

Forschungsdatenzentrum

der Bundesagentur für Arbeit
im Institut für Arbeitsmarkt-
und Berufsforschung

FDZ

FDZ-Methodenreport

02/2015

DE

Methodische Aspekte zu Arbeitsmarktdaten

IAB-Beschäftigtenbefragung

Projekt „Arbeitsqualität und wirtschaftlicher Erfolg:
Panelstudie zu Entwicklungsverläufen in deutschen
Betrieben – Personenbefragung“

Holger Schütz,
Reiner Gilberg,
Christian Dickmann,
Helmut Schröder



Bundesagentur für Arbeit

Methodenbericht

IAB-Beschäftigtenbefragung - Projekt „Arbeitsqualität und wirtschaftli- cher Erfolg: Panelstudie zu Entwicklungs- verläufen in deutschen Betrieben“ – Personenbefragung“

Autoren: Holger Schütz, Reiner Gilberg, Christian Dickmann, Helmut Schröder

infas Institut für angewandte
Sozialwissenschaft GmbH

Friedrich-Wilhelm-Straße 18
D-53113 Bonn
Tel. +49 (0)228/38 22-0
Fax +49 (0)228/31 00 71
info@infas.de
www.infas.de

Bericht an

IAB
Regensburger Straße 104
90478 Nürnberg

Projekt

4927
Bonn, Mai 2013
Su/Sc

Vorgelegt von

infas Institut für angewandte
Sozialwissenschaft GmbH
Friedrich-Wilhelm-Straße 18
53113 Bonn

Kontakt

Dr. Helmut Schröder
Bereichsleitung Sozialforschung

Tel. +49 (0)228/38 22-406
Fax +49 (0)228/310071
E-Mail h.schroeder@infas.de

Autoren

Dr. Holger Schütz, Senior-Projektleiter
Dr. Reiner Gilberg, Abteilungsleiter Statistik
Christian Dickmann, Senior-Projektleiter Datenaufbereitung
Dr. Helmut Schröder, Bereichsleiter Sozialforschung

Der Inhalt dieses Berichts darf ganz oder teilweise nur mit unserer schriftlichen Genehmigung veröffentlicht, vervielfältigt, gedruckt oder in Informations- und Dokumentationssystemen (information storage and retrieval systems) gespeichert, verarbeitet oder ausgegeben werden.

Vorbemerkung

Im Rahmen der Studie „Arbeitsqualität und wirtschaftlicher Erfolg: Panelstudie zu Entwicklungsverläufen in 600 deutschen Betrieben“ hat das IAB das infas – Institut für angewandte Sozialwissenschaft GmbH mit der Durchführung der Beschäftigtenbefragung beauftragt. Die als Panel angelegte Studie erfasst, wie sich betriebliche Personal- und Geschäftsstrategien in der unternehmerischen Entwicklung und in der beruflichen Entwicklung der Beschäftigten auswirken.

Der vorliegende Bericht stellt zunächst das Erhebungsdesign und Erhebungsinstrument der Beschäftigtenbefragung vor (Kapitel 1) und beschreibt danach die Zusammensetzung der Stichprobe (Kapitel 2). Im dritten Kapitel wird das Erhebungsinstrument inklusive der Pretestergebnisse vorgestellt. Die Feldorganisation und -durchführung sind Gegenstand von Kapitel 4. Kapitel 5 stellt die Feld- und Ausschöpfungsergebnisse dar. Anschließend werden die Datenprüfungs- und Datenlieferungsprozesse thematisiert (Kapitel 6). Im letzten, siebten Kapitel werden die Methodik und Ergebnisse der Selektivitätsanalyse und Gewichtung referiert.

infas - Institut für angewandte Sozialwissenschaft GmbH

Inhaltsverzeichnis

1 Erhebungsdesign	6
1.1 Einordnung des Projekts	6
1.2 Studiendesign im Überblick	7
2 Stichprobenkonzept und -auswahl	9
2.1 Grundgesamtheit und Stichprobenziehung	9
2.2 Umfang der Stichprobe	10
2.2.1 Preteststichprobe	10
2.2.2 Stichprobe der Haupterhebung und Tranchierung	10
2.2.3 Tranchierung der Einsatzstichproben	10
3 Erhebungsinstrument	11
3.1 Inhalte und Struktur des Fragebogens	11
3.2 Ansprache der Zielperson und Screening	12
3.3 Fragebogenprogrammierung	14
3.4 Pretest	15
4 Felddurchführung der Haupterhebung	16
4.1 Feldzeit und Feldsteuerung	16
4.2 Interviewerschulungen	17
4.3 Interviewereinsatz	17
4.4 Dokumentation des Feldverlaufs	19
4.5 Kontaktierung im Feldverlauf	20
4.6 Supervision zur Qualitätssicherung	21
5 Feldergebnis	22
5.1 Final Outcome	22
5.2 Outcome Rates	26
5.3 Maßnahmen zur Verbesserung der Adressausschöpfung	27
5.3.1 Adressaktualisierung	27
5.3.2 Konvertierung	27
5.4 Paneleinwilligungen und Zuspielungsbereitschaft	27
5.5 Interviewdauern	28
5.6 Interviewabbrüche	28
6 Datenprüfung und -lieferung	29
6.1 Definition gültig realisierter Interviews	29
6.2 Prüfroutinen im Fragenprogramm	30
6.3 Datenprüfung und -aufbereitung nach Feldende	30
6.4 Datenlieferung	31
7 Selektivität und Gewichtung	31
7.1 Wahrscheinlichkeit für die Einsatzstichprobe	32
7.2 Wahrscheinlichkeit für die realisierte Stichprobe	34
7.3 Adjustierte Gewichtungsfaktoren	37
7.4 Gesamtselektivität	38

Tabellenverzeichnis

Tabelle 1	Das Studiendesign im Überblick	8
Tabelle 2	Tranchierung der Stichprobe	11
Tabelle 3	Struktur und Inhalt des Erhebungsinstruments	12
Tabelle 4	Screeningausfälle nach Gründen	13
Tabelle 5	Eingesetzte Interviewer nach Geschlecht	17
Tabelle 6	Eingesetzte Interviewer nach Altersgruppen	18
Tabelle 7	Eingesetzte Interviewer nach Schulabschluss	18
Tabelle 8	Vollständig realisierte Interviews pro Interviewer	19
Tabelle 9	Anzahl vollständige Interviews pro Interviewer, gruppiert	19
Tabelle 10	Kontakthäufigkeit der realisierten Interviews	20
Tabelle 11	Kontakthäufigkeit, gruppiert	21
Tabelle 12	Kontakthäufigkeit der realisierten Interviews	21
Tabelle 13	Feldergebnis (Final Outcome), Welle 1	24
Tabelle 14	Outcome Rates	26
Tabelle 15	Panelbereitschaft und Zuspielbereitschaft	27
Tabelle 16	Interviewdauern in Minuten	28
Tabelle 17	Interviewdauern in Minuten, gruppiert	28
Tabelle 18	Interviewabbrüche nach Fragenbereich im Interview, Häufigkeiten	29
Tabelle 19	Ausfallmodellierung Einsatzstichprobe (logistische Regression, odds ratios)	32
Tabelle 20	Ausfallmodellierung realisierte Stichprobe für die Einsatzstichprobe (logistische Regression, odds ratios)	35
Tabelle 21	Kennziffern für die Gewichtungsfaktoren	37
Tabelle 22	Ausfallmodellierung realisierte Stichprobe für die Bruttostichprobe (logistische Regression, odds ratios)	38

1 Erhebungsdesign

1.1 Einordnung des Projekts

Im Rahmen eines umfänglicheren Forschungsvorhabens hat das Institut für Arbeitsmarkt- und Berufsforschung (IAB) das Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ konzipiert und infas mit der Durchführung der Erhebungsarbeiten für eine Beschäftigtenbefragung beauftragt.

Das Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ untersucht in einem Längsschnittansatz von zwei Erhebungswellen, wie sich betriebliche Personal- und Geschäftsstrategien in der unternehmerischen Entwicklung und in der beruflichen Entwicklung ihrer Beschäftigten auswirken. In einem Panelansatz werden geschätzte 600 Betriebe als auch die Beschäftigten mittels Surveys als auch mit Hilfe von Prozess produzierten Daten der BA (IEB, BeH) beobachtet. Die Studie ist als Employer-Employee-Linked-Study angelegt. Die Befragungsdaten von Betriebs- und Beschäftigtenbefragung können miteinander verknüpft werden. Dies erlaubt eine Mehrebenenanalyse, um Entwicklungen und Effekte des Betriebs von denen der beschäftigten Personen zu trennen bzw. um Interaktionseffekte zu modellieren. Die Zustimmung der Betroffenen vorausgesetzt, ist außerdem eine Verknüpfung mit Daten der Beschäftigtenhistorik und der übrigen IEB-Quellen vorgesehen.

Ein solches Forschungsdesign wurde mit einer Fallzahl von rd. 150 Unternehmen bereits erfolgreich von einer Forschungsgemeinschaft von IAB, RWI und infas im Rahmen des Projektes WeLL – Weiterbildung als Bestandteil Lebenslangen Lernens erfolgreich durchgeführt¹. Bei jener Studie stand die Weiterbildung im Mittelpunkt. Untersucht wurde, wie sich das Weiterbildungsverhalten und die Effekte der Weiterbildung der Beschäftigten in denselben Unternehmen unter gegebenen, in der Betriebsbefragung ermittelten Rahmenbedingungen unterscheiden. Mittels vier Erhebungswellen konnten Längsschnittdaten von Verhaltens- und Ergebnisvariablen gemessen und mit den Längsschnittinformationen aus den Prozessdaten der BA verknüpft werden. Dies ermöglichte die Trennung von betrieblichen und individuellen Faktoren und Effekten in Bezug auf das Thema Weiterbildung.

Die Studie „Arbeitsqualität und wirtschaftlicher Erfolg“ operiert mit einem vergleichbaren methodischen Ansatz. Im Zentrum steht allerdings eine wesentlich schwieriger zu messende Wirkung der betrieblichen Personalpolitik auf den unternehmerischen Erfolg und die Entwicklung der Beschäftigten. Die Studie hat das Potenzial, Ergebnisse über nachhaltige Effekte betrieblicher Arbeitsökono-

¹ Vgl. u.a. http://fdz.iab.de/de/Integrated_Establishment_and_Individual_Data/panel_well.aspx; Bender, S., M. Fertig, K. Görlitz, M. Huber, S. Hummelsheim, P. Knerr, A. Schmucker und H. Schröder (2008), WeLL – Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens. RWI Materialien 45. Essen; Bender, S., M. Fertig, K. Görlitz, M. Huber, S. Hummelsheim, P. Knerr, A. Schmucker und H. Schröder (2010), WeLL – Berufliche Weiterbildung als Bestandteil Lebenslangen Lernens – Bisherige Arbeiten und Perspektiven. RWI Materialien 61. Essen

mie und –gestaltung zu produzieren wie sie in dieser Form in Deutschland noch nicht vorliegen.

1.2 Studiendesign im Überblick

Die Panelstudie „Arbeitsqualität und wirtschaftlicher Erfolg“ zielt auf die Untersuchung der Frage ab, wie sich betriebliche Personal- und Geschäftsstrategien in der unternehmerischen Entwicklung und in der beruflichen Entwicklung der Beschäftigten auswirken. Die operative Durchführung der darin enthaltenen Beschäftigtenbefragung liegt in der Verantwortung von infas. In allen Phasen der ersten Welle dieser Panelstudie stand infas dabei in engem Kontakt und Austausch mit dem Auftraggeber IAB sowie den kooperierenden Partnerinstituten ZEW Mannheim und Universität zu Köln.

Die Beschäftigtenbefragung ist in eine Employer-Employee-Linked Study eingebettet. Beteiligt sind Betriebe mit einer Mindestgröße von 50 Beschäftigten, die sich zuvor im Rahmen der IAB-Betriebspanelerhebung 2012 beteiligt haben, bei der Betriebsleitungen bzw. Personalverantwortlichen befragt wurden. Im Anschluss an diese Betriebsbefragung galt es, eine Stichprobe von insgesamt 7.500 Beschäftigten aus diesen Betrieben zu befragen.

Alle sozialversicherungspflichtig Beschäftigten dieser Betriebe sind in der Beschäftigtenhistorik (BeH) als Grundgesamtheit zum Stichtag 31.12. registriert. Für die Auswahl in diesem Vorhaben hat das IAB die Datei zum Stand 31.12.2011 zugrunde gelegt.

Bei dieser Welle des Betriebspanels nahmen insgesamt 869 Betriebe mit mehr als 50 Mitarbeiterinnen und Mitarbeitern teil. Alle Beschäftigten aus diesen Betrieben bildeten die Auswahlgesamtheit für die Personenbefragung der Studie „Arbeitsqualität und wirtschaftlicher Erfolg“.

Die wesentlichen Eckpunkte des Studiendesigns der Beschäftigtenbefragung sind in Tabelle 1 zusammengefasst.

Tabelle 1 Das Studiendesign im Überblick

Grundgesamtheit	Sozialversicherungspflichtig beschäftigte Arbeitnehmer
Auswahlgesamtheit	Sozialversicherungspflichtig beschäftigte Arbeitnehmer aus 869 deutschen Betrieben, die an der Befragungswelle 2012 des Betriebspanels teilgenommen haben
Stichprobe	Pretest: 500 Erwerbspersonen Haupterhebung: 869 Betriebe, 30.703 Erwerbspersonen
Erhebungsmethode	Computergestützte Telefoninterviews (CATI)
Durchführungszeitraum	Haupterhebung: 3.12.2012 bis 3.4.2013 Pretest: 19.11.2012 bis 20.11.2012
Auswertbare Interviews	Haupterhebung: 7.508 Interviews Pretest: 46 Interviews
Interviewdauer	Haupterhebung: Ø 30,2 Minuten
Anzahl Interviewer	Haupterhebung: 236
Schulungsmethode für Interviewer	<ul style="list-style-type: none"> – Grundschulung – Studienspezifische persönliche Schulung aller Interviewer – Schriftliche Studieninformation
Qualitätssicherung im Telefonstudio	<ul style="list-style-type: none"> – Supervision u. bedarfsweise Interviewernachschulungen – Telefonhotline und Projekt-Email – Fallweise Nachverfolgung von Beschwerden durch Zielpersonen
Gewichtung	– Designgewichtung und Kalibrierung für die Gesamtstichprobe
Datenaufbereitung	Datenprüfung: <ul style="list-style-type: none"> – Prüfung der Datenablage vor Feldbeginn – Datenprüfung der realisierten Interviews
Datenlieferung und Dokumentationen	<ul style="list-style-type: none"> – Alle Datensätze in STATA-Format – Befragungsdatensatz, ungewichtet nach Feldende – Methodendatensatz – Enddatensatz, gewichtet – Methodenbericht; mit gewichteten Tabellierungen
Besonderheiten der Studie	Verknüpfung von Betriebsbefragungsdaten und Personenbefragungsdaten (Employer-Employee-Linkage Design)

Quelle: infas

2 Stichprobenkonzept und -auswahl

2.1 Grundgesamtheit und Stichprobenziehung

Alle sozialversicherungspflichtig Beschäftigten der 869 Betriebe, für die eine Jahresmeldung bei der Beschäftigtenhistorik des IAB (BeH) zum 31. Dezember 2011 vorlag, die also zu diesem Stichtag in einem der Betriebe beschäftigt waren, wurden für die Stichprobenziehung berücksichtigt. In der Betriebsbefragung wurden nur Betriebe mit mindestens 50 sozialversicherungspflichtig Beschäftigten befragt. Landwirtschaftliche Betriebe, gemeinnützige Organisationen und der öffentliche Dienst waren hierbei nicht miteinbezogen. Da der Rücklauf aus dem Betriebspanel sehr viel höher als ursprünglich erwartet ausfiel, konnten Daten von 869 Betrieben statt wie ursprünglich geplant von 600 Betrieben als Basis für die Beschäftigtenbefragung herangezogen werden. Vor diesem Hintergrund und aufgrund eines entsprechenden Votums des Auftraggebers BMAS wurde zwischen dem IAB und infas im November 2012 beschlossen, für die Haupterhebung die ursprüngliche Zielgröße von $n = 5.000$ Interviews auf $n = 7.500$ zu erhöhen.

Die Auswahlgesamtheit umfasste insgesamt 43.823 Personen. Für 12.909 von diesen lag keine Telefonnummer vor. Auch unter Nutzung der Adressfactory der Deutschen Post und elektronischen Telefonverzeichnissen war es für diese nicht möglich, eine Telefonnummer zu ermitteln. Von den verbleibenden Telefonnummern waren 211 Angaben doppelt vorhanden, so dass nach Dublettenbereinigung 30.703 Adressen für die Haupterhebung zur Verfügung standen (weiteres unter Abschnitt 2.2.2). Für diese Adressen wurden alle Telefonnummern mittels Adressfactory und den elektronischen Telefonverzeichnissen überprüft und aktualisiert.

Das IAB lieferte die Stichprobe mit folgenden Merkmalen an infas:

- systemfreie Personen-ID
- Alter
- Geschlecht
- Stellung im Beruf
- Tagesentgelt
- systemfreier Betriebsidentifikator
- Bundesland des Arbeitsortes
- Wirtschaftszweig des Betriebs
- Anzahl der zu befragenden Beschäftigten pro Betrieb.

Die zuletzt genannte Schichtungsvariable der Interviews je Betrieb wurde nicht zur Quotierung eingesetzt, sondern ausschließlich zur Feldkontrolle der Zufallsstichprobe verwendet. So wurde vor der Ziehung der dritten Einsatztranche (s.u.) die Variable herangezogen, um die Güte der bisher realisierten Interviewverteilung zu kontrollieren.

2.2 Umfang der Stichprobe

2.2.1 Preteststichprobe

Zu Projektbeginn wurde mit dem IAB verabredet, eine von der Haupterhebungstichprobe gesonderte Preteststichprobe ziehen zu lassen, die auch für die zweite Befragungswelle 2014 genutzt werden kann. Diese vom IAB gelieferte Preteststichprobe umfasste Adressdaten für 817 Personen. Für die Adressen der Preteststichprobe wurden die mitgelieferten und fehlenden Telefonnummern mit der Adressfactory der Deutschen Post und elektronischen Telefonverzeichnissen aktualisiert und recherchiert. Von den 817 gelieferten Adressdaten konnte für 628 eine Telefonnummer recherchiert bzw. aktualisiert werden. Von diesen 628 Adressdaten wurden 500 für den Pretest eingesetzt und standen somit insgesamt zur Verfügung (Bruttostichprobe). Die 500 Adressdaten wurden proportional auf vier Betriebsgrößenklassen (jeweils 125 Personen) verteilt.

2.2.2 Stichprobe der Haupterhebung und Tranchierung

Für die Haupterhebung wurde vom IAB zunächst eine Stichprobe im Umfang von 37.831 Adressen geliefert, von denen nach Durchführung einer Telefonnummernrecherche und der Bereinigung von Dubletten 28.620 Adressen mit Telefonnummern als einsatzfähige Stichprobe verblieben. Von diesen Adressen wurden zunächst zwei Drittel (18.145 Adressen) als erste Tranche der Einsatzstichprobe angeschrieben und im Feld eingesetzt, das verbleibende Drittel (10.475 Adressen) wurde als zweite Tranche unmittelbar nach dem Jahreswechsel ins Feld gebracht. Mit dieser Stichprobe sollten gemäß dem ursprünglichen Vertragsumfang 5.000 Interviews realisiert werden. Tatsächlich war es möglich, mit dieser Stichprobe sogar über 7.000 Interviews umzusetzen, also bereits relativ nahe an die Aufstockungszielmarke von 7.500 Interviews insgesamt zu gelangen. Nichtsdestotrotz wäre diese Stichprobe etwas zu knapp ausgefallen, um damit die Gesamtinterviewfallzahl zu erreichen, weshalb im Februar 2013 infas beim IAB eine kleine Zusatzstichprobe mit rund 2.200 Fällen anforderte. Als dritte Tranche der Einsatzstichprobe ergaben sich 2.083 Adressen mit Telefonnummern, die Ende Februar 2013 angeschrieben und ab März antelefoniert wurden.

2.2.3 Tranchierung der Einsatzstichproben

Um eine maximale Ausschöpfung der Stichprobe sicherzustellen, wurden die Adressen nicht alle zeitgleich, sondern in Tranchen eingesetzt. Die ursprüngliche Einsatzstichprobe von 28.620 Adressen kam in zwei Tranchen zum Einsatz, die dritte Tranche bildeten die Adressen der Stichprobenergänzungslieferung, sofern eine gültige Telefonnummer vorlag. Dies traf für 2.083 Fälle zu. Mit knapp sieben Prozent aller Einsatzadressen macht die dritte Tranche insgesamt nur einen geringen Teil der Gesamtstichprobe aus (Tabelle 2).

Tabelle 2 Tranchierung der Stichprobe

Tranche	absolut	relativ (%)	Einsatz ab...
1	18.145	59,1	Dezember 2012
2	10.475	34,1	Januar 2013
3	2.083	6,8	März 2013
Summe	30.703	100	

Quelle: infas

3 Erhebungsinstrument

3.1 Inhalte und Struktur des Fragebogens

Der Fragebogen für die CATI-Erhebung wurde unter Leitung des Instituts für Arbeitsmarkt und Berufsforschung (IAB) in enger Zusammenarbeit mit dem Zentrum für europäische Wirtschaftsforschung (ZEW), dem Lehrstuhl für ABWL und Personalwirtschaftslehre der Wirtschafts- und Sozialwissenschaftlichen Fakultät der Universität zu Köln und infas inhaltlich konzipiert und entwickelt. Die Vorlage wurde sodann in der Verantwortung und Federführung von infas in ein feldfähiges Befragungsinstrument überführt.

Der Fragebogen in der Erwerbspersonenbefragung im Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ enthält ein breitgefächertes Fragespektrum gemäß der Vielfalt und Komplexität der zu berücksichtigenden Themenaspekte. Tabelle 3 fasst die Inhalte des Fragebogens überblicksartig zusammen.

Tabelle 3 Struktur und Inhalt des Erhebungsinstruments

Frageblock	Inhalt
100: Erwerbstätigkeit	Berufliche Stellung Zeitliche Befristung des Beschäftigungsverhältnisses vertraglich geregelte und reale Arbeitszeit Heimarbeit
200: Personalentwicklung	Weiterbildungsmaßnahmen Mitarbeitergespräche Weiterentwicklung und Beförderungen Arbeitsplatzsicherheit
300: Arbeitsbedingungen und Arbeitsbelastungen	Job-Charakteristika Work-Life-Balance
400: Vergütung	Leistungs- und erfolgsabhängige Gehaltsbestandteile Sonderzahlungen Betriebliche Altersversorgung
500: Bindung, Werte und Unternehmenskultur	Unternehmenskultur Betriebsbindung / Commitment Kollegialität Arbeitszufriedenheit Einkommenszufriedenheit Gerechtigkeit Emotionales Engagement bei der Arbeit
600: Persönlichkeit und Einstellungen	Vertrauen Ungerechtigkeitssensibilität Persönlichkeitsmerkmale (Big Five) Risikoverhalten
700: Gesundheit	Aktueller Gesundheitszustand Krankheitstage Wohlbefinden
800: Soziodemographie	Bildungsabschluss Familienstand Haushaltsgröße Kinder Migrationshintergrund Brutto-/Nettolohn

Quelle: infas

3.2 Ansprache der Zielperson und Screening

Gemäß guter und standardmäßiger Praxis stellten sich die Telefoninterviewer zu Beginn des telefonischen Kontakts zunächst mit Namen und der Nennung des Arbeitgebers vor, um sodann unter Nennung des Studientitels und des Auftraggebers BMAS Studie die Teilnahme an dem Interview anzufragen. In diesem Kontext wurde auch auf das briefliche Anschreiben verwiesen, das diesen Anruf bereits angekündigt hatte. Um teilnehmerbezogene Selbstselektionen möglichst gering zu halten, wurde in der brieflichen und telefonischen Ansprache der Zielpersonen die Studie unter dem allgemeineren Titel „Arbeit und Betrieb“ kommuniziert. Auch die Erläuterung der Studieninhalte war recht allgemein gehalten.

ten und verzichtete bewusst auf die Nennung analytisch relevanter Dimensionen, war aber zugleich mit dem Hinweis auf die Untersuchung der Arbeitsbedingungen von Beschäftigten so interessant gehalten, dass dadurch zur Mitwirkung an der Befragung motiviert werden konnte.

Nach dem eigentlichen Interviewstart wurden dann zunächst Geburts- und Geschlechtsangaben erfragt und diese programmtechnisch mit den vorliegenden Angaben aus der Stichprobe verglichen und überprüft, dass sich tatsächlich die richtige, allein zulässige Zielperson am Apparat befand und an der Studie beteiligt. Ergaben sich im Rahmen dieses Screenings Abweichungen, wies die Programmsteuerung die Interviewer zu Rückfragen an, um die Angaben ggf. zu korrigieren und sie im System zu erfassen. Handelte es sich nach Abgleich mit der Stichprobeninformation nicht um die richtige Zielperson, so war das Interview an dieser Stelle zu beenden.

Überprüft wurde auch, ob die Zielperson weiterhin bei demselben Betrieb tätig war wie im Dezember 2011, worauf die Stichprobeninformation basierte. War die Zielperson dort unterdessen nicht mehr erwerbstätig, konnte das Interview nicht geführt werden. Screeningausfälle aufgrund dieses Kriteriums umfassten 843 Personen (inklusive einer „weiß nicht“-Nennung). Die Gesamtübersicht zu den Screeningausfallgründen findet sich in Tabelle 4.

Tabelle 4 Screeningausfälle nach Gründen

	abs.	in %
ZP aktuell nicht erwerbstätig	249	21,2
Geburtsjahr verweigert	23	2,0
Geburtsjahr weiß nicht	2	0,2
Abweichung Geburtsdatum von Stichprobe: Es handelt sich nicht um die ZP	4	0,3
Noch im selben Betrieb wie im Dezember 2011 beschäftigt: Nein	842	71,8
Noch im selben Betrieb wie im Dezember 2011 beschäftigt: Weiß nicht	1	0,1
Berufliche Stellung: Beamter	2	0,2
Berufliche Stellung: Selbständiger oder Freiberufler	6	0,5
Berufliche Stellung: Mithelfender Familienangehöriger	2	0,2
Hotline- oder email-Rücklauf: ZP meldet, nicht zur Zielgruppe zu gehören	41	3,5
Insgesamt	1172	100

Quelle: Personenbefragung (Welle 1)

3.3 Fragebogenprogrammierung

Für die Programmierung von Fragebögen verwendet infas die CATI-/CAPI-Software ODIN von NIPO. Die vielfältigen Möglichkeiten, die bei Programmierungen mit dieser Software berücksichtigt werden können, tragen wesentlich zur Erleichterung der Durchführbarkeit von telefonischen Interviews sowie zur Vollständigkeit und Qualität der erhobenen Daten bei. Fehlerquellen werden minimiert und der Umfang nachträglicher Prüf- und Bereinigungsarbeiten wird deutlich reduziert.

Die Umsetzung komplexer Filter- und Steuerungsstrukturen ist dabei sehr gut möglich. Durch die Programmierung aller Filtersprünge wird der Interviewer automatisch durch das Instrument geführt, subjektive Filterfehler werden vollständig ausgeschlossen. Die Interviewer können sich auf diese Weise vollständig auf das Vorlesen der Fragen und das Notieren der Antworten der Zielpersonen konzentrieren.

Die bei CATI-Erhebungen im Hintergrund der Befragung laufende Programmierung erlaubt außerdem eine Reihe von Kontrollen und Steuerungen während des Interviews. Diese Vorteile konnten zum Teil auch bei der Durchführung der Beschäftigtenbefragung genutzt werden:

- *Farbleitsystem*: Durch die farbliche Gestaltung der Texte auf dem Bildschirm wird die Interviewführung erleichtert. Fest definierte Farbuordnungen lassen die Interviewer jederzeit eindeutig den vorzulesenden Interviewtext von Intervieweranweisungen sowie anderen Funktionen unterscheiden
- *Zielgruppenspezifische Steuerung von Fragen*: Fragen, die zielgruppenspezifisch gestellt werden sollten, können programmtechnisch gesteuert werden. Dies erleichtert sowohl den Ablauf der Befragung als auch die Koordination des CATI-Feldes.
- *Rotation und Randomisierung von Itembatterien*: Bei Itembatterien können die Items je nach Bedarf rotiert oder zufallsgesteuert eingespielt werden. Zudem kann die Einblendung von einzelnen Items einer Itembatterie individuell über Filterbedingungen gesteuert werden, so dass z.B. eine Verknüpfung von zuvor erfassten Antworten mit weiteren Bedingungen vorgenommen werden kann, die dann das Einblenden der betreffenden Fragen steuert. In dieser Befragung wurden auf Wunsch der beteiligten Auswertungsinstitute keine Rotationen und Randomisierungen eingerichtet.
- *Rangprüfungen*: Wertebereiche können auf Plausibilität geprüft und mit anderen Angaben verglichen werden. Mit der Prüfung von Wertebereichen (Programmierung gültiger Wertebereiche) wird der spätere Bereinigungsaufwand deutlich verringert. Dies gilt z.B. für Jahres- und Prozentangaben. Solche Prüfungen wurden entsprechend den Festlegungen des Auftraggebers in die programmierte Fassung des Fragebogens eingearbeitet und kamen bei dieser Befragung relativ häufig zum Einsatz.
- *Zeitpunktspezifische Steuerung von Wertebereichsprüfungen*: Der Wertebereich einer Variable kann in Abhängigkeit von einem bestimmten Zeitpunkt gesteuert werden. Dieses Feature war für diese Befragung nicht relevant.

Der programmierte Fragebogen wurde dann in einem ersten Prüfschritt bei infas ausführlich getestet. Dazu wurden virtuelle Fälle über das System erzeugt, für die eine Überprüfung der Datenablage stattfand. Darüber hinaus erfolgten manuelle Testeingaben speziell geschulter Fragebogentester, die die korrekte Umsetzung anhand der Programmiervorlage überprüften (Fragensukzession und Filterführung, Texteinblendungen, Hilfstexte etc.).

3.4 Pretest

Der Pretest im Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ fand vom 19.11. bis 20.11.2012 statt. Im Rahmen dieses Pretests wurde die Feldfähigkeit des Erhebungsinstrumentes unter realen Feldbedingungen getestet. Überprüft wurden insbesondere die Verständlichkeit und Akzeptanz der Fragen bei den Zielpersonen sowie die korrekte technische Umsetzung der Fragebogenvorlage. Außerdem galt es, die durchschnittliche Dauer der Erhebung bzw. einzelner Fragebogenmodule zu ermitteln, um ggf. die Länge des Fragebogens für die Haupterhebung noch anpassen zu können.

Im Rahmen einer zufälligen Stichprobenzuordnung (Split-Half-Design) wurde außerdem überprüft, an welcher Stelle im Fragebogen die Frage zur Zuspielungsbereitschaft von BA-Daten die besseren Zustimmungswerte erhält. Hier standen zwei Auffassungen und entsprechend zwei Umsetzungsoptionen zur Prüfung an: In der einen Variante gegen Ende des Fragebogens und in direktem Anschluss an die Panelfrage, in der zweiten Variante abgekoppelt von der Panelfrage und ungefähr in der Mitte des Fragebogens an einer thematischen Übergangsstelle platziert. Der Pretest ergab im Ergebnis, dass die Abfrage der Zuspielungsbereitschaft in direktem Anschluss an die Frage zur Panelbereitschaft etwas höhere Zustimmungswerte erzielte als im Falle der Platzierung in der Mitte des Interviews. Daher wurde für die Haupterhebung entschieden, Panel- und Zuspielungsbereitschaft zusammen gegen Interviewende einzuholen.

Die Interviews wurden zentral im infas-Telefonstudio in Bonn durchgeführt. Im Vorfeld fand eine persönliche Schulung der Pretest-Interviewer durch die infas-Projektleitung statt. Insgesamt wurden zehn Interviewer/innen mündlich geschult.

An den zwei Feldtagen wurden 46 Pretest-Interviews durchgeführt. Die Dauer der Interviews betrug durchschnittlich 29,8 Minuten; das kürzeste Interview dauerte 21, das längste 49 Minuten.

Im infas-Telefonstudio besteht für Kunden und Projektleiter die Möglichkeit, den Verlauf von Interviews direkt zu verfolgen. In einem eigens dafür eingerichteten Raum können Interviews mitgehört und die Eingaben der Interviewer mitverfolgt werden. Diese Gelegenheit wurde von allen beteiligten Forschungsinstituten (IAB, ZEW, Uni Köln) gemeinsam mit der infas-Projektleitung am 20. November 2012 genutzt, so dass Auffälligkeiten in den mitgehörten Interviews unmittelbar miteinander diskutiert und bestimmte Änderungen schon vorsondiert abgestimmt werden konnten.

Der Pretest konnte die Feldfähigkeit und Funktionalität des Fragebogens im Wesentlichen bestätigen. Für die Hauptstudie wurde das Erhebungsinstrument in wenigen Punkten angepasst und optimiert.

Die zwischen infas und Auftraggeber abgestimmten Änderungen wurden dann für die Haupterhebung in die programmierte Fassung übernommen. Anschließend wurden die Änderungen sowohl durch infas als auch durch den Auftraggeber noch einmal getestet.

4 Felddurchführung der Haupterhebung

4.1 Feldzeit und Feldsteuerung

Die Feldzeit in der Haupterhebung dauerte vom 3. Dezember 2012 bis zum 3. April 2013.

Die Zielpersonen jeder Tranche erhielten vor der telefonischen Kontaktaufnahme ein Anschreiben zusammen mit einer Datenschutzerklärung zugesandt. Das Anschreiben informierte über die Ziele der Studie und warb um die Mitwirkung an der Befragung. Des Weiteren enthielt das Anschreiben für etwaige Rückfragen die kostenfreie Servicetelefonnummer von infas, eine projektspezifische Email-Adresse sowie auch eine Telefonnummer der zuständigen Kontaktperson des IAB. Gemäß der geltenden Datenschutzbestimmungen war dem Anschreiben ein Datenschutzblatt beigelegt, das über die anonymisierte Verwendung, Nichtweitergabe und spätere Löschung der erhobenen Daten informiert. Die Datenschutzerklärung war von allen Datenschützern der beteiligten Institute unterzeichnet. Die Homepage von infas bot den Zielpersonen zudem eine weitere Informationsmöglichkeit, da dort eine Kurzbeschreibung der Studie hinterlegt war.

Alle mitwirkenden Befragungspersonen der ersten Welle erhielten wenige Wochen nach dem Abschluss des Interviews ein persönlich adressiertes Dankschreiben für die Teilnahme an der Studie. Bei den panelbereiten Teilnehmern der Studie war dem Dankschreiben zudem ein Adressänderungsblatt beigelegt, um etwaige Adressänderungen einzutragen und an infas zurückzuschicken. Von dieser Möglichkeit der schriftlichen Adressaktualisierung machten bis Mitte Mai 2013 bereits 75 Personen Gebrauch, was die Sinnhaftigkeit dieser Maßnahme der Panelpflege unterstreicht.

Während des laufenden Feldes der Haupterhebung nutzt infas standardmäßig ein Reportingsystem, das tagesaktuell den Stand der Feldarbeit mit unterschiedlichen Kennziffern abbildet. Mit Hilfe dieses Reportings werden u.a. Ausschöpfungsstatistiken und andere Kennziffern errechnet, die wichtige Hinweise zur Qualitätseinschätzung des laufenden Feldes liefern. Neben der Abbildung der realisierten Interviews und der anderen Kontaktstatus (keine Kontakte, Verweigerungen, usw.) im Rahmen der Stichprobenbearbeitung sind darin u.a. auch Indikatoren für die durchschnittliche Interviewdauer oder Kontakthäufigkeiten enthalten. Diese Informationen dienen der laufenden Feldkontrolle (Monitoring)

und ermöglichen es, nach Bedarf und zeitnah die Feldarbeit nachzusteuern und anzupassen, wenn dies erforderlich oder geboten erscheint.

4.2 Interviewerschulungen

Bei infas werden alle Interviewer in einer Grundschulung in den Techniken des standardisierten Interviewens unterrichtet. In regelmäßigen Abständen werden diese Kenntnisse aufgefrischt. Darüber hinaus gehört es zum guten Standard jeder infas-Studie, jeweils projektbezogene Schulungen der eingesetzten Interviewer vorzunehmen. Auch im Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ wurden entsprechend die Interviewer zum Pretest sowie vor Feldstart der Hauptstudie durch die Projektleitung in persönlich-mündlichen Schulungen projektspezifisch geschult. Hierzu gehörten zunächst Informationen zu den wichtigsten Rahmendaten der Studie (Herkunft der Stichprobe, Zielgruppe der Befragung, Feldzeit, Interviewdauer etc.).

Im Anschluss daran wurde als Kernstück der Schulung der Fragebogen anhand eines durch die Projektleitung entwickelten Beispielfalls gemeinsam mit den Interviewern durchgearbeitet. Der Beispielfall war so konzipiert, dass er das ganze Fragenprogramm des Instruments abdeckte. Im Anschluss an die Schulung wurde den Interviewern die Möglichkeit gegeben, den Fragebogen selbständig in einer Testversion des Instruments noch einmal durchzugehen. Ergänzend zur mündlichen Schulung erhielt jeder Interviewer ein zusammenfassendes Papier mit den wichtigsten Informationen zur Studie zum möglichen Nachschlagen. Die Pretest-Schulung wurde in Anwesenheit einer IAB-Vertreterin durchgeführt.

4.3 Interviewereinsatz

Im Projekt „Arbeitsqualität und wirtschaftlicher Erfolg“ haben insgesamt 236 Interviewerinnen und Interviewer mindestens ein Interview vollständig realisiert. Durch den Einsatz einer vergleichsweise großen Zahl von Interviewern war es möglich, die nachträglich um 2.500 Fälle erhöhte Interviewzahl in einer Feldzeit von insgesamt vier Monaten zu realisieren.

Die Interviews wurden von ungefähr gleich vielen Männern wie Frauen durchgeführt (vgl. Tabelle 5). Der Frauenanteil an allen Interviewern ist geringfügig höher als der Männeranteil.

Tabelle 5 Eingesetzte Interviewer nach Geschlecht

Geschlecht	Absolut	Prozent
Männlich	114	48,3
Weiblich	122	51,6
Insgesamt	236	100,0

Quelle: eigene Berechnungen

Die eingesetzten Interviewerinnen und Interviewer verteilten sich über alle Altersgruppen (vgl. Tabelle 6). Zwei Drittel aller Interviewer entfallen auf die zwei jüngsten Alterskohorten von 18 bis 34 Jahre. Die mittleren Altersgruppen von 35 bis 44 Jahre bzw. 45 bis 54 Jahre decken zusammen einen Anteil von einem Fünftel aller Interviewer ab. Lediglich sieben Prozent der Interviewer waren zwischen 55 und 64 Jahre alt. Die restlichen drei Prozent sind in der Altersklasse 65 Jahre und älter.

Tabelle 6 Eingesetzte Interviewer nach Altersgruppen

Alter	Absolut	in Prozent
18 bis 24 Jahre	104	44,1
25 bis 34 Jahre	58	24,6
35 bis 44 Jahre	19	8,1
45 bis 54 Jahre	30	12,7
55 bis 64 Jahre	17	7,2
65 Jahre und älter	8	3,4
Insgesamt	236	100,0

Quelle: eigene Berechnungen

Das Bildungsniveau der eingesetzten Interviewer ist im Vergleich zum Durchschnitt der Bevölkerung überdurchschnittlich hoch (vgl. Tabelle 7). Zusammengekommen verfügten gut 70 Prozent der eingesetzten Interviewer zumindest über eine Fachhochschulberechtigung. Knapp 18 Prozent der Interviewer haben die Mittlere Reife. Lediglich knapp vier Prozent verfügen über einen Haupt- oder Volksschulabschluss bzw. über einen Abschluss an einer Polytechnischen Oberschule und rund sieben Prozent der Interviewer haben einen anderen (z.B. ausländischen) Schulabschluss.

Tabelle 7 Eingesetzte Interviewer nach Schulabschluss

Höchster Schulabschluss	Absolut	in Prozent
Hauptschul-/Volksschulabschluss/ POS	9	3,8
Mittlere Reife	42	17,8
Fachhochschulreife, Abitur/Hochschulreife	164	69,5
sonstiger Schulabschluss	16	6,8
keine Angabe	5	2,1
Insgesamt	236	100,0

Quelle: eigene Berechnungen

Im Durchschnitt führte jeder der 236 Telefoninterviewer/innen rund 32 Interviews (Standardabweichung 31,6) (vgl. Tabelle 8). Die minimale Anzahl liegt bei

einem vollständigen Interview. Maximal wurden von einem Interviewer 174 Interviews (rund 2 Prozent aller Interviews) durchgeführt.

Tabelle 8 Vollständig realisierte Interviews pro Interviewer

	Anzahl
Durchschnittliche Interviewanzahl pro Interviewer	32
Maximale Interviewanzahl pro Interviewer	174
Minimale Interviewanzahl pro Interviewer	1
Standardabweichung	31,6
Anzahl Interviewer mit mindestens einem realisierten Interview	236

Quelle: eigene Berechnungen

Dabei führten gut die Hälfte aller Interviewer ein bis zwanzig Interviews (vgl. Tabelle 9). Ein Viertel der Interviewer erbrachten jeweils zwischen 21 und 40 Interviews und das verbleibende Viertel hatte höhere Interviewzahlen pro Kopf. Immerhin insgesamt 17 Prozent der Interviewer/innen haben dabei jeweils 60 und mehr Interviews absolviert.

Tabelle 9 Anzahl vollständige Interviews pro Interviewer, gruppiert

Interviews	Absolut	in Prozent
1 bis 10 Interviews	69	29,2
11 bis 20 Interviews	43	18,2
21 bis 40 Interviews	58	24,6
41 bis 60 Interviews	26	11,0
61 bis 100 Interviews	30	12,7
über 101 Interviews	10	4,2
Insgesamt	236	100,0

Quelle: eigene Berechnungen

4.4 Dokumentation des Feldverlaufs

Um ein hohes Ausschöpfungsziel zu erreichen und rasch auf mögliche Feldprobleme reagieren zu können, muss der Feldprozess jederzeit überschaut und vollständig transparent gehalten werden. Ein zentrales Instrument dieser notwendigen kontinuierlichen Feldkontrolle bildet das Monitoring einer CATI-Kontaktdatei, in der sämtliche Kontaktversuche (und die Kontaktanzahl je Fall) gespeichert werden. Die Analyse dieser Kontaktdateien stellt Transparenz über den Feldprozess her und erlaubt eine kontinuierliche Bewertung des Feldstands.

Die Grundlage für die Rücklaufbearbeitung bildet das infas-Sample-Management-System (iSMS), das die Informationen der Stichprobe in einer Da-

tenbank verwaltet. Darin liegen alle Telefonnummern der Ausgangsstichprobe in einer zentralen Tabelle. Aus dieser Stichprobe können weitere Tabellen sowohl als Import- als auch als Exporttabellen erstellt werden. Exporttabellen werden benötigt, um unter anderem Einsatzstichproben für das Telefonstudio zu definieren. Der Tabellenaufbau und -umfang wird an die jeweiligen Studienbedürfnisse angepasst.

Im Lauf der Erhebung der „IAB-Beschäftigtenbefragung“ wurde für jeden Kontakt bzw. Kontaktversuch der jeweilige Bearbeitungsstatus festgehalten. So entstand im Feldverlauf eine Historie für jede zu kontaktierende Adresse der Stichprobe. Jeder dieser Kontakte wird im Adressverwaltungssystem (iSMS) protokolliert, unabhängig davon, ob der Kontakt erfolgreich war oder nicht. Durch dieses stetige Nachhalten und die Dokumentation der Ausfallgründe ist tagesaktuell der Bearbeitungsstand und -status aller Adressen abrufbar und somit eine Bewertung des Feldstandes möglich.

4.5 Kontaktierung im Feldverlauf

Die Zielpersonen wurden montags bis freitags in der Zeit von 13 bis 21 Uhr kontaktiert und samstags zwischen 10 Uhr und 18 Uhr. War die Möglichkeit der Kontaktaufnahme im laufenden Feld nicht gegeben oder wurde das Interview aus diversen Gründen verweigert, wurde dies entsprechend in eine der Subkategorien systematischer und neutraler Ausfälle verbucht.

Insgesamt wurden im Rahmen der Studie über die gesamte Feldzeit über eine halbe Million Kontaktversuche unternommen. Bezogen auf die gesamte Bruttostichprobe wurde jede eingesetzte Telefonnummer durchschnittlich siebzehnmals kontaktiert bis der finale Feldstand feststand. Die maximale Kontaktanzahl einer Telefonnummer liegt bei 137 Kontaktversuchen (vgl. Tabelle 10).

Tabelle 10 Kontakthäufigkeit der realisierten Interviews

Kontakthäufigkeit	Anzahl
Durchschnittliche Kontakthäufigkeit	17,0
Minimum Kontakthäufigkeit	1
Maximum Kontakthäufigkeit	137
Standardabweichung	25,1
Fälle	30642

Basis: Kontaktdaten/Methodendatensatz; nur Fälle mit mindestens einem Kontaktversuch

Die Anzahl der Kontaktversuche streut allerdings beträchtlich, für den Großteil der Stichprobe lag die Kontakthäufigkeit zwischen 1 und 20 (Tabelle 11).

Tabelle 11 Kontakthäufigkeit, gruppiert

	abs.	in %
1-10 Kontakte	19906	64,8
11-20 Kontakte	4068	13,2
21-50 Kontakte	3233	10,5
51-100 Kontakte	3175	10,3
Mehr als 100 Kontakte	321	1,0
Insgesamt	30703	100

Basis: Kontaktdaten/Methodendatensatz

Zur Realisierung eines Interviews waren durchschnittlich acht Kontakte erforderlich (vgl. Tabelle 12).

Tabelle 12 Kontakthäufigkeit der realisierten Interviews

Kontakthäufigkeit	Anzahl
Durchschnittliche Kontakthäufigkeit	8,0
Minimum Kontakthäufigkeit	1
Maximum Kontakthäufigkeit	96
Standardabweichung	9,5
Fälle	7508

Basis: Kontaktdaten/Methodendatensatz

4.6 Supervision zur Qualitätssicherung

Durch den Einsatz qualitätssichernder Maßnahmen wird bei jeder CATI-Studie gewährleistet, dass die Regeln des standardisierten Interviews eingehalten werden und der Ablauf der Studie insgesamt reibungslos verläuft.

Ein wichtiges Element der Qualitätssicherung bildet die Supervision des infas-Telefonstudios, die durch ein speziell geschultes Supervisorenteam geleistet wird. Zielsetzung der Supervision ist es, durch kontinuierliche Betreuung und persönliche Ansprechbarkeit eine hohe Qualität der Befragungsdaten zu erhalten und die Ausschöpfung einer Studie durch adäquate Kontaktaufnahme zu maximieren. Das Supervisionsteam ist mit den methodischen Anforderungen komplexer sozialwissenschaftlicher Projekte und der CATI-Befragungstechnik bestens vertraut. Für die Zwecke dieser Studie erhielten die Supervisoren zudem eine studienspezifische Schulung. Gemäß der Arbeitsweise im Telefonstudio

arbeiten die Mitarbeiterinnen und Mitarbeiter des Supervisionsteam in einem Schichtbetrieb, der gewährleistet, dass pro Schicht jeweils ca. vier bis fünf Supervisoren anwesend und ansprechbar sind.

Zum Kerngeschäft und den häufigsten Aufgaben der Supervision gehörten auch bei dieser Studie die Kontrolle durch Mithören der Interviews im Telefonstudio und Überprüfung der Eingaben durch Beobachtung auf der Supervisionsmaske. Dies wird ermöglicht durch die Übertragung der Anzeigen und Eingaben des Interviewermonitors auf den Supervisionsarbeitsplatz. Bei Rückmeldungen von Problemen, oder Auffälligkeiten bzw. Fehlern der Interviewer/innen erfolgten in der Regel direkt im Anschluss an das Interview Rückkopplungsgespräche und Korrekturhinweise durch die Supervision. Ggf. leisteten die Supervisoren also persönliche Interviewernachschulungen. Die Notwendigkeit dazu war bei dieser Studie aber recht gering. Wenig Nachfrage bestand bei dieser CATI-Welle auch an Klärungen von inhaltlichen Fragen oder Problemen durch die infas-Projektleitung, die durch die Supervision in solchen Fällen regelmäßig eingeschaltet wird. In der Gesamtschau verlief die Supervision für diese Studie also reibungslos und ohne nennenswerte Probleme oder Auffälligkeiten.

5 Feldergebnis

5.1 Final Outcome

Differenzierte Berechnungen des Feldergebnisses und der Stichprobenausschöpfung orientieren sich bei infas an den Standards der *American Association for Public Opinion Research (AAPOR)*.² Danach wird der *Final Outcome* in folgende Kategorien ausdifferenziert:

- Zielgruppenzugehörigkeit unbekannt (UE – Unknown Eligibility)
- Nicht Zielgruppe (NE – Not Eligible)
- Nonresponse – Nicht erreicht (NR-NC – Nonresponse-Non-Contact)
- Nonresponse – Nicht befragbar (NR-NA – Nonresponse-Not able)
- Nonresponse – Sonstiges (NR-O – Nonresponse-Other)
- Nonresponse – Nicht teilnahmebereit (NR-R – Nonresponse-Refusal)
- Vollständig realisiertes Interview (I – Complete Interview)
- Unvollständig realisiertes Interview (IP – Partial Interview)

Die Bearbeitungscode des *Final Outcome* bilden den endgültigen Bearbeitungsstatus einer Adresse ab. Hierzu wird zwischen den Feldcodes innerhalb eines Kontaktverlaufs eine Priorisierung vorgenommen. In der Regel gilt dabei der Status aus dem letzten Sprachkontakt zum Haushalt, auch wenn danach noch vergebliche Kontaktversuche unternommen worden sein sollten. So werden etwa mündliche Terminvereinbarungen oder persönlich erteilte Teilnahmever-

² Vgl. American Association for Public Opinion Research (2011): Standard Definitions. Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised Version 2011.

weigerungen als Feldendstatus ausgewiesen, selbst wenn danach ein Interviewer noch vergeblich versucht haben sollte, die Zielperson zu erreichen.

In die AAPOR-Kategorie „Zielgruppenzugehörigkeit unbekannt“ (UE) werden Nummern eingeordnet, in denen trotz mehrfacher Kontaktversuche kein Sprachkontakt hergestellt werden konnte. In diesen Fällen konnte daher die Zugehörigkeit zur Zielgruppe nicht geklärt werden. Unter der Kategorie „Nicht Zielgruppe“ (NE) sind insb. Screeningausfälle verbucht, d.h. Personen, die die Mitwirkungskriterien für die Befragung nicht erfüllten. Außerdem sind hier auch verstorbene oder ins Ausland verzogene Zielpersonen einsortiert.

Die Gruppe der „Nonresponses“ (NR) umfasst mit Ausnahme der Fälle Nonresponse-Nicht erreicht (NR-NC) alle Ausfälle von Personen, bei denen die Zielgruppenzugehörigkeit definitiv festgestellt werden konnte. Die Nonresponse-Ausfälle werden insgesamt differenziert nach „nicht erreicht“, „sonstiges“, „nicht befragbar“ und „nicht teilnahmebereit“.

Unter den nicht erreichten Zielpersonen (NR-NC) sind alle Personen verbucht, mit denen unter der vorhandenen und geschalteten Nummer kein telefonischer Sprachkontakt hergestellt werden konnte.

Unter Nonresponse-Sonstiges (NR-O) befinden sich alle Personen, bei denen in einem Kontakt (mit der Zielperson oder einen anderen Haushaltsmitglied) festgestellt wurde, dass sie zur Zielgruppe der Studie gehören. Es wurde dann jedoch nicht unmittelbar ein Interview realisiert, sondern z.B. ein Wiederanruf zu einem günstigeren Zeitpunkt vereinbart. Die anschließenden weiteren Kontaktversuche blieben dann jedoch erfolglos. Außerdem fallen in diese Kategorie alle Personen, die nach Auskunft einer Kontaktperson während der gesamten Feldzeit nicht erreichbar sind. Zudem sind Zielpersonen, die für ein Telefoninterview nicht ausreichende Deutschkenntnisse besaßen, in dieser Rubrik verbucht.

In die Kategorie „nicht befragbar“ fallen Personen, die zur Zielgruppe gehören, aber aus gesundheitlichen Gründen (Krankheit, körperliche Einschränkung oder Behinderung) nicht an einem Interview teilnehmen können. Bei dieser CATI-Befragung war diese Kategorie nicht gefüllt. Die Gruppe der „nicht teilnahmebereiten“ Personen umfasst schließlich alle Personen, die trotz intensiver Bemühungen der Interviewer nicht zu einem Interview bereit waren. Diese Kategorie beinhaltet hier 13 unterschiedliche Begründungen für die Nichtteilnahme an einem Interview. In der Gruppe der realisierten Interviews werden sowohl die vollständig als auch die unvollständig realisierten Interviews ausgewiesen. Bei den unvollständigen, d.h. ungültigen Interviews handelt es sich um unterbrochene Interviews, die jedoch bis zum Feldende nicht wieder aufgenommen werden konnten.

Entsprechend dieser Kategorisierung ergibt sich das in Tabelle 13 dargestellte Feldergebnis.

Tabelle 13 Feldergebnis (Final Outcome), Welle 1

	abs.	in %	abs.	in %	abs.	in %
	I	II	III	IV	V	VI
Bruttostichprobe	30703	100				
Zielgruppenzugehörigkeit unbekannt (UE)	6283	20,5				
Kein Anschluss	3695	12,0				
unter Anschluss nur Fax/Modem	142	0,5				
falsche Telefonnummer/ ZP unter Anschluss unbekannt	1901	6,2				
ZP wohnt da nicht mehr/neue Anschrift unbekannt	545	1,8				
ZP nicht in Zielgruppe (NE) / außerhalb Grundgesamtheit	1234	4,0				
ZP nicht in Zielgruppe	1172	3,8				
ZP ins Ausland verzogen	4	0,0				
ZP verstorben / Betrieb erloschen	58	0,2				
Bereinigtes Brutto (I): Zielgruppe	23186	75,5	23186	100		
Nonresponse - Nicht befragbar (NR-NA)	144	0,5	144	0,6		
Nonresponse - Nicht erreicht (NR-NC)	11	0,0	11	0,0		
Nicht abgehoben/nicht erreicht	7	0,0	7	0,0		
Anrufbeantworter	3	0,0	3	0,0		
Besetzt/ Anschluss vorübergehend nicht erreichbar	1	0,0	1	0,0		
Nonresponse - Sonstiges (NR-O)	4264	13,9	4264	18,4		
Wiederanruf vereinbart	2773	9,0	2773	12,0		
erneutes Anschreiben auf Wunsch	63	0,2	63	0,3		
definitiver Termin	259	0,8	259	1,1		
ZP in Feldzeit nicht zu erreichen	652	2,1	652	2,8		
keine Verständigung möglich/ Deutsch nicht ausreichend	517	1,7	517	2,2		
Bereinigtes Brutto (II): Zielperson erreicht und befragbar	22006	71,7	22006	94,9	22006	100
Nonresponse - Nicht teilnahmebereit (NR-R)	11208	36,5	11208	48,3	11208	50,9
ZP verweigert grundsätzlich	2989	9,7	2989	12,9	2989	13,6
ZP verweigert: keine Zeit/ dauert zu lange/ wird zu viel	650	2,1	650	2,8	650	3,0
ZP verweigert: will nicht am Telefon befragt werden	535	1,7	535	2,3	535	2,4
ZP verweigert: krank	58	0,2	58	0,3	58	0,3
ZP verweigert Start / keine Einwilligung zum Interview	105	0,3	105	0,5	105	0,5
ZP verweigert neue Anschrift	1	0,0	1	0,0	1	0,0
Abbruch im Fragebogen	177	0,6	177	0,7	177	0,8
legt sofort auf	2083	6,8	2083	9,0	2083	9,5
kein Zugang zu ZP/ Zugang verhindert/ Teilnahme untersagt	1459	4,8	1459	6,3	1459	6,6
KP verweigert jegliche Auskunft	108	0,4	108	0,5	108	0,5
ZP verweigert: kein Interesse/ Thema	1897	6,2	1897	8,2	1897	8,6
ZP verweigert: Datenschutzgründe/ zu persönlich	280	0,9	280	1,2	280	1,3
ZP verweigert: sonstige Gründe	876	2,9	876	3,8	876	4,0
Interviews, insgesamt	7559	24,6	7559	32,6	7559	34,3
Interviews, gültige	7508	24,5	7508	32,4	7508	34,1
Interviews, ungültige	51	0,1	51	0,2	51	0,2

Basis: CATI-Befragung, Feldendstand (Final Outcome)

Tabelle 13 belegt zunächst, dass ein gutes Viertel der Bruttostichprobe für die Zwecke der Befragung wegfiel, weil entweder die Zielgruppenzugehörigkeit nicht geklärt werden konnte oder aber die Zielperson ausgescreent werden musste. Der Anteil der darüber hinaus nicht erreichten Zielpersonen ist insgesamt sehr klein. Dies geht auf eine Bereinigung im Zuge der Erstellung des Final Outcomes zurück. Fälle, die mindestens 16 erfolglose Kontakte hintereinander und keinen realisierten Sprachkontakt aufwiesen, wurden als nicht geschaltete Nummern gewertet und daher von der Kategorie ‚Kein Kontakt‘ (NR-NC) auf die die Kategorie ‚Zielgruppenzugehörigkeit unbekannt‘ (UE) und den Rücklaufcode ‚Kein Anschluss‘ umgesetzt.³

Ein nicht ganz unbeträchtlicher Teil von rund 10 Prozent der Zielpersonen lag auf Wiedervorlage für erneuten Anruf, konnte aber während der Feldzeit nicht mehr realisiert werden. Diese Größenordnung ist aber insgesamt und im Vergleich zu anderen Studien nichts Ungewöhnliches. Der Anteil der Verweigerer beträgt über ein Drittel der Bruttostichprobe und die Hälfte der erreichten und befragbaren Zielpersonen. Am stärksten fallen dabei die Gründe „grundsätzliche Verweigerung“, „legt sofort auf“ und „kein Interesse am Thema“ ins Gewicht. Interviewabbrüche machen insgesamt nur einen kleinen Teil aus und bilden etwa 1,5 Prozent des gesamten Nonresponse ab. Für eine Befragung von Beschäftigten kann das Feldergebnis insgesamt durchaus positiv gewertet werden.

Ausschöpfungsraten geben an, welcher Adress- oder Stichprobeneinsatz nötig war, um eine bestimmte Anzahl an Interviews zu realisieren. Mit Blick auf diese Adressausschöpfung ist festzuhalten, dass bezogen auf das Gesamtbrutto auf knapp 25 Prozent ein Interview entfällt. Der Wert verbessert sich deutlich auf über 32 Prozent, wenn die Berechnung sich nur auf die Fälle der Stichprobe bezieht, die mit hoher Sicherheit zur Zielgruppe gehören. Diese Netto-Ausschöpfungsquote entspricht im Wesentlichen der *Response Rate 5* nach AAPOR.

Das Ergebnis fällt noch positiver aus, wenn sich die Berechnung im Nenner nur auf die Fälle bezieht, die tatsächlich erreicht wurden und auch befragbar waren. Personen, die überhaupt nicht (Fälle NR-NC) oder aber in der Feldzeit nicht zu erreichen waren, wurden hier ebenso wenig berücksichtigt wie die Personen, deren Deutschkenntnisse für ein Telefoninterview nicht ausreichten. Die Interviewausschöpfung erhöht sich auf über ein Drittel (34 Prozent) aller interviewfähigen Personen, mit denen ein Sprachkontakt realisiert wurde.

Im folgenden Unterabschnitt werden weitere Ergebnisquoten auf Grundlage des *Final Outcome* vorgestellt. Diese Ergebnisquoten orientieren sich alle an den AAPOR-Berechnungsmethoden.

³ Erfahrungsgemäß handelt es sich bei diesen Nummern um nicht funktionierende oder im Haushalt nicht geschaltete Nummern (z.B. von ISDN-Anschlüssen). Auswertungen über alle infas-Studien hinweg belegen, dass die Wahrscheinlichkeit, ein Interview zu realisieren, ab 16 Kontaktversuchen ohne irgendeinen Sprachkontakt gegen Null geht.

5.2 Outcome Rates

Neben der Ausschöpfungsquote informieren weitere Kennziffern über wichtige Dimensionen des Feldergebnisses. So messen Kooperationsraten grundsätzlich den Anteil der Interviews in Relation zu allen erfolgreich kontaktierten Einheiten der Zielgruppe. Kontaktraten messen dagegen den Anteil aller Fälle, mit denen überhaupt ein Sprachkontakt im Rahmen der Studie zustande kam. Üblicherweise rekurren Kontaktraten dabei auf Haushalte als Bezugseinheit, also üblicherweise im Rahmen von Haushaltbefragungen. Für die Zwecke der vorliegenden Studie werden Kontaktraten auf der Ebene von Zielpersonen vorgelegt. Verweigerungsrate (*refusal rates*) geben schließlich die Anteile der Personen an, die die Mitwirkung an einer Studie oder Befragung zurückweisen. Die einschlägige Literatur (vgl. u.a. AAPOR 2011) hält für diese Kennziffern sehr ausdifferenzierte und unterschiedliche Definition bereit. Wir haben uns hier aus Gründen der Anschaulichkeit mit Ausnahme der Kooperationsrate jeweils auf die bereinigte Bruttostichprobe I (vgl. Tabelle 13) als Nenner bezogen.

Tabelle 14 fasst die Werte dieser spezifisch definierten *Outcome Rates* zusammen.

Tabelle 14 Outcome Rates

	Berechnungsbasis	in %
Ausschöpfungsquote	Interviews/ bereinigtes Brutto	32,4
Kontaktrate	Σ Interviews, Verweigerungen, NR-Sonstiges / bereinigtes Brutto	99,1
Verweigerungsrate	Verweigerungen/ bereinigtes Brutto	48,3
Kooperationsrate	Interviews / Σ Interviews, Refusals, auskunftsfähige Kontakte	34,4

Basis: CATI-Befragung, Feldendstand (Final Outcome). Die Berechnungen erfolgten in Anlehnung an AAPOR-Definitionen, sind aber nicht vollständig identisch.

Die ausgewiesene Kooperationsrate setzt hingegen die realisierten Interviews in Bezug zur Summe aus gültigen Interviews, den Verweigerungen und sonstigen Kontakten. Dabei wurden hier die Personen herausgerechnet, die zum Kontaktierungszeitpunkt faktisch nicht auskunftsfähig waren, selbst bei möglicher Kooperationsbereitschaft. Dies betrifft die dauerhaft Erkrankten (NR-NA), die Personen ohne ausreichende Deutschkenntnisse für ein Telefoninterviews sowie diejenigen, die in der Feldzeit nicht zu erreichen waren.

Die Werte stellen der ersten Welle dieser Beschäftigtenbefragung insgesamt ein gutes Zeugnis aus. Ein Drittel aller Zielpersonen hat insgesamt kooperiert und ein auswertungsfähiges Interview gegeben. Der Anteil der Verweigerungen liegt unter 50 Prozent der Zielpersonen. Bei der Kontaktrate liegt der Wert nahe am Maximum.

5.3 Maßnahmen zur Verbesserung der Adressausschöpfung

5.3.1 Adressaktualisierung

Ende Januar 2013 wurde für n=495 Fälle eine Rufnummernaktualisierung dergestalt durchgeführt, dass eine in der Stichprobe vorhandene zweite Telefonnummer freigeschaltet wurde.

5.3.2 Konvertierung

Um eine optimale Ausschöpfung der Adressen zu gewährleisten, wurde zum Februar 2013 eine Konvertierungsstudie gestartet, um alle bisherigen sogenannten weichen Verweigerer im Rahmen einer neuerlichen Kontaktierung doch noch zur Mitwirkung an der Studie zu bewegen und ein Interview zu geben. In der Praxis wurden dazu Personen mit den Rücklaufcodes ‚kein Zugang zur Zielperson‘, ‚Kontaktperson verweigert Auskunft‘, ‚legt sofort auf‘, ‚krank‘, ‚kein Interesse‘, ‚keine Zeit‘ in eine separate Teilstudie umgesetzt und diese Adressen sodann gezielt mit überdurchschnittlich versierten Interviewerinnen und Interviewern kontaktiert. Auch ein Teil der Fälle mit dem Rücklaufcode ‚Zielperson in Feldzeit nicht zu erreichen‘ wurde nochmals überprüft.

Die Konvertierungsstichprobe startete mit 7.651 Adressen und wurde mehrmals aufgefüllt. Insgesamt flossen 8.205 Adressen in die Konvertierung, mit denen noch einmal 641 Interviews realisiert werden konnten. Das ist ein zufriedenstellendes Ergebnis und entspricht 8,5 Prozent aller gültigen Interviews und knapp acht Prozent der eingesetzten Konvertierungsstichprobe.

5.4 Paneleinwilligungen und Zuspielungsbereitschaft

Von den 7.508 gültigen Interviews willigten 7.054 Personen ein, dass ihre Adresse für eine erneute Kontaktierung zum Zwecke eines zweiten Interviews gespeichert wird. Das entspricht einer Panelbereitschaft von 94 Prozent. Mit einer Zuspielung von BA-Daten erklärten sich 6.195 Personen einverstanden, was einer Zustimmungquote von 82,5 Prozent entspricht.

Tabelle 15 Panelbereitschaft und Zuspielbereitschaft

		Zuspielbereitschaft		
		erteilt	nicht erteilt	Insgesamt
Panelbereitschaft	erteilt	6051 (80,6%)	1003 (13,3%)	7054 (94%)
	nicht erteilt	144 (1,9%)	310 (4,1 %)	454 (6,0%)
	Insgesamt	6195 (82,5%)	1313 (17,5)	7508 (100%)

Quelle: Personenbefragung, Welle 1

5.5 Interviewdauern

Die durchschnittliche Interviewdauer in der Haupterhebung bewegt sich mit 30 Minuten genau im Rahmen der für das Projekt geplanten und kalkulierten Zeiten (Tabelle 16). 55 Prozent aller Interviews bewegen sich in diesem Durchschnittsbereich, ein weiteres Drittel der Interviews überschreitet die Länge von 40 Minuten nicht (Tabelle 17).

Tabelle 16 Interviewdauern in Minuten

	Mean	Min	Max	Stddev
Interviewdauer	30,2	16,3	77,5	6,9

Quelle: Personenbefragung, Welle 1

Auf im Einzelfall schwierige Gesprächsverläufe verweist die Maximallänge von eineinviertel Stunden (Tabelle 16). Allerdings ist die Anzahl der Fälle mit sehr langer Interviewdauer über einer Stunde sehr überschaubar (Tabelle 17). Interviewdauern zwischen (über) 40 und 60 Minuten betreffen acht Prozent aller Interviews.

Tabelle 17 Interviewdauern in Minuten, gruppiert

	abs.	in %
1 bis 20 Minuten	142	1,9
Über 20 bis 30 Minuten	4097	54,6
Über 30 bis 40 Minuten	2630	35,0
Über 40 bis 50 Minuten	527	7,0
Über 50 bis 60 Minuten	91	1,2
Über 60 Minuten	21	0,3
Insgesamt	7508	100

Quelle: Personenbefragung, Welle 1

5.6 Interviewabbrüche

Insgesamt kam es mit 167 Fällen zu nur einer insgesamt kleinen Anzahl an Abbrüchen durch die Befragungsperson. Dieser niedrige Wert belegt die in der Gesamtschau hohe Akzeptanz der Befragung durch die Zielpersonen. Dabei entfällt ein Zehntel der Abbrüche auf frühzeitige Abbrüche schon im Intro (d.h. vor dem eigentlichen Interviewstart) sowie im Rahmen des Screenings, bei der Kontrolle der Geburtsangaben und der Betriebszugehörigkeit. Da weitere vier Zehntel der

Abbrüche auf die Blöcke 100 und 200 entfallen, sind demnach mehr als die Hälfte aller Abbrüche im vorderen Bereich des Fragebogens zu verbuchen.

Tabelle 18 Interviewabbrüche nach Fragenbereich im Interview, Häufigkeiten

	abs.	in %
Abbruch im Intro	11	6,6
Fragenblock Screening	5	3,0
Fragenblock 100	19	11,4
Fragenblock 200	50	29,9
Fragenblock 300	26	15,6
Fragenblock 400	2	1,2
Fragenblock 500	43	25,7
Fragenblock 600-800	11	6,6
Insgesamt	167	100

Quelle: Personenbefragung, Welle 1

Interessanterweise treten allein 43 der 50 Abbrüche im Block 200 bei der Frage 208 auf, d.h. der Wahrnehmung und Bewertung der betrieblichen Praxis in Bezug auf Personalentwicklung und Beförderungen. Dies ist die erste Frage mit komplexeren und inhaltlich durchaus auch kritischen Einschätzungsfragen. Es darf plausibel vermutet werden, dass ein Teil der Abbrecher/innen an dieser Stelle die Frageitems als entweder zu heikel oder aber als zu schwierig empfanden und daher zur ‚Notbremse‘ des Abbruchs griffen. Auch bei den komplexen Einschätzungsfragen zu Arbeitsbedingungen und –belastung (F301) und zur Work-Life-Balance (F303) treten gehäuft Abbrüche auf, aber auch im Fragenkomplex zu Bindung, Werten und Unternehmenskultur (Block 500). Dabei sind es wieder die komplexeren und ggf. heiklen Fragen zur Unternehmenskultur (F501) und Kollegialität/Mobbing (F502), bei denen ein Abbruch erfolgte.

6 Datenprüfung und –lieferung

6.1 Definition gültig realisierter Interviews

Insgesamt wurden 7.559 Interviews realisiert. Im Zuge der Datenprüfung wurden 41 Fälle als ungültig deklariert, weil das Alter in Jahren zwischen Stichprobenangaben und Befragungsangaben um jeweils mehr als ein Jahr voneinander abwich. Offenbar waren hier irrtümlich namensgleiche Personen einer älteren bzw. jüngeren Generation befragt worden. Bei einem dieser Fälle wich auch das Geschlecht zwischen Stichprobe und Befragung ab. Zehn weitere Interviews

waren begonnen, dann jedoch unterbrochen worden, ohne dass es während der Feldzeit zu einer Wiederaufnahme und Beendigung des Interviews kam. Diese Teilinterviews wurden nach Prüfung als nicht auswertbar, also als ungültig deklariert. Damit verbleiben 7.508 gültig realisierte Fälle im Auswertungsdatensatz. Unter diesen 7508 Fällen sind zehn Fälle, bei denen das Alter zwischen Stichproben- und Befragungsangaben um maximal ein Jahr voneinander abweicht. Diese kleinen, nicht ungewöhnlichen Abweichungen sind wahrscheinlich auf fehlerhafte Eingaben im Survey oder in der Ursprungsdatei zurückzuführen und tolerabel.

6.2 Prüfroutinen im Fragenprogramm

Inhaltliche Plausibilitätsprüfungen wurden – wie bereits in Kapitel 3 erläutert – nach Absprache mit dem Auftraggeber bereits im CATI-Instrument berücksichtigt. Diese umfassten hier insbesondere zahlreiche Rangeprüfungen, da in dem Instrument diverse offene numerische Abfragen enthalten waren.

6.3 Datenprüfung und -aufbereitung nach Feldende

Die ersten Schritte der Datenprüfung begannen bereits vor der eigentlichen Feldphase. Nachdem der abgestimmte Fragebogen programmiert war, wurde die korrekte Ablage der Daten geprüft. Dabei wurde der gesamte Filterverlauf im programmierten Instrument nachvollzogen. Für alle Fragen wurde zudem auf die Einhaltung des vorgesehenen Wertebereichs geachtet, damit bei geschlossenen Skalen lediglich die in der Fragebogenvorlage vorgesehenen Codes auftreten können. Bei offenen Zahlenangaben durften die erhobenen Werte nur in der vorgegebenen Range auftreten.

Nach der Erhebung werden die Daten weiteren Prüfungen unterzogen. Mit Plausibilitätstests wurden die realisierten Eingaben bei den offenen Werten überprüft. Infas identifizierte einige wenige mögliche Inplausibilitäten und wies das IAB bei der Datenlieferung auf diese Punkte als Entscheidungshilfe für etwaige Datenbereinigungsschritte hin. Diese Punkte betrafen:

- Zehn Fälle, bei denen die Zahl der Tage, an denen an Kursen teilgenommen wurde (Frage F203) und die Zahl der Krankheitstage in den letzten 12 Monaten (Frage F702) zusammen bei über 200 Tagen lagen.
- In 19 Fällen wird die Aufnahme einer betrieblichen Altersversorgung (F 407) mit einem Alter unter 15 bzw. erst ab dem 60. Lebensjahr begonnen. Zumindest bei den Personen unter 15 Jahren erscheint die Angabe unplausibel.
- In sechs Fällen wurden Frageitems zu Kollegialität beantwortet, im Folgeitem jedoch geantwortet, dass gar keine Kollegen/Vorgesetzten vorhanden sind.
- In Frage 808 sind für n=108 Fälle auch Kinder ab 14 Jahren aufgenommen worden (statt nur Kinder unter 14 Jahren). Außerdem wurde in drei Fällen die Sortierreihenfolge nicht eingehalten, also nicht mit der Auflistung des ältesten Kindes begonnen.
- In 10 Fällen fielen mutmaßliche Tippfehler und Inkonsistenzen bei den numerischen Angaben bei Sonderzahlungen (Fragen 401c und 401d) sowie bei Brutto- und Nettoverdienst auf.

6.4 Datenlieferung

Die Datenlieferung der noch ungewichteten Befragungsdaten mit 7.508 gültigen Fällen erfolgte Anfang April 2013, d.h. sehr zeitnah nach Feldende.

Der Methodendatensatz wurde am 23. April 2013 ausgeliefert und enthält nach Abstimmung mit dem IAB die Merkmale

- LFD/ID der gesamten Bruttostichprobe
- IAB-ID
- Letzter Rücklaufcode für alle Bruttofälle
- Anzahl Kontaktversuche
- die Stichprobenmerkmale, wie vom IAB geliefert
- Tranchenkennung
- Kennzeichen für auswertbare Querschnittsfälle
- Kennzeichen für Zugehörigkeit zum Längsschnitt
- Interviewdatum.

Den finalen, gewichteten Datensatz erhielt das IAB im Mai 2013. Dabei wurde auch ein Tabellenband mitgeliefert, der die Befragungsergebnisse in Form gewichteter Häufigkeitsauszählungen in übersichtlicher, handlicher Form zusammenfasst.

7 Selektivität und Gewichtung

Basis für die Berechnung der Gewichtungsfaktoren waren vom IAB gelieferte Auswahlwahrscheinlichkeiten für die Personen der Bruttostichprobe. Diese wurden berechnet als Produkt aus der Auswahlwahrscheinlichkeit der Betriebe und der Auswahlwahrscheinlichkeit der Beschäftigten in den ausgewählten Betrieben. Die Designgewichte entsprechen der reziproken Auswahlwahrscheinlichkeiten für die Bruttostichprobe.

Diese Designgewichte wurden anschließend über eine zweistufige Ausfallmodellierung mittels Logit-Modellen adjustiert. Mit Hilfe dieser Modelle erfolgte die Berechnung von Wahrscheinlichkeiten, die zur Adjustierung der Designgewichte verwendet wurden. Im Ergebnis erteilen die Modelle Auskunft über Selektivitäten der realisierten Stichprobe. Ein solcher modellbasierter Ansatz ist hier möglich, da weitreichende Informationen über die Bruttostichprobe vorliegen.

Betrachtet werden zwei potentielle Selektionsstufen. Die erste liegt darin begründet, dass nicht für jede Person der Stichprobe eine Telefonnummer vorliegt bzw. zu ermitteln ist. Die zweite Selektion ergibt sich im Feldeinsatz. Personen, die trotz entsprechend langer Feldzeit nicht zu erreichen sind und Personen, die eine Teilnahme an der Befragung ablehnen, fallen im Erhebungsprozess aus. Diesen Selektionsstufen entsprechend erklärt das erste Logit-Modell die Wahrscheinlichkeit eines Einsatzes im Feld. Das zweite Logit-Modell erklärt hingegen die Teilnahme an der Befragung im Falle eines Einsatzes im Feld.

7.1 Wahrscheinlichkeit für die Einsatzstichprobe

Im ersten Schritt wurde für jede Person der Bruttostichprobe die Wahrscheinlichkeit berechnet, in die Einsatzstichprobe zu gelangen. Ein Ausfall von der Bruttostichprobe zur Einsatzstichprobe ergibt sich im Falle einer fehlenden Telefonnummer, da nur Personen mit einer Telefonnummer befragt werden konnten. Berechnet wurde ein Logit-Modell mit den für alle Personen der Bruttostichprobe zur Verfügung stehenden unabhängigen Variablen:

- Arbeitsort Bundesland
- Geschlecht
- Altersgruppe
- Vollzeit-/Teilzeitstelle
- Wirtschaftszweig (WZ08-Abschnitt)
- Tagesentgelt (Quantile)

Tabelle 19 Ausfallmodellierung Einsatzstichprobe (logistische Regression, odds ratios)⁴

Abhängige Variable: Y = 1: Einsatzstichprobe, Y = 0: nicht in Einsatzstichprobe	odds ratio	p-Wert
Arbeitsort Bundesland		
Schleswig-Holstein	Referenz	
Hamburg	0,58	0,00
Niedersachsen	1,11	0,22
Bremen	0,84	0,04
Nordrhein-Westfalen	0,67	0,00
Hessen	0,73	0,00
Rheinland-Pfalz	0,67	0,00
Baden-Württemberg	0,79	0,00
Bayern	0,76	0,00
Saarland	0,85	0,13
Berlin	0,42	0,00
Brandenburg	0,77	0,03
Mecklenburg-Vorpommern	0,88	0,25
Sachsen	0,67	0,00
Sachsen-Anhalt	0,86	0,09
Thüringen	0,78	0,01

⁴ Die exponierten Regressionskoeffizienten einer logistischen Regression lassen sich als sog. „odds ratios“ interpretieren. Bei kategorialen unabhängigen Variablen kann ein odds ratio direkt als Wahrscheinlichkeitsverhältnis der einzelnen Ausprägungen dieser Variablen gegenüber der Referenzkategorie interpretiert werden. So bedeutet beispielsweise ein odds ratio von 2,0 auf der Variable Geschlecht bei den Frauen mit der Referenzkategorie „Männer“, dass die Wahrscheinlichkeit bei den Frauen doppelt so groß ist wie bei den Männern. Ein odds ratio kleiner eins kann als negativer Zusammenhang interpretiert werden. Ein Wert von 0,5 bei den Frauen bedeutet, dass die Wahrscheinlichkeit bei den Frauen geringer ist, nämlich nur halb so groß wie bei den Männern. Umgekehrt ausgedrückt wäre die Wahrscheinlichkeit der Männer bei diesem Beispiel also doppelt so groß wie die der Frauen, was sich auch rechnerisch ermitteln lässt, indem eins durch das relative Risiko dividiert wird: $1/0,5 = 2,0$.

<i>Fortsetzung Tabelle 19</i>	odds ratio	p-Wert
Geschlecht		
Männer	Referenz	
Frauen	0,70	0,00
Altersgruppe		
15-24 Jahre	Referenz	
25-34 Jahre	0,89	0,06
35-44 Jahre	0,69	0,00
45-54 Jahre	0,70	0,00
55-59 Jahre	0,75	0,00
60-81 Jahre	0,72	0,00
Vollzeit/Teilzeit		
Vollzeit	Referenz	
Teilzeit	0,75	0,00
fehlende Angabe	1,05	0,09
Wirtschaftszweig (WZ08 Abschnitt)		
B: Bergbau und Gewinnung von Steinen und Erden	Referenz	
C: Verarbeitendes Gewerbe	0,98	0,92
D: Energieversorgung	1,19	0,42
E: Wasserversorgung; Abwasser-, Abfallentsorgung	0,73	0,16
F: Baugewerbe	1,47	0,06
G: Handel; Instandhaltung und Reparatur von Kraftfahrzeugen	0,92	0,67
H: Verkehr und Lagerei	0,85	0,43
I: Gastgewerbe	0,76	0,30
J: Information und Kommunikation	0,76	0,18
K: Erbringung von Finanz- und Versicherungsdienstleistungen	1,19	0,39
L: Grundstücks- und Wohnungswesen	0,83	0,49
M: Erbringung von freiberufl., wissensch. und techn. Dienstleist.	0,89	0,57
N: Erbringung von sonstigen wirtschaftlichen Dienstleistungen	1,77	0,01
P: Erziehung und Unterricht	3,61	0,10
Q: Gesundheits- und Sozialwesen	0,88	0,53
R: Kunst, Unterhaltung und Erholung	0,45	0,01
S: Erbringung von sonstigen Dienstleistungen	1,24	0,32
Tagesentgelt gruppiert (Quantile)		
1,55 - 63,09 Euro	Referenz	
63,10 - 88,28 Euro	0,66	0,00
88,29 - 110,46 Euro	0,46	0,00
110,47 - 144,38 Euro	0,33	0,00
144,39 - 717,29 Euro	0,28	0,00
Konstante	10,12	0,00
n	43580	
Pseudo-R ²	0,05	

Insgesamt sind die Unterschiede zwischen der Bruttostichprobe und der Einsatzstichprobe (mit Telefonnummer) gering. Dies zeigt sich zum einen an den eher geringen odds ratios und zum anderen an der mit einem Pseudo-R² von rd. 0,05 geringen Erklärungskraft des Modells. Dennoch ist der Ausfall an dieser Stelle nicht völlig zufällig, sondern es zeigen sich einige systematische Zusammenhänge.

Im Vergleich zur Referenzkategorie Schleswig-Holstein ist in allen Bundesländern mit Ausnahme von Niedersachsen der Anteil der Personen mit einer (recherchierten) Telefonnummer, die in der Einsatzstichprobe verwendet werden konnten, geringer. Insbesondere in den beiden Stadtstaaten Hamburg und Berlin ist der Ausfall auf Grund fehlender, nicht recherchierbarer Telefonnummern erwartungsgemäß höher.

Frauen haben im Vergleich zu Männern eine etwas geringere Wahrscheinlichkeit, in die Einsatzstichprobe zu gelangen. Dies gilt in etwa gleicher Größenordnung auch für die Teilzeitbeschäftigten im Vergleich zu den Vollzeitbeschäftigten.

Beim Alter zeigt sich, dass die Altersgruppen ab 25 Jahren aufwärts alle etwa die gleiche Wahrscheinlichkeit für das Vorliegen einer Telefonnummer haben. Im Vergleich zur Referenzkategorie der 15-24-Jährigen ist diese bei den älteren Altersgruppen durchgängig geringer. Dies bedeutet, dass für die 15-24-Jährigen häufiger als für die anderen Altersgruppen eine Telefonnummer vorliegt.

Die Wahrscheinlichkeit in die Einsatzstichprobe zu gelangen variiert zudem moderat über die Wirtschaftszweige. Zu berücksichtigen sind hier die zum Teil sehr geringen Fallzahlen, so dass auch hohe odds ratios statistisch nicht signifikant sind.

Beim Tagesentgelt zeigt sich ein im Vergleich zu den anderen Merkmalen relativ starker linearer Effekt. Je höher das Tagesentgelt, desto geringer ist die Wahrscheinlichkeit des Vorliegens einer Telefonnummer.

7.2 Wahrscheinlichkeit für die realisierte Stichprobe

Personen, für die eine Telefonnummer in der Auswahlgesamtheit vorlag oder für die eine Telefonnummer recherchiert werden konnte, bildeten die Einsatzstichprobe. Für diese Personen der Einsatzstichprobe wurde im zweiten Schritt die Wahrscheinlichkeit eines realisierten Interviews berechnet. Der Ausfall von der Einsatzstichprobe zur realisierten Stichprobe ist hier bedingt durch Nichterreichbarkeit oder eine fehlende Teilnahmebereitschaft (Verweigerung) aus unterschiedlichen Gründen. Berechnet wurde auch im zweiten Schritt ein Logit-Modell mit den für alle Personen der Einsatzstichprobe zur Verfügung stehenden unabhängigen Variablen.

Tabelle 20 Ausfallmodellierung realisierte Stichprobe für die Einsatzstichprobe (logistische Regression, odds ratios)

Abhängige Variable: Y = 1: realisierte Stichprobe, Y = 0: in Einsatzstichprobe, aber nicht realisiert	odds ratio	p-Wert
Arbeitsort Bundesland		
Schleswig-Holstein	Referenz	
Hamburg	1,06	0,69
Niedersachsen	1,25	0,02
Bremen	1,01	0,91
Nordrhein-Westfalen	1,12	0,19
Hessen	1,09	0,36
Rheinland-Pfalz	1,18	0,09
Baden-Württemberg	1,11	0,27
Bayern	1,30	0,00
Saarland	0,83	0,19
Berlin	1,04	0,72
Brandenburg	1,25	0,08
Mecklenburg-Vorpommern	1,50	0,00
Sachsen	1,34	0,00
Sachsen-Anhalt	1,39	0,00
Thüringen	1,58	0,00
Geschlecht		
Männer	Referenz	
Frauen	1,04	0,24
Altersgruppe		
15-24 Jahre	Referenz	
25-34 Jahre	0,95	0,50
35-44 Jahre	1,10	0,21
45-54 Jahre	1,43	0,00
55-59 Jahre	1,46	0,00
60-81 Jahre	1,24	0,02
Vollzeit/Teilzeit		
Vollzeit	Referenz	
Teilzeit	1,34	0,00
fehlende Angabe	0,88	0,00
Wirtschaftszweig (WZ08 Abschnitt)		
B: Bergbau und Gewinnung von Steinen und Erden	Referenz	
C: Verarbeitendes Gewerbe	1,13	0,62
D: Energieversorgung	1,35	0,27
E: Wasserversorgung; Abwasser-, Abfallentsorgung	0,98	0,93
F: Baugewerbe	1,12	0,66
G: Handel; Instandhaltung und Reparatur von Kraftfahrzeugen	1,14	0,58
H: Verkehr und Lagerei	1,04	0,89

<i>Fortsetzung Tabelle 20</i>	odds ratio	p-Wert
I: Gastgewerbe	0,41	0,02
J: Information und Kommunikation	1,24	0,41
K: Erbringung von Finanz- und Versicherungsdienstleistungen	1,46	0,13
L: Grundstücks- und Wohnungswesen	0,74	0,41
M: Erbringung von freiberufl., wissensch. und techn. Dienstleist.	1,20	0,49
N: Erbringung von sonstigen wirtschaftlichen Dienstleistungen	0,91	0,71
P: Erziehung und Unterricht	1,15	0,83
Q: Gesundheits- und Sozialwesen	1,18	0,51
R: Kunst, Unterhaltung und Erholung	0,56	0,32
S: Erbringung von sonstigen Dienstleistungen	0,99	0,98
Tagesentgelt gruppiert (Quantile)		
1,55 - 63,09 Euro	Referenz	
63,10 - 88,28 Euro	1,07	0,15
88,29 - 110,46 Euro	1,22	0,00
110,47 - 144,38 Euro	1,49	0,00
144,39 - 717,29 Euro	1,96	0,00
Konstante	0,15	0,00
n	30703	
Pseudo-R ²	0,02	

Die Unterschiede zwischen der realisierten Stichprobe und der nicht realisierten Einsatzstichprobe sind offensichtlich sehr gering. Die Erklärungskraft des Modells ist mit einem Pseudo-R² von 0,02 ausgesprochen niedrig. Es zeigen sich nur wenige, schwach ausgeprägte systematische Effekte. Diese sind nahezu durchgängig gegenläufig gegenüber dem ersten Ausfallmodell. Die Wahrscheinlichkeit eines realisierten Interviews ist also in den Gruppen höher, bei denen die Wahrscheinlichkeit, in die Einsatzstichprobe zu gelangen niedriger ist und umgekehrt.

Hinsichtlich des Regionalmerkmals Bundesland sind die Unterschiede in der Realisierungswahrscheinlichkeit gering; es ergeben sich niedrige odds ratios. Interessanterweise sind die odds ratios bei diesem Ausfallmodell in allen Bundesländern gegenüber der Referenzkategorie Schleswig-Holstein höher.

Hinsichtlich der Geschlechtszugehörigkeit ergeben sich keine Teilnahmeeffekte und auch für die Wirtschaftszweigzugehörigkeit zeigen sich keine nennenswerten Unterschiede. Einzig bei Beschäftigten des Gastgewerbes ist eine vergleichsweise geringere Realisierungswahrscheinlichkeit feststellbar.

In Bezug auf das Beschäftigtenmerkmal Alter ergibt sich für alle älteren Altersgruppen eine höhere Wahrscheinlichkeit einer Interviewrealisierung gegenüber der Referenzkategorie. Auch bei den Teilzeitbeschäftigten ist die Wahrscheinlichkeit eines realisierten Interviews im Gegensatz zum ersten Ausfallmodell höher.

Das Merkmal Höhe des Tagesentgelts von Beschäftigten erbringt auch im zweiten Ausfallmodell lineare und zugleich die stärksten Effekte. Diese Effekte sind allerdings exakt gegenläufig zum ersten Modell. Mit steigendem Tagesentgelt steigt auch die Wahrscheinlichkeit eines realisierten Interviews.

7.3 Adjustierte Gewichtungsfaktoren

Im Anschluss an die Berechnung der Modelle wurde das Produkt der vorhergesagten Wahrscheinlichkeiten beider Modelle berechnet und die Designgewichte durch diese geschätzte Teilnahmewahrscheinlichkeit dividiert. Dadurch ergeben sich die modifizierten bzw. adjustierten Designgewichte, die dem Datensatz zugespielt wurden. Da es sich hierbei um Gewichtungsfaktoren und nicht um Hochrechnungsfaktoren handelt, wurden die Gewichte fallzahlnormiert, d. h. die Summe der Gewichte entspricht der Fallzahl, der Mittelwert der Gewichtungsfaktoren ist entsprechend 1.

Zur Beurteilung der Gewichte wird im Folgenden neben dem Mittelwert und der Standardabweichung auch das Effektivitätsmaß angegeben (E). Das Effektivitätsmaß E basiert auf der Varianz des Gewichtungsfaktors. Das Effektivitätsmaß gibt in Prozent der realisierten Fallzahl an, wie groß bei Verwendung des Gewichts die effektive Fallzahl (n') bei einem passiven Merkmal ist, das mit den aktiven Merkmalen nicht korreliert. Die effektive Fallzahl entspricht der Anzahl Befragter, die bei einer uneingeschränkten Zufallsauswahl gegeben die Varianz des Merkmals in der Stichprobe den gleichen Stichprobenfehler produziert hätte. Das Effektivitätsmaß drückt das Verhältnis von n zu n' als Prozentwert aus und wird berechnet über:

$$E = (n' / n) * 100, \text{ wobei: } n' = ((\sum_i g_i)^2 / \sum_i g_i^2)$$

Tabelle 21 Kennziffern für die Gewichtungsfaktoren

Mittelwert	1,00
Standardabweichung	1,03
Minimum	0,12
Maximum	12,12
Effektivitätsmaß	48,57%
effektive Fallzahl	3.647

Das Effektivitätsmaß ist mit rund 50 Prozent vergleichsweise gering. Dies beruht allerdings ausschließlich auf dem Designgewicht, dessen Effektivitätsmaß auch bereits bei rd. 48 Prozent liegt. Die Adjustierung anhand der Ausfallmodellierung verändert die Varianz der Gewichte faktisch nicht. Verändert wird lediglich der Range der Gewichtungsfaktoren (bei den Designgewichten von 0,17 – 8,6). Diese nur unwesentliche Veränderung der Designgewichte durch die Ausfallmodellierung erklärt sich durch die, über beide Ausfallmodelle betrachtet, sehr geringen Selektivitätseffekte.

Zudem ist auch festzuhalten, dass aufgrund der hohen Fallzahl der realisierten Stichprobe die effektive Fallzahl immer noch groß, der korrigierte Stichprobenfehler trotz des Designeffekts entsprechend gering ist.

7.4 Gesamtselektivität

Die Ausfallmodellierung erfolgte für die Berechnung der Wahrscheinlichkeiten zur Adjustierung der Gewichte zweistufig. Da sich bei den beiden Modellen gegenläufige Ergebnisse zeigen, erfolgte zusätzlich eine Modellierung der Gesamtselektivität. Dazu wurde ein Logit-Modell berechnet, in dem der Ausfall von der Bruttostichprobe zur realisierten Stichprobe, d. h. für alle Fälle der Bruttostichprobe die Wahrscheinlichkeit eines realisierten Interviews berechnet wurde. Verwendet wurden erneut alle für die Bruttostichprobe zur Verfügung stehenden Merkmale.

Tabelle 22 Ausfallmodellierung realisierte Stichprobe für die Bruttostichprobe (logistische Regression, odds ratios)

Abhängige Variable: Y = 1: realisierte Stichprobe, Y = 0: in Bruttostichprobe, aber nicht realisiert	odds ratio	p-Wert
Arbeitsort Bundesland		
Schleswig-Holstein	Referenz	
Hamburg	0,85	0,20
Niedersachsen	1,25	0,02
Bremen	0,95	0,58
Nordrhein-Westfalen	0,95	0,58
Hessen	0,97	0,74
Rheinland-Pfalz	0,97	0,76
Baden-Württemberg	1,03	0,72
Bayern	1,12	0,18
Saarland	0,78	0,07
Berlin	0,73	0,01
Brandenburg	1,09	0,48
Mecklenburg-Vorpommern	1,33	0,01
Sachsen	1,11	0,25
Sachsen-Anhalt	1,24	0,03
Thüringen	1,36	0,00

<i>Fortsetzung Tabelle 22</i>	odds ratio	p-Wert
Geschlecht		
Männer	Referenz	
Frauen	0,92	0,01
Altersgruppe		
15-24 Jahre	Referenz	
25-34 Jahre	0,94	0,36
35-44 Jahre	0,99	0,91
45-54 Jahre	1,28	0,00
55-59 Jahre	1,35	0,00
60-81 Jahre	1,16	0,08
Vollzeit/Teilzeit		
Vollzeit	Referenz	
Teilzeit	1,21	0,00
fehlende Angabe	0,91	0,01
Wirtschaftszweig (WZ08 Abschnitt)		
B: Bergbau und Gewinnung von Steinen und Erden	Referenz	
C: Verarbeitendes Gewerbe	1,10	0,67
D: Energieversorgung	1,44	0,16
E: Wasserversorgung; Abwasser-, Abfallentsorgung	0,85	0,56
F: Baugewerbe	1,25	0,35
G: Handel; Instandhaltung und Reparatur von Kraftfahrzeugen	1,10	0,68
H: Verkehr und Lagerei	0,96	0,88
I: Gastgewerbe	0,39	0,01
J: Information und Kommunikation	1,03	0,89
K: Erbringung von Finanz- und Versicherungsdienstleistungen	1,50	0,09
L: Grundstücks- und Wohnungswesen	0,70	0,31
M: Erbringung von freiberufl., wissensch. und techn. Dienstleist.	1,11	0,67
N: Erbringung von sonstigen wirtschaftlichen Dienstleistungen	1,01	0,96
P: Erziehung und Unterricht	1,57	0,47
Q: Gesundheits- und Sozialwesen	1,10	0,69
R: Kunst, Unterhaltung und Erholung	0,41	0,12
S: Erbringung von sonstigen Dienstleistungen	1,09	0,73
Tagesentgelt gruppiert (Quantile)		
1,55 - 63,09 Euro	Referenz	
63,10 - 88,28 Euro	0,94	0,21
88,29 - 110,46 Euro	0,96	0,44
110,47 - 144,38 Euro	1,01	0,90
144,39 - 717,29 Euro	1,18	0,00
Konstante	0,16	0,00
n	43580	
Pseudo-R ²	0,01	

Die Effekte der beiden für die Berechnung der Gewichtungsfaktoren geschätzten Ausfallmodelle heben sich in der Tat nahezu vollständig auf. Die Erklärungskraft des Modells ist mit einem Pseudo- R^2 von 0,01 außerordentlich gering. Zudem zeigen sich bei keinem Merkmal odds ratios in einer nennenswerten Größenordnung. Im Vergleich zur Bruttostichprobe liefert die realisierte Stichprobe demnach nahezu identische Verteilungen bei den betrachteten Merkmalen.

Impressum

FDZ-Methodenreport 2/2015

Herausgeber

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

Redaktion

Stefan Bender, Dagmar Theune

Technische Herstellung

Dagmar Theune

Rechte

Nachdruck - auch auszugsweise - nur mit
Genehmigung des FDZ gestattet

Bezugsmöglichkeit

http://doku.iab.de/fdz/reporte/2015/MR_02-15.pdf

Internet

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

Rückfragen zum Inhalt an:

Dr. Helmut Schröder
Bereichsleitung Sozialforschung
infas Institut für angewandte
Sozialwissenschaft GmbH
Friedrich-Wilhelm-Straße 18
53113 Bonn

Tel. +49 (0)228/38 22-406
Fax +49 (0)228/310071
E-Mail h.schroeder@infas.de