

Stichprobenziehung bei Erhebungen zu den Finanzen privater Haushalte im Euroraum

Dieser Beitrag ist der zweite einer Reihe von Artikeln zu den statistischen Herausforderungen der Forschung zu Finanzen und Ausgaben der privaten Haushalte im Euroraum. Er beschäftigt sich mit der Stichprobenziehung bei entsprechenden Mikrodaterhebungen und insbesondere mit dem sogenannten *Oversampling*, dem überproportionalen Miteinbeziehen bestimmter Bevölkerungsgruppen, das aufgrund der starken Vermögenskonzentration besonders für Erhebungen zu den Finanzen privater Haushalte eine wichtige Rolle spielt. Im ersten Teil des Beitrags werden mögliche Strategien und Verfahren bei der Stichprobenziehung erläutert. Im zweiten Teil werden verschiedene Implementierungsmöglichkeiten der Stichprobenziehung anhand von drei Fallbeispielen, dem deutschen sozioökonomischen Panel (SOEP), dem US-amerikanischen Survey of Consumer Finances (SCF) und dem spanischen Survey of Household Finances (EFF), präsentiert. Ein Fernziel für die regelmäßigen Befragungen des Eurosystems in Österreich sollte eine statistische Kooperation des Bundesministeriums für Finanzen, der Statistik Austria und der Oesterreichischen Nationalbank (OeNB) sein. Dabei ist besonders auf Anonymität und Datenvertraulichkeit zu achten.

Pirmin Fessler,
Peter Mooslechner,
Martin Schürz¹

1 Einleitung

Dies ist der zweite von vier Beiträgen für die Quartalspublikation *Statistiken – Daten & Analysen* zu den statistischen Herausforderungen, die sich bei Befragungen zu Finanzen und Ausgaben privater Haushalte im Euroraum stellen. Das Eurosystem plant, in den nächsten Jahren auf regelmäßiger Basis – zumindest alle drei Jahre – eine weitreichende Datenerhebung zu Finanzen und Ausgaben privater Haushalte (*Household Finance and Consumption Survey – HFCS*) durchzuführen. Die OeNB beabsichtigt die Durchführung der österreichischen Haushaltsbefragung für das Frühjahr 2010 in enger Kooperation mit der Deutschen Bundesbank. Der HFCS wird eine der wenigen Erhebungen weltweit sein, die Einkommens-, Ausgaben-, Vermögens- und Verschuldungsdaten der privaten Haushalte (im Folgenden Haushalte genannt) gemeinsam erheben. Die Notwendigkeit einer Mikrodaterhebung wurde bereits im ersten Artikel der Serie (Fessler, Mooslechner und Schürz, 2009) erläutert. Zudem bietet die Webseite des Household

Finance and Consumption Networks der EZB weitreichende Informationen zum Projekt und zur Nützlichkeit der daraus resultierenden Daten (www.ecb.int/home/html/researcher_hfcn.en.html). Der vorliegende Beitrag widmet sich der Frage der Stichprobenziehung und des *Oversampling* (Überrepräsentation einer Teilpopulation).

In Kapitel 2 werden statistische Probleme beim Stichprobendesign analysiert. In Kapitel 3 wird über Erfahrungen von drei Ländern – Deutschland, den USA und Spanien (Deutsches Institut für Wirtschaftsforschung – DIW, Fed und Banco de España) – referiert, die bereits seit langem erfolgreich Haushaltsbefragungen zu Finanzen durchführen.

2 Stichprobenziehung

Die Stichprobenziehung ist von zentralem Einfluss für die Qualität einer Befragung.

Das Stichprobendesign (Sample Design) beinhaltet die Regeln für den Selektionsprozess der Stichprobe. Aus einer Grundgesamtheit von Einheiten

¹ pirmin.fessler@oebn.at, peter.mooslechner@oebn.at, martin.schuerz@oebn.at

der Zielpopulation werden Einheiten für die Stichprobe gewählt. Die primäre Untersuchungseinheit im HFCS ist, wie in fast allen anderen Erhebungen zu Finanzen, der Haushalt. Die Menge aller Haushalte in einem Land stellt die Grundgesamtheit dar, aus der die Stichprobe gezogen wird.

Im Allgemeinen basieren nahezu alle Stichprobendesigns vergleichbarer Erhebungen auf einem sogenannten *probabilistischen Sampling*. Dabei besteht für jede Einheit der Grundgesamtheit eine positive Wahrscheinlichkeit in die Stichprobe einbezogen zu werden. Diese Wahrscheinlichkeiten sollten im Idealfall im Vorhinein bekannt sein, um valide Gewichtungen vornehmen zu können. Dies kann jedoch nur selten erreicht werden, da es kaum Datenbanken gibt, in denen tatsächlich jeder einzelne Haushalt erfasst wird. In den meisten Ländern gibt es Haushalte, die nicht in Telefon- bzw. Adressdatenbanken aufscheinen und/oder nicht im Melderegister erfasst sind. Des Weiteren verändert sich die Haushaltsstruktur und es kann nicht davon ausgegangen werden, dass die Grundgesamtheit der Haushalte zum Zeitpunkt der Ziehung identisch ist mit jener am Ende des Befragungszeitraums.

