	(0.24)
Universität	1.068	(0.48)	1.042	(0.29)
Selbsteinschätzung Lesen und Schreiben	0.953	(-0.73)	0.975	(-0.36)
Selbsteinschätzung Mathematik	1.053	(0.89)	1.030	(0.46)
Erwerbsstatus (Ref. 35-49 Std. erwerbstätig)				
bis 19 Std. erwerbstätig	1.116	(0.69)	1.037	(0.22)
20-34 Std. erwerbstätig	0.944	(-0.37)	0.976	(-0.15)
50 und mehr Std. erwerbstätig	0.782*	(-2.13)	0.755*	(-2.28)
arbeitslos	0.852	(-0.82)	0.804	(-1.05)
in Bildung / Wehrdienst	1.001	(0.00)	0.996	(-0.02)
nicht erwerbstätig	1.294	(1.18)	1.318	(1.22)
Haushaltsform (Ref. alleine/WG)				
nur mit Partner	1.411*	(2.28)	1.491**	(2.71)
nur mit Kind(ern)	1.324	(1.46)	1.300	(1.20)
mit Partner und Kind(ern)	1.549**	(3.22)	1.571***	(3.30)
bei Eltern	2.278***	(5.65)	2.208***	(4.90)
PAPI: Interviewererfahrung (in Jahren)			1.219***	(4.08)
PAPI: Anzahl Interviews pro Int.			0.983***	(-6.85)
PAPI: Anzahl Interviews pro Tag			0.862**	(-2.74)
PAPI: Geschlecht Interviewer: Frau			0.803	(-1.27)
PAPI: Alter Interviewer			1.003	(0.27)
PAPI: Bildung Interviewer (Ref. Ausbildung)				
keine Angabe			0.422	(-1.73)
ungelernt			1.670	(1.91)
Abitur			0.563*	(-2.48)
FH / Universität			0.821	(-0.91)
Pseudo R ²	0.026		0.114	
Chi ²	154***		401***	

Logistische Regressionen, Odds Ratios und z-Statistik, robuste Standardfehler (geclustert nach PAPI-Interviewer); Signifikanzniveaus: * p<0.05, ** p<0.01, *** p<0.001; N = 5.813 Teilnehmer, 115 PAPI-Interviewer.
Quelle: ALWA-Methodendaten und SUF, eigene Berechnungen

4.3 Kooperationsbereitschaft

Im dritten und letzten Schritt wird untersucht, welche Merkmale die Realisierungswahrscheinlichkeit des Face-to-Face-Interviews – oder, anders gesagt, die Wahrscheinlichkeit, das Interview nicht zu verweigern – beeinflussen. Grundgesamtheit dieser Analyse sind alle PAPI-bereiten Studienteilnehmer, die von den PAPI-Interviewern erreicht wurden und keine neut-

ralen Ausfälle darstellen. Ins erste Modell werden wieder ausschließlich Befragtenmerkmale einbezogen. Ins zweite Modell werden zusätzlich Merkmale des vorhergehenden telefonischen Interviews aufgenommen. In den Modellen 3a und 3b werden schließlich Merkmale der PAPI-Interviewer berücksichtigt, und zwar wieder in absoluter und relativer Form.

Betrachtet man Modell 1, dann fällt auf, dass Frauen eine signifikant geringere Realisierungswahrscheinlichkeit aufweisen als Männer und dass diese mit zunehmendem Alter der Befragten steigt. Analog zur Teilnahmebereitschaft beeinflussen auch hier das Bildungsniveau und die selbst eingeschätzten Kompetenzen die Teilnahme, wobei an der Haustür nur die Mathematikkompetenzen von Bedeutung sind. Schließlich führen Personen, die bis 19 Stunden in der Woche erwerbstätig sind, und Arbeitslose eher ein Interview als die Referenzgruppe der Vollzeitbeschäftigten (Watson und Wooden 2009). Die Haushaltsform übt im Gegensatz dazu keinen Effekt auf die Realisierungswahrscheinlichkeit des Interviews aus. Die Einflüsse der genannten Merkmale lassen sich erstens mit zeitlichen Spielräumen erklären (Alter, Erwerbsform, Arbeitslosigkeit), zweitens mit Vorbehalten, einen fremden Interviewer in die eigene Wohnung zu lassen (Geschlecht), und drittens mit einem größeren Interesse Höhergebildeter an den Inhalten des Interviews und geringeren Ängsten gegenüber den angekündigten Aufgabenheften. An der Tür, wo das Befragungsintrou weitaus weniger standardisiert erfolgt als im telefonischen Interview bei der Frage nach der Interviewbereitschaft, war die Nennung von Mathematikaufgaben anscheinend bedeutsamer für die Realisierungschance als die Nennung von Leseaufgaben.

Neben soziodemografischen Merkmalen der Befragten beeinflussen auch ihre Erfahrungen im vorangegangenen Telefoninterview die Wahrscheinlichkeit, sich am Face-to-Face-Interview zu beteiligen (Modell 2). Befragte, deren Antwortbereitschaft vom CATI-Interviewer als schlecht eingestuft worden war, beteiligen sich mit geringerer Wahrscheinlichkeit als Befragte mit einer durchgehend hohen Antwortbereitschaft. Ebenso führen stärkere Ermüdung und viele Verweigerungen im telefonischen Interview zu einer geringeren Beteiligungswahrscheinlichkeit.

Deutliche Effekte üben schließlich auch die Merkmale der PAPI-Interviewer aus (Modelle 3a und 3b). Mit steigender Erfahrung steigt die Realisierungswahrscheinlichkeit, umgekehrt sinkt sie mit zunehmender Zahl bereits realisierter ALWA-Interviews. Diese beiden Effekte können einerseits mit besseren Strategien erfahrener Interviewer erklärt werden, andererseits mit Feldzeiteffekten und mangelnder Motivation, ‚schwierige‘ Fälle am Ende der Feldzeit zu bearbeiten (Watson und Wooden 2009). Soziodemografische Merkmale der Interviewer spielen demgegenüber kaum eine Rolle. So ist es für die Realisierungschancen unerheblich, ob Männer oder Frauen, junge oder ältere Interviewer an der Tür stehen. Einzig das Bildungsniveau der Interviewer ist von Bedeutung: Die Referenzgruppe mit maximal Realschulabschluss und einer Berufsausbildung ist erfolgreicher als Interviewer aus anderen Bildungsgruppen, insbesondere Personen ohne Berufsausbildung sowie Personen mit Abitur und ohne akademische Bildung. In relativer Hinsicht spielen Merkmale der Interviewer schließlich keine Rolle bei der Erklärung der Realisierungschance des PAPI-Interviews.

Tabelle 8: Wahrscheinlichkeit der Teilnahme am PAPI-Interview

	Modell 1		Modell 2		Modell 3a		Modell 3b	
	OR	z	OR	z	OR	z	OR	z
Frau	0.804**	(-2.76)	0.797**	(-2.93)	0.789**	(-2.94)	0.792**	(-2.94)
Alter (in Jahren)	1.017***	(3.87)	1.019***	(4.25)	1.019***	(4.31)	1.027***	(3.67)
Migrationshintergrund (Ref. kein)								
selbst zugewandert	0.878	(-1.10)	0.895	(-0.93)	0.871	(-1.16)	0.887	(-1.01)
Eltern zugewandert	0.975	(-0.20)	0.981	(-0.15)	0.973	(-0.22)	0.985	(-0.12)
neue Bundesländer	0.965	(-0.20)	0.956	(-0.25)	0.973	(-0.16)	0.913	(-0.56)
ländliche Region (bis 20.000 EW)	1.233	(1.57)	1.219	(1.46)	1.237	(1.61)	1.255	(1.65)
mehrere Wohnorte	0.825	(-1.10)	0.845	(-0.97)	0.810	(-1.20)	0.837	(-1.01)
Bildung (Ref. Mittlere Reife + Ausb.)								
ungelernt / Hauptschule + Ausb.	0.774**	(-2.64)	0.782*	(-2.47)	0.802*	(-2.17)	0.788*	(-2.33)
Abitur	1.190	(1.62)	1.175	(1.52)	1.209+	(1.79)	1.165	(1.29)
Fachhochschule	1.105	(0.86)	1.110	(0.89)	1.124	(1.00)	1.057	(0.31)
Universität	1.402***	(3.29)	1.402***	(3.30)	1.465***	(3.78)	1.363	(1.91)
Lesen- und Schreibkompetenzen	1.047	(0.95)	1.038	(0.77)	1.049	(0.99)	1.045	(0.92)
Mathematikkompetenzen	1.165***	(4.33)	1.166***	(4.34)	1.152***	(3.86)	1.160***	(4.07)
Erwerbsstatus (Ref. 35-49 Std. erw.)								
bis 19 Std. erwerbstätig	1.447**	(2.98)	1.462**	(3.05)	1.505**	(3.23)	1.491**	(3.13)
20-34 Std. erwerbstätig	1.096	(0.74)	1.113	(0.86)	1.132	(0.98)	1.102	(0.78)
50 und mehr Std. erwerbstätig	1.007	(0.07)	1.011	(0.11)	1.000	(0.00)	1.003	(0.03)
arbeitslos	1.995***	(3.42)	2.068***	(3.60)	2.104***	(3.82)	2.086***	(3.74)
in Bildung / Wehrdienst	1.274	(1.93)	1.249	(1.75)	1.281	(1.88)	1.290	(1.91)
nicht erwerbstätig	0.996	(-0.03)	1.023	(0.14)	1.069	(0.43)	1.048	(0.30)
Haushaltsform (Ref. alleine/WG)								
nur mit Partner	0.839	(-1.39)	0.836	(-1.43)	0.829	(-1.49)	0.837	(-1.42)
nur mit Kind(ern)	1.094	(0.53)	1.091	(0.51)	1.075	(0.43)	1.072	(0.40)
mit Partner und Kind(ern)	1.013	(0.12)	1.002	(0.02)	0.992	(-0.07)	0.993	(-0.07)
bei Eltern	0.998	(-0.02)	0.992	(-0.06)	0.976	(-0.17)	1.016	(0.11)
Ermüdung Befragter			0.956	(-1.89)	0.951*	(-2.06)	0.952*	(-2.03)
schlechte Antwortbereitschaft			0.677*	(-2.54)	0.685*	(-2.42)	0.674*	(-2.54)
Anzahl Verweigerungen			0.939***	(-3.32)	0.943**	(-2.91)	0.941**	(-3.07)
PAPI: Int.erfahrung (in Jahren)					1.128**	(2.86)	1.118*	(2.51)
PAPI: Anzahl Interviews pro Int.					0.995**	(-2.92)	0.995**	(-2.62)
PAPI: Anzahl Interviews pro Tag					0.951	(-0.84)	0.952	(-0.79)
PAPI: Geschlecht Int.: Frau					1.012	(0.08)		
PAPI: Alter Interviewer					1.001	(0.15)		
PAPI: Bildung Int. (Ref. Ausb.)								
keine Angabe					0.292	(-1.94)		
ungelernt					0.524***	(-3.75)		
Abitur					0.656*	(-2.29)		
FH / Universität					0.781	(-1.40)		
PAPI: Int.+ Bef. gleiches Geschl.							1.021	(0.35)
PAPI: rel. Alter Int. (Ref. +/-10 Jahre)								
Int. min. 11 Jahre jünger							0.691	(-0.96)
Int. min. 11 Jahre älter							1.212	(1.42)
PAPI: rel. Bildung Int. (Ref. gleich)								
Int. niedriger gebildet							1.062	(0.39)
Int. höher gebildet							1.001	(0.01)
Pseudo R ²	0.020		0.025		0.045		0.038	
Chi ²	149***		208***		264***		238***	

Logistische Regressionen, Odds Ratios und z-Statistik, robuste Standardfehler (geclustert nach PAPI-Interviewer); Signifikanzniveaus: * p<0.05, ** p<0.01, *** p<0.001; N = 5.210 Teilnehmer, 115 PAPI-Interviewer.
Quelle: ALWA-Methodendaten und SUF, eigene Berechnungen

5 Zusammenfassung und Schlussfolgerungen

Im vorliegenden Beitrag wurden vorhandene Daten über Studienteilnehmer und Nichtteilnehmer genutzt, um die Selektivitätsprozesse zu untersuchen, die bei der Durchführung der IAB-Befragung ALWA entstanden sind. Neben den üblichen Ausfallprozessen, welche aus Nichterreichbarkeit und Nichtteilnahme resultieren, waren bei dieser Studie zwei weitere Ausfallprozesse von zentralem Interesse. Erstens konnten nicht für alle Adressen aus der Einwohnermeldeamtsstichprobe Telefonnummern für die CATI-Befragung gefunden werden. Zweitens entstanden zusätzliche Verzerrungen durch das zweistufige Design von ALWA. Im Anschluss an die telefonische Befragung der Zielpersonen, in deren Fokus die Erhebung der Soziodemographie und des retrospektiven Lebensverlaufs stand, wurden in einem persönlichen Interview die kognitiven Grundkompetenzen der Studienteilnehmer getestet, die sich im Telefoninterview bereit erklärt hatten, an einem persönlichen Interview teilzunehmen.

Unsere Analysen zur Selektivität der CATI-Befragung zeigen, dass die Daten hinsichtlich dreier Merkmale – Migrationshintergrund, Alter und Bildung deutliche Verzerrungen gegenüber den Referenzdaten des Mikrozensus aufweisen. Die Ursache für diese Verzerrungen ist zum großen Teil in der Telefonnummernrecherche für die Adressen der Einwohnermeldeamtsstichprobe zu suchen. Die nicht-zufällige Verbreitung von Mobilfunkanschlüssen und nichtgelisteten Festnetzanschlüssen, insbesondere für die sehr mobile Altersgruppe zwischen 26 und 30 Jahren in Großstädten, scheint hier eines der Hauptprobleme zu sein.

Künftige Befragungen sollten, so unser Fazit, also nach Möglichkeit nicht mehr ausschließlich telefonisch durchgeführt werden – zumindest wenn sie auf Einwohnermeldeamtsstichproben oder auf konventionellen Telefonstichproben basieren. Diese Schlussfolgerung wurde in der Erwachsenenetappe des Nationalen Bildungspanels (NEPS), der Nachfolgestudie von ALWA, bereits umgesetzt. Die in dieser Studie neu gezogenen Teilnehmer wurden von Beginn an in einem CATI-CAPI-Mix befragt. Dafür wurde die Ausgangsstichprobe in zwei Substichproben aufgeteilt: eine mit und eine ohne recherchierbare Telefonnummern. Die Mitglieder der ersten Substichprobe wurden im ersten Interview telefonisch, die der zweiten in einem persönlichen Interview angegangen. In späteren Interviews können die Studienteilnehmer dann selbst entscheiden, in welchem Modus sie befragt werden möchten.

Das zweite zentrale Ergebnis unserer Analysen ist, dass durch das zweistufige Design von ALWA zusätzliche Selektivität nur in relativ geringem Ausmaß entstanden ist – zumindest, was soziodemographische Merkmale der Befragten angeht. Bei den drei Ausfallprozessen der Teilnahmebereitschaft, der Erreichbarkeit und der Kooperationsbereitschaft zeigen sich zwar einige signifikante Effekte, in der Summe ergeben sich jedoch nur messbare Abweichungen zwischen CATI- und PAPI-Teilnehmern hinsichtlich des Alters, der formalen Bildung und der Selbsteinschätzung ihrer Lese- und Mathematikkompetenzen: PAPI-Teilnehmer sind älter, höher gebildet und sie schätzen sich als bessere Leser, Schreiber und Rechner ein als die Gruppe aller CATI-Teilnehmer (Tabelle 9). Ob die Verzerrungen in punkto formaler Bildung und der Selbsteinschätzung von Kompetenzen mit einem anderen Design zu vermeiden wären, erscheint im Rückblick fraglich. So besteht zwar theoretisch die Möglichkeit, bei

der Abfrage der Bereitschaft zum persönlichen Interview das Befragungsthema weniger explizit zu benennen, als das in ALWA getan wurde, aber das würde erstens ethische Probleme aufwerfen und zweitens die Selektivitätsprobleme nur verschieben. So zeigte sich in der NEPS-Erwachsenenbefragung, in der Kompetenztest und Befragung gemeinsam in einem persönlichen Interview durchgeführt werden, dass niedrig gebildete und kompetenzarme Befragte vermehrt auf Ausweichstrategien zurückgreifen, um den Tests zu entgehen: entweder sie verweigern das persönliche Interview, nehmen aber später an einem telefonischen Konvertierungsinterview teil, sie verweigern die Testteilnahme oder sie brechen die Tests ab.

Tabelle 9: Randverteilungen der CATI- und PAPI-Befragten

	CATI	PAPI	
Frau	50.6%	50.0%	
Alter			
18-24 Jahre	20.0%	18.0%	
25-29 Jahre	8.9%	8.0%	
30-34 Jahre	8.5%	8.2%	
35-39 Jahre	15.3%	14.0%	
40-44 Jahre	21.7%	23.6%	
45-50 Jahre	25.5%	28.1%	
Migrationshintergrund			
keiner	82.8%	84.1%	
selbst zugewandert	8.3%	7.4%	
Eltern zugewandert	8.9%	8.6%	
neue Bundesländer	18.4%	19.5%	
ländliche Region (bis 20.000 EW)	12.5%	13.2%	
mehrere Wohnorte	2.5%	2.4%	
Bildung			
ungelernt / Hauptschule + Ausbildung	18.1%	14.9%	
Mittlere Reife + Ausbildung	30.8%	28.5%	
Abitur	18.7%	19.0%	
Fachhochschule	10.4%	11.0%	
Universität	22.0%	26.5%	
Lese- und Schreibkompetenzen	4.00	4.08	
Mathematikkompetenzen	3.72	3.79	
Erwerbsstatus			
bis 19 Std. erwerbstätig	9.8%	10.5%	
20-34 Std. erwerbstätig	12.8%	12.7%	
35-49 Std. erwerbstätig	38.4%	37.3%	
50 und mehr Std. erwerbstätig	15.1%	16.2%	
arbeitslos	5.1%	5.6%	
in Bildung / Wehrdienst	11.7%	10.9%	
nicht erwerbstätig	7.2%	6.7%	
Haushaltsform			
alleine / WG	16.0%	16.2%	
nur mit Partner	16.2%	15.5%	
nur mit Kind(ern)	4.5%	5.3%	
mit Partner und Kind(ern)	46.6%	48.0%	
bei Eltern	16.7%	15.0%	
N	10177	4037	Quelle: ALWA-Methodendaten und SUF, eigene Berechnungen

Größere Selektivität im persönlichen Interview wurde durch die Erfahrungen erzeugt, die die Befragten im telefonischen Interview gemacht haben. So wirken sich Ermüdung, schlechte Antwortbereitschaft und Unsicherheit der Befragten im Telefoninterview negativ auf die Bereitschaft zur Teilnahme am persönlichen Interview aus und in geringerem Maße beeinflussen sie auch die Kooperationsbereitschaft. Für die Erreichbarkeit und die tatsächliche Teilnahme am PAPI-Interview spielen daneben Merkmale der Interviewer eine Rolle, insbesondere ihre Erfahrung und die Anzahl der Interviews, die sie bereits durchgeführt haben. Dieses Ergebnis zeigt, wie wichtig gut geschulte und erfahrene Interviewer und die Feldsteuerung in einem Face-to-Face-Feld sind. Im internationalen Vergleich werden in Deutschland für Interviewerschulungen sehr wenige Ressourcen zur Verfügung gestellt. Gleichzeitig steigt die Bedeutung von persönlichen Interviews angesichts der großen Selektivität telefonischer Befragungen und der wachsenden Verbreitung von Mixed-Mode-Designs.

Letztlich ziehen wir aus dem Einsatz des zweistufigen Designs in ALWA eine positive Bilanz. Die zusätzlichen Ausfallprozesse durch die getrennte Durchführung von Kompetenztests in einem persönlichen Interview haben zwar zu einer erheblichen Verringerung der Fallzahlen geführt, aber kaum zu einem Verlust der Repräsentativität der Daten hinsichtlich soziodemographischer Merkmalen der Befragten. Dem gegenüber steht das Kosten-Nutzen-Verhältnis des zweistufigen Designs. Zum einen wurden die Befragten durch die Trennung der Erfassung des Lebenslaufes in der telefonischen Befragung und der Kompetenztests im nachgelagerten persönlichen Interview erheblich entlastet und zum anderen war die Studie wesentlich kostengünstiger als eine reine Face-to-Face-Befragung.

Literatur

Aronson, Elliot; Wilson, Timothy; Akert, Robin: Sozialpsychologie, 6. Aufl. München: Pearson Studium.

Abraham, Katharine G.; Maitland, Aaron; Bianchi, Suzanne M. (2006): Nonresponse in the American Time Use Survey. Who Is Missing from the Data and How Much Does It Matter? In: Public Opinion Quarterly, 70: 676-703.

Antoni, Manfred (2013): Linking survey data with administrative employment data: The case of the German ALWA survey. FDZ Methodenreport 03/2013 (im Erscheinen). Nürnberg: IAB.

Antoni, Manfred; Drasch, Katrin; Kleinert, Corinna; Matthes, Britta; Ruland, Michael; Trahms, Annette (2010): Arbeiten und Lernen im Wandel. Teil I: Überblick über die Studie. FDZ Methodenreport 05/2010. Nürnberg: IAB.

Blom, Annelies G.; Lynn, Peter; Jäckle, Annette (2008): Understanding Cross-National Differences in Unit Non-Response: The Role of Contact Data. ISER Working Paper No. 2008-01. Colchester: University of Essex.

Brehm, John (1993): The Phantom Respondents. Opinion Surveys and Political Representation. Ann Arbor: The University of Michigan Press.

Busse, Britta; Fuchs, Marek (2012): The Components of Landline Telephone Survey Coverage Bias. The Relative Importance of No-Phone and Mobile-Only Populations. In: Quality and Quantity, 46: 1209-1225.

De Leeuw, Edith D. (2005): To Mix or Not to Mix Data Collection Modes in Surveys. In: Journal of Official Statistics, 21: 233-255.

Dillman, Don (2009): Some Consequences of Survey Mode Changes in Longitudinal Surveys. In: Peter Lynn (Hrsg.), Methodology of Longitudinal Surveys (S. 127-140). Chichester: John Wiley and Sons.

Fitzgerald, John; Gottschalk, Peter; Moffitt, Robert (1998): The Impact of Attrition in the Panel Study of Income Dynamics on Intergenerational Analysis. In: Journal of Human Resources, 33: 300-344.

Gilberg, Reiner; Hess, Doris; Prussog-Wagner, Angela; Steinwede, Angelika (2011): Arbeiten und Lernen im Wandel. Teil III: Methodenbericht. FDZ Methodenreport 10/2011. Nürnberg: IAB.

Granato, Nadia (1999): Die Befragung von Arbeitsmigranten: Einwohnermeldeamt-Stichprobe und telefonische Erhebung? In: ZUMA-Nachrichten, 45: 44-60.

Goyder, John (1987): The Silent Minority. Nonrespondents on Sample Surveys. Boulder: Colorado Westview Press.

Heerwergh, Dirk; Abts, Koen; Loosveldt, Gert (2007): Minimizing Survey Refusal and Non-contact Rates: Do Our Efforts Pay Off? In: Survey Research Methods, 1: 3-10.

Kleinert, Corinna; Matthes, Britta; Jacob, Marita (2008): Die Befragung „Arbeiten und Lernen im Wandel“. Theoretischer Hintergrund und Konzeption. IAB-Forschungsbericht 05/2008. Nürnberg: IAB.

Kleinert, Corinna; Yamamoto, Kentaro; Wölfel, Oliver; Gilberg, Rainer (2012): Working and Learning in a Changing World. Part VI: Literacy and Numeracy Skills – Test Design, Implementation, Scaling and Statistical Models for Proficiency Estimation. FDZ Methodenreport 10/2012. Nürnberg: IAB.

Kreuter, Frauke; Lemay, Michael; Casas-Cordero, Carolina (2007): Using Proxy Measures of Survey Outcomes in Post-Survey Adjustments. Examples from the European Social Survey (ESS). In: American Statistical Association (Hrsg.), Proceedings of the Survey Research Methods Section, S. 3142-3149. Alexandria.

Kreuter, Frauke; Müller, Gerrit; Trappmann, Mark (2010): Nonresponse and Measurement Error in Employment Research. Making Use of Administrative Data. In: Public Opinion Quarterly, 74: 880-906.

Koch, Achim (1997): Teilnahmeverhalten beim ALLBUS 1994. Soziodemographische Determinanten von Erreichbarkeit, Befragungsfähigkeit und Kooperationsbereitschaft. In: Kölner Zeitschrift für Soziologie und Sozialpsychologie, 49: 98-122.

Koch, Achim (1998): Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist. Ausschöpfungsquoten und Stichprobenverzerrungen in Allgemeinen Bevölkerungsumfragen. In: ZUMA-Nachrichten, 42: 66-90.

Lengerer, Andrea; Bohr, Jeanette; Janßen, Andrea (2005): Haushalte, Familien und Lebensformen im Mikrozensus – Konzepte und Typisierungen. ZUMA-Arbeitsbericht 2005/05. Mannheim: ZUMA.

Lepkowski, James M.; Couper, Mick P. (2002): Nonresponse in the Second Wave of Longitudinal Household Surveys. In: Robert M. Groves, Don A. Dillman, John L. Eltinge and Roderick J.A. Little (Hrsg.), Survey Nonresponse (S. 259-272). Chichester: John Wiley and Sons.

Lipps, Oliver (2012): Using Information from Telephone Panel Surveys to Predict Reasons for Refusal. In: MDA – Methoden, Daten, Analysen, 6: 3-20.

Lynn, Peter; Clarke, Paul (2002): Separating Refusal Bias and Non-Contact Bias: Evidence from UK National Surveys. In: The Statistician, 51: 319-333.

Lynn, Peter; Clarke, Paul; Martin, Jean; Sturgis, Patrick (2002): The Effects of Extended Interviewer Efforts on Nonresponse Bias. In: Robert M. Groves, Don A. Dillman, John L. Eltinge, Roderick J.A. Little (Hrsg.): Survey Nonresponse (S. 135-147). New York: Wiley.

Nicoletti, Cheti; Buck, Nicolas H. (2004): Explaining Interviewee Contact and Co-operation in the British and German Household Panels. In: Manfred Ehling, Ulrich Rendtel (Hrsg.), Harmonization of Panel Surveys and Data Quality (S. 143-166). Wiesbaden: Statistisches Bundesamt.

Nöthen, Manuela (2005): Von der „traditionellen Familie“ zu „neuen Lebensformen“. Neuerungen in der Familienberichterstattung des Mikrozensus. In: Wirtschaft und Statistik, 1/2005: 25-40.

Schnell, Rainer (2012): Survey-Interviews. Methoden standardisierter Befragungen. Wiesbaden: VS Verlag für Sozialwissenschaften.

Schnell, Rainer; Esser, Elke; Hill, Paul B. (2008): Methoden der empirischen Sozialforschung. München, Wien: Oldenbourg Verlag.

Schräpler, Jörg-Peter (2000): Was kann man am Beispiel des SOEP bezüglich Nonresponse lernen? ZUMA Nachrichten, 46: 117-149.

Watson, Nicole; Wooden, Mark (2009): Identifying Factors Affecting Longitudinal Survey Response. In: Peter Lynn (Hrsg.), Methodology of Longitudinal Surveys (S. 157-181). Chichester: John Wiley and Sons.

Impressum

FDZ-Methodenreport 2/2013

Herausgeber

Forschungsdatenzentrum (FDZ)
der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt- und Berufsforschung
Regensburger Str. 104
90478 Nürnberg

Redaktion

Stefan Bender, Heiner Frank

Technische Herstellung

Heiner Frank

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2013/MR_02-13.pdf

Internet

<http://fdz.iab.de/>

Rückfragen zum Inhalt an:

Corinna Kleinert
Institut für Arbeitsmarkt und
Berufsforschung (IAB)
Regensburger Str. 104
90478 Nürnberg
Tel. +49 (0)911/179-3075
<mailto:corinna.kleinert@iab.de>