

methoden daten analysen

ZEITSCHRIFT FÜR EMPIRISCHE SOZIALFORSCHUNG

mda

2007, Jahrgang 1, Heft 1



Guido Mehlkop und Rolf Becker Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen

Karl Salentin Die Aussiedler-Stichprobenziehung

Oliver Lipps Attrition in the Swiss Household Panel

Britta Matthes, Maïke Reimer und Ralf Künstler Techniken und Werkzeuge zur Unterstützung der Erhebung retrospektiver Längsschnittdaten

Herausgegeben von *Christof Wolf*
Marek Fuchs
Bärbel Knäuper
Steffen Kühnel

Methoden – Daten – Analysen. Zeitschrift für Empirische Sozialforschung

Die Zeitschrift wird herausgegeben von der Gesellschaft Sozialwissenschaftlicher Infrastruktureinrichtungen e.V. (GESIS). Die GESIS ist eine Einrichtung der Leibniz-Gemeinschaft.

Herausgeber: Christof **Wolf** (Mannheim, geschäftsführend), Marek **Fuchs** (Kassel), Bärbel **Knäuper** (Montreal), Steffen **Kühnel** (Göttingen)

Wissenschaftlicher

Beirat: Hans-Jürgen **Andreß** (Köln), Andreas **Diekmann** (Zürich), Sabine **Häder** (Mannheim), Udo **Kelle** (Marburg), Dagmar **Krebs** (Giessen), Frauke **Kreuter** (College Park, Maryland), Edith **de Leeuw** (Utrecht), Norbert **Schwarz** (Ann Arbor)

Redaktion: Rolf Porst
GESIS-ZUMA
Postfach 12 21 55
68072 Mannheim
Tel.: 0621 – 1246 228
E-Mail: mda@gesis.org
Internet: www.gesis.org/Publikationen/Zeitschriften/MDA/

Die MDA deckt alle Fragestellungen aus dem Bereich der Empirischen Sozialforschung ab, insbesondere aus dem Bereich der Umfragemethodik. Im Vordergrund stehen Artikel, welche die methodischen und/oder statistischen Kenntnisse der Profession erweitern, sowie Beiträge, die sich mit der Anwendung der Methoden der Empirischen Sozialforschung in der Forschungspraxis beschäftigen, oder solche, in denen ein statistisches Verfahren exemplarisch angewandt wird. Obwohl der Schwerpunkt auf Umfragemethoden liegt, sind Beiträge zu anderen methodischen Bereichen willkommen.

Alle Beiträge, die zur Veröffentlichung in der MDA eingereicht werden, werden von mindestens zwei unabhängigen Gutachtern blind begutachtet.

Der Nachdruck von Beiträgen ist nach Absprache möglich. Die MDA erscheint dreimal im Jahr und stehen als Printversion und online zur Verfügung. Die Registrierung für den Bezug der MDA erfolgt über die Webseiten der GESIS:

http://www.gesis.org/Publikationen/Zeitschriften/MDA/MDA_subscribe.asp

Druck: Concordia-Druckerei König oHG, Mannheim-Sandhofen
Gedruckt auf chlorfrei gebleichtem Papier.

ISSN 1864-6956

1. Jahrgang 2007 © GESIS, Mannheim, Juni 2007

Inhalt

2 Editorial

FORSCHUNGSBERICHTE

5 Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen zu kriminellen Handlungen. Theoretische Überlegungen und empirische Befunde eines Methodenexperiments
Guido Mehlkop und Rolf Becker

25 Die Aussiedler-Stichprobenziehung
Kurt Salentin

45 Attrition in the Swiss Household Panel
Oliver Lipps

69 Techniken und Werkzeuge zur Unterstützung der Erinnerungsarbeit bei der computergestützten Erhebung retrospektiver Längsschnittdaten
Britta Matthes, Maike Reimer und Ralf Künster

REZENSIONEN

93 Methoden der Sozialforschung. Andreas Diekmann (Hg.), 2006.
Michael Braun

94 Evaluationsforschung und Qualitätsentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Reinhard Stockmann, 2006.
Jürgen H. P. Hoffmeyer-Zlotnik und Natalja Menold

ANKÜNDIGUNGEN

97 Einführung in familien- und haushaltsbezogene Analysen mit dem Mikrozensus

98 Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt

100 Hinweise für unsere Autorinnen und Autoren

Editorial

Liebe Leserin, lieber Leser,

in Ihren Händen liegt die erste Ausgabe der neuen Zeitschrift *Methoden, Daten, Analysen: Zeitschrift für Empirische Sozialforschung* – kurz *MDA*. Der Schwerpunkt der Zeitschrift liegt auf den Methoden der Umfrageforschung, also auf Veröffentlichungen zur Erhebung, Verarbeitung, Dokumentation und Analyse von Umfragedaten. Im Vordergrund stehen Artikel, welche die methodischen und/oder statistischen Kenntnisse der Profession erweitern, sowie Beiträge, die sich mit der Anwendung der Methoden der Empirischen Sozialforschung in der Forschungspraxis beschäftigen, oder solche, in denen ein statistisches Verfahren exemplarisch angewandt wird.

Die *MDA* ist hervorgegangen aus den beiden nunmehr eingestellten *GESIS*-Zeitschriften *ZA-Informationen* und *ZUMA-Nachrichten*. Diese beiden Publikationen hatten sich in den letzten Jahren zu wichtigen Diskussionsforen für die deutsche Methodenforschung entwickelt und dabei die wissenschaftlichen Standards der Beiträge deutlich angehoben. Durch die Zusammenfassung dieser Publikationen in eine Zeitschrift für Empirische Sozialforschung, die von einem unabhängigen Gremium herausgegeben und von einem international besetzten wissenschaftlichen Beirat beraten wird, wird diese Entwicklung konsequent fortgesetzt. Die *MDA* ist die Zeitschrift für das interdisziplinäre und zusehends internationale Feld der Methodenforscher und Methodenanwender. Sie wird von der *GESIS* als Dienstleistung für die scientific community getragen und ist zuallererst fachlichen Standards und einem konzisen thematischen Zuschnitt verpflichtet.

Die Zeitschrift wird herausgegeben von Christof Wolf (*GESIS*, geschäftsführender Herausgeber), Marek Fuchs (Universität Kassel), Bärbel Knäuper (McGill University, Montreal) und Steffen Kühnel (Universität Göttingen). Die eingereichten Manuskripte werden in einem zweistufigen Verfahren beurteilt. Zunächst entscheiden die Herausgeber, ob der Beitrag thematisch in die Zeitschrift passt und den einschlägigen Qualitätsstandards genügt. Passiert ein Beitrag diese Stufe wird er zwei Gutachtern in anonymisierter Form zur Bewertung vorgelegt. Auf Basis dieser Gutachten entscheiden die Herausgeber über Annahme, Überarbeitung oder Ablehnung des Manuskripts. Letztlich entscheidet allein wissenschaftliche Exzellenz über die Veröffentlichung oder Ablehnung eines Beitrags. Dafür garantiert das Herausbergremium, das seine Entscheidung auf Basis der Gutachten unabhängig von anderen Gesichtspunkten fällt.

Über die Ausrichtung und Entwicklung der Zeitschrift wacht ein hochkarätig besetzter internationaler Beirat, dem Hans-Jürgen Andreß, Andreas Diekmann, Sabine Häder, Udo Kelle, Dagmar Krebs, Frauke Kreuter, Edith de Leeuw und Norbert Schwarz angehören.

Beiträge können in deutscher und englischer Sprache eingereicht und publiziert werden, so dass der Wirkungskreis der Zeitschrift nicht auf den deutschsprachigen Raum beschränkt ist. Er wird sich jedoch auf den deutschsprachigen Bereich konzentrieren und stellt somit eine Ergänzung zu internationalen Publikationen wie dem *Journal of Official Statistics*, *Public Opinion Quarterly*, *Quality and Quantity*, *Social Science Research*, *Sociological Methods and Research*, *Survey Methodology* oder *Survey Research Methods* dar.

Die wissenschaftliche Umfrage ist längst fachübergreifend von Bedeutung: Psychologie, politische Wissenschaft, Soziologie, Gesundheitsforschung, Wirtschaftsforschung, Human-Geographie oder Gerontologie sind Fächer, in denen standardisierte Befragungen heute eine bedeutende Rolle spielen. Die methodischen Grundlagen wissenschaftlicher Umfragen sind somit ein interdisziplinäres Feld. Die MDA versteht sich daher als interdisziplinäres Forum für Beiträge zu allen Belangen der Methodologie standardisierter Befragungen, die zur Entwicklung der „Wissenschaft von der Umfrage“ beitragen, indem sie das breit vorhandene Erfahrungswissen zur Durchführung standardisierter Befragungen durch theoretische Ansätze – wie etwa dem Total Survey Error oder dem kognitionspsychologischen Ansatz – aufbereiten, systematisieren und ergänzen.

Die Beiträge in diesem Heft verdeutlichen die Ausrichtung der MDA. Guido Mehlkop und Rolf Becker untersuchen vor dem Hintergrund der Austauschtheorie und der Theorie der rationalen Wahl in einem experimentellen Design den Einfluss materieller Anreize auf die Rücklaufquoten postalischer Umfragen. Kurt Salentin beschreibt, wie man eine Zufallsstichprobe von Aussiedlern ziehen kann, obwohl ein eigenes Register für diese sozialpolitisch wichtige Gruppe fehlt. Oliver Lipps, der das Ausmaß und die Selektivität der Panelausfälle im Schweizer Haushalt-Panel untersucht, trägt mit diesem Beitrag zu unserem Wissen über diese besonders wichtige Form der Sammlung von Längsschnittdaten bei. Britta Matthes, Maike Reimer und Ralf Künstler schließlich beschreiben ein Programm, das die Inkonsistenzen bei der Erhebung von Retrospektivangaben deutlich vermindern und somit die Datenqualität dieser Erhebungsform entscheidend verbessern kann.

Neben den wissenschaftlichen Originalbeiträgen wird jedes Heft der Zeitschrift einen Rezensionsteil enthalten, der über einschlägige Neuerscheinungen informiert. In dieser Ausgabe bespricht Michael Braun einen von Andreas Diekmann herausgegebenen Sammelband zu den Methoden der Empirischen Sozialforschung; Jürgen Hoffmeyer-Zlotnik und Natalja Menold besprechen einen Band von Reinhard Stockmann zur Evaluationsforschung.

Wenn Sie ein Manuskript bei der MDA einreichen möchten, senden Sie dieses bitte an mda@gesis.org. Informationen zu den erwarteten Standards finden Sie am Ende des Heftes. Sollten Sie Interesse an einer Tätigkeit als Gutachter oder Rezensent für die Zeitschrift haben, können Sie uns dies ebenfalls unter mda@gesis.org mitteilen.

Wenn Sie unsere neue Methodenzeitschrift beziehen möchten, können Sie sich sowohl für den Bezug der Printausgabe als auch für den Bezug eines elektronischen Hinweises zum Erscheinen eines neuen Heftes registrieren. Für beides steht Ihnen ein Formular auf der Seite www.gesis.org/Bestellen/Publikationen/index.htm zur Verfügung.

Zum Schluss noch ein besonderer Dank an Rolf Porst, der die nicht immer leichte Aufgabe als Redakteur der Zeitschrift hervorragend erfüllt, und an Christa von Briel, die unsere Arbeit durch ihren unermüdlichen Einsatz bei der Erstellung der Druckvorlagen zu einem erfolgreichen Abschluss gebracht hat.

CHRISTOF WOLF * MAREK FUCHS * BÄRBEL KNÄUPER * STEFFEN KÜHNEL

Zur Wirkung monetärer Anreize auf die Rücklaufquote in postalischen Befragungen zu kriminellen Handlungen

Theoretische Überlegungen und empirische Befunde eines Methodenexperiments

The effects of monetary incentives on the response rate in mail surveys on self-reported criminal behavior

Theory and empirical results of an experiment

Guido Mehlkop und Rolf Becker

Zusammenfassung

Um Antwortverweigerungen bei postalischen Befragungen zu minimieren, wird zunehmend empfohlen, materielle Anreize in Form von geldwerten Geschenken für die Befragten einzusetzen. Aus Sicht der Austauschtheorie und eines aufgeklärten Rational Choice-Ansatzes sollten diese Geschenke den Nutzen aus einer Teilnahme für die Befragten erhöhen, die Kosten kompensieren bzw. ein Gefühl der Verpflichtung zur Teilnahme erzeugen. Erstaunlicherweise gibt es im deutschsprachigen Raum nur wenige Studien zu diesem Zusammenhang. In diesem Beitrag wird über ein Methodenexperiment in Dresden berichtet. Dabei wurde im Rahmen des Pretests zu einer größeren Bevölkerungsumfrage dem Fragebogen einer zufällig ausgewählten Versuchsgruppe eine 5-Euro-Banknote beigelegt. Einer gleich großen und ebenfalls zufällig ausgewählten Kontrollgruppe wurde dieses Geschenk vorenthalten. Während in der Versuchsgruppe rund 52% der Angeschriebenen antworteten, waren dies in der Kontrollgruppe nur 28%. Dieser Unterschied ist signifikant. Er unterstreicht die Bedeutung von monetären Anreizen für die Rücklaufsteigerung in postalischen Befragungen.

Abstract

In order to minimize non-response in mail surveys it is recommended to use monetary incentives for the respondents. From the perspective of both the theory of social exchange and of rational choice it is assumed that such an incentive significantly increases the utility of participation for the respondents and compensates them for their costs. Further, it is assumed that the incentive raises feelings of obligation to fill out the questionnaire. However, the number of studies referring to this topic is still limited, especially in the German-speaking countries. In this paper, we report on the results of an experiment carried out in Dresden. In the context of a population survey, the experimental group received an incentive (5-euro bill) along with the questionnaire while the control group received only the questionnaire. While about 52% of those in the experimental group returned the questionnaire, only 28% of those in the control group did. This significant difference confirms the importance of prepaid monetary incentives in order to increase the response rate in mail surveys.

1 Einleitung¹

Bei postalischen Befragungen zu heiklen Themen wie Kriminalität ist vor allem dann mit einer hohen Verweigerungsrate („nonresponse“) zu rechnen, wenn die Befragten selbst über ihr kriminelles Handeln berichten sollen (Becker 2006; Becker/Günther 2004). Aus Sicht der Rational Choice- und der Austauschtheorie liegt dies daran, dass viele Probanden die mit dem Ausfüllen des Fragebogens verbundenen Kosten höher einschätzen als ihren aus der Teilnahme resultierenden persönlichen Nutzen. Die Beteiligung an einer wissenschaftlichen Untersuchung ist vor diesem theoretischen Hintergrund nur dann zu erwarten, wenn der subjektive Nutzen die erwarteten Kosten übertrifft (Esser 1974). Kosten aus der Teilnahme an einer Befragungen zu sensiblen Themen können z. B. aus der Preisgabe privater Informationen, der Bekanntgabe „wahrer“ Angaben bei heiklen Fragen, aus der Erwartung von Sanktionen und aus den mit der Befragung verbundenen zeitlichen Investitionen und kognitiven Anstrengungen entstehen (De Leeuw 2001). Dem stehen mögliche Nutzenaspekte gegenüber, die z. B. aus dem Wunsch nach Artikulation der eigenen Meinung und Identität, dem Bedürfnis nach Unterstützung wissenschaftlicher Forschung oder aus Interesse am Thema der Befragung bzw. aus dem Wunsch nach sozialer Anerkennung als „gute“ Befragungsperson resultieren können (Esser 1986; Groves/Singer/Corning 2000: 302–303).

Es ist im Rahmen von postalischen Befragungen – auch aus entscheidungs- und handlungstheoretischer Sicht – jedoch nicht davon auszugehen, dass alle Probanden die für sie entstehenden Kosten und Nutzen eindeutig einzuschätzen und gegeneinander abzuwägen vermögen. Viele Probanden dürften zu Beginn einer Befragungssituation einer Teilnahme im Gegenteil eher unentschlossen gegenüber stehen. Es wird deshalb – und hierbei vor allem von Dillman (1978, 2000) – immer wieder darauf hingewiesen, dass zusätzlich zur intrinsischen Motivation der erwartete Nutzen aus der Teilnahme an einer postalischen Befragung durch materielle Anreize (so genannte „Incentives“) erhöht werden kann. Materielle Anreize scheinen besonders dann eine Rücklaufsteigerung zu bewirken und vormals Unentschlossene oder auch Ablehnende zur Mitwirkung motivieren zu können, wenn sie im Voraus, und ohne an Bedingungen geknüpft zu sein, gegeben werden. Viele Befragte fühlen sich durch sie verpflichtet, mit der Teilnahme an der Untersuchung und der Beantwortung der (heiklen) Fragen im Sinne der Reziprozitätsnorm (Gouldner 1960) eine adäquate Gegenleistung zum erhaltenen „Geschenk“ zu erbringen (Mauss 1925; Groves/Singer/Corning 2000: 303).² Grundsätzlich können diese Anreize in Form von beigelegtem Bargeld oder Geschenken (wie Kugelschreiber, Telefonkarten, Lotterielose o.ä.) dargereicht werden.

1 Wir danken den Herausgebern und zwei anonymen Gutachtern für wertvolle Kritik und Anmerkungen. Weiterhin bedanken wir uns bei der Deutschen Forschungsgemeinschaft für die finanzielle Förderung des Forschungsprojektes „Kriminalität in Dresden“.

2 Materielle Anreize sollten aus diesen Gründen auch in Panelstudien mit schriftlichem Fragebogen (und insbesondere bei wiederholten Befragungen zu heiklen Themen) eingesetzt werden, da die Teilnehmer und Teilnehmerinnen motiviert werden müssen, sich mindestens zwei Mal an der Untersuchung zu beteiligen (Rendtel 1995). Eine systematische Panelmortalität („unit nonresponse“) würde zur Verzerrung der Stichprobe führen und die Qualität der Daten erheblich einschränken (Grilliches/Hall/Hausman 1978; Little/Rubin 1987).

Während in angelsächsischen Raum die Wirkung beider Formen von Incentives relativ gut untersucht wurde (und sich dort Bargeld als die wirkungsvollere Form herausgestellt hat, siehe Church 1993; Dodd 1998; Furse/Steward/Rados 1981; Groves/Singer/Corning 2000; James/Bolstein 1990; Yammarino/Skinner/Childers 1991; Yu/Cooper 1983), liegen für den deutschsprachigen Raum vor allem Untersuchungen über die Wirkung geldnaher Incentives, nicht aber von Bargeld vor (Arzheimer/Klein 1998; Diekmann/Jann 2001; Harkness et al. 1998; Porst 1999). Dieser Beitrag versucht, diese Lücke zu schließen. Unseres Wissens ist unsere Studie die erste im deutschsprachigen Raum, die die Wirkung von beim Erstkontakt dargebotenen, rein monetären Anreizen untersucht.

Diese Studie geht somit der Frage nach, ob die rücklaufsteigernde Wirkung monetärer Anreize auch im deutschsprachigen Raum nachzuweisen ist. Wenngleich hier davon ausgegangen wird, dass die Reziprozitätsnorm nach dem Erhalt eines Geschenkes kulturunabhängig ist, ist diese Frage letztlich empirisch zu beantworten. Eine weitere offene Frage ist, ob die Beigabe von materiellen Incentives auch bei postalischen Umfragen zu heiklen Themen wie der (anonymen) Preisgabe eigener krimineller Handlungen die gewünschten Effekte bringt, und schließlich ist auch die Frage noch nicht beantwortet, ob die monetären Anreize für einige Gruppen stärker wirken als für andere Gruppen. In der vorliegenden Untersuchung kann dies allerdings nur ansatzweise für Geschlechterunterschiede getestet werden.³

In den folgenden Abschnitten sollen die oben formulierten Argumente für die Verwendung von Incentives ausgeführt und auf der Grundlage des bisherigen Wissensstandes darlegt werden, wie wichtig materielle Anreize zur Maximierung des Rücklaufs bei postalischen Befragungen sind. Im ersten Abschnitt gehen wir detailliert auf theoretische Erklärungen zur Wirkung von Incentives in postalischen Befragungen ein. Im zweiten Abschnitt werden mehrere empirische Studien und Methodenexperimente vorgestellt, welche die Bedeutung materieller Anreize demonstrieren. Schließlich wollen wir im dritten Abschnitt Ergebnisse eines eigenen Methodenexperiments und Schlussfolgerungen daraus vorstellen. Ziel des Experiments war es, herauszufinden, ob mit dem Fragebogen verschickte Geldbeträge („prepaid monetary incentives“) die Antwortbereitschaft in postalischen Befragungen zum *eigenen delinquenten Verhalten* zu erhöhen vermögen. Im vierten und letzten Abschnitt erfolgt die Diskussion der Befunde, aus der Schlussfolgerungen für die empirische Sozialforschung gezogen werden.

3 Es liegen zwar in der Forschungsliteratur einige Hinweise dazu vor, dass sich Incentives günstig auf die Teilnahmebereitschaft von üblicherweise eher untervertretenen Bevölkerungsgruppen auswirken (Martin/Abreu/Winters 2001; Martinez-Ebers 1997; Nederhof 1983; Singer 1998; Singer/van Hoewyk/Maher 2000) – allerdings sind diese Ergebnisse spärlich und teilweise widersprüchlich. Umgekehrt liegen keine Befunde vor, die zeigen, dass die Beigabe von Incentives die Stichprobenzusammensetzung dramatisch verzerrt (Arzheimer/Klein 1998; Berry/Kanouse 1987; Furse/Stewart 1982; James/Bolstein 1990; Shettle/Mooney 1999; Warriner et al. 1996). Auch in dieser Studie können wir diese Frage nicht letztlich klären, finden aber zumindest keine Anzeichen dafür, dass die Incentives unterschiedlich auf die Geschlechter wirken.

2 Theoretische Grundlagen: Rational Choice- und Austauschtheorie

Generell gilt, dass postalische Befragungen mittels standardisierter Fragebogen den höchsten (Item- und Unit-) Nonresponse aufweisen (Schräpler 2001: 13). Die Verallgemeinerbarkeit der Ergebnisse einer Studie und unter Umständen auch die Qualität der erhobenen Daten nehmen mit wachsender Kluft zwischen angestrebter und tatsächlich nutzbarer Fallzahl ab,⁴ wobei dieses Problem des Nonresponse bei postalischen Befragungen oft als besonders virulent angesehen wird (vgl. etwa Schnell/Hill/Esser 1999; kritisch dazu Reuband 2001: 308). Somit bedarf es Strategien zur Maximierung der Rücklaufquote (vgl. dazu Klein/Porst 2000). Durch den Einsatz personalisierter Erinnerungsschreiben nach der Methode von Dillman (1978, 2000) können die Rücklaufquoten in postalischen Befragungen erhöht werden. Auch der Einsatz von materiellen Anreizen für die Befragten wird in diesem Zusammenhang diskutiert. Doch nicht nur die reine Rücklaufquote ist zu beachten, sondern auch, ob es signifikante Unterschiede zwischen Teilnehmern und Verweigerern gibt (Koch 1998). So wird oft angenommen, dass besonders relativ alte und sehr junge Menschen (allerdings aus unterschiedlichen Gründen) sowie schlechter Gebildete Schwierigkeiten mit postalischen Befragungen haben (vgl. die Diskussion in Reuband 2001: 307). Man kann auch davon ausgehen, dass für erwerbstätige Personen mit wenig Freizeit die Opportunitätskosten der Zeit zum Ausfüllen eines Fragebogens hoch sind. Sozial selektive Verweigerungen beeinträchtigen in sozialwissenschaftlichen Analysen die Qualität der Stichprobe und die Daten (Rendtel 1989; Pötter/Rendtel 1993).

Zur Maximierung der Responseraten wird in der empirischen Sozialforschung der Einsatz von Geld als Anreiz bei postalischen Befragungen zunehmend befürwortet und in der Praxis durchgeführt (Dillman 2000: 167; Singer 1998: 7). In postalischen Befragungen scheinen extrinsische Anreize besonders angezeigt, da die Befragten die (teilweise recht umfangreichen) Fragebögen allein ausfüllen müssen. Es fehlen Interviewerinnen und Interviewer, die die Befragten motivieren und bei Schwierigkeiten unterstützen können (Singer 1998: 9). In dem von uns durchgeführten Experiment, auf das wir noch eingehen werden, sollten die Befragten einen voll standardisierten Fragebogen mit 77 Fragen (ca. 10 DIN-A 4 Seiten) ausfüllen. Der Fragebogen beinhaltete auch Fragen zum kriminellen Verhalten der Befragten.⁵ Bei solch persönlichen (und rechtlich relevanten) Fragen sollte die Antwortbereitschaft nochmals geringer sein, als bei Befragungen zu weniger sensiblen Themen.

4 Grundsätzlich kann man zwischen zwei Nachteilen geringer Ausschöpfungsquoten unterscheiden. Zum einen sind dadurch die Möglichkeiten der multivariaten Analyse und der statistischen Absicherung der Befunde eingeschränkt. Zum zweiten ist die Verallgemeinerbarkeit der Befunde auf die Grundgesamtheit beeinträchtigt, sollte der Ausfall systematisch sein.

5 Folgende Delikte wurden im Fragebogen thematisiert: Ladendiebstahl, Versicherungsbetrug, Steuerbetrug, Unfallflucht und Schwarzfahren. Die Befragten sollten angeben, ob sie diese Delikte bereits in der Vergangenheit begangen haben, ob sie die Delikte gegebenenfalls in Zukunft begehen würden, ob Personen aus ihrem näheren Bekanntenkreis diese Delikte begehen. Ferner wurden nach Einstellungen zu kriminalitätsrelevanten Normen und Gesetzen sowie nach der Einschätzung von Nutzen aus kriminellen Handlungen, erwarteter Strafhöhe, erwarteter Entdeckungswahrscheinlichkeit und weiterer Variablen im Sinne der Rational Choice-Theorie gefragt. Am Ende des Fragebogens befand sich ein relativ langer demographischer Teil.

Die theoretischen Grundlagen und Argumente für die positiven Effekte materieller Anreize werden dabei aus der Rational Choice-Theorie bzw. der mit dieser verwandten Austauschtheorie abgeleitet. Beide Ansätze vermuten prinzipiell deutliche positive Effekte materieller Anreize auf die Bereitschaft, an einer Befragung teilzunehmen und die Fragen (möglichst vollständig und wahrheitsgemäß) zu beantworten. Im Detail gibt es jedoch wichtige Unterschiede, die hier erörtert werden. Generell kann man aus entscheidungstheoretischer Sicht die Befragungssituation – wie jede andere soziale Situation auch (Esser 1986) – als sozialen Prozess sehen, in dem der Befragte als Akteur die Situation interpretiert, die ihm zur Verfügung stehenden Handlungsalternativen evaluiert und letztlich diejenige auswählt, die ihm den höchsten (subjektiv erwarteten) Nutzen realisiert (Esser 1991: 55ff.; Schräpler 2001: 2). Dabei kann die Befragungssituation aus Sicht des Befragten in drei Stufen eingeteilt werden: Zuerst muss die jeweilige Frage verstanden werden, dann müssen zweitens die zur Verfügung stehenden Handlungsalternativen evaluiert werden (etwa Antwort oder Verweigerung, richtige oder falsche Angaben, Abbruch des Interviews, etc.). Drittens muss sich der Befragte für eine konkrete Handlung entscheiden (Riphahn/Serfling 2002: 4). In der Befragungssituation entstehen dem Befragten Kosten (Konzentration, Offenbarung seiner Einstellungen oder Handlungen sowie Opportunitätskosten für die Zeit, die benötigt wird, um den Fragebogen auszufüllen).

Riphahn/Serfling (2002: 5) verweisen darauf, dass die Schlüsselkosten in einer Befragungssituation die Preisgabe privater Informationen und die Anstrengungen zur Beantwortung der Fragen sind. Der Nutzen, welcher dem Akteur aus der Beteiligung an der Befragung zukommt, kann im Interesse am Thema der Befragung liegen oder beispielsweise auch darin, seine Einstellungen kund zu tun, sowie soziale Anerkennung zu erlangen (dies möchten wir den intrinsischen Nutzen nennen, vgl. Esser 1991: 55). Übersteigen jedoch die Kosten den intrinsischen Nutzen, so wird der Befragte von der Teilnahme absehen. Somit können (zusätzlich zur intrinsischen Motivation) extrinsische Motivationen die Kosten-Nutzen Bilanzierung zugunsten des Nutzens beeinflussen, oder in den Worten von Hill/Willis (2001: 416):

„The model implies a participation decision rule in which the sample member responds if and only if the expected monetary evaluation of the utility of the interview process plus the incentive payment exceeds the expected time costs of the interview.“ (vgl. auch Riphahn/Serfling 2002: 8)

Dies gilt besonders für die Gruppe der Unentschlossenen (vgl. Schräpler 2001: 5). Menschen, die aufgrund eines Interesses am Thema oder wegen des Bedürfnisses, ihre Einstellungen zu artikulieren motiviert sind, bedürfen solcher extrinsischer Motivationen nicht. Ebenso ist es ziemlich unwahrscheinlich, dass Menschen, die von vornherein Befragungen ablehnend gegenüberstehen, durch äußere Motivationen umgestimmt werden können (siehe auch Groves/Singer/Corning 2000: 301). Es ist davon auszugehen, dass die Gruppe der Unentschlossenen gerade beim Thema „Kriminalität“ recht groß ist, weil die Befragten sich darüber unsicher sind, ob sie Informationen über ihr eigenes kriminelles Handeln preisgeben sollen und deswegen die Schwelle der Teilnahmebereitschaft höher ist, als bei Befragungen zu weniger sensiblen Themen. Eine extrinsische Motivation in Form von materiellen Anreizen ist bei postalischen Befragungen auch deswegen wichtig, weil schriftlich

kontaktierten Probanden die Antwortverweigerung leichter fällt als Probanden in mündlichen Befragungen bei Anwesenheit eines Interviewers. Somit muss der Akteur einen Nutzen für sich sehen, der durch zusätzliche Anreize garantiert scheint. Da bei postalischen Befragungen in der Regel die Anonymität garantiert ist und wegen der Abwesenheit von Interviewern auch keine unmittelbaren Sanktionen zu befürchten sind, dienen materielle Anreize in diesem Kontext vor allem der Kompensation anfallender Partizipationskosten zu Gunsten einer höheren Teilnahme.⁶ Die Literatur zu Nonresponse stimmt weitgehend darin überein, dass die Befragten – wenn sie prinzipiell zur Beteiligung bereit sind – in der Regel die meisten Fragen beantworten (Schräpler 2001: 12).⁷ Somit gilt es, die Schwelle zur Beteiligung zu überschreiten, wobei der Schwellenwert durch materielle Anreize gesenkt werden kann.

Abgesehen von den bisher diskutierten kompensatorischen Effekten materieller Anreize für zeitbezogene Opportunitätskosten werden seitens der Austausch- bzw. Reziprozitätstheorie weitere Argumente für materielle Anreize ins Feld geführt. Durch Incentives soll ein Gefühl der Verpflichtung zur angemessenen Gegenleistung – also der Reziprozität – bei den Befragten geweckt werden. Dillman (2000: 153) geht sogar so weit zu sagen, dass Incentives bei den Befragten Vertrauen gegenüber den Forschern bzw. der Studie generieren können:

„Sending the money with the questionnaire is a goodwill gesture that puts the sponsor and questionnaire in a positive light and sets the stage for the respondent to reciprocate with an appropriate gesture of completing the questionnaire.“ (Dillman 2000: 168)

Während die Rational Choice- und die Austauschtheorie den Nutzen materieller Anreize prinzipiell betonen, herrscht keine Einigkeit darüber, wie Incentives die Befragten zur Beantwortung der Fragen bringen und *auf welche Art* sie eingesetzt werden sollen. Aus Sicht der restriktiven Rational Choice-Theorie ist keineswegs sicher, dass ein dem Fragebogen beigelegtes Geschenk die Rücklaufquote erhöht. Eine strikt rationale Person (die nicht durch das Thema an sich bereits hinreichend interessiert ist, an der Befragung teilzunehmen) behält das Geschenk und enthält sich der Teilnahme. Anders sollte es sein, wenn ein Geschenk in Aussicht gestellt und nur bei nachweisbarer Beteiligung ausgehändigt wird. Dann nämlich geht das versprochene Geschenk zusätzlich zum Nutzen aus dem Interesse an der Befragung in die Nutzenkalkulation des Befragten ein. Übersteigt der Gesamtnutzen die erwarteten Gesamtkosten der Befragung, dann werden diese Personen sehr wahrscheinlich an der Befragung teilnehmen (Diekmann/Jann 2001: 21). Somit prognostiziert die strikte Rational Choice-Theorie den Erfolg versprochener, aber nicht direkt beigelegter Geschenke.

Anders sollte es sich verhalten, wenn man die Austausch- bzw. Reziprozitätstheorie zugrunde legt. Diese basiert auf Überlegungen von u. a. Blau (1964) und Homans (1961) und wurde erstmals von Dillman (1978) zur Erklärung des Teilnahmeverhaltens in postalischen Befragungen

6 Materielle Anreize können als Entscheidungshilfe gesehen werden: Die Investitionen beim Ausfüllen des Fragebogens werden anerkannt und belohnt. Der persönliche Nutzen bzw. die Motivation zur Teilnahme werden erhöht (Krebs 1995: 118).

7 Wobei festzuhalten ist, dass lückenhafte Datensätze in vielen Studien eine Realität darstellen.

herangezogen. Die Grundidee des modifizierten austauschtheoretischen Ansatzes von Dillman (2000) besteht darin, dass sich viele Befragte bei Erhalt des Fragebogens unsicher darüber sind, ob sie überhaupt teilnehmen sollen und dass diese unentschlossenen Personen zur Teilnahme motiviert werden können, wenn ihnen die Teilnahme so leicht wie möglich gemacht wird, wenn sie sich zur Teilnahme verpflichtet fühlen und sie den Eindruck gewinnen, dass sich die Teilnahme lohnt bzw. ihnen wenigstens nicht schadet (vgl. Schnell 1997: 251). Die Interview-Situation wird als sozialer Austausch gesehen zwischen dem Forscher, der Informationen begehrt, welche im Besitz des Befragten sind und dem Befragten, der entscheidet ob, und wenn ja, welche Informationen er preisgibt (Hill/Willis 2001: 418). Hier erzeugt die Vorleistung durch ein beigelegtes Geschenk ein Gefühl der Verpflichtung, ebenfalls eine Leistung zu erbringen, sprich an der Befragung teilzunehmen (siehe etwa Diekmann/Jann 2001: 21), selbst wenn die Antworten aus Sicht der Befragten kostspielig, weil peinlich, schmerzhaft oder kognitiv schwierig sind (Hill/Willis 2001: 418). Krebs (1995: 119) etwa vermutet – im Gegensatz zur restriktiven Rational Choice-Hypothese – dass das Ausloben eines Geschenkes *nur bei Teilnahme* die Bereitschaft den Fragebogen auszufüllen tendenziell senkt, weil dies den Effekt der Vorleistung neutralisiert und so keine Reziprozität produziert wird. Daher ist anzunehmen, dass das Gefühl der Verpflichtung (ausgelöst durch eine Vorleistung beim *Erstkontakt*) die Bereitschaft zur Teilnahme erhöht.

Zusammenfassend kann aus Sicht der Rational Choice- und der Austauschtheorie gefolgert werden, dass sich durch materielle Anreize die Rücklaufquote und die Bereitschaft, auf alle Fragen zu antworten, in postalischen Befragungen signifikant erhöhen lassen sollte.

Uneinigkeit herrscht zwischen den beiden theoretischen Ansätzen darüber, ob versprochene oder direkt beigelegte Incentives eine höhere Wirkung erzielen. Dies ist letztlich eine empirische Frage. Deswegen sollen im folgenden Teil empirische Studien und Experimente vorgestellt werden, die die verschiedenen Varianten des Einsatzes von Incentives getestet haben und zum Grossteil belegen, dass der Einsatz von Incentives bei postalischen Befragungen zu einer signifikanten Steigerungen der Responserate führt, ohne dass dies die Qualität der Daten und die Teilnahmebereitschaft in späteren Untersuchungen gefährdet (Berger 2006).

Letztlich kann man auch die Frage stellen, ob die Beigabe monetärer Anreize die Teilnahmebereitschaft generell erhöht, oder ob bestimmte Gruppen stärker motiviert werden als andere Gruppen. Somit würde der Effekt der Incentives von anderen Merkmalen der Befragten abhängen bzw. mit diesen interagieren. Groves/Singer/Corning (2000: 301-302) etwa vermuten, dass monetäre Incentives stärker bei Befragten wirken, die nur gering in Sozialkapital generierende Netzwerke eingebunden sind („low community involvement“), da diese Befragten per se geringere Verpflichtungen gegenüber ihren Mitmenschen empfinden und so die oben angeführte Reziprozitätsthese nicht zuträfe. Die Klärung dieser Frage setzt allerdings umfassende Kenntnis über die Verweigerer voraus, über die wir in unserer Studie nicht verfügen.

3 Der positive Effekt materieller Incentives auf die Rücklaufquote in postalischen Befragungen – empirische Belege

Diekmann/Jann (2001: 19) berichten von zwei postalischen Befragungen, die Diekmann mit anderen Mitarbeitern in den 1990er Jahren in der Schweiz durchgeführt hat. Dabei wurden Telefonkarten im Wert von 10 Schweizer Franken als Geschenk mit versandt. Sie informieren über Rücklaufquoten, die mit Werten von 88 bzw. 84% weit über den Erwartungen lagen und schließen daraus, „... dass der Versand der Telefonkarten an diesem Ergebnis einen nicht unerheblichen Anteil hatte“ (ebenda). Leider war es in diesen beiden Studien nicht möglich, auch Kontrollgruppen aus Befragten zu bilden, die keine Geschenke erhalten haben.

Deswegen haben Diekmann/Jann (2001) ein weiteres Experiment mit zwei Versuchsgruppen und einer Kontrollgruppe durchgeführt (Nettostichprobe $N = 576$). Der ersten Versuchsgruppe wurde im Rahmen einer postalischen Befragung eine Telefonkarte im Wert von 10 Schweizer Franken bei Zurücksenden des ausgefüllten Fragebogens versprochen. Die Probanden in der zweiten Versuchsgruppe erhielten diese Telefonkarte bereits mit dem Fragebogen und den Probanden in der Kontrollgruppe wurde weder eine Karte beigelegt noch ein Geschenk versprochen. Die Rücklaufquote bei der Gruppe, der das Geschenk lediglich versprochen wurde, betrug nur 58,1%, im Gegensatz zu 72,4% in der zweiten Gruppe, die das Geschenk von Anfang an erhielt (vgl. Diekmann/Jann 2001: 23). Im Vergleich mit der Kontrollgruppe (Rücklaufquote 62,7%) ist die Rücklaufquote bei einem beigefügten Geschenk um 9,7 Prozentpunkte signifikant höher. Diekmann/Jann (2001: 25) schließen daraus, dass versprochene, an die Teilnahme geknüpfte Geschenke *nicht* zur Steigerung der Rücklaufquote beitragen – dass dem Erstkontakt beigelegte Geschenke die Rücklaufquote jedoch deutlich und statistisch signifikant erhöhen.

Singer (1998) berichtet von einem Experiment aus Detroit (USA), in welchem die Einstellungen zur Sterbehilfe mittels einer postalischen Befragung erhoben wurden. Die Hälfte der angeschriebenen Probanden erhielt mit dem Ankündigungsschreiben zu der Studie eine Banknote im Wert von fünf US-Dollar, die andere Hälfte erhielt kein Präsent. Die Rücklaufquote in der Experimentalgruppe betrug 65,1% und war damit signifikant höher als die Rücklaufquote in der Kontrollgruppe ohne Geschenk (41,4%).

Aufschlussreicher als die oben zitierten einzelnen Experimente bzw. Befragungen sind Meta-Analysen, die die Ergebnisse zahlreicher Studien zusammenfassen, vergleichen und analysieren. So hat Singer (1998) 37 experimentelle Studien über den Einsatz materieller Anreize untersucht.⁸ Die Ergebnisse dieser Meta-Analyse fasst Singer (1998: 10ff.) wie folgt zusammen: Incentives haben allgemein einen starken positiven Effekt auf die Rücklaufquoten. In Regressionsanalysen (abhängige Variable: Differenz der Rücklaufquoten zwischen Versuchs- und Kontrollgruppen) er-

8 Bei den 37 Studien, die der Meta-Analyse von Singer (1998) zu Grunde liegen, handelt es sich allerdings nicht um postalische, sondern um mündliche und telefonische Befragungen. Die Anreize wurden im Voraus gegeben, aber nicht einem Fragebogen beigelegt.

wies sich ihr Effekt als positiv, signifikant und linear – d. h. es gibt in den von Singer (1998) berichteten Studien keinen abnehmenden Grenznutzen materieller Anreize (siehe auch Church 1993; Trussell/Lavrakas 2004; Yu/Cooper 1983).⁹ Doch andererseits gibt es auch Studien, die einen abnehmenden Grenznutzen monetärer Incentives bzw. eine Stagnation der Rücklaufquote ab einem bestimmten Wert der Incentives berichten (Armstrong 1975; Fox/Crask/Kim 1988; James/Bolstein 1992; Martin/Abreu/Winters 2001; Warriner et al. 1996).¹⁰ Ein weiteres Ergebnis der Meta-Analyse von Singer (1998) ist, dass Anreize, die im Vorhinein gegeben werden, bessere Ergebnisse realisieren als Entschädigungen, die lediglich bei nachweisbarer Teilnahme ausgezahlt werden. Dillman (2000: 153) berichtet, dass vor allem jüngere Menschen, die ebenfalls in Studien ohne Incentives regelmäßig hohe Verweigerungsraten aufweisen, durch materielle Anreize zur Teilnahme an der Befragung motiviert werden können. Dies bedeutet, dass Menschen, denen die Mühe der Teilnahme normalerweise als zu hoch erscheint, durch die Incentives in ihrer Abwägung von Kosten und Nutzen signifikant beeinflusst werden. Es bedeutet jedoch nicht, dass diese Befragten nur des Geldes wegen den Fragebogen *aufs Geratewohl* ausfüllen. Im Gegenteil – die Vergabe von Geschenken an die Befragten realisiert eine höhere Ausschöpfung, eine bessere Zusammensetzung der Stichprobe bei nicht schlechterer Qualität der erhobenen Daten (Singer 1998: 25). Doch wie bereits in der Einleitung angesprochen, ist auch zu diesem Sachverhalt die Forschungslage nicht eindeutig, da auch Singer (1998) auf sechs Studien verweist, in denen kein Einfluss der Incentives auf die Zusammensetzung der Stichprobe gefunden werden konnte.

Eine weitere oft zitierte Meta-Analyse wurde von Church (1993) erstellt. Er analysierte insgesamt 38 experimentelle und quasi-experimentelle Studien, wobei diese qua der Anwendung der Incentives in vier Subgruppen unterteilt werden konnten: Experimente, in denen a) Geld oder b) andere materielle Incentives *direkt* den Fragebögen als Vorleistung beigelegt wurden und auf der anderen Seite Designs, in denen c) Geld oder d) materielle Anreize *versprochen* wurden, wenn und nur wenn der ausgefüllte Fragebogen zurückgesandt wurde. Betrachtet man die 38 Studien insgesamt, so wurde mit Hilfe der im vornherein beigelegten *monetären* Incentives die Rücklaufquote im Vergleich mit der Kontrollgruppe ohne Präsent durchschnittlich um 19,1% pro Studie erhöht, also um fast ein Fünftel. In 89% der von Church (1993: 68) untersuchten Studien konnten im Vergleich mit den jeweiligen Kontrollgruppen die Rücklaufquoten durch die beigelegten Gaben – egal ob lediglich versprochen oder im Vorhinein beigelegt – erhöht werden. Betrachtet man allerdings nur die Studien, in denen die finanziellen Anreize erst bei Rücksendung des Fragebogens

9 Laut Hill/Willis (2001: 422) ist die Teilnahmebereitschaft auch eine positive lineare Funktion aus der Höhe der materiellen Anreize – dies gilt vor allem bei Panel-Befragungen, bei denen der (zeitliche) Aufwand zur Beantwortung der Fragen ja periodisch wiederkehrt und deswegen immer wieder kompensiert werden muss. Wegen der Linearität des Effektes ist nicht anzunehmen, dass an einem bestimmten Punkt der Grenznutzen der Incentives erreicht ist, dass also weitere Incentives in der neuen Welle nicht mehr zur Teilnahme führen. Interessanterweise ist übrigens der Effekt der Incentives der einzige lineare Effekt. Andere Faktoren wie Erinnerungsschreiben zeigen kurvilineare Effekte, weisen also einen abnehmenden Grenznutzen auf. Somit erscheinen diese Incentive-Zahlungen besonders bei Zeitreihen- bzw. Panelstudien sinnvoll (vgl. Hill/Willis 2001: 426, Fußnote 7).

10 Dieser Frage kann allerdings im Rahmen dieses Experimentes nicht weiter nachgegangen werden, da die Höhe der beigelegten Geldbeträge nicht variiert.

ausgezahlt wurden, so gab es keine signifikanten Unterschiede zwischen Versuchs- und Kontrollgruppe (Church 1993: 71).¹¹ Church folgerte daraus (1993: 73): „..., only incentives provided with the initial mailing of the survey instrument had any significant or meaningful positive impact on response rates ...“. Auch die Art der materiellen Anreize machen einen Unterschied: Wurden monetäre Anreize verwendet (als ‚prepaid incentives‘), so stieg die Rücklaufquote um durchschnittlich 19% im Vergleich zu der Kontrollgruppe, die keinerlei Geschenke erhielt. Wobei die Rücklaufquote mit nicht-monetären Gaben mit fast acht Prozent geringer ausfiel (Church 1993: 62).

Die Verwendung von Bargeld statt anderer materieller Anreize (wie Telefonkarten oder Kugelschreiber) scheint aus zwei Gründen sinnvoll. Erstens sprechen die Studien von Church (1980), Hansen (1980) und Simmons/Wilmot (2004) für die Überlegenheit monetärer im Vergleich zu geldnahen Incentives. Bargeld ist fungibel und kann von den Befragten individuell verwendet werden. Materielle Anreize, wie Kugelschreiber, Schokolade usw. werden von unterschiedlichen Befragten als unterschiedlich wertvoll angesehen (Groves/Singer/Corning 2000: 301). In den 1990er Jahren konnten Telefonkarten in sinnvoller Weise als Anreize verwendet werden (Diekmann/Jann 2001). Dieses Vorgehen erscheint jedoch heute weniger brauchbar, da durch die weite Verbreitung von Mobiltelefonen Telefonkarten für öffentliche Fernsprecher unattraktiv erscheinen. Zweitens, die Beigabe von nicht-monetären Geschenken (wie Kugelschreibern) verursacht wegen des Gewichtes höhere Portokosten bzw. die Kugelschreiber in den Briefumschlägen machen eine maschinelle Sortierung der Postsendungen unmöglich. Auch die Verwendung von Verrechnungsschecks verursacht einen deutlich höheren Mehraufwand – für die Organisatoren der Studie ebenso wie für die Befragten selbst (siehe Dillman 2000: 168–169).

In einer weiteren US-amerikanischen Studie mit 378 Befragten aus Detroit fanden Groves/Singer/Corning (2000) ebenfalls, dass die Beigabe monetärer Anreize in Form einer 5-US-Dollar-Banknote beim Erstkontakt die Rücklaufquote deutlich erhöht. Während die Rücklaufquote in der Kontrollgruppe ohne Anreize rund 41% betrug, retournierten rund 65% der Befragten, die Anreize erhalten hatten, den Fragebogen. In dieser Studie wurde auch getestet, ob die Effekte der Anreize zwischen Befragten mit verschiedenen Merkmalen variieren.¹² Die Befragten wurden zusätzlich aufgrund ihrer Einbindung in soziale Netzwerke in zwei Gruppen unterschieden. In der Gruppe mit wenig ausgeprägten Netzwerken betrug der Rücklauf ohne beigelegten monetären Anreiz 21%, mit beigelegtem Anreiz 63%. Von den Befragten mit stark ausgeprägten Netzwerken sendeten ohne Anreiz 50% den Fragebogen zurück, wurde ein Anreiz beigelegt, stieg die Teilnahmebereitschaft auf 66%. Wenngleich die Autoren daraus folgern, dass die Wirkung der Incentives mit anderen Merkmalen interagiert, so ist dennoch festzuhalten, dass selbst unter den Befragten

11 Dillman (2000: 167) berichtet von einer Studie mit drei Versuchsgruppen, in der *versprochene* Geschenke im Wert von 50 US-Dollar eine Rücklaufquote von 57% realisieren konnten, aber ein *beigelegter* 1-US-Dollar-Schein 64% und ein *beigelegter* 5-Dollar-Schein sogar 71%. Dieser Befund spricht ebenfalls für beigelegte Geschenke.

12 Dieses Methodenexperiment wurde als follow-up Studie durchgeführt, d. h. für alle kontaktierten Befragten im Methodenexperiment lagen bereits umfangreiche Informationen vor. Dies lässt es auf der einen Seite zwar zu, die Merkmale der Verweigerer im Experiment näher zu bestimmen. Auf der anderen Seite kann hier nicht der Effekt der bereits erfolgten ersten Teilnahme kontrolliert werden.

mit starken Netzwerken (und einer daraus gefolgerten stärkeren Reziprozitätsverpflichtung) die Beigabe der Incentives die Teilnahmebereitschaft deutlich erhöhte. Schließlich war die Teilnahmebereitschaft mit Incentives in beiden Gruppen annähernd gleich hoch (63 bzw. 66%).

Der Erfolg von im Vorhinein beigelegten monetären Incentives kann somit für den angelsächsischen Raum als gut belegt angesehen werden. In der deutschsprachigen sozialwissenschaftlichen Forschung wurde bisher jedoch keine systematische Forschung mit *im Voraus gegebenen finanziellen* Anreizen durchgeführt. Die vorhandenen experimentellen Studien im deutschsprachigen Raum untersuchten vor allem die Wirkung von *geldnahen* Incentives (z. B. Arzheimer/Klein 1989; Harkness/Mohler/Schneid/Christoph 1998; Porst 1999; Diekmann/Jann 2001, die Briefmarken oder Telefonkarten beigelegt hatten). Die Frage, ob monetäre Incentives auch in anderen kulturellen Kontexten als dem angelsächsischen Raum wirken, kann letztlich auch nur empirisch beantwortet werden. Es gibt aber einige theoretische Argumente, die einen kultur-unabhängigen Effekt nahe legen. So bezeichnet Marcel Mauss (1984: 19) den jedem Geschenk innewohnenden Zwang zur Reziprozität, also zur Gegenleistung als „Felsen [...], auf denen unsere Gesellschaften ruhen“. Mauss (1984: 163) beschreibt die Erwartung von Reziprozität somit als anthropologische Konstante in menschlichen Transaktionen (vgl. auch Oswald 2004). Wenngleich Reziprozität somit auch generell als anthropologische Konstante angesehen werden kann, so gibt es auch Argumente für die unterschiedliche Ausprägung des Grades der Reziprozität zwischen Gesellschaften. Groves/Singer/Corning (2000) etwa operationalisieren das Verpflichtungsgefühl (als Grundlage der Reziprozität) über ehrenamtliches Engagement in Vereinen, Parteien und Organisationen und nehmen an, dass dieses „community involvement“ aus dem Gefühl der bürgerlichen Verpflichtung („civic duty“, S. 302) heraus die Teilnahme an Befragungen per se wahrscheinlicher macht (siehe auch Dillman 1978; Gallup 1940; Goyder 1987). Bei Menschen mit ohnehin hohem „community involvement“ sollten die monetären Anreize weniger stark wirken als bei Menschen mit niedrigem „community involvement“ (Groves/Singer/Corning 2000: 303). Dies lässt die Schlussfolgerung zu, dass die Wirkung von monetären Anreizen zwischen Gesellschaften mit unterschiedlichem mittleren „community involvement“ variiert. Beispielsweise wird das ehrenamtliche Engagement in den USA als stärker als in Deutschland beschrieben (Birnkraut 2003: 31). Somit scheint es angezeigt, den Einfluss der monetären Anreize im deutschsprachigen Raum zu untersuchen.

4 Methodenexperiment zur Wirkung eines monetären Anreizes in einer postalischen Befragung zu selbstberichtetem delinquentem Verhalten

Forschungshintergrund und experimentelles Design: Das hier vorgestellte Methodenexperiment wurde im Rahmen des Pretests für eine postalische Befragung von Bürgerinnen und Bürgern in der sächsischen Landeshauptstadt Dresden durchgeführt. In dieser von der DFG finanzierten Studie sollen bei der Hauptuntersuchung rund 3.500 Dresdnerinnen und Dresdner zur Kriminalität im

Allgemeinen und zur eigenen Delinquenz im Besonderen befragt werden. Es handelt sich um eine Querschnittsanalyse, die in starker Anlehnung an die Total bzw. Tailored Design Method von Dillman (2000) durchgeführt wird.¹³ Die Befragung soll Daten liefern, mit denen das modifizierte Rational Choice-Modell kriminellen Handelns von Mehlkop/Becker (2004) überprüft werden kann.

Der im Methodenexperiment verwendete Fragebogen wurde aufgrund der Erkenntnisse eines vorgängig durchgeführten kognitiven Pretests überarbeitet¹⁴, bevor er unter sehr ähnlichen Bedingungen eingesetzt wurde, die auch für die eigentliche Hauptuntersuchung vorgesehen sind (der einzige Unterschied besteht darin, dass im Experiment nicht mit einem Erinnerungsschreiben nachgefasst wird, in der Hauptuntersuchung ist jedoch ein Erinnerungsschreiben vorgesehen). Das Experiment wurde mit zwei Gruppen durchgeführt. Hierzu wurden 200 Personen aus der Stichprobe für die Hauptuntersuchung gezogen, die uns vom Einwohner- und Standesamt der Landeshauptstadt Dresden zur Verfügung gestellt wurde.¹⁵ Die Bildung der beiden gleich großen Untersuchungsgruppen erfolgte per Losentscheid. Die erste Gruppe erhielt mit dem Anschreiben und Fragebogen sofort den Anreiz von 5 Euro zugeschiedt. Die zweite Gruppe bildete die Kontrollgruppe. Ihre Mitglieder

- 13 In seiner viel beachteten *Tailored Design Method* schlägt Dillman (2000: 150ff.) fünf Elemente vor, die zur Maximierung des Rücklaufes führen. Erstens, einen „nutzerfreundlichen“ Fragebogen (einfach zu verstehende und eindeutig zu beantwortende Fragen; ein klares und ansprechendes Layout; ein überschaubarer Umfang; einleitende Fragen, die das Interesse des Befragten wecken usw.). Zweitens, wiederholte Kontaktaufnahmen mit den Befragten. Das dritte Element besteht aus einem adressierten und mit einer Briefmarke versehenen Rückumschlag. Dies erleichtert das Rücksenden. Das vierte Element bildet die Personalisierung der Korrespondenz (namentliche Anrede der Befragten). Als letztes Element zur Maximierung der Rücklaufquote nennt Dillman (2000: 153) finanzielle Anreize, wie beispielsweise eine 5-Dollar-Banknote. Der Unterschied zwischen der früheren Total und der neueren Tailored Design Method besteht vor allem in der Aufnahme der finanziellen Anreize als obligatorischem (statt optionalem) Element, der Frankierung mit echten Briefmarken (früher ebenfalls nur optional) sowie dem Informationsbrief als erste Kontaktaufnahme (in der alten Methode war dieser Erstkontakt nicht vorgesehen und es wurde gleich der Fragebogen verschickt; vgl. Dillman 2000: 153–4). Zusätzlich sollen Presseberichte die potentiellen Befragten über das Thema informieren und zur Teilnahme motivieren. Wir haben uns stark an Dillman orientiert, aber den Erstkontakt mit Informationsbrief und die echten Briefmarken weggelassen. Die „Sächsische Zeitung“ (mit einer Auflage von ca. 319.000 Exemplaren die größte Tageszeitung Dresdens) veröffentlichte am 8. August 2005 einen Bericht über die Befragung bzw. ein Interview mit Guido Mehlkop. Am 18. August 2005 wurde zeitgleich mit der Versendung der Fragebögen an die Stichprobe des Pretests ein zweiter, kleinerer Artikel in der Sächsischen Zeitung veröffentlicht, der nochmals die Befragung ankündigte und die Anonymität betonte. Im Zuge der Hauptuntersuchung werden weitere Berichte in dieser Tageszeitung folgen.
- 14 Zur Optimierung des Fragebogens sollten Pretests im Sinne des „Zwei-Phasen-Pretestings“ durchgeführt werden (siehe Prüfer/Rexroth 2000). Die erste Phase besteht aus einem kognitiven Pretest (Pretest I) und die zweite Phase aus einem herkömmlichen Pretest (Standard-Pretest II). Die kognitiven Techniken im Pretest I sollen erhellen, wie Befragte Begriffe und Formulierungen des Fragebogens verstehen und aufgrund welcher Überlegungen ihre Antworten zustande kommen, d. h. die Teilnehmer des ersten Pretests werden explizit und unabhängig von den eigentlichen Antworten zur Offenlegung ihrer Gedankengänge, Verständnisprobleme usw. aufgefordert. Dazu wurde der Fragebogen einerseits 15 Personen verschiedenen Alters und Berufes vorgelegt und deren Anmerkungen und Anregungen zu Frageformulierungen und Antwortkategorien notiert. Dabei ging es vornehmlich um die Verständlichkeit von Fragen und Antwortmöglichkeiten und darum, ob die Antwortmöglichkeiten erschöpfend sind. Ferner sollte herausgefunden werden, wie lange die Befragten zum Ausfüllen des Fragebogens brauchen. Andererseits wurde der Fragebogen verschiedenen Experten der Technischen Universität Dresden sowie ZUMA in Mannheim zur Begutachtung vorgelegt. An dieser Stelle möchten wir uns bei Michael Braun, Peter Graeff, Michael Häder, Karl-Siegbert Rehberg und Ekkart Zimmermann für wertvolle Hinweise bedanken.
- 15 Die Stadt Dresden hat aus dem Einwohnermelderegister eine Stichprobe von 4.000 Dresdner Bürgerinnen und Bürgern gezogen. Auswahlkriterium war, dass die Befragten über 18 Jahre alt sind und ihren Hauptwohnsitz in Dresden haben. Aus dieser Liste haben wir für den Pretest 200 Personen ausgewählt, d. h. aus der alphabetisch aufsteigenden Liste wurde jede 20. Person ausgewählt. Diese wurden dann aus der Gesamtstichprobe entfernt, um doppelte Anschreiben zu vermeiden.

erhielten keinen Anreiz. Es ließ sich auf diese Weise unter kontrollierten Bedingungen testen, ob die materiellen Anreize einen signifikanten Einfluss auf die Rücklaufquote ausüben.

Die 200 Fragebögen für den experimentellen Pretest wurden im August 2005 verschickt. Beide Gruppen erhielten per Standardpost einen Umschlag mit einem personalisierten Anschreiben, eine Anleitung zum Ausfüllen des Fragebogens, den Fragebogen, eine adressierte und frankierte Antwortpostkarte und einen ebenfalls adressierten und frankierten Rückumschlag. Das Anschreiben, die Postkarte und der Rückumschlag waren in beiden Gruppen identisch.¹⁶ An die Anleitung für die Gruppe mit monetärem Anreiz hatten wir lediglich ein Postskriptum angefügt, dass auf die 5-Euro-Banknote hinwies: *„P.S. In den Unterlagen finden Sie eine 5 Euro Banknote. Dieses Geld ist als kleines Dankeschön und als Aufwandsentschädigung für Ihre Mühe gedacht. Egal ob Sie sich dazu entscheiden, den Fragebogen auszufüllen oder nicht, das Geld können Sie auf jeden Fall behalten.“*

Da die Fragebögen anonym zurückgesendet wurden, musste aus in ihnen ersichtlich sein, ob sie durch jemanden aus der Gruppe mit monetärem Anreiz oder aus der Gruppe ohne vorheriges Geldgeschenk ausgefüllt worden waren. Zu diesem Zweck hatten wir auf der letzten Seite des Fragebogens eine Danksagung beigefügt, die je nach Gruppenzugehörigkeit variierte. Bei der Gruppe mit Geldgeschenk lautete der Satz: *„Wir möchten Ihnen sehr herzlich für Ihre Mitarbeit an dieser wichtigen Studie danken!“*. Bei der Gruppe ohne Incentives hatten wir das *„sehr herzlich“* gegen ein *„ganz herzlich“* ausgetauscht. Bis auf das Postskriptum in der Anleitung und diesen abschließenden Satz gab es keinerlei Unterschiede. Somit sind Unterschiede in der Rücklaufquote zwischen den Gruppen ausschließlich auf das Geldgeschenk zurückzuführen und nicht auf andere Einflussfaktoren.

Empirische Befunde zum Rücklauf: Die Fragebögen waren Mitte August 2005 bei den Befragten eingegangen.¹⁷ Wie geplant wurde die Feldphase nach drei Wochen abgeschlossen und es wurde nicht durch Erinnerungsschreiben nachgefasst. Von den insgesamt 199 Probanden wurden 79 ausgefüllte Fragebögen zurückgesandt – dies entspricht einer Gesamt-Rücklaufquote von 39,7%.¹⁸ Die Rücklaufquote setzt sich wie folgt zusammen: Von den Befragten, die kein Geld erhielten, schickten 28 Personen (also 28%) den ausgefüllten Fragebogen zurück. Hingegen sandten 51 Personen (also 51,5%), die 5 Euro erhalten hatten, den Fragebogen ausgefüllt zurück. Fünf Personen aus der Experimentalgruppe haben den Fragebogen unausgefüllt, aber zusammen mit der 5-Euro-Banknote zurückgeschickt. Eine Person hat den ausgefüllten Fragebogen zusammen mit der 5-Euro-Banknote retourniert. 42 Personen haben das Geldgeschenk somit ohne Gegenleistung behalten.

16 Die Befragten wurden gebeten, den ausgefüllten Fragebogen anonym mit Rückumschlag zurück zu senden und getrennt davon die Antwortpostkarte mit ihrem Namen zu versehen und ebenfalls an uns zurück zu senden. So konnten wir einerseits die Anonymität der Antworten garantieren und andererseits den Rücklauf kontrollieren.

17 Zur Kontrolle des Posteingangs haben wir parallel zum Anschreiben an die Befragten die Unterlagen auch an die in Dresden wohnenden Mitarbeiter des Projekts geschickt. Deren Rückmeldungen gehen selbstverständlich nicht in den hier berichteten Rücklauf bzw. in die Auswertungen ein.

18 Eine Briefsendung kam mit dem Vermerk „unbekannt verzogen“ zurück. Eine Ersatzperson wurde nicht angeschrieben. Die weggezogene Person gehörte zur Gruppe, die das Geldgeschenk erhalten hatte. Die Gesamtstichprobe des Pretests reduzierte sich dadurch auf 199 Befragte.

Der in Tabelle 1 dokumentierte Zusammenhang zwischen Geldbeigabe und Antwortbereitschaft ist hoch signifikant ($p = 0,001$). Der sehr deutliche Unterschied spricht für den Einsatz monetärer Anreize. Es sei nochmals betont, dass diese Rücklaufquoten in einer Befragung zu einem heiklen Thema, mit einem relativ langen Fragebogen und ohne Nachfassen erreicht wurden.

Tabelle 1 Teilnahmebereitschaft in Abhängigkeit von der Vergabe eines monetären Incentives (in Klammern: Zeilenprozente)

	Teilnahme	Verweigerung
Kein Geld beigelegt	28 (28%)	72 (72%)
Geldgeschenk beigelegt	51 (51,5%)	48 (48,5%)
Gesamt	79 (39,7%)	120 (60,3%)

Chi-Quadrat = 11,491 ($p = 0,001$, 2-seitig); $df = 1$; $\Phi = 0,240$ ($p = 0,001$)

In einem zweiten Schritt wollen wir analysieren, ob die monetären Anreize den Rücklauf unabhängig von bestimmten Merkmalen der Befragten erhöhen, oder ob das beigelegte Geld bei Befragten mit bestimmten Merkmalen einen anderen (höheren oder niedrigeren) Einfluss auf die Antwortbereitschaft hat als bei Befragten mit anderen Merkmalen. Dieses Unterfangen ist deswegen schwierig, weil man wie fast immer bei Studien zum Antwortverhalten keinerlei Informationen über diejenigen Befragten hat, die ihre Teilnahme verweigern (Koch 1998). Da die Verweigerer den Fragebogen nicht zurückgeschickt haben, können wir auch nicht auf deren demographische Daten zurückgreifen. Die einzige Information, die wir im vorliegenden Experiment über die Verweigerer besitzen, ist deren Geschlecht. Weil wir aus den Vornamen in der Adressenliste wissen, wie viele Frauen und Männer wir angeschrieben haben und die retournierten Fragebögen eine Frage nach dem Geschlecht der Befragungsperson beinhalten, konnten wir aus der Differenz der insgesamt angeschriebenen Frauen und Männer in der Experimental- und Kontrollgruppe und der Zahl der Frauen und Männer, die Fragebögen retourniert haben, bestimmen, wie viele der angeschriebenen Frauen und Männer aus den respektiven Gruppen geantwortet haben.¹⁹

19 Weiterhin kennen wir den Wohnort der Befragten (Postleitzahlenbereich und Strasse). Theoretisch wäre es somit möglich, vom Wohnort auf das ungefähre Einkommen und das Alter der Befragten zu schließen, indem man es an das Durchschnittseinkommen und das Durchschnittsalter der Bewohner des Stadtteils approximiert. Da die Postleitzahlbereiche in Dresden jedoch nicht mit den Stadtteilen übereinstimmen, für die die Statistikstelle der Stadt Dresden die relevanten Daten erhebt, erscheint eine solche Approximation als zu ungenau.

Tabelle 2 Teilnahmebereitschaft nach Geschlecht und Einsatz von Incentives
(in Klammern: Zeilenprozente)

	Männer			Frauen		
	Teilnahme	Verweigerung	N	Teilnahme	Verweigerung	N
Fragebogen mit Geldgeschenk	21 (48,8%)	22 (51,2%)	43	30 (53,6%)	26 (46,4%)	56
Fragebogen ohne Geldgeschenk	13 (28,3%)	33 (71,7%)	46	15 (27,8%)	39 (72,2%)	54
Gesamt	34 (38,2%)	55 (61,8%)	89 (100%)	45 (40,9%)	65 (59,1%)	110 (100%)

Insgesamt haben wir 110 Frauen und 89 Männer angeschrieben. Die Versuchsgruppe, deren Mitglieder einen Geldschein im Wert von 5 Euro erhalten haben, setzt sich dabei aus 56 Frauen und 43 Männern zusammen. Die Kontrollgruppe ohne monetäres Geschenk besteht aus 54 Frauen und 46 Männern (siehe Tabelle 2). Betrachtet man die gesamte Zahl der Verweigerer unabhängig vom Geldgeschenk, so stellt sich heraus, dass prozentual geringfügig mehr Männer (Verweigerungsquote: 62% aller Männer) als Frauen (59%) die Teilnahme an der Studie verweigert haben. Der Zusammenhang zwischen Geschlecht und Antwortverweigerung ist jedoch nicht signifikant (Chi-Quadrat = 0,151 ($p = 0,771$ 2-seitig); $df = 1$; Phi = 0,028 ($p = 0,698$)).

Aufschlussreicher sind allerdings die Verweigerungsraten, wenn man die Gruppen zusätzlich nach Erhalt des Geldes unterscheidet. Von den angeschriebenen Männern, denen zum Fragebogen ein Geldschein beigelegt wurde, verweigerten rund 51% die Teilnahme. In der Kontrollgruppe, deren Mitglieder kein Geldpräsent erhalten haben, beträgt der Anteil männlicher Verweigerer hingegen rund 72%. Bei den Frauen sieht der Unterschied in der Verweigerungsrate zwischen der Versuchs- und der Kontrollgruppe sehr ähnlich aus. Wurde von Anfang an Geld beigelegt, so verweigerten 46% der Frauen die Teilnahme. Ohne den Einsatz eines monetären Anreizes steigt die Verweigerungsquote ebenso deutlich wie bei den Männern auf 72%.

Tabelle 3 Vorhersage der Teilnahme am Pretest
(Odds Ratio, geschätzt mit logistischer Regression)

Männer	0,911
Monetärer Anreiz	2,727***
Nagelkerke R ²	0,077
N	199

Chi-Quadrat-Wert des Likelihood-Ratio Tests: 11,72 ($p = 0,003$);

*** $p \leq 0,001$

Ob die Geldbeigabe die Verweigerungsrate unabhängig vom Geschlecht minimiert, kann auch mittels einer logistischen Regression überprüft werden. Die abhängige Variable ist hier die Teilnahme am Pretest. Die unabhängigen Variablen sind das Geschlecht (Referenzkategorie: Frauen) und die

Beigabe von Incentives. Die in Tabelle 3 dokumentierten Schätzungen zeigen, dass das Geschlecht *ceteris paribus* keinen signifikanten Einfluss auf die Teilnahmebereitschaft ausübt. Ganz anders gestaltet sich die Sache, wenn man den Einfluss des monetären Anreizes auf die Teilnahmebereitschaft betrachtet und dabei das Geschlecht kontrolliert. Es zeigt sich, dass monetäre Incentives die Antwortrate in statistisch signifikanter Weise erhöhen. Aus Tabelle 2 wird ersichtlich, dass die Verweigerungsrate ohne Geldpräsent bei Männern und Frauen bei knapp über 70% liegt. Beim Einsatz monetären Anreizes liegt die Verweigerungsrate hingegen bei Frauen bei 46% und bei den Männern bei 51%. Obwohl dieser Unterschied nicht signifikant ist ($\Phi = 0,047$, $p = 0,640$) scheint es so, dass Frauen auf die Incentives etwas stärker reagieren. Dies eröffnet Raum für die Frage, ob Frauen allenfalls stärker auf Incentives mit reziprokem Verhalten reagieren bzw. die Norm der Reziprozität stärker verinnerlicht haben als Männer. Dieser Frage soll allerdings hier nicht weiter nachgegangen werden, da wir keinen signifikanten Geschlechtsunterschied festgestellt haben.

5 Zusammenfassung und Diskussion

Theoretische Überlegungen zur Teilnahmeverweigerung bei postalischen Befragungen aus Sicht der Austausch- bzw. der Rational Choice-Theorie und einige Studien und Experimente zu diesem Thema lassen es plausibel erscheinen, dass beim Erstkontakt beigelegte (monetäre) Geschenke die Responseraten deutlich erhöhen. Für den anglo-amerikanischen Raum ist die positive Wirkung im Voraus gegebener monetärer Anreize gut belegt. Hingegen liegen für den deutschsprachigen Raum nur wenige Studien vor und diese verwenden zudem geldnahe Anreize statt Bargeld. Somit fehlen bislang im deutschsprachigen Raum Studien, welche die Wirkung monetärer Anreize nachweisen konnten. Des Weiteren ist die Beigabe monetärer Anreize natürlich ein ernstzunehmender Kostenfaktor. Es ist letztlich eine empirische Frage, ob materielle Geschenke wirklich den Rücklauf erhöhen und damit die mit ihnen verbundenen Mehrkosten kompensieren können. Ein von den Autoren durchgeführtes Methodenexperiment sollte dazu Antworten liefern.

Das Experiment wurde im Rahmen des Pretests zu einer postalischen Bevölkerungsbefragung in Dresden zum Thema Kriminalität und kriminelles Handeln durchgeführt. Einer Versuchsgruppe von 100 zufällig ausgewählten Erwachsenen wurde zusammen mit dem Fragebogen eine 5-Euro-Banknote als Geschenk beigelegt. Die Probanden in einer gleich großen Kontrollgruppe erhielten keinen monetären Anreiz. Da der Fragebogen und das Anschreiben für beide Gruppen identisch waren, sind unterschiedliche Responsraten ausschließlich auf die Incentives zurückzuführen. Die erzielten Rücklaufquoten der beiden Gruppen unterscheiden sich signifikant voneinander. Während nur 28% der Befragten aus der Kontrollgruppe (ohne Anreiz) einen ausgefüllten Fragebogen zurückgesandt haben, haben 52% der Befragten aus der Versuchsgruppe (mit Anreiz) geantwortet. Hierbei ist zu bedenken, dass diese Rücklaufquoten ohne Nachfassen in Form eines Erinnerungsschreibens erzielt wurden. Bisherige Erfahrungen legen es aber nahe, dass auf ein Erinnerungsschreiben eine ähnliche Rücklaufquote wie auf das Erstanschreiben zu erwarten ist.

Somit wäre im günstigen Falle für die geplante postalische Bevölkerungsbefragung in Dresden eine Responserate von insgesamt fast 75% zu erwarten (vgl. auch Hippler 1988; Mangione 1998). Die signifikant höhere Rücklaufquote beim Einsatz von Geschenken würde damit die Kosten für die Incentives kompensieren. Unsere Ergebnisse zeigen, dass die Beigabe monetärer Anreize zum Fragebogen (also unabhängig davon, ob die Befragten tatsächlich antworten) auch im deutschsprachigen Raum die Teilnahmebereitschaft signifikant erhöht. Dies spricht gegen eine kulturspezifische Wirkung von Incentives.

Die Ergebnisse des Methodenexperimentes legen es auch nahe, dass monetäre Anreize bei postalischen Befragungen zu relativ heiklen Themen, wie der Preisgabe eigenen kriminellen Handelns, die Rücklaufquote signifikant erhöhen. Man kann somit vermuten, dass die monetären Incentives nicht nur unabhängig von den kulturellen Rahmenbedingungen, sondern auch von der thematischen Ausrichtung der Befragung wirken. Diese Ergebnisse müssen jedoch durch weitere Forschung abgesichert werden. Auch zur Interaktion von Incentives zu anderen Untersuchungselementen (z. B. Thema der Untersuchung, Anforderungsgrad für die Befragten, Anzahl der Kontaktaufnahmen) liegen für unseren kulturellen Kontext erst wenige Informationen vor.

Im Rahmen des Methodenexperimentes konnte ebenfalls ansatzweise überprüft werden, ob die beigelegten Anreize bestimmte Bevölkerungsgruppen mehr zur Teilnahme motivieren als andere. Wenn dies so wäre, dann sollten die potentiellen Vorteile durch die Erhöhung der Rücklaufquote sehr genau gegen die Nachteile abgewogen werden, die durch eine mögliche Verzerrung der Stichprobe entstehen können. In unserem Experiment hatten wir nur Informationen über das Geschlecht der Verweigerer. Das beigelegte Geldgeschenk erhöhte die Kooperationsbereitschaft von Männern und Frauen gleichermaßen. Zwar werden die Frauen durch die Geldgabe eher zur Teilnahme motiviert als die Männer; aber die Unterschiede zwischen den Geschlechtern sind nicht signifikant. Das Methodenexperiment unterstreicht insgesamt die Nützlichkeit monetärer Geschenke in postalischen Befragungen: Die Teilnahmebereitschaft wird überzufällig erhöht, weil die Kosten der Teilnahme für die Befragten kompensiert werden. Dabei scheinen die Geschenke nicht zu einem „sample selection bias“ bezüglich des Geschlechts zu führen.

Trotz dieser eindrücklichen Ergebnisse besteht noch ein großer Forschungsbedarf. Der generell positive Einfluss monetärer Anreize auf die Teilnahmebereitschaft kann zwar als gesichert angesehen werden, aber es bleiben weiterhin einige Fragen unbeantwortet: Es ist z. B. noch offen, ob monetäre Anreize einige Bevölkerungsgruppen mehr ansprechen als andere (im vorgestellten Experiment kann dies für das Geschlecht verneint werden). Um diese Frage zu beantworten, benötigt man demographische Informationen über die Verweigerer. Gerade dies stellt eine große Anforderung an das Design von Methodenexperimenten dar. Noch eingehender untersucht werden sollte auch, welchen Einfluss der Wert und die Beschaffenheit eines materiellen Anreizes auf den Rücklauf in postalischen Befragungen besitzen. Besonders interessant ist in diesem Zusammenhang die Frage, ob die Teilnahmebereitschaft linear mit dem Wert der Geschenke steigt oder ob ein abnehmender Nutzen bis zu einem Punkt hin zu erwarten ist, ab dem ein höheres Incentive die Antwortbereitschaft nicht weiter erhöht. Dazu sind Methodenexperimente geeignet, in denen es

mehrere Kontrollgruppen von hinreichender Größe gibt, denen monetäre Anreize in unterschiedlicher Höhe dargeboten werden. Schließlich muss weitere Forschung auf diesem Gebiet die Fragen beantworten, ob monetäre Anreize auch in weniger umfangreichen und anforderungsreichen Befragungen die Teilnahmebereitschaft erhöhen und ob der Effekt der Anreize auch in anderen kulturellen Kontexten als im anglo-amerikanischen und deutschsprachigen Raum festzustellen ist (etwa in Asien).

Literatur

- Armstrong, J., 1975: Monetary Incentives in Mail Surveys. *Public Opinion Quarterly* 39: 111–116.
- Arzheimer, K. und M. Klein, 1998: Die Wirkung materieller Anreize auf den Rücklauf einer postalischen Panelbefragung. *ZA-Information* 43: 6–31.
- Becker, R. und R. Günther, 2004: Selektives Antwortverhalten bei Fragen zum delinquenten Handeln. Eine empirische Studie über die Wirksamkeit der ‚sealed envelope technique‘ bei selbst berichteter Delinquenz mit Daten des ALLBUS 2000. *ZUMA-Nachrichten* 54: 39–59.
- Becker, R., 2006: Selective Response to Questions about Delinquency. *Quality & Quantity* 40: 483–498.
- Berger, F., 2006: Zur Wirkung unterschiedlicher materieller incentives in postalischen Befragungen. Ein Literaturbericht. *ZUMA-Nachrichten* 58: 81–100.
- Berry, S. H. und D. E. Kanouse, 1987: Physician Response to a Mailed Survey. An Experiment in Timing of Payment. *Public Opinion Quarterly* 51: 102–114.
- Birnkrant, G., 2003: Ehrenamt in kulturellen Institutionen im Vergleich zwischen den USA und Deutschland. Dissertation. Pädagogische Hochschule Ludwigsburg.
- Blau, P., 1964: *Exchange and Power in Social Life*. New York: Wiley.
- Church, A. H., 1993: Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis. *Public Opinion Quarterly* 5: 62–79.
- De Leeuw, E. D., 2001: Reducing Missing Data in Surveys: An Overview of Methods. *Quality & Quantity* 35: 147–160.
- Diekmann, A. und B. Jann, 2001: Anreizformen und Ausschöpfungsquoten bei postalischen Befragungen. Eine Prüfung der Reziprozitätshypothese. *ZUMA-Nachrichten* 48: 19–27.
- Dillman, D. A., 2000: *Mail and Internet Surveys: The Tailored Design Method*. Second Edition. New York: Wiley.
- Dillman, D. A., 1978: *Mail and Telephone Surveys. The Total Design Method*. New York: Wiley.
- Dodd, T., 1998: Incentive Payments on Social Surveys: A Summary of Recent Research. *Survey Methodology Bulletin* 43: 23–27.
- Esser, H., 1974: Der Befragte. S. 107–145 in: J. van Koolwijk u.a. (Hg.), *Erhebungsmethoden: Die Befragung. Techniken der empirischen Sozialforschung* (4. Band). München: Oldenbourg.
- Esser, H., 1986: Können Befragte lügen? Zum Konzept des „wahren Wertes“ im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 38: 314–336.
- Esser, H., 1991: Alltagshandeln und Verstehen. Zum Verhältnis von erklärender und verstehender Soziologie am Beispiel von Alfred Schütz und „Rational Choice“. Tübingen: Mohr.
- Fox, R., M. Crank und J. Kim, 1988: Mail Survey Response Rate: A Meta-Analysis of Selected Techniques for Inducing Response. *Public Opinion Quarterly* 52: 467–491.
- Furse, D. H. und D. W. Stewart, 1982: Monetary Incentives Versus Promised Contribution to Charity: New Evidence on Mail Survey Response. *Journal of Marketing Research* 19: 375–380.
- Furse, D. H., D. W. Stewart und D. L. Rados, 1981: Effects of Foot-In-The-Door, Cash Incentives, and Follow-ups on Survey Response. *Journal of Marketing Research* 18: 473–478.
- Gallup, G., 1940: *The Pulse of Democracy: The Public Opinion Poll and How it Works*. New York: Simon & Schuster.
- Goyder, J., 1987: *The Silent Minority*. Boulder: Westview.
- Grilliches, Z., B. H. Hall und J. A. Hausman, 1978: Missing Data and Self-selection in Large Panels. *Annales De L'Insee* 30–31: 138–176.

- Groves, R. M., E. Singer und A. Corning, 2000: Leverage-Saliency Theory of Survey Participation. Description and Illustration. *Public Opinion Quarterly* 64: 299-308.
- Hansen, R. A., 1980: A self-perception Interpretation of the Effect of Monetary and Nonmonetary Incentives on Mail Survey Respondent Behavior. *Journal of Marketing Research* 17: 77-83.
- Harkness, J., P. Mohler, M. Schneid und B. Christoph, 1998: Incentives in Two German Mail Surveys 1996/97 Et 1997. S. 201-218 in: A. Koch und R. Porst (Hg.), *Nonresponse in Survey Research*. ZUMA, Mannheim: ZUMA-Nachrichten Spezial 4.
- Hill, D. H. und R. J. Willis, 2001: Reducing Panel Attrition. A Search for Effective Policy Instruments. *The Journal of Human Resources* 3: 416-438.
- Hippler, H.-J., 1988: Methodische Aspekte schriftlicher Befragungen: Probleme und Forschungsperspektiven. *planung & analyse* 6: 244-248.
- Homas, G. C., 1961: *Social Behavior. Its Elementary Forms*. New York, N.Y.: Harcourt, Brace & World.
- James, T. und R. Bolstein, 1990: The Effect of Monetary Incentives and Follow-up Mailings on the Response Rate and Response Quality in Mail Surveys. *Public Opinion Quarterly* 54: 346-361.
- Klein, S. und R. Porst, 2000: Mail Surveys. Ein Literaturbericht. ZUMA-Technischer Bericht, 10/2000.
- Koch, A., 1998: Wenn „mehr“ nicht gleichbedeutend mit „besser“ ist: Ausschöpfungsquoten und Stichprobenverzerrungen in allgemeinen Bevölkerungsumfragen. *ZUMA-Nachrichten* 42: 66-93.
- Krebs, D., 1995: Selbstselektion: Demographisches oder attitudinales Problem. *ZA-Information* 36: 114-125.
- Little, R. J. und D. B. Rubin, 1987: *Statistical Analysis with Missing Data*. New York: Wiley.
- Mangione, T. W., 1998: Mail Surveys. S. 399-427 in: L. Bickman und D. J. Rog (Hg.), *Handbook of applied social research methods*. Thousand Oaks: Sage.
- Martin, E., D. Abreu und F. Winters, 2001: Money and Motive: Effects of Incentives on Panel Attrition in the Survey of Income and Program Participation. *Journal of Official Statistics* 17: 267-284.
- Martinez-Ebers, V., 1997: Using Monetary Incentives with Hard-to-reach Populations in Panel Surveys. *International Journal of Public Opinion Research* 9: 77-86.
- Mauss, M., 1984 (franz. Orig. 1925): *Die Gabe. Form und Funktion des Austauschs in archaischen Gesellschaften*. Frankfurt: Fischer.
- Mehlkop, G. und R. Becker, 2004: Soziale Schichtung und Delinquenz. Eine empirische Anwendung eines Rational Choice-Ansatzes mit Hilfe von Querschnittsdaten des ALLBUS 1990 und 2000. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 56: 95-126.
- Nederhof, A. J., 1983: The Effects of Material Incentives in Mail Surveys: Two Studies. *Public Opinion Quarterly* 47: 103-111.
- Oswald, M., 2004: *Gabe und Gewalt. Studien zur Logik und Poetik der Gabe in der frühhöfischen Erzählliteratur*. Göttingen: Vandenhoeck & Ruprecht.
- Porst, R., 1999: Thematik oder Incentives? Zur Erhöhung der Rücklaufquoten bei postalischen Befragungen. *ZUMA-Nachrichten* 45: 72-87.
- Pötter, U. und U. Rendtel, 1993: Über Sinn und Unsinn von Repräsentativitätsstudien. *Allgemeines Statistisches Archiv* 77: 269-280.
- Prüfer, P. und M. Rexroth, 2000: Zwei-Phasen-Pretesting. S. 203-219 in: P. Mohler und P. Lüttinger (Hg.), *Querschnitt. Festschrift für Max Kaase*. Mannheim: Zentrum für Umfragen, Methoden und Analysen.
- Rendtel, U., 1989: Über den Einfluss der Panelselektivität auf Längsschnittanalysen. *Vierteljahreshefte zur Wirtschaftsforschung* 58: 45-61.
- Rendtel, U., 1995: *Panelausfälle und Panelrepräsentativität*. Frankfurt am Main: Campus.
- Reuband, K.-H., 2001: Möglichkeiten und Probleme des Einsatzes postalischer Befragungen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 53: 307-333.
- Riphahn, R.T. und O. Serfling, 2002: *Item Non-Response on Income and Wealth Questions*. IZA-Working Paper. Bonn: IZA.
- Schnell, R., 1997: *Nonresponse in Bevölkerungsumfragen*. Opladen: Leske+Budrich.
- Schnell, R., P. B. Hill und E. Esser, 1999: *Methoden der empirischen Sozialforschung*. München: Oldenbourg.
- Schräpler, J.-P., 2001: *Respondent Behaviour in Panel Studies. A Case Study of the German Socio-Economic Panel (GSOEP)*. DIW-Discussion Paper 244. Berlin: DIW.
- Shettle, C. und G. Mooney, 1999: Monetary Incentives in US Government Surveys. *Journal of Official Statistics* 5: 231-250.
- Simmons, E. und A. Wilmot, 2004: Incentive Payments on Social Surveys: A Literature Review. *Survey Methodology Bulletin* 53: 1-11.

- Singer, E., 1998: Incentives for Survey Participation: Research on Intended and Unintended Consequences. ZUMA-Nachrichten 42: 7-29.
- Singer, E., J. Van Hoewyk und M. Maher, 2000: Experiments with Incentives in Telephone Surveys. Public Opinion Quarterly 64: 171-188.
- Trussell, N. und P.J. Lavrakas, 2004: The Influence of Incremental Increases in Token Cash Incentives on Mail Survey Response. Public Opinion Quarterly 68: 349-67.
- Warriner, K., J. Goyder, H. Gjertsen, P. Hohner und K. McSpurren, 1996: Charities, No; Lotteries, No; Cash, Yes: Main Effects and Interactions on a Canadian Incentives Experiment. Public Opinion Quarterly 60: 542-562.
- Yammarino, F.J., S.J. Skinner und T.L. Childers, 1991: Understanding Mail Survey Response Behavior. A Meta-Analysis. Public Opinion Quarterly 55: 613-639.
- Yu, J. und H. Cooper, 1983: A Quantitative Review of Research Design Effects on Response Rates to Questionnaires. Journal of Marketing Research 20: 36-44.

Korrespondenzadressen: Guido Mehlkop
Technische Universität Dresden
Institut für Soziologie
Mommsenstraße 13
01069 Dresden
Guido.Mehlkop@tu-dresden.de

Rolf Becker
Universität Bern
Institut für Erziehungswissenschaften
Abteilung Bildungssoziologie
Muesmattstraße 27
CH-3012 Bern
Rolf.Becker@edu.unibe.ch

Die Aussiedler- Stichprobenziehung

Sampling *Aussiedler*

Kurt Salentin

Zusammenfassung

Deutschstämmige Aussiedler bilden die größte Zuwanderergruppe in der Bundesrepublik. Obwohl Integrationsprobleme bekannt sind, fehlt es an amtstatistischen und Erhebungsdaten, weil eindeutige Identifikationsmerkmale bislang fehlen. Vor- und Nachteile bisheriger Stichprobentechniken für Primärerhebungen (Random Route, Zuganglisten in Übergangsheimen, Schneeballstichproben, onomastische Verfahren) werden erörtert. Der Beitrag stellt dann zum einen ein duales Verfahren vor, das Aussiedler in Einwohnermeldedateien und anderen Registern (a) an spezifischen Doppelstaatsangehörigkeiten und (b) an Geburtsorten erkennt. Er schildert Details des Verfahrens. Anhand einer Anwendung auf Geburtsorte im Bielefelder Melderegister und verschiedener statistischer Quellen wird die Ausschöpfung und der erwartbare Anteil falsch positiver Erkennungen (Beifang) abgeschätzt. Gegenwärtig ist eine Ausschöpfung von ca. 95% der Zielgruppe zu erwarten. Je nach Herkunftsland muss mit Beifang im einstelligen Prozentbereich gerechnet werden, für Polen ist historisch bedingt mit einem erhöhten Anteil zu rechnen. Auf die technisch aufwendige Identifikation des Geburtsortes sollte nicht verzichtet werden, weil ca. die Hälfte der Aussiedler nur die deutsche Staatsangehörigkeit besitzt. Zum anderen wird die Nutzung aussiedlerspezifischer Vor- und Zunamen zur Stichprobenziehung erörtert. In einer Melderegisterstichprobe nach Geburtsort überwiegen deutsche Namen. Allerdings existieren sowohl deutsche Vor- als auch Nachnamen, die Aussiedler relativ gut von autochthonen Deutschen differenzieren. Bis auf ein späteres Einreisejahr und gewisse Abweichungen in Netzwerk- und Kontaktmustern unterscheiden sich die Träger herkunftslandtypischer Namen nicht von den Trägern deutscher Namen.

Abstract

Aussiedler (resettlers of German descent) are the largest immigrant group in contemporary Germany. Though economic and social integration difficulties on their part are known, pertinent survey and official statistical data are still lacking due to the absence of clear-cut criteria discerning them from the autochthonous population. The article discusses advantages and disadvantages of *Aussiedler* sampling techniques hitherto used (random route, address lists kept in transit camps, snowball sampling, name-based techniques). An alternative method based on (a) dual citizenship patterns and (b) specific birthplaces in the population register is then introduced and discussed. Estimates of target population exhaustion and the rate of false identification are calculated from a pilot implementation in the city of Bielefeld. Currently, an exhaustion of 95% may be expected with a rate of false identification of below 10%, varying between former countries of residence. Though cumbersome in technical terms, the birthplace criterion is superior to citizenship-based identification, as about half the *Aussiedler* population cannot be discerned via combinations of German and second citizenships. Further, sampling based on names typical to the group is tested. Though *Aussiedler* names predominantly sound German, a range of names exists that distinguish *Aussiedler* fairly well from autochthonous Germans. Apart from particularities concerning the period of immigration and network and social contact patterns, bearers of German names vs. names typical for countries of former residence do not display significantly different social characteristics.

1 Zielgruppe und bisherige Stichprobenverfahren¹

Der vorliegende Beitrag soll Möglichkeiten und Grenzen der verfügbaren sozialwissenschaftlichen Stichprobenverfahren für deutschstämmige Aussiedler in der Bundesrepublik aufzeigen. Drei Umstände gebieten eine solche Erörterung. 1. Aussiedler bilden hier die größte Einwanderergruppe. Zwischen dem 1.1.1950 und dem 31.12.2004 sind nach Angaben des Bundesverwaltungsamts 4.446.360 Aussiedler einschließlich ihrer Angehörigen ins Bundesgebiet eingereist; seit 1989 sind es 2.823.724 Personen.² Alleine der Umfang der 2.298.938 aus der ehemaligen UdSSR Stammen dieses Zeitraums erreicht beinahe die Stärke der ca. 2.430.000 Türkischstämmigen, die am 31.12.2004 im Bundesgebiet als Ausländer registriert sind (1.764.318 Personen) oder als Eingebürgerte leben (ca. 666.000 Personen).³ 2. In den Sozial- und Bildungswissenschaften herrscht weitgehend Einvernehmen darüber, dass die Platzierung vieler Aussiedler, insbesondere der Spätaussiedler seit 1993, auf den sozialen und ökonomischen Statushierarchien aufgrund sprachlicher und qualifikatorischer Defizite sowie anderer migrantentypischer Handicaps prekär ist (z. B. Seifert 2001). 3. Gleichzeitig bestehen in Wissenschaft und Politik erhebliche Kenntnisdefizite, da Aussiedler als Deutsche in der amtlichen Statistik, etwa der Bildungs-, Sozialhilfe- und mit Ausnahmen auch der Arbeitsmarktstatistik, nicht gesondert ausgewiesen werden. In der quantitativen Sozialforschung gelten Aussiedler als schwer erreichbar, weil Auswahlverfahren, die sich nicht an klaren Kriterien wie der Staatsangehörigkeit orientieren können, keine repräsentativen Aussagen erwarten lassen.

Als Grundgesamtheit sollen hier alle nach dem Bundesvertriebenengesetz als Aussiedler oder Spätaussiedler anerkannten Personen gelten.⁴ Wie immer in Fragen der Migrations- und Minderheitenforschung hängt es von der Fragestellung ab, ob auch die Nachkommen einbezogen werden sollen, d.h. ob eine über die juristische Kategorie hinausgehende Definition sinnvoll ist.

Generell besitzen die Kommunalverwaltungen bestmögliche Ausgangsbedingungen zur Identifikation des Personenkreises mit Migrationshintergrund, da ihnen mehr Informationen zur Verfügung stehen als die Menge der durch das Melderecht zur Übermittlung an Dritte freigegebenen Variablen. Sie sind daher u. a. in der Lage, Aussiedler zu erkennen sowie allgemein aus dem Haushaltskontext einen Migrationshintergrund auch bei solchen Personen zu erschließen, die keine direkten Migrationsindikatoren wie ausländische Geburtsorte oder Staatsangehörigkeiten

1 Der Autor dankt dem Amt für Bürgerberatung in Verbindung mit dem Amt für Stadtplanung, Statistik und Wahlen der Stadt Bielefeld für die Bereitstellung dreier anonymisierter Melderegisterauszüge.

2 Bundesverwaltungsamt, www.bva.bund.de/imperia/md/content/abteilungen/abteilungiii/iiist/28.pdf, S. 2. (13.12.2005).

3 Quellen: Statistische Jahrbücher seit 1970, diverse schriftliche Mitteilungen des Statistischen Bundesamts.

4 Nach § 2 (3) in Verbindung mit § 1 (1) des Gesetzes unterscheiden sich Aussiedler von *anderen* Vertriebenen dadurch, dass sie die deutschen Siedlungsgebiete „nach Abschluß der allgemeinen Vertriebensmaßnahmen“ verlassen haben. Die administrative Unterscheidung zwischen Aussiedlern und Spätaussiedlern beruht im wesentlichen auf der Ausreise aus der ehemaligen Sowjetunion und dem Baltikum vor dem 31.12.1992. Nichtdeutsche Ehegatten und Abkömmlinge von Aussiedlern werden nach § 4 (3) mit der Einreise in die Bundesrepublik Spätaussiedler. Da die Verreibungen in der Mitte der 1950er Jahre endeten, stellt sich die Frage nach einer Differenzierung von Aussiedlern und anderen Vertriebenen nur bei den heute über 50Jährigen.

besitzen (aus juristischen Gründen nur bei Minderjährigen). Einige deutsche Städte haben begonnen, durch manuelle Einzelfallprüfung den Zuwanderungshintergrund ihrer Einwohner vollständig zu bestimmen (Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration 2005: 551; ergänzt um mündliche Auskünfte eines Vertreters des Deutschen Städtetages vom 16.1.2006). Doch leider wird diese sehr präzise Informationsquelle auf absehbare Zeit nicht für wissenschaftliche Zwecke nutzbar sein. Einerseits sind es meist nicht die Einwohnermeldeämter, sondern die Planungs-, Statistik- und Stadtforschungsämter, die diese wertvolle Pionierarbeit leisten. Im Gegensatz zum Meldewesen verfügt dieser Funktionsbereich über keinerlei Rechtsgrundlage einer Weitergabe von Individualdaten, so dass Forscher allenfalls kleinräumige Aggregatdaten werden nutzen können. (Bereits die juristische Absicherung der Nutzung von Meldedaten durch Planungsämter ist nicht unkompliziert.) Andererseits ist derzeit unklar, ob man sich auf ein einheitliches Konstruktionsverfahren für den Migrationshintergrund mit vergleichbaren Ergebnissen verständigen und inwiefern es angesichts des enormen Aufwands zu einer flächendeckenden Anwendung des Verfahrens kommen wird. Bis auf weiteres müssen Sozialwissenschaftler deshalb andere Wege einschlagen.

Bisher sind im Rahmen quantitativer Studien mehrere Zugangsstrategien für Aussiedler verfolgt worden (auf qualitative Studien kann hier nicht eingegangen werden).

1. Das nach zwei Wellen (1991, 1994) nicht fortgesetzte Aussiedlerpanel des Bundesinstituts für Bevölkerungsforschung (Mamney/Schiener 1996: 146ff., 1998: 11ff.) stellte in nach regionaler Schichtung ausgewählten Übergangsheimen eine Stichprobe von Neuzuwanderern zusammen. In ähnlicher Weise verfuhr Swiaczny (2000) in einer auf Mannheim beschränkten Lokalstudie. Mamney/Schiener diskutierten und verwarfen Alternativen: Ein zentrales Aussiedlerregister existiert nicht. In den Einwohnerregistern der kommunalen Meldebehörden ist kein personenbezogenes Merkmal „Aussiedler“ vorhanden. Eine Adressliste des mit dem Lastenausgleichsverfahren befassten Bundesausgleichsamts konnte aus Datenschutzgründen nicht genutzt werden. Das von Befragungsinstituten sonst bevorzugte Random-Route-Verfahren kam (und kommt auch heute) wegen des geringen Anteils von Aussiedlerhaushalten nicht in Frage. Man sah keine Möglichkeit, anstelle einer Zugangsstichprobe eine repräsentative Bestandsstichprobe einschließlich früherer Zugangskohorten zu erstellen. Weil der Erstkontakt bei einer Stichprobe aus Personen ohne eigene Wohnung unternommen wurde, stellte sich enorme Panelmortalität (über 50% von Welle 1 zu Welle 2) ein.
2. Die Aussiedlerhaushalte in der Zuwandererstichprobe des Sozio-oekonomischen Panels (SOEP, s. Schupp/Wagner 1995; Haisken-DeNew/Frick 2005) wurden vom durchführenden Institut zuvor bei anderweitiger Feldarbeit registriert, so dass kein erhöhter Random-Route-Aufwand entstand. Forschern ohne ständige Feldtätigkeit in der Wohnbevölkerung steht dieses Verfahren nicht zur Verfügung. Bei Aussiedler-Analysen bereitet das SOEP zudem unvermeidliche Probleme. Reichen die Daten für manche Forschungszwecke aus, sind sie als Grundlage für Sekundäranalysen doch immer mit inhaltlichen Einschränkungen verbunden. Daneben erzeugt die überwiegend statische Stichprobenanlage des Panels⁵ für eine sich

5 Nur Ausgründungen aus sowie individuelle Zugänge zu bestehenden Haushalten erweitern die Stichprobe.

durch kontinuierliche Zuwanderung dynamisch verändernde Gruppe Abbildungsdefizite eigener Art. So sind etwa die letzten Aussiedlerkohorten mit auch für diese Herkunftsgruppe vergleichsweise schlechten Deutschkenntnissen, unzureichender beruflicher Qualifikation und erhöhtem Armutsrisiko (Engels 2000) im SOEP zwangsläufig kaum vertreten.

3. Auf das manchmal als Notlösung gewählte Schneeballverfahren soll hier wegen dessen bekannter Unsicherheiten (Gabler 1992) nicht ausführlich eingegangen werden. Für die SOEP-Zuwandererstichprobe wurde parallel zu dem gerade beschriebenen Verfahren mit dem Schneeballsystem experimentiert. Schupp/Wagner (1995) beschreiben die Verzerrungen, die aufgrund ungleicher Netzwerkstrukturen auftraten und aufgrund derer dieses Verfahren nicht weiter eingesetzt wurde.
4. Für Untersuchungen mit eigenständiger Stichprobenziehung versprechen onomastische (namensbasierte) Erkennungsverfahren (Humpert/Schneiderheinze 2000, 2002) Abhilfe, weil viele Aussiedler russische Vor- und Familiennamen tragen. So deuten Namen wie Juri Idimetschenko recht eindeutig auf russische Herkunft hin. Doch schon bei Viktoria Kamrowskow treten Zweifel auf, und einen Helmut Kühnemeier oder Andreas Altstether kann kein Erkennungsalgorithmus als Aussiedler ausweisen.⁶ Die Methode leidet unter dem in § 94 des Bundesvertriebenengesetzes festgelegten Recht der Aussiedler, deutschsprachige Formen ihres Vor- und Familiennamens und ggf. sogar neue Vornamen anzunehmen. Eine hohe Ausschöpfung wird bei dieser Methode möglicherweise mit hohem Beifang (falsch positive Erkennung) nichtdeutscher osteuropäischer Personennamen erkauft, die nicht zu Aussiedlern gehören, während eine hohe Treffsicherheit nur bei geringer Ausschöpfung erreichbar ist. Trotzdem könnte das namensbasierte Verfahren in Verbindung mit Screenerfragen dann für eine einfache und günstige Stichprobenziehung etwa aus dem Telefonbuch gut geeignet sein, wenn die Ausschöpfung kein Kriterium und gleichzeitig gewährleistet ist, dass die Namenscharakteristik als Determinante der Ziehungswahrscheinlichkeit nicht mit sozial-kulturellen Eigenschaften oder anderen Kriteriumsmerkmalen korreliert. Diesen Fragen wendet sich Abschnitt 3 zu.
5. Für Jugendlichen- und Schülerbefragungen wurden häufig Schulen bzw. Schulklassen als Auswahlrahmen benutzt (z. B. Strobl/Kühnel 2000; Babka v. Gostomski 2001; Wetzels et al. 2001). Dabei dienen Schulstatistiken über Aussiedleranteile zur Steuerung der Auswahl von Schulen. Einrichtungen mit geringem Anteil werden aus forschungsökonomischen Erwägungen häufig unterrepräsentiert, und die Präzision der auf Lehrereinschätzungen beruhenden Hintergrunddaten ist, wie allgemein bei prozessproduzierten Daten, manchmal kaum klar zu beurteilen. Da dieses Verfahren sich nicht auf die Aussiedlerpopulation übertragen lässt, sollen weitere Einzelheiten hier übersprungen werden.

In diesem Beitrag wird zunächst ein alternatives Verfahren vorgestellt, das durch Nutzung der Staatsangehörigkeiten und des Geburtsorts im Melderegister eine hohe Ausschöpfung bei geringem Beifang realisiert. Wegen der Nutzung von Ortsnamen (Toponymen) ist hier auch von einer toponomastischen Methode die Rede. Das Melderegister erwies sich bei einer ausführlichen Erörterung denkbarer Zuwanderer-Auswahlgesamtheiten (Salentin 1999) als beste verfügbare

6 Es handelt sich um fiktive Namen, die weder in den verwendeten Stichproben noch im aktuellen Telefonbuch der Bundesrepublik existieren. Jede Ähnlichkeit mit realen Personen wäre rein zufällig.

Grundlage. Das Verfahren ist aber auch auf andere Datenbestände anwendbar, sofern Geburtsorte zugänglich sind. Das Bundesamt für Migration und Flüchtlinge hat für eine Aussiedlerbefragung 2006 teilweise Geburtsorte als Kriterium zur Auswahl aus Melderegistern verwendet (Haug/Sauer 2007).⁷

2 Toponomastische Stichprobenziehung

2.1 Grundlagen

Weil sie als Deutsche im Sinn des Grundgesetzes weitgehend beim Grenzübertritt ihren deutschen Pass erhalten, gehen Aussiedler einerseits zu keinem Zeitpunkt in die amtliche Ausländerstatistik ein. Eine Ausnahme bilden nichtdeutsche Ehegatten der Deutschstämmigen, die mit diesen seit weniger als drei Jahren verheiratet sind, sowie sonstige mitreisende Familienangehörige. Sie bleiben Ausländer und müssen sich einbürgern lassen; über ihre Zahl existieren keine verlässlichen Angaben. Als deutsche Staatsangehörige sind Aussiedler wie in den meisten Statistiken andererseits scheinbar auch in Melderegistern nicht zu identifizieren. Das Melderegister bietet jedoch bisher unbeachtete Identifikationsmöglichkeiten unter Nutzung des Merkmals der zweiten Staatsangehörigkeit und des Geburtsorts. Obwohl bereits vor der Reform des Staatsangehörigkeitsrechts im Jahr 1999 die Bundesrepublik dem Prinzip der Vermeidung von Mehrstaatlichkeit folgte, wurden erhebliche Anteile der Gruppe der Aussiedler unter Hinnahme doppelter Staatsangehörigkeit eingebürgert. Nach Auskunft des Bundesministeriums des Inneren liegen dazu keine verlässlichen Zahlen vor, doch weil die Beibehaltung oder Wiederannahme der früheren Staatsangehörigkeit für Aussiedler bis ins Jahr 1999 rechtlich nicht zu beanstanden war, ist offensichtlich ein beträchtlicher Teil der Aussiedler damals Mehrstaater geworden. (Dies ist auch einer Pressemitteilung des früheren Innenministers Schily vom 4. August 2002 zu entnehmen, der zufolge zwischen 1993 und 2000 1,2 Millionen Aussiedler die deutsche Staatsangehörigkeit erhalten haben, ohne ihre frühere aufgeben zu müssen).⁸ 1994 besaßen im Aussiedlerpanel des Bundesinstituts für Bevölkerungsforschung drei Jahre nach Einreise 63% der Haushaltsvorstände aus Polen, Rumänien und der ehemaligen Sowjetunion neben der deutschen eine weitere Staatsangehörigkeit (Mamme/Schiener 1998: 37). Zum Anteil der Mehrstaater unter den Aussiedlern liegen amtlicherseits auch für die Zeit nach 2000 keine Publikationen vor. Selbst, wenn er wegen der geänderten Gesetzeslage zurückgegangen ist, dürfte gegenwärtig eine beträchtliche Zahl der Aussiedler an der Kombination bestimmter Staatsangehörigkeiten zu erkennen sein, etwa Bundesrepublik und Russische Föderation, Bundesrepublik und Kasachstan etc.⁹ In Teil 3.3 erfolgt eine Schätzung des Anteils.

7 Der Autor dankt Frau Sonja Haug für die Überlassung eines Vortragsmanuskripts und mündliche Informationen. Leider ist ein ausführlicher Forschungsbericht zur Studie bis zum Ende des MDA-Begutachtungsverfahrens nicht publiziert.

8 Bundesministerium des Inneren, http://www.bmi.bund.de/cln_028/nn_662928/Internet/Content/Nachrichten/Archiv/Pressemitteilungen/2002/08/Schily_zum_modernen_Id_89621_de.html (16.2.2006).

9 Während seit der Novelle des Staatsangehörigkeitsrechts die *Annahme* einer fremden Staatsangehörigkeit für Deutsche rechtswidrig ist, ist die *Beibehaltung* der früheren Staatsangehörigkeit, die den meisten Ausländern bei der Einbürgerung verweigert wird, bei Aussiedlern nicht zu beanstanden.

Es verbleibt aber ein Personenkreis, der seine frühere Staatsangehörigkeit aufgegeben hat, also nur die deutsche Staatsangehörigkeit besitzt, die sog. Einstaater. Er ist wiederum an Geburtsorten in Herkunftsgebieten der Aussiedler zu erkennen und kann anhand dieser Geburtsorte identifiziert werden, weil nach dem Melderechtsrahmengesetz des Bundes und den Meldegesetzen der Länder Geburtsorte zu den zulässigen Merkmalen einer sog. Melderegister-Gruppenauskunft, also einer Stichprobe, gehören. Die Auswahl nach Geburtsort ist allerdings technisch aufwendiger als eine statische Filterung nach einem Merkmal wie dem Alter oder der Staatsangehörigkeit, weil keine simple Regel wie Staatsangehörigkeit = Russland zugrunde gelegt werden kann. Es sind Schreibfehler und -varianten (Nowoneschinka, Nowoneshenka, Nowoneshinka), heute ungebrauchliche deutsche Ortsnamen (Grünfeld in Kirgistan, Alexanderhöh in Russland) sowie Verwaltungsgliederungszusätze (Pobotschnoje Geb. Saratow/Kasachstan) zu beachten. Ohne hier im Detail die auftretenden technischen Probleme abhandeln zu können, muss festgestellt werden, dass ein gewisser Programmieraufwand entsteht. Der Aufwand für die Zusammenstellung von Ortsnamenlisten ist so lange überschaubar, wie keine vollständige Ausschöpfung erzielt werden muss, weil bereits wenige Geburtsorte eine große Personenzahl identifizieren (s. u.). Mit entsprechendem Rechercheinsatz kann aber auch ein solches Maximalziel weitgehend realisiert werden. Dabei ist es nicht notwendig, Synonyme und Schreibvarianten zusammenzufassen, jedoch muss eine möglichst eindeutige Zuordnung eines Orts zu einem Gebiet sichergestellt werden. Dazu gehört der Zugriff auf internationale Ortsdatenbanken.¹⁰

2.2 Ausschöpfung der Grundgesamtheit

Identifikation der Geburtsorte

Anhand eines Melderegisterauszugs wird im Folgenden abgeschätzt, zu welchem Grad das Verfahren in der Lage ist, Aussiedler zu identifizieren, ohne dabei fälschlicherweise Nicht-Aussiedler zu entdecken. Eine *exakte* Aussage zur Ausschöpfung wird kaum jemals möglich sein, weil der Umfang der Einstaater innerhalb der lokalen Aussiedler-Grundgesamtheit unbekannt ist, es sein denn, man könnte die Geburtsorte aller Deutschen einschließlich der aus der ganzen Welt stammenden Eingebürgerten lückenlos bestimmen. Die Größenordnung der erreichbare Ausschöpfung lässt sich aber umreißen. Hier geschieht dies in einer Pilotanwendung der Toponomastik auf Mehrstaater in Verbindung mit der publizierten Migrationsstatistik. Zunächst liegt uns ein Bestand 90.000 bundesdeutscher, 733 russischer, 697 kasachischer, 126 kirgisischer, 1136 rumänischer und 2.435 polnischer Ortsnamen (einschließlich der Schreibvarianten und älterer Bezeichnungen) vor. Untersucht wurden damit die anonymen Daten der über 20.000 Bielefelder Einwohner, die die erste oder zweite Staatsangehörigkeit Russlands, Kasachstans, Kirgisistan, Polens oder Rumäniens,

10 Z. B. der GEOnet Names Server der National Geospatial-Intelligence Agency (NGA) der USA, nationale Statistikbehörden und Postverwaltungen.

also der Hauptherkunftsländer der Aussiedler¹¹, besitzen, mit dem Ziel, nur durch den Geburtsort die Herkunftsländer zu identifizieren. Die Geburtsorte gehören ausnahmslos zu Ausländern (15,8%) oder Mehrstaatern (84,2%¹²). Die folgende Tabelle 1 weist die durch Datenbankanalyse maschinell ermittelten Geburtsländer nach aktueller Nationalität aus. Die Nationalitäten fassen die Erststaatsangehörigkeit der Ausländer und die Zweitstaatsangehörigkeit der deutschen Doppelstaater zusammen. (Letztere besitzen als erste Staatsangehörigkeit immer die deutsche.) An dieser Stelle ist es nicht möglich, auch deutsche Einstaater mit Geburtsort in einem Herkunftsländ wie Kasachstan – also mutmaßliche Aussiedler – zu ermitteln. Dies folgt in Abschnitt 2.3 auf Grundlage eines zweiten Datensatzes.

Tabelle 1 Staatsangehörigkeiten und identifizierte Geburtsorte

	Ausländische Staatsangehörigkeit										Gesamt N
	Polen		Rumänien		Russland		Kasachstan		Kirgisistan		
	N	%	N	%	N	%	N	%	N	%	
Geburtsort											
BRD	1808	20,5	71	13,2	836	13,7	536	13,4	246	19,5	3497
Polen	5092	57,7	0	0,0	3	0,0	2	0,0	1	0,1	5098
Rumänien	1	0,0	410	76,4	0	0,0	0	0,0	0	0,0	411
Russland	2	0,0	0	0,0	2568	42,1	264	6,6	131	10,4	2965
Kasachstan	0	0,0	0	0,0	515	8,4	1985	49,5	75	5,9	2575
Kirgisistan	3	0,0	0	0,0	130	2,1	41	1,0	516	40,8	690
Sonstige	40	0,5	1	0,2	154	2,5	79	2,0	37	2,9	311
uneindeutig	1125	12,7	53	9,9	397	6,5	212	5,3	81	6,4	1868
unbekannt	759	8,6	2	0,4	1501	24,6	894	22,3	177	14,0	3333
Gesamt	8830	100	537	100	6104	100	4013	100	1264	100	20748

Quelle: Melderegister der Stadt Bielefeld, Stand 13.12.2005; eigene Analysen.

Es zeigen sich Unterschiede, aber auch Gemeinsamkeiten zwischen den Herkunftsländern. Für 57,7% der polnischen Staatsangehörigen (einschließlich der Deutsch-Polen) ließ sich ein Geburtsort in Polen ermitteln. 20,5% sind in der Bundesrepublik geboren (darauf wird gleich zurückzukommen sein). 12,7% der Orte lassen sich keinem Land zuordnen, weil ihr Name in mehreren Ländern auftritt. Ein Beispiel für diesen Fall ist „Eichenau“, das es sowohl im Landkreis Fürstfeldbruck in Bayern als auch als alte Bezeichnung für Katowice-Mala Dabrowka (Woiwodschaft Slaskie/Oberschlesien) in Polen gibt. 8,6% der Geburtsorte sind bisher nicht ländercodiert. Die Geburtsorte der anderen Staatsangehörigkeiten sind seltener uneindeutig, weil in den betreffenden Ländern deut-

11 Aus den UdSSR-Nachfolgestaaten, Rumänien und Polen stammen 94% aller Aussiedler (Quelle: s. Fußnote 1). Die Ukraine als wichtiges aus der Sowjetunion hervorgegangenes Herkunftsländ bleibt hier unberücksichtigt.

12 Anteil der Mehrstaater: Polen 80,8%; Rumänien 67,8%; Russland 85,1%; Kasachstan 91,0%; Kirgisistan 88,9%.

sche Bezeichnungen die Ausnahme bilden. Ortsnamen treten aber oft mehrfach in Russland, der Ukraine, Kasachstan und anderen Staaten auf. Fasst man die Nachfolgestaaten der Sowjetunion zusammen, sinkt der Anteil uneindeutiger Zuordnungen unter 1%. Gleichzeitig sind die Geburtsorte der russischen, kasachischen und kirgisischen Staatsangehörigen häufiger unidentifiziert, was zum einen an den vielfältigen Varianten der ursprünglich kyrillisch geschriebenen Namen in lateinischer Schrift und zum anderen an der im Vergleich mit den ehemaligen deutschen Siedlungsgebieten in Polen, besonders in Schlesien, wesentlich schlechteren Dokumentation der Migrationsgeschichte liegt.¹³ Auffällig zeigt sich die Verschlingung der (bekanntlich nicht immer freiwilligen) Wanderungsverläufe der Deutschstämmigen innerhalb der Sowjetunion: 10,4% der kirgisischen Staatsangehörigen sind in Russland geboren, während von den 2.575 in Kasachstan Geborenen 264 Personen (6,6%) aktuell einen Pass der russischen Föderation besitzen.

Die Nachkommen der Aussiedler der ersten Generation sind selbstverständlich nicht an ihrem Geburtsort zu erkennen, der ja definitionsgemäß im Inland liegt. Dagegen haben offenbar auch viele Kinder der Aussiedler mehrfache Staatsangehörigkeiten inne, wie die Analyse der Geburtsorte ergibt. Von den in Bielefeld gemeldeten Staatsangehörigen der Russischen Föderation (Ein- und Doppelstaater) sind 14% in der Bundesrepublik geboren. Dieser Effekt kommt nicht durch fehlerhafte Ortsnamenerkennung zustande, denn alleine 12% geben Bielefeld als Geburtsort an. Die Anteile sind für die anderen Staatsangehörigkeiten teilweise sogar noch deutlich größer. Es ist damit sichergestellt, dass zumindest auf absehbare Zeit Aussiedler-Stichprobenziehungen nach Staatsangehörigkeit auch für die Folgegeneration möglich sein werden, wengleich keine Aufschlüsse darüber vorliegen, welche Ausschöpfung der zweiten Generation auf diese Weise erreicht wird und welche strukturellen Unterschiede zwischen Ein- und Mehrstaaten bestehen.

Die theoretische Obergrenze der erzielbaren Erkennungsleistung ist erreicht, wenn ein vollständiges Ortsverzeichnis zugrunde gelegt werden kann. Realistisch ist es dagegen anzunehmen, dass die Lage der Hälfte der zur Zeit unbekanntenen Orte mit vertretbarem Aufwand noch recherchiert werden kann.¹⁴ Damit lässt sich aus Tabelle 1 errechnen: 95,7% der Geburtsorte in Polen, 87,7% der Orte in Russland etc. und über die Staaten gewichtet gemittelt 92,0% aller Aussiedler-Geburtsorte lassen sich lokalisieren. Wenn nun etwa die Hälfte der Aussiedler bereits als Doppelstaater zu erkennen ist (dazu Abschnitt 2.3) und von der anderen Hälfte, den deutschen Einstaatern, 92% mit dem Geburtsortsverfahren ermittelt werden können, lässt sich durch Kombination von Staatsangehörigkeit und Geburtsort eine Erkennungsrate von gut 95% aller Aussiedler realisieren.

13 Zum früheren deutschen Verwaltungswesen in Schlesien und persönlichen Biographien finden sich zahlreiche Druckschriften und Internetseiten.

14 Hilfreich wäre die Einbeziehung der 101.418 Orte umfassenden Liste typischer Aussiedler-Geburtsorte des Bayerischen Landeskriminalamtes, die das Bundesamt für Migration und Flüchtlinge nutzte (Haug/Sauer 2007: 8).

Falsch positive Identifikation (Beifang)

Welche Personen erkennt das Verfahren fälschlicherweise? Diese Frage interessiert vornehmlich aus ökonomischen Gründen, nämlich im Hinblick auf überflüssige Kontakte, denn letztlich lässt sich ja eine Beeinträchtigung der realisierten Stichprobe durch Zielgruppenfremde mit Screenerfragen leicht vermeiden. Doppelstaater, die gleichzeitig Deutsche und Angehörige eines Herkunftslandes von Aussiedlern sind, ohne selbst deutscher Abstammung zu sein, dürften aufgrund der restriktiven deutschen Einbürgerungsrechts quantitativ unbedeutend bleiben. Eher stellen sonstige Personen ein Problem dar, die in der UdSSR oder ihren Nachfolgestaaten, Polen und Rumänien geboren sind. Hier ist zwischen Ausländern und deutschen Einwohnern zu unterscheiden.

Ausländer stellen nur insofern ein Problem dar, als der vollständige Verzicht auf sie den Ausschluß der oben erwähnten Aussiedler-Haushaltsangehörigen (frisch verheiratete Ehegatten, sonstige Mitreisende) nach sich zieht. Am Beispiel der russischen Nationalität kann man ablesen, dass in diesem – unter inhaltlichen Gesichtspunkten jedoch nicht sonderlich relevanten – Fall erheblicher Beifang zu erwarten ist. Am 31.12.2004 lebten knapp 179.000 Personen russischer Nationalität in der Bundesrepublik.¹⁵ Darunter sind neben ehemaligen Angehörigen der Roten Armee etc. hauptsächlich jüdische Kontingentflüchtlinge aus den GUS-Staaten. Die Zahl letzterer liegt zwischen den 219.604 für die Gruppe registrierten Einreisen und Aufnahmezusagen (bis Ende 2004) und den 111.811 am 18.4.2005 im Ausländerzentralregister (AZR) als Kontingentflüchtlinge verzeichneten GUS-Staatsangehörigen (Haug/Schimany 2005: 8). Die genaue Zahl ist unbekannt, weil viele Juden in Drittstaaten weiterreisen; viele inzwischen Eingebürgerte bleiben dem AZR verborgen. Offensichtlich ist von deutlich mehr Personen ohne als mit Beziehung zu einem Aussiedler-Haushalt auszugehen.

Etwas schwieriger ist der Beifang unter Eingebürgerten zu quantifizieren. Die Eingebürgertenstatistik enthält seit 2000 keine Aussiedler mehr. Für frühere Zeiträume ist nach Auskunft des Bundesministeriums des Inneren¹⁶ zwar nicht sichergestellt, dass alle Aussiedler enthalten sind, aber zumindest ein großer Teil, was die Bestimmung des Umfangs nichtdeutscher Einbürgerungen erschwert. Eine Schätzung ist dennoch möglich. Bis 1988 wurden den Statistischen Jahrbüchern zufolge jährlich maximal 1.000 Personen aus der UdSSR eingebürgert. Seit 2000 sind es im Mittel 4.100 jährlich. Dagegen waren es von 1989 bis 1999 im Durchschnitt 118.800 pro Jahr. Offensichtlich erklärt sich die große Zahl in dieser Periode nur durch die enthaltenen Aussiedler. Nehmen wir nun an, dass seit 1985 im Mittel nur ca. 4.000 Personen nichtdeutscher Abstammung jährlich die deutsche Staatsbürgerschaft verliehen wurde – davor fanden kaum Einbürgerungen statt – und alle diese Personen noch im Bundesgebiet leben. (Der Versuch einer weitergehenden Quantifizierung der Abwanderung für diesen Personenkreis wie auch für Aussiedler führt zu weit, siehe aber die Überlegungen bei Mamme/Schiener (1996) und Engels (2000: 2).) Dann stünden 2,3 Mio. Aussiedlern aus der ehemaligen UdSSR ca. 75.000 sonstige Eingebürgerte gegenüber, und eine Beifangquote von 3% wäre folglich zu erwarten.

15 Statistisches Bundesamt, <http://www.destatis.de/basis/d/bevoe/bevoetab10.php> (3.7.2007).

16 Telefonische Auskunft der Abteilung M 2.5.

Ein ähnlicher Überschlag ergibt für polnischstämmige Aussiedler bei ca. 80.000 Eingebürgerten einen Beifang von 5,2%. Wegen geschichtlicher Besonderheiten stellt Polen allerdings in zweifacher Hinsicht einen Sonderfall dar, weshalb diese Quote stark irreführend ist. Eine erste, nicht quantifizierbare, aber vermutlich kleine Störgröße bilden in diesem Fall Personen, die in den ehemaligen deutschen Ostgebieten im heutigen Polen geboren wurden, ihren Wohnsitz aber in der Vorkriegszeit als Binnenmigranten im Deutschen Reich spontan gen Westen verlassen und nicht erst, wie es §1 (1) des Bundesvertriebenengesetzes zur Erfüllung der Vertriebeneneigenschaft fordert, „im Zusammenhang mit den Ereignissen des zweiten Weltkrieges infolge Vertreibung, insbesondere durch Ausweisung oder Flucht, verloren“ haben. Zweitens sind in der unmittelbaren Nachkriegszeit mehrere Millionen Vertriebene ohne Einbürgerung in die Bundesrepublik und in die Deutsche Demokratische Republik gelangt (Faulenbach 2002: 44). Dort fanden sie u. a. wegen der vergleichbaren Vorkriegssituation und des Lastenausgleichsgesetzes Startvoraussetzungen vor, die denen der alteingesessenen Bevölkerung nicht unähnlich waren. Es wäre unangemessen, sie zugleich mit den Aussiedlern ab den achtziger Jahren in eine Stichprobe zu ziehen, und ein aufwendiges Screening wird daher unvermeidlich. Wegen der eher geringen Zahl der zwischen 1960 und 1986 eingereisten Personen¹⁷ nehmen sich in den Geburtskohorten ab 1960 allerdings die Effizienzprobleme weniger dramatisch aus als in früheren. Nichtsdestoweniger legt diese Überlegung es nahe, dem uneingeschränkt funktionierenden Doppelstaaterprinzip für Aussiedler aus Polen der Vorzug gegenüber dem Geburtsortsprinzip zu verleihen.

2.3 Ist die Toponomastik verzichtbar?

Das führt zu der Frage, ob das Kriterium der Doppelstaatigkeit für eine Stichprobenziehung aus dem Melderegister ausreicht bzw. inwiefern auf die ja viel aufwendigere Auswertung des Geburtsorts durch das toponomastische Erkennungsverfahren verzichtet werden kann. Dies wäre allenfalls bei einem kleinen Anteil deutscher Einstaater der Fall. Um diese Frage zu beantworten, wurde ein zweiter Datensatz herangezogen. Er enthält die Geburtsorte und Staatsangehörigkeiten aller nicht in Bielefeld selbst geborenen Personen im Bielefelder Melderegister, also auch die Geburtsorte aller Deutschen, die in Mannheim, New York, Timbuktu oder Omsk geboren sind.¹⁸ Diese Datenbasis schließt alle Aussiedler im juristischen Sinn ein, auch aktuell deutsche Einstaater, da diese im Ausland geboren sein müssen. Die eindeutig in den wichtigsten Herkunftstaaten der Aussiedler geborenen Personen wurden identifiziert. Uneindeutige Gebietszuordnungen und gänzlich unbekannte Geburtsorte bleiben nun unberücksichtigt. Für die Proportionen der Staatsangehörigkeiten dürfte dies aber nicht von erheblicher Bedeutung sein.

Die Zahl der möglichen Kombinationen von Geburtsländern und Staatsangehörigkeiten ist unübersichtlich. So ist es möglich, dass in einem Ort im heutigen Kasachstan geborene Deutsch-

17 Vgl. Fußnote 2.

18 187.840 von 331.151 Bielefeldern sind außerhalb der heutigen Stadtgrenzen geboren (56,7%).

stämmige einen russischen Pass besitzen, z. B. weil sie noch zu Sowjetzeiten den Wohnsitz gewechselt haben. Zur Vereinfachung wurde eine Zusammenfassung zu der technischen Variablen *Staatigkeit* vorgenommen (siehe Tabelle 2). Staatigkeit bedeutet die Kombination von Staatsangehörigkeiten und hat hier drei Ausprägungen. Unter *Einstaatern* sind Personen mit nur einer Staatsangehörigkeit, hier der deutschen, zu verstehen. *Mehrstaater* besitzen mindestens, aber fast ausnahmslos genau, zwei Staatsangehörigkeiten, von denen hier immer eine die deutsche ist. Die zweite ist mehrheitlich, aber nicht ausschließlich, die des Geburtslandes. Alle Kombinationen wie die obige werden damit erfasst. Die Kategorie *Ausländer* definiert sich durch das Fehlen der deutschen Staatsangehörigkeit und umfasst sowohl ausländische Ein- als auch Mehrstaater (letztere etwa mit polnischem und gleichzeitig französischem Pass).

Tabelle 2 Staatigkeit nach Geburtsland

(1) Geburtsland	(2) dt. Ein- staater*	(3) % der Stichprobe	(4) % der Deutschen	(5) dt. Mehr- staater*	(6) % der Stichprobe	(7) % der Deutschen	(8) Ausländer*	(9) % der Stichprobe	(10) Summe
Russland	2.278	41,1	45,1	2.770	50,0	54,9	490	8,8	5.538
Polen	11.603	69,4	74,4	3.983	23,8	25,6	1.137	6,8	16.723
Kasachstan	2.785	49,6	52,5	2.516	44,8	47,5	310	5,5	5.611
Kirgisistan	1.054	57,1	61,0	674	36,5	39,0	119	6,4	1.847
Rumänien	274	39,9	51,8	255	37,2	48,2	157	22,9	686
Summe	17.994	59,2	63,8	10.198	33,5	36,2	2.213	7,3	30.405

Quelle: Melderegister der Stadt Bielefeld, Stand 17.3.2006; eigene Auswertungen.

* siehe Anmerkungen im Text.

Die Fallzahlen in Tabelle 2 sind zum einen auf die gesamte Stichprobe prozentuiert. Zunächst ist zu erkennen, dass der Anteil der Ausländer, abgesehen von Rumänien, nur im einstelligen Prozentbereich liegt (Spalte 9). Zum anderen ist für deutsche Ein- und Mehrstaater der Prozentanteil bezogen auf die Summe dieser beiden Ausprägungen angegeben (Spalten 4 u. 7). Sie entspricht abgesehen von sehr wenigen nicht eingebürgerten Aussiedlern sowie eingebürgerten Nicht-Aussiedlern der Zahl der Aussiedler (Ausnahme: Polen, s.o.). Russlanddeutsche Aussiedler sind demnach mehrheitlich Doppelstaater (54,9) während gut die Hälfte der kasachischen, kirgisischen und rumänischen Aussiedler nicht mehr die Staatsangehörigkeit des Herkunftslandes innehat (51,8% bis 61,0%, Spalte 4). Ob der Verzicht auf deutsche Einstaater zu qualitativen Verzerrungen führt, weil sie systematisch andere sozialstrukturelle Merkmale aufweisen als Doppelstaater, kann sicher letztlich nur substantielle Forschung zeigen. Zumindest bis zum Beweis des Gegenteils und auf jeden Fall unter dem Gesichtspunkt der quantitativen Ausschöpfung muss das Fazit aber lauten, dass das toponomastische Verfahren unerlässlich ist.

2.4 Administrative und juristische Implementationsvoraussetzungen, Kosten

Obwohl das toponomastische (wie auch das onomastische) Verfahren nicht an einen speziellen Auswahlrahmen gebunden ist, wird doch die Auswahl aus dem kommunalen Melderegister der häufigste Anwendungsfall sein. Die Mitwirkung der Meldeämter ist daher eine zentrale Implementationsvoraussetzung. Die juristischen Grundlagen sind insofern gegeben, als 1. zumindest Universitäten und vergleichbare Forschungseinrichtungen zu den „sonstigen öffentlichen Stellen“ zu zählen sind, die das Melderecht als Empfänger von Gruppenauskünften vorsieht, 2. die Daten zur Erfüllung der Forschungsaufgabe erforderlich sind, weil Aussiedler vorrangig am Geburtsort zu erkennen sind, und 3. die benötigten Merkmale in den Meldegesetzen auch explizit als mitteilbar genannt werden. Dass hier Listen von Geburtsorten als Input verwendet werden, dürfte, juristisch betrachtet, nicht von Bedeutung sein, da ja auch konventionelle Melderegisterstichproben auf – allerdings wesentlich kürzere – Listen von Staatsangehörigkeiten zurückgreifen; es ist zu bedenken, dass ja sogar Auskünfte über namentlich genannte Einzelpersonen melderechtlich unter bestimmten Voraussetzungen erlaubt sind. Allerdings wird man dies im Einzelfall Amtsleitern und örtlichen Datenschutzbeauftragten erläutern müssen. Ablehnungen des Verfahrens wurden auf sechs bisher gestartete Anfragen an Städte hin nicht erteilt. Dass zumindest simple melderegisterbasierte Stichproben bundesweit möglich sind, zeigt der ALLBUS (Haarmann et al. 2006).

Die technische Realisierung dürfte indes oft ein Problem darstellen, weil die Programmierung des Selektionsalgorithmus an die Grenzen der Möglichkeiten der datenführenden Dienststellen stößt. Auch die vereinfachte Version, der starre 1:1-Abgleich eines Bestands von Ortsnamen mit den lokal gespeicherten Geburtsorten, verursacht in Ämtern und beauftragten Rechenzentren nicht unerheblichen Programmieraufwand, da es sich um ein zwar keineswegs generell in der Datenbankprogrammierung, jedenfalls aber im Verwaltungsvollzug ungewöhnliches Verfahren handelt. Die Kosten für den Zusatzaufwand wird in der Regel der Forscher tragen müssen, weshalb das Verfahren in jedem Fall merklich teurer als herkömmliche Melderegisterstichproben sein muss. Leider ist an dieser Stelle keine allgemein gültige Kostenschätzung möglich, weil die Kalkulationsmodelle der Länder und Kommunen zu stark variieren. So sieht Hamburg einen Korridor von 100,- bis 12.500,- Euro je Gruppenauskunft vor, während die Gebührenordnung Niedersachsens eine Pauschale von 14,80 bis 35,70 Euro zuzüglich 3 bis 15 Cent je ausgewählten Einwohner und 0,1 bis 3 Cent je registrierten Einwohner festlegt.¹⁹ Für eine Stichprobe von N=2.000 bei einem Einwohnerbestand von 200.000 Personen bleiben die Kosten ohne Berücksichtigung des Programmieraufwands im günstigsten Fall unter 300, im ungünstigsten steigen sie über 6.300 Euro. Die Rechnungsmodelle der anderen Bundesländer sind ähnlich flexibel, wobei z. B. in Baden-Württemberg zumindest die Komponente für die gemeldeten, aber nicht ausgewählten Einwohner ent-

19 http://hh.juris.de/hh/MeldeGebO_HA_1993_P1.htm, http://cdl.niedersachsen.de/blob/images/C1736588_L20.pdf (18.4.2006). Die Gebührenordnungen zahlreicher Gebietskörperschaften sind im Internet verfügbar. Weitere Details würden den verfügbaren Rahmen sprengen.

fällt.²⁰ Die Kosten schwanken auch innerhalb eines Bundeslandes. Verwaltungen mancher Städte erlassen Universitäten die Gebühren oder rabattieren sie erheblich (wie Stuttgart). Realistisch erscheint dem Autor der Durchschnittsbetrag von 200 bis 1.000 Euro je Kommune, den Orte wie Dortmund, Stuttgart und Hamburg auf Anfrage nannten; leider sind auch Ausreißer von über 10.000 Euro bekannt.

Dies ist, neben der bisher herrschenden Unklarheit über Referenzparameter, sicher einer der Gründe, warum ein naheliegendes - wenn man will, eigentlich triviales - Verfahren in der Vergangenheit kaum eingesetzt worden ist. Der Ressourcenaufwand für Flächenstichproben ist hoch, allenfalls sind zweistufige Verfahren mit wenigen Kommunen auf einer ersten Auswahlstufe vorstellbar, wie sie das Bundesamt für Migration und Flüchtlinge (Haug/Sauer 2007: 5) implementierte. Für bundesweit repräsentative Erhebungen werden die Kosten in der jetzigen kommunalen Struktur des Meldewesens wohl immer prohibitiv bleiben.

3 Onomastische Stichprobenziehung

3.1 Numerische Beschreibung

Die Stichprobenziehung auf der Grundlage spezifischer Namen verdient eine genaue Prüfung, weil sie in Verbindung etwa mit dem Telefonbuch als Auswahlgrundlage eindeutige Vorzüge aufweist: Sie ist kostengünstig, leicht zu implementieren und problemlos auf beinahe beliebige Flächeneinheiten ausdehnbar. Im folgenden wird geprüft, inwiefern die in der Einleitung angedeuteten möglichen Nachteile diese Vorteile zunichte machen.

Als Grundlage dient eine im Juni 2004 für eine Studie zu Diskriminierungserfahrungen aus dem Einwohnermelderegister der Stadt Bielefeld gezogene Zufallsstichprobe von 1732 Aussiedlern (neben anderen wichtigen Herkunftsgruppen) im Alter von 18 bis 25 Jahren. Ziehungskriterium war für sie ein Geburtsort in einem Herkunftsland von Aussiedlern. Die Verteilung nach Geburtsland (Russland 47,9%, Kasachstan 33,8%, Kirgisistan 11,3%, Polen 6,8%, sonstige Länder 0,3%) weist ein leichtes Übergewicht aus Russland stammender Personen auf, weil bei der Stichprobenziehung ein sehr frühes Ortsregister mit niedriger Ausschöpfung der anderen Herkunftsländer zum Einsatz kam. Es ist auch nicht auszuschließen, dass in der Bruttostichprobe, für die nur teilweise Befragungsergebnisse zur Validierung vorliegen, auch eine kleine Zahl von Nicht-Aussiedlern, v. a. aus Polen, enthalten ist.

Die Vor- und Nachnamen in der Stichprobe wurden nach ihrer Typik in die Kategorien 1. *deutsch*, 2. *für das Geburtsland typisch* sowie 3. *unbekannt oder für ein Drittland typisch* eingeordnet. Dabei konnte nur auf Allgemeinwissen zurückgegriffen werden, weil keine elabo-

20 Der Einfluß der Einwohnerzahl auf die Kosten einer Meldeamtsstichprobe hat im ALLBUS zu einer Anpassung des Stichprobendesigns geführt, vgl. Haarmann et al. 2006: 55, FN. 21.

rierte Quelle wie die von Huschka/Gerhards/Wagner (2005) genutzte Personennamen-Datei der Personennamen-Beratungsstelle der Universität Leipzig zur Verfügung stand. Da die Daten aus dem Melderegister stammen, soll im Interesse der Vertraulichkeit die Zuordnung nicht im Detail aufgeführt werden. Oft enthalten herkunftslandtypische fremde Varianten der Namen die Zeichenfolgen *ks* anstelle von *x*, *ei* oder *ej* anstelle von langem *e* oder *é*, *dr* statt *der*, *z* oder *sz* statt *s*, *ija* oder *ja* anstelle von *ia* etc. Vor allem wurden in der Bundesrepublik nur selten vorkommende Namen russischen oder polnischen Ursprungs als typisch für Russland, Polen etc. klassifiziert. Zwischen den Ländern unterscheiden sich die Verteilungen der Namen-Typik nicht sehr stark. Lediglich Personen aus Polen tragen häufiger Nachnamen des Herkunftslandes sowie andererseits deutsche Vornamen als der Rest der Stichprobe. Eine Kreuztabelle von Vor- und Nachnamen der Gesamtstichprobe (Tabelle 3) zeigt, dass deutsche Namen eindeutig überwiegen (63,7% der Vor-, 77,8% der Nachnamen). Mehr als die Hälfte (52,4% der Personen) trägt sowohl einen deutschen Vor- als auch Nachnamen.

Tabelle 3 Namenstypik der Aussiedler

Vorname	Nachname						Summe	
	Deutsch		Herkunftsland		Unklar		N	%
	N	%	N	%	N	%		
Deutsch	907	52,4	161	9,3	35	2,0	1103	63,7
Herkunftsland	419	24,2	151	8,7	22	1,3	592	34,2
Unklar	22	1,3	12	0,7	3	0,2	37	2,1
Summe	1348	77,8	324	18,7	60	3,5	1732	100,0

Quelle: Melderegister der Stadt Bielefeld, Juni 2004; eigene Analysen.

Die Verteilung der Nachnamen ist deutlich flacher als die der Vornamen. Es wurden einschließlich der Schreibvarianten 1112 verschiedene Nach-, aber nur 382 Vornamen ermittelt. Deutsche Vornamen kommen durchschnittlich 5,0 mal, fremde 4,4 mal vor, deutsche Nachnamen 1,8 mal, fremde 1,1 mal. (Wegen der unterschiedlichen Besetzung der Herkunfts-Kategorien sind diese Werte nicht direkt vergleichbar.) Die Hitliste der häufigsten Vornamen wird von Al. (96 Vorkommen), Vi. (64), An. (57) bzw. den weiblichen Namen Ol. (48), Ta. (39) und An. (37) angeführt. In der Menge der 38 in der Stichprobe mehr als zehnmals vorkommenden Vornamen befinden sich 26 für den deutschen Sprachraum allgemein typische. Somit ist abzusehen, dass eine Stichprobe auf Vornamenbasis nicht unerheblichen Beifang erzielen würde. Da sich lokale Clusterungen, etwa durch Kettenmigration, kaum auf die Vor-, aber nachhaltig auf die Nachnamenverteilung auswirken, soll letztere hier nicht aufgefächert werden. Es muss genügen zu sagen, dass sich unter den 16 häufigsten die im deutschen Sprachraum allgemein sehr verbreiteten Familiennamen Schm. (231.000 mal im Telefonbuch), Br. (52.000 mal), Pe. (38.000 mal), Schr. (63.000 mal), Wo. (62.000 mal) und Be. (91.000 mal) befinden.

3.2 Abschätzung der Trennschärfe

Für eine grobe Abschätzung der onomastisch erzielbaren Trennschärfe wurden die Namen der Stichprobe mit dem Bestand der privaten Anschlüsse in einem maschinenlesbaren Telefonbuch verglichen. Idealerweise müssten die Aussiedlernamen einem bundesweiten Bestand (nicht mit einer Bielefelder Stichprobe, weil dabei örtliche Besonderheiten zusätzliche Probleme bereiten würden) aus der Zeit vor Beginn der Einreisewelle aus den deutsch besiedelten Gebieten in Mitteleuropa und der ehemaligen Sowjetunion gegenübergestellt werden, was aber technisch nicht zu realisieren ist. Daher wurde eine frühe Telefon-CD aus dem Jahr 1997 benutzt. Sie enthält 33 Mio. private Anschlüsse. 3 Mio. Vornamen sind nicht vorhanden oder abgekürzt. Im Jahr 1997 war allerdings ein erheblicher Teil der Aussiedler schon eingereist, wobei hingegen ihre Neigung, sich auch ins Telefonbuch eintragen zu lassen, unbekannt ist. Man kann annehmen, dass durch die zeitweilige Heimunterbringung viele Aussiedler erst mit Verzögerung in die Telefonbücher eingehen. Es kommt bei dieser Vorgehensweise durch einige Unbekannte sicher zu Schätzfehlern, die aber hingenommen werden können, da nur eine Größenordnung in Erfahrung gebracht werden soll. Ein einfaches Maß der Trennschärfe ist das Verhältnis der Personenzahl eines gegebenen Namens in der Stichprobe zur Zahl der auf einen solchen Namen gemeldeten Telefonanschlüsse. Die zentrale Strategie der Onomastik besteht darin, bevorzugt Namen mit einem hohen Quotienten zu verwenden wie Ir., der sich 36 mal in der Stichprobe und fast 5.500 mal im Telefonbuch befindet, und nicht etwa Pe., den es zwar auch 27 mal in der Aussiedlerstichprobe, aber vor allem 532.000 mal im Telefonbuch gibt. Daher wurden alle Namen nach dem Trennschärfequotienten sortiert. Dann wurde ermittelt, wie sich mit schlechter werdender Trennschärfe die Zahl gefundener Telefonanschlüsse verändert (dazu Abbildungen 1 und 2).

In Abbildung 1 ist gegen die Ausschöpfung der Bielefelder Stichprobe (x-Achse) abgetragen, wie viele Telefonanschlüsse mit den für die jeweilige Ausschöpfung notwendigen Vornamen gefunden werden (gestrichelte Linie, linke y-Achse). Wie man sieht, ist das Verfahren für eine Ausschöpfung bis ca. 60% relativ effizient. Die Vornamen von 60% der Stichprobe existieren 441.000 mal im Telefonbuch. Bei einer angenommenen Haushaltsgröße von durchschnittlich 2,5 Personen und einer Telefonbuch-Eintragsquote von 33% (beides grobe Schätzungen) würden somit 3,3 Mio. Personen erreicht. Dies entspricht der Größenordnung der Zahl aller zwischen 1950 und Ende 1996 eingereisten Aussiedler (3,2 Mio.), unter denen Personen aus der ehemaligen UdSSR und aus Polen ja den weitaus größten Anteil haben. Zweifelsohne ist es unrealistisch, nun anzunehmen, mit dem Verfahren würde eine Trennschärfe von annähernd 100% erzielt, weil die Zahlen fast gleichauf liegen. Doch andererseits kann man schließen, dass selbst bei Annahme ungünstigerer Parameter (Stichprobeneffekte, geringere Eintragsdichte der Aussiedler etc.) die Trennschärfe nicht maßlos sinkt, was den Screening-Aufwand prohibitiv in die Höhe schnellen ließe. Abbildung 1 zeigt allerdings auch, dass hauptsächlich fremde Namen durch hohe Trennschärfe hervorgehoben sind (obere Linie, rechte y-Achse). Einige seltene deutsche Doppelnamen erweisen sich als hochgradig trennscharf, weshalb der Anteil deutscher Namen anfangs noch eher hoch ist und dann fällt.

Abbildung 1 Trennschärfe der Vornamen

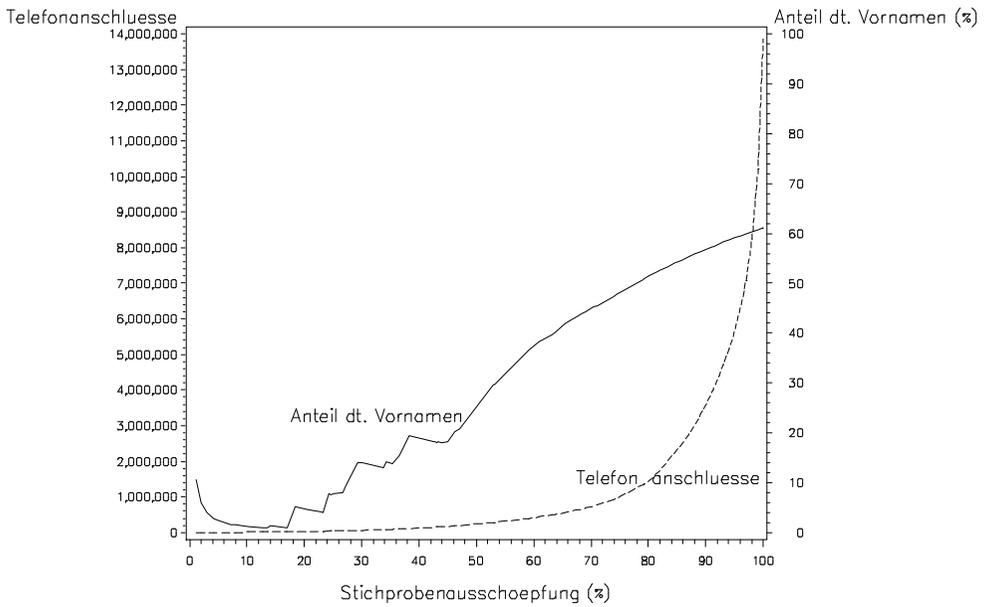
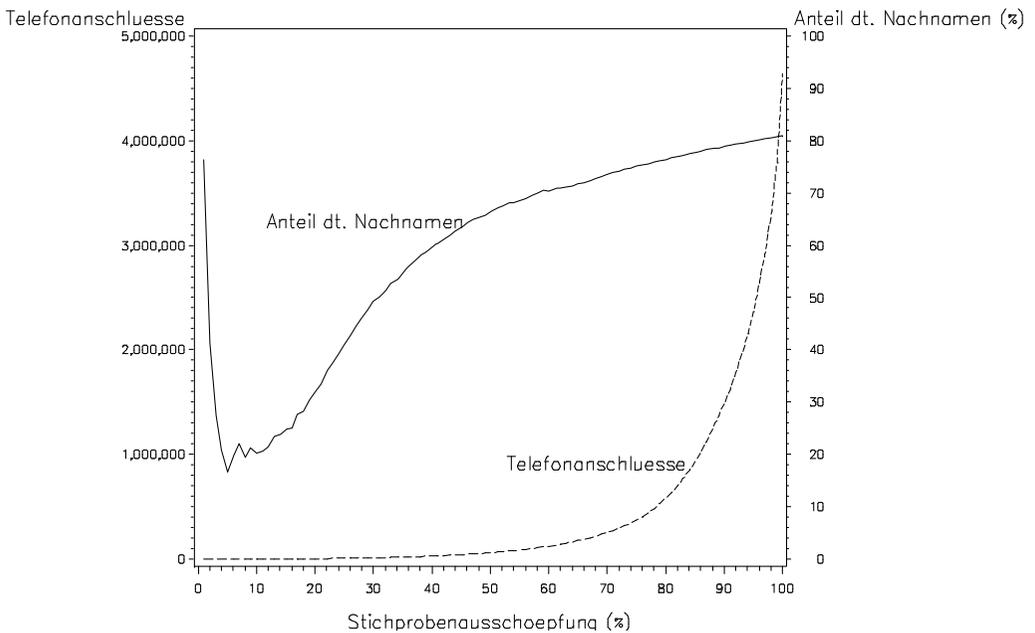


Abbildung 2 Trennschärfe der Nachnamen



Bei 60%iger Ausschöpfung liegt der Anteil deutscher Namen bei nur 38,2%. Um keine Disparität entstehen zu lassen, müsste also eigentlich die Ausschöpfung erhöht werden, was wiederum zu Lasten der Effizienz ginge.

Nachnamen erweisen sich in dieser Auswertung als trennschärfer (siehe Abbildung 2). Auch die Verzerrung zu Lasten deutscher Namen ist weniger ausgeprägt. Es besteht allerdings, wie oben erwähnt, die Gefahr, dass dies durch Stichprobenfehler zustande kommt. Örtliche Ballungen einzelner Clans führen zu einem hohen Trennschärfemaß bei geringer absoluter Ausbeute, weil die Namen nur in Bielefeld vorkommen. Für überregionale Stichproben sind diese Namen nutzlos. Die Bilanz verschlechtert sich, wie oben gezeigt, bei Verwendung häufiger Familiennamen. Daher soll auf den Einsatz der Nachnamen nicht weiter eingegangen werden.

3.3 Unterscheiden sich Aussiedler mit fremden Namen von denen mit deutschen Namen?

Solange nicht belegt ist, dass sich die Träger deutscher und fremder Namen unterscheiden, spricht nichts gegen die Nutzung überwiegend russischer und polnischer Namen. Man könnte in diesem Fall auf das aufwendige und voraussetzungsvolle Geburtsortsverfahren verzichten und mit einer namensorientierten Auswahl z. B. aus dem Telefonbuch auch zum Ziel kommen, was unter forschungsökonomischen Gesichtspunkten nicht ganz uninteressant wäre. Mit Ergebnissen der Befragung auf Grundlage obiger Stichprobe sollen dazu Anhaltspunkte gesammelt werden.

Betrachtet wird der Vorname. Neben Differenzen im Antwortverhalten kann ein Effekt durch selektives Unit-Non-Response zustande kommen. Für 62 Adressen liegen Gründe der Nicht-Teilnahme vor. Der Anteil der Verweigerer ist unter den Personen mit fremdem höher (62%) als mit deutschen Namen (46%), der Zusammenhang ist aber bei der gegebenen Fallzahl nicht signifikant. Die Betrachtung konzentriert sich daher auf einen Vergleich zwischen 47 Fällen mit deutschen und 35 Fällen mit fremden Namen. Der auffälligste Unterschied ist im durchschnittlichen Einreisezeitpunkt gegeben, der bei deutschem Namen im März 1991 und damit fast zwei Jahre früher als bei fremdem Namen (Februar 1993) liegt. (Das Lebensalter unterscheidet sich wegen der gesetzten Stichprobenkriterien nicht.) Dieser Effekt entspricht der bekannten Tatsache, dass in den späteren Einreisekohorten der Anteil russisch-deutscher Mischehen stark angestiegen ist, während in der frühen Zuwanderungsphase eher Menschen kamen, die ihre ethnische Spezifität als Deutsche bewahrt hatten. In den wichtigsten sozialstrukturellen Dimensionen (Bildung, Einkommen, Erwerbsstatus, Sprachfähigkeit) ließen sich keine signifikanten Unterschiede ermitteln, obwohl für Träger deutscher Namen erhöhte Anteile Vollzeitbeschäftigter und niedrigere Arbeitslosenanteile berechnet wurden. Signifikante Unterschiede (t-Test, $\alpha=5\%$) stellen sich hingegen in Einstellungen und Netzwerkaktivitäten heraus. Wer einen deutschen Vornamen trägt, hat eher das Gefühl, mit autochthonen Deutschen viele Gemeinsamkeiten zu besitzen. Er hat mehr deutsche Vereinskameraden und trifft Freunde aus dem Herkunftsland häufiger (übrigens ebenfalls autochthone deutsche Freunde, aber nicht statistisch signifikant). In der Diskriminierungserfahrung ist

kein signifikanter Unterschied zu entdecken. Resümierend läßt sich festhalten, dass abgesehen von der längeren Aufenthaltsdauer in dieser Auswertung nur punktuelle Differenzen in Einstellungsangelegenheiten sichtbar werden und insgesamt noch keine Strukturabweichungen und damit keine fundamentalen Gründe gefunden wurden, die eine exakte Parität der Namensverteilung zur Repräsentativität einer Stichprobe erforderlich machen. Dies gilt natürlich unter dem Vorbehalt der begrenzten Tragfähigkeit der hier benutzten Stichprobe und muss unter Umständen zukünftig revidiert werden.

4 Vorläufige Bewertung und Ausblick

Die beschriebenen Verfahren zur Identifikation von Aussiedlern aufgrund von Staatsangehörigkeiten und Geburtsort und ggf. Namen sind zwar aufwendig, weil sie historisch-geographische Recherchen voraussetzen und ein gutes Maß nicht-trivialer Datenbankprogrammierung erfordern und weil in den beteiligten Kommunen Datenschutzerfordernisse in höherem Maß als bei üblichen Stichprobenziehungen geprüft werden müssen, wobei zudem beträchtliche Kosten anfallen. Auf lokaler Ebene aber handelt es sich um eine vergleichsweise sehr präzise Methode, die eine Ausschöpfung der Zielgruppe von 95% bei sehr geringem Beifang erwarten lässt.

Entworfen wurde diese Technik zur Stichprobenziehung. Zwei weitere Anwendungsfelder sind aber denkbar:

1. Auch in anderen Datenbeständen, etwa im System der Sozialversicherungen, werden Geburtsorte registriert, wenn auch nicht immer in Verbindung mit mehreren Staatsangehörigkeiten. Die Selektion nach Geburtsort dürfte einen tieferen Einblick in die strukturellen Verteilungen und die Risiken, denen Aussiedler in der Bundesrepublik ausgesetzt sind, bieten, als mit der bisherigen Statistik möglich war.
2. Kommunale Statistiken, etwa zur residentuellen Segregation von Zuwanderern, leiden bekanntermaßen unter den Beschränkungen des Merkmals Staatsangehörigkeit. Das toponomastische Verfahren vermag hier für deutschstämmige Aussiedler Abhilfe zu schaffen.

Stellt die Toponomastik eine Methode *comme il faut* dar, so wird die Forschung oft größere Flächeneinheiten abdecken und deshalb alleine aus befragungsökonomischen Gründen auf die Onomastik zurückgreifen müssen. Dabei sind zwischen Trennschärfe, Ausschöpfung und Repräsentativität Kompromisse einzugehen, mit denen man aber nach derzeitigem Erkenntnisstand wohl leben kann. Es müsste möglich sein, auch mit Listen typischer Aussiedlernamen vor allem aus dem Telefonbuch passable überregionale Stichproben zu generieren.

Eine besondere Herausforderung stellt der Nachwuchs der Aussiedler im juristischen Sinn dar, der für die Sozialwissenschaften nicht weniger relevant sein darf als die im Ausland Geborenen, weil etwa Sprachdefizite im elterlichen Haushalt nicht von heute auf morgen abgelegt werden. Das Geburtsort-Prinzip ist definitionsgemäß für die Stichprobenziehung der zweiten Generation ungeeignet, während Doppelstaatigkeit offensichtlich in gewissem Umfang weiterhin vorliegt. Da sich vermutlich auch Präferenzen für Vornamen rasch ändern, ruhen die Hoffnungen vor allem

darauf, noch eine Weile typische Nachnamen identifizieren zu können. Auch diese werden durch Eheschließungen zunehmend an Trennschärfe verlieren. Vielleicht ist aber der *race-relations-cycle* irgendwann in der Zukunft, wenn auch dieses letzte Kriterium unbrauchbar geworden ist, soweit fortgeschritten, dass sich niemand mehr für Aussiedlernachkommen interessieren muss.

Literatur

- Babka von Gostomski, Ch., 2001: Adressenscreening über Schulen: Erfahrungen beim IKG-Jugendpanel. ZA-Information 49: 29-48.
- Beauftragte der Bundesregierung für Migration, Flüchtlinge und Integration, 2005: 6. Bericht über die Lage der Ausländerinnen und Ausländer in der Bundesrepublik Deutschland. http://www.bundesregierung.de/Content/DE/Publikation/IB/Anlagen/ausl_C3_A4nderbericht-6-teil-III,property=publicationFile.pdf (27.6.2007).
- Engels, D., 2000: Aussiedlerinnen und Aussiedler in der Sozialhilfe. Köln: Institut für Sozialforschung und Gesellschaftspolitik (ISG).
- Faulenbach, B., 2002: Die Vertreibung der Deutschen aus den Gebieten jenseits von Oder und Neiße: Zur wissenschaftlichen und öffentlichen Diskussion in Deutschland. Aus Politik und Zeitgeschichte B 51-52: 44-54.
- Gabler, S., 1992: Schneeballverfahren und verwandte Stichprobendesigns. ZUMA-Nachrichten 31: 47-69.
- Haarmann, A., E. Scholz, M. Wasmer, M. Blohm und J. Harkness, 2006: Konzeption und Durchführung der „Allgemeinen Bevölkerungsumfrage der Sozialwissenschaften“ (ALLBUS) 2004. Mannheim: ZUMA.
- Haisken-DeNew, J. und J. Frick, (Hg.), 2005: Desktop Companion to the German Socio-Economic Panel (SOEP): Version 8.0. Berlin: Deutsches Institut für Wirtschaftsforschung.
- Haug, S. (unter Mitarb. v. P. Schimany), 2005: Jüdische Zuwanderer in Deutschland. Nürnberg: Bundesamt für Migration und Flüchtlinge. Working Papers 3/2005.
- Haug, S. und L. Sauer, 2007: Zuwanderung und Arbeitsmarktintegration von Aussiedlern und Spätaussiedlern. Ergebnisse einer Befragung zur Evaluation des Wohnortzuweisungsgesetzes. Vortragsmanuskript für den Workshop „Integrationschancen von Spätaussiedlern“, Donnerstag, 29.3. – Freitag, 30.3.2007, in der Bundesagentur für Arbeit, Institut für Arbeitsmarkt- und Berufsforschung. Nürnberg: Bundesamt für Migration und Flüchtlinge.
- Humpert, A. und K. Schneiderheinze, 2000: Stichprobenziehung für telefonische Zuwandererumfragen – Einsatzmöglichkeiten der Namenforschung (Onomastik). ZUMA-Nachrichten 47: 36-63.
- Humpert, A. und K. Schneiderheinze, 2002: Stichprobenziehung für telefonische Zuwandererumfragen – Praktische Erfahrungen und Erweiterung der Auswahlgrundlage. S. 187-208 in: S. Gabler und S. Häder (Hg.): Telefonstichproben – Methodische Innovationen und Anwendungen in Deutschland. Münster: Waxmann.
- Humpert, A., 2004: Erfahrungen mit Personennamen zur Bildung von Stichproben für Betriebsbefragungen. ZUMA-Nachrichten 54: 141-153.
- Huschka, D., J. Gerhards und G. Wagner, 2005: Messung und Analyse des sozialen Wandel anhand der Vergabe von Vornamen: Aufbereitung und Auswertung des SOEP. Dokumentation der Datenbasis und der Vercodung. Berlin: Institut für Soziologie der Freien Universität Berlin. http://userpage.fu-berlin.de/~gerhards/dokumentefrei/projektdoku_vornamen.pdf (3.7.2007).
- Mammey, U. und R. Schiener, 1996: Das bib-Aussiedlerpanel – methodische Überlegungen zur Repräsentativität. Zeitschrift für Bevölkerungswissenschaft 21: 145-171.
- Mammey, U. und R. Schiener, 1998: Zur Eingliederung der Aussiedler in die Gesellschaft der Bundesrepublik Deutschland: Ergebnisse einer Panelstudie des Bundesinstituts für Bevölkerungsforschung. Opladen: Leske + Budrich (Schriftenreihe des Bundesinstituts für Bevölkerungsforschung 25).
- Salentin, K., 1999: Die Stichprobenziehung bei Zuwandererbefragungen. ZUMA-Nachrichten 45: 115-135.
- Seifert, W., 2001: Gutachten für die „Unabhängige Kommission Zuwanderung“ beim Bundesministerium des Inneren zum Thema: Berufliche Integration von Zuwanderern in Deutschland. Berlin: Bundesministerium des Inneren.

- Schupp, J. und G. Wagner, 1995: Die Zuwanderer-Stichprobe des Sozio-oekonomischen Panels (SOEP). Vierteljahrshefte zur Wirtschaftsforschung 64: 16-25.
- Strobl, R. und W. Kühnel, 2000: Dazugehörig und ausgegrenzt. Analysen zu Integrationschancen junger Aussiedler. Weinheim/München: Juventa.
- Swiaczny, F., 2000: Innerstädtische Migration von Aussiedlern: Räumliches Verhalten und Netzwerke als Ursache für Konfliktpotentiale am Beispiel der Stadt Mannheim. Sozialwissenschaften und Berufspraxis 23: 61–85.
- Wetzels, P., D. Enzmann, E. Mecklenburg und Ch. Pfeiffer, 2001: Jugend und Gewalt: Eine repräsentative Dunkelfeldanalyse in München und acht anderen deutschen Städten. Baden-Baden: Nomos.

Korrespondenzadresse: Kurt Salentin
Institut für interdisziplinäre Konflikt- und Gewaltforschung
Universität Bielefeld
33501 Bielefeld
kurt.salentin@uni-bielefeld.de

Attrition in the Swiss Household Panel

Attrition im Schweizer Haushalt Panel

Oliver Lipps

Abstract

In this paper, we analyze magnitude and possible selectivity of attrition in first wave respondents in the Swiss Household Panel (SHP), from wave two (2000) through wave seven (2005). After comparing attrition of first wave respondents with that of other panel surveys, we proceed to model selectivity of attrition in two steps: we first build separate wave-to-wave models, and second a longitudinal all-wave model. The latter model includes wave interaction effects. The first models allow for tracing of selectivity development, i.e. whether an initial selectivity might compensate or cumulates over time, the second to assessing the effects of the covariates in a specific wave, controlling for the base attrition effect. In particular it allows for the analysis of consequences due to discrete fieldwork events.

Our results support the findings in the literature: attritors are in general the younger people and the males, foreigners, the socially and politically "excluded", i.e. those who show little social and political interest and participation, those who are mostly dissatisfied with various aspects in their life, and those who live in households with high unit nonresponse, and who exhibit a worse reporting behavior. This pattern is rather cumulative than compensating over panel waves. Excessive attrition in two waves presumably caused by two discrete events in the panel is not particularly selective. Still existing variation in selective attrition is worth to be further explored.

Zusammenfassung

In diesem Artikel werden Ausmaß und Selektivität der Panelattrition im Schweizer Haushalt Panel zwischen den Wellen zwei (2000) und sieben (2005) untersucht. Die Stichprobenbasis bilden die Teilnehmer der ersten Welle (1999). Nach einem Vergleich der Höhe und der Selektivität der Attrition mit anderen Haushaltpanels wird die Selektivität in zwei Schritten modelliert: zunächst durch Übergangsmodelle, in denen jeweils zwei aufeinanderfolgende Wellen verwendet werden, und schließlich durch ein Längsschnittmodell unter Verwendung aller Wellen. Letzteres enthält Wellendummies. Die Übergangsmodelle dienen dazu, Selektivität zu identifizieren und weiterzuerfolgen, d.h. zu erkennen, ob sich eine initiale Selektivität im Laufe der Zeit verstärkt oder ausgleicht. Mit dem Längsschnittmodell können spezifische Effekte von Kovariaten in einzelnen Wellen untersucht werden, wobei die Basisattrition kontrolliert wird. Insbesondere können die Auswirkungen einzelner Panelereignisse analysiert werden.

Die Ergebnisse bestätigen bisherige Erkenntnisse aus der Literatur: Personen, die aus dem Panel aussteigen sind tendenziell jünger und männlich, Ausländer, mit geringem sozialen und politischen Interesse und Engagement, die mit verschiedenen Lebensaspekten Unzufriedenen, die in Haushalten mit hohem Unit-Nonresponse lebenden und diejenigen mit schlechterer Befragungsqualität. Dieses Muster bleibt im Lauf des Panels bestehen. Die hohe Attrition in zwei Panelwellen, die vermutlich von singulären Ereignissen herrühren, ist nicht übermäßig selektiv.

1 Introduction¹

The major purpose of a household panel survey, in which individuals in households are surveyed repeatedly over waves, is to represent the real dynamics in the sample population. If individuals other than non-sample cases drop out of the panel ("attrition"²), this has at first obvious consequences on the longitudinal sample size. However, other than making analyses impossible due to cell sizes becoming too small after some waves, or merely producing higher standard errors in descriptive statistics, a selective attrition may in addition lead to wrong conclusions of important measures under consideration. For instance, in a recent analysis, Stocké and Stark (2005) show with data from the Eurobarometer³ that due to listwise deletion of individuals because of item nonresponse the share of persons going to a vote is around 9% higher than including the whole sample. Because the mechanisms leading to item nonresponse in a cross-sectional survey or to attrition in a panel survey might originate from similar factors (Loosveldt/Pickery/Billiet 2002, and the literature review below), we generally suspect a selective attrition in longitudinal surveys. E.g. in the British Household Panel Survey (BHPS), several refusal conversion techniques are being applied, showing that converted refusals have characteristics which are partly distinct from easier to convince respondents (Burton et al. 2004). Thus the characteristics of the members of a panel might well change after a longer time period.

Attrition is usually modelled and predicted with the help of standard socio demographic variables collected in a former wave. E.g., in the German GSOEP the size of the community a respondent lives is significant for the odds of a successful contact (Spiess/Kroh 2005). In the European Household Panel (EHP) the individual longitudinal panel response can be explained to a good extent by the *socio-demographic variables* age, employment status (i.e., full-time vs. not), and partnership (Nicholetti/Perrachi 2005). However, it is important to note that the socio-demographics are "fallible: they are correlates, not causes of the survey participatory behaviour" (Groves/Couper 1996: 81). This is emphasized also by Stoop (2005), who specify these causes for (non)cooperation: "social isolation, social participation, ..., interest in societal well-being, doing voluntary work, political interest and knowledge, ..., electoral participation, the type of sponsor, and attitudes towards surveys" (p. 126). Therefore, if available, variables measuring political interest and social participation (e.g. Pickery/Loosveldt/Carlton 2001), and item nonresponse (INR) on difficult (Loosveldt/Pickery/Billiet 2002) or sensitive (Schräpler 2004) questions to include motiva-

1 This work is supported by the Swiss National Science Foundation. I wish to thank my colleagues at the SHP for their valuable comments. Eric Graf suggested to conduct an in-depth correlation analysis, which led to the finally used aggregated health and satisfaction parameters. I would like to thank three anonymous reviewers for helpful comments. Any responsibility for the analysis or interpretations presented in this paper is however with the author.

2 With the term attrition we refer to all drop outs of a panel survey, i.e. refusals (non-cooperation) and non contacts of all *interview eligible* individuals (i.e. all who continue to be part of the sample: all who did not decease, are not being institutionalized, or for whom a valid reason for a proxy interview is given; see for the latter in the case of the SHP <http://www.swisspanel.ch/project/participation/index.php?lang=en&pid=53>).

3 http://ec.europa.eu/public_opinion/index_en.htm (3.7.2007).

tional factors are usually used in analyses on attrition. Groves et al. give the theoretical concepts to explain unit nonresponse in surveys (2004: 176):

"The theoretical perspectives that are most commonly applied to survey participation include

- '*opportunity costs*' (which is) based on the notions that busy persons disproportionately refuse to be interviewed because the costs of spending time away from any other pursuits is more burdensome than for others ...
- '*social isolation*', which influence persons at the high and low ends of the socio-economic spectrum to refuse survey requests ...
- '*topic interests*' (and motivation), which fuel hypotheses about how the interested disproportionately responding ...
- '*oversurveying*' that suggests fatigue from survey requests."

In addition to the cross-sectional factors, there is in addition the longitudinal aspect that the reasons to attrite from a panel need not be constant: De Keulenaer (2004) analyzing attrition in the Panel Study of Belgian Households, finds "that the effects of SES variables decrease with additional wave ..., while the effects of the variables describing the *interview situation* increase ..." (p. 5).

In the literature on attrition evidence is reported for all issues but to opportunity costs. As to *opportunity cost* related variables, one might e.g. think of interview time. In the US-Panel Survey on Income Dynamics (PSID), although proved to be negatively correlated with panel attrition, interview length has found not to be causally related to attrition (Zabel 1998). In the Household, Income and Labour Dynamics in Australia (HILDA) panel survey, a non significant relation exists between attrition and interview time (Watson/Wooden 2004). The latter state in this regard that "Interview time, ... is a product of instrument length, respondent interest in the survey, and respondent difficulty with the questions. Consequently, we expect the interview length to comprise a mix of respondents, some of whom found the interview very difficult and others who enjoyed the experience." (p. 302). Moreover, different interview length in the same survey is largely a matter of filter complexity: those who find themselves in more filters, usually have a specific socio-demographic status. This applies especially to individuals who are in the labour force, and have therefore to answer a number of work related questions. It can thus be expected that interview length correlates with the socio-demographic status.

As to *social isolation* factors, Watson and Wooden (2004), using the HILDA panel survey, find that attritors after one wave are more likely to have reported lower life satisfaction levels, are more likely renter of a flat rather than owner of a house or a flat. Also housing conditions have been shown to be significant in terms of attrition in other large panel surveys (Watson and Wooden 2004, Fitzgerald et al. 1998, Zabel 1998, Gray et al. 1996). However, in the HILDA survey, when controlling interview situation in the first wave, these indicators largely lose explanatory power.

As bad health prevents often from active social participation, we subsume *health* conditions under this category: Gray et al. (1996) analyzing attrition in a survey of health and life style in Great Britain, state that "those characteristics which were found to be related to attrition ... are

smokers, the less sporting and those who did not feel a part of the community" (p. 171). However, there was "little or no relationship between the health and psychological variables and a person's conscious decision to participate ..." (p. 182).

With respect to *topic interests* and *motivational* factors, in cross-sectional analyses little political interest and social participation and many "don't know" answers are shown to be strongly correlated with little political knowledge and weakly pronounced attitudes of the individuals as well as a general negative attitude toward surveys (Stocké/Stark 2005). Such (non)response behavior can be considered as resulting from a lack of cognitive efforts and disinterest toward the survey (Pickery/Loosveldt/Carton 2001; Loosveldt/Pickery/Billiet 2002, Schräpler 2004, Stocké/Stark 2005). The within household response rate is a good indicator for household specific motivation as according to Watson and Wooden (2004), in the HILDA and in the BHPS surveys, "in line with the results ... for the BHPS, we see that coming from a partially responding household is a major risk factor for non-participation at the next wave" (p. 302).

Loosveldt/Carton (1997), analyzing the respondent's decision to participate in the second wave of a Flemish election panel survey, find that the ability to provide an answer during the first interview plays a crucial role. Respondents who have problems to answer the questions and are difficult to interview are more likely to refuse to be interviewed in the second wave. The respondent's knowledge about the surveyed aspects and reasons for a survey is therefore crucial for panel participation. This correlates with the experience of a "pleasant" interview during the first panel wave, which seems to be an important issue for further panel stay (Loosveldt/Pickery/ Billiet 2002).

As to the answers given to subjective categorical questions, Pickery/Loosveldt (2004) view the non-occurrence of at least some extreme category answers as proving a low interview quality, because "One can expect that even respondents without a pronounced opinion will use the extreme response options now and then, especially when different scales are considered. If they do not, they probably do not expend the effort required. ... this is a form of satisficing" (p. 9). Similar results can be concluded with respect to the use of the midpoint category (see e.g. Tourangeau/Rasinski 1988). Scherpenzeel (2002) considers an excessive use of middle category answers as a proof of a low motivation to conduct the survey.

Regarding *panel specific aspects*, in the US Panel Study of Income Dynamics (PSID), the cross sectional representativeness seems to remain "roughly intact" (Fitzgerald et al. 1998: 251). However there is evidence that attrition is correlated with higher levels of "turnover and stability in earnings, marital status, and geographical mobility" (op. cit., p. 296). Also Watson/Wooden (2004) find more attritors among those in the HILDA panel who change their address more often.

With few exceptions, in the literature, only two panel waves are used in order to analyze panel attrition.⁴ As we are especially interested in the stability of the attrition variables across waves, we model the successive transitions between waves, but do also built a longitudinal all-wave model.

4 One exception is Rendtel (2002), who uses econometric models to analyze whether attrition is missing at random.

Specifically the article is organized as follows: First, we present the Swiss Household Panel (SHP), the data used for the analysis. To get a better impression on magnitude and selectivity of nonresponse and attrition in the SHP we describe the first wave nonresponse process, before we compare the distributions of the first wave respondents and its stayer subset after five waves with those in the BHPS and the ECHP. Next the modelling variables used are described in more detail. In a first modelling approach, these variables are used in a year to year attrition analysis. Here we will in particular investigate which variables are significant and whether the significant variables are cumulative rather than compensative as regards to attrition. Next a longitudinal model is built and analyzed. Most importantly, wave interaction terms will show whether certain covariates deviate from the base attrition model in specific panel waves. The results are discussed before the last section concludes. Note that we do not consider interviewer effects on nonresponse in this work.⁵

2 Data: The Swiss Household Panel

In this work we use data from the Swiss Household Panel (SHP), an ongoing, nationwide, yearly conducted, centralised CATI panel survey on the Swiss residential population. Questions are about household composition and socio-demographics, health, well being and attitudes, politics, social networks, and economics. The SHP started in 1999 with slightly more than 5000 households. In the SHP, in each year first the household composition together with the relationships between all household members, and the basic socio-demography is asked of the household reference person in the grid questionnaire. The household reference person is an adult who is sufficiently knowledgeable of the household characteristics, including the household finances. The grid questionnaire completion takes three to ten minutes, depending on household size and complexity of relationships. Then, a household related questionnaire is to be completed (about 10 minutes), again by the reference person. After this information is given, each household member from the age of 14 on has to complete his/her own individual questionnaire (about 35 minutes).

2.1 First wave nonresponse

In Switzerland, survey analysts and researchers face comparatively high nonresponse rates⁶, and subsequent attrition in the case of a panel survey. In the first wave (1999), questionnaire response in the SHP is as follows, in the different steps (MIS-Trend 2000):

⁵ An attrition analysis using the SHP together with interviewer characteristics can be found in Lipps (2006b).

⁶ See for a cross-country comparison of the response rates of e.g. the European Social Survey (ESS) Stoop 2005. Switzerland ranks at the very bottom within all countries involved in the ESS. A probable reason for the high nonresponse rate might be over-surveying in Switzerland, see Budowski/Scherpenzeel (2005) for the special case of the SHP. Note that similar to Nicoletti/Peracchi (2005), if we talk about nonresponse in the first wave, we distinguish between noncontact and refusal.

1. out of the 14,174 (gross) addresses drawn from the national telephone register and called by the interviewers, 1,025 were no valid telephone numbers (fax etc.). This leaves 13,149 net addresses.
2. out of the 13,149 net addresses, 1,065 could not be reached (i.e. 8.1% non contact rate).
3. out of the 12,084 contacted households, 2,712 (22.5%) are non-sample cases (i.e. business lines, language problems, etc.), 2,309 of the remaining 9,366 households (24.7%) refused to complete the grid questionnaire.
4. out of the 7,057 households who completed the grid questionnaire, 1,062 (15.0%) refused to complete the household main questionnaire.
5. out of the remaining 5,995 households, in 921 households (15.4%) all individuals refused to complete their individual questionnaire.

This leaves us with a household net response rate of 48.6% (5,074 "completed" households of 10,437 sample households), i.e. the grid and household questionnaires are completed and at least one household member filed his/her individual questionnaire.

On the individual level, according to the screening results from the household grid questionnaire, there are 10,293 individuals living in the 5,074 participating households. Of these, 921 (9.0%) are non-sample cases (language problems, illness, etc.). Of the remaining 9,372 persons, 1,573 (16.8%) refused, leaving a sample of 7,799 respondents.

Because, apart from the geographical region, there is no information on the gross sample, sample selectivity can in principle only be calculated using information of the households who completed the grid questionnaire. Based on the screened households, it can be shown that foreign households are underrepresented. Within households, males, younger individuals, and again persons with foreign nationality answer to the survey in a worse way (Cornali/Vonlanthen 2001).

In all it can be assumed that due to nonresponse, in the first SHP wave especially foreign individuals are underrepresented to a quite strong degree.

2.2 Attrition in the SHP compared with other panel surveys

Despite various measures to motivate panel participants (Budowski/Scherpenzeel 2005), the SHP faces a relatively high attrition of around 17% per year. This figure is higher than in other well established large (mostly CAPI) panels like the German SOEP (e.g. Kroh/Spiess 2005), the British BHPS (e.g. Burton et al. 2004) or the US PSID⁷. However we talk about attrition in a comparatively restrictive longitudinal manner: we only include persons who already answered the individual questionnaire in the first wave, and thus – other than sometimes done in other panels – do not include any new entrants into the panel⁸. Nevertheless, e.g. the first sample in the SOEP faces a longitudinal attrition of 10% during the transition from wave 1 (in 1984) to wave 2, 7% from

7 See the various articles in the Journal of Human Resources 33 (2), Special Issue: Attrition in Longitudinal Surveys, 1998.

8 These concern new household members who are therefore not original or longitudinal panel members (Naud/Latouche 2001).

wave 2 to wave 3, 5% from wave three to wave 4, and declines to around 2% in the long term (Kroh/Spiess 2005: 21). Similarly, the PSID has an attrition of 12% from wave 1 (in 1968) to wave 2, and about 3% in the years to follow (Fitzgerald et al. 1998: 254). Also in the BHPS, after an attrition of longitudinal first wave (in 1991) respondents of around 12%, attrition reduced to less than 5% from wave three on (Burton et al. 2004: 4).

We first compare some important distributions of the SHP total original and five waves stayers sample, with those from the BHPS (Lynn et al. 2006). The BHPS started in 1991 and was an example for the SHP with respect to design and content. However the BHPS interviews are conducted face-to-face. In order to have comparable samples, we compare the total 1999 sample persons with those who validly report during all five waves between 1999 and 2003 in the SHP. Also we drop all individuals who become (known) nonsample cases within this time period. We have the following sample sizes:

Sample	SHP	BHPS
# Respondents eligible in all waves until wave 5	7654	10264
# Participants first wave, with all waves until wave 5	3891	7246
First 5 wave response rate [%]	50.8	70.6

The attrition in the SHP is almost 20% points higher than in the BHPS.

Sex	SHP all	SHP 5w	BHPS all	BHPS 5w
Proportion of men [%]	43.6	42.3	47.7	46.2

With about 1.4% point differences between the men's share in the total and the stayer sample, the differences are about the same in both surveys.

Age [%]	SHP all	SHP 5w	BHPS all	BHPS 5w
14 (BHPS: 16) - 24	15.2	10.4	15.9	14.6
25-34	19.0	17.8	19.1	20.0
35-44	23.2	25.7	17.5	19.1
45-54	18.9	20.9	14.5	15.0
55-64	12.2	13.7	12.9	13.3
65-74	8.1	8.5	12.1	11.9
75+	3.4	3.0	8.0	6.1

The differences with respect to (starting) age groups are considerable between the SHP 1999 stayer and total sample on one hand, and the BHPS 5 wave respondents and total sample on the other. In the SHP the selection is mostly due to the youngest age class. This may be an indicator of problems with tracing of movers by the survey agency.

Marital Status [%]	SHP all	SHP 5w	BHPS all	BHPS 5w
Divorced or separated	6.7	6.7	5.7	5.8
Living as couple	7.4	6.6	6.3	6.4
Married	57.8	63.4	58.3	61.0
Never married	23.7	19.1	20.7	18.6
Widowed	4.5	4.3	9.1	8.2

The differences in the marital status between the 1999 five waves stayer sample and the 1999 total sample in the SHP on one hand, and the 1991 total sample and those, who responded all waves until wave five in the BHPS on the other, while higher in the SHP, are not tremendous.

Education [%]	SHP all	SHP 5w	BHPS all	BHPS 5w
High (Degree)	31.2	34.6	23.3	25.8
Middle (Level, Other)	47.5	47.3	41.4	42.2
Low/ No Qualification	21.3	18.1	34.7	31.4

The SHP and BHPS differences in percentage points between the total and stayer samples with respect to education are about the same.

Household Size [%]	SHP all	SHP 5w	BHPS all	BHPS 5w
1 PHH	17.4	16.8	13.7	13.1
2 PHH	30.0	29.5	33.4	33.7
3 PHH	15.5	14.6	20.4	20.4
4+ PHH	37.2	39.1	32.5	32.8

Compared to the BHPS, larger households are slightly overrepresented in the SHP after five waves.

Household Income: Quintiles	SHP all	SHP 5w	BHPS all	BHPS 5w
Lowest	20.0	16.7	20.0	17.5
2	20.0	19.8	20.0	19.2
3	20.0	20.7	20.0	20.8
4	20.0	21.1	20.0	20.9
Highest	20.0	21.4	20.0	21.6

As to income quintiles, the SHP distribution is not very different from that of the BHPS.

General Health	SHP all	SHP 5w	BHPS all	BHPS 5w
Very well	33.5	34.0	28.2	29.2
Well	50.4	51.7	45.0	45.7
Average	13.8	12.7	18.6	17.7
Bad	2.0	1.6	6.2	5.6
Very bad	0.2	0	2.0	1.7

Both distribution differences are similar.

Qualitatively similar results to those of the BHPS are reported from other panel surveys (e.g. the US PSID Fitzgerald et al. 1998). Generally, it appears that – given the much higher attrition in the SHP – higher attrition in the SHP does not automatically mean higher (cross sectional) selectivity. One exception is the higher drop out of younger persons.

In the following, we compare the most important socio-demographic attrition ratios (proportion in total sample / proportion in stayer sample after five waves) with those of the ECHP. For the latter, the attrition ratios from the middle 80 percentile countries are shown, omitting the lower and upper 10%. The ratios of the first and the last waves are depicted, where last wave means between third and fifth wave, depending on ECHP country.

Figure 1 Attrition ratios (proportion in total sample / proportion in stayer sample after 5 waves) in the European Community Household Panel, Swiss Household Panel, and British Household Panel Survey

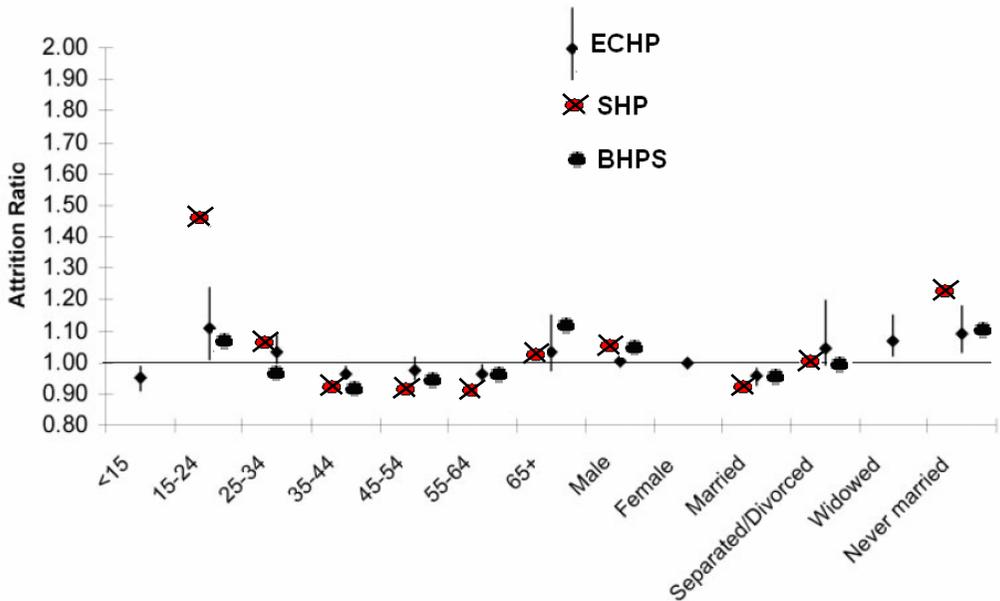


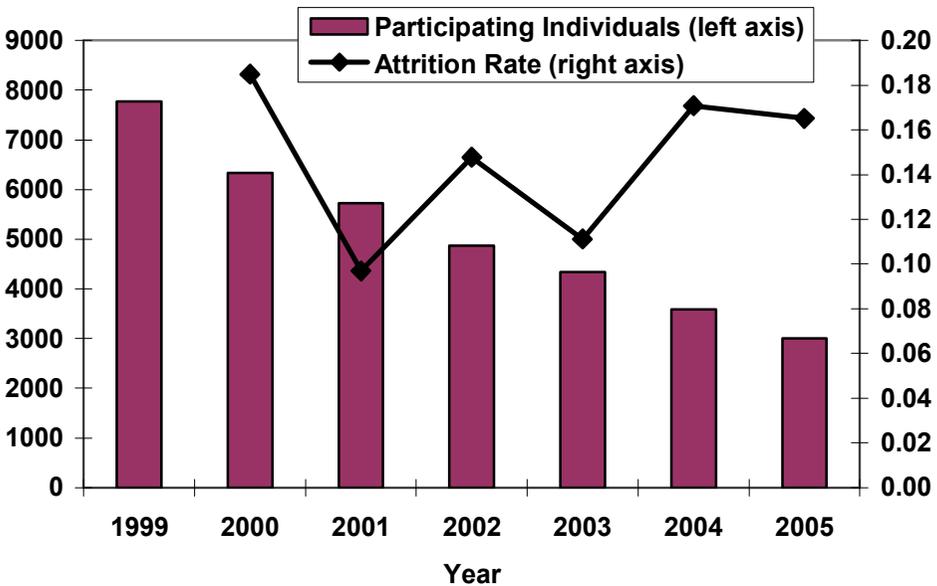
Figure taken from: Watson (2003), Figure 1.

Again, the problem to keep younger people in the sample becomes apparent.

3 Panel Survival in the SHP

The longitudinal sample in the SHP, which is still interview eligible in the next wave, started in 1999 with 7769 individuals. In 2000, this sample size drops to 6333, to 5719 in 2001, to 4874 in 2002, to 4332 in 2003, 3592 individuals in 2004, and finally ends up with 2999 in 2005. We depict the SHP wave specific number of participating individuals together with the attrition rates in Figure 2.

Figure 2 Number of participating eligible first wave individuals and attrition rates in the Swiss Household Panel, wave 2 (2000) through wave 7 (2005)



The main problem of the SHP is that the attrition rate does not decline: Although the attrition in the second wave is comparatively high, it is not outrageous, especially taking into account an "oversurveying" effect prevalent in Switzerland (Zimmermann/Joye 2003) with harmful effects on response rates and presumably attrition.

We like to assess the consequences of the two major panel specific "events" in the history of the SHP:

- The biographical questionnaire, a P&P self completion sent to the respondents in May 2002 (Budowski/Scherpenzeel 2005). A small part of the sample was used for the 2001 pretest, but this concerned only around one sixth of the whole sample.
- To notify the respondents of a duration of the panel survey of five years, before the first wave started. This was the time horizon of the initial project funding, which the respondents were deliberately told (Budowski/Scherpenzeel 2005).

Although the 2001 pretest of the biographical questionnaire did not show any adverse effects toward the SHP CATI response rate in the next wave (Scherpenzeel et al. 2002), we are now in a position to analyze effects from the biographical main questionnaire. In terms of the "five wave announcement", we expect a negative influence on those who were difficult to convince to take part in the first five waves, and who might have an argument to drop out in the sixth wave. Remarks noted by the interviewers while asking the household grid questionnaire during the sixth wave like "In the beginning it has been told that the survey lasts five years and I already took part for five waves but have enough now" supports this hypothesis.

Apart from the expected high attrition in the second wave (2000) we find consequently two peaks in Figure 2, in 2002 and in 2004. These coincide with the "events" in the SHP. We therefore find the expected higher attrition rates supposedly due to the biographical questionnaire and the "five wave announcement". The all wave model to be discussed later will give more insight in the covariate influence on attrition in these particular two waves.

In the following attrition analysis we combine all kinds of possible drop-outs, i.e. due to refusals, non contacts, or others (De Keulenaer 2004). The reason for this decision is first that in a CATI survey like the SHP, from the second wave on, non-contact is a minor problem compared to refusal at least regarding its magnitude. This is also due to the easier respondent tracing by the survey agency based on the information from previous waves. Secondly, a true refusal may easily be confused with a noncontact: this may be because the respondent is alerted by the advance letter announcing the survey call some days before and therefore does not answer the phone once the expected telephone number appears on the display. Moreover, it happens sometimes that a respondent Y disowns another eligible member X in the household, telling that X is not available, on X's behalf. What is usually coded noncontact is a true refusal. In addition, a metadata comparison of the characteristics of those who refuse and those who did not continue to respond due to other reasons shows that both groups are not very different with the exception of age and household size (Gray et al. 1996, Table 1). We nevertheless include covariates which can be expected to be strongly related to noncontact rather than refusal, such like age, or whether the household expects to move in the next 12 months, if available. With respect to the former, we already realized that young adults drop out of the SHP to a higher extent which points to problems of tracing moving households.

4 Independent Variables

We model attrition with the help of a bundle of *socio-demographic* variables on one hand, including the household composition (number of adults 18 years or older, number of children under the age of 18), sex, age⁹ in seven groups (14-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+), whether the individual is married, education on a binary scale (median cut), and finally the employment status, i.e. full time or part time employed, unemployed, or out of the labour force.

We analyze the influence of having Swiss citizenship, and the interview language. The latter variable is also dichotomized, as within the three SHP survey languages German, French and Italian, about 73% of the Swiss residents speak German as first language (Swiss Federal Statistical Office 2005). We further distinguish between the seven Swiss regions "Lake of Geneva", "Middeland", "North-West Switzerland", "East Switzerland", "Central Switzerland", "Zurich", and "Ticino". In addition we investigate the size of the community in which the respondent lives, measured by a binary variable which amounts to 1 if the respondent lives in a large or middle sized community ("Grosszentrum" or "Mittelzentrum"), and 0 else.

Also whether the household moved since the last panel wave is taken into account. Moving is an important predictor of the odds to making contact with a household. However it is a priori not clear whether moving has an influence on response, once contact is established. In the HILDA survey at least, the likelihood of obtaining an interview is independent of moving. In order to assess whether the survey institute successfully tracks households, the intention to move is taken into account in the attrition model as well.

In order to measure panel motivation using information of the degree of *social isolation*, we include house ownership as this variable was found to discriminate grid response (Naud/Latouche 2001). In order to approach disposable income – a direct use of income is not possible because the missing values are not (yet) imputed in the SHP, a variable which counts various potential bad states of the dwelling (6 at maximum) as well as flat or house ownership is used.

Next we investigate a bundle of satisfaction variables; all satisfaction variables are equally measured on an 11pt scale, where 0 means completely unsatisfied, and 10 means completely satisfied:

- satisfaction with the financial situation of the household
- satisfaction with living together with the other household members or with living alone, respectively
- satisfaction with the activities in the free time
- satisfaction with the amount of leisure time

These satisfaction measures are standardised and then combined into an aggregate parameter.¹⁰

⁹ Age is measured in 1999 and kept constant afterwards, i.e. we measure initial age.

¹⁰ Worth mentioning is that because of the aggregation, we encounter only few missing values due to item non-response: a variable containing a missing value is simply not considered for the individual concerned for the mean value calculation. The mean parameter consists then of fewer variables.

Health reasons are very often stated when people argue why they do not like to continue to take part. In order to avoid collinearity, we combine the following (standardised) health related variables into one aggregate parameter ("health problems") which is the mean of the following variables:

- the degree to which the self rated health is bad
- whether at least on one day during the last years health problems were encountered
- whether there are impediments in everyday activities due to health problems
- satisfaction with health
- whether medications are needed
- the degree to which the person feels anxious in terms of her health

The extent to which health has improved during the last 12 months is not added to this aggregated health parameter, but kept as a single *health change* measure.

As to the *topic interest and motivational* factors, we investigate a series of variables which are a measure of societal and political interest and participation. We consider in more detail the variables, which are also originally measured on an 11pt scale

- interest in politics
- in degree to which the respondent has trust in the federal government (compare Stocké/ Stark 2005)
- the self assessment of his or her political influence
- and the number of times one would go to ten possible votes.

These variables enter a common parameter "political interest".

In order to first distinguish the rather theoretical parameter political interest from actual or potential political participation and because a correlation analysis shows that the variables between the two groups correlate only rather weakly, we define a second aggregated political parameter "political participation", which is the mean of the following standardised variables:

- the extent one is willing to take part at demonstrations
- the degree one would take part at strikes
- the extent one would participate at boycotts

We furthermore include the political left-right political orientation. Also this variables is standardised. We further consider whether the individual is active in voluntary work and does sports at least once a week.

As *survey related motivation variables*, we include the household response rate from the preceding wave. Furthermore, we introduce the dummy "reference person" from the last year in multi-adult households, as Lipps (2006b) shows that response propensities of former reference persons are much higher than of other persons in the household.

Next, we include variables measuring the interview quality of the preceding wave. The variables indicate whether at least one of the 23 subjective attitude questions present in all waves has not been validly answered (subjective questions noanswer), the number of extreme and middle category answers to these questions, as well as not providing enough information for the data editor to be able to calculate total individual income.

As to the extreme categories, we suspect that not only an underuse, but also an overuse of extreme categories may be a form of satisficing and therefore include the number of all extremes in the data. The same is done for the midpoint (cat 5) category.

In order to avoid deletion of records due to missing values other than the validly coded INR categories "don't know" or "no answer", the (few) missing values are recoded to the respective modal value.

5 Single Waves Models

The single wave models aim to getting an idea which person groups do especially attrite in which wave. In each single wave model, we include all longitudinal persons who give a valid interview in the first of the two transition waves under consideration and are still interview eligible in the second wave. As we are especially interested in the identification of even weakly significant attrition variables, we build logistic forward regression models with a comparatively small inclusion level equal to .01. As mentioned above, we consider whether the household intends to move within the next twelve months in these single wave models. This variable is not available in the years 2000 and 2002. To interpret the coefficients and their significance notice that they are not comparable due to the different sample sizes of the models.

The results of the forward regression models are depicted in the left columns of Table 1 and Table 2. To facilitate an overview, we denote in Table 1 only the significant variables, with a plus sign if the variable correlates positively with attrition in the wave considered, or a minus sign if the variable correlates negatively with attrition in this wave. The full coefficients for the significant variables at the 1% level are listed in Table 2.

We first compare significant variables in the different variable groups possibly susceptible for attrition. Evidently there are large differences: within *socio-demography* especially moving factors and (young) age are crucial. Nationalities other than Swiss show higher attrition rates, but only in the second wave. Spatial issues thus are not of major importance as long as they do not measure the main language regions. In the *social isolation* variable group, only satisfaction is important, while housing does not play a role at all. *Topic interest and motivation* play a major role in all waves, as well as *survey status*. That health problems are of minor importance in each wave is surprising, because health problems are often a main reason for individual refusal, as far as reasons for dropping out are given in the grid questionnaire.

A second immediate notion is that if an attrition variable is significant in several waves, attrition is always cumulative rather than compensating. This is consistent with findings in the literature (e.g. Gray et al. 1996). In addition the magnitudes are rather stable over time, as can be seen in Table 2. Thirdly we notice that the pseudo R^2 statistics range between .05 are thus not tremendously high. This shows that, although we included many covariates trying to explain attrition variation across different variable groups derived from theory, there are still other reasons responsible.

In the following, we identify and interpret the selective effects in the single transition models.

As to the *socio-demographic* variables, there are no particularly surprising results. The number of adults in a household does not discriminate attrition until the year 2004. In this year, persons living in households with two (or more) adults attrite way above average. This may be due to the above stated "five wave announcement". Households with several adults may be more sensitive to such announcements, perhaps more so because they face a higher response burden due to having to answer several individual questionnaires.

For the variable number of children in the household on the other hand, we find negative effects as regards to attrition in the first and the third transition. An explanation is that households with children have a more stable lifestyle, and are much easier to be found and contacted. After four waves however, the remaining households in the panel and the remaining family households seems to be in "balance" insofar, as there is no difference in the attrition rate from wave five on.

The same stability argument applies for the middle aged groups between 50 and 59 years of age, and for the married persons; for the former there is a negative attrition effect in the second and especially the fourth wave, which further cumulates in the sixth wave. Very severely however, the adolescent individuals and particularly the younger adults attrite to a very high degree constantly from the second wave on. With respect to age distribution, this may lead to an "overaging" of the panel sample in later panel waves. This high attrition is an alarming signal regarding tracing of moving respondents, as moving is more prevalent in this age group. This finding is strengthened by the very high attrition rate among those who intend to move within the next 12 months. As can be seen in Table 2, the odds to attrite of the "willing to move" people do not decrease over waves; i.e., tracing efforts undertaken are not increasingly efficient.

As to the older cohort aged 70 and older, we encounter a higher attrition in the second wave which does not continue later. This is surprising, as due to increasingly bad health, deceases, institutionalisation, or simply a too high response burden, one might expect a constantly higher attrition in this age group. Their higher stability in life seems to play a higher role than a higher "natural" transition probability towards ineligibility.

Also unanticipated is that education or being full time employed or jobless is never significant. This may be a result from a conflict between the higher interest in the topic by higher educated or full time employed people and their tougher time budget. That the unemployed do not show a higher attrition rate could not be expected from the literature. In fact the attrition is as high as for the reference group, which consists of those who are not in the labour force. In this respect, both interest in the survey and response burden seem to be important factors. The contrary may hold for the part time employed: they show a smaller attrition in the third and especially in the sixth wave, as response burden may play a minor role for this group.

Investigating nationality, Swiss citizens attrite to a smaller extent, as can be expected from other panel surveys. Swiss-German speaking people show a high risk to attrite only in wave five.

This partially complements the slightly higher attrition rate of the East-Swiss who all belong to the German part of Switzerland in the third wave. Second, it may result from an institution effect: the "five wave announcement" may have been communicated more explicitly in the German speaking CATI centre compared to the French speaking centre.

Not surprisingly, the intention to move decreases the odds of a contact and therefore increases the probability of attrition. Households who moved since the last interview show a significantly higher participation rate: this is probably due to the fact that the propensity to move right again is smaller for these households.

Concerning *social isolation* factors, housing variables, which can be used as weak proxies for wealth (ownership) and income (state of housing), do not play a role at all. This comes as a surprise. Similarly, neither health problems nor health improvements largely affect attrition in the subsequent wave. This fits well to the non-increasing attrition of the elderly after the second wave and shows that health per se does not affect attrition. On the other hand, the aggregated satisfaction parameter is significant through the first waves.

The most interesting effects stem from the variables measuring *topic interest and motivation*: expectedly, being interested in politics and potentially participating in political or societal activities largely decrease the odds to attrite. This is partially also true for involvement in voluntary work. Having a left rather than a conservative political orientation on the other hand has no effects. Overall, the lowest attrition can be expected from those with a high political interest. This pattern holds virtually during all panel waves.

The timing of the influence is also interesting: those who are involved in voluntary work exhibit a decreased attrition only from wave four on, whereas individuals who participate in politics show a particularly low attrition in the second and third waves. Perhaps the latter do act more immediately and to a stronger degree.

Regarding the data quality characteristics of the interviews preceding the wave under investigation, especially those who use many extreme categories answers to the subjective questions have highly positive odds to attrite from wave three on. In fact, this intra-individually rather stable variable¹¹ seems to be an excellent measure to assess attrition in the next wave, and does not correlate high with the other variables considered. A possible interpretation for the strong effect on attrition is that giving a high number of extreme categories answers is an indicator of little substantive interest in the survey. A similar albeit much weaker result holds for those who give a high number of middle category answers. The latter effect is however prevalent only in the fifth panel wave.

Not giving a valid answer to at least one of the subjective questions has no significant effect at all; however a not valid answer to one of the income questions has strong positive attrition effects in the second wave. This should have positive consequences on the income nonresponse in subsequent waves.

11 The correlation coefficient between two waves never falls below the value of .5.

Finally a high within household response rate leads to a constantly highly significant lower attrition in the next wave. This confirms that other household members' disposition in a preceding panel wave has strong impacts on the own participation behavior (Lipps 2006).

Similarly one might wonder why the reference person of the previous wave, who has to answer the household grid and preferably the household questionnaire, attrites to a smaller extent than other persons in the household. On one hand, these tasks speak in favour of the hypotheses that her commitment to the panel is from the beginning stronger than those of other individuals. However, all characteristics are measured in the year before the possible attrition under investigation occurs: It may equally be true that the previous year's reference person suffered such a high response burden due to not only having completed the grid and household questionnaire, but also her own individual questionnaire that she is more likely to attrite. Nevertheless we find a cumulative negative attrition of the reference person, so the first hypothesis is confirmed.

We note in addition that the 2002 and the 2004 ("panel events") models differ only in some minor respects from the other models: as to the "five year announcement", in 2004 there is by and large a higher attrition especially by those who live together with other adults, a lower for part time employed and again persons engaged in voluntary work. A possible interpretation may be that especially those with a higher household specific survey burden have been especially bothered by continuing the panel despite the "five year announcement". However these findings should be interpreted carefully.

6 All Waves Model

In the second modelling step we are especially interested in testing for duration dependence, i.e. all things equal, we try to identify person groups for whom the base attrition rate differs and, in addition, whether this base rate is different in single waves or even shows a monotonous development over time. Second we like to identify potential effects of the two panel "events" biographic survey" and "five wave announcement" in wave four (2002) and wave six (2004), respectively, in the context of *all* panel waves.

In the SHP the decision is made to keep all persons in the sample who temporary refuse to take part, i.e., refusal only in one wave. Once an individual refuses for two consecutive waves, (s)he is excluded from further contact attempts. Because we like to keep all interview eligible individuals including the temporary attritors in the analysis, we cannot apply a true survival model.¹² Nevertheless, it is important to control for the clustering of waves within individuals when duration dependence is tested¹³ (Zabel 1998). Therefore we apply a two-level random effects model, thus treating the single individual residuals as random variables. This controlling constitutes the

12 See Lipps (2006a) for an attrition analysis of the SHP using only individuals with monotonous participation patterns.

13 I.e., all individuals contribute to the same extent, irrespective of their panel participation duration.

main difference vis-à-vis the single wave models, where significances depend on the sample sizes, which vary considerably between waves. Note in addition that we cannot estimate a random effect forward all wave model. Thus in this stage all variables are included, not only the significant ones. This may have implication on co-varying variables.

The qualitative results of the all wave model are depicted in the right columns of Table 1, the odds ratios of all coefficients in the right columns of Table 2.

Looking at the modelling results at the right columns of Table 1, it becomes clear that only few of the variables considered are significant in terms of a wave specific deviation from the base attrition rate. This means that only few person characteristics which discriminate attrition change over waves. In addition the significance of these "deviance variables" reaches the 1‰ level only in the case of the satisfaction measure in wave six. This reflects the stability of the base attrition rate over the waves, i.e. time is cumulative rather than compensating. E.g. people who are engaged in voluntary work attrite to a lesser extent constantly in every wave. That the high initial attrition does not significantly decrease over time can be seen by the insignificant wave specific effects.

There are some wave specific peculiarities vis-à-vis the results of the single-wave models. E.g., being married or part time employed has transition specific effects in the single wave models. However these effects completely vanish in the all wave model. These variables are captured in other now included variables, see the above remark about co-varying independent variables. Conversely, having health problems, then without effects, do significantly increase the odds of attrition in the all wave model. Overall however, in the all wave model, the base effect is "stabilised".

As we are especially interested in the effects of the panel events "biographical questionnaire" and "five wave announcement", we check the 2002 and 2004 columns in more detail. As to the fourth wave attrition effects, we identify a comparatively higher attrition by those who show a higher political participation, or are reference persons. Both effects affect the highly negative base attrition with respect to these characteristics in a positive way. This shows that person groups, who generally show smaller attrition rates, are particularly deterred by the additional burden. Regarding the "five year announcement" effects on attrition in 2004, we find a highly significant increased attrition rate for those who are more satisfied with various aspects in their lives. Again this particular wave effect compensate for the small base attrition rate of more satisfied individuals. These persons may have also attrited to a higher rate due to the higher response burden, because of their tougher disposable time budget.

Overall we find that individuals who generally show a higher panel loyalty attrite to a higher extent due to the "events" considered. This may prove their sensitivity toward additional survey burdens whose possible effects on future attrition rates should be carefully analyzed before they are implemented.

7 Summary and Conclusions

In this article we analyze attrition in the Swiss Household Panel (SHP) from wave two (2000) through wave seven (2005). We only include individuals, who already completed the individual questionnaire in the first wave, and are still interview eligible in the wave under consideration.

First, comparing attrition in the SHP with that in the British Household Panel and the European Community Household Panel, we show that although attrition is comparatively high in the SHP, it is not particularly selective with respect to important socio-demographic or -economic variables. However the problem to keep younger individuals in the sample is a challenge for the SHP – the more that this is the person group which moves and forms new households, thereby maintaining the (cross-sectional) sample size in the panel survey.

We use the characteristics of the respondent in the preceding wave as covariates to model independent logistic forward regression transition models in a first step. We find that the younger people, those without children in the households, those who intend to move, foreigners, those living in households with interview refusing participants or showing little interest and (potential) participation in politics and society, those dissatisfied with various aspects in life, and those who are not reference person in their household attrite to a higher degree than their respective counterparts. Only very weak effects are due to spatial aspects and the housing situation of the respondents, as well as their health status or physical activities exerted.

These findings are in line with attrition analysis results from other European Panel surveys (Buck et al. 2003; Watson 2003) and confirms the "social exclusion" theory (Groves/Couper 1998; Groves et al. 2004; Stoop 2005). In addition we find strong effects from interview quality characteristics in that those who exhibit an extreme answering behavior (overuse of extreme categories answers) or tend to refuse to answer income questions also attrite to a higher rate. An assumed satisficing behavior (overuse of middle category answers) or proving little cognitive effort invested in the answers (not answering the subjective questions) is not necessarily related to a higher attrition. The effects are in all models cumulative rather than compensative.

In the second step an all wave random effects model with wave interaction terms is estimated. We find that the base model is rather similar to a synopsis of the single wave models, and that the effects of the wave interactions are minor. In particular, although attrition in the waves with the two specific panel "events"¹⁴ is higher, there are neither significantly increased attrition effects from the two wave dummies nor a particularly increased attrition by special person groups. The small deviations from the base attrition in the two "event" years are comparatively small compensatory effects: those who generally show a smaller attrition are positively affected and vice versa. In general however, the mostly insignificant wave interaction effects show that the panel participants are affected to more or less a similar degree.

14 First the biographic survey in 2002 and second the assumed end of the panel due to communicating to the respondents that the panel is initially funded for five years at the start of the survey in 1999.

Our models have a comparatively small explanatory power. This shows that other than easily measurable factors used here from completed household or individual interviews also play an important role with respect to attrition behavior. The quality of the contacts of interviewers with respondents prior to an interview can give some hints for future research: E.g., while the respondent socio-demography is significant for the outcome of the first contact with an interviewer, Groves/Couper show that it loses its predictive power for those requiring more than one contact to obtain a final disposition (1996: 74, 1998: 255). First analyses to better understand the process leading to attrition are currently under way (see Lipps 2007), who uses call data from the SHP.

In order to reduce the high attrition the SHP, a couple of measures were taken for the 2006 wave (MIS-Trend 2007), partially based on the findings in this article: most importantly, an incentive experiment has been implemented among panel households¹⁵ in order to increase motivation. In addition, unlike the rule adopted in the SHP not to ask households who refused to participate for two consecutive waves, households, who did not answer during the 2004 and the 2005 waves, have been approached all the same. Finally, the notion of completed households has been changed¹⁶. First results show that these measures decreased attrition in the 2006 wave to a considerable extent.

Reference

- Budowski, M. and A. Scherpenzeel, 2005: Encouraging and Maintaining Participation in Household Surveys: The Case of the Swiss Household Panel. *ZUMA Nachrichten* 56: 10–36.
- Burton, J., H. Laurie and P. Lynn, 2004: The Long-Term Effectiveness of Refusal Conversion Procedures on Longitudinal Surveys. *ISER Working Papers* 2004–11.
- Cornali, A. and C. Vonlanthen, 2001: Description de la pondération de l'échantillon du Panel Suisse de ménages. *Swiss Household Panel Working Paper* 4_01, Neuchâtel.
- De Keulenaer, 2004: Socio-Economic Status Bias in Survey Nonresponse: An Analysis of the Panel Study of Belgian Households (PSBH). *Proceedings of the 57th WAPOR Annual Conference*, May 12–13, Phoenix, Arizona.
- Fitzgerald, J., P. Gottschalk and R. Moffitt, 1998: An Analysis of Sample Attrition in Panel Data: The Michigan Panel Study of Income Dynamics. *J. Human Resources* 33, Special Issue: Attrition in Longitudinal Surveys: 251–299.
- Gray, R., P. Campanelli, K. Deepchand and P. Prescott-Clarke, 1996: Exploring Survey Nonresponse: the Effect of Attrition on a Follow-up of the 1984–85 Health and Life Style Survey. *The Statistician* 45: 163–183.
- Groves, R. and M. Couper, 1996: Contact-Level Influences on Cooperation in Face-to-Face Surveys. *Journal of Official Statistics*, 12: 63–83.
- Groves, R. and M. Couper, 1998: *Nonresponse in Household Interview Surveys*. New York: Wiley.
- Groves, R. M., F. J. Fowler, M. P. Couper, J. M. Lepkowski, E. Singer and R. Tourangeau, 2004: *Survey Methodology*. Wiley Series in Survey Methodology.
- Kroh, M. and M. Spiess, 2005: Documentation of Sample Sizes and Panel Attrition in the German Socio Economic Panel (SOEP) 1984–2004. *Data Documentation* 6, DIW Berlin.