Um die Grundgesamtheit trotzdem möglichst vollständig zu erfassen, bieten sich verschiedene Strategien an: Manche Erhebungen basieren auf einem sogenannten *Area Probability Sampling (AP)*, bei dem geografisch Gebiete oder etwa Straßenzüge ausgewählt werden und dann innerhalb dieser Adressen nach dem Zufallsprinzip gezogen werden. Oder aber es werden auf Subebenen Random-Route-Verfahren eingesetzt, bei denen die Interviewer nach einer vorgegebenen Wegbeschreibung zufällige Haushalte wählen. Andere

Erhebungen greifen auf Adressdaten (z. B. Telefonbücher) oder Registerdaten (Melderegister) zurück. Die gewählten Herangehensweisen hängen stark von den in einem Land verfügbaren Daten ab. So werden in Deutschland und Österreich häufig Melderegisterdaten verwendet. Aber auch geografische Datensätze (wie etwa beim Pre-Test des HFCS der Deutschen Bundesbank oder dem US-Survey of Consumer Finances – SCF) kommen häufig zum Einsatz. Je nach Qualität der verfügbaren Datenbanken können unterschiedliche Methoden in den einzelnen Ländern optimal sein. Zudem variieren die Kosten je nach angewandter Methode. Ein Random-Route-Verfahren ist etwa mit deutlich höheren Kosten verbunden als eine Ziehung aus einem bereits vorhandenen Melderegister. Bei der Verwendung von Adressdateien von Meinungsforschungsinstituten besteht die Gefahr, dass es sich um Personendaten handelt. Sollte ein entsprechender Abgleich nicht funktionieren, werden einzelne Haushalte mehr als einmal in der Grundgesamtheit aufscheinen. Zudem ist es wahrscheinlich, dass auf vorrekrutierte Haushalte zurückgegriffen wird, da diese Institute ihre Datenbanken auf Basis von vorangegangenen Erhebungen erweitern. Am erfolgversprechendsten in Bezug auf Haushaltsdaten für Österreich wären eine Ziehung aus dem zentralen Melderegister oder ein Random-Route-Verfahren; Letzteres ist aber kostenintensiv.

Bei Vermögensumfragen stellt sich ein weiteres Problem. Es ist von einer hohen Variabilität auszugehen. Um die Vermögensungleichheit angemessen abbilden zu können, wäre eine sehr große Stichprobenanzahl notwendig, denn die Genauigkeit der Schätzungen steigt nur langsam mit der Stichprobengröße

an.² Dies macht Umfragen aber teuer. Die externen Kosten der Umfragen hängen in starkem Ausmaß von der Stichprobengröße ab. Die Länge des Fragebogens spielt in diesem Zusammenhang hingegen eine untergeordnete Rolle.

Ist es möglich, Bevölkerungsgruppen zu identifizieren, für die die Varianz der Vermögensvariablen besonders hoch ist, kann durch überproportionale Berücksichtigung dieser Bevölkerungsgruppe die Gesamtstichprobengröße kleiner gehalten werden. Dadurch kann bei Schätzungen auch eine niedrigere Varianz, das heißt eine präzisere Schätzung, erzielt werden.

Die zu wählende Stichprobengröße hängt in der Praxis davon ab, wie der Gesamtumfragefehler (die Summe der Fehler auf allen Stufen der Umfragedurchführung) bei gegebenem Budget am effizientesten minimiert werden kann.

2.1 Problem der hinreichenden Erfassung der Vermögenskonzentration

Die Vermögenskonzentration ist international beträchtlich (Tabelle 1). Dies

führt dazu – wenn alle Haushalte mit derselben Wahrscheinlichkeit in die Stichprobe gezogen werden – dass für besonders vermögende Haushalte am rechten Rand der Verteilung nur sehr unpräzise Schätzungen möglich sind. Weil diese Haushalte aber einen überproportionalen Anteil am Vermögen halten, sind sie für eine Analyse besonders relevant; Tabelle 1 veranschaulicht dieses Problem. In den USA hält das oberste Prozent der Vermögenden etwa ein Drittel der gesamten Vermögensbestände aller privaten Haushalte. Wird diese Gruppe nur proportional zu ihrem Bevölkerungsanteil in die Stichprobe einbezogen, hätte dies zur Folge, dass die Anzahl der Beobachtungen im Vergleich zu ihrer Relevanz in Bezug auf das Vermögen zu gering wäre. Innerhalb dieser Gruppe ist die Variabilität zudem am höchsten, da die Vermögensverteilung am rechten Rand besonders steil wird.

Eine überproