
**Beeinflussen Interviewereffekte die
Einkommensangabe in Befragungen?
Eine Analyse mithilfe administrativer Daten**

Adrian Reichert



Masterarbeit

Institut für Statistik

Ludwig-Maximilians-Universität München

Betreuung: Prof. Dr. Frauke Kreuter

Barbara Felderer

05. Februar 2015

Zusammenfassung

Daten, die aus Befragungen gewonnen werden, können aus den verschiedensten Gründen Fehler und Verzerrungen aufweisen. Dies gilt ganz besonders bei der Frage nach dem Einkommen der Befragten. Eine Möglichkeit der Quantifizierung von Messfehlern im Einkommen liegt im Vergleich der Befragungsangaben mit als zuverlässig erachteten Aufzeichnungen aus administrativen Daten. Derartige Untersuchungen erfreuen sich immer größerer Popularität. Nach gegenwärtigem Forschungsstand wurden die Messfehler im Einkommen in der Regel auf soziodemographische Merkmale der Befragten zurückgeführt. Ob und in welchem Umfang die Interviewer einen Einfluss auf die Messfehler im Einkommen ausüben, wurde bisher jedoch ausgeklammert.

Die Verknüpfung von Befragungsdaten aus dem Panel Arbeitsmarkt und soziale Sicherung des Instituts für Arbeitsmarkt- und Berufsforschung mit administrativen Daten der Bundesagentur für Arbeit bietet eine geeignete Datenbasis um Zusammenhänge zwischen Eigenschaften der Interviewer und der Einkommensangabe der Befragten systematisch zu untersuchen. Mithilfe einer Mehrebenenanalyse wird gezeigt, dass Interviewereffekte die Einkommensangabe in Befragungen signifikant beeinflussen. Diese Effekte fallen jedoch insbesondere im Vergleich zum Einfluss der Interviewer auf andere Fehler- und Verzerrungsquellen im Befragungsprozess, wie beispielsweise die totale oder partielle Verweigerung der Teilnahme oder die Verweigerung der Erlaubnis zur Verknüpfung der Befragungsdaten mit administrativen Daten, weniger stark ins Gewicht.

Inhaltsverzeichnis

1	Einführung	1
2	Literaturüberblick	4
3	Daten	8
3.1	Integrierte Erwerbsbiographien	8
3.2	Panel Arbeitsmarkt und soziale Sicherung	11
3.3	Datenverknüpfung und Fallsektion	14
3.4	Datenqualität und Selektivitäten	19
3.4.1	Integrierte Erwerbsbiographien	19
3.4.2	Panel Arbeitsmarkt und soziale Sicherung	20
4	Methoden	28
5	Ergebnisse	35
5.1	Deskriptive Analyse der Daten	35
5.2	Regressionsanalysen	42
6	Zusammenfassung	48
	Literaturverzeichnis	50
	Anhang	54

1 Einführung

Das Einkommen als maßgeblicher Einflussfaktor vieler persönlicher, wirtschaftlicher und gesellschaftlicher Handlungen ist seit jeher ein wichtiges Forschungsgebiet in den Wirtschafts- und Sozialwissenschaften. In den meisten Fällen müssen für die Gewinnung von Daten über das Einkommen einer Zielpopulation repräsentative Befragungen durchgeführt werden. Je nach Größe der Stichprobe und Umfang des Bedarfs nach weiteren Informationen über die Zielpopulation können dabei ein hoher Aufwand und hohe Kosten entstehen. Der aktuelle Forschungsstand zeigt zudem, dass Befragungsdaten eine Reihe von Fehler- und Verzerrungsquellen aufweisen können. Darunter fallen unter anderem Messfehler, die z.B. durch fehlerhafte Erinnerung der Befragten, durch den Einfluss des Interviewers oder die Gestaltung der Fragen auftreten können. Auch der überproportionale Ausfall bestimmter Untergruppen der Zielpopulation bei der Teilnahme an der Befragung oder der Beantwortung bestimmter Fragen kann zu Verzerrungen führen. Bei der Interpretation von Studien, die auf Befragungsdaten basieren, ist es daher entscheidend, die Qualität der Befragungsdaten und deren Einschränkungen zu kennen.

Im Fall von Messfehlern, ergibt sich eine Abweichung zwischen dem von der befragten Person angegebenen Wert und dem tatsächlichen Wert. Da der tatsächliche Wert in aller Regel weder erhoben noch beobachtet werden kann, muss zur Bestimmung des Messfehlers eine Referenzdatenquelle gefunden werden, die dem tatsächlichen Wert möglichst nahe kommt. Die Daten, die die tatsächlichen Eigenschaften von Personen am besten abbilden, sind dabei in der Regel die Aufzeichnungen von Behörden und weiteren administrativen Einrichtungen, die zur gesetzmäßigen Erbringung ihrer Leistungen in unterschiedlichem Umfang Daten erheben müssen. Diese Daten werden aufgrund ihrer hohen Relevanz sorgfältiger geprüft und

meist auch ohne Zutun und Einfluss der betreffenden Person erhoben. Zu derartigen Einrichtungen gehören in Deutschland beispielsweise die Krankenkassen, die Rentenversicherung und die Bundesagentur für Arbeit. Aufgrund der Sensibilität und des Schutzbedarfs der hier gespeicherten persönlichen Daten ist der Zugang zu den administrativen Daten dieser Einrichtungen allerdings stark eingeschränkt. Das „Panel Arbeitsmarkt und soziale Sicherung“ des Instituts für Arbeitsmarkt- und Berufsforschung ermöglicht einen solchen Vergleich mit administrativen Daten. Bei Zustimmung der Befragten können die Befragungsdaten mit den Erwerbsbiographien der Befragten, die bei der Bundesagentur für Arbeit vorliegen, verknüpft werden. So entsteht ein Datensatz, der mit der Wahl der administrativen Daten als Referenz eine Analyse der Messfehler im Einkommen ermöglicht.

Es existieren einige Studien, in denen die Messfehler im Einkommen anhand der Verknüpfung von Befragungsdaten und administrativen Daten untersucht werden. Diese beschränken sich jedoch meist auf die Betrachtung, ob die Messfehler im Einkommen von der Höhe des Einkommens selbst abhängen, und die Hinzunahme soziodemographischer Merkmale der Befragten. Eine weitere mögliche Einflussquelle für Messfehler im Einkommen, nämlich das Verhalten und die Eigenschaften des Interviewers bei der Befragung, wurde von derartigen Studien hingegen bisher ausgeklammert. In anderen Bereichen der Survey-Forschung, beispielsweise bei der Untersuchung von Antwortverweigerungen oder der Verweigerung einer Verknüpfung der Befragungsdaten mit administrativen Daten, gibt es jedoch eine große Zahl von Studien zu Interviewereffekten.

Ziel dieser Untersuchung ist es, den Einfluss der Interviewer auf den Messfehler bei den Einkommensangaben der Befragten zu modellieren. Dazu werden die Messfehler zunächst durch ein Modell, das nur Merkmale der Befragten selbst enthält, erklärt. Anschließend wird dieses Modell um Merkmale der Interviewer und Interaktionseffekte zwischen Befragten und Interviewern erweitert und auf eine signifikant größere Erklärungskraft hinsichtlich der Messfehler im Vergleich zum Ausgangsmodell getestet. Durch eine Mehrebenenanalyse mit Fixed-Effects- und Random-Effects-Modellierung für die Interviewereffekte wird schließlich die Korrelation der Messfehler innerhalb der Befragten eines Interviewers (Intra-Klassen-Korrelation) bestimmt.

Die Arbeit ist folgendermaßen aufgebaut: In Kapitel 2 wird zunächst der aktuelle Forschungsstand zu Messfehlern im Einkommen und zu Interviewereffekten bei Befragungen überblicksweise gezeigt. Anschließend werden in Kapitel 3 die beiden Datenquellen und der Verknüpfungsprozess beschrieben, sowie einige Qualitätsaspekte des verknüpften Datensatzes erörtert. Kapitel 4 beschreibt die Methoden der Untersuchung, Kapitel 5 zeigt deren Ergebnisse und interpretiert diese. Zuletzt werden die Ergebnisse in Kapitel 6 zusammengefasst und in den aktuellen Forschungsstand eingeordnet.

2 Literaturüberblick

Messfehler bei der Angabe des Einkommens in Befragungen wurden bereits in einer Vielzahl von Studien untersucht. Einen Überblick über mögliche Ursachen von Messfehlern im Einkommen geben Moore et al. (2000). Besondere Aufmerksamkeit soll an dieser Stelle Studien geschenkt werden, die die Messfehler mithilfe der Zuspiegelung administrativer Daten betrachten (Duncan und Hill 1985; Bound und Krueger 1991; Bound et al. 1994; Bollinger 1998; Pedace und Bates 2000; Stinson 2002; Kapteyn und Ypma 2007; Bricker und Engelhardt 2008; Gottschalk und Huynh 2010; Kim und Tambourini 2014).

Bound und Krueger (1991) legen anhand eines aus der US-amerikanischen CPS-Befragung und den administrativen Daten der Sozialversicherungsbehörde (SSA) verknüpften Datensatzes nahe, dass es sich bei den Messfehlern im Einkommen einer Person nicht um Messfehler klassischer Art, also Messfehler, deren Höhe unabhängig vom Wert der gemessenen Variable selbst ist, handelt. Die Messfehler sind in dieser Untersuchung zwar um den Nullpunkt zentriert, es besteht jedoch ein negativer Zusammenhang zwischen dem Messfehler im Einkommen und der Einkommenshöhe. Dies impliziert, dass Personen mit niedrigem Einkommen bei der Einkommensangabe tendenziell eher mehr angeben, als sie tatsächlich verdienen, wohingegen Personen mit hohem Einkommen eher weniger angeben, als sie tatsächlich verdienen. Dieser Effekt wird von den Autoren als Mean-Reverting Measurement Error bezeichnet.

Der Mean-Reverting Effekt bei Messfehlern im Einkommen wird von einer Reihe weiterer Studien bestätigt (Bound et al. 1994; Bollinger 1998; Bricker und Engelhardt 2008). Außerdem wurden in einigen Studien auch soziodemographische Merkmale hinsichtlich ihres Einflusses auf den Messfehler untersucht. Die Bildung der Befragten zeigte hier meist keinen oder nur geringen Zusammenhang mit dem Messfehler (Bound und Krueger 1991; Pedace und Bates

2000; Bricker und Engelhardt 2008), während für das Geschlecht der Befragten Effekte gezeigt werden konnten (Bound und Krueger 1991; Bollinger 1998), teilweise auch in Interaktion mit dem Alter (Gottschalk and Huynh 2005).

Soziodemographische Merkmale der Befragten werden auch von Kim und Tambourini (2014) anhand von Daten aus der US-amerikanischen SIPP-Befragung, wiederum verknüpft mit administrativen Daten der SSA, untersucht. Auch hier zeigt sich wieder ein Mean-Reverting Effekt im Messfehler des Einkommens. Im Gegensatz zu anderen Studien betrachten die Autoren den Einfluss von soziodemographischen Merkmalen auf den Messfehler des Einkommens auch im Hinblick auf das Einkommen selbst, indem sie ihre Stichprobe in fünf Einkommensquantile unterteilen und die Einflüsse für jede Untergruppe separat bestimmen. So zeigen sie beispielsweise, dass Afroamerikaner im Vergleich zu Weißen bei niedrigem Einkommen in ihrer Befragungsangabe stärker übertreiben, bei hohem Einkommen jedoch stärker untertreiben. Diese Effekte hätten sich ohne die Berücksichtigung von Einkommensquantilen im Mittel gegenseitig aufgehoben. Neben der ethnischen Zugehörigkeit befinden sich unter den weiteren Variablen, die bei Berücksichtigung der Einkommensquantile signifikante Effekte zeigen, Bildung, Arbeitsplatzsicherheit (gemessen durch das Vorhandensein einer Befristung des Arbeitsvertrages und einen Jobwechsel innerhalb des letzten Kalenderjahres) und betriebliche Altersvorsorge der Befragten. Zwischen den Geschlechtern wird jedoch kein Unterschied hinsichtlich dieser Variablen festgestellt.

In all den zuvor genannten Studien wird der Messfehler als Abweichung der Befragungsangabe vom Wert in den administrativen Daten definiert. Damit wird implizit vorausgesetzt, dass die Einkommensaufzeichnungen in den administrativen Daten als Referenzwert den korrekten Wert abbilden oder zumindest nur in einem so geringen Umfang Fehler aufweisen, dass diese vernachlässigt werden können. Diese Annahme wird jedoch nicht in allen Untersuchungen getroffen (Kapteyn und Ypma 2007; Stinson 2002). Kapteyn und Ypma (2007) modellieren zusätzlich zum Messfehler in der Befragung eine Fehlerstruktur in den administrativen Daten. Als Quelle ihrer modellierten Fehler führen die Autoren falsche Verknüpfungen an, die während des Prozesses der Generierung administrativer Daten entstehen können, falls dafür, wie in deren Studie, Informationen aus verschiedenen Quellen verknüpft werden müssen. Zur Ana-

lyse verwenden sie den schwedischen Teil der Daten aus der europäischen SHARE-Befragung, verknüpft mit administrativen Daten, die über die Sozialversicherungsnummer aus Zensusdaten und Einkommensaufzeichnungen zusammengefügt werden. Auf dieser Datenbasis kommen die Autoren zu dem Schluss, dass erstens kein Mean-Reverting Effekt vorliegt und zweitens in den administrativen Daten Verzerrungen in beträchtlichem Umfang vorliegen. Die Ergebnisse der in den vorhergehenden Absätzen benannten Studien könnten also dadurch verzerrt sein, dass die Abbildung der Realität durch die administrativen Daten vorausgesetzt wird und diese als Referenz zur Bestimmung von Messfehlern dienen. Nichtsdestotrotz werden auch in der hier angestrebten Analyse der Interviewereffekte auf das Einkommen die administrativen Daten als Referenz verwendet.

Der zweite Forschungsbereich, auf den diese Untersuchung aufbaut, ist der Einfluss von Interviewern auf die Teilnahme der Zielpersonen an Befragungen und deren Angaben. Von einem Interviewereffekt wird immer dann gesprochen, wenn sich die von einem bestimmten Interviewer in einer bestimmten Variable erhobenen Werte ähnlicher sind als Werte, die von anderen Interviewern in den gleichen Variablen erhoben werden.

Einen Überblick über den aktuellen Forschungsstand der verschiedenen Quellen und Ausprägungen von Interviewereffekten geben West et al. (2013), wie im Folgenden kurz zusammengefasst wird. Mögliche Ursachen für Interviewereffekte liegen beispielsweise im unterschiedlichen Nachfrageverhalten der Interviewer bei Unklarheiten (Mangione et al. 1992) oder im Auftreten von sozialer Erwünschtheit, die sich in der Interaktion zwischen Befragtem und Interviewer ergeben kann (Schnell 1997). Viele Studien untersuchen, ob die Art der Frage, die dem Befragten gestellt wird, mit der Höhe von auftretenden Interviewereffekten zusammenhängt. Dabei kann zwischen faktischen Fragen zu Tatsachen und Fragen zu Einstellungen des Befragten unterschieden werden. Schnell und Kreuter (2005) geben einen Überblick über die unterschiedlich ausfallenden Ergebnisse diverser Untersuchungen zu diesem Thema. Sie kommen zu dem Ergebnis, dass es keine eindeutigen Kriterien gibt, nach denen sich Variablen hinsichtlich ihrer Anfälligkeit für Interviewereffekte einordnen lassen. Interviewereffekte ließen sich aber durch sorgfältige Einweisung der Interviewer und standardisierte Interviewabläufe verringern.

Eine Möglichkeit der Messung der Interviewereffekte stellt der Intraklassenkorrelationskoeffizient (ICC) dar, der von Kish (1962) als Anteil der Varianz, die durch die Interviewer entsteht, an der Gesamtvarianz (Summe der Varianz ohne Einfluss des Interviewers und der Varianz durch den Interviewer) definiert wird. Dieses Maß wird seither in vielen Studien verwendet um die Höhe der Interviewereffekte zu quantifizieren. In aktuellen Studien werden Interviewereffekte häufig auf ihre Auswirkungen auf Nonresponse (O’Muircheartaigh und Campanelli 1999; Lipps und Pollien 2011; Sala et al. 2012) und die Zustimmung der Befragten zur Zuspiegelung von administrativen Daten (Beste 2011; Sakshaug und Couper 2012; Korbmacher und Schröder 2013) hin untersucht.

Besonderes Augenmerk soll an dieser Stelle auf die Studie von Korbmacher und Schröder (2013) gelegt werden, da sie in der Methodik Gemeinsamkeiten mit der in dieser Arbeit angestrebten Analyse der Interviewereffekte auf den Messfehler des Einkommens aufweist. Die Autoren untersuchen die Interviewereffekte auf die Zustimmungswahrscheinlichkeit zur Zuspiegelung administrativer Daten anhand des deutschen Teils der SHARE-Befragung. Sie verwenden ein logistisches Mehrebenenmodell für Befragte (erste Ebene) und Interviewer (zweite Ebene) mit der Zustimmungswahrscheinlichkeit als abhängiger Variable und Random-Effects für die Interviewer. Zur Kontrolle wurden Variablen nach den Ergebnissen früherer Studien ausgewählt. Diese beinhalten eine Reihe von soziodemographischen Merkmalen der Befragten, Kennzeichen der Interaktion zwischen Interviewern und Befragten, sowie Merkmale des Interviewers selbst. All diese Variablen werden schrittweise ausgehend von einem leeren Modell ohne weitere Variablen zu einem vollen Modell mit Nutzung aller Variablen hinzugefügt. Die Intraklassenkorrelation liegt in dieser Untersuchung mit 41%. Das impliziert, dass Interviewer einen großen Einfluss auf die Entscheidung der Befragten für oder gegen eine Zustimmung zur Zuspiegelung administrativer Daten ausüben.

3 Daten

Alle Analysen wurden mit einem verknüpften Datensatz aus Befragungsdaten und administrativen Daten durchgeführt. Auf Befragungsseite wurden dazu Daten aus dem Panel Arbeitsmarkt und soziale Sicherung des Instituts für Arbeitsmarkt- und Berufsforschung (IAB) verwendet. Für die administrativen Daten konnte auf die Integrierten Erwerbsbiographien zurückgegriffen werden, die ebenfalls vom IAB stammen. Die folgenden Abschnitte sollen einen Überblick über Aufbau und Erhebung beider Datenquellen und die Durchführung ihrer Verknüpfung geben. Anschließend wird der verknüpfte Datensatz deskriptiv analysiert und hinsichtlich von Datenqualität und Selektivitätsprozessen durch die Auswahl bestimmter Fälle beurteilt.

3.1 Integrierte Erwerbsbiographien

Die Bundesagentur für Arbeit speichert als Sozialversicherungsträger in Deutschland Daten über sozialversicherungspflichtige Beschäftigungen, Arbeitslosengeldbezug, Bezug von Grundsicherung, Maßnahmenteilnahme und Arbeitslosenmeldungen, soweit diese zur Erbringung ihres gesetzmäßigen Auftrags erforderlich sind. Dazu gehören die Berechnung der zu leistenden Beiträge durch Beschäftigte und die Auszahlung von Transferleistungen an Anspruchsberechtigte. Als Forschungsinstitut der Bundesagentur für Arbeit kann das IAB die Sozialversicherungsdaten der Bundesagentur für Arbeit zu wissenschaftlichen Zwecken nutzen. Vor der Aufbereitung durch das IAB liegen die Daten zunächst getrennt nach ihrer Informationsquelle vor:¹

¹Einen umfassenden Überblick über Quellen und Aufbereitung der IEB geben Vom Berge et al. (2013) sowie das Handbuch Integrierte Erwerbsbiographien des IAB (nur zum internen Gebrauch, hier in der Version 11.00.01-131009 vom 14.04.2014 verwendet).

- In der Beschäftigten-Historik (BeH) werden die Beschäftigungsmeldungen der Arbeitgeber für abhängig Beschäftigte an die Sozialversicherungen erfasst.
- Die Leistungsempfänger-Historik (LeH) umfasst die Zeiträume, in denen eine Person Lohnersatzzahlungen von der Bundesagentur für Arbeit erhält.
- Die Maßnahmenteilnehmer-Historik (MTH) umfasst Zeiträume, in denen eine Person an Maßnahmen der aktiven Arbeitsmarktförderung, z.B. Arbeitsbeschaffungsmaßnahmen oder Weiterbildungsmaßnahmen, teilnimmt.
- Die Leistungs-Historik-Grundsicherung (LHG, XLHG) umfasst Zeiträume, in denen eine Person Leistungen nach dem SGB II erhält.
- Die Arbeitssuchenden-Historik (ASU, XASU) umfasst Zeiträume, in denen sich eine Person bei einer Arbeitsagentur als arbeitssuchend gemeldet hat.

Diese fünf Quellen werden jeweils auf Personenebene anhand der Sozialversicherungsnummer der Rentenversicherung zu den Integrierten Erwerbs-Biographien (IEB) zusammengefasst. Die Verknüpfung dieser Informationen ermöglicht einen umfassenden Überblick über die Teilnahme am Arbeitsmarkt und den Bezug von Sozialleistungen einer Person im Zeitverlauf. Für den Vergleich von Einkommen in den IEB mit Befragungsdaten sind ausschließlich die Informationen aus der BeH verwendbar, da dies die einzige Quelle ist, die Auskunft über die Einkünfte der betreffenden Person gibt. Die Quellen LeH, MTH, LHG/XLHG und ASU/XASU enthalten hingegen keine Angaben über die Höhe der Unterstützungsleistungen, die die Person von der Bundesagentur für Arbeit oder der für sie zuständigen Kommune erhält.

Die Basis der BeH bilden die Arbeitgeber-Meldungen nach dem integrierten Meldeverfahren für Kranken-, Renten- und Arbeitslosenversicherung. Jeder Arbeitgeber in Deutschland ist verpflichtet, sozialversicherungspflichtige, abhängige Beschäftigungsverhältnisse an die Sozialversicherungsträger zu melden. Darunter fallen auch geringfügig Beschäftigte und mithelfende Familienangehörige, jedoch nicht Beamte, Selbständige und Studierende. Jede sozialversicherungspflichtige Beschäftigung wird zunächst der zuständigen Krankenkasse gemeldet, welche

die Daten an die Rentenversicherung übermittelt, von wo sie an die Bundesagentur für Arbeit als Träger der Arbeitslosenversicherung weitergegeben werden.

Jede sozialversicherungspflichtige Beschäftigung wird mindestens einmal im Jahr gemeldet. Wenn die Beschäftigung durchgehend vom 1. Januar bis zum 31. Dezember eines Jahres anhält, wird zum Ende des Jahres eine Jahresmeldung mit den entsprechenden Informationen erstellt. Bei Beendigung eines bestehenden oder Neuaufnahme eines neuen Arbeitsverhältnisses innerhalb eines Jahres wird eine End- bzw. Aufnahmemeldung erstellt. Außerdem werden Änderungen an einem andauernden Beschäftigungsverhältnis als neue Meldungen aufgenommen, sofern sich Eigenschaften verändern, die einen Einfluss auf die Höhe der Sozialversicherungsbeiträge haben. Dazu gehören beispielsweise Entgeltänderungen, Änderungen in der Arbeitszeit, Krankenkassenwechsel oder Unterbrechungen durch Elternzeit. Aufgrund der verschiedenen Meldungsarten und bei mehreren gleichzeitigen Beschäftigungen können sich die einzelnen Meldungen daher für eine Person überschneiden.

Die übermittelten Merkmale der Beschäftigten enthalten insbesondere tagesgenaue Angaben zu Beginn und Ende der Beschäftigungsepisode, verschiedene Merkmale zur eindeutigen Identifikation des Arbeitnehmers (z.B. Name, Geburtstag, Geschlecht, Sozialversicherungsnummer) und Angaben zum gesamten Bruttoentgelt innerhalb der Beschäftigungsepisode. In den IEB wird dieses Gesamtentgelt gleichmäßig auf den Episodenzeitraum verteilt und als Brutto-Tagesentgelt ausgewiesen. Der Episodenzeitraum wird dabei nicht in Arbeits-, sondern in Kalendertagen gemessen.² Bruttoeinkommen, die über der Beitragsbemessungsgrenze der Arbeitslosenversicherung liegen, werden dabei zensiert und lediglich in Höhe der Bemessungsgrenze ausgewiesen.

Bei der Aufbereitung der Meldungen zu wissenschaftlichen Zwecken durch das IAB werden ausschließlich Jahresmeldungen, Abmeldungen, Unterbrechungsmeldungen und Einmalzahlungen berücksichtigt, da nur diese Meldungsarten Angaben zum Entgelt beinhalten. Beschäftigungsanmeldungen sind nicht mit Angaben zum Entgelt versehen und werden nicht berücksichtigt. Außerdem werden Meldungen als Duplikate entfernt, sofern sie dieselben Einträge in den Merkmalen Versicherungsnummer, Betriebsnummer, Beginndatum, Enddatum,

²Für das ganze Jahr wird von 365 Kalendertagen ausgegangen.

Entgelt und Abgabegrund aufweisen.

3.2 Panel Arbeitsmarkt und soziale Sicherung

Das Panel Arbeitsmarkt und soziale Sicherung (PASS) ist eine als Panel angelegte, jährlich wiederholte Haushaltsbefragung des IAB.³ In der ersten Erhebungswelle (2006/2007) wurden 18954 Personen aus 12794 Haushalten zu ihrer Teilnahme am Arbeitsmarkt, Arbeitslosigkeit, ihren Lebensumständen und zu persönlichen Einstellungen befragt.⁴ Aufgrund der Ausrichtung des Panels auf die Forschung im deutschen Sozialleistungsbereich, sind im PASS Leistungsempfängerhaushalte überrepräsentiert. Etwas mehr als die Hälfte der Haushalte wurde direkt aus einer Leistungsempfängerstichprobe der Bundesagentur für Arbeit gezogen (BA-Stichprobe), die restlichen Haushalte wurden aus einer allgemeinen geschichteten Bevölkerungsstichprobe mit überproportionalem Anteil an Haushalten mit geringem Einkommen (Microm-Stichprobe) gezogen.

Aufgrund der Auslegung des PASS als Panel wurden die Teilnehmer der Befragung in den folgenden Erhebungswellen um die Teilnahme an einer erneuten Befragung gebeten, sofern sie bei der ersten Befragung ihr Einverständnis zu einer Wiederholungsbefragung geäußert hatten. Wie bei allen Panel-Studien kam es jedoch auch beim PASS zu einer nennenswerten Verringerung der Teilstichproben im Zeitverlauf. In der zweiten Erhebungswelle nahmen nur 48% der Befragten aus der BA-Stichprobe und 65% der Befragten aus der Microm-Stichprobe an der Wiederholungsbefragung teil. Um die Repräsentativität der Stichprobe aufrecht zu erhalten, wurden in jeder weiteren Erhebungswelle zusätzlich Auffrischungstichproben gezogen. In allen Erhebungswellen wurden Zugänge aus Leistungsempfängerhaushalten neu im PASS aufgenommen, in der fünften Erhebungswelle (2011) zusätzlich auch wieder Personen einer allgemeinen Bevölkerungsstichprobe, die aus den Einwohnermelderegistern gezogen wurde.

Die ausgewählten Haushalte erhielten zunächst per Brief die Ankündigung eines telefonischen Kontaktversuchs zur Befragung (CATI - „computer assisted telephone interview“). Als Anreiz zur Teilnahme wurde eine Aufwandsentschädigung in Höhe von zehn Euro je realisier-

³Einen umfassenden Überblick über Erhebung und Aufbereitung der PASS-Befragungsdaten geben die Methoden- und Datenreporte des FDZ (Berg et al. 2012; Jesske und Schulz 2012; Bethmann et al. 2013)

⁴Die Response-Rate in der ersten Erhebungswelle des PASS lag bei 30.5%

tem Interview in Aussicht gestellt.⁵ Falls ein Haushalt nach mehrfachen Kontaktversuchen telefonisch nicht erreichbar war, wurde versucht, die Befragung direkt im Haushalt in Anwesenheit des Interviewers (CAPI - „computer assisted personal interview“) durchzuführen. Eine persönliche Befragung war zudem auch bei ausdrücklichem Wunsch des Haushalts möglich. Der überwiegende Teil der Interviews wurde auf Deutsch durchgeführt. Um auch Immigranten mit eingeschränkten Sprachkenntnissen die Teilnahme an der Befragung zu ermöglichen, wurden zusätzlich die Sprachen Englisch (nur in der ersten Erhebungswelle), sowie Türkisch und Russisch (in allen Erhebungswellen) angeboten.

Die Befragung wurde zunächst auf Haushaltsebene durchgeführt. Falls der Haushalt Sozialleistungen nach dem zweiten Buch des Sozialgesetzbuchs bezog, wurde dazu der Antragsteller ausgewählt. In Haushalten, die keine Sozialleistungen bezogen, wurde dazu diejenige Person ausgewählt, die nach eigener Einschätzung am besten mit den Haushaltsangelegenheiten vertraut war. Die Fragen zum Haushalt umfassten unter anderem Informationen zu Haushaltszusammensetzung, Wohnsituation und Sozialleistungsbezug.

Anschließend wurden alle Mitglieder des Haushalts, die älter als 15 Jahre waren, auf Personenebene befragt. Für Personen, die 65 Jahre oder älter waren, wurde ein verkürzter „Seniorenfragebogen“ verwendet. Schwerpunkt der Personenbefragung waren die Teilnahme am Arbeitsmarkt und der Bezug von Sozialleistungen, also die Erwerbsbiographie der Befragten. Im Fokus standen auch persönliche Einstellungen, soziale Herkunft und Bildung.

Ab der vierten Erhebungswelle (2010) wurden die Befragten gebeten, alle arbeitsmarktrelevanten Episoden, die sie seit dem Beginn des Jahres 2008 durchlaufen hatten, jeweils einzeln und monatsgenau mit Beginn- und Enddatum anzugeben.⁶ Ziel war es dabei, die persönlichen Erwerbsbiographien der Befragten nachzubilden. Es wurde zwischen den folgenden Episodenarten unterschieden:

- Erwerbstätigkeit mit einem Einkommen von mehr als 400 Euro
- Arbeitslosigkeitsmeldung (umfasst auch die Teilnahme an Maßnahmen der

⁵Panelbereite Teilnehmer der vorangegangenen Welle erhielten den Anreiz in Höhe von zehn Euro mit dem Anschreiben unabhängig von ihrer erneuten Teilnahme bereits vor Zustandekommen des Interviews.

⁶Bei wiederholt Befragten wurden die Informationen aus den vorangegangenen Erhebungswellen fortgeschrieben.

Bundesagentur für Arbeit)

- Schulische Ausbildung
- Berufliche Ausbildung, Lehre oder Studium
- Wehrdienst oder Zivildienst
- Hausfrau/Hausmann
- Mutterschutz, Erziehungsurlaub oder Elternzeit
- Rentner, Pensionär oder im Vorruhestand

Auch beim Einkommen der Befragten unterscheidet sich die Erhebungsmethode in den einzelnen Erhebungswellen. In der ersten Erhebungswelle wurde das monatliche Einkommen nur für die Haupterwerbstätigkeit der Befragten erhoben. Von der zweiten bis zur vierten Erhebungswelle sollten die Befragten ihr monatliches Einkommen summiert über alle gleichzeitigen nicht-geringfügigen Tätigkeiten angeben. Ab der fünften Erhebungswelle wurde das monatliche Einkommen dann einzeln für jede angegebene Beschäftigungsepisode, die über den Zeitpunkt der Befragung hinaus fortgeführt wurde, erfasst. Falls die Befragten die genaue Angabe ihres Einkommens ablehnten, wurden sie um die Angabe in Kategorien mit einer Breite von 500 Euro gebeten. Durch die einzelne Erhebung aller Episoden lassen sich bei Vorliegen mehrerer gleichzeitiger Beschäftigungsverhältnisse die jeweiligen Beiträge separieren. Die Erwerbsverläufe im PASS ab der fünften Erhebungswelle weisen somit die gleiche Struktur wie die IEB auf.

Gemäß der Datenschutzbestimmungen der Bundesrepublik Deutschland ist das Einverständnis der befragten Personen zwingende Voraussetzung für die Verknüpfung ihrer Angaben. Die Befragten wurden daher um ihre Zustimmung gebeten und auf ihr Recht zum Widerruf der Zustimmung hingewiesen. Der überwiegende Teil der Befragten war mit der Zusammenführung der Daten einverstanden, je nach Erhebungswelle lag die Zustimmungsrate zwischen 76% und 87%.