- 15 A random quarter of households received 1.) no incentives (control group), 2.) stamps unconditionally, 3.) participation at a lottery or 4.) a donation to be donated to a charity institution. 3.) and 4.) applied on the individual level after completion of the individual interview (conditional). Also incentives for interviewers have been implemented depending on collective response rate achievements.
- 16 Before the 2006 wave, a household has been considered „complete“ if the household grid and at least one individual interview was filled. From wave 2006 on, completeness means that all interview eligible household members complete their interview. The idea was to increase cooperation among those who are not reference persons in the household.

- Lipps, O., 2006: Analysis of Panel Participation in Couples using Interviewer Characteristics and the Partner's Behaviour. Swiss Household Panel Working Paper 3_06, Neuchâtel.
- Lipps, O., 2006a: Attrition and Motivation in the Swiss Household Panel. Swiss Household Panel Working Paper 2_06, Neuchâtel.
- Lipps, O., 2006b: Attrition of and within Households in the Swiss Household Panel – Roles of the Reference Person, Other Persons, and the Interviewer in the Course of Time. Swiss Household Panel Working Paper 6_06, Neuchâtel.
- Lipps, O., 2007: Cooperation in Centralised CATI Household Panel Surveys – A Contact-based Multilevel Analysis to Examine Interviewer, Respondent, and Fieldwork Process Effects. Submitted.
- Loosveldt, G. and A. Carton, 1997: Evaluation of Nonresponse in the Belgian Election Panel Study '91-'95. Proceedings of the Survey Research Methods Section, American Statistical Association.
- Loosveldt, G., J. Pickery and J. Billiet, 2002: Item Nonresponse as a Predictor of Unit Nonresponse in a Panel Survey. *Journal of Official Statistics* 18: 545-557.
- Lynn, P., N. Buck, J. Burton and H. Laurie, 2006: Quality Profile: British Household Panel Survey, Version 2, Waves 1 to 13, 1991-2003. Institute for Social and Economic Research, University of Essex.
- MIS-Trend, 2000: Statistiques finales. Unpublished Field Report of the SHP wave 1 (1999) data collection, Lausanne.
- MIS-Trend, 2007: Panel Suisse de Ménages. Rapport technique réalisé pour Le Panel Suisse de Ménages – Neuchâtel. Unpublished Field Report, Lausanne.
- Naud, J.-F. and M. Latouche, 2001: Weighting of the Swiss Household Panel. Wave 2 Detailed Description. Swiss Household Panel Working Paper 8_01, Neuchâtel.
- Nicoletti, C. and F. Peracchi, 2005: Survey response and survey characteristics: Micro-level evidence from the European Community Household Panel. *Journal of the Royal Statistical Society: Series A* 168/4: 763-781.
- Pickery, J. and G. Loosveldt, 2002: A Multilevel Multinomial Analysis of Interviewer Effects on Various Components of Unit Nonresponse. *Quality and Quantity* 36: 427-437.
- Pickery, J. and G. Loosveldt, 2004: A Simultaneous Analysis of Interviewer Effects on Various Data Quality Indicators. *Journal of Official Statistics* 20: 77-89.
- Pickery, J., G. Loosveldt and A. Carton, 2001: The Effects of Interviewer and Respondent Characteristics on Response Behaviour in Panel Surveys. *Sociological Methods & Research* 29: 509-523.
- Rendtel, U., 2004: Attrition in Household Panels: A Survey. Chintex (The Change from Input Harmonisation to Ex-post Harmonisation in National Samples of the European Community Household Panel – Implications on Data Quality) Working Paper 22.
- Scherpenzeel, A., 2002: Interviewer Effects in data from the Swiss Household Panel. Swiss Household Panel Working Paper 2_06, Neuchâtel.
- Scherpenzeel, A., E. Zimmermann, M. Budowski, R. Tillmann, B. Wernli and A. Gabadinho, 2002: Experimental pre-test of the biographical questionnaire. Swiss Household Panel Working Paper 8_02, Neuchâtel.
- Schräpler, J., 2004: Respondent Behavior in Panel Studies. A Case Study for Income Nonresponse by Means of the German Socio Economic Panel (SOEP). *Sociological Methods and Research* 33: 118-156.
- Stocké, V. and S. Stark, 2005: Stichprobenverzerrung durch Item-Nonresponse in der International vergleichenden Politikwissenschaft (with English Summary). Sonderforschungsbereich 504, Working Paper 05_43, University of Mannheim, Germany.
- Stoop, I., 2005: The Hunt for the Last Respondent: Nonresponse in Sample Surveys. PhD Thesis, Utrecht University.
- Swiss Federal Statistical Office, 2005: Eidgenössische Volkszählung 2000 – Sprachenlandschaft in der Schweiz. Neuchâtel, April 2005.
- Tourangeau R. and K. A. Rasinski, 1988: Cognitive Processes underlying Context Effects in Attitude Measurements. *Psychological Bulletin* 103: 299-314.
- Watson, D., 2003: Sample Attrition between Wavers 1 and 5 in the European Community Household Panel. *European Sociological Review* 19: 361-378.
- Watson, N. and M. Wooden, 2004: Sample Attrition in the HILDA Survey. *Australian Journal of Labour Economics* 7: 293-308.
- Zabel, J. E., 1998: An Analysis of Attrition in the Panel Study of Income Dynamics and the Survey of Income and Program Participation with an Application to a Model of Labor Market Participation. *Journal of Human Resources* 33, Special Issue: Attrition in Longitudinal Surveys: 479-506.
- Zimmermann, E. and D. Joye, 2003: Item Nonresponse: Ignorance or Refusal? Paper presented for the 2003 Workshop "Item Non-response and data quality in large social surveys", October 2003, Basel, Switzerland.

Appendix

Table 1 "Qualitative" influence on attrition

	Single Wave Models (single transitions)						All Wave Model (indiv. clustering controlled)					
	2000	2001	2002	2003	2004	2005	base	+2001	+2002	+2003	+2004	+2005
Wave effect (only all wave model)							Ref.					
SOCIO-DEMOGRAPHY												
Number of adults in household					++							
Number of kids in household	--		--				--					
Male	++						++					
age14-19	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Reference category					
age20-29	++	++	+									
age30-39						++	--					
age40-49							--					
age50-59	-		--		-		--					
age60-69			-				--					
age70-	+											
Married		--	-									
Education high												
Full time employed												
Part time employed		-			--							
Unemployed												
Swiss Citizen	--						--					
Language Swiss German				++						+		
Lake Geneva	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Reference category					
Middleland										Only base effect		
North-West CH										Only base effect		
Zurich										Only base effect		
East-CH		+								Only base effect		
Central CH										Only base effect		
Ticino										Only base effect		
Lives in Urban Centre												
Household moved within last 12 months	--	--	--	--	--	--	--					
Household has intention to move within next 12 months	.	++	.	++	++	++
SOCIAL ISOLATION												
House bad												
House owner												
Satisfaction with various aspects	--	-	-				--				++	
HEALTH												
Health Problems							++	-				
Health improved during last 12 months												
Do_sports				--						-		
TOPIC INTEREST AND MOTIVATION												
Engaged in voluntary work			-		-		-					
(Potential) Political Participation	--	--					--		+			
Political Left orientation							-					
Political interest	--		--	--	--	--	--					
Subjective questions noanswer												
Number of extreme categories		++	++	+			++					
Number of middle categories				+								
Income noanswer	++						++					
Response rate within household	--	--	--	--	--	--	--					
SURVEY STATUS OF INDIVIDUAL												
Person is Reference Person (only more-adult-HH)	--	--		--	-	-	--		+			
N	7769	6333	5719	4874	4332	3592	N=32619, p=0.20 (Intra-cluster coeff.)					
MCFadden Pseudo R ²	0.068	0.064	0.057	0.053	0.045	0.055	-					
LR chi ²	506	331	309	212	209	176	Wald chi ² (Deg. Freedom=180)					1485

"+"=positive, 1% significance level, ++=positive, 1% significance level, "-"=negative, 1% significance level, --=negative, 1% significance level. "-"= n.a. Single wave models: forward regression. All wave model: only significant (1%) effects indicated.

Table 2 "Quantitative" influence on attrition

	Single Wave Models (single transitions)						All Wave Model (indiv. clustering controlled)					
	2000	2001	2002	2003	2004	2005	base	+2001	+2002	+2003	+2004	+2005
Wave effect (only all wave model)							Ref.	.83	.88	.38	1.17	.97
SOCIO-DEMOGRAPHY												
Number of adults in household					1.19		1.05	.94	.97	1.02	1.13	1.02
Number of kids in household	.89		.84				.85	1.01	.98	1.06	1.15	1.01
Male	1.25						1.19					
age14-19	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Reference category					
age20-29	1.57	1.74	1.38				.96					
age30-39						.64	.57					
age40-49							.54					
age50-59	.78		.67		.75		.42					
age60-69			.65				.46					
age70-	1.48						.75					
Married		.72	.77				.99	.78	.86	.87	1.01	.96
Education high							.95	.94	.88	.91	.94	1.03
Full time employed							1.09	1.06	1.15	1.15	.98	1.11
Part time employed		.74			.71		1.1	.82	1.08	1.06	.71	.9
Unemployed							1.61	1	.69	1.27	1.23	.76
Swiss Citizen	.65						.61	1.15	1.27	1.36	1.23	1.66
Language Swiss German				1.58			1.23	1.14	.84	1.41	1.08	.73
Lake Geneva	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Reference category					
Middleland							.97					Only base effect
North-West CH							.95					Only base effect
Zurich							.90					Only base effect
East-CH		1.40					1.05					Only base effect
Central CH							.87					Only base effect
Ticino							.95					Only base effect
Lives in Urban Centre							.98	1.29	.98	.99	.98	.82
Household moved within last 12 months	.39	.38	.26	.21	.20	.18	.33	1.45	.77	.83	.71	.66
Household has intention to move within next 12 months	.	1.28	.	1.39	1.26	1.32
SOCIAL ISOLATION												
House bad							.98	1.03	1.06	1.06	.96	1.07
House owner							.83	1.25	1.11	1.07	.95	1.1
Satisfaction with various aspects	.80	.82	.79				.75	1.06	1.06	1.12	1.44	1.23
HEALTH												
Health Problems							1.31	.71	.95	.95	.77	.92
Health improved during last 12 months							.99	1.01	1.01	1.04	1.05	.99
Do_sports				.74			1.07	.99	.99	.74	.8	.98
TOPIC INTEREST AND MOTIVATION												
Engaged in voluntary work			.77		.78		.86	1.01	.88	.89	.88	1.06
(Potential) Political Participation	.85	.84				.82	.84	1.04	1.19	1.14	1.12	.96
Political Left orientation							.89	1.05	1.16	1.04	1.16	1.06
Political interest	.77		.82	.79	.75	.74	.82	1.07	1.01	1.04	.9	.9
Subjective questions noanswer							.97	1.15	1.21	1.21	1.08	1.33
Number of extreme categories		1.05	1.05	1.03			1.03	1.01	1.03	1.01	.96	1
Number of middle categories				1.07			1.01	1.01	.98	1.05	1.03	.97
Income noanswer	1.35						1.42	.84	.79	.79	.83	.8
Response rate within household	.14	.14	.23	.22	.30	.20	.14	.84	1.2	1.55	1.68	1.16
SURVEY STATUS OF INDIVIDUAL												
Person is Reference Person (only more-adult-HH)	.56	.66		.63	.77	.72	.6	1.17	1.52	1.22	1.32	1.26
N	7769	6333	5719	4874	4332	3592	N=32619, p=.20 (Intra-cluster coeff.)					
MCFadden Pseudo R ²	.068	.064	.057	.053	.045	.055	-					
LR chi ²	506	331	309	212	209	176	Wald chi ² (Deg. Freedom=180)					1485

odds ratios. "-" = n.a. Single wave models: only significant (1%) effects from forward regression model included.

Contact Address:

Oliver Lipps
Swiss Household Panel
4, Espace de l'Europe
P.O. Box 688
CH-2002 Neuchâtel
oliver.lipps@unine.ch

Techniken und Werkzeuge zur Unterstützung der Erinnerungsarbeit bei der computergestützten Erhebung retrospektiver Längsschnittdaten

Techniques and tools to support respondents' memory process in computer-aided retrospective life course interview

Britta Matthes, Maike Reimer und Ralf Künster

Zusammenfassung

Erhebungen von retrospektiven Lebensverlaufsangaben zeigen, dass Befragte mehr oder weniger große Schwierigkeiten beim Erinnern und Datieren von Ereignissen und Episoden aus ihrem Leben haben. In früheren Lebensverlaufsbefragungen des Max-Planck-Instituts für Bildungsforschung (German Life History Study) wurden die Angaben der Befragten deshalb nach Ende des Interviews einer sorgfältigen und arbeitsintensiven Edition unterzogen. Um den Editionsaufwand deutlich zu verringern und gleichzeitig die Qualität der Angaben der Befragten schon während der Erhebung zu verbessern, haben wir uns zunächst mit der Frage beschäftigt, wie autobiographisches Erinnern und Datieren aus gedächtnispsychologischer Sicht funktioniert. Die sich daraus ergebenden Konsequenzen haben wir bei der Neuentwicklung eines Instrumentes zur Erhebung retrospektiver Lebensverlaufsangaben in computergestützten, telefonischen Längsschnittbefragungen („TrueTales“) umgesetzt. Dieses Instrument enthält verschiedene Techniken und Werkzeuge, mit denen die Interviewer die Erinnerungsarbeit der Befragten wesentlich intensiver und effektiver unterstützen können als bisher möglich.

Abstract

Surveys of retrospective life history data show that respondents face various degrees of difficulty when recalling and dating autobiographical events and episodes. The data of previous life course interviews conducted by the Max-Planck-Institute for Human Development (German Life History Study) were therefore edited through time- and labor-intensive processes. In order to reduce these immense efforts during editing and to improve data quality while conducting the interviews, we investigated how autobiographical recalling and dating works from the cognitive psychological point of view. We then put our findings into practice developing an instrument called "TrueTales". This instrument allows the collection of retrospective life course data in computer-aided telephone interviews. Different techniques and tools are applied so that interviewers can support respondents' memory processes more intensively and effectively than previously possible.

1 Probleme bei der Erhebung von Lebensverlaufsdaten

1979 betrat die Forschergruppe um Karl Ulrich Mayer mit der standardisierten retrospektiven Erfassung von Lebensverläufen erhebungstechnisches Neuland. Seither wurden die Erhebungsmethoden von retrospektiven Längsschnittdaten in nahezu jeder Lebensverlaufsbefragung des Max-Planck-Instituts für Bildungsforschung überprüft und weiter entwickelt (Mayer/Brückner 1989; Brückner 1993; Brückner 1994; Brückner/Mayer 1995; Brückner 1995; Brückner/Mayer 1998; Solga 2001; Goedicke/Lichtwardt/Mayer 2004; Matthes/Lichtwardt/Mayer 2004; Hillmert et al. 2004). Das wichtigste Ziel dieser Bemühungen war es, mit möglichst geringem zeitlichem und sachlichem Aufwand intersubjektiv vergleichbare Lebensverlaufsdaten vollständig und konsistent zu erheben.

Retrospektivbefragungen sind jedoch vor allem mit dem Problem konfrontiert, dass den Befragten sowohl in Bezug auf das Erinnern als auch auf das Datieren von Ereignissen Fehler unterlaufen (Reimer 2001, 2005). Verlaufsanalysen sind jedoch auf eine möglichst lückenlose, konsistente und valide Rekonstruktion von Ereignisgeschichten angewiesen. Die erhobenen Lebensverlaufsdaten wurden deshalb immer einer sorgfältigen, zeit- und arbeitsintensiven Datenedition unterzogen, bei der die Lückenlosigkeit, Genauigkeit, Plausibilität und synchrone Konsistenz überprüft wurde. Die intensiv geschulten Editeure entschieden über jeden einzelnen Fall nach festgelegten Regeln und rekonstruierten bzw. korrigierten mangelhafte Angaben aufgrund von Informationen aus anderen Interviewteilen. In vielen Fällen mussten die Befragten aber erneut kontaktiert werden, weil anderweitig kein angemessener Umgang mit den identifizierten Inkonsistenzen gefunden werden konnte. Für die Erhebung der Lebensverläufe der westdeutschen Geburtsjahrgänge 1929-31, 1939-41 und 1949-51 (LV I) schätzen Mayer/Brückner (1989), dass die Edition noch einmal so viel Zeit und Ressourcen in Anspruch genommen hatte wie die eigentliche Erhebung. Auch im Teilprojekt „Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland (LV-West 64/71)“ wurde pro Fall im Schnitt noch einmal das 1,3-Fache der Interviewdauer für Edition und Nachrecherche aufgewendet. Dieser enorme Aufwand wird wegen der damit verbundenen hohen Datengenerierungskosten in Frage gestellt. Außerdem ist umstritten, ob eine über die Korrektur erhebungstechnischer Fehler hinausgehende Datenedition überhaupt sinnvoll ist (Hillmert 2002). Da bei der Datenedition scheinbar inkonsistente Angaben korrigiert werden, besteht prinzipiell die Möglichkeit, dass ungewöhnliche Lebensverläufe „geglättet“, Daten also fälschlicherweise korrigiert werden. Ediert man die Daten jedoch nicht, können die unterbleibenden Korrekturen zu Fehlern in den Ergebnissen führen, denn die identifizierten Inkonsistenzen könnten auch auf Erinnerungsfehlern oder Missverständnissen beruhen. Das Dilemma der Datenedition liegt darin, dass sich fast nie eindeutige Regeln ableiten lassen, welche der vorhandenen Informationen valide sind.

Ganz besonders problematisch wird es, wenn Lebensverläufe im Paneldesign erhoben werden, weil hier nicht nur die Konsistenz der Lebensverläufe innerhalb der aktuell erhobenen Panelwelle gewährleistet werden muss, sondern übergreifend auch die Konsistenz zu den in der vorhergehenden Panelwelle erhobenen Lebensverläufen. In vielen Panel-Studien wird versucht, die Ereig-

nisgeschichte zu vervollständigen, indem Informationen für einen bestimmten Zeitraum doppelt (jeweils in den aufeinander folgenden Panelwellen) erhoben werden (z. B. LV-Ost Panel, British Household Panel Survey (BHPS)). In der Praxis zeigt sich jedoch, dass die zeitlich überlappend erhobenen Informationen häufig nicht identisch sind. Es treten Abweichungen in den Werten einzelner Variablen auf (z. B. in der Bezeichnung von beruflichen Tätigkeiten), die Beginn- oder Enddaten können unterschiedlich sein, oder es werden für den gleichen Zeitraum gänzlich verschiedene Ereignisgeschichten berichtet. Damit ist die Konstruktion vollständiger und konsistenter Lebensverläufe nicht oder nur mit großem Editions Aufwand möglich. Eine andere Variante, mit diesem Episodenanschlussproblem in Panel-Befragungen umzugehen besteht darin, nach Veränderungen seit dem letzten Interview zu fragen (z. B. Socio-oeconomic Panel (SOEP)). Hier wird von der befragten Person erwartet, dass sie sich ohne Hilfestellung korrekt an die Informationen erinnert, die sie in der vorhergehenden Panelwelle gegeben hat. Gab es im Lebensverlauf mehrere Veränderungen seit dem letzten Interview, wird eine Zuordnung konkretisierender Angaben schwierig oder unmöglich.

Bei Vorbereitung auf die jüngste Lebensverlaufsbefragung des Max-Planck-Instituts für Bildungsforschung „Frühe Karrieren und Familiengründung: Lebensverläufe der Geburtskohorte 1971 in Ost- und Westdeutschland (LV-Panel 71)“, einer Wiederholungsbefragung der ost- und westdeutschen Geburtskohorte 1971 (vgl. Matthes/Mayer 2006), hatten wir uns das Ziel gesetzt, Strategien zu entwickeln, mit den Erinnerungsschwierigkeiten der Befragten auf neuartige Weise umzugehen, so dass dadurch die Datenqualität gesteigert und eine herkömmliche Dateneditionsphase überflüssig werden würde. Der vielversprechendste Weg erschien uns, retrospektive Erinnerungsfehler schon im Interview zu vermeiden und Inkonsistenzen in Zusammenarbeit zwischen Befragten und Interviewern bereits während der Datenerhebung aufzuklären, statt diese erst in einer nachträglichen Edition der Daten zu korrigieren. Daher beschäftigten wir uns zunächst mit der Frage, wie das autobiographische Erinnern aus gedächtnispsychologischer Sicht funktioniert. Darauf gehen wir im Folgenden zunächst kurz ein. Anschließend erläutern wir, welche Konsequenzen sich daraus für die Gestaltung eines Instruments zur Erhebung von retrospektiven Lebensverlaufsangaben ergeben und wie wir diese in dem Erhebungsinstrument „TrueTales“ umgesetzt haben. Abschließend wird dargestellt, ob das neue Erhebungsinstrument wirklich zu besseren Lebensverlaufsdaten führt.

2 Wie funktioniert autobiographisches Erinnern? Gedächtnispsychologische Befundlage

Befunde der Gedächtnispsychologie sprechen dafür, dass das Erinnern an autobiographische Ereignisse und Ereignisgeschichten anders funktioniert als deren Datierung (vgl. Reimer 2001). Deshalb wenden wir uns im Folgenden zunächst der Frage zu, wie autobiographische Ereignisse erinnert werden und skizzieren anschließend, auf welche Weise diese zeitlich verortet werden.

2.1 Wie werden autobiographische Ereignisse erinnert?

Erinnerungen an Ereignisse und Episoden sind im Langzeitspeicher des Gedächtnisses als ein Netzwerk miteinander verbundener Repräsentationen organisiert. Diese Repräsentationen können über drei verschiedene Strukturen der Erinnerungsorganisation miteinander verknüpft werden (vgl. Barsalou 1988; Belli 1998; Conway/Pleydell-Pearce 2000). Dementsprechend gibt es drei Erinnerungspfade mit deren Hilfe die im Gedächtnis abgespeicherten Informationen abgerufen werden können:

1. *Themengebundene Erinnerungspfade.* Bei der Erinnerung von Informationen, die innerhalb eines Themenbereiches abgespeichert sind, werden die Strukturen der Erinnerungsorganisation hierarchisch vom Allgemeinen zum Speziellen miteinander verbunden (z. B. „Meine Zeit im Unternehmen X“ – „Der erste Arbeitstag“ – „Mittagessen mit den neuen Kollegen“).
2. *Themenübergreifende Erinnerungspfade.* Handelt es sich um themenübergreifende Erinnerungen, werden diese über parallel liegende Strukturen der Erinnerung auf den gleichen oder auf verschiedenen Abstraktionsebenen organisiert. Die Erinnerung des arbeitsbezogenen Ereignisses „Weihnachtsfeier im Büro“ kann z. B. mit dem partnerschaftlichen Ereignis „anschließendes Treffen mit dem Partner in einer Bar“ über einen parallelen Ablauf erfolgen.
3. *Sequenziell organisierte Erinnerungspfade.* Sequenziell strukturierte Erinnerungsorganisationen verlaufen entlang der kausal-temporären Abfolge von Ereignissen, z. B. „Arbeitsplatz im Unternehmen X antreten“ – „Verlust des Arbeitsplatzes wegen Verkleinerung des Unternehmens“ – „arbeitslos sein“.

In der Gedächtnispsychologie konnte belegt werden, dass Ereignisse oder Ereignissequenzen, die lange andauern, emotional bedeutsam oder folgenschwer sind, eine höhere Wahrscheinlichkeit haben, korrekt abgerufen zu werden als solche, die kurz andauern, emotional neutral sind und keine merklichen Folgen für den Erinnernden haben. So sind Ereignisse, die eine zentrale Bedeutung für die Lebensgeschichte haben, mit anderen Ereignissen im Gedächtnis eng verbunden und werden deshalb mit hoher Genauigkeit über lange Zeit erinnert. Bei kurzen und eher belanglosen Ereignissen besteht hingegen die Gefahr, dass sie im Interview nicht angesprochen werden (Barclay 1986, Bluck/Habermas 2000).¹ Für retrospektiv berichtete Ereignisabfolgen bedeutet dies, dass die Befragten eher weniger Veränderung und mehr Kontinuität angeben und Ereignisse wie z. B. Umzüge innerhalb einer Stadt, kleinere Veränderungen des Arbeitsplatzes innerhalb eines Unternehmens und kurze Phasen der Arbeitslosigkeit nicht berichten (Lampinen et al. 2000). Generell durchsuchen Menschen auf einen Reiz wie z. B. Antwortenweisungen, Fragen oder Nachfragen in einem Interview das gesamte Netzwerk der Repräsentationen entlang der hierarchischen, parallelen und sequenziellen Erinnerungsstrukturen nach Informationen, die in ausreichendem Maße zu dem empfangenen Reiz passen (Conway 1996; Shum/Rips 1999; Sudman/Bradburn/Schwarz 1996).

1 Nach wie vor ist in der Gedächtnispsychologie umstritten, ob einzigartige, seltene und untypische Ereignisse eher vergessen werden, weil sie schema-inkonsistent sind (Bluck/Habermas 2000), oder doch besser erinnert werden können, weil sie stärker Aufmerksamkeit auf sich zu ziehen vermögen (Brewer 1988).

2.2 Wie werden autobiographische Ereignisse datiert?

Die Befunde der Gedächtnispsychologie sprechen dafür, dass exakte Datumsangaben und Zeiträume nicht direkt in der Erinnerung repräsentiert sind, sondern durch rekonstruktive Inferenzprozesse ermittelt werden. Typischerweise bedienen sich Menschen zweier unterschiedlicher Inferenzprozesse, um sich daran zu erinnern, wann ein Ereignis stattfand bzw. wie lange es andauerte.

Rekonstruktionen über den Kontext. Wenn die gewünschte Information im Langzeitspeicher mit anderen Informationen in hohem Maße verknüpft ist und/oder schon wiederholt abgerufen wurde, entspricht die Erinnerung einem schnellen und unkomplizierten „Wiederfinden“. Dezentralere Informationen rekonstruieren die Befragten jedoch aus ihrem biografischen Kontext, indem sie sich weitere interne Anhaltspunkte bilden, die sie durch das Netzwerk ihrer Repräsentationen leiten („Habe ich danach noch eine Ausbildung gemacht? Hm, nachdem mein Sohn zur Schule kam – ja, da waren diese Wochenendseminare in Buchhaltung, die ich besucht habe – etwa ein halbes Jahr lang.“). Nur sehr wenige bedeutende biografische Punkte, die häufig erinnert werden, so genannte „landmark events“ (z. B. Hochzeiten, Geburts- oder Feiertage), sind zusammen mit Kalenderdaten im Gedächtnis abgespeichert (vgl. Friedman 1993; Larsen/Thompson/Hansen 1996; Shum 1998). Die Rekonstruktion des individuellen biografischen Kontexts wird daher vor allem bei der Erinnerung untergeordneter Ereignisse und Kalenderdaten wichtig, die schwieriger zu erinnern sind. („Ich begann im Unternehmen X zu arbeiten, nachdem ich im Mai 1990 geheiratet hatte.“) Menschen benutzen die hierarchischen, parallelen und sequenziellen Abrufpfade der Erinnerungsorganisation, um Repräsentationen zu verbinden und benutzen dabei alle Arten zeitlicher Verknüpfung zwischen Ereignissen, wie z. B. „früher als“, „danach“ und „seither vergangene Zeit“ (Huttenlocher/Hedges/Prohaska 1988). Die erinnerte Information wird häufig mit weiteren Erinnerungen abgeglichen, um sicher zu gehen, dass sie in Übereinstimmung mit anderen und bedeutenden Ereignissen erinnert wurde (Huttenlocher/Hedges/Bradburn 1990; Larsen/Thompson 1995; Loftus/Marburger 1983).

Individuelle Erinnerungsstrategien. Die Art und Weise, wie eine Person ihr Gedächtnis organisiert und welcher Erinnerungsstrategien sie sich bedient, ist zum einen sozialisatorisch geprägt und zum anderen in hohem Maße individuell bedingt. In unserer Gesellschaft, in der Lebensverläufe stark durch institutionelle Regulationen beeinflusst werden, existieren eine Vielzahl von Ereignissen, die für die Mehrheit der Befragten von Bedeutung sind, wie z. B. Übergänge im Schul- und Ausbildungsverlauf. Auf individueller Ebene kann der Kontext, in dem eine Person ein Ereignis erinnert, jedoch stark variieren. Während ein Befragter einen Umzug im Kontext seines Zusammenzugs mit dem Partner erinnert, hat ein anderer einen ähnlichen Umzug im Kontext eines Tätigkeitswechsels in eine andere Stadt berichtet. Eine Person rekonstruiert den Zeitpunkt eines Tätigkeitswechsels über die Geburt ihres Kindes, während eine andere dies aus dem Wissen ableitet, dass seither drei Jahre vergangen sind.

3 Das Erhebungsinstrument „TrueTales“

Die Befunde der Gedächtnispsychologie machen deutlich, dass man in Interviews üblicherweise sowohl den verschiedenen Erinnerungsstrategien als auch den rekonstruktiven Inferenzprozessen der Befragten nur unzureichend entgegen kommt. In der Regel werden kontextualisierte Rekonstruktionsprozesse weder erlaubt noch unterstützt. Vielmehr wird davon ausgegangen, dass alle Befragten sämtliche Informationen ohne oder mit nur wenig systematischer Hilfe im Gedächtnis abrufen können. Das macht insbesondere solche Lebensverläufe anfällig für Erinnerungsfehler, die viele Übergänge zwischen untypischen und eher kurzen Episoden aufweisen oder die nicht gut in die Logik einer Lebensgeschichte eingebunden sind. So erinnern Frauen in der Regel Familienereignisse und Umzüge besser als Männer (Auriat 1993; Poulain/Riandey/Firdion 1991; Thompson et al. 1998), sind aber häufig schlechter im Erinnern von Arbeitsmarktereignissen (Elias 1991), weil Werdegänge von Frauen stärker von Bewegungen zwischen und Parallelitäten von (geringfügiger oder Teilzeit-) Erwerbstätigkeit und Haushaltsführung bzw. Kindererziehung gekennzeichnet sind als die von Männern. Lebensverläufe mit vielen untypischen Ereignissen sind offenbar schwieriger zu erinnern und daher für eine Erfassung fehleranfälliger. Werden für alle Befragten identische Anweisungen, Fragen und Nachfragen verwendet, fällt es den Befragten häufig schwer, einen Bezug zu ihrer individuellen biografischen Konstellation herzustellen. Die starke Standardisierung verhindert es, auf individuelle Erinnerungsstrategien der Befragten während des Interviews auch individuell einzugehen. Die standardisierten Hinweisreize sind darüber hinaus unterschiedlich effektiv, da die gleiche Information für unterschiedliche Personen auf verschiedene Weise abgebildet und verknüpft ist.

Erinnerungsfehler können aber durch effektive, die Erinnerungsarbeit der Befragten unterstützende Interviewtechniken und -werkzeuge vermieden werden. Aufbauend auf den Ergebnissen der Gedächtnispsychologie und unter Berücksichtigung der langjährigen Erfahrungen bei der Erhebung und Edition von Lebensverlaufsdaten am Max-Planck-Institut für Bildungsforschung, Berlin, sind wir zu der Einsicht gelangt,

- dass nur eine konsequente Modularisierung, in der innerhalb der jeweiligen Themenbereiche chronologisch vorgegangen wird, zur Erhebung vollständiger Ereignisgeschichten führt;
- dass es bei modular erhobenen Ereignisgeschichten notwendig ist, nach der Datenerhebung – aber noch während des Interviews – eine Datenrevision vorzunehmen, in der zeitliche Inkonsistenzen visualisiert und – falls nötig – korrigiert werden können;
- dass durch Einspielen relevanter autobiographischer Abruf-Hinweisreize in die standardisierten Fragen Ereigniskontexte gezielt aktiviert werden können;
- dass die streng standardisierte Erhebungsweise durch eine flexible Unterstützung der Erinnerungsfähigkeit der Befragten ergänzt werden muss;
- dass diese Ergänzung höhere Anforderungen an die Interviewer stellt und somit auch Konsequenzen für ihre Schulung und Feldbetreuung hat.

Diese Erkenntnisse werden im Folgenden für das Erhebungsinstrument „TrueTales“ konkretisiert, das bei der jüngsten Lebensverlaufsbefragung des Max-Planck-Instituts für Bildungsforschung

„Frühe Karrieren und Familiengründung: Lebensverläufe der Geburtskohorte 1971 in Ost- und Westdeutschland (LV-Panel 71)“ zum Einsatz gekommen ist.

3.1 Konsequente Modularisierung und sequentielles Erinnern

Zur Erhebung von Lebensverlaufsdaten können zwei prinzipiell unterschiedliche Strategien eingesetzt werden, die jeweils auf spezifischen Annahmen beruhen, wie Befragte ihre Lebensverläufe möglichst vollständig, zuverlässig und konsistent erinnern und berichten: die sequenzielle und die modulare Vorgehensweise.

Die sequenzielle Erhebung. Die sequenzielle Befragungsstrategie geht davon aus, dass vollständige Lebensgeschichten am besten in der chronologischen Abfolge der Ereignisse und Episoden erinnert werden, in der sie stattgefunden haben. Bei dieser Verfahrensweise beginnt die Erhebung an einem festgelegten Startzeitpunkt (z. B. Geburt des Befragten, Enddatum der letzten Panel-Welle, historisches Ereignis). Zunächst wird der Status der Person zu diesem Zeitpunkt erfasst, z. B. „Was haben Sie im Januar 2007 gemacht?“. Damit für alle Episoden des gleichen Status sämtliche relevanten Details identisch erhoben werden, gibt es Wiederholungsschleifen, in denen die Befragten für jede berichtete Episode des gleichen Statustyps die gleichen Fragen gestellt bekommen. Am Ende der Wiederholungsschleife wird gefragt, ob sich der Status verändert hat, und wenn ja, wann. Hat sich der Status nicht verändert, endet die Wiederholungsschleife. Gab es eine Statusveränderung wird gefragt, in welchen Status der Befragte übergegangen ist, z. B. „Was haben Sie direkt im Anschluss daran gemacht?“. Dies wiederholt sich so lange, bis alle erinnerten Statusveränderungen chronologisch bis zum aktuellen Interviewzeitpunkt als Episoden erfasst sind.

Der Vorteil dieser Methode liegt in der lückenlosen Erhebung der Ereignisgeschichte. Ein wesentlicher Nachteil ist jedoch, dass die chronologische Erhebung von Längsschnittdaten das Vergessen oder Fehlerinnern paralleler und sich überschneidender Ereignisse und Episoden (z. B. zwei parallel ausgeübte Erwerbstätigkeiten) begünstigt, und die Tendenz der Befragten fördert, ihre Lebensverläufe durch die Ausblendung kurzer, für sie unwesentlich erscheinender Episoden zu „glätten“.

Die modulare Erhebung. Das modulare Verfahren zur Erhebung von Längsschnittdaten versucht die Nachteile der sequenziellen Methode zu vermeiden, indem das ganze Leben mehrfach in den verschiedenen Lebensbereichen durchlaufen wird. Äquivalent zur Festlegung relevanter Statuspositionen bei der sequenziellen Erhebung werden hierbei die relevanten Lebensbereiche in so genannten Modulen organisiert. Auch hier gibt es Wiederholungsschleifen, in denen sämtliche relevante Details für alle Episoden eines Lebensbereichs identisch erhoben werden. Das Besondere an der modularen Erhebung ist, dass die Befragten innerhalb eines Moduls beliebig viele Episoden aus einem Lebensbereich berichten können, selbst wenn die Episoden zeitlich parallel verlaufen, z. B. ein Fernstudium, das parallel zu einer betrieblichen Ausbildung absolviert wird. Innerhalb eines Moduls wird die Erinnerung über die themengebundenen Abrufpfade unterstützt, indem am

Ende jeder Wiederholungsschleife nach weiteren Episoden in diesem Lebensbereichs gefragt wird (z. B. „Haben Sie außer dieser Ausbildung noch eine weitere Ausbildung gemacht?“). Durch diese themenbereichsspezifische Rekonstruktion der Ereignisgeschichte wird das Erinnern kurzer bzw. paralleler Episoden unterstützt.

Insbesondere die *Kombination der sequentiellen und modularen Erhebungsweise*, d. h. die Unterteilung der Lebensgeschichtlichen Ereignisabfolge in mehrere thematisch getrennte Lebensbereiche (Modularisierung) und innerhalb dieser eine chronologische Abfrage (Sequenzialisierung), kommt den Eigenschaften der autobiographischen Gedächtnisdatenbank sowohl beim Erinnern als auch beim Datieren entgegen. Dabei konkurrieren zeitliche und inhaltliche Strukturen nicht. Vielmehr liegt einer erfolgreichen Gedächtnissuche und einer bestmöglichen Datierung ein flexibler Einsatz von inhaltlichen und zeitlichen Verankerungen zugrunde.

Die Kombination der sequentiellen und modularen Erhebungsweise bei der retrospektiven Erfassung von Lebensverläufen ist nicht neu. Schon in der ersten Lebensverlaufserhebung des Max-Planck-Instituts für Bildungsforschung wurden die Lebensverlaufs-Episoden in Modulen und innerhalb der Module in chronologischer Reihenfolge erhoben. In diesen Erhebungen wurde jedoch der Lebensverlauf nach der ersten Erwerbstätigkeit als quasi kontinuierlicher Erwerbsverlauf, der maximal durch Arbeitslosigkeits- oder Nichterwerbsphasen unterbrochen werden konnte, verstanden. Dem steht der Befund zunehmender Komplexität der Bildungs- und Erwerbsgeschichten gegenüber (z. B. Parallelität von Bildung und Beschäftigung oder Beschäftigung im Kontext von Nichterwerbsphasen (z. B. Elternzeit)). Im Ergebnis dieser Erhebungsweise werden einerseits kurze oder parallele Bildungs- oder Erwerbsaktivitäten und kurze Arbeitslosigkeitsphasen nicht berichtet. Der Bildungs- und Erwerbsverlauf wird geglättet, weil die Erhebung sich an einem „Hauptaktivitätsstrang“ durch das Leben der Befragten hangelt. Schul- oder Ausbildungsphasen, die zwischen Erwerbsepisoden stattgefunden haben, mussten doppelt berichtet werden: sowohl als Episoden im Schul- bzw. Ausbildungsmodul als auch als Lückenaktivitäten zwischen Erwerbsepisoden im Erwerbsmodul. Solche Doppelterfassungen von Episoden im Interview sind dem Befragten nicht nur schwierig zu vermitteln, sie führen nach dem Interview häufig auch zu Zuordnungsschwierigkeiten der identischen, aber in unterschiedlichen Modulen erfassten Episoden. Dem kann nur durch eine *konsequentere Modularisierung* entgegengewirkt werden, indem alle interessierenden Teilbereiche des Bildungs- und Erwerbsverlaufs, also auch Arbeitslosigkeit oder Wehrdienst bzw. Erziehungszeiten in eigenen Längsschnittmodulen erfasst werden, so dass die Befragten in jedem dieser Module erneut chronologisch durch ihr Leben gehen müssen.

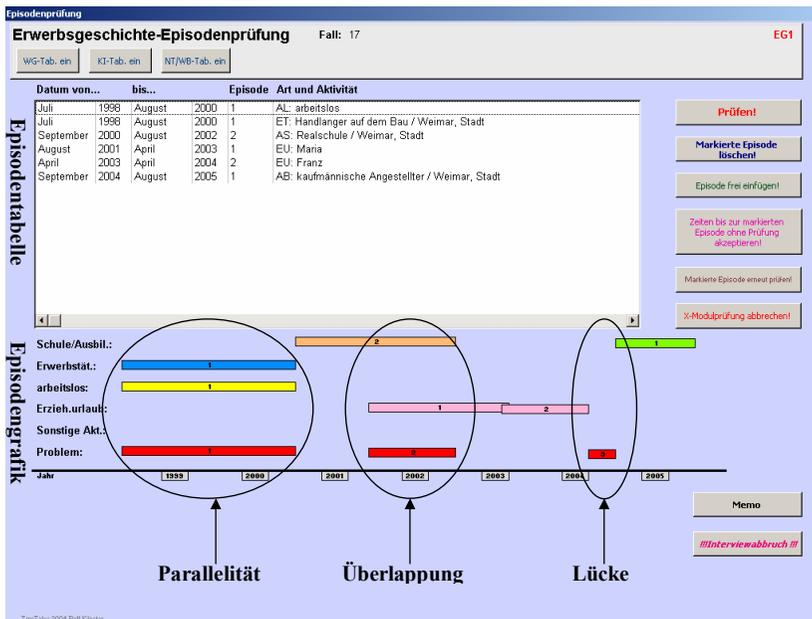
3.2 Datenrevision während des Interviews

Das zentrale gedächtnispsychologische Problem, das bei dieser Form der Kombination von Modularisierung und Sequenzialisierung entsteht, ist die zeitliche Verankerung einzelner Episoden in die chronologische Reihenfolge des gesamten Lebensverlaufs. Deshalb fällt den Befragten bei Anwendung dieses Vorgehens insbesondere das Datieren von Episoden schwer. Lücken und Inkonsistenzen zwischen den

verschiedenen Lebensbereichen können von den Interviewern nur sehr schwer erkannt und dementsprechend nachgefragt werden. Um dieses Problem zu lösen, haben wir das Interview in zwei Teile unterteilt: a) die Datenerhebung, in der die Fragen gestellt und die Angaben eingegeben werden, und b) die Datenrevision, in der zeitliche Inkonsistenzen visualisiert und korrigiert werden können.

In einem eigenen Modul zur Datenrevision (Prüf- und Ergänzungsmodul) wird die zeitliche Konsistenz der Datumsangaben aus allen Modulen, in Zusammenarbeit mit den Befragten, auf ihren Gesamtzusammenhang überprüft und nötigenfalls korrigiert. Dieses Prüf- und Ergänzungsmodul ist in Anlehnung an den biografischen Kalender aufgebaut (u. a. Belli 1998; Belli/Shay/Stafford 2001; Caspi et al. 1996; Vaart 2002; Zouwen/Dijkstra/Vaart 1993; Axinn/Pearce/Ghimire 1999; Yoshihama et al. 2002; Belli et al. 2004). Dabei sind die einzelnen Episoden in den verschiedenen Lebensbereichen entlang der chronologischen Zeitachse sowohl tabellarisch als auch grafisch dargestellt (vgl. Abbildung 1).

Abbildung 1 Prüf- und Ergänzungsmodul



Da in telefonischen Interviews biographische Kalender nur für die Interviewer sichtbar sind, können den Befragten die sich (aus der Sicht des Wissenschaftlers) ergebenden Inkonsistenzen (wie Parallelitäten, Überschneidungen, Lücken) nur durch Nachfragen und Anbieten spezifischer Abruf-Hinweisreize geklärt werden. Dem Interviewer müssen die Inkonsistenzen also auf den ersten Blick deutlich werden. Deshalb werden diese in der Grafik zusätzlich mit einem eigenen Balken, dem so genannten Problembalken, hervorgehoben.

Die Lebensverlaufsdaten werden im Prüf- und Ergänzungsmodul einer rekursiven Konsistenzprüfung unterzogen, d. h. nach jeder Veränderung einer Episode (Änderung des Anfangs- bzw. Enddatums, Löschen oder Einfügen einer Episode) wird erneut getestet, ob weiterhin Inkonsistenzen bestehen. Dieser rekursive Prozess erfolgt solange, bis alle Inkonsistenzen beseitigt oder geklärt sind. Dadurch wird gewährleistet, dass die korrigierten bzw. nachträglich erhobenen Episoden wirklich zu vollständigen Lebensverläufen führen.

Um Lücken schließen zu können bzw. nachträglich berichtete Episoden vollständig, d. h. mit allen Detailinformationen erfassen zu können, bietet „TrueTales“ die Möglichkeit, in die entsprechenden Module zurück zu springen bzw. ein Lückenaktivitätsmodul aufzurufen. Hierzu entscheidet der Interviewer anhand der Angaben des Befragten, welches Modul zur Erfassung der fehlenden Episode aufgerufen werden muss. Dieses Modul wird dann mit allen modulspezifischen Fragen bis zur vollständigen Erhebung der Episode durchlaufen, um danach zur aktualisierten Datenrevision zurückzukehren. Darüber hinaus haben wir, um Auswertungsprobleme durch Datierungsfehler zu verringern, Algorithmen entwickelt, die auch Episoden mit teilweise oder vollständig fehlenden Datumsangaben in die zeitliche Prüfung mit einzubeziehen. Außerdem haben wir Strategien zum Umgang mit solchen Episoden konzipiert, die nicht eindeutig zeitlich verortet werden können.

Das Prüf- und Ergänzungsmodul stellt nicht nur eine Form der Sofortedition dar, die gegenüber herkömmlichen Editionsverfahren Kosten und Zeit spart, sondern erleichtert auch die Erinnerungsarbeit, indem die Befragten aktiv unterstützt werden, parallele und sequenzielle Rekonstruktionspfade für das Abrufen von Informationen zu nutzen.

Weil die in den verschiedenen Modulen erhobenen Episoden erst in einem eigenen Schritt nach ihrer systematischen, modularen Erhebung zu einem Lebensverlauf zusammengesetzt werden und die dabei entstehenden Lücken und Inkonsistenzen zusammen mit dem Befragten aufgeklärt werden, ergeben sich intersubjektiv besser vergleichbare, vollständigere und konsistentere Lebensverlaufsdaten als beim Einsatz biographischer Kalender, bei denen es den Befragten überlassen ist, in welcher inhaltlichen und zeitlichen Reihenfolge die Ereignisse berichtet werden.

3.3 Kontextualisierung der Erinnerungsarbeit

Ein Instrument kann jedoch nicht nur durch seinen grundlegenden Aufbau erinnerungsunterstützend wirken, sondern vor allem auch durch die Formulierung der Fragen bzw. durch Bereitstellung von Hilfsmitteln, die es erlauben, den individuellen Kontext der befragten Person sinnvoll in das Interview einzubinden. Da die Frage in einem telefonischen Interview der zentrale Abruf-Hinweisreiz ist, müssen – um die Befragten bei ihrer komplexen Erinnerungsarbeit bestmöglich zu unterstützen und Missverständnissen zwischen Interviewer und Befragtem entgegenzuwirken – die zu erinnernden Referenzmengen und -ereignisse genau definiert sein. Können die Interviewer Angaben zu Ereignissen, Zeiten und Personen, die im Verlauf des Interviews bereits erfasst wurden, als Kontexte für die zu erinnernden Informationen anbieten, gelingt es den Befragten leichter, ihre Lebensgeschichte zu strukturieren und eigenständig Abruf-Hinweisreize zu

generieren. Dabei hat sich in verschiedenen Studien gezeigt, dass Abruf-Hinweisreize, die sich individueller Referenzen bedienen, also auf Ereignisse und Zeiträume aus dem Leben der befragten Person beziehen (Geburt eines Kindes, Abschluss einer Ausbildung), die Erinnerungsleistung besser unterstützen als öffentliche Referenzen, wie z. B. ein politisches Ereignis (Fall der Mauer, Bundestagswahl) (Larsen/Thompson 1995; Shum 1998). Im Anschluss an diese Überlegungen haben wir viele Fragen kontextualisiert, indem wir während des Interviews auf bereits vorhandene autobiografische Angaben des Befragten zurückgegriffen und diese in Form von Einblendetexten in die standardisierten Fragen eingespielt haben. Diese autobiografischen Angaben stammen entweder aus den aktuell erfassten Angaben des laufenden Interviews (interviewintern) oder aus den Daten der vorhergehenden Erhebungswelle (panelwellenübergreifend).

Interviewintern: Kontextualisierte Abruf-Hinweisreize werden in „TrueTales“ u. a. bei Inkonsistenzen im Prüf- und Ergänzungsmodul eingesetzt, um eindeutig vermitteln zu können, auf welche Ereignisse und Zeiten sich diese beziehen. Wurde im Prüf- und Ergänzungsmodul der „Prüfen“-Button angeklickt, wird menügesteuert in chronologischer Reihenfolge eine bestehende Inkonsistenz dem Befragten mittels einer standardisierten Frageformulierung eingeblendet. Der Fragetext richtet sich nach der Art der Inkonsistenz (Lücke, zeitliche Überschneidung oder unklarer Episodenanschluss wegen fehlender Zeitangaben) und bezieht wesentliche Informationen der dabei beteiligten Episoden ein, so dass der Interviewer dem Befragten eindeutig rückmelden kann, welcher Episodenübergang gemeint ist und um welche Problemlage es sich handelt (vgl. Abbildung 2).

Abbildung 2 Interviewinterne Kontextualisierung im Prüf- und Ergänzungsmodul

Episodenprüfung
Erwerbsgeschichte-Episodenprüfung Fall: 17 EG1

Wig-Tab. ein KI-Tab. ein NT/WB-Tab. ein

Datum von...	his...	Episode	Art und Aktivität
Juli 1998	August 1999	1	AL: arbeitslos
August 1999	April 2000	1	ET: Handlanger auf dem Bau / Weimar, Stadt
September 2000	August 2002	1	AS: Realschule / -8
August 2001	April 2003	1	EU: Maria
April 2003	April 2004	2	EU: Franz
September 2004	März 2005	1	AB: kaufmännische Angestellter / Weimar, Stadt

Schule/Ausbil.: [Timeline visualization]

Erwerbstat.: [Timeline visualization]

arbeitslos.: [Timeline visualization]

Erzieh.urlaub.: [Timeline visualization]

Sonstige Akt.: [Timeline visualization]

Problem.: [Timeline visualization]

Jahr [1999] [2000] [2001] [2002] [2003] [2004] [2005]

So wie ich das bisher notiert habe, ist hier eine Lücke. Ich habe hier notiert, dass Sie die Erwerbstätigkeit <<Handlanger auf dem Bau / Weimar, Stadt>> bis April 2000 gemacht haben und ab September 2000 in der Schule <<Realschule / -8>> waren). Da ist eine Lücke. Habe ich das richtig notiert und Sie haben dazwischen etwas anderes gemacht (z.B. einen Schulabschluss nachgeholt, eine Aus- oder Weiterbildung absolviert, eine Erwerbs- oder Nebenerwerbstätigkeit ausgeübt, sich um den Haushalt gekümmert)?

Buttons: [Prüfen!], [Markierte Episode löschen], [Episode frei einfügen!], [Zurück bis zur markierten Episode ohne Prüfung akzeptieren!], [Markierte Episode erneut prüfen!], [X-Modulprüfung abbrechen!], [Meno], [Interviewabbruch #], [JA!], [NEIN!], [Keine Angabe!]

TrueTales 2004 Ralf Kauter

Abhängig von der Antwort des Befragten und von der Art des Problems werden dann Episoden vollständig nacherfasst, Korrekturen der Zeitangaben vorgenommen, Überschneidungen als korrekt akzeptiert, die zeitliche Lagerung der Episoden spezifiziert oder lediglich registriert, dass sich der Befragte nicht genau erinnern kann oder will.

Um unmissverständlich anzuzeigen, auf was oder wen sich eine Frage bezieht, greifen wir während der Datenerhebung aber auch auf von dem Befragten bereits genannte Wohnorte, Ausbildungen, Erwerbstätigkeiten und die dazugehörigen Zeiten sowie auf die Namen von Partnern und Kindern² zurück.

Panelwellenübergreifend: Panelerhebungen zeigen immer wieder, dass es für die Befragten sehr schwierig ist, sich ohne weiteres exakt an ihre Aktivitäten zum Zeitpunkt des letzten Interviews zu erinnern. Um einen Anschluss zweier Panelwellen zu gewährleisten, greifen wir in „TrueTales“ direkt auf Informationen aus dem letzten Interview zurück und bieten diese den Befragten als Gedächtnisanker an. Durch das Einspielen von Referenzereignissen aus der Zeit des letzten Interviews gelingt es den Befragten leichter, inhaltlich und zeitlich einen lückenlosen Anschluss an die vorherige Befragung zu finden. Die Informationen aus der vorherigen Befragung haben den Vorteil, dass sie frei von Erinnerungsfehlern sind, da sie sich auf Angaben beziehen, die für die Befragten zum Interviewzeitpunkt der vorherigen Befragung aktuell waren. Wir nennen diese – zum Zeitpunkt des letzten Interviews andauernden – Episoden „Aufsetzepisoden“ und nutzen die Zeitangaben und Detailvariablen dieser Aufsetzepisode als Grundlage und Startinformation für die Fortschreibung der ersten Episode des jeweiligen Moduls in der aktuellen Befragung. Damit die Interviewer sofort die eingespielten Informationen erkennen, haben wir diese mit speziellen Zeichen markiert. So fragen wir für die Fortschreibung der Erwerbsgeschichte: „In unserem letzten Interview haben wir notiert, dass Sie im <<September 1998>> als <<Maurer>> gearbeitet haben. Bis wann haben Sie das gemacht?“. Befragte, die zum Zeitpunkt des letzten Interviews nicht erwerbstätig waren, werden zu Beginn des Erwerbsmoduls an die letzte Erwerbstätigkeit vor diesem Aufsetzzeitpunkt erinnert und dann nach weiteren Erwerbstätigkeiten gefragt, z. B. „Die letzte Erwerbstätigkeit, die wir notiert haben, war <<Maurer>>. Waren Sie seit <<Datum letztes Interview>> noch einmal erwerbstätig?“. Da wir bei Befragten, die im letzten Interview angaben, noch nie erwerbstätig gewesen zu sein, keine Gedächtnisanker setzen können, fragen wir lediglich: „Waren Sie seit <<Datum letztes Interview>> erwerbstätig?“. Erst nachdem der Befragte angibt, wann die Aufsetzepisode endete, kann eine neue Episode angelegt, oder falls keine weiteren Episoden dieses Typs vorhanden sind, können Episoden aus dem nächsten Lebensbereich berichtet werden.³

2 Weil die Angaben für jedes Kind sehr früh im Interview erhoben werden, ihre Betreuungsgeschichte aber erst später, ist es sehr wichtig, die Namen der Kinder zu erfragen, um mit der Namensnennung bei den Befragten den eindeutigen Bezug zum jeweiligen Kind herstellen zu können.

3 Wenn man auf die originalen Angaben der Befragten aus dem letzten Interview zugreift, kommt es fast nie vor, dass Befragte der Aufsetzepisode widersprechen. Sollten allerdings edierte Aufsetzepisoden fortgeschrieben werden, muss man damit häufiger rechnen. Bestreitet ein Befragter, dass die Aufsetzepisode zum Zeitpunkt des letzten Interviews tatsächlich angedauert hat, wird pragmatisch davon ausgegangen, dass die Aufsetzepisode zum Zeitpunkt des letzten Interviews endete. Begründet werden kann dieses pragmatische Vorgehen damit, dass Angaben über die aktuelle Situation weniger mit Erinnerungsfehlern behaftet sind als retrospektive Angaben.

3.4 Flexible Unterstützung der Erinnerungsarbeit

Ein computergestütztes retrospektives Lebensverlaufsinterview muss grundsätzlich einer großen Bandbreite von individuellen Lebenslaufbesonderheiten gerecht werden, so dass die Interviewer in der Regel streng standardisiert verfahren können, indem das Interview entlang des filterbasierten Aufbaus des Erhebungsinstruments geführt wird. Weil standardisierte Abruf-Hinweisreize im Lebensverlaufsinterview jedoch häufig allein nicht ausreichen, um die Erinnerung an Ereignisse und deren zeitliche Einordnung anzuregen, sollten die Interviewer jedoch von der streng standardisierten Befragungstechnik abweichen können. Wir haben uns sowohl während der Datenerhebung als auch bei der Datenrevision bemüht, die Standardisierung so weit wie möglich aufrecht zu erhalten. Aber gerade in Lebensverlaufsinterviews gibt es häufig Situationen, in denen der Interviewer flexibel auf die Äußerungen der Befragten eingehen muss, z. B. wenn der Befragte selbst Klärungsbedarf äußert („Ich habe da noch ein Praktikum gemacht, wollen Sie das auch wissen?“) oder um Unterstützung bei der Rekonstruktion von Datumsangaben bittet („Ich hab da noch eine Ausbildung angefangen, weiß aber nicht mehr wann das war.“), wenn dem Befragten auffällt, dass er eine falsche Angabe gemacht hat oder wenn der Befragte eine Episode eines bereits abgeschlossenen Moduls „nachliefert“. Aber auch wenn dem Interviewer Inkonsistenzen auffallen, z. B. wenn ein Befragter ein Referendariat als Erwerbsepisode interpretiert und nicht, wie im Sinne der Studienziele, als Ausbildung; oder wenn ein Befragter mehrere parallele Vollzeit-Erwerbsepisoden berichtet, sollte der Interviewer von den standardisierten Frageformulierungen so weit wie nötig abweichen dürfen.

Durch unterstützende Äußerungen und Nachfragen des Interviewers (Sondieren) sollte sowohl das Erinnern als auch das Datieren erleichtert werden. Dementsprechend unterscheiden wir inhaltliches und zeitliches Sondieren: Inhaltliches Sondieren soll helfen zu entscheiden, welche Episoden in welchem Modul aufzunehmen sind und welche Kriterien erfüllt sein müssen, um eine Lebensphase als eigenständige Episode zu erfassen, aber auch welche Antwortoption aus einer Liste für die Beschreibung eines Tatbestandes am ehesten zutrifft. Zeitliches Sondieren soll dagegen dem Befragten helfen, sich korrekt an Start- und Enddatum sowie an die Reihenfolge von Episoden zu erinnern.

Damit die Interviewer sondieren können, brauchen sie erstens Informationen, die sie den Befragten bei Erinnerungsproblemen als Abruf-Hinweisreize zur Verfügung stellen können. Der Interviewer kann hierzu dem Befragten mitteilen, welche Ereignisse und welche dazugehörigen Start- und Enddaten aus dem gleichen Lebensbereich stammen (themengebundener, sequenzieller Abruf). Reicht diese Maßnahme nicht aus, muss auf bereits erhobene Angaben aus anderen Lebensbereichen (themenübergreifender Abruf) zurückgegriffen werden. So kann es z. B. hilfreich sein, bei der Erfassung von Erziehungsurlaubsphasen die Geburtsdaten der Kinder als Rekonstruktionshilfe zur Verfügung zu stellen. Deshalb haben wir in „TrueTales“ ein Werkzeug implementiert, mit dessen Hilfe die Befragten ihre Angaben zu einer bestimmten Episode mit anderen Ereignissen – auch aus anderen Lebensbereichen – absichern können. Durch Anklicken kann jederzeit während des Interviews für jeden zentralen Lebensbereich eine Tabelle eingeblendet werden, in der ausgewählte Angaben zu Ereignissen, Zeiten und Personen, die im Interview bereits erfasst wurden, als Kontexte für weitere zu erinnernde Informationen aufgelistet sind. Bei Bedarf kann der Interviewer somit die

Abruf-Hinweisreize den individuellen Bedürfnissen und Präferenzen des Befragten anpassen. Indem er dem Befragten bereits erinnerte Ereignisse und Datumsangaben ins Gedächtnis ruft, kann er wahlweise die sequenziellen, themengebundenen und auch themenübergreifenden Abrufpfade stimulieren.

Zweitens brauchen die Interviewer Werkzeuge, mit denen sie angemessen auf die Sondierungsergebnisse reagieren können. In „TrueTales“ haben wir den Interviewern eine Reihe von Werkzeugen zur flexiblen Behebung von Inkonsistenzen bzw. Korrektur von Angaben zur Verfügung gestellt. Um eine möglichst hohe Standardisierung während der Datenerhebung zu gewährleisten, erlaubt das Frageprogramm bei der Datenerhebung jedoch Abweichungen von den Frageformulierungen und der Fragenabfolge nur in sehr begrenztem Umfang. So können zwar mit dem „Zurück“-Button zuvor gemachte Angaben aufgerufen und korrigiert werden. Dieses „Zurückblättern“ im Fragebogen ist jedoch auf die aktuell berichtete Episode beschränkt.⁴ Auch kann ein abgeschlossenes Modul während der Datenerhebung nicht wieder aufgerufen werden, um Episoden zu ergänzen. Dagegen sind die flexiblen Eingriffsmöglichkeiten im Prüf- und Ergänzungsmodul vielfältiger (vgl. Abbildung 3).

Abbildung 3 Flexible Werkzeuge im Prüf- und Ergänzungsmodul

Datum von...	bis...	Episode	Art und Aktivität
Juli 1998	August 2000	1	AL: arbeitslos
Juli 1998	August 2000	1	ET: Handlanger auf dem Bau / Weimar, Stadt
September 2000	August 2002	2	AS: Realschule / Weimar, Stadt
August 2001	April 2003	1	EU: Manis
April 2003	April 2004	2	EU: Franz
September 2004	August 2005	1	AB: kaufmännische Angestellter / Weimar, Stadt

4 In modular strukturierten Längsschnitterhebungen gestaltet sich das Zurückblättern problematisch, da der Aufruf der einzelnen Eingabemasken nicht linear, sondern rekursiv erfolgt, d. h. bestimmte Maskensequenzen werden in Abhängigkeit von der Episodenanzahl mehrfach durchlaufen, so dass die „Entfernung“ zur benötigten Information in Abhängigkeit von den bereits erhobenen Lebensbereichen und deren Episodenanzahl sehr groß sein kann. Der unbegrenzte Rückgriff mit Hilfe eines „Zurück“-Buttons erscheint daher allein aus zeitlichen Gründen nicht sinnvoll. Darüber hinaus ist der Gefahr vorzubeugen, dass der Interviewer in der komplexen Interviewstruktur durch ein völlig freies Vor- und Zurückgehen die Übersicht verliert.

So wurde z. B. der „Episode frei einfügen“-Button geschaffen, weil es nur so möglich ist, ein bislang nicht erfasstes paralleles Ereignis zu erheben. Oder wenn eine Episode fälschlicherweise doppelt eingegeben wurde (z. B. das Referendariat als Ausbildung und Erwerbstätigkeit) kann mit Hilfe des „Markierte Episode löschen“-Buttons die falsch erfasste Episode gelöscht werden. Wenn einem Befragten erst sehr spät im Interview auffällt, dass er sich bei den Zeitangaben generell um ein Jahr geirrt hat, kann im Prüf- und Ergänzungsmodul durch Doppelklick auf die entsprechende Episode eine Maske aufgerufen werden, in der die Zeitangaben korrigiert werden können. Und weil es nicht auszuschließen ist, dass in der Datenrevision in Einzelfällen komplexe und zahlreiche Inkonsistenzen auftreten, die Interviewer und Befragte in ihrer Problemlösungskompetenz überfordern, haben wir die Buttons „Zeiten bis zur markierten Episode ohne Prüfung akzeptieren!“ und „Prüfung abbrechen!“ als pragmatischen Notausstieg aus der Datenrevision konzipiert.

3.5 Konsequenzen für die Interviewerausbildung

Die Erhebung von Lebensverläufen mit einem solch komplexen computerbasierten Instrument, das nicht nur standardisierte Elemente enthält, sondern auch flexibel gesteuert werden kann und in das überdies die Datenedition integriert ist, stellt hohe Anforderungen an die Interviewer. Die Interviewer müssen sehr gut auf ihre kommunikativen Aufgaben vorbereitet und in ihrer Kommunikationsfähigkeit intensiv geschult werden. Sie sollten deshalb weniger als „Reiz-Lieferanten“, sondern vielmehr als kompetente „Feld-Agenten“ fungieren. Vor diesem Hintergrund sollte bei der Feldarbeit auf folgende Prinzipien Wert gelegt werden:

Kompetenz. Da „TrueTales“ die Interviewer mit weiter reichenden Freiheiten ausstattet als es für gewöhnlich in standardisierten Interviews der Fall ist, müssen sie sowohl ein umfassendes Verständnis von den Erhebungskonzepten, den Zielen und der Datenstruktur als auch von den kognitiven und kommunikativen Komplikationen des Datenerhebungsprozesses haben. Die Interviewer sollten deshalb nicht nur inhaltlich geschult, sondern auch mit den – bei den Befragten ablaufenden – Erinnerungsprozessen vertraut gemacht werden und lernen, effektive Nachfragen zu stellen. Darüber hinaus erfordern die langen und komplizierten Interviews ein hohes Maß an Aufmerksamkeit auch von Seiten der Befragten. Dementsprechend empfiehlt es sich, den Interviewern nicht nur beizubringen, mit welchen Strategien sie Personen zur Teilnahme an der Studie bewegen, sondern auch zu vermitteln, wie sie die Befragten während des Interviews immer wieder zu einer konzentrierten Mitarbeit motivieren können.

Einbindung. Für die Sicherung der Datenqualität ist ein kontinuierlicher Austausch zwischen Forschern und Interviewern unerlässlich, weil dadurch einerseits die kontinuierliche Beratung der Interviewer sowie die Überprüfung ihrer Kompetenzen sichergestellt und andererseits Erfahrungen der Interviewer für die Verbesserung des Erhebungsinstrumentes, der Erhebungskonzepte sowie der Organisation der Interviews genutzt werden können. Durch Supervision kann eine fortlaufende Unterstützung und Kontrolle der Interviewer während der Datenerhebung gewährleistet werden. Aber die im Feld auftretenden – im Interviewerhandbuch oft nicht behandelten – Schwierigkeiten

sollten von den Forschern in während der Feldphase stattfindenden Interviewerbesprechungen aufgegriffen und eine praktikable Lösung gefunden werden. Im Gegenzug erfahren die Forscher von den Interviewern, inwiefern die Erhebungskonzepte den tatsächlichen Lebensverläufen zuwider laufen und welche Konstellationen nicht bedacht wurden.

Verantwortlichkeit. Den Interviewern sollte vermittelt werden, dass sie für die Qualität der Lebensverlaufsdaten, die sie in den Interviews erheben, Verantwortung tragen und übernehmen. Sie sollten in die Lage versetzt werden, eventuelle Fehler, die sie im Interview verursacht haben oder die durch Missverständnisse zwischen ihnen und den Befragten entstanden sind, selbst zu beheben oder zumindest derart zu dokumentieren, dass sich die Daten später problemlos korrigieren lassen. Um dies sicher zu stellen, sollte ein geschulter Projektmitarbeiter stichprobenartig ausgewählte Interviews eines jeden Interviewers auf inhaltliche und kommunikative Aspekte hin überprüfen, z. B. ob der Interviewer von der Standardisierung nicht zu weit abgewichen ist, die Befragten bei der Erinnerungsarbeit angemessen unterstützt wurden oder der Interviewer über das ganze Interview hinweg auf den Befragten einging und dennoch neutral blieb. Durch ein regelmäßiges Feedback sollten die Interviewer einerseits auf eine qualitativ hochwertige Arbeitsweise kontrolliert werden und andererseits die Möglichkeit bekommen, Probleme in der Interviewführung intensiv zu besprechen, um ihnen beim nächsten Interview begegnen zu können.

4 Hat sich „TrueTales“ bewährt?

Um zu untersuchen, ob das Erhebungsinstrument „TrueTales“ es den Befragten erleichtert, sich an Ereignisse zu erinnern und diese korrekt zu datieren, haben wir eine Methodenstudie „Kognition und Kommunikation bei der Erhebung retrospektiver Längsschnittdaten (LV-CC)“ (vgl. Reimer/Matthes/Mayer 2006) durchgeführt. Eine Verbesserung der Datenqualität lässt sich nur im direkten Vergleich mit herkömmlichen Erhebungsinstrumenten überprüfen. Deshalb haben wir die Methodenstudie als Split-Ballot-Studie angelegt, bei der einerseits ein Erhebungsinstrument zum Einsatz kam, das im Wesentlichen die in der letzten Lebensverlaufsbefragung des Max-Planck-Instituts für Bildungsforschung implementierten Techniken und Werkzeuge bei der Datenerhebung nachahmt (LV-CC Kurz). Andererseits wurde ein Erhebungsinstrument eingesetzt, das „TrueTales“ mit allen beschriebenen Techniken und Werkzeugen entspricht (LV-CC Lang). LV-CC Kurz enthielt keine Einblendetabellen, es wurde nicht auf die Angaben aus dem letzten Interview zurückgegriffen (Aufsetzepisoden), d. h. die Befragten wurden z. B. lediglich gefragt, ob sie im Januar 1993 erwerbstätig waren, und es konnte keine Datenrevision vorgenommen, sondern lediglich Lücken gefüllt werden.

In der Methodenstudie wurden 600 Frauen und Männer der 1964er Geburtskohorte befragt, die bei der letzten Lebensverlaufsbefragung des Max-Planck-Instituts für Bildungsforschung (LV-West 64/71) ihre Panelbereitschaft erklärt hatten. Die Befragten wurden mittels einer Zufallsauswahl der Erhebung mit LV-CC Kurz und LV-CC Lang zugewiesen. Durch die doppelte Erhebung

eines bereits beim letzten Interview erhobenen Zeitraumes haben wir einen Überlappungszeitraum (1993 bis zum letzten Interviewzeitpunkt) geschaffen, für den wir die Übereinstimmung der Angaben und somit die jeweils erreichte Datenqualität messen können (vgl. Abbildung 4).

Abbildung 4 Überlappungszeitraum für eine fiktive Person

Jahr	...	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	
LV-CC					Erwerbstätigkeit 1			Arbeitslosigkeit 1			Erwerbstätigkeit 2							
LV-64/71	Erwerbstätigkeit 1							Arbeitslosigkeit 1										

Dabei betrachten wir als die validesten uns zur Verfügung stehenden Informationen für den Überlappungszeitraum die edierten Daten der LV-West 64/71. Durch einen Vergleich dieser Daten mit den Daten, die wir einerseits mit LV-CC Kurz und andererseits mit LV-CC Lang gewonnen haben, ist es möglich, die Effektivität des Einsatzes der beiden Erhebungsinstrumente in einem ökologisch zuverlässigen Setting (also mit tatsächlichen Interviewern und Befragten), in einem kontrollierten experimentellen Design (in dem andere Einflussfaktoren konstant gehalten oder randomisiert werden können) und mit einem einheitlichen Bewertungsmaßstab (an dem die Datengüte gemessen werden kann) zu überprüfen. Die folgende Designskizze liefert einen kompakten Überblick über Konzeption und Rahmendaten der Methodenstudie (vgl. Tabelle 1).

Tabelle 1 Designskizze der Methodenstudie

	LV-CC Kurz	LV-CC Lang („True Tales“)
Stichprobe	Ausgangsstichprobe: 1964 geborene Personen, die in der Lebensverlaufsstudie LV-West 64/71 ihre Panelbereitschaft erklärt hatten: N = 1.349	
Kontaktierte Fälle	N = 448	N = 448
Realisierte Fälle	N = 299	N = 301
Erhebungsmethode	Computergestützte telefonische Interviews (CATI)	
Erhebungszeitraum	Juli 2004 - August 2004	Juni 2004 - August 2004
Durchschnittliche Interviewdauer	17,7 Minuten	19,2 Minuten
Feldsteuerung	11 Interviewer/-innen, die intensiv geschult und ausgebildet wurden (vgl. Matthes/Reimer/Mayer 2006)	

Inhaltlich ist die Methodenstudie eng an das übliche Fragenprogramm früherer Teilprojekte der Deutschen Lebensverlaufsstudie (GLHS) angelehnt. Da der Fokus dieser Studie vor allem auf der Nennung sowie der richtigen Datierung von Episoden lag, wurde jedoch zum einen auf einige Module komplett und zum anderen innerhalb der ausgewählten Module auf einige, für die methodologische Zielsetzung dieser Studie weniger relevante Fragen verzichtet.

Erinnerungsfehler führen leicht dazu, dass Ereignisse und Übergänge seltener als tatsächlich aufgetreten berichtet werden. Wenn „TrueTales“ Erinnerungsfehlern entgegenwirkt, müssten die Ereignisabfolgen in den mit LV-CC Lang geführten Interviews mehr Episoden pro Modul enthalten als mit LV-CC Kurz. Die Anzahl der in LV-CC Kurz und LV-CC Lang jeweils berichteten Episoden lässt sich jedoch nicht direkt vergleichen, da die Befragten in der LV-West 64/71 unterschiedlich viele Episoden aufweisen, d. h. die Interpretation der Anzahl der Episoden muss vor dem Hintergrund der Anzahl der in der LV-West 64/71 berichteten Episoden erfolgen. In der folgenden Tabelle ist deshalb die Anzahl der von den mit LV-CC Kurz befragten Personen berichteten Episoden der Anzahl der von diesen Personen in der LV-West 64/71 berichteten Episoden gegenübergestellt. Haben in der LV-West 64/71 299 Personen 37 Arbeitslosigkeitsepisoden berichtet, so berichteten diese Personen in der LV-CC Kurz nur 34 Arbeitslosigkeitsepisoden, während die Anzahl berichteter Arbeitslosigkeitsepisoden von 42 in der LV-West 64/71 auf 63 in der LV-CC Lang steigt (vgl. Tabelle 2).

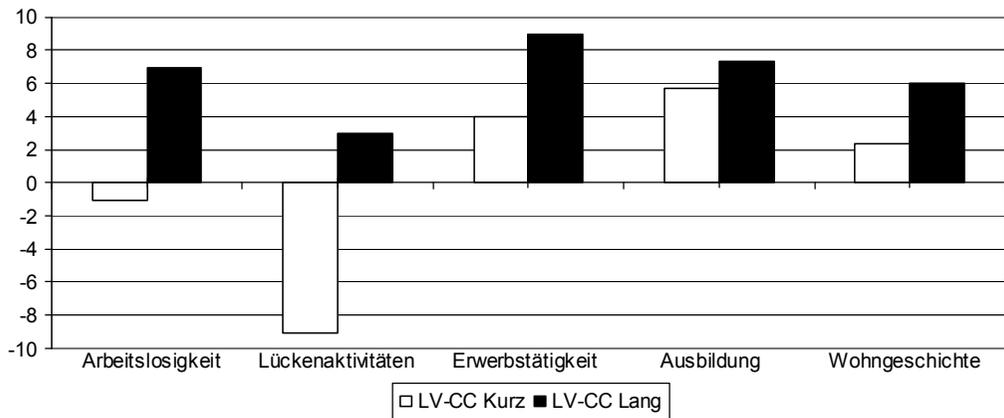
Tabelle 2 Anzahl der berichteten Episoden

Modul	Befragte der LV-CC Kurz		Befragte der LV-CC Lang	
	In LV-West 64/71	In LV-CC Kurz	In LV-West 64/71	In LV-CC Lang
Arbeitslosigkeit	37	34	42	63
Lückenaktivitäten	171	144	144	153
Erwerbstätigkeit	414	426	471	498
Ausbildung	46	63	57	79
Wohngeschichte	388	395	393	411
Befragte	299		301	

Auf den ersten Blick fällt auf, dass sowohl in LV-CC Kurz als auch in LV-CC Lang häufig mehr Episoden berichtet wurden als in der LV-West 64/71. Dieser Tatbestand lässt sich damit erklären, dass in LV-CC die Fragen, die maßgeblich dafür sind, ob eine weitere Episode erfasst wird oder nicht, sehr genau spezifiziert wurden. In der LV-West 64/71 wurde z. B. gefragt: „Haben Sie gleich anschließend oder auch später eine weitere berufliche Ausbildung gemacht, eine unterbrochene Ausbildung fortgesetzt oder ein Studium aufgenommen?“. In der LV-CC wurde dagegen gefragt: „Haben Sie nach {Einspielung letztes Interviewdatum} (noch) eine Ausbildung gemacht? Denken Sie dabei bitte auch an Volontariate, Praktika, Meisterlehrgänge, Anerkennungsjahre, Referendariate, Trainees, Promotionen, Habilitationen und Facharztweiterbildungen.“ Weil vorgelesen werden musste, was der Befragte unter einer Ausbildung zu verstehen hatte, wurde eine höhere Episodenanzahl erreicht.

Da wir in LV-CC Kurz und LV-CC Lang den gleichen Fragestimulus verwendet haben, kann man das Ergebnis, dass bei der Verwendung von LV-CC Lang in allen Modulen die Steigerung der berichteten Episodenanzahl höher war als bei der Verwendung von LV-CC Kurz, als Beweis für die Vorzüge der erinnerungsunterstützenden Befragung mit Hilfe von „TrueTales“ interpretieren. In der folgenden Abbildung ist dargestellt, um wie viel seltener bzw. häufiger Episoden in den verschiedenen Modulen pro Person für den Überlappungszeitraum mit LV-West 64/71 berichtet wurden (vgl. Abbildung 5).

Abbildung 5 Veränderung der Anzahl berichteter Episoden gegenüber LV-West 64/71



Insbesondere Arbeitslosigkeits- und Lückenepisoden können mit LV-CC Lang besser erinnert werden als mit LV-CC Kurz. Dieser Vorteil tritt insbesondere für die Befragten deutlich hervor, die komplexe Biografien mit weniger konventionellen Episoden (z. B. Arbeitslosigkeit) bzw. häufigere und kürzere Episoden haben. So profitieren Frauen deutlich stärker von der in LV-CC Lang umgesetzten Befragungstechnik als Männer. In allen Modulen, aber ganz besonders im Modul Lückenaktivitäten sowie Erwerbstätigkeiten werden von Frauen mit LV-CC Lang mehr Episoden berichtet als mit LV-CC Kurz (vgl. Tabelle 3).

Tabelle 3 Anzahl der berichteten Episoden bei Frauen

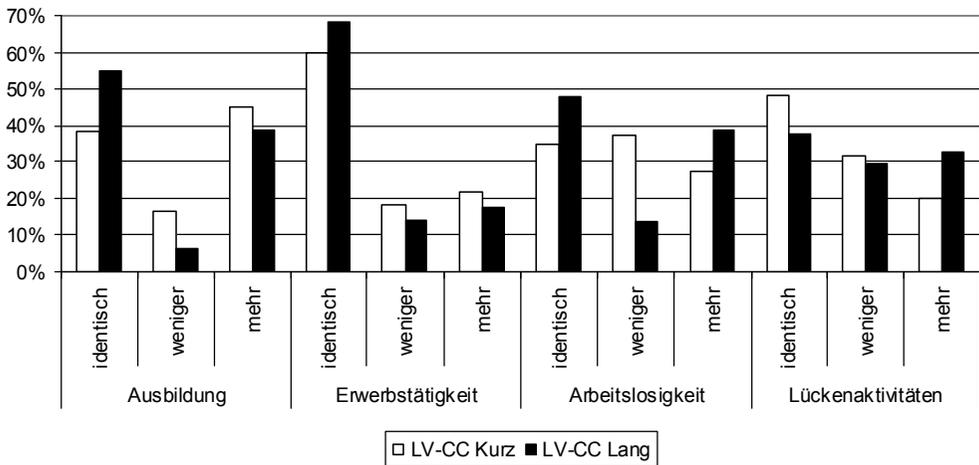
Modul	Befragte Frauen der LV-CC Kurz		Befragte Frauen der LV-CC Lang	
	In LV-West 64/71	In LV-CC Kurz	In LV-West 64/71	In LV-CC Lang
Arbeitslosigkeit	24	15	20	28
Lückenaktivitäten	153	135	125	148
Erwerbstätigkeit	165	172	197	215
Ausbildung	12	18	25	33
Wohngeschichte	196	189	196	202
Befragte	150		146	

Dies ist ein Hinweis darauf, dass „TrueTales“ insbesondere für die Biografien von Frauen besser funktioniert, die in der Regel häufiger Veränderungen und atypische Episoden aufweisen und die daher in Bezug auf Erinnerungsprobleme systematisch benachteiligt sind.

Die Überlegenheit von LV-CC Lang wird aber auch im direkten Vergleich der Angaben der Befragten aus dem früheren und dem späteren Interview deutlich. Die Anzahl der mit LV-CC Lang erhobenen Ausbildungs- und Erwerbsepisoden stimmt in höherem Maße mit der Anzahl der im

LV-West 64/71-Interview berichteten überein als die mit LV-CC Kurz erhobenen Episoden. Auch dass mit LV-CC Lang mehr Ausbildungs- und Lückenepisoden berichtet wurden, deutet darauf hin, dass LV-CC Lang die Erinnerungsarbeit der Befragten besser unterstützt als LV-CC Kurz (vgl. Abbildung 6).

Abbildung 6 Übereinstimmung der individuellen Anzahl berichteter Episoden gegenüber LV-West 64/71



Von besonderem Interesse ist, wie häufig dabei die Werkzeuge zur Bereitstellung individueller biografischer Abruf-/Hinweisreize als auch die standardisierten und flexiblen Tests und Korrekturmöglichkeiten im Prüf- und Ergänzungsmodul zum Einsatz gekommen sind. Sowohl im Datenerhebungs- als auch Datenrevisionsteil nutzten die Interviewer die Werkzeuge, um den Befragten Informationen zur Nutzung paralleler Erinnerungspfade anzubieten und um Lücken und Inkonsistenzen zu identifizieren und zu korrigieren (vgl. Tabelle 4).

Tabelle 4 Häufigkeit der Verwendung der in LV-CC Lang zur Verfügung gestellten Werkzeuge

Werkzeug	Häufigkeit	
Einblendetabelle:	270	
Standardisierte Prüfung:	bei Lückenproblemen	187
	bei Überschneidungsproblemen	184
	bei unklaren Zeiten	11
Flexibler Dateneingriff:	Episode löschen	37
	Episode einfügen	38

In 270 der insgesamt 301 mit LV-CC Lang geführten Interviews verwendeten die Interviewer die Einblendetabellen. Im Prüf- und Ergänzungsmodul wurde die filtergesteuerte standardisierte Prüf- und Ergänzungsroutine 382 mal aufgerufen: 187 mal, um Lückenergebnisse nachzuerheben, 184 mal, um zu prüfen, ob eine Überschneidung zwischen verschiedenen Episoden korrekt ist, und 11 mal, um fehlende Datumsangaben am Beginn bzw. Ende einer Episode zu klären. Darüber hinaus wurde in 75 Interviews von der Möglichkeit zum flexiblen Entfernen oder Einfügen von Episoden Gebrauch gemacht.

Mit dem Einsatz von „True Tales“ in dieser Methodenstudie konnten wir also nachweisen, dass die neu entwickelten Techniken und Werkzeuge zur Unterstützung der Erinnerungsarbeit sowohl die Qualität der erhobenen Lebensverlaufsdaten verbessert als auch die Anschlussgenauigkeit von Informationen aus unterschiedlichen Panelwellen erhöht. Aber auch die Interviewer empfanden den Einsatz von „TrueTales“ deutlich angenehmer, weil das Interview besser steuerbar war und die Kommunikationssituation mit den Befragten insgesamt als gelungener eingeschätzt wurde.

