

Interviewtechniken bei Erhebungen zu den Finanzen privater Haushalte im Euroraum

Pirmin Fessler,
Peter Mooslechner,
Martin Schürz¹

Das Eurosystem plant, in den nächsten Jahren auf regelmäßiger Basis – wenigstens alle drei Jahre – weitreichende Datenerhebungen zu den Finanzen und Ausgaben privater Haushalte durchzuführen. Der HFCS (Household Finance and Consumption Survey) wird eine der wenigen Erhebungen weltweit sein, die Einkommens-, Ausgaben-, Vermögens- und Verschuldungsdaten gemeinsam erheben. Neben einem angemessenen Fragebogendesign und einer repräsentativen Stichprobenziehung wird besonderes Augenmerk auf die Rolle der Interviewer gelegt. Der vorliegende Beitrag beschäftigt sich mit der Bedeutung der Interviewer bei Mikrodatenerhebungen im Kontext der geplanten HFCS-Erhebungen im Euroraum.

1 Einleitung

Das Eurosystem plant, in den nächsten Jahren auf regelmäßiger Basis – wenigstens alle drei Jahre – weitreichende Datenerhebungen zu den Finanzen und Ausgaben privater Haushalte (im Folgenden Haushalte genannt) durchzuführen. Der HFCS (Household Finance and Consumption Survey) wird eine der wenigen Erhebungen weltweit sein, die Einkommens-, Ausgaben-, Vermögens- und Verschuldungsdaten gemeinsam erheben. Damit wird eine Vielzahl ökonomischer Forschungsfragen besser zu analysieren sein. Die Oesterreichische Nationalbank (OeNB) plant die Realisierung der österreichischen Haushaltsbefragung für das Frühjahr 2010. Die vorliegende Analyse ist der dritte Beitrag einer Artikelserie zu methodischen und statistischen Fragen im Zusammenhang mit Erhebungen zu den Finanzen und dem Konsum privater Haushalte im Euroraum.²

Eine ungenügende Datenqualität bei Haushaltserhebungen kann sich unter anderem aus einem schlechten Fragebogendesign, einer nicht ausreichenden Dokumentation zu den einzelnen Fragen oder aus einer fehlenden Bereitschaft mancher Haushalte, an der

Befragung teilzunehmen, ergeben. Bei Haushaltsbefragungen nehmen Interviewer eine besonders wichtige Stellung ein. Interviewer sind in zweierlei Hinsicht bedeutsam: erstens bei der Stichprobenziehung und zweitens in Bezug auf die Kooperationsbereitschaft der Befragten. Sie können die Antwortbereitschaft der Befragten erhöhen und den Frage-Antwort-Prozess strukturieren. Sie zeichnen die Antworten auf und/oder geben diese in ein Computerprogramm ein. Die Interviewer prüfen die Ergebnisse und übermitteln diese dann an die Projektleitung des Meinungsforschungsinstituts.

Diese Vielzahl an Aufgaben birgt Fehlermöglichkeiten. Es kann einerseits zu systematischen Fehlern kommen; so können Messfehler, die Differenz zwischen dem erhobenen Wert einer Variablen und dem wahren, aber unbekanntem Wert der Variablen in der Grundgesamtheit, entstehen. Andererseits können auch nichtsystematische Fehler entstehen, die sich in einer größeren Varianz der Schätzer in Bezug auf mögliche Replikationen widerspiegeln. Dies beeinflusst die Replizierbarkeit von Ergebnissen negativ. Das Verhalten der Interviewer kann grund-

¹ Pirmin.Fessler@oebn.at, Peter.Mooslechner@oebn.at, Martin.Schuerz@oebn.at. Die Autoren danken Karin Wagner und Siegfried Zottel für wertvolle Kommentare.

² Die im vorliegenden Beitrag verwendeten Fachbegriffe wurden in den ersten beiden Studien der Serie besprochen; siehe „Statistische Herausforderungen der Forschung zu Finanzen privater Haushalte im Euroraum“ (Heft Q1/09) und „Stichprobenziehung bei Erhebungen zu den Finanzen privater Haushalte im Euroraum“ (Heft Q2/09).

sätzlich den *Coverage Error*, den *Sampling Bias*, den *Non-Response-Error* und den *Measurement Error* beeinflussen.

Die wichtige Stellung der Interviewer im Prozess der Haushaltsbefragung wird noch unterstrichen durch die spezielle Sensitivität des Themas der Haushaltsfinanzen (Kennickell, 2005; 2006; 2007). Fragen zum privaten Vermögen von Haushalten weisen einige Spezifika auf: Sie werden ungern beantwortet. Die Diskretion der Bevölkerung hinsichtlich Angaben zum privaten Vermögensbesitz ist in vielen Ländern – und Österreich zählt vermutlich zu dieser Gruppe – weitreichend. Sie erfordern teilweise einen hohen Wissensstand. Es wird oft eine Reihe von Unterlagen (Verträge usw.) benötigt. Da es bei Erhebungen zu Haushaltsfinanzen besonders um genaue Zahlenangaben zu Ausgaben sowie zum Einkommen und Vermögen geht, ist eine korrekte Aufzeichnung der angegebenen Werte durch die Interviewer eine Schlüsselaufgabe. So sind beispielsweise die Fehlermöglichkeiten bei den Angaben zur Einkommensvariablen zahlreich: nicht vorhandene Angaben zu bestimmten Einkommenskomponenten (z. B. Vermögenseinkommen) oder zur Dauer des Bezugs aus einer Einkommensquelle. Im Extremfall fehlt nicht nur eine Angabe über die Höhe des bezogenen Einkommens, sondern es erfolgt auch keine Zuordnung zu den vorgegebenen Einkommensintervallen.

Kapitel 2 beschäftigt sich mit dem Thema standardisierter Interviews. In Kapitel 3 wird auf die Rolle des Interviewers im Allgemeinen sowie auf mögliche Fehlerquellen, Strategien zu deren Vermeidung und auf das Interviewmanagement im Speziellen eingegangen. Kapitel 4 beschäftigt sich mit der Paradataerhebung im HFCS und Kapitel 5 mit einigen Erfahrungen aus der OeNB-Immobilienvermögenserhe-

bung 2008. In Kapitel 6 werden Schlussfolgerungen gezogen.

2 Standardisierte Interviews

Um eine konsistente Messung der abgefragten Variablen zu gewährleisten, verwenden sehr viele Erhebungen sogenannte standardisierte Interviews. Der Fragetext einschließlich der Antwortmöglichkeiten und die Reihenfolge der Fragen sind vorgegeben. Dabei sollen die Befragten dieselben Fragen unter möglichst konsistenten Bedingungen beantworten, damit diese einheitlich interpretiert werden können. Ziel ist, diese Bedingungen so effektiv zu gestalten und zu dokumentieren, dass eine Replizierbarkeit prinzipiell gewährleistet ist.

Dabei spielt die Wahl der Erhebungsmethode eine bedeutende Rolle. Prinzipiell steht eine Vielzahl verschiedener Erhebungsmethoden zur Verfügung, die sich hinsichtlich der Kontrollmöglichkeiten und der damit verbundenen Kosten unterscheiden. Vereinfacht gilt, je besser die Kontrollmöglichkeiten, desto höher die Kosten.

1. CAPI (Computer-Assisted Personal Interviewing). Eine computergestützte, persönliche Interviewerbefragung stellt das Instrument mit den höchsten Kontrollmöglichkeiten dar und ist auch mit den höchsten Kosten verbunden. Dabei liest der Interviewer die Fragen von einem Computerbildschirm vor und gibt dann die Antworten direkt in den PC ein. Das hat den Vorteil, dass einerseits das Fragenprogramm unterschiedliche Verläufe nehmen kann, der Interviewer jederzeit eingreifen und Erklärungen geben kann und dabei auf gespeicherte Informationen zurückgreifen kann. Zudem können durch die CAPI-Software anhand von Bezügen zwischen ver-

schiedenen Antworten bereits während des Interviews laufend Plausibilitätschecks durchgeführt werden. Aus den eingegebenen Daten entsteht direkt der Datensatz.

2. CATI (Computer-Assisted Telephone Interviewing) ist das Pendant zu CAPI. Der Interviewer besucht aber den Respondenten nicht persönlich, sondern das Interview wird telefonisch geführt.
3. ACASI (Audio Computer-Assisted Self-Interviewing). Der Respondent gibt die Antworten direkt in eine Computermaske ein. Auch Webumfragen können auf diese Art gestaltet sein.
4. PAPI (Paper And Pencil Interviewing). Der Respondent füllt einen Fragebogen auf Papier aus.

In sehr vielen Fällen werden bei Interviews CATI- oder Webumfragen durchgeführt; auch PAPI ist häufig noch im Einsatz. Bei einer wissenschaftlichen Erhebung im Umfang des HFCS ist ein CAPI-Ansatz aber angezeigt, da es sich um einen sehr ausführ-

lichen und sehr komplexen Fragebogen mit sehr vielen unterschiedlichen Frageprogrammverläufen handelt. Den Interviewern kommt dabei eine besondere Rolle zu, da sie bei Unsicherheiten, vermuteten Unklarheiten oder erkennbaren Missverständnissen eingreifen müssen, aber gleichzeitig den Befragten nicht in seiner Beantwortung beeinflussen dürfen. Um die Anzahl dieser notwendigen Eingriffe möglichst klein zu halten, und vorab ein möglicherweise differierendes Interviewerverhalten als Fehlerquelle gering zu halten, wird eine Reihe von Informationen und Verhaltensregeln vorgegeben und Fragen derart gestellt, dass es möglichst nicht zu Rückfragen kommt, die Erklärungsbedarf des Interviewers erzeugen. Auf die Rolle des Interviewers wird in Kapitel 3 detaillierter eingegangen.

Bei der Einleitung zur Umfrage müssen bereits eine Reihe von Informationen gegeben und/oder den Befragten diverse Materialien zur Verfügung gestellt werden. Die Befragten benötigen jedenfalls folgende Informationen (Kasten 1).

Kasten 1

Seitens der zur Befragung ausgewählten Haushalte kann es eine Reihe von Bedenken geben, die gegen eine Teilnahme sprechen:

- Warum ist die Umfrage wichtig?
- Warum soll der Haushalt an der Umfrage teilnehmen?
- Von wem wurde die Umfrage beauftragt?
- Wie wurden Name und Anschrift des Umfrageteilnehmers ermittelt?
- Warum kann an der Umfrage anstelle des Befragten nicht ein anderer Haushalt/ein Haushalt, der nicht über ein statistisches Verfahren ermittelt wurde, teilnehmen?
- Welche Institutionen werden die Umfragedaten verwenden?
- Welche Themen werden in der Umfrage abgedeckt?
- Wie lange wird die Befragung ungefähr dauern?
- Welche Hilfestellungen für Befragte sind vorgesehen?

Drei Argumente sind vonseiten der Interviewer anzuführen: Den befragten Haushalten muss verdeutlicht werden, dass sie viele andere Haushalte repräsentieren und daher ihre Auskünfte besonders wichtig für geldpolitische Entscheidungen sind. Die nationale und die europäische Bedeutung der Erhebung soll herausgestrichen werden. Unerlässlich ist die Zusicherung der Vertraulichkeit.

Bei der Beantwortung der Fragen können gegebenenfalls auch Unterlagen weiterhelfen, die von Banken, Versicherungen, vom Arbeitgeber oder von anderen Stellen stammen. In der Umfrage geht es in erster Linie um die finanzielle Lage der Haushalte. Viele Fragen verlangen nach einer quantitativen Beantwortung, das heißt nach einer exakten Angabe eines Betrags in Euro. Wenn dies nicht möglich sein sollte, weil die Interviewten den Betrag nicht genau kennen oder angeben wollen, kann auch eine Bandbreite für den Betrag genannt werden.

Trotz der Standardisierung und laufender Plausibilitätschecks des CAPI-Programms bleiben manche Sachverhalte unklar und bedürfen einer wiederholten Kontaktierung des Respondenten oder sogar eines erneuten Besuchs des Interviewers. Im Fall der Umfrage der Banco de España wurden rund 67% der Respondenten erneut kontaktiert (Bover, 2004). Bei den meisten erfolgte dies telefonisch, manche wurden jedoch auch erneut besucht.

2.1 Anreize

Anreize spielen sowohl für den Respondenten als auch für den Interviewer eine große Rolle im Erhebungsprozess. Was den Respondenten betrifft, so wird einerseits durch Vorabinformation mittels eines Schreibens des Zentralbankgouverneurs versucht, die Relevanz der Teilnahme an der Umfrage zu erklären und Vertrauen aufzubauen. Studien zeigen, dass sich derartige Ankündigungsschreiben positiv auf die Kontakt- und Kooperationsbereitschaft auswirken (Kennickell, 2007); das verringert das Non-Response-Verhalten. Andererseits werden weitere Anreize für die Teilnahme geboten. Bei der OeNB-Immobilienvermögenserhebung 2008 war eine freiwillige Gewinnspielteilnahme möglich. Die Deutsche Bundesbank setzte bei ihrem Pre-Test des HFCS einen monetären Anreiz in Form einer Goldmünze im Wert von 10 EUR.

Neben Anreizen für die Respondenten muss auch auf die Anreizstruktur

für die Interviewer Rücksicht genommen werden. Ein Fehlanreiz ergibt sich, wenn Interviewer nur daran interessiert sind, vollständig durchgeführte Interviews abzugeben. Dann werden sie sich um jene Fälle mehr bemühen, die ihnen Erfolg versprechend und jene vermeiden, die schwierig scheinen. Das kann durch Vorgaben des Meinungsforschungsinstituts leicht verhindert werden. Zudem muss die Bezahlungsstruktur so gestaltet werden, dass Interviews möglichst entsprechend dem Aufwand der Interviewer abgegolten werden und nicht nur nach abgeschlossenen Fällen bezahlt wird. Das würde dazu führen, dass potenziell kürzere Fälle (z. B. Mieter) bevorzugt würden.

3 Rolle des Interviewers³

Interviewer sind im Allgemeinen nicht neutral im Prozess der Datenerhebung. Während früher die Evaluierung der Interviewertätigkeit nur im Hinblick auf ausgefüllte Fragebögen erfolgte, wird nun dem Einfluss der Interviewer auf die Befragungsergebnisse mehr Augenmerk geschenkt.

Erfahrene Interviewer weisen generell bessere Antwortraten auf als weniger erfahrene Kollegen. Das gilt sowohl für *Unit-* als auch für *Item-Non-Response*. Allerdings dürften erfahrene Interviewer weniger sorgfältig beim Vorlesen von Fragestellungen und bei der Einhaltung des Protokolls sein.

Eine aktuelle Untersuchung zum Einfluss der Interviewer auf die Befra-

³ Einen detaillierten Überblick zu diesem Thema, der auch eine maßgebliche Grundlage für diesen Beitrag darstellt, liefern unter anderem Groves et al. (2004).

gungsergebnisse zeigt, dass viele Attribute der Interviewer wichtig sind. Im Rahmen des Sozioökonomischen Panels (SOEP) wurde eine Befragung aller aktuell 586 eingesetzten Interviewer durchgeführt (94% Antwortrate). Die Befragten gaben Selbsteinschätzungen zu Werten sowie zu Einstellungen wie Kreativität, Gewissenhaftigkeit, Geselligkeit, Verträglichkeit und Neurotizismus ab (Weinhardt et al., 2009). In einer logistischen Regression zeigte sich, dass Offenheit die Chancen auf Rundungsfehler signifikant erhöhte. Spezifische Anforderungen an Interviewer sind eine ausgeprägte Kontaktfähigkeit und Einfühlungsvermögen. Ein freundliches Auftreten und eine entspannte Durchführung des Interviews sind besonders wichtig.

Bei Umfragen zu Haushaltsfinanzen kommen im Allgemeinen Interviewermanuals zum Einsatz. Auch den Interviewern im Rahmen des HFCS wird ein Interviewermanual zur Verfügung gestellt werden.

3.1 Interviewerfehler

Der Kommunikationsprozess zwischen dem Interviewer, der die Befragung durchführt, und der Zielperson einer Befragung ist mitentscheidend für die Qualität der resultierenden Daten. Gründe für systematische Verzerrungen, die durch den Einsatz von Interviewern hervorgerufen werden können, sind:

- *Persönlichkeits-Bias*: Die Zielperson wird durch die Persönlichkeit der Interviewer beeinflusst (Aussehen, Stimme usw.). Allein die Anwesenheit einer Person kann dazu führen, dass die Befragten eher „sozial erwünschte“ Antworten geben. Auch mangelnde Auskunftsbereitschaft kann zu unbefriedigenden Ergebnissen führen. Die Zielperson

will möglichst schnell aus der Stresssituation der Befragung herauskommen und beantwortet dann Fragen, ohne wirklich mental involviert zu sein.

- *Verhaltens-Bias*: Die Zielperson wird durch das Verhalten der Interviewer beeinflusst, z. B. durch Bemerkungen des Interviewers oder durch Gestik und Mimik bei persönlichen Befragungen.
- *Selektive Wahrnehmung*: Jeder Mensch nimmt seine Umwelt und Äußerungen von anderen Menschen in spezifischer Weise wahr. Bei nicht eindeutigen Formulierungen werden eigene Interpretationen herangezogen. So können z. B. bei offenen Fragen die gleichen Antworten der Befragten von verschiedenen Interviewern anders verstanden und notiert werden. Die selektive Wahrnehmung stellt meist eine unbewusste Tendenz des Interviewers dar, die gegebenen Antworten in Richtung seiner eigenen Erwartungen „verfälscht“ wahrzunehmen.
- *Erfahrungsbias*: Bezeichnet die unterschiedliche Beantwortung, die sich durch die Erfahrung des Interviewers ergibt.

Messungen des Interviewerfehlers anhand von Vergleichen mit PAPI- oder CATI-Ansätzen sind relativ einfach durchzuführen. Bei komplexen Umfragen wird der Vergleich allerdings schwierig, weil angenommen werden muss, dass einige Fragen ohne die Hilfe des Interviewers im CATI- oder PAPI-Fall wohl nicht vollständig von allen Respondenten verstanden werden. Neben diesen systematischen Fehlern, die durch den Einsatz von Interviewern zustande kommen können, können auch nichtsystematische Fehler auftreten, die die Varianz der Schätzer vergrößern.

3.2 Interviewervarianz

Interviewervarianz bezeichnet jenen Teil der Variabilität der Schätzungen auf Basis der Befragungen, der sich aufgrund des Einsatzes unterschiedlicher Interviewer ergibt. Es handelt sich um nichtsystematische Fehler, die keinen Einfluss auf den Erwartungswert (bei einer ausreichenden Anzahl von Replizierungen), wohl aber auf die Variabilität der Schätzer haben. Eine Erhebung wird nach dieser Definition als eine mögliche Realisierung einer Vielzahl an möglichen Erhebungen betrachtet. Wenn verschiedene Interviewer unterschiedliche Antworten hervorrufen, führt das dazu, dass bei Replizierungen auch unterschiedliche Schätzer erzeugt würden. Die Variabilität dieser Schätzer ist die Interviewervarianz.

Ähnlich wie bei einer Cluster-Bildung im Sample Design wird beim Einsatz desselben Interviewers für ein zusätzliches Interview weniger an neuen unabhängigen Informationen generiert als beim Einsatz eines anderen Interviewers. Bei einer Cluster-Bildung im Sample Design ist das deshalb der Fall, weil davon ausgegangen wird, dass die Beobachtungen innerhalb eines Clusters homogener sind als die Grundgesamtheit und daher eine zusätzliche Beobachtung außerhalb des Clusters mehr neue Information über die Grundgesamtheit enthält. Beim Interviewereinsatz kann analog argumentiert werden: Da Interviewer die Respondenten beeinflussen, beinhaltet ein zusätzliches Interview von einem gleichwertigen anderen Interviewer oder einem Interviewer, der die Antworten gar nicht beeinflussen würde, mehr neue Information über die Grundgesamtheit als ein weiteres Interview vom selben Interviewer. Es existiert also ein Zielkonflikt zwischen Erfahrung der Interviewer und geringer Interviewervarianz.

Die Messung der Interviewervarianz gestaltet sich schwierig. Angenommen, es werden zwei Interviewer für zwei verschiedene Stadtteile eingesetzt und es kommt zu unterschiedlichen Antworten. Das kann aufgrund tatsächlicher durchschnittlicher Unterschiede in den Attributen der Respondenten zustande kommen oder aufgrund einer möglichen Beeinflussung durch die unterschiedlichen Attribute der Interviewer. Notwendig wäre eine zufällige Zuweisung der Respondenten zu einem Interviewer. Diese ist in den meisten Befragungen aus Kostengründen nicht möglich, da es zu enormen Wegzeiten der Interviewer kommen würde. Ein weiterer negativer Effekt einer zufälligen Zuordnung kann aus anderen Gründen entstehen. So kann sich etwa das Non-Response-Problem vergrößern, wenn Interviewer in einer anderen Region eingesetzt werden als in jener, in der sie selbst leben (wenn z. B. Interviewer aus Wien Respondenten in Vorarlberg befragen würden).

Das einfachste Modell zur Messung der Interviewervarianz ist jenes von Kish (1962):

$$y_{ij} = \phi_i + \underbrace{\varphi_j + \varepsilon_{ij}}_{u_{ij}}$$

Wobei y_{ij} den gemessenen und ϕ_i den wahren Wert darstellen. Der gesamte Fehler u_{ij} teilt sich in den systematischen Fehler, der durch den Interviewer zustande kommt (φ_j) und einen zufälligen Fehler des Respondenten, wenn dieser von diesem Interviewer interviewt wird, auf. Die Subskripte beziehen sich auf den Respondenten (i) bzw. auf den Interviewer (j). Der Erwartungswert von φ_j über alle Interviewer j sollte in Bezug auf eine Messung der Interviewervarianz null sein. Der Fokus liegt auf der Variabilität der entstehenden Schätzer bei möglichen Replikationen. Das Modell kann auch

zur Messung des Interviewerfehlers herangezogen werden, der sich dann in einem Erwartungswert von φ_j ungleich null ausdrücken sollte. Die Interviewervarianz vergrößert sich, je stärker φ_j über die Interviewer variiert. Um dies zu messen, wird die Korrelation innerhalb der von einem Interviewer durchgeführten Fälle errechnet. Diese wird als ρ bezeichnet.⁴ Der Vorteil bei Kish's Ansatz ist, dass dazu einfache Schätzer aus der Varianzanalyse verwendet werden können. ρ liegt immer im geschlossenen Intervall $[-(1/(m-1)), 1]$ wobei m die Anzahl der durchgeführten Interviews pro Interviewer darstellt. Fälle mit wenigen Interviewern oder Interviews führen zu Ergebnissen knapp kleiner null. Werte nahe null bedeuten, dass Antworten nicht stark von den Interviewern beeinflusst werden, das heißt, dass verschiedene Interviewer innerhalb der Variabilität aus der Stichprobe die gleichen Antworten bekommen. Höhere Werte weisen auf einen stärkeren Einfluss der Interviewer hin. Die Stärke des Einflusses, das heißt jener Anteil der Varianz, der als Interviewervarianz bezeichnet werden kann, ergibt sich als $Var_{Interviewer} = 1 + \rho(m-1)$. Eine derartige Vorgehensweise ist aber nur dann zulässig, wenn die Respondenten den Interviewern per Zufallsverfahren zugewiesen wurden, sodass der Effekt des einzelnen Interviewers von dem Effekt der Attribute seiner Respondenten unterschieden werden kann.

Das Modell zeigt deutlich, dass die durchschnittliche Anzahl an Interviews pro Interviewer starken Einfluss auf die Interviewervarianz hat. Je mehr Inter-

views pro Interviewer, desto stärker wird $Var_{Interviewer}$ über die Multiplikation von ρ mit m erhöht. Auch hier besteht wieder ein Trade-off zwischen der gesammelten Erfahrung der Interviewer während der Feldphase und der Interviewervarianz. Von Frage zu Frage wird sich der Einfluss der Interviewer und damit die Interviewervarianz unterscheiden. Je mehr Fragen im Fragebogenprogramm einen Eingriff des Interviewers erfordern (Erklärungen, Nachfragen usw.), desto höher wird die Interviewervarianz ausfallen. Zudem scheint es, dass auch bestimmte Respondentengruppen (z. B. ältere Personen) anfälliger auf eine Beeinflussung durch den Interviewer sind.

3.3 Strategien zur Vermeidung von Interviewerfehlern und Interviewervarianz

Strategien zur Vermeidung von Interviewerfehlern können, wenn sie auf alle Interviewer greifen, auch dazu beitragen, die Interviewervarianz zu verringern. Dabei spielt vor allem die Motivation eine große Rolle. Ob der Interviewer dem Respondenten vermitteln kann, dass er präzise Antworten benötigt und ihn auch dazu motivieren kann, diese zu geben, wird das Ergebnis stark beeinflussen.

Um die Interviewervarianz zu verringern, muss ein konsistentes Verhalten der Interviewer angestrebt werden. Dabei sind grundsätzlich vier Kanäle zu erwähnen:

1. Ein möglichst gleiches, neutrales und zielorientiertes Verhalten im Umgang mit allen Respondenten ist anzustreben. Die Prozeduren der

⁴ $\rho = [(V_b - V_w)/m] / [(V_b - V_w)/m + V_w]$

V_b stellt dabei den „between mean square“ in einer einfaktoriellen Varianzanalyse mit dem Interviewer als Faktor dar.

V_w stellt den „within mean square“ in einer Varianzanalyse dar.

m ist die Anzahl der durchgeführten Interviews pro Interviewer.

Umfrage und ihre Ziele sind klar und verständlich darzustellen.

2. Die Auswahl der Instrumente (Laptop, Antwortkarten, vorgefertigte Erklärungen und Routinen), die den Interviewern zur Verfügung stehen, sowie das Training und Management der Interviewer.
3. Die Fragen müssen so gewählt werden, dass sie zu möglichst wenigen Situationen führen, in denen der Interviewer eingreifen muss. Insbesondere das Nachfragen bei inadäquaten Antworten sollte vorzugsweise vermieden werden. Sollte es in einzelnen Situationen aber unumgänglich sein, ist eine einheitliche und neutrale Vorgangsweise anzustreben. Ein wortgenaues Vorlesen der Fragen, eine genaue Einhaltung der Fragenreihenfolge und eine wörtliche Aufnahme von Antworten auf offene Fragen sind Vorbedingungen.
4. Die Kontrolle der Interviewertätigkeit soll sicherstellen, dass durch nicht korrekt abgewickelte Interviews keine Verfälschung der Untersuchungsergebnisse erfolgt. Die Methode und die Häufigkeit der Kontrolle müssen dokumentiert werden.

Ein Interview ist kein *normales* Zwiegespräch, sondern erfordert Neutralität des Interviewers. Das bedeutet unter anderem, dass Meinungsäußerungen seitens der Interviewer unterbleiben müssen (z. B. „Da bin ich nicht Ihrer Meinung“ oder „Da stimme ich Ihnen zu“). Die Befragten dürfen in ihren Antworten nicht beeinflusst werden, und die Antworten müssen mit Nachdruck eingefordert werden. Insbesondere müssen Interviewer darauf achten, dass sie bei notwendigem Nachfragen im Fall inadäquater Antworten der Respondenten diese nicht in eine Antwortrichtung beeinflussen oder ihnen bestimmte Antwortkategorien

nahelegen. Wird etwa nach der Anzahl der Erbschaften gefragt und ein Respondent antwortet mit „einige“, muss seitens der Interviewer nachgefragt werden. Unpassend wäre die Nachfrage „Etwa drei Mal?“, eine neutrale Nachfrage wäre z. B. „Wie oft genau?“

Zusammenfassend ist das Ziel, ein konsistentes Verhalten der Interviewer in Bezug auf ihre unterschiedlichen Respondenten zu erreichen.

3.4 Interviewermanagement

Unter Interviewermanagement werden im Allgemeinen die Interviewerselection, das Interviewertraining, die Interviewersupervision und die Zuweisung eines Stichprobenanteils an die Interviewer verstanden.

Betreffend die *Interviewerselection* gibt es relativ wenig Literatur bezüglich deren Einfluss auf die Qualität der Ergebnisse. Es scheint jedoch unbestritten, dass die Qualität der Interviewer eine große Rolle spielt, weshalb auch auf den Selektionsprozess bei der Interviewerauswahl geachtet werden muss. Bei Interviewern handelt es sich meist um Teilzeitbeschäftigte, die abends und an den Wochenenden arbeiten müssen, da die meisten Befragten zu diesen Zeiten am ehesten im Haushalt anzutreffen sind. Artikulationsfähigkeit ist von großer Bedeutung. Zusätzlich bedarf es beim Einsatz von CAPI auch Grundkenntnissen der EDV und bei der Thematik des HFCS der Fähigkeit, die wissenschaftlichen Ziele der Umfrage und das Vertrauen in die Anonymität der Erhebung gut vermitteln zu können. Interviewer mit Erfahrung in ähnlichen wissenschaftlichen Projekten – wie EU-SILC (Statistics on Income and Living Conditions) – sind grundsätzlich zu bevorzugen. Auch eine gewisse Achtsamkeit auf Details wird bei einer derartigen umfassenden

Erhebung wichtig sein. Insgesamt stehen die Kommunikationsfähigkeit und die Gewissenhaftigkeit in Bezug auf die Vorgaben der Durchführung im Vordergrund.

Das *Interviewertraining* ist von besonderer Relevanz, wenn es darum geht, möglichst konsistentes Interviewerverhalten zu generieren. Dabei zeigt sich, dass ein Interviewertraining von mindestens einem Tag unerlässlich ist, damit sich neue Interviewer in der Interviewsituation in zufriedenstellender Weise gemäß den Vorgaben verhalten (Bover, 2004; Kennickell, 2006). Die Trainingsprogramme beinhalten im Allgemeinen eine Einführung in die Ziele der Erhebung, richtige Fragestellung, richtiges Nachfragen auf inadäquate Antworten, richtiges Aufzeichnen der Antworten, insbesondere bei offenen Fragen und allgemeine administrative Abläufe sowie die richtige Bedienung der CAPI-Software. Zusätzlich können praktische Interviews unter Aufsicht geübt und besprochen werden sowie Strategien zu einer effizienten Abarbeitung der zugewiesenen Adressen erarbeitet werden. Insbesondere richtiges Nachfragen bei inadäquaten Antworten erfordert gutes Training und Übung.

Die *Interviewersupervision* ist vor allem während der Feldphase von großer Bedeutung. Dank der Möglichkeiten von CAPI können schon während des Interviews viele Plausibilitätschecks durchgeführt und Inkonsistenzen vermieden werden. Zudem können durch eine schnelle Übertragung von fertiggestellten Interviews an das Umfrageinstitut Unregelmäßigkeiten frühzeitig entdeckt werden. Werden bei einem Interviewer überproportional viele fehlende Angaben entdeckt, kann nachgeforscht werden, ob es möglicherweise mit seinem Verhalten zusammenhängt. Weiters kann durch laufende

stichprobenartige Kontrollen festgestellt werden, ob Interviews tatsächlich stattgefunden haben bzw. ob bestimmte Vorgaben eingehalten wurden.

Die *Zuweisung eines Stichprobenanteils* an die Interviewer ist für die Interviewervarianz von großer Bedeutung. Je größer der Anteil, desto höher die mögliche Interviewervarianz. Unterscheiden sich allerdings die Interviewer hinsichtlich ihrer Interviewqualität – korreliert möglicherweise positiv mit Erfahrung – kann es Sinn machen, weniger Interviewer einzusetzen und ihnen größere Anteile zuzuweisen.

Eine hohe Motivation der Interviewer ist jedenfalls das Fundament für eine erfolgreiche Erhebung. Der Interviewer sollte sich als Teil des gesamten wissenschaftlichen Projekts verstehen und sich bewusst sein, dass er maßgeblich für die Qualität der daraus resultierenden Daten und damit auch für deren Anwendbarkeit in den Analysen mitverantwortlich ist.

4 Paradata

Das Verhalten von Interviewern in den konkreten Gesprächssituationen kann nicht direkt kontrolliert werden. Technisch wäre es zwar denkbar, dass die Interviews aufgezeichnet werden, aber dies würde die Antwortbereitschaft senken (fehlende Wahrung der Privatsphäre) und den administrativen Aufwand extrem erhöhen.

Die Fragen an die Interviewer (nach dem realisierten Interview) dienen als Indikator für die Zuverlässigkeit der Angaben des Befragten. Im Fall des HFCS werden Einschätzungen hinsichtlich der wahrgenommenen Richtigkeit der Einkommens- und Vermögensangaben, der Fähigkeit Fragen zu verstehen bzw. Euro-Beträge richtig anzugeben, des Ausdrucksvermögens, des generellen Interesses an der Haushaltsbefragung und der Zögerlichkeit im Ant-

wortverhalten abgefragt. Weitere Fragen beziehen sich auf die Innenausstattung der Immobilie des Befragten und auf die Verwendung von Unterlagen.

Zudem wird in einer offenen Frage auf spezifische Verdachtsmomente der Interviewer geachtet. Gab es vermutlich falsche Angaben und/oder Wiederholungen? Wie stark war das Interesse an den Inhalten der Erhebung? War der Respondent vor oder nach dem Interview misstrauisch? Waren andere Personen anwesend?

Grundsätzlich gilt, Paradata helfen entscheidend bei der Imputation und der Auswertung der Datensätze, insbesondere auch bei der Forschung zur Erhebungsmethode selbst.

5 Erfahrungen mit Interviewern im Rahmen der Immobilienvermögenserhebung

Im Rahmen der OeNB-Immobilienvermögenserhebung 2008 wurde auch den Wahrnehmungen der Interviewer besonderes Augenmerk geschenkt. Im Einsatz waren bundesweit 98 Interviewer (Wagner und Zottel, 2009).

Unter anderem gab es Fragen an die Interviewer zur Einschätzung der Immobilie, der Wohngegend und der Interviewsituation. Befragten aus bildungsfernen Schichten wurde von den Interviewern in deutlich geringerem Ausmaß vertraut. Dies mag damit zusammenhängen, dass Unterschiede im sprachlichen Ausdrucksvermögen zum Interviewer auftreten, dass sie möglicherweise ein direkteres Auftreten haben oder sich weniger überzeugend präsentieren als besser gebildete Interviewte. Weibliche Interviewer misstrauten häufiger – die Gründe sind unklar – als ihre männlichen Kollegen.

Die Einschätzung der Wohngegend eignet sich für Unit-Non-Response-Analysen, da diese Informationen für

befragte und auch für verweigernde Haushalte vorhanden sind. Ergeben sich schwierige Situationen bei Befragungen, sollten die Interviewer grundsätzlich das Ziel haben, diese zu vermeiden bzw. einzugrenzen. Daher hat die Fed ein striktes Feedback-System für die Interviewer aufgebaut. Die Daten werden im Prozess der Datenerhebung geprüft, und das Feedback wird direkt an die Interviewer gegeben. Arthur Kennickell, Projektdirektor des Survey of Consumer Finances (SCF) der Fed und Konsulent der EZB beim Projekt der zukünftigen euroraumweiten Haushaltsbefragungen, betont stets die maßgebliche Bedeutung des Interviewertrainings und auch die EZB folgt diesen erfahrungsbasierten Argumenten. Im Rahmen des HFCS werden die Notenbanken intensiv in das Interviewertraining eingebunden sein.

6 Schlussfolgerungen

Interviewer nehmen eine wichtige Position bei Mikrodatenerhebungen ein. Besonders bei komplexen Erhebungen bedarf es eines persönlichen Interviews, um gegebenenfalls intervenieren zu können, nachzufragen und Erklärungen zu geben. Durch dieses Eingreifen kann allerdings ein Interviewerfehler entstehen. Je nach Design der Erhebung kann die Haltung des Interviewers über unterschiedliche Kanäle die Erhebungsqualität stark beeinflussen. Das Verhalten der Interviewer kann grundsätzlich Auswirkungen auf den *Coverage Error*, den *Sampling Bias*, den *Non-Response-Error* und den *Measurement Error* haben. Besonders wichtig sind Interviewer hinsichtlich der Kooperationsbereitschaft der Befragten. Sie können die Antwortbereitschaft der Befragten erhöhen.

Verzerrungen entstehen jedoch nicht nur durch Interviewer, sondern auch durch die Befragten. So erfolgen unzutreffende Antworten z. B. bei Angaben

zur Vermögenshöhe. Hier werden falsche Antworten gegeben, etwa aus Scham bei armen Menschen bzw. aus Angst vor Neid bei reichen Menschen. Freiwillige Haushaltsbefragungen haben analytische Grenzen, die bei der Datenverwendung in der wissenschaftlichen Arbeit stets einzubeziehen sind.

Literaturverzeichnis

- Bover, O. 2004.** The Spanish Survey of Household Finances (EFF). Description and Methods of the 2002 Wave. Documentos Ocasionales 409. Banco de España.
- Deaton, A. 1985.** Panel Data from Time Series of Cross Sections. In: Journal of Econometrics 30. 109–126.
- DIW. 1998.** Auswertung von Statistiken über die Vermögensverteilung in Deutschland.
- Durrant, G. B., R. Groves, L. Staetsky und F. Steele. 2005.** Effects of Interviewer Attitudes and Behaviors on Refusal in Household Surveys. WP M09/05. University of Southampton.
- EZB. 2006.** Household Finance and Consumption Network. Euro Area Questionnaire on Household Finance and Consumption. www.ecb.int/home/html/researcher_hfcn.en.html
- Fessler, P., P. Mooslechner und M. Schürz. 2009.** Statistische Herausforderungen der Forschung zu Finanzen privater Haushalte im Euroraum. In: Statistiken Q1/09. Wien: OeNB. 57–66.
- Fessler, P., P. Mooslechner und M. Schürz. 2009.** Stichprobenziehung bei Erhebungen zu den Finanzen privater Haushalte im Euroraum. In: Statistiken Q2/09. Wien: OeNB. 51–62.
- Groves, R., F. J. Fowler, M. P. Couper, J. M. Leprowski, E. Singer und R. Tourangeau. 2004.** Survey Methodology. Wiley Publishing. New Jersey.
- Kennickell, A. B. 1998.** Multiple Imputation in the Survey of Consumer Finances. Vortrag beim Joint Statistical Meeting. Dallas, TX.
www.federalreserve.gov/Pubs/oss/oss2/papers/impute98.pdf
- Kennickell, A. B. 2000.** Wealth Measurement in the Survey of Consumer Finances: Methodology and Directions for Future Research.
www.federalreserve.gov/pubs/oss/oss2/papers/measurement.pdf
- Kennickell, A. B. 2005.** The Good Shepherd: Sample Design and Control for Wealth Measurement in the Survey of Consumer Finances. Präsentation anlässlich der Luxembourg Wealth Study Conference. Perugia. Jänner.
www.federalreserve.gov/Pubs/oss/oss2/papers/sampling.perugia05.2.pdf
- Kennickell, A. B. 2006** Who's Asking? Interviewers, Their Incentives, and Data Quality in Field Surveys. www.federalreserve.gov/pubs/oss/oss2/papers/iariw.2006.7.pdf
- Kennickell, A. B. 2007.** Look and Listen, But Don't Stop: Interviewers and Data Quality in the 2007 SCF. Proceedings of the Survey Research Methods Section. American Statistical Association. www.amstat.org/Sections/Srms/Proceedings/y2007/Files/JSM2007-000648.pdf
- Kennickell, A. B. 2008.** The Bitter End? The Close of the 2007 SCF Field Period. Proceedings of the Survey Research Methods Section. American Statistical Association.
www.amstat.org/sections/srms/Proceedings/y2008/Files/300522.pdf
- Kish, L. 1962.** Studies of Interviewer Variance for Attitudinal Variables. In: Journal of the American Statistical Association 57. 92–115.
- Survey of Consumer Finances. 2007.** Field Interviewer Manual 2007. NORC (National Opinion Research Center) at the University of Chicago.
- Survey of Consumer Finances. 2007.** Questionnaire Manual 2007. NORC at the University of Chicago.

Wagner, K. und S. Zottel. 2009. OeNB-Immobilienvermögenserhebung 2008 der privaten Haushalte – Beschreibung, verwendete Methoden und Bewertungsansätze. In: Statistiken Q4/09. Wien: OeNB.

Weinhardt, M., J. Schupp und F. Kreuter. 2009. The Impact of Interviewers' Personality on Measurement Error. www.diw.de/sixcms/detail.php/237557