3.3 Datenverknüpfung und Fallselektion

Zur Erstellung eines verknüpften Datensatzes wurden der Personendatensatz sowie Informationen über die eingesetzten Interviewer der fünften PASS-Erhebungswelle aus dem Jahr 2011 und die IEB mit Stand Ende Dezember 2011⁷ verwendet. Insgesamt wurden in der fünften Erhebungswelle 15607 Personen aus 10235 Haushalten befragt. Davon wiesen 5764 Personen nur ein über den Interviewzeitpunkt hinaus andauerndes Beschäftigungsverhältnis auf⁸ und waren bereit ihr Einkommen anzugeben. 131 Personen wollten ihr Einkommen nicht angeben oder gaben an, dieses nicht zu kennen. Unter den 5764 Personen mit Einkommensangabe zogen es 452 (7.8%) vor, ihr Einkommen nicht exakt, sondern lediglich in Kategorien einer Breite von 500 Euro angeben. Mit 5312 Personen (92.2%) war jedoch der Großteil der Befragten mit Einkommensangabe bereit, dieses auch genau anzugeben. Diese Personengruppe stellt die Basis für die Datenverknüpfung und Fallselektion zur Erstellung des verknüpften Datensatzes zur Analyse der Interviewereffekte dar. Tabelle 3.1 zeigt die Entwicklung der Stichprobengröße im Laufe der einzelnen Verknüpfungs- und Selektionsschritte.

Die Zustimmung zur Verknüpfung von Befragungsangaben und administrativen Daten war unter den PASS-Teilnehmern, die ihr Einkommen genau angegeben hatten und nur ein Beschäftigungsverhältnis zum Interviewzeitpunkt aufwiesen, mit 4975 Personen (93.8%) sehr hoch. Diese Teilnehmer wurden anhand der Merkmale Name, Geschlecht, Adresse und Geburtsdatum in den IEB identifiziert. Wie von Antoni und Bethmann (2014) beschrieben, wurde den PASS-Teilnehmern aus der BA-Stichprobe ihr IEB-Konto direkt über ihre Bedarfsgemeinschaft zugeordnet. Bei Scheitern der direkten Zuordnung, sowie bei allen Teilnehmern aus der Bevölkerungsstichprobe wurden exakte Übereinstimmungen der beiden Datenquellen in verschiedenen Kombinationen der zuvor genannten Eigenschaften identifiziert (exaktes Record Linkage). Anschließend wurde die Verknüpfung bei Personen ohne exakte Übereinstimmung⁹ auf Basis von Ähnlichkeitsmaßen (probabilistisches Record Linkage) durchgeführt.

⁷IEB v10.00.00, April 2014.

⁸Zur Sicherstellung von eindeutigen Verknüpfungen auf Beschäftigungsepisodenebene wurden ausschließlich Personen mit nur einem Beschäftigungsverhältnis zum Interviewzeitpunkt betrachtet.

⁹Eine fehlende exakte Übereinstimmung könnte beispielsweise durch fehlerhafte Schreibweise der Merkmale in einer der beiden Datenquellen, veraltete Einträge in den IEB, fehlerhafte Eintragungen des Interviewers oder Falschangaben in PASS bedingt sein.

Tabelle 3.1: Datenverknüpfung und Fallselektion

Selektionsschritt	Verbleibende Fälle	Verbleibender Anteil je Schritt
Gesamtstichprobe der fünften Erhebungswelle	15607	
Personen mit Einkommensangabe und nur einem Beschäftigungsverhältnis	5764	36,9%
Personen mit genauer Einkommensangabe	5312	92,2%
Zuspielerlaubnis vorhanden	4975	93,7%
Zuordnung zu IEB (Personenebene)	4358	87,6%
Eindeutige IEB-PASS-Relation (Personenebene)	4171	95,7%
Erfolgreiche Zuordnung auf Episodenebene	3620	86,8%
Einkommen unter Beitragsbemessungsgrenze	3488	96,4%
Manuelle Entfernung unplausibler Verknüpfungen und Interviewereintragen	3421	98,1%
Entfernung von Personen mit fehlenden Werten in mindestens einer der Analysevariablen	3075	89,9%
Interviewer hat mehr als 5 Interviews in der Endstichprobe durchgeführt	2694	87,6%

Dadurch konnte insgesamt 4358 PASS-Befragten (87.6%) ein IEB-Konto zugeordnet werden. Nach Abzug von nicht eindeutigen Zuordnungen (Fälle, in denen mehrere PASS-Konten einem IEB-Konto oder mehrere IEB-Konten einem PASS-Konto zugeordnet wurden) verblieben 4171 eindeutige Verknüpfungen auf Personenebene. Um das Beschäftigungseinkommen aus dem PASS und den IEB vergleichen zu können musste die Verknüpfung für jede Person auf Episodenebene vorgenommen werden. Voraussetzung dafür war, dass, wie beim PASS, in den IEB nur eine Beschäftigung zum Interviewzeitpunkt verzeichnet war. Diese eindeutige Verknüpfung auf Episodenebene wurde bei 3620 Personen erreicht.

Die erfolgreichen eindeutigen Verknüpfungen wurden als Basis für die Auswahl einer geeigneten Stichprobe zur Analyse der Interviewereffekte verwendet. 3.6% der erfolgreich verknüpften Befragten mussten aufgrund der Zensierung der Bruttoeinkommen in den IEB oberhalb der Beitragsbemessungsgrenze der Arbeitslosenversicherung (2011: 5500 Euro in den alten Bundesländern, 4800 Euro in den neuen Bundesländern) ausgeschlossen werden. Wegen unplausibler Eintragungen des Interviewers, unplausibler PASS-IEB-Verknüpfungen oder Beamtenstatus des Befragten¹⁰ wurden im nächsten Schritt 1.9% der Personen manuell entfernt. Um die Vergleichbarkeit der verschiedenen Modelle zu den Interviewereffekten durch eine konstante Stichprobe sicherzustellen, konnten nur Personen mit vollständigen Angaben in allen relevanten Analysevariablen (3075 Personen bzw. 89.9%) herangezogen werden. Um zufällige Effekte zu beschränken, wurden im letzten Selektionsschritt nur noch Personen beibehalten, die von einem Interviewer befragt wurden, der mindestens sechs Interviews in der verbleibenden Stichprobe durchgeführt hatte. Somit stehen für die Analyse der Interviewereffekte insgesamt 2694 Personen zur Verfügung. Dies entspricht 17.4% aller Befragten der fünften Erhebungswelle des PASS.

Mangels Verfügbarkeit derartiger Angaben in den IEB wurden die meisten soziodemographischen und beschäftigungskennzeichnenden Merkmale der Befragten ausschließlich aus dem PASS entnommen.¹¹ Die IEB wurden neben dem Bruttoeinkommen lediglich als Quelle für

¹⁰Beamte sind mit ihrer Haupttätigkeit grundsätzlich nicht in den IEB verzeichnet, hier ist daher von einer Falschzuordnung mit einer im Interview nicht angegebenen Nebentätigkeit oder einer falschen Verknüpfung auf Personenebene auszugehen.

¹¹Die meisten soziodemographischen Angaben, z.B. Bildung, Familienstand oder Anzahl der Kinder, sind in den IEB ausschließlich für Leistungsempfänger vorhanden, da diese Daten von der Bundesagentur für Arbeit nur zur Bestimmung von Leistungsansprüchen und zur Vermittlung benötigt werden.

Geschlecht, Alter und Bundesland des Wohnorts der Befragten herangezogen. Der verknüpfte Datensatz enthält folgende Befragtenmerkmale:

- Geschlecht des Befragten
- Alter des Befragten
- Bruttoeinkommen (PASS): Angegebenes monatliches Einkommen des Befragten vor Steuern und Sozialabgaben, bei Verweigerung der genauen Angabe wurde um eine Angabe in Kategorien (Breite 500 Euro) gebeten.
- Bruttoeinkommen (IEB): Einkommen des Befragten vor Steuern und Sozialabgaben wie aus den administrativen Daten. Das in den IEB verzeichnete Tageseinkommen wurde in das Monatseinkommen umgerechnet.
- Interviewsprache: Die Interviews wurden in der fünften Erhebungswelle in den Sprachen Deutsch, Russisch und Türkisch durchgeführt.
- Migrationshintergrund: Bei Zuzug nach Deutschland des Befragten selbst oder mindestens eines Elternteils.
- Panelbereitschaft: Gibt an, ob der Befragte bereit ist, in der nächsten Erhebungswelle erneut an der Befragung teilzunehmen.
- Familienstand des Befragten (verheiratet/eingetragene Partnerschaft oder alleinstehend)
- Kinder: Gibt an, ob der Befragte Kinder hat.
- Schulbildung: Höchster erreichter Schulabschluss des Befragten. Es wird zwischen Abitur, Realschulabschluss, Hauptschulabschluss, sowie keinem oder sonstigem Abschluss unterschieden.
- Stellung im Beruf: Gibt an, ob der Befragte als Arbeiter oder Angestellter einzuordnen ist.
- Befristung im Arbeitsverhältnis: Gibt an, ob das Arbeitsverhältnis des Befragten befristet ist.

- Stichprobenkennzeichen: Gibt an, ob der Befragte über die BA-Stichprobe (oder deren Aufstockung) oder über die Bevölkerungstichprobe (oder deren Auffrischung) in das Panel aufgenommen wurde.
- Jobwechsel: Gibt an, ob der Befragte innerhalb der letzten drei Monate vor dem Interview seinen Job gewechselt hat.
- Kurzzeittätigkeit: Kennzeichen für Personen, deren Erwerbstätigkeit eine Dauer von weniger als einem Monat aufwies.
- Interviewmodus: Gibt an, ob das Interview telefonisch (CATI) oder persönlich im Haushalt des Befragten (CAPI) durchgeführt wurde.
- Wohnort: Wohnort in den alten oder neuen Bundesländern.

Außerdem stehen Informationen über den Interviewer und die Interviewsituation zur Verfügung, die zur Analyse der Interviewereffekte herangezogen werden:

- Zuordnungsvariable zwischen Befragtem und Interviewer
- Geschlecht des Interviewers
- Alter des Interviewers
- Interviewererfahrung: Anzahl der Jahre, die der Interviewer zum Interviewzeitpunkt beim Erhebungsinstitut ifas gearbeitet hat.¹²
- Schulbildung des Interviewers: Höchster erreichter Schulabschluss des Interviewers. Es wird zwischen Abitur, Realschulabschluss, Hauptschulabschluss, sowie keinem oder sonstigem Abschluss unterschieden.
- Teilnahme des Interviewers in der Vorwelle
- Gleiche Interviewer/Befragten-Konstellation in der vierten Erhebungswelle
- Anzahl der durchgeführten Interviews des Interviewers in der fünften Erhebungswelle

¹²Das Erhebungsinstitut ifas führt seit der vierten Erhebungswelle die Befragungen im PASS durch. Über frühere Tätigkeitszeiten des Interviewers bei anderen Erhebungsinstituten stehen keine Informationen zur Verfügung.

3.4 Datenqualität und Selektivitäten

Die Validität der Ergebnisse einer Untersuchung hängt entscheidend von der Qualität der verwendeten Daten ab. Zudem muss geprüft werden, ob die Ergebnisse der Befragung durch die systematische Teilnahmeverweigerung bestimmter Personengruppen verzerrt sind. Eine weitere Quelle potenzieller Verzerrungen stellen Selektivitätsprozesse dar, die sich aus der wie in Abschnitt 3.3 beschriebenen Auswahl bestimmter Untergruppen aus der Gesamtstichprobe ergeben können. Darunter fallen insbesondere das Einverständnis der Befragten zur Verknüpfung der Datenquellen und die Reduktion der Analyse auf Personen, die bereit waren ihr Einkommen exakt anzugeben. Diese Aspekte sollen im Folgenden diskutiert werden.

3.4.1 Integrierte Erwerbsbiographien

Wegen der Verwendung der IEB als Referenz für das tatsächliche Einkommen einer Person zum Vergleich mit ihrer Befragungsangabe gelten hier besonders strenge Qualitätsanforderungen. Jacobebbinghaus und Seth (2007) stufen die Daten aus den IEB als sehr zuverlässig ein.¹³ Dies trifft ganz besonders auf Variablen zu, die nicht nur für statistische Zwecke erhoben wurden, sondern von der Bundesagentur für Arbeit direkt zur Erfüllung ihrer gesetzmäßigen Aufgaben benötigt werden. Darunter fällt bei den Beschäftigungsmeldungen vor allem die Höhe des Bruttoeinkommens selbst, da sie die Berechnungsgrundlage für die Beiträge zur Arbeitslosenversicherung und die Ansprüche bei Arbeitslosigkeit darstellt. Hier kann daher von einer sehr hohen Genauigkeit ausgegangen werden.

Falls es sich bei den Episoden nicht um ununterbrochene und über das ganze Jahr andauernde Beschäftigungen handelt, die in den IEB als Jahresmeldungen verbucht werden, können die gemeldeten Episodenzeiträume Ungenauigkeiten aufweisen.¹⁴ Unpräzise Angaben dieser Art wirken sich auf die Einkommenswerte in den IEB aufgrund der zeitraumbezogenen Ausweisung als (Kalender-) Tageseinkommen aus. Der Effekt wird dabei umso stärker, je kürzer das Beschäftigungsverhältnis schon zum Interviewzeitpunkt andauerte. Aus diesem

¹³Die Autoren beziehen ihre Aussagen auf den Datensatz SIAB des IAB. Dieser stellt eine Teilstichprobe der IEB dar.

¹⁴Mögliche Ungenauigkeiten bei den Episodenzeiträumen sind im Detail im Handbuch Integrierte Erwerbsbiographien des IAB beschrieben.

Grund wurde eine Kontrollvariable für einen Jobwechsel innerhalb der letzten drei Monate vor dem Interviewzeitpunkt erstellt. Dieser Fall betraf jedoch mit 8.4% der Analyse-Stichprobe nur einen geringen Teil der Befragten. Insgesamt ist daher von einer hohen Genauigkeit der Einkommenswerte aus den IEB auszugehen.

Durch Verzögerungen beim Meldeverfahren, kann es dazu kommen, dass Beschäftigungsmeldungen durch den Arbeitgeber erst mit einer Verspätung von mehr als einem Jahr eintreffen.¹⁵ Die IEB werden jedoch nicht bei jedem Eintreffen einer einzelnen neuen Beschäftigungsmeldung, sondern gebündelt für jedes Kalenderjahr nach 6, 12, 18 und 36 Monaten aktualisiert. Die hier verwendeten IEB der Version 10.00.00 vom April 2014 erhalten daher für das Jahr 2011 nur Beschäftigungsmeldungen, die im Zeitraum von bis zu 18 Monaten später (bis zum 30. Juni 2013) eingetroffen sind. Ein Teil der gescheiterten Verknüpfungen auf Episodenebene lässt sich darauf zurückführen. In diesen Fällen wurde zwar das richtige IEB-Konto identifiziert und erfolgreich auf Personenebene verknüpft, aufgrund der noch nicht eingegangenen Beschäftigungsmeldung konnte jedoch das administrative Einkommen zum Interviewzeitpunkt nicht festgestellt werden. Daher kann die Abweichung zwischen administrativem Einkommen und Befragungseinkommen bei diesen Personen nicht ermittelt werden, so dass sie aus der Analyse ausgeschlossen werden müssen.

3.4.2 Panel Arbeitsmarkt und soziale Sicherung

Bei der Verknüpfung der PASS-Daten mit den IEB zum Zweck des Einkommensvergleichs verdient das Auftreten dreier möglicher Quellen für Verzerrungen besondere Aufmerksamkeit:

- Verzerrungen durch Teilnahmeverweigerung bestimmter Personengruppen an der Befragung (Unit Nonresponse Bias)
- Verzerrungen durch die ausschließliche Auswahl von Personen mit exakter Einkommensangabe (Item Nonresponse Bias)
- Verzerrungen durch Verweigerung des Einverständnisses zur Datenverknüpfung durch bestimmte Personengruppen (Consent Bias)

¹⁵Siehe dazu auch Vom Berge et al. (2013).

Durch Teilnahmeverweigerung von Zielpersonen einer Befragung (Unit Nonresponse Bias) kommt es zu Verzerrungen in den Daten, wenn diese sich auf bestimmte Personengruppen konzentriert und dadurch bestimmte Personengruppen im Vergleich zur Zielpopulation der Befragung über- bzw. unterrepräsentiert sind. Schnell et al. (2010) zeigen den Einfluss der Teilnahmeverweigerung im PASS für die Leistungsempfänger-Teilstichprobe in der ersten Erhebungswelle auf.¹⁶ Die Autoren betrachten die Variablen Alter, Geschlecht, Lohn, Familienstand, Zahl der Kinder, Zeit seit dem letzten Job und Gesamtarbeitslosigkeitszeit hinsichtlich des Auftretens von Unit Nonresponse Bias. Sie kommen zu dem Ergebnis, dass trotz der geringen Teilnahmequote von 28.7% der Zielpersonen in dieser Teilstichprobe bei keiner der betrachteten Variablen außer der Haushaltsgröße und der Gesamtarbeitslosigkeitszeit, die beide bei der Analyse der Interviewereffekte nicht berücksichtigt werden, ein signifikanter Unit Nonresponse Bias auftritt.

Item Nonresponse Bias entsteht dann, wenn Befragte die Antworten auf bestimmte Fragen nicht zufällig, sondern nach einem dahinter liegenden Muster verweigern. Würden sich beispielsweise Bezieher hoher Einkommen deutlich öfter entscheiden, ihr Einkommen nicht anzugeben als Bezieher niedriger Einkommen, würde es bei der Verwendung des angegebenen Einkommens als Ziel- oder Einflussvariable einer Untersuchung zu Verzerrungen der Ergebnisse kommen. In dieser Untersuchung der Interviewereffekte ist das Einkommen als Zielvariable gleichzeitig auch eine Selektionsvariable. Da der Vergleich beider Einkommenswerte nur bei Vorliegen in beiden Quellen möglich ist, müssen Personen, die über kein Einkommen verfügen, Personen deren Einkommen nicht in den IEB zu finden ist, und Personen, die die Angabe ihres Einkommens verweigert hatten, aus der Untersuchung ausgeschlossen werden. Außerdem wurden keine Personen betrachtet, die ihr Einkommen nicht genau, sondern in Kategorien angegeben hatten. Aus diesem Grund soll die Untersuchungsgruppe im Folgenden anhand der Ausprägungen einiger soziodemographischer Merkmale, die später unter anderem zur Erklärung der Abweichungen herangezogen werden, mit der Gesamtstichprobe des PASS,

¹⁶Der Nonresponse Bias kann nur für die Leistungsempfängerstichprobe bestimmt werden, da bei dieser Personengruppe im Gegensatz zur Bevölkerungsstichprobe auch bei Nichtteilnahme an der Befragung Informationen über die Zusammensetzung der Bedarfsgemeinschaften vorliegen.

der Unterstichprobe aller Beschäftigten¹⁷ und den Personen, die die Antwort auf die Einkommensfrage verweigert, oder ihr Einkommen nur kategorial angegeben hatten¹⁸, verglichen werden. Tabelle 3.2 gibt hierzu einen Überblick.

Die Variablen Geschlecht, Alter und Migrationshintergrund der Befragten zeigen lediglich geringe Unterschiede in der Häufigkeit ihrer Ausprägungen über die fünf Stichproben. Von der Interpretation des geringfügig höheren Anteils an Frauen unter den kategorialen Einkommensangaben und des höheren Anteils an Migranten unter den Angabeverweigerern wird an dieser Stelle wegen der geringen Fallzahlen in beiden Gruppen abgesehen.

Größere Unterschiede bestehen bei der Bildung der befragten Personen. Die Gesamtstichprobe, die - bedingt durch das Studiendesign des PASS - Leistungsempfänger überrepräsentiert, weist gegenüber den anderen Stichproben, die jeweils nur Beschäftigte enthalten, einen höheren Anteil von Personen mit geringer schulischer Bildung (kein Abschluss, bzw. Hauptschulabschluss) auf. Diese Selektivität rührt von den schlechteren Chancen von geringer Gebildeten auf den Arbeitsmarkt her und ist naturgemäß bei der Untersuchung von Einkommen aus Beschäftigung nicht zu vermeiden. Dementsprechend fällt in der Gesamtstichprobe auch der Anteil der Befragten aus der Leistungsempfängerstichprobe im Vergleich zu den Beschäftigungsstichproben höher aus.

Ein Vergleich hinsichtlich des Einkommens ist nur zwischen den Beschäftigungsstichproben möglich. Auffallend sind hier die deutlich höheren Werte bei der Stichprobe der Gesamtbeschäftigten gegenüber den verknüpften Stichproben. Dieser Effekt ist hauptsächlich auf das Verwerfen aller Beobachtungen oberhalb der Beitragsbemessungsgrenze zurückzuführen. Einen Hinweis darauf, dass hiermit einige Extremwerte mit größerem Einfluss auf den Einkommensdurchschnitt entfernt wurden, gibt die Entwicklung des Medianeinkommens, das durch die Bereinigung weniger stark sinkt. Da jedoch in den IEB oberhalb der Beitragsbemessungsgrenze keine Informationen über das administrative Einkommen vorliegen, musste diese Bereinigung trotz der entstehenden Selektivitäten durchgeführt werden. Außerdem scheinen Bezieher höherer Einkommen weniger oft bereit zu sein, ihr Einkommen anzugeben. Die ad-

¹⁷An dieser Stelle wurden alle Beschäftigten betrachtet, die zum Interviewzeitpunkt laut eigener Angabe in PASS genau ein Beschäftigungsverhältnis aufwiesen. Siehe dazu auch Spalte 2 in Tabelle 3.2

¹⁸Auch für diese beiden Personengruppen wurden zum Vergleich verknüpfte Stichproben erstellt, um das administrative Einkommen in den IEB zu vergleichen. Siehe dazu auch die Spalten 4 und 5 in Tabelle 3.2.

Tabelle 3.2: Vergleich der Teilstichproben anhand soziodemographischer Merkmale

			Verknüpfte Stichproben		
	PASS Gesamtstich- probe Welle 5	Beschäftigte	Genauere Einkommensangabe	Kategoriale Einkommens- angabe	Keine Einkommensangabe
Anzahl der Personen	15607	5764	2694	208	45
Geschlecht					
Weiblich	52,9%	48,2%	52,3%	56,2%	53,3%
Männlich	47,1%	51,8%	47,7%	43,8%	46,7%
kA	0%				
Durchschnittsalter	45	43	42	43	44
Migrationshintergrund					
Ja	24,2%	21,6%	21,8%	22,1%	33,3%
Nein	73,2%	76,5%	78,2%	76,0%	66,7%
kA	2,6%	1,9%		1,9%	
Bildungsabschluss					
Abitur	24,0%	31,8%	25,9%	24,5%	26,7%
Realschule	33,0%	39,9%	45,2%	40,9%	37,8%
Hauptschule	33,7%	25,8%	26,6%	29,8%	33,3%
Kein Abschluss	5,0%	2,3%	2,3%	4,8%	2,2%
kA/sonstige	4,3%	0%			
Stichprobe:					
BA	54,5%	45,5%	48,9%	46,2%	40%
Microm	45,5%	54,5%	51,1%	53,8%	60%
Zuspieldbereitschaft	81,9%	92,2%	100%	100%	100%
Durchschnittseinkommen (PASS)	n/a	2242	1960	1847	n/a
Medianeinkommen (PASS)	n/a	1900	1800	1750	n/a
Durchschnittseinkommen (IEB)	n/a	n/a	2071	2154	2281
Medianeinkommen (IEB)	n/a	n/a	1853	1968	1853
Messfehler (Mittel)	n/a	n/a	-112	-307	n/a
Messfehler (Median)	n/a	n/a	-86	-236	n/a

ministrativen Einkommenswerte fallen dementsprechend bei den Antwortverweigerungen und den kategorialen Angaben höher aus.

Zuletzt soll untersucht werden, inwiefern es durch die zur Verknüpfung erforderliche Zuspäherlaubnis zu Selektivitäten kommt (Consent Bias). Dies ist der Fall, wenn sich die Untergruppen der Zustimmenden und der Verweigernden systematisch hinsichtlich von Variablen der Befragung oder der administrativen Daten unterscheiden. Auftreten und Ausmaß derartiger Verzerrungen wurden von Sakshaug und Kreuter (2012) anhand einer mit den IEB verknüpften Stichprobe der Erwerbslosen aus der ersten Erhebungswelle des PASS diskutiert. Neben dem Consent Bias wurden dort auch der Nonresponse Bias und der Measurement Error Bias (Abweichung der Befragungsangabe vom entsprechenden Wert in den administrativen Daten) für eine Reihe von Variablen¹⁹ berechnet und verglichen. Es wurde nur bei den Variablen Alter und Migrationshintergrund ein signifikanter Consent Bias festgestellt, der jeweils gering ausfiel. Die restlichen Variablen, darunter auch das in dieser Untersuchung betrachtete Einkommen, zeigten keinen signifikanten Consent Bias. Im Vergleich zu Nonresponse Bias und Measurement Error Bias fiel der Consent Bias geringer aus. Die Unterschiede waren dabei insbesondere für die Variablen Einkommen und Migrationshintergrund signifikant.

Beste (2011) betrachtete anhand der Gesamtstichprobe des PASS aus der ersten Erhebungswelle, ob bestimmte Variablen einen Einfluss auf die Zustimmungswahrscheinlichkeit hatten.²⁰ Dabei zeigten Geschlecht, Alter und Bildung keinen signifikanten Einfluss, höheres Bruttoeinkommen und das Vorliegen eines Beschäftigungsverhältnisses führten jedoch zu einer höheren Zustimmungswahrscheinlichkeit.²¹ Befragte, die nach Deutschland eingewandert waren, erteilten dagegen seltener die Zustimmung. Außerdem wiesen mehrere Merkmale des Interviewers einen signifikanten Zusammenhang mit der Zustimmungsentscheidung auf, darunter Alter, Bildungsabschluss und Geschlecht des Interviewers. Insgesamt fiel der Anteil der durch die Befragten- und Interviewervariablen erklärten Gesamtvarianz im Zustimmungsver-

¹⁹Die betrachteten Variablen umfassten Alter, Migrationshintergrund, Sozialleistungsbezug, Behindertenstatus, Erwerbsstatus und Einkommen.

²⁰Im Gegensatz zu Sakshaug und Kreuter (2012) verwendete Beste (2011) ausschließlich die Angaben aus dem PASS.

²¹Dies unterstreicht auch den Unterschied in der Zustimmungswahrscheinlichkeit zwischen der Gesamtstichprobe und der Beschäftigtenstichprobe (81.9% gegenüber 92.2%) im in dieser Arbeit analysierten Datensatz (siehe Tabelle 3.2).

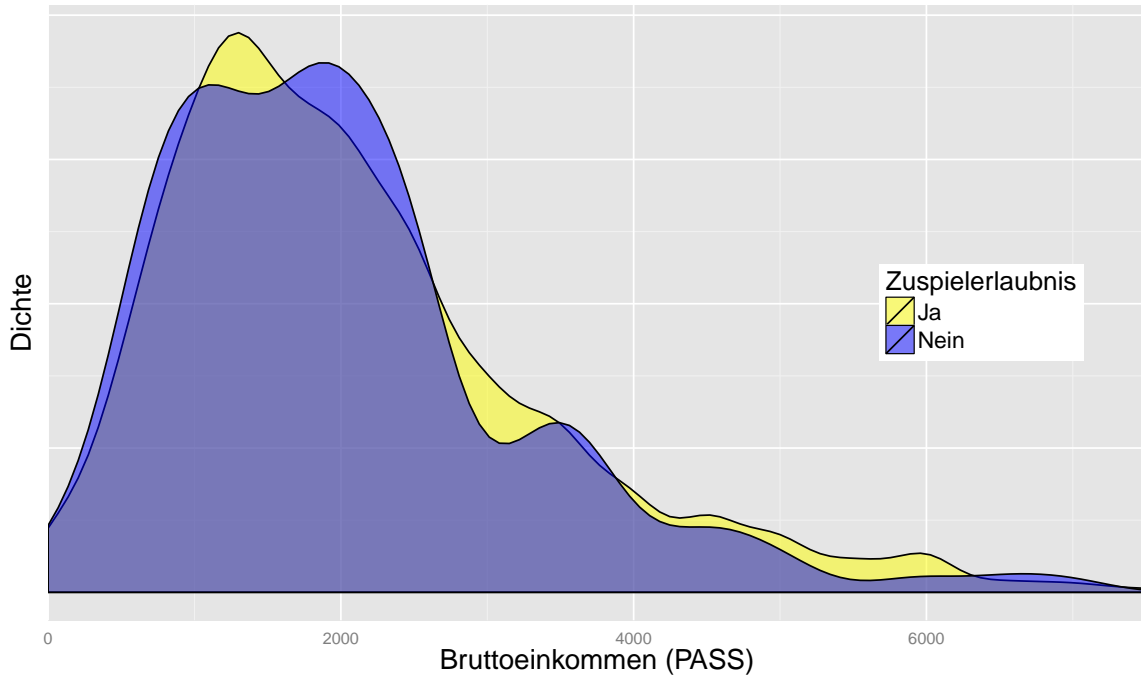
halten jedoch eher gering aus, so dass Beste (2011) von geringen Selektivitätseffekten durch den Zustimmungsprozess ausgehen konnte.

Tabelle 3.3 fasst eine Auswahl der Unterschiede zwischen Zustimmenden und Verweigernden der Beschäftigtenstichprobe zusammen, die wie in Abschnitt 3.3 beschrieben als Basis für den Verknüpfungs- und Selektionsprozess verwendet wurde. Es fällt auf, dass der Anteil der Männer unter den Verweigernden größer ist als unter den Zustimmenden. Das Durchschnittsalter zeigt keine Unterschiede zwischen den beiden Gruppen, während beim Bildungsniveau und beim Vorliegen eines Migrationshintergrunds geringe Unterschiede bestehen. Der von Beste (2011) gezeigte Effekt der höheren Zustimmungswahrscheinlichkeit bei höherem Einkommen tritt auch hier auf. Sowohl das Durchschnitts- als auch das Medianeinkommen liegt bei den Zustimmenden höher als bei den Verweigernden. Dies lässt sich auch anhand von Abbildung 3.1 erkennen. Dort sind die Verteilungen des im PASS angegebenen Bruttoeinkommens für die vorliegende Stichprobe dargestellt. Die Durchführung eines Kolmogorow-Smirnow-Test für die Gleichheit der Verteilungen ergibt, dass die Verteilungen sich nicht signifikant unterscheiden ($p > 0.05$).

Tabelle 3.3: Einverständnis zur Verknüpfung in der Beschäftigtenstichprobe

	Verknüpfungserlaubnis	Keine Verknüpfungserlaubnis
N	5312	452
Geschlecht:		
Weiblich	48,7%	42,2%
Männlich	51,3%	57,8%
Durchschnittsalter	43	43
Migration		
Ja	21,2%	25,6%
nein	76,8%	72,7%
kA	2,0%	1,7%
Bildungsabschluss		
Abitur	31,7%	32,6%
Realschule	40,4%	34,3%
Hauptschule	25,5%	29,5%
Kein Abschluss	2,2%	3,6%
kA/sonstige	0,2%	
Stichprobe		
BA	44,9%	53,2%
Microm	55,1%	46,8%
Durchschnittseinkommen (PASS)	2253	2098
Medianeinkommen (PASS)	1900	1800

Abbildung 3.1: Bruttoeinkommen im PASS und Zustimmung zur Zuspielung



Zusammenfassend kann festgestellt werden, dass sich die größten Selektivitäten, sei es durch das Erteilen bzw. Nicht-Erteilen der Verknüpfungserlaubnis, durch den Prozess der Verknüpfung an sich, oder durch die Auswahl der für die Analyse relevanten Personen, im Einkommen selbst zeigen. Dieser Umstand ist aufgrund der Fragestellung der Untersuchung unvermeidbar und muss daher bei der Interpretation der Ergebnisse berücksichtigt werden.

4 Methoden

Ziel dieser Untersuchung ist es, den Einfluss der Interviewer auf die Befragungsangabe bezüglich des Einkommens und damit den Messfehler zum tatsächlichen Einkommen zu bestimmen. Daher soll zunächst der Messfehler definiert werden. Der Notation von Kim und Tambourini (2014) folgend, wird davon ausgegangen, dass sich das in der Befragung angegebene Einkommen aus dem wahren Einkommen des Befragten und dem Messfehler ergibt, der aufgrund der in Kapitel 2 beschriebenen Ursachen entstehen kann:

$$y_i^{Befragung} = y_i + u_i, \quad (4.1)$$

wobei $y_i^{Befragung}$ das angegebene Einkommen einer Person i , y_i das tatsächliche Einkommen und u_i den Messfehler darstellen. Da in dieser Untersuchung - wie in einer Vielzahl anderer Studien auch - das Einkommen aus administrativen Daten als Referenz für das wahre Einkommen verwendet wird, ergibt sich im verknüpften Datensatz aus dem PASS und den IEB für den Messfehler:

$$u_i = y_i^{PASS} - y_i^{IEB} \quad (4.2)$$

Der Messfehler setzt sich also aus der Differenz aus dem im PASS angegebenen Einkommen y_i^{PASS} und dem in den IEB verzeichneten Einkommen y_i^{IEB} zusammen. Alle in dieser Untersuchung verwendeten Einkommensgrößen verstehen sich als monatliches Bruttoeinkommen in Euro. Unter Anwendung von (4.2) stellt der Messfehler also die absolute monatliche Differenz in Euro dar. Damit wird beispielsweise eine Befragungsangabe von 1100 Euro bei einem tatsächlichen Einkommen von 1000 Euro genauso bewertet wie eine Befragungsangabe

von 5100 Euro bei einem tatsächlichen Einkommen von 5000 Euro.

Der Messfehler als Hauptgegenstand der Untersuchung wird in allen Modellen als zu erklärende Variable verwendet. Im ersten Schritt wird zunächst ausschließlich der Einfluss soziodemographischer Variablen des Befragten sowie Merkmale der Interviewsituation (Interviewsprache und Befragungsmodus) auf den Messfehler mit Hilfe linearer Regression bestimmt. Das im weiteren beschriebene Modell weist folgende Struktur auf (Fahrmeier et al. 2007):

$$u_i = \alpha + \beta x_i + \epsilon_i \quad (4.3)$$

Der Messfehler eines Befragten i setzt sich also aus dem allgemeinen Intercept α , den Beiträgen des Kovariablenvektors x_i und seiner dazugehörigen Koeffizienten β , sowie aus dem zufälligen Residuum ϵ_i zusammen. Der Kovariablenvektor enthält alle Variablen, die für das jeweilige Modell als erklärende Variablen ausgewählt werden. Die Koeffizienten α und β werden im Statistik-Programm R mit Hilfe der Funktion lm bestimmt. Diese Funktion verwendet dazu die Methode der kleinsten Quadrate (OLS), die die Koeffizienten so wählt, dass die Quadratsumme der Residuen minimiert wird. Der Modellierung liegen folgende Annahmen zugrunde:

- Die Residuen ϵ_i stellen eine Zufallsvariable mit Erwartungswert Null dar: $E(\epsilon_i) = 0$
- Die Residuen sind zwischen zwei Personen i und j nicht korreliert: $Cov(\epsilon_i, \epsilon_j) = 0$
- Die Varianz der Residuen ist konstant (Homoskedastizität): $Var(\epsilon_i) = E(\epsilon_i^2) = \sigma^2$
- Es existieren keine exakten linearen Beziehungen zwischen den Kovariablen. Bei Vorliegen einer solchen Beziehung wären die Koeffizienten nicht eindeutig identifizierbar.

Daraufhin werden dem Modell Merkmale der Interviewer und Interaktionseffekte mit den Befragten als Kovariablen hinzugefügt. Es wird mittels eines Likelihood-Ratio Tests überprüft, ob dieses Modell gegenüber dem Basis-Modell eine signifikant höhere Erklärungskraft bezüglich des Messfehlers aufweist. Anschließend wird auf Basis des Modells mit Befragtenmerkmalen, Interviewermerkmalen und Interaktionseffekten eine Variablenselektion durchgeführt. Ziel ist es hierbei, das minimale Modell zu finden, das nur die Variablen enthält, die

einen signifikanten Beitrag zur Erklärung des Messfehlers leisten. Die Variablenselektion wird mit Hilfe der Funktion *stepAIC* mit der Option *backward* im Statistik-Programm *R* durchgeführt.¹ Dabei werden ausgehend vom vollen Modell schrittweise alle Variablen entfernt, die keinen signifikanten Beitrag zur Erklärung leisten. Kriterium für die Entfernung von Variablen aus dem Modell ist die Minimierung des Akaike-Informationskriterium (AIC), das auf der Loglikelihood basiert.

Die im ersten Schritt zur Erklärung der Messfehler verwendeten Variablen wurden auf Basis vorheriger Studien, sowie aufgrund eigener Überlegungen ausgewählt. Im Folgenden werden für die einzelnen Variablen die Motivation der Auswahl und der erwartete Einfluss auf den Messfehler dargelegt.

- **Administratives Einkommen:** Einige Studien zeigen, dass das Einkommen selbst einen großen Einflussfaktor auf die Messfehler beim Einkommen darstellt. Viele davon kommen zu dem Ergebnis, dass Personen mit hohem Einkommen in der Befragung einen zu niedrigen Wert und Personen mit niedrigem Einkommen einen zu hohen Wert angeben (siehe Kapitel 2, Mean-Reverting Measurement Error).
- **Alter der Befragten:** Gottschalk und Huynh (2005) zeigen, dass bei jüngeren Befragten stärkere Messfehler im Einkommen auftreten.
- **Geschlecht der Befragten:** Vorhergehende Untersuchungen zeigen einen Zusammenhang zwischen dem Geschlecht und den Messfehlern (Pedace und Bates 2000). Männliche Befragte scheinen in der Befragung öfter ein zu geringes Einkommen anzugeben.
- **Migrationshintergrund:** Kim und Tambourini (2014) zeigen in ihrer Untersuchung, dass Befragte mit Migrationshintergrund² häufiger ein zu geringes Einkommen angeben.
- **Schulbildung der Befragten:** Kim und Tambourini (2014) zeigen Effekte der Bildung auf das Einkommen, die sich jedoch nicht konstant über die Einkommensquantile verhalten.
- **Familienstand / Kinder**

¹Die Ergebnisse wurden durch Verwendung der Option *forward* überprüft. Die vorgeschlagenen Modelle unterschieden sich bei beiden Optionen nicht.

²Hier definiert als Geburt außerhalb der USA.

- Befristung im Arbeitsverhältnis / Jobwechsel / kurzes Beschäftigungsverhältnis: Kim und Tambourini (2014) zeigen, dass Beschäftigungsinstabilität einen signifikanten Einfluss auf den Messfehler hat, da diese zu häufigerer Unter- bzw. Übertreibung in der Befragungsangabe führt.
- Stellung im Beruf: Effekte für die Stellung im Beruf könnten über eine Korrelation mit dem Einkommen oder der Beschäftigungsstabilität zustande kommen. Hier ist daher auf Multikollinearität zu achten.
- Interviewsprache: Personen, deren Deutschkenntnisse zur Beantwortung und zum Verständnis der Fragen nicht ausreichen, könnten der Befragung skeptischer gegenüberstehen und zu stärker verzerrten Angaben neigen.
- Befragungsmodus: Holbrook et al. (2003) zeigen, dass Befragte bei telefonischen Interviews misstrauischer sind und häufiger sozial erwünschte Antworten geben als in persönlichen Interviews. Andererseits könnte die größere Interaktion in persönlichen Interviews zu größeren Interviewereffekten führen.
- Teilnahme des Befragten in Vorwelle/Panelbereitschaft: Die erneute Teilnahme bzw. die Bereitschaft dazu zeigt, dass der Befragte Vertrauen in die Befragung hat und bereit ist, wahrheitsgemäße Informationen zu geben. Die Messfehler könnten daher bei dieser Personengruppe geringer ausfallen.
- Gleiche Befragter-Interviewer-Konstellation in der Vorperiode: Wenn der Befragte und der Interviewer beide an der Vorwelle teilgenommen haben und der Befragte den Interviewer wiedererkennt, könnte das Vertrauen in den Interviewer und die Befragung insgesamt höher ausfallen. Da bei telefonischer Befragung ein Wiedererkennen wesentlich unwahrscheinlicher als bei persönlicher Befragung ist, wird diese Variable in Interaktion mit dem Befragungsmodus untersucht.
- Interviewer bereits in Vorperiode tätig: Durch die Teilnahme an der vorherigen PASS-Erhebungswelle, ist der Interviewer mit den meisten Fragen bereits vertraut und könnte eine Routine entwickelt haben.

- Schulbildung der Interviewer: Höhere Bildung des Interviewers könnte dazu führen, dass dieser dem Befragten gegenüber seriöser wirkt, und gleichzeitig über eine höhere Fähigkeit verfügt, Verständnisfragen des Befragten zufriedenstellend zu beantworten.
- Geschlecht des Interviewers: Es ist denkbar, dass sich das Geschlecht des Interviewers auf die Angabe sozial erwünschter Antworten auswirkt. Es wird daher auch die Interaktion mit dem Geschlecht des Befragten betrachtet.
- Alter des Interviewers: Ältere Interviewer verfügen möglicherweise über mehr Erfahrung und werden für seriöser gehalten. Daher könnte der Messfehler bei ihnen geringer ausfallen. Das Alter der Interviewer wird nicht nur isoliert, sondern auch in Interaktion mit dem Alter des Befragten betrachtet.

Im zweiten Schritt werden zusätzlich Interviewereffekte miteinbezogen. Unter der Annahme, dass Interviewereffekte existieren, sind die Daten durch eine hierarchische Struktur mit 2 Ebenen, auf denen die Messfehler entstehen können, gekennzeichnet. Die erste Ebene stellen dabei die Befragten selbst dar, die zweite Ebene die Interviewer. Für den Messfehler von Person i , die von Interviewer j befragt wurde, wird daher folgende Struktur angenommen:³

$$u_{ij} = \alpha + \beta x_{ij} + v_j + \epsilon_{ij} \quad (4.4)$$

Gegenüber Modell (4.3) aus dem ersten Schritt, ist hier nun zusätzlich ein interviewerspezifischer Effekt v_j enthalten. Eine einfache Möglichkeit der Bestimmung von v_j besteht in dessen direkter Aufnahme in die Schätzgleichung. So wird ein separater Achsenabschnitt für jeden Interviewer durch die lineare Regression im OLS-Verfahren als Fixed-Effect bestimmt. Da der hier vorliegende Datensatz jedoch 268 Interviewer umfasst, führt die Aufnahme der Interviewereffekte als Fixed-Effects dazu, dass 267 zusätzliche Koeffizienten geschätzt werden müssen. Zusätzlich zu den Voraussetzungen von Modell (4.4) darf das Residuum ϵ_{ij} nicht mit den Interviewereffekten v_j korreliert sein.

Eine weitere Möglichkeit besteht in der Auffassung von v_j als Zufallsvariable. Dann kann

³Für weitere Informationen zu Modellierung und Interpretation von Mehrebenenmodellen siehe Goldstein (1999)

der Effekt je Interviewer als zufälliger Achsenabschnitt (Random-Effect) aufgefasst werden. In diesem Fall wird im Vergleich zum Fixed-Effects-Modell kein separater Achsenabschnitt für jeden Interviewer, sondern die Varianz, die auf der Ebene der Interviewer entsteht, geschätzt. Damit stehen hier nicht die absoluten Effekte der einzelnen Interviewer im Fokus, sondern die Gesamtvarianz. In diesem Zusammenhang ist es daher möglich, die Intraklassenkorrelation (siehe Kapitel 2) als Anteil der durch die Interviewer auf der zweiten Ebene verursachten Varianz an der Gesamtvarianz (erste und zweite Ebene) zu bestimmen.

Die Annahmen im Random-Effects-Modell sind im Vergleich zum Fixed-Effects-Modell einschränkender, da v_j nicht mit dem Kovariablenvektor der Befragten x_{ij} korreliert sein darf. Das bedeutet, dass die Interviewer hinsichtlich der Merkmale der Befragten zufällig aus einer Gesamtpopulation aller Interviewer zugeteilt worden sein müssten. Dies ist in der Realität jedoch schwer umzusetzen. Auch im PASS wird eine zufällige Zuordnung aus verschiedenen Gründen nicht erreicht, beispielsweise wegen der Durchführung aller Befragungen in einer Region von einem einzigen Interviewer.

Um zu überprüfen, ob die Interviewereffekte trotzdem keine Korrelation mit dem Kovariablenvektor x_{ij} aufweisen, ob die zusätzliche Annahme des Random-Effects-Modells also erfüllt ist, wird ein Hausman-Test durchgeführt. Dieser vergleicht die für die erste Ebene bestimmten Koeffizienten der Befragtenmerkmale aus dem Fixed-Effects-Modell mit denen des Random-Effects-Modells. Falls die zusätzliche Annahme des Random-Effects-Modells verletzt wird, führt dies zu Verzerrungen und es ergibt sich ein Unterschied zu den Koeffizienten des Fixed-Effects-Modells. Die Nullhypothese des Hausman-Tests besteht darin, dass das Random-Effects-Modell und das Fixed-Effects-Modell keinen Unterschied in ihren geschätzten Koeffizienten aufweisen. Wird die Nullhypothese verworfen, bedeutet dies also, dass die zusätzliche Annahme des Random-Effects-Modell verletzt wird und dieses Modell daher verzerrte Ergebnisse liefert. Falls die Nullhypothese nicht verworfen werden muss, ist das Random-Effects-Modell zulässig und weist eine größere Effizienz als das Fixed-Effects-Modell auf.

Zusammenfassend lautet das Vorgehen wie folgt: Zunächst wird der Messfehler allein durch Befragten- und Interviewermerkmale, sowie deren Interaktionseffekte dargestellt und das mi-

nimale Modell mit der höchsten Erklärungskraft identifiziert. Anschließend werden auf dessen Basis die Interviewereffekte in Form von Fixed- und Random-Effects-Modellen miteinbezogen. Die Annahme der Unkorreliertheit zwischen Interviewereffekten und Kovariablen im Random-Effects-Modell wird überprüft, indem Fixed- und Random-Effects-Modelle gegeneinander getestet werden. Um die Signifikanz der Interviewereffekte zu überprüfen, werden Fixed- und Random-Effects-Modelle mit den linearen Modellen aus dem ersten Schritt verglichen.

5 Ergebnisse

In diesem Kapitel werden die Ergebnisse der Untersuchung vorgestellt und diskutiert. Vor der Anwendung der in Kapitel 4 vorgestellten Regressionsverfahren zur Quantifizierung der Interviewereffekte wird der vorliegende Datensatz hinsichtlich seiner Zusammensetzung aus Befragten und Interviewern und der Verteilung der Messfehler deskriptiv analysiert.

5.1 Deskriptive Analyse der Daten

Die Zusammensetzung des Analysedatensatzes wurde bereits in Abschnitt 3.4.2 (siehe Tabelle 3.2, Spalte 3) anhand der wichtigsten Merkmale der Befragten thematisiert. Daneben enthält der Analysedatensatz eine Reihe von weiteren Variablen, die in der Regressionsanalyse als Einflussfaktoren auf den Messfehler des Einkommens verwendet werden. Deren Ausprägungen sind in Tabelle 5.1 zusammengefasst.

Die Beschäftigungsverhältnisse sind bei einem sehr großen Anteil der Befragten stabil. Die drei Indikatoren für unsichere Beschäftigungsverhältnisse (Befristung, Beschäftigungsdauer weniger als ein Monat und Jobwechsel) weisen nur eine geringe Häufigkeit auf. Die Interviews wurden mehrheitlich telefonisch durchgeführt, die Sprache des Interviews war abgesehen von wenigen Einzelfällen Deutsch. Beinahe alle Befragten zeigten sich zu einer erneuten Befragung in der nächsten Erhebungswelle bereit, der Anteil der Wiederholungsteilnehmer (Teilnahme in mindestens einer der ersten vier Erhebungswellen) liegt bei 74%. Derselbe Interviewer wie in der vierten Erhebungswelle wurde knapp 16% der Befragten zugewiesen.¹

¹Aufgrund des Wechsels des Erhebungsinstitutes nach der dritten Erhebungswelle, kann dieselbe Befragten-Interviewer-Konstellation wie in der fünften Erhebungswelle nur bei Teilnehmern der vierten Welle bestehen.

Tabelle 5.1: Weitere Merkmale der Beschäftigtenstichprobe

Befristung im Arbeitsverhältnis		Kurze Erwerbstätigkeit (< 30 Tage)	
Ja	17.1%	Ja	0.2%
Nein	82.9%	Nein	99.8%
Stellung im Beruf		Jobwechsel in den letzten 3 Monaten	
Arbeiter	30.0%	Ja	8.4%
Angestellter	70.0%	Nein	91.6%
Familienstand		Kinder	
Alleinstehend	51.6%	Ja	70.2%
Verheiratet/eingetragene Partnerschaft	48.4%	Nein	29.8%
Befragungsmodus		Interviewsprache	
CATI	62.0%	Deutsch	98.7%
CAPI	38.0%	Russisch	1.1%
		Türkisch	0.2%
Teilnahme am PASS in vorheriger Welle		Panelbereitschaft	
Ja	74.2%	Ja	98.8%
nein	25.8%	Nein	1.2%
Gleicher Interviewer in der letzten Welle			
Ja	15.6%		
Nein	84.4%		

Abbildung 5.1: Monatliches Bruttoeinkommen im PASS und den IEB

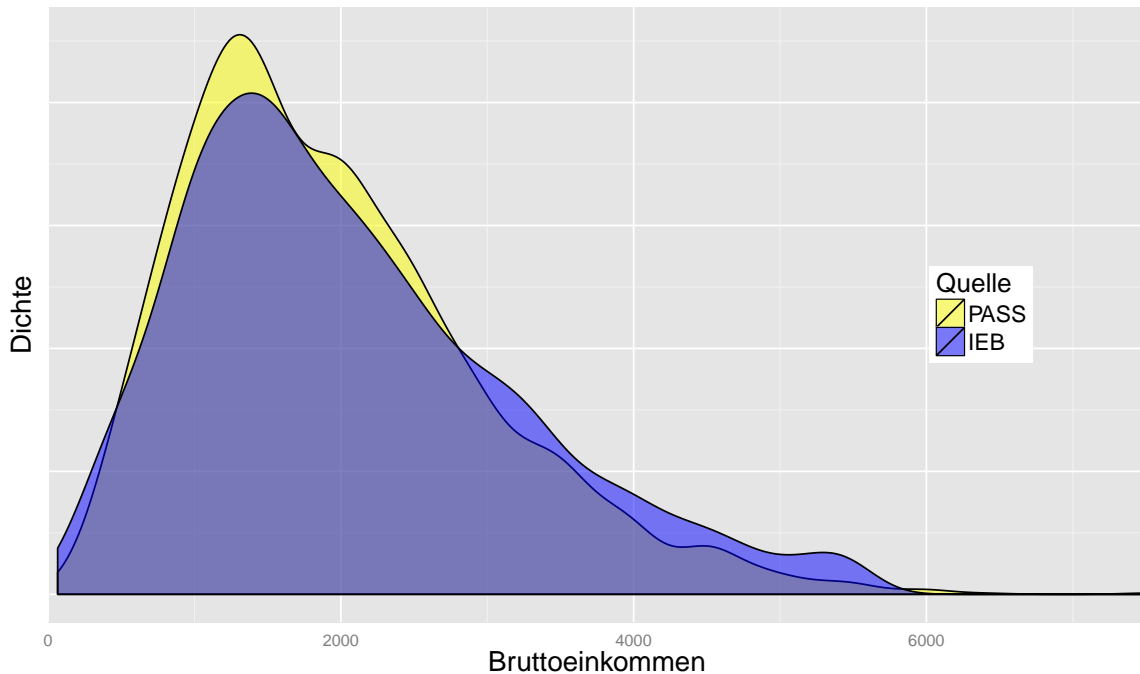
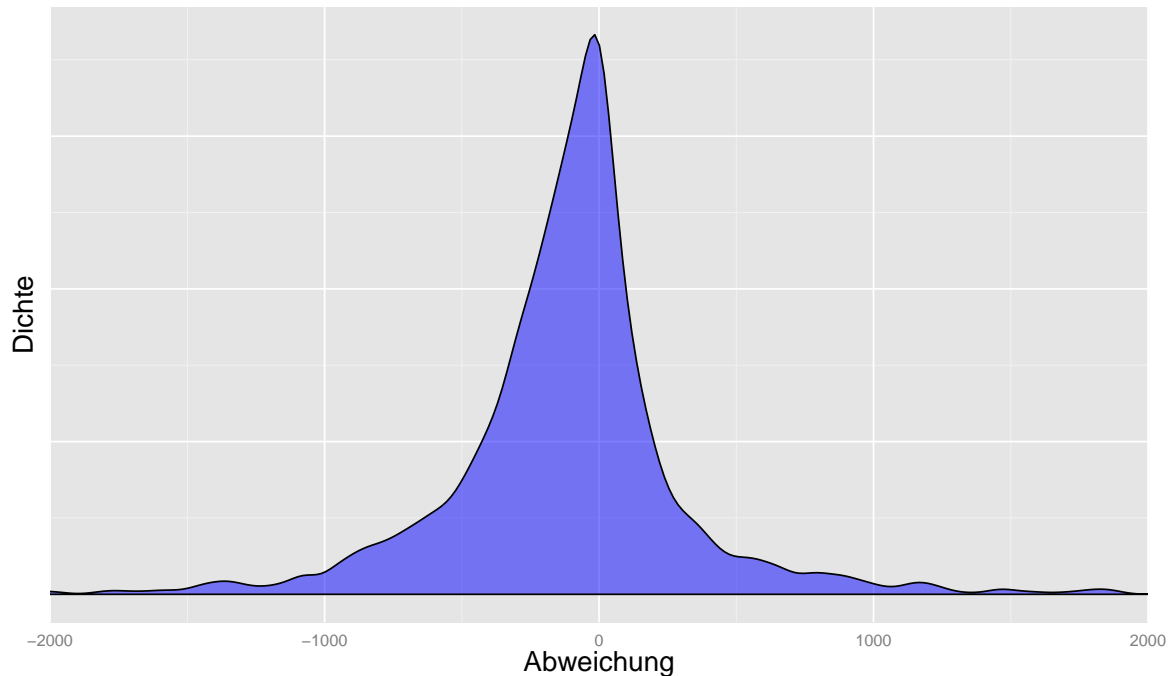


Abbildung 5.1 zeigt die Verteilungen von Befragungsangaben und administrativen Werten des Bruttoeinkommens in Gegenüberstellung. Hier lässt sich die Auswirkung der Bereinigung der Stichprobe um Personen, deren administratives Einkommen über der Beitragsbemessungsgrenze liegt, erkennen. Oberhalb der Beitragsbemessungsgrenze treten deshalb nur vereinzelt Werte im PASS auf, die unter der Annahme der IEB als Referenz per Definition mindestens den Abstand zur Bemessungsgrenze als Messfehler beinhalten.² Im Einkommensbereich unterhalb der Bemessungsgrenze fällt auf, dass die IEB, abgesehen von einem leichten Überhang bei den Einkommen unter 500 Euro, einen größeren Anteil im höheren Einkommensbereich (zwischen 2800 und 5500 Euro) aufweisen. Im Gegensatz dazu sind im PASS mehr Werte im Bereich zwischen 500 und 2800 Euro zu finden.

Dementsprechend ist die Differenz zwischen dem PASS- und dem IEB-Einkommen im Mittel über alle Befragten negativ. Dies unterstreicht auch die Verteilung der Messfehler in Abbildung 5.2. Im Gegensatz zu anderen Untersuchungen des Messfehlers im Einkommen (z.B.

²Die Graphik wurde am rechten Rand bei 7500 Euro abgeschnitten. Über dieser Grenze liegen nur wenige Werte, der Maximalwert unter den Befragungsangaben liegt bei 9000 Euro.

Abbildung 5.2: Verteilung des Messfehlers



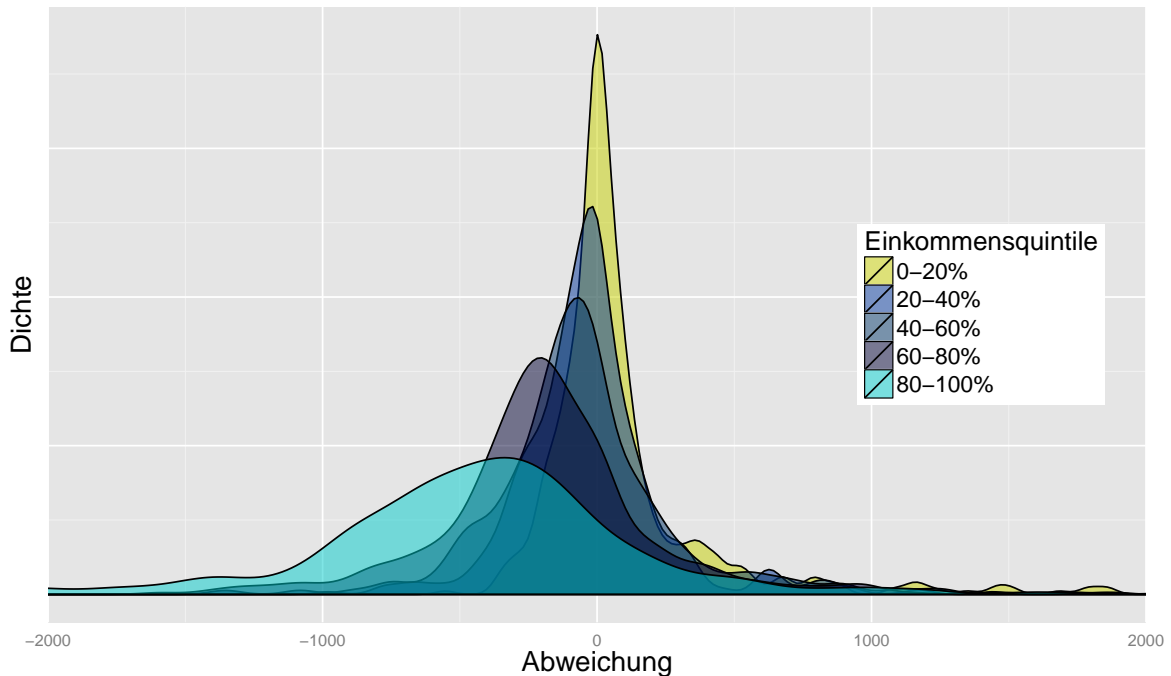
Bound und Krueger 1991) ist dieser im vorliegenden Datensatz nicht um den Wert 0 zentriert. Arithmetisches Mittel und Median (siehe Tabelle 3.2) sind jeweils im negativen Bereich angesiedelt.

Einen Anhaltspunkt dafür, ob es sich bei diesem Messfehler um einen Mean-Reverting Measurement Error handelt, kann ein graphischer Vergleich hinsichtlich von Einkommensquantilen geben. Dazu wurden die Befragten anhand ihres Einkommens in den administrativen Daten in 5 Einkommensgruppen (Quintile) unterteilt.³ Abbildung 5.3 zeigt die Verteilungen der Messfehler in den Einkommensquantilen.

Zunächst fällt auf, dass sich der Messfehler mit steigendem Einkommen immer weiter in den negativen Bereich bewegt. Dieses wichtige Ergebnis vieler bisheriger Studien kann auf Basis dieser Daten daher bestätigt werden. Auch hier lassen sich Anzeichen für das Übertreiben von niedrigen Einkommensbeziehern in der Befragung (gleichbedeutend mit einem großen positiven Messfehler), wie beispielsweise in der Untersuchung von Bound und Krueger (1991),

³Quintil 1: 0 bis 1092 Euro monatliches Bruttoeinkommen; Quintil 2: 1092 bis 1573 Euro; Quintil 3: 1573 bis 2154 Euro; Quintil 4: 2154 bis 3022 Euro; Quintil 5: 3022 bis 5500 Euro

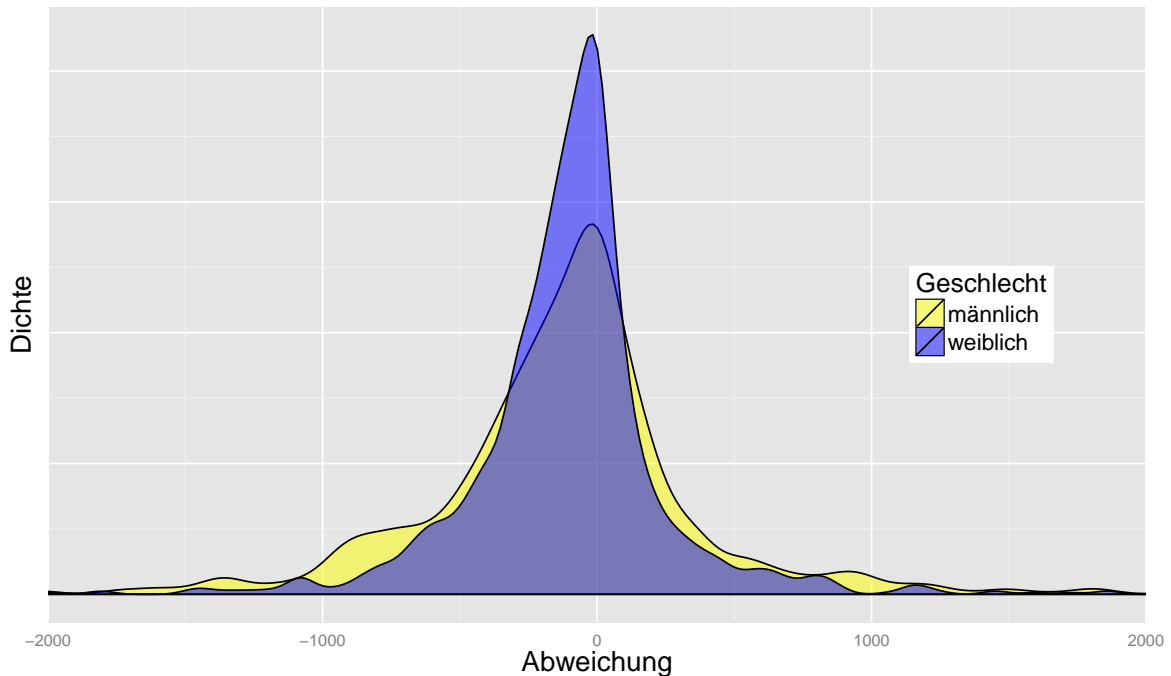
Abbildung 5.3: Verteilung des Messfehlers nach Einkommensquintilen



erkennen. Das erste Quintil weist (als Einziges) einen positiven Mittelwert im Messfehler auf, dieser liegt bei 96 Euro. Bemerkenswert ist auch die Tatsache, dass sich die Varianz der Messfehler von den niedrigeren zu den höheren Einkommensgruppen vergrößert. Dafür ist die Tatsache verantwortlich, dass die Varianz der Befragungsangaben mit zunehmendem Einkommen steigt.

Aufmerksamkeit verdient auch die Unterscheidung des Messfehlers hinsichtlich des Geschlechts der Befragten (siehe Abbildung 5.4). Der Messfehler für männliche Befragte weist sowohl im negativen, als auch im positiven Bereich eine größere Varianz auf. Zur Überprüfung, ob die Verteilungen von Männern und Frauen signifikant voneinander abweichen, wurde ein Kolmogorow-Smirnow-Test durchgeführt. Aufgrund des Testergebnisses ($p < 0.05$) kann davon ausgegangen werden, dass sich die Verteilungen der Messfehler im Einkommen für die beiden Geschlechter signifikant unterscheiden. Ähnliche Vergleiche der Messfehler wurden außerdem hinsichtlich der Variablen Stichprobe (BA bzw. Microm), Befragungsmodus und Wiederholungsteilnahme durchgeführt (siehe Abbildung A.1, A.2, A.3 im Anhang). Für alle diese Variablen ergibt der Kolmogorow-Smirnow-Test keinen signifikanten Unterschied in den

Abbildung 5.4: Verteilung des Messfehlers nach Geschlecht



Verteilungen des Messfehlers.

Außerdem soll noch ein Blick auf die Zusammensetzung der Interviewer geworfen werden. Tabelle 5.2 gibt einen Überblick über die Merkmale des Teils der Interviewer aus der fünften Erhebungswelle, der die Personen in der Analyse-Stichprobe befragt hat. Die Interviewer sind im Mittel etwas älter als die Befragten, das Geschlechterverhältnis ist dem der Befragten sehr ähnlich. Mit einer Abiturquote von 69% weisen die Interviewer im Vergleich zu den Befragten deutlich öfter eine höhere Schulbildung auf. Bezüglich der Interviewerfahrung liegen nur die Anzahl der Arbeitsjahre beim Erhebungsinstitut infas vor. Im Mittel sind die Interviewer dort seit knapp vier Jahren beschäftigt, nur ein geringer Teil ist seit weniger als einem Jahr beschäftigt. Mehr als die Hälfte der in der fünften Erhebungswelle eingesetzten Interviewer haben bereits an der vierten Erhebungswelle teilgenommen. Die Bandbreite der durchgeführten Interviews je Interviewer in der gesamten fünften Erhebungswelle bewegt sich zwischen einem Interview als Minimum und 178 Interviews als Maximum. Wie in Abschnitt 3.3 dargestellt, beschränkt sich diese Analyse jedoch nur auf Interviewer, die mindestens sechs Interviews im Analysedatensatz durchgeführt haben.

Tabelle 5.2: Zusammensetzung der Interviewer in der Analytestichprobe

Geschlecht		Durchschnittsalter	47
Männlich	51.3%	Minimum	19
Weiblich	48.7%	Maximum	81
Höchster Schulabschluss		Durchschnittliche	
Abitur	69.0%	Interviewerfahrung (Jahre)	3.75
Realschule	20.9%	Interviewerfahrung >= 1 Jahr	98.1%
Hauptschule	10.1%	Interviewerfahrung >= 5 Jahre	30.9%
Teilnahme an der Vorwelle		Durchschnittliche Anzahl Interviews Welle 5	49
Ja	53%	Minimum	16
Nein	47%	Maximum	178

5.2 Regressionsanalysen

In diesem Abschnitt werden die Ergebnisse der in Kapitel 4 beschriebenen Regressionsmodelle zusammengefasst und interpretiert. Im ersten Schritt wurden für den Messfehler als abhängige Variable mithilfe der Funktion *lm* im Statistik-Programm *R* drei lineare Modelle aufgestellt. Die Ergebnisse der Parameterschätzungen und das unkorrigierte bzw. korrigierte R^2 als Maß der Erklärungskraft der drei Modelle werden in Tabelle 5.3 zusammengefasst. Modell 1 stellt dabei das Basis-Modell mit Befragtenmerkmalen als Kovariablen dar. In Modell 2 wurden zusätzlich Merkmale des Interviewers, der Interviewsituation und Interaktionseffekte aufgenommen. Ausgehend von Modell 2 stellt Modell 3 das minimale nach Variablenselektion vorgeschlagene Modell dar.

Bei der Betrachtung der Ergebnisse von Modell 1 fällt als Erstes auf, dass die wenigsten der erklärenden Variablen einen signifikanten Effekt zeigen. Eine der signifikanten Variablen ist das administrative Einkommen an sich, das einen negativen Koeffizienten aufweist. Das bedeutet, dass die Höhe des Messfehlers im Einkommen vom Einkommen selbst abhängt und es mit steigendem Einkommen zu einer immer stärkeren Untertreibung durch die Befragten kommt. Damit bestätigt sich die Vermutung, die sich im vorherigen Abschnitt schon aus der graphischen Analyse der Einkommensquintile (siehe Abbildung 5.3) ergab. Zusätzlich zum linearen Effekt des Einkommens wurde auch ein quadratischer Term zur Berücksichtigung nicht-linearer Effekte ins Modell aufgenommen. Dieser zeigt sich zwar signifikant, übt jedoch aufgrund seines im Vergleich zum linearen Term betragsmäßig sehr geringen Koeffizienten keinen großen Einfluss auf den Messfehler aus. Weitere signifikante Effekte zeigen das Geschlecht der Befragten (Frauen untertreiben ihr Einkommen im Vergleich zu Männern), das Vorliegen eines Migrationshintergrunds (Personen mit Migrationshintergrund übertreiben eher) und die Stellung im Beruf (Angestellte übertreiben eher als die Referenzkategorie der Arbeiter).

In Modell 2 wurden zusätzlich persönliche Merkmale, Erfahrung und Routine des Interviewers, sowie die Interviewsituation mit Befragungsmodus und Befragungssprache als erklärende Variablen ins lineare Modell aufgenommen. Außerdem wurden Interaktionseffekte zwischen Interviewer und Befragtem für Alter, Geschlecht und eine mögliche Bekanntschaft des In-

Tabelle 5.3: Ergebnisse der linearen Regressionsmodelle

Variable	Parameterschätzung		
	Modell 1	Modell 2	Modell 3
Intercept	293.4***	271.3**	282.5***
Einkommen (IEB)	-0.250***	-2.538***	-0.248***
Quadrat des Einkommens (IEB)	0.00001*	0.00001*	0.00001*
Geschlecht: Weiblich	-135.2***	-126.8***	-134.8***
Alter	0.672	-7.673	
Migrationshintergrund	52.02*	56.12**	52.79**
Schulbildung			
Hauptschulabschluss	24.45	22.80	25.58
Mittlere Reife	51.80	48.19	51.88
Abitur	106.8	102.9	108.5
Familienstand: Verheiratet	--8.376	-7.585	
Kinder	-23.59	-22.39	
Keine Panelbereitschaft	-116.0	-122.8	-112.5
Teilnahme in Vorwelle	-9.706	-1.222	
Befristetes Arbeitsverhältnis	-29.13	-32.45	-34.50
Jobwechsel in den letzten 3 Monaten	-27.57	-29.64	
Kurze Erwerbstätigkeit (< 30 Tage)	101.9	92.83	
Stellung im Beruf: Angestellter	52.06**	55.44**	53.93**
Befragungsmodus: CAPI		-12.74	
Interviewsprache			
Russisch		174.2	
Türkisch		14.97	
Geschlecht des Interviewers: Weiblich		35.29	24.29
Alter des Interviewers		8.173	
Schulbildung des Interviewers			
Mittlere Reife		47.30	
Abitur		24.26	
Teilnahme des Interviewers an Vorwelle		21.21	
Anzahl der Interviews (Welle 5)		-21.08	
Interviewererfahrung (Jahre)		-0.450	
Interviewer bekannt: CAPI		-24.36	
Interviewer bekannt: CATI		-95.57	
Geschlechtsinteraktion		-19.54	
Altersdifferenz Interviewer-Befragter		8.228	
R ² / korrigiertes R ²	17.3%/16.8%	17.7%/16.5%	17.3%/16.9%

Signifikanz: p < 0.05 ***, p < 0.01 ***, p < 0.001 ****

interviewers aus der Vorwelle aufgenommen. Keine dieser neu hinzugefügten Variablen zeigt sich jedoch signifikant. In den übrigen Variablen ergeben sich nur wenige Unterschiede. Nach wie vor haben das Einkommen selbst, Geschlecht, Migrationshintergrund und die Stellung im Beruf einen signifikanten Einfluss auf den Messfehler.

Um zu überprüfen, ob Modell 2 als Ganzes trotzdem über mehr Aussagekraft verfügt als Modell 1, wurde ein Likelihood-Ratio-Test durchgeführt. Es erstaunt wenig, dass auch dieser die Wahl von Modell 1 nahe legt ($p > 0.05$). Das korrigierte R^2 für Modell 2 sinkt marginal im Vergleich zu Modell 1, da die Bestrafung für die zusätzlich zu schätzenden Parameter stärker ins Gewicht fällt als deren zusätzliche Erklärungskraft.

Anschließend wurde auf Basis von Modell 2 durch einen Variablenselektionsmechanismus dasjenige Modell 3 für den Messfehler gesucht, in dem möglichst wenige Parameter bestimmt werden müssen, und welches gleichzeitig durch die Aufnahme weiterer Variablen nicht mehr signifikant verbessert werden kann. Dazu wurde die im Statistik-Programm *R* implementierte Funktion *stepAIC* verwendet. Außer den schon in Modell 1 und 2 signifikanten Variablen, wurden lediglich die Schulbildung der Befragten, befristetes Arbeitsverhältnis, Verweigerung einer wiederholten Befragung und das Geschlecht des Interviewers beibehalten, wobei diese vier Variablen isoliert betrachtet keinen signifikanten Einfluss auf den Messfehler im Einkommen aufweisen.

Nachdem in den Modellen 1 bis 3 gezeigt wurde, dass einzelnen Variablen der Interviewer kein signifikanter Einfluss auf den Messfehler zukommt, wurden die Interviewereffekte im zweiten Schritt mit der in Kapitel 4 beschriebenen Mehrebenenstruktur von Befragten (erste Ebene) und Interviewern (zweite Ebene) durch Fixed- und Random-Effects-Modelle analysiert. Als erklärende Variablen auf der ersten Ebene kam in beiden Fällen die Auswahl aus Modell 3 zum Einsatz. So kann durch den Vergleich der beiden Modelle mit dem linearen Modell 3 überprüft werden, ob die Mehrebenenstruktur für den Messfehler relevant ist. Tabelle 5.4 zeigt die Parameterschätzungen für das Fixed-Effects-Modell (Modell 4) und das Random-Effects-Modell (Modell 5), zusätzlich ist auch Modell 3 nochmals zum Vergleich dargestellt.

Wie in Kapitel 4 beschrieben basiert Modell 4 auf der gewöhnlichen linearen Regression,

Tabelle 5.4: Ergebnisse des Fixed- und des Random Effects-Modells

	Modell 3 Lineares Modell	Modell 4 (Fixed Effects-Modell)	Modell 5 (Random Effects-Modell)
Intercept	282.5***	590.0***	279.9
Monatliches IEB-Einkommen	-0.248***	-0.253***	-0.249
Quadrat des monatl. IEB-Einkommens	0.00001*	0.00001*	0.00001*
Geschlecht: weiblich	-134.8***	-138.3***	-135.2
Migrationshintergrund	52.79**	54.95*	60.31**
Schulbildung			
Hauptschulabschluss	25.58	47.47	31.46
Mittlere Reife	51.88	72.44	55.79
Abitur	108.5	136.4*	113.9
Stellung im Beruf: Angestellter	53.93**	50.36*	52.44
Keine Panelbereitschaft	-112.5	-68.23	-109.1
Befristetes Arbeitsverhältnis	-34.50	-36.34	-35.68
Geschlecht des Interviewers: Weiblich	24.29	-304.2	25.40
R ² / korrigiertes R ²	17.3%/16.9%	25.8%/19.1%	
Intraklassenkorrelation (ICC)			2.8%

Signifikanz: $p < 0.05$ '*', $p < 0.01$ '***', $p < 0.001$ '****'

wobei für jeden Interviewer ein separater Achsenabschnitt geschätzt wird. Bei 268 Interviewern in der Stichprobe müssen daher 267 zusätzliche Parameter gegenüber Modell 3 bestimmt werden. Der Übersichtlichkeit halber und auch weil die absolute Höhe der Achsenabschnitte der Interviewer für die Interpretation der Interviewereffekte im Allgemeinen nicht relevant ist, werden die separaten Achsenabschnitte nicht in Tabelle 5.4 aufgeführt. Für Modell 4 wird genau wie für Modell 3 das unkorrigierte bzw. korrigierte R^2 angegeben. Hier zeigt sich, dass durch die Aufnahme der 267 Achsenabschnitte das unkorrigierte R^2 stark steigt (von 17.3% auf 25.8%), das korrigierte R^2 jedoch aufgrund der Bestrafung für die zusätzliche Komplexität des Modells nur unterproportional steigt (von 16.9% auf 19.1%). Um beurteilen zu können, ob Modell 4 gegenüber Modell 3 vorzuziehen ist, wurde ein Likelihood-Ratio-Test durchgeführt. Das Ergebnis ($p < 0.05$) zeigt, dass die Modellierung mit Fixed-Effects insgesamt über eine signifikant größere Erklärungskraft als das rein lineare Modell verfügt.

Das Random-Effects-Modell (Modell 5) hingegen fasst, wie in Kapitel 4 beschrieben, die Heterogenität durch die Interviewer als Zufallsvariable für jeden Interviewer auf. Die Schätzung von 267 interviewerspezifischen Achsenabschnitten ist daher nicht mehr nötig. Es wird lediglich der Varianzbeitrag der Interviewer zur Gesamtvarianz, also die Intraklassenkorrelation bestimmt. Die Berechnung erfolgte im Statistik-Programm *R* mit Hilfe der Funktion *lmer*, die dem Prinzip der Maximum-Likelihood-Schätzung folgt. Auch dieses Modell wurde gegen das rein lineare Modell 3 getestet. Dazu wurde ein simulationsbasierter Likelihood-Ratio-Test verwendet, wie er im Paket *RLRsim* in *R* implementiert ist (Scheipl et al. 2007). Das Testergebnis ($p < 0.05$) zeigt, dass auch das Random-Effects-Modell eine höhere Erklärungskraft für den Messfehler zeigt als Modell 3.

Zuletzt wurde außerdem getestet, ob die zusätzliche Annahme des Random-Effects-Modells, also dass die Random-Effects mit den erklärenden Variablen der ersten Ebene unkorreliert sind, verletzt wird. Wie bereits ein kurzer Blick auf die Parameterschätzungen von Modell 4 und Modell 5 nahe legt, unterscheiden sich die Koeffizienten der meisten Variablen nicht substantiell, insbesondere bei den Variablen, die einen signifikanten Einfluss auf den Messfehler haben. Die Nullhypothese des verwendeten Hausman-Tests, dass keine Unterschiede zwischen den Parameterschätzungen der beiden Modelle 4 und 5 bestehen, wird nicht abgelehnt

($p > 0.05$). Das Random-Effects-Modell ist also gültig und effizienter als die Fixed-Effects-Modellierung.

Zusammenfassend lässt sich durch den Vergleich der Modelle 1 bis 5 feststellen, dass die Mehrebenenstruktur für den Messfehler relevant ist. Die zur Verfügung stehenden Variablen über die Interviewer tragen kaum zur Erklärung des Messfehlers bei (siehe Modell 2), es handelt sich also um unbeobachtete Heterogenität. Die Likelihood-Ratio-Tests zeigen, dass die Fixed-Effects- und Random-Effects-Modellierung signifikant zur Erklärung der Gesamtvarianz beiträgt. Die Random-Effects-Modellierung ist aufgrund der Gültigkeit ihrer Annahmen anwendbar und der weniger effizienten Fixed-Effects-Modellierung vorzuziehen. Mit einer Intraklassenkorrelation von 2.8% fallen die Effekte trotzdem eher gering aus. Im Vergleich zur Frage nach der Zuspiegelungserlaubnis von administrativen Daten, bei der aktuelle Studien ICC-Werte zwischen 28% (Beste 2011) und 41% (Korbmacher und Schröder 2013) nachweisen, ist die Befragungsangabe zum Einkommen daher deutlich weniger von Interviewereffekten beeinflusst.

Bezüglich des Einflusses der soziodemographischen Merkmale der Befragten auf den Messfehler zeigt sich, dass die Koeffizienten der signifikanten Variablen trotz unterschiedlicher Modellierungen mit oder ohne Einbeziehung der Interviewereffekte kaum voneinander abweichen. Für das Einkommen selbst kann ein negativer Zusammenhang mit dem Messfehler gezeigt werden. Diese Beobachtung stimmt mit den Ergebnissen des Großteils der Literatur zur Untersuchung der Messfehler im Einkommen (z.B. Bound und Krüger 1991; Kim und Tambourini 2014) überein. Außer für das Einkommen selbst, konnten in allen Modellen signifikante Untertreibungen in der Einkommensangabe bei Frauen, sowie signifikante Übertreibungen bei Befragten mit Migrationshintergrund und Angestellten (im Vergleich zu Arbeitern) festgestellt werden.

6 Zusammenfassung

Ziel dieser Untersuchung war es, den Einfluss der Interviewer auf die Messfehler bei der Angabe des Einkommens in Befragungen zu bestimmen. Dies war unter der Annahme möglich, dass administrative Daten als Referenz für die Bestimmung des Messfehlers dienen können. So konnte aus Befragungsdaten, Daten zu den durchführenden Interviewern und administrativen Daten von hoher Qualität ein verknüpfter Datensatz zur Analyse erstellt werden. In einem zweistufigen Prozess wurde der Messfehler in der Befragungsangabe zum Einkommen zunächst durch erhobene Merkmale der Befragten und Interviewer und schließlich durch Interviewereffekte mit Hilfe einer Mehrebenenanalyse modelliert.

Die Ergebnisse zeigen, dass signifikante Interviewereffekte auftreten. Die im Rahmen der Befragung zusätzlich gesammelten Interviewermerkmale tragen nicht zur Erklärung des Messfehlers bei. Die Mehrebenenstruktur ist für die Zusammensetzung des Messfehlers jedoch relevant. Die Intraklassenkorrelation als Maß der unbeobachteten Heterogenität fällt mit 2.8% trotzdem eher gering aus. Aktuelle Studien zu Interviewereffekten bei der Zustimmung der Befragten zur Zuspiegelung ihrer administrativen Daten zeigen deutlich höhere Intraklassenkorrelationen mit Werten, die teilweise bis zu 40% erreichen. Es ist deshalb davon auszugehen, dass der Befragte bei der Angabe seines Einkommens vom Interviewer weniger beeinflusst wird, als bei anderen Entscheidungen in Befragungen.

Außerdem konnten die Ergebnisse eines Großteils der Studien, die bisher die Messfehler im Einkommen durch die Verknüpfung von Befragungsdaten und administrativen Daten analysiert hatten, bestätigt werden. Auch in der vorliegenden Untersuchung kam es in höheren Einkommensbereichen dazu, dass die befragten Personen weniger als ihr tatsächliches Einkommen angaben. Demgegenüber waren die Angaben von Beziehern niedrigerer Einkommen

tendenziell genauer, im untersten Einkommensbereich wurden im Mittel leichte Übertreibungen des Einkommens festgestellt. Des Weiteren zeigten sich im Messfehler Effekte des Geschlechts, beim Vorliegen eines Migrationshintergrunds oder in der Stellung im Beruf.

Die Ergebnisse sind jedoch mit Vorsicht zu interpretieren. So basiert die Modellierung des Messfehlers auf der Annahme, dass administrative Daten als Referenz für den wahren Wert des Einkommens der Befragten verwendet werden können. Zwar gelten die verwendeten administrativen Daten der Bundesagentur für Arbeit insbesondere für das Einkommen als sehr zuverlässig, Fehler können jedoch nie ganz ausgeschlossen werden. Für die Untersuchung der Interviewereffekte stand zudem nur eine begrenzte Anzahl von Eigenschaften der Interviewer zur Verfügung. Einstellungen des Interviewers oder Merkmale der Interviewsituation wie Antwortzeiten des Befragten, Zahl der gestellten Gegenfragen oder Gesamtdauer des Interviews, wie sie in anderen Studien zu Interviewereffekten verwendet werden konnten.

Literaturverzeichnis

Antoni, Manfred; Bethmann, Arne: PASS-Befragungsdaten verknüpft mit administrativen Daten des IAB 1975–2010 (PASS-ADIAB 7510). FDZ-Datenreport 03/2014, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Berg, Marco et. al.: Codebuch und Dokumentation des Panel Arbeitsmarkt und soziale Sicherung (PASS). FDZ-Datenreport 06/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Beste, Jonas: Selektivitätsprozesse bei der Verknüpfung von Befragungs- mit Prozessdaten. FDZ-Methodenreport 09/2011, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Bethmann, Arne; Fuchs, Benjamin; Wurdack, Anja: User Guide Panel Study Labour Market and Social Security“ (PASS): Wave 6. FDZ-Datenreport 07/2013, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Bollinger, C. R. (1998): Measurement Error in the Current Population Survey: A Nonparametric Look. *Journal of Labor Economics*, 16, 576-94.

Bound, J., Brown C., Duncan G. J. und Rodgers W. L. (1994): Evidence on the Validity of Cross-Sectional and Longitudinal Labor Market Data. *Journal of Labor Economics*, 12, 345-68.

Bound, J. und Krueger A. B. (1991): The Extent of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs Make a Right?. *Journal of Labor Economics*, 9, 1-24.

Bricker, J. und Engelhardt G. V. (2008): Measurement Error in Earnings Data in the Health and Retirement Study. *Journal of Economic and Social Measurement*, 33, 39-61.

Duncan, G. J. und Hill D. H. (1985): An Investigation of the Extent and Consequences of Measurement Error in Labor-Economic Survey Data. *Journal of Labor Economics*, 3, 508-32.

Fahrmeir, L., Kneib, T., und Lang, S.M. (2007). *Regression*. Springer DE.

Goldstein, H. (1999): *Multilevel Statistical Models*. Third Edition.

Gottschalk, P., und Huynh M. (2005): Validation Study of Earnings Data in the SIPP—Do Older Workers Have Larger Measurement Error? CRR Working Paper 2005-07, Center for Retirement Research at Boston College, Chestnut Hill, MA.

Gottschalk, P., und Huynh M. (2010): Are Earnings Inequality and Mobility Overstated? The Impact of Nonclassical Measurement Error. *The Review of Economics and Statistics*, 92, 302-15.

Holbrook, A.L., Green, M.C., and Krosnick, J.A. (2003): Telephone Versus Face-to-Face Interviewing of National Probability Samples with Long Questionnaires: Comparisons of Respondent Satisficing and Social Desirability Bias. *Public Opinion Quarterly*, 67, 79–125.

Jacobebbinghaus, P., und Seth, S. (2007): The german integrated employment biographies sample IEBS. *Schmollers Jahrbuch. Zeitschrift für Wirtschafts- und Sozialwissenschaften*, 127, 335- 342.

Jesske, Birgit; Schulz, Sabine: Methodenbericht Panel Arbeitsmarkt und Soziale Sicherung PASS: 5. Erhebungswelle – 2011. FDZ-Methodenreport 11/2012, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

Kapteyn, A. und Ypma J. (2007): Measurement Error and Misclassification: A Comparison of Survey and Administrative Data. *Journal of Labor Economics*, 25, 513-551.

Kim, C., und Tamborini, C. R. (2014): Response Error in Earnings: An Analysis of the Survey of Income and Program Participation Matched With Administrative Data. *Sociological Methods & Research*, 43, 39-72.

Kish, L. (1962). Studies of Interviewer Variance for Attitudinal Variables. *Journal of the American Statistical Association*, 57, 92–115.

Korbmacher J. M., und Schroeder M. (2013): Consent when Linking Survey Data with Administrative Records: The Role of the Interviewer. *Survey Research Methods*, 7(2), 115-131.

Lipps, O. und Pollien, A. (2011): Effects of Interviewer Experience on Components of Non-

response in the European Social Survey. *Field Methods*, 23(2), 156-172.

Mangione, T. W., Fowler, F. J., und Louis, T. A. (1992). Question Characteristics and Interviewer Effects. *Journal of Official Statistics*, 8, 293–307.

Moore, J. C., Stinson, L. L. und Welniak E. J. (2000): Income Measurement Error in Surveys: A Review. *Journal of Official Statistics*, 16, 331-61.

O’Muircheartaigh, C. und Campanelli, P. (1999): A Multilevel Exploration of the Role of Interviewers in Survey Non-Response. *Journal of the Royal Statistical Society, Series A*, 162(Part 3), 437–446.

Pedace, R. und Bates N. (2000): Using Administrative Data to Assess Earnings Reporting Error in the Survey of Income and Program Participation. *Journal of Economic and Social Measurement*, 26, 173-92.

Sakshaug, J. W., Couper M. P., Ofstedal M. B., und Weir D. R. (2012): Linking Survey and Administrative Records: Mechanisms of Consent. *Sociological Methods & Research*, 41(4), 535-569.

Sakshaug, J. W., und Kreuter F. (2012): Assessing the Magnitude of Non-Consent Biases in Linked Survey and Administrative Data. *Survey Research Methods*, 6(2), 113-122.

Sala, E., Burton J., und Knies G. (2012): Correlates of Obtaining Informed Consent to Data Linkage: Respondent, Interview, and Interviewer Characteristics. *Sociological Methods & Research*, 41(3), 414-439.

Schnell, R. (1997). *Nonresponse in Bevölkerungsumfragen: Ausmaß, Entwicklung und Ursachen*. Opladen: Leske + Budrich.

Schnell, R., Gramlich, T., Mosthaf, A., und Bender, S. (2010): Using complete administration data for nonresponse analysis: The PASS survey of low-income households in Germany. *Proceedings of Statistics Canada Symposium 2010*.

Schnell, R. und Kreuter, F. (2005): Separating Interviewer and Sampling-Point Effects. *Jour-*

nal of Official Statistics, 21, 389–410.

Stinson, M. (2002): Estimating measurement error in SIPP annual job earnings: A comparison of census survey and SSA administrative data. U.S. Census Bureau Technical Report TP-2002-24, Suitland, MD.

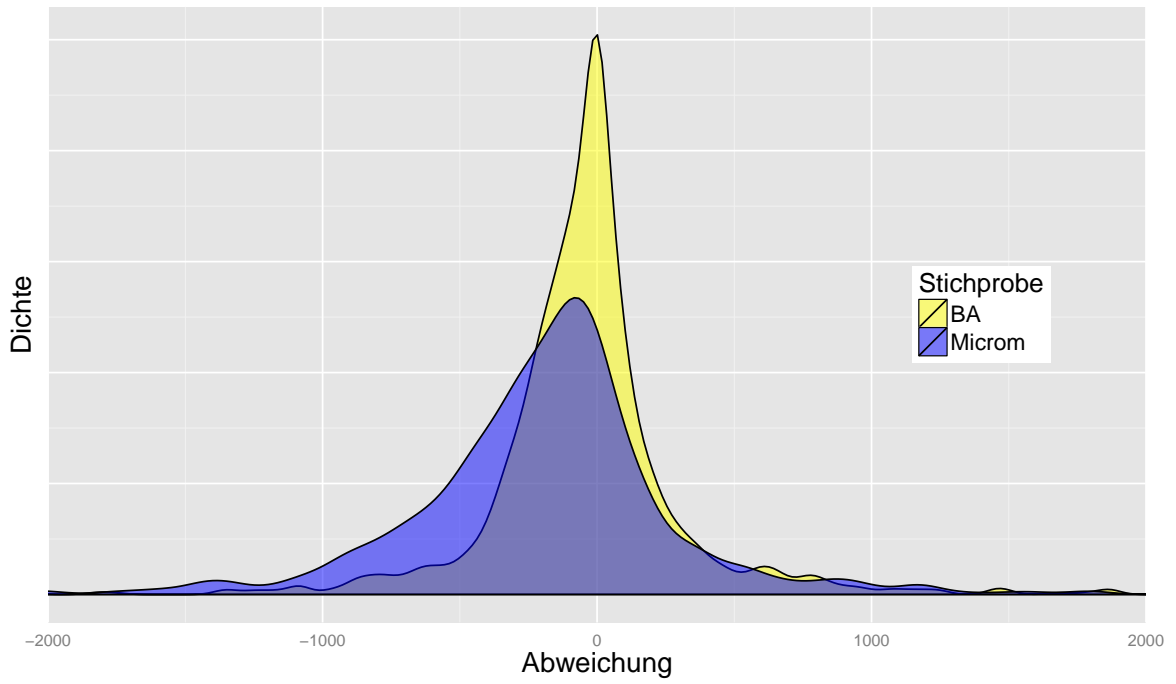
Vom Berge, Philipp; König, Marion; Seth, Stefan: Stichprobe der Integrierten Arbeitsmarktbiografien (SIAB) 1975 - 2010. FDZ-Datenreport 01/2013, Institut für Arbeitsmarkt- und Berufsforschung, Nürnberg.

West B., Kreuter F., und Jaenichen U. (2013): “Interviewer” Effects in Face-to-Face Surveys: A Function of Sampling, Measurement Error, or Nonresponse?. *Journal of Official Statistics*, 29, 277-297.

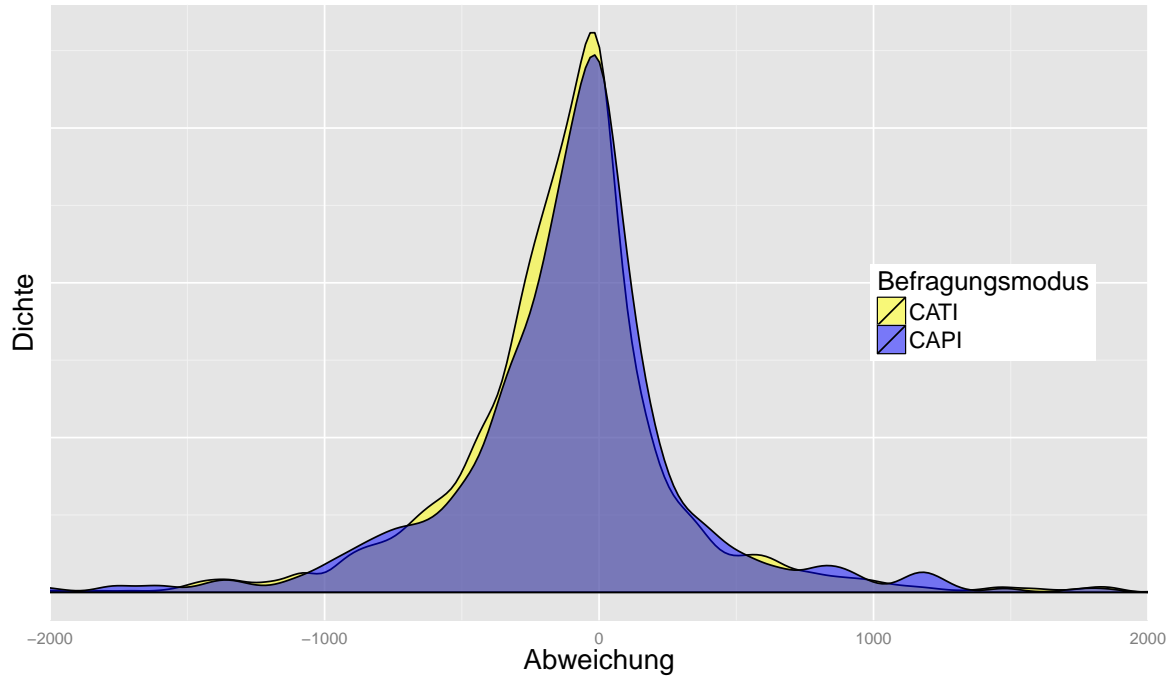
Anhang

Weitere Abbildungen

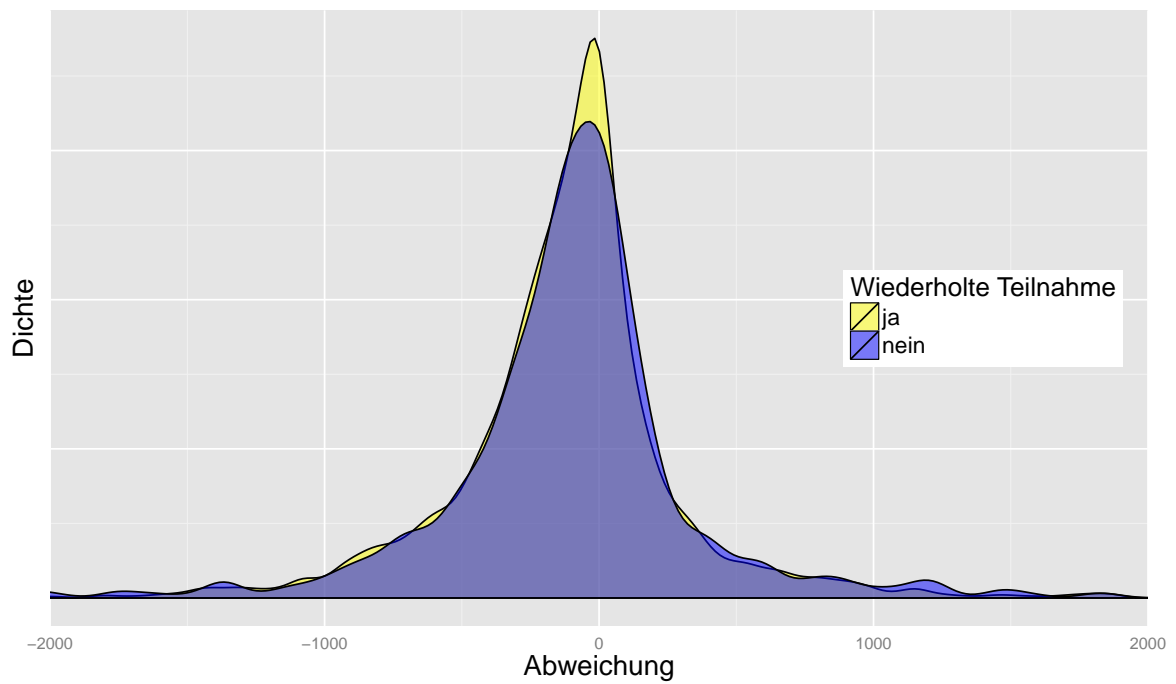
A.1 Verteilung des Messfehlers nach Herkunftsstichprobe



A.2 Verteilung des Messfehlers nach Befragungsmodus



A.3 Verteilung des Messfehlers bei Wiederholungsteilnahme



Erklärung zur Urheberschaft

Hiermit versichere ich, die vorliegende Masterarbeit selbstständig und lediglich unter Benutzung der angegebenen Quellen und Hilfsmittel verfasst zu haben.

München, 05. Februar 2015

Adrian Reichert