5 Zusammenfassung

Das Ziel der Entwicklung von „TrueTales“ war es, auf Erinnerungsfehlern der Befragten beruhende Inkonsistenzen bei der Fortschreibung bzw. bei der Erhebung von retrospektiven Lebensverlaufsdaten zu minimieren und damit deren Editions Aufwand deutlich zu verringern. Aufbauend auf Erkenntnissen aus der Gedächtnispsychologie, wie Ereignisse erinnert und datiert werden, und auf langjährigen Erfahrungen bei der Erhebung und Edition von retrospektiven Lebensverlaufsdaten haben wir verschiedene Techniken und Werkzeuge zum Einsatz gebracht, mit denen die Interviewer in computergestützten, telefonischen Befragungen die Erinnerungsarbeit der Befragten besser unterstützen können und sich damit die Qualität der Angaben der Befragten bereits während der Datenerhebung verbessert:

- Zur Vermeidung von Erinnerungsfehlern hat sich als besonders ertragreich herausgestellt, die modulare und sequentielle Form der Datenerhebung zu kombinieren. Durch konsequente Modularisierung wird erreicht, dass die Befragten parallele Bildungs- oder Erwerbsaktivitäten und kurze Arbeitslosigkeitsphasen häufiger berichten. Innerhalb der Module sollte dann chronologisch vorgegangen werden, um das sequentielle Vorgehen beim Erinnern zu unterstützen.
- Um die mit dem mehrfachen chronologischen Durchlaufen der verschiedenen Lebensbereiche verbundenen Probleme beim Datieren zu kompensieren, sollte die Befragung in zwei Teile geteilt werden: die Datenerhebung, in der die Fragen gestellt und die Angaben erfasst werden, und die Datenrevision, in der zeitliche Inkonsistenzen visualisiert und – falls nötig – korrigiert werden können.
- Sowohl bei der Datenerhebung als auch bei der Datenrevision sollten Techniken und Werkzeuge zum Einsatz gebracht werden, die die Befragten bei der Erinnerungsarbeit unterstützen, wie das Einspielen autobiographischer Angaben der Befragten aus der vorherigen Panelwelle in die standardisierten Einstiegsfragen (Aufsetzepisoden) zur Verringerung von

Episodenanschlussfehlern; das Einblenden autobiographischer Angaben aus dem aktuellen Interview als Kontext für die zu erinnernde Informationen; die Bereitstellung von Einblendetabellen zur Unterstützung paralleler und sequentieller Erinnerungsstrategien der Befragten.

- In beiden Teilen sollten sowohl filtergesteuerte, standardisierte Frageprozeduren, als auch Werkzeuge für ein flexibles Eingehen auf die Angaben der Befragten enthalten sein. Eine kontrollierte Abweichung von der streng standardisierten Befragungstechnik kommt den Erinnerungsprozessen häufig entgegen und trägt zur Verbesserung der Kommunikationssituation im Interview bei. Dabei können kalenderartige Instrumente die aktive und klärende Kommunikation unterstützen, da sich der Interviewer flexibel und konkret auf das beziehen kann, was der Befragte bereits über sein Leben angegeben hat. Darüber hinaus erhält die Interviewsituation dadurch den Anschein eines „normalen Gesprächs“, was sich positiv auf die Motivation der Befragten auswirkt, auch an längeren Interviews mitzuwirken.

Die Ergebnisse der eigens für die Evaluation von „TrueTales“ durchgeführten Methodenstudie LV-CC wie auch Erfahrungen bei der Wiederholungsbefragung der ost- und westdeutschen Geburtskohorte 1971 (LV-Panel 71) belegen, dass diese Techniken und Werkzeuge Erinnerungsfehler bei der retrospektiven Erhebung von Ereignisgeschichten vermeiden helfen. Insbesondere bei der Erhebung komplexer bzw. unkonventioneller Ereignisgeschichten profitiert man bei Verwendung von „TrueTales“.

„TrueTales“ wurde von uns konzipiert für die Wiederholungsbefragung der 1971 in Ost- und Westdeutschland Geborenen, in der ein retrospektives Intervall von etwa 11 Jahren erfasst wird. Das Erhebungsinstrument lässt sich aber jedem anderen Zeitintervall und anderen Forschungszielen anpassen. Für die Belange der Mediennutzungs- oder der Zeitbudgetforschung können sogar für einen Zeitraum von einer Woche (in Tagen untergliedert) oder auch nur eines Tages (in Stunden oder Doppelstunden untergliedert) Ereignisgeschichten und Aktivitätsabfolgen beim Eintrag auf Konsistenz und Lückenlosigkeit geprüft werden. Strategie und Technologie bieten darüber hinaus große Potentiale für die Erstellung von selbstauszufüllenden Fragebögen, insbesondere im Online-Bereich. In der GLHS hat sich seit jeher bestätigt, dass Menschen, die über persönlich bedeutsame Aspekte ihres Lebens oder ihrer Biographie Auskunft geben sollen und sich als individuelle Gesprächspartner ernst genommen fühlen, bereit sind, sehr lange Interviews zu geben oder komplexe Fragebögen mit dem entsprechenden Aufwand an Zeit und Sorgfalt auszufüllen.

Literatur

- Auriat, N., 1993: My wife knows best: A comparison of event dating accuracy between the wife, the husband, the couple and the Belgium population register. *Public Opinion Quarterly* 57: 165-190.
- Axinn, W., L. D. Pearce und D. Ghimire, 1999: Innovations in life history calendar applications. *Social Science Research* 28: 243-264.
- Barclay, C. R., 1986: Schematization of autobiographical memory. S. 82-99 in: D. C. Rubin (Hg.): *Autobiographical Memory*. Cambridge: Cambridge University Press.
- Barsalou, L. W., 1988: The content and organization of autobiographical memories. S. 193-243 in: U. Neisser und E. Winograd (Hg.): *Remembering reconsidered: Ecological and traditional approaches to the study of memory*. New York: Cambridge University Press.

- Belli, R. F. 1998: The structure of autobiographical memory and the Event History Calendar: Potential improvements in the quality of retrospective reports in surveys. *Memory* 6: 383-406.
- Belli, R. F., E. H. Lee, F. P. Stafford und C. H. Chou, 2004: Calendar and Question-List Survey Methods: Association Between Interviewer Behaviors and Data Quality. *Journal of Official Statistics* 20: 185-218.
- Belli, R. F., W. Shay und F. Stafford, 2001: Event history calendars and question list surveys: A direct comparison of interviewing methods. *Public Opinion Quarterly* 65: 45-74.
- Bluck, S. und T. Habermas, 2000: The Life Story Schema. *Motivation & Emotion* 25: 121-147.
- Brewer, W. F., 1988: Memory for randomly sampled autobiographical events. S. 21-90 in: U. Neisser und E. Winograd (Hg.): *Remembering reconsidered: Ecological and traditional approaches to the study of memory*. New York: Cambridge University Press.
- Brückner, E., 1993: Lebensverläufe und gesellschaftlicher Wandel - Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1919-1921. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 44.
- Brückner, E., 1994: Erhebung ereignisorientierter Lebensverläufe als retrospektive Längsschnittkonstruktion. S. 38-69 in: R. Hauser, N. Ott und G. Wagner (Hg.): *Mikroanalytische Grundlagen der Gesellschaftspolitik*. Berlin: Akademie-Verlag.
- Brückner, E. und K. U. Mayer, 1998: Collecting Life History Data. Experiences From the German Life History Study. S. 152-181 in: J. Z. Giele und G. Elder (Hg.): *Methods of Life Course Research: Qualitative and Quantitative Approaches*. Sage: Thousands Oaks.
- Brückner, H., 1995: Surveys don't lie, people do? An analysis of data quality in a retrospective life course study. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 50.
- Brückner, H. und K. U. Mayer, 1995: Lebensverläufe und gesellschaftlicher Wandel -Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1954-1956 und 1959-1969. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 48.
- Caspi, A., T. E. Moffitt, A. Thornton, D. Freedman, J. W. Amell, H. L. Harrington, J. Smeijers und P.H. Silva, 1996: The Life History Calendar: A research and clinical assessment method for collecting retrospective event-history data. *International Journal of Methods in Psychiatric Research* 6: 101-114.
- Conway, M. A., 1996: Autobiographical knowledge and autobiographical memory. S. 67-93 in: D. C. Rubin (Hg.): *Remembering our past: Studies in autobiographical memory*. Cambridge/Mass.: Cambridge University Press.
- Conway, M. A. und C. W. Pleydell-Pearce, 2000: The construction of autobiographical memories in the self-memory system. *Psychological Review* 107: 261-288.
- Elias, P., 1991: Methodological, statistical and practical issues arising from the collection and analyses of work history information by survey techniques. *Bulletin de Méthodologie Sociologique* 31: 3-31.
- Friedman, W. J., 1993: Memory for the time of past events. *Psychological Bulletin* 113: 44-66.
- Goedicke, A., B. Lichtwardt und K. U. Mayer, 2004: *Dokumentationshandbuch Ostdeutsche Lebensverläufe im Transformationsprozess (LV-Ost Panel)*. Berlin: Max-Planck-Institut für Bildungsforschung. Materialien aus der Bildungsforschung 75.
- Hillmert, S., 2002: Edition von Lebensverlaufsdaten: zur Relevanz einer systematischen Einzelfallbearbeitung bei standardisierten Befragungen. *ZUMA-Nachrichten* 51: 120-140.
- Hillmert, S., R. Künster, P. Spengemann und K. U. Mayer, 2004: *Projekt Ausbildungs- und Berufsverläufe der Geburtskohorten 1964 und 1971 in Westdeutschland, Dokumentation*. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 78.
- Huttenlocher, J., L. V. Hedges und N. M. Bradburn, 1990: Reports of elapsed time: Bounding and rounding processes in estimation. *Journal of Experimental Psychology, Learning, Memory and Cognition* 16: 196-213.
- Huttenlocher, J., L. V. Hedges und V. Prohaska, 1988: Hierarchical organization in ordered domains: Estimating the dates of events. *Psychological Review* 95: 471-484.
- Lampinen, J. M., J. M. Faries, J. S. Neuschatz und M. P. Togli, 2000: Recollection of things schematic: The influence of scripts on recollective experience. *Applied Cognitive Psychology* 14: 543-554.
- Larsen, S. F. und C. P. Thompson, 1995: Reconstructive memory in the dating of personal and public events. *Memory and Cognition* 23: 780-790.
- Larsen, S. F., C. P. Thompson und T. Hansen, 1996: Time in autobiographical memory. S. 129-154 in: D. C. Rubin (Hg.): *Autobiographical memory*. Cambridge: Cambridge University Press.
- Loftus, E. F. und W. Marburger, 1983: Since the eruption of Mt. St. Helens, has anyone beaten you up? Improving the accuracy of retrospective reports with landmark events. *Memory and Cognition* 11: 114-120.

- Matthes, B., B. Lichtwardt und K. U. Mayer, 2004: Dokumentationshandbuch Ostdeutsche Lebensverläufe im Transformationsprozess (LV-Ost 71). Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 76.
- Matthes, B. und K. U. Mayer, 2006: Dokumentation des Projekts Frühe Karrieren und Familiengründung: Lebensverläufe der Geburtskohorte 1971 in Ost- und Westdeutschland (LV-Panel 71). Berlin: Max-Planck-Institut für Bildungsforschung.
- Mayer, K. U. und E. Brückner, 1989: Lebensverläufe und Wohlfahrtentwicklung: Konzeption, Design und Methodik der Erhebung von Lebensverläufen der Geburtsjahrgänge 1929-31, 1939-41, 1949-51. Berlin: Max-Planck-Institut für Bildungsforschung, Materialien aus der Bildungsforschung 35.
- Poulain, M., B. Riandey und J.-M. Firdion, 1991: Enquête Biographique et Registre Belge de Population: Une Confrontation des Données. *Population* 1: 89-104.
- Reimer, M., 2001: Die Zuverlässigkeit des autobiographischen Gedächtnisses und die Validität retrospektiv erhobener Lebensverlaufsdaten: Kognitive und erhebungspragmatische Aspekte. Berlin: Max-Planck-Institut für Bildungsforschung.
- Reimer, M., 2005: Autobiografisches Gedächtnis und retrospektive Datenerhebung: Die Rekonstruktion und Validität von Lebensverläufen (Studien und Berichte 70). Berlin: Max-Planck-Institut für Bildungsforschung.
- Reimer, M., B. Matthes und K. U. Mayer, 2006: Dokumentation des Projekts Kognition und Kommunikation bei der Erhebung retrospektiver Längsschnittdaten (LV-CC). Berlin: Max-Planck-Institut für Bildungsforschung.
- Shum, M. S., 1998: The role of temporal landmarks in autobiographical memory processes. *Psychological Bulletin* 124: 423-442.
- Shum, M. S. und L. J. Rips, 1999: The respondent's confession: Autobiographical memory in the context of surveys. S. 95-109 in: M. G. Sirken, D. J. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur und R. Tourangeau (Hg.): *Cognition and Survey Research*. New York: Wiley.
- Solga, H., 2001: Longitudinal surveys and the study of occupational mobility: panel and retrospective design in comparison. *Quality and Quantity* 35: 291-309
- Sudman, S., N. M. Bradburn und N. Schwarz, 1996: Thinking about answers: The application of cognitive processes to survey methodology. San Francisco: Jossey-Bass.
- Thompson, C. P., D. J. Herrmann, D. Bruce, J. D. Read, D. G. Payne und M. P. Toggia, 1998: Autobiographical memory: Theoretical and applied perspectives. Mahwah, N.J.: Erlbaum.
- Vaart, W. van der, 2002: The time-line: The effects of an experimental aided recall technique in a real life survey: Paper presented at the International Conference on Questionnaire Development, Evaluation and Testing Methods. Charleston: South Carolina.
- Yoshihama, M., K. Clum, A. Crampton und B. Gillespie, 2002: Measuring the lifetime experience of domestic violence: Application of the life History Calendar method. *Violence & Victims* 17: 297-317.
- Zouwen, J. van der, W. Dijkstra und W. van der Vaart, 1993: Effects of measures aimed at increasing the quality of recall data. *Bulletin de Méthodologie Sociologique* 39: 3-19.

Korrespondenzadresse: Britta Matthes
Institut für Arbeitsmarkt- und Bildungsforschung
Forschungsbereich 7: Bildung und Beschäftigung, Lebensverläufe
Weddingstraße 20 - 22
90478 Nürnberg
Britta.Matthes@iab.de

Rezensionen



ANDREAS DIEKMANN (Hg.), 2006: Methoden der Sozialforschung. Kölner Zeitschrift für Soziologie und Sozialpsychologie. Sonderheft 44/2004. Wiesbaden: VS Verlag für Sozialwissenschaften. ISBN 3-531-14362-X, 613 Seiten, 52,90 EUR.

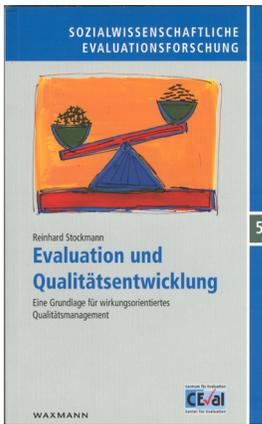
Eine Einführung in die Methoden der Sozialforschung – wie der Titel vielleicht nahelegen mag – bietet dieser Sonderband der Kölner Zeitschrift für Soziologie und Sozialpsychologie nicht. Absicht war ebenfalls nicht, einen umfassenden Überblick über die Verfahren der empirischen Sozialforschung zu bieten. Vielmehr sollten – in ausgewählten Bereichen, aber nicht „flächendeckend“ (so fehlt etwa die interkulturell vergleichende Sozialforschung) – innovative Entwicklungen aufgezeigt und die herkömmliche Praxis kritisch hinterfragt werden. Behandelt werden insbesondere solche neueren Entwicklungen, die noch keinen Eingang in die Einführungsliteratur in die Methoden der Empirischen Sozialforschung gefunden haben.

Das Buch beginnt mit einer – unbedingt lesenswerten – Einleitung von Andreas Diekmann, in der zunächst die Entwicklung der empirischen Sozialforschung in den letzten Jahrzehnten pointiert beschreiben und damit der Hintergrund für die einzelnen Beiträge gesetzt wird: Die Veränderung bzw. Weiterentwicklung der Methoden der Datenerhebung, der Analysierbarkeit großer Datensätze und der statistischen Analysemethoden durch technologische Innovationen. Im Anschluss daran skizziert er die einzelnen Beiträge aus 6 inhaltlichen Bereichen: „Wissenschaftstheoretische Grundlagen“, „Stichproben, Gewinnung von Daten, Datenqualität“, „Sozialer Kontext und Netzwerke“, „Zeitbezogene Daten“, „Spezielle Probleme der Datenanalyse“ sowie „Modellbildung und Simulation“. Die entsprechenden Ausführungen sind knapp, verfehlen aber ihre Funk-

tion nicht: den Lesern als Orientierung für das Weiterlesen zu dienen. Schließlich zieht Diekmann auch schon eine Bilanz aus den vorliegenden Beiträgen: Wie soll es mit der empirischen Sozialforschung weitergehen? Wo liegen die Risiken und Chancen? Er plädiert in diesem Zusammenhang für eine Erhöhung der Standards bei der Datenerhebung, eine ausgeprägtere „Replikationskultur“ (Wiederholung von Untersuchungen an neuen Daten), mehr Metaanalysen zum Vergleich von Forschungsergebnissen (um einen kumulativen Erkenntniszuwachs zu sichern) sowie mehr systematische Methodenexperimente.

Es ist bemerkenswert, dass es den meisten Autoren gelungen ist, auf jeweils relativ wenigen Seiten nicht nur die neuesten Entwicklungen aufzuzeigen bzw. die gängige Praxis kritisch zu diskutieren, sondern dies auch in einer Weise zu tun, die auch für Nicht-Fachleute in den jeweiligen Spezialgebieten durchaus zugänglich ist. Dabei werden jedoch nicht überall Lösungen vorgeschlagen, die bereits jetzt in den Kanon des „gesicherten“ Wissens der Sozialforschung aufgenommen werden könnten. Dazu befinden sich einige der Beiträge zu nahe an der Front der Forschung, so dass erst die Zeit zeigen wird, wie erfolgversprechend bzw. tragfähig sich einige der neueren Entwicklungen und Perspektiven erweisen. Dies trifft sowohl auf die Bereiche zu, die zu einem großen Teil wissenschaftsimmanent vorangetrieben werden (wie bei statistischen Verfahren) als auch dort, wo die methodischen Innovationen überwiegend von den gesellschaftlichen und technologischen Veränderungen „getrieben“ werden (im negativen wie im positiven Sinne), wie bei den Techniken der Datenerhebung (wobei die Entwicklungen im Bereich des Internets zur Zeit eher zu den positiven, die im Bereich des Telefons zu den negativen externen Faktoren zu zählen sind).

Die einzelnen Kapitel wird jeder mit unmittelbarem Gewinn lesen können, der in dem jeweiligen Gebiet oder in einem angrenzenden Feld lehrt oder forschet. Der gesamte Band eignet sich darüber hinaus auch als Grundlage für Lehrveranstaltungen für Fortgeschrittene, die einen Überblick über neuere Entwicklungen der Methoden der Sozialforschung bieten wollen.



REINHARD STOCKMANN, 2006: Evaluation und Qualitätswentwicklung. Eine Grundlage für wirkungsorientiertes Qualitätsmanagement. Sozialwissenschaftliche Evaluationsforschung Band 5. Münster: Waxmann. ISSN 1861-244 X, ISBN 3-8309-1621-3, 375 Seiten, 29,90 EUR.

Ausgangspunkt der Darstellung ist die Qualitätswentwicklung in Nonprofit-Organisationen. Um die Qualität von Programmen und Leistungen von Nonprofit-Organisationen beurteilen zu können, sind einerseits relevante Daten, andererseits geeignete Konzepte und Instrumente notwendig. Die Qualitätsmanagementkonzepte für Unternehmen eignen sich wenig für den Nonprofitbereich, da in Bezug auf Zielgruppen (Kunden), Wettbewerbssituation und organisationale Rahmenbedingungen deutliche Unterschiede bestehen. An Konzepten und Instrumenten wird sinnvoll und angemessen Erscheinendes oftmals aus verschiedenen „Steinbrüchen“ (S. 13) besorgt. Sinnvoll ist es jedoch, auf die eigenen Bedürfnisse zugeschnittene Konzepte und Instrumente zu entwickeln. Hierzu will das vorliegende Buch beitragen.

Ziel ist es, „eine theoretisch fundierte Evaluationskonzeption und eine Methode zu entwickeln, die (1.) für die Gewinnung von leistungs- und wirkungsbezogenen Daten genutzt werden kann, aus der (2.) Bewertungskriterien abgeleitet werden können, die sich für Evaluation insbesondere im Nonprofitsektor eignen und denen (3.) ein multidimensionales Kriterienset zur Beurteilung der Qualität von Nonprofit-Organisationen und ihrer Wirkungen möglich sein“ (S. 17) sollte. Darüber hinaus soll sich die „hier entwickelte Evaluationskonzeption prinzipiell für alle Tätigkeitsfelder und alle Phasen der Leistungserstellung von der Planung, über die Durchführung und über ein eventuelles Fördernde hinaus ... einsetzen“ lassen (S. 17). Weiteres Ziel: „... die konzeptionellen Ausarbeitungen, die parzelliert in verschiedenen Publikationen und ... Evaluationsberichten dokumentiert sind, zu vereinheitlichen, zu-

sammenzuführen und so weiterzuentwickeln, dass sie generell ... genutzt werden können“ (S. 21).

Das Buch ist in drei Bereiche gegliedert: Zunächst werden die zentralen Begriffe diskutiert (Kapitel 2), dann die für eine Instrumentenentwicklung zentralen Theorien erörtert (Kapitel 3) und schließlich der Evaluationsleitfaden und die für dessen Anwendung notwendigen Methoden und Techniken vorgestellt.

Kapitel 2, „Qualitätsmanagement und Evaluation“, beginnt mit der Diskussion des Begriffes „Qualität“ und deren Normierung. Sodann wird hinterfragt, ob bei steigender Qualitätsorientierung Konzepte des Qualitätsmanagements, die für Profit-Organisationen entwickelt wurden, für Nonprofit-Organisationen anwendbar sind. Nach dem Begriff der „Qualität“ wird der Begriff der „Evaluation“ diskutiert und verschiedene Aufgaben und Funktionen von „Evaluation“ vorgestellt, mit denen sich Daten sammeln und bewerten lassen. Ein größerer Exkurs betrachtet Monitoring im Vergleich zum Controlling. Es folgt ein Vergleich von mehr oder weniger handlungsoffenen Konzepten des Qualitätsmanagements und der Evaluation, in dem Gemeinsamkeiten und Unterschiede herausgearbeitet werden. Das Votum des Autors geht dahin, „eine theoretische Konzeption und ein(en) methodische(n) Ansatz für ein wirkungsorientiertes Monitoring- und Evaluationssystem“ aufzustellen (S. 96).

In Kapitel 3 wird eine „theoretische Konzeption der Evaluation“ erarbeitet. Hierbei ist der Begriff der „Wirkung“, als abhängige Variable, zunächst zu definieren und in seinen Dimensionen abzugrenzen. Dann folgt das Vorstellen der vier theoretischen Ansätze, die für ein wirkungsorientiertes Konzept zur Evaluation von Leistungsangeboten und Programmen herangezogen werden und einander komplementär ergänzen: Zur Betrachtung von Wirkungen von (chronologisch ablaufenden) Programmen im Prozessverlauf wird der theoretische Ansatz der Lebensverlaufsfor- schung herangezogen. Da Programme aber über Organisationen abgewickelt werden, werden zweitens organisationstheoretische Konzepte benötigt, um sowohl organisationsinterne Zusammenhänge als auch Aspekte der Beziehung einer Organisation mit ihrer Umwelt (S. 118) zu diskutieren. Drittens werden Konzepte der Innovations- und Diffusionsforschung betrachtet, da eine Programmintervention zu Innovation führen soll, die wiederum eine möglichst hohe Verbreitung finden soll. Abschließend wird viertens

ein Konzept von Nachhaltigkeit diskutiert, um Langfristigkeit und Reichweite von Innovation und deren verändernde Wirkungen auf das System und sich ändernde Umweltbedingungen zu messen.

Um nach möglichst theoriebasierten Kriterien einen Evaluationsleitfaden zu strukturieren und Kriterien für die Qualitätsbestimmung von Programmen zu erlangen, werden die theoretischen Konzepte zu unterschiedlichen Analyseperspektiven zusammengefasst und die Bewertungsfelder und -kriterien zur Konstruktion eines Evaluationsleitfadens aus Verlaufsmodell, Organisations- und Innovations-/Diffusionsforschung sowie einem multidimensionalen Nachhaltigkeitskonzept abgeleitet. Das Ergebnis sind Bewertungskriterien und vier Qualitätsdimensionen für die Bewertung der Leistungen und Wirkungen von Non-profit-Organisationen.

Kapitel 4 behandelt die Methodik der Evaluation und ihre Anwendung mit dem Muster-Evaluationsleitfaden als Kernstück. Dieser Muster-Evaluationsleitfaden, entwickelt aus den in Kapitel 3 dargestellten theoretischen Überlegungen, stellt ein universelles Grundmuster dar. Er lässt sich aus Sicht seines Konstrukteurs über Ergänzungen und Erweiterungen von Analysefragen auf alle Programme oder institutionalisierten Leistungsangebote anpassen. Der Leitfaden wird zunächst von der Gegenstandsbeschreibung, über Fragen zum Programmablauf, zu internen Organisationsstrukturen und externen Interventionsbereichen und der Programmqualität vorgestellt. Sodann werden Bearbeitungs- und Bewertungsverfahren und der Aufbau einer Indikatorendatenbank diskutiert.

Das Methodenkapitel beginnt mit einer Diskussion der Vor- und Nachteile unterschiedlicher Untersuchungsdesigns. Es folgt das Vorstellen der wichtigsten Erhebungsmethoden für das Erfassen von Wirkungen und kausaler Zusammenhänge. Dann wird ein Multimethodenansatz vorgeschlagen, um Schwächen einzelner Erhebungsinstrumente durch Stärken anderer auszugleichen (S. 178). Abgeschlossen wird das methodische Kapitel durch einerseits eine Diskussion der Anforderungen und Schwierigkeiten bei der praktischen Durchführung der Evaluation (Kap. 4.5) und andererseits praktische Hinweise für die Planung und Durchführung einer Evaluation (Kap. 4.6).

Das Buch endet mit einer zusammenfassenden Bewertung des oben dargestellten und diskutierten Evaluationsansatzes.

Durch die theoretische Verknüpfung des Qualitätsbegriffes mit dem Evaluationskonzept gelingt es, ein neues Verständnis der Evaluation zu entwickeln. Eine Evaluation ermittelt demnach nicht nur die Wirkung eines Programms, sondern bewertet es anhand von umfassenden Qualitätskriterien. Der im Kapitel 4.2 vorgestellte Evaluationsleitfaden dient der Programmbewertung anhand der Programmanalyse (diese umfasst Programmkonzept, Ressourcen, Umwelt- und Kontextbedingungen, Verlauf des Programms), Analyse der Wirkungsfelder (sowohl bezogen auf die Zielgruppe als auch auf die Trägerorganisation) und einer Bewertung der Programmqualität. Der Autor grenzt sich mit diesem Konzept von Evaluationsprogrammen ab, die ohne vorab definierte Zielkriterien arbeiten oder nur auf die erwartete Wirkung des Programms abzielen. Der Anspruch des Leitfadens ist dabei, sowohl intendierte als auch nicht intendierte nachhaltige Wirkungen und deren Größen feststellen zu können. Ungenügend bearbeitet wird in diesem Zusammenhang die Frage der Anwendung des Leitfadens, um diese anspruchsvollen Ziele erreichen zu können. Zwar wird auf die Bedeutung des Erhebungsdesigns (Kap. 4.4.1) hingewiesen, bei dem die Vergleiche zu mehreren Messzeitpunkten und mit einer Kontrollgruppe erfolgen sollten. Dennoch kommt die Problematik der konkreten Umsetzung zu kurz. Die methodische Herausforderung besteht gerade in der Ermittlung von Wirkungen und Wirkungsgrößen mit Hilfe eines angemessenen Erhebungsdesigns. Die vom Autor gegebenen Anwendungsbeispiele (im Kap. 5) zeigen die Verwendung des Leitfadens größtenteils für ex-post Evaluationen, die keine Urteile über die Wirkung und Wirkungsgröße zulassen. In Bezug auf die Verwendung konkreter Erhebungsmethoden wird ein kurzer Abriss (im Kap. 4.4.2) gegeben, der weniger die Anwendung des Leitfadens erklärt, sondern schlagwortartig Vorgehensweisen aneinanderreihet, deren Effekte nur dann einzuschätzen sind, wenn man die Methoden in Theorie und Praxis kennt und teils auch die Wertungen der zitierten Experten einzuordnen in der Lage ist. Da der optimale Einsatz einer Methode abhängig ist vom Erhebungsdesign und von der Fragestellung, ergibt sich häufig ohnehin die Notwendigkeit, einem Multimethodenansatz zu folgen.

Das vorliegende Buch erarbeitet ein umfassendes theoriebasiertes Evaluationskonzept für die Bewertung der Leistungserbringung in Nonprofit-Organisationen. Diese erfolgt durch eine Analyse des Programmkonzeptes, des Ablaufs, der Erhebung von internen (organisationalen) und externen Wirkungen und deren Nachhaltigkeit. Insofern kann der entwickelte Leitfaden Informationssuche organisieren und strukturieren und als eine Ausgangsbasis für Qualitätsmanagement im Nonprofit-Bereich verwendet werden. Es ist hier gelungen, fundierte Dimensionen und Kriterien aufzustellen. Erforderlich ist jedoch eine kritische Auseinandersetzung mit der Frage der Ermittlung von Programmwirkungen. Zusätzlich darf der erforderliche Operationalisierungsaufwand bei einer konkreten Evaluation nicht unterschätzt werden.

JÜRGEN H.P. HOFFMEYER-ZLOTNIK UND NATALIA MENOLD

Mitteilungen

Einführung in familien- und haushaltsbezogene Analysen mit dem Mikrozensus

24. - 25. Oktober 2007, GESIS-ZUMA, Mannheim

Familien- und haushaltsbezogene Analysen mit dem Mikrozensus

Der Mikrozensus bietet aufgrund seiner Anlage als Haushaltsstichprobe – befragt werden alle in einem Haushalt wohnenden Personen – ein wichtiges Datenpotenzial für haushalts- und familienbezogene Analysen mit sozialstrukturellem Schwerpunkt. Ziel dieses Workshops ist es, einen Einblick in dieses Potenzial zu vermitteln, aber auch vorhandene Grenzen aufzuzeigen. Hierfür wird am ersten Tag ein Überblick in die spezifischen Besonderheiten des Mikrozensus mit dem Schwerpunkt ‚Haushalt und Familie‘ gegeben. Am zweiten Tag erfolgt eine Vertiefung dieser Kenntnisse anhand von ausgewählten Analysebeispielen. Sofern von Seiten der Teilnehmer Interesse an der Diskussion von spezifischen Inhalten oder Analyseproblemen besteht, bitten wir, uns diese bis zum 31.8.2007 mitzuteilen (heike.wirth@gesis.org).

Der Workshop wird in enger Kooperation mit dem Statistischen Bundesamt (Gruppe VIII C – Mikrozensus, Arbeitskräftestichprobe, Haushalt und Familie) durchgeführt.

Interessenten werden gebeten, sich beim GESIS-ZUMA-Tagungssekretariat anzumelden:

workshop-mannheim@gesis.org

Tel.: 0621-1246-221

Anmeldeschluss: 22. September 2007

Teilnahmebeitrag: 120 €, Studenten 80 €

Maximale Teilnehmerzahl: 16

Kontakt:

Dr. Heike Wirth

GESIS-ZUMA

Postfach 12 21 55

68072 Mannheim

Tel.: 0621-1246-269

heike.wirth@gesis.org

<http://www.gesis.org/Dauerbeobachtung/GML/index.htm>

5. Nutzerkonferenz „Forschung mit dem Mikrozensus: Analysen zur Sozialstruktur und zum Arbeitsmarkt“

16. – 17. November 2007, Mannheim

Analysen zur Sozialstruktur und zum Arbeitsmarkt

Die regelmäßig von GESIS-ZUMA und dem Statistischen Bundesamt gemeinsam durchgeführten Nutzerkonferenzen wenden sich an ForscherInnen, die mit dem Mikrozensus arbeiten oder daran interessiert sind. Die Ziele der Nutzerkonferenzen sind die Präsentation und Diskussion neuer Forschungsergebnisse und die Förderung des Erfahrungsaustauschs der ForscherInnen untereinander und mit den statistischen Ämtern als Datenproduzenten.

Die diesjährige Nutzerkonferenz steht im Zeichen des fünfzigsten Jubiläums des Mikrozensus, der in Deutschland im Oktober 1957 erstmals durchgeführt wurde. Neben dem enormen Stichprobenumfang besteht eine der großen Stärken des Mikrozensus in der regelmäßigen Durchführung mit weitgehend identischen Fragenprogrammen. Für die empirischen Sozial- und Wirtschaftswissenschaften ist dies insbesondere für Untersuchungen zu Fragen des sozialen und wirtschaftlichen Wandels von Interesse.

Die inhaltlichen Schwerpunkte liegen auf migrations- und arbeitsmarktspezifischen Aspekten. Im Vordergrund stehen zeitvergleichende Analysen und die Potenziale, die der Mikrozensus 2005 mit der Möglichkeit zur Identifikation des Migrationshintergrundes für die Forschung bietet. Daneben werden in den Beiträgen auch spezifische methodische Aspekte bei der Verwendung von Mikrozensusdaten in der Sozial- und Wirtschaftsforschung diskutiert.

Die Konferenz gliedert sich in acht Themenblöcke mit 21 Einzelvorträgen:

- Einführend werden aus Sicht der amtlichen Statistik und der Forschung ein Rückblick auf 50 Jahre Mikrozensus und künftige Entwicklungsmöglichkeiten gegeben.
- Das Thema Migration und Arbeitsmarkt wird in zwei Sitzungen unter soziologischen und ökonomischen sowie unter zeitlich und europäisch vergleichenden Perspektiven diskutiert. Ergänzend werden in zwei weiteren Sitzungen migrations- und familienspezifische Fragestellungen und methodische Aspekte behandelt.
- Mit den im Mikrozensus an zentraler Stelle stehenden Themen „Erwerbstätigkeit und Sozialsystem“ beschäftigt sich eine Sitzung, der Bereich „Familie und Arbeitsmarkt“ ist Gegenstand einer weiteren Sitzung.
- Die Konferenz endet mit einer Abschlussdiskussion und der Gelegenheit, die während der Tagung aufgeworfenen Fragen und Erfahrungen aus der Nutzung der Mikrozensusdaten in der Forschung abschließend zu diskutieren.

Das Konferenzprogramm mit Hinweisen zur Anreise und Übernachtungsmöglichkeiten sowie die Ergebnisse früherer Nutzerkonferenzen sind abrufbar unter:

<http://www.gesis.org/Dauerbeobachtung/GML/Service/Veranstaltungen/Index.htm>

Interessenten werden gebeten, sich unter dem Kennwort „*Konferenz Mikrozensus*“ beim GESIS-ZUMA-Tagungssekretariat anzumelden.

Anmeldeschluss: 15. Oktober 2007

Teilnahmebeitrag: 90 €, Studenten 60 €

GESIS-ZUMA-Tagungssekretariat

Postfach 12 21 55

68072 Mannheim

workshop-mannheim@gesis.org

Tel.: 0621-1246-221

Hinweise für unsere Autorinnen und Autoren

Methoden – Daten – Analysen (MDA) veröffentlicht Beiträge aus dem Bereich der Empirischen Sozialforschung, insbesondere aus dem Bereich der Umfragemethodik. Im Vordergrund stehen Artikel, welche die methodischen und/oder statistischen Kenntnisse der Profession erweitern, sowie Beiträge, die sich mit der Anwendung der Methoden der Empirischen Sozialforschung in der Forschungspraxis beschäftigen, oder solche, in denen ein statistisches Verfahren exemplarisch angewandt wird. Obwohl der Schwerpunkt auf Umfragemethoden liegt, sind Beiträge zu anderen methodischen Bereichen willkommen. Die Artikel sollen für eine breite Leserschaft von Wissenschaftlern und Praktikern im Bereich der Empirischen Sozialforschung verständlich sein.

Manuskripte, die bereits an anderer Stelle veröffentlicht sind oder gleichzeitig anderen Publikationsorganen zur Veröffentlichung angeboten worden sind, werden grundsätzlich nicht berücksichtigt. Eine spätere Veröffentlichung eines in der MDA erschienenen Beitrages ist möglich, sofern an exponierter Stelle auf die Erstveröffentlichung des Beitrages in der MDA hingewiesen wird.

Jeder Beitrag, der zur Veröffentlichung in MDA eingereicht wird, wird zunächst von den Herausgebern danach bewertet, ob er für eine Veröffentlichung grundsätzlich in Frage kommt.

Falls die Herausgeber einer Veröffentlichung grundsätzlich ablehnend gegenüber stehen, werden die Autoren unter Angabe von Gründen für diese Entscheidung informiert.

Falls die Herausgeber zur Ansicht gelangen, dass der Beitrag grundsätzlich zur Veröffentlichung in Frage kommt, wird er anonymisiert an mindestens zwei unabhängige Gutachter verschickt, die um eine Stellungnahme gebeten werden. Im Zweifelsfall wird ein drittes Gutachten eingeholt.

Wird ein Beitrag nach Beschluss der Herausgeber in das Begutachtungsverfahren gegeben, erfolgt die abschließende Entscheidung über ein Manuskript auf der Basis der Gutachten durch die Herausgeber. Im Falle einer Ablehnung erhalten die Autoren eine ausführliche Begründung für die Ablehnung. Wird eine Überarbeitung eines Beitrages für erforderlich gehalten, erhalten die Autoren detaillierte Überarbeitungshinweise.

Unabhängig vom Ergebnis des Begutachtungsverfahrens werden die Autoren von der Entscheidung durch die Redaktion per E-Mail informiert.

Die folgenden Regeln sind bei der Abfassung von Manuskripten zu beachten:

Manuskripte müssen per E-Mail (mda@gesis.org) eingereicht werden. Der Umfang der Manuskripte soll inklusive Leerzeichen alles in allem nicht mehr als 70.000 Zeichen betragen.

Den Beiträgen sind Abstracts in Deutsch und Englisch (jeweils ca. 15 Zeilen) voranzustellen. Auch der Titel des Beitrages ist in Deutsch und Englisch einzureichen.

Um die Anonymität der Beiträge zu wahren, darf in einem Manuskript nur der Titel des Beitrages enthalten sein, nicht aber Namen oder Anschriften der Autoren; Name und Anschrift der Autoren müssen, gemeinsam mit dem Titel des Beitrages, auf einer separaten Seite eingereicht werden.

Beiträge sind mit dem Dezimalklassifikationssystem zu untergliedern (1 - 2 - 2.1 - 2.2 - 3 usw.). Die Gliederungstiefe geht dabei höchstens auf *eine* Stelle nach dem Punkt.

Tabellen enthalten Tabellenummer und Titel im Tabellenkopf, Abbildungen werden analog behandelt.

Grafiken sind mittels gängiger Grafiksoftware zu erstellen. Ist eine spezielle Grafiksoftware erforderlich, übernimmt der Autor/die Autorin die endgültige Formatierung der Grafiken in eigener Regie.

Bei der Erstellung von Tabellen und Grafiken ist zu berücksichtigen, dass der Satzspiegel 13,5 cm (Breite) x 19 cm (Höhe) beträgt. Die Grafiken sind als jpeg- oder tif-Dateien zu liefern; Auflösung mindestens 150 dpi bei Schwarz-Weiss-Grafiken, 300 dpi bei farbigen Grafiken.

Die Beiträge sind unter Wahrung der gültigen Rechtschreibregelungen (neue Rechtschreibung) zu erstellen.

Werden in einem Beitrag empirische Daten verwandt, muss die Möglichkeit der Replikation bestehen. Im Falle einer Veröffentlichung in der MDA erklären sich die Autoren daher schriftlich bereit, Dritten auf deren Anfrage hin die Daten und Programmroutinen zur Verfügung zu stellen.

Anmerkungen und Fußnoten sind mit der Fußnotenfunktion des Schreibprogrammes (im Normalfalle Word) zu erstellen; bitte nicht gesondert formatieren. Fußnoten sind nur für inhaltliche Kommentare vorzusehen, nicht für bibliographische Hinweise.

Literaturhinweise im Text sind nach den folgenden Mustern aufzuführen: Müller (2002) – Müller (2002: 75) – (vgl. Müller 2002: 75) – (Müller 2002; Mayer/Müller/Schulze 2003).

Das Literaturverzeichnis ist wie folgt zu gestalten:

Buchveröffentlichungen:

Strobl, R. und W. Kühnel, 2000: Dazugehörig und ausgegrenzt. Analysen zu Integrationschancen junger Aussiedler. Weinheim/München: Juventa.

Zeitschriftenbeiträge:

Fuchs, M. und M. Sixt, 2007: Zur Nachhaltigkeit von Bildungsaufstiegen. Soziale Vererbung von Bildungserfolgen über mehrere Generationen. Kölner Zeitschrift für Soziologie und Sozialpsychologie 59: 1–29.

Beiträge in Büchern:

Braun, M. und I. Borg, 2004: Berufswerte im zeitlichen und im Ost-West-Vergleich. S. 179–199 in: R. Schmitt-Beck, M. Wasmer und A. Koch (Hg.): Sozialer und politischer Wandel in Deutschland. Analysen mit ALL-BUS-Daten aus zwei Jahrzehnten. Wiesbaden: VS-Verlag für Sozialwissenschaften.

Internetquellen:

Stadtmüller, S. und R. Porst, 2005: Zum Einsatz von Incentives bei postalischen Befragungen. ZUMA How-to-Reihe, Nr. 14 (Mannheim: Zentrum für Umfragen, Methoden und Analysen, ZUMA). http://www.gesis.org/Publikationen/Berichte/ZUMA_How_to/Dokumente/pdf/how-to14rp.pdf (29.06.2007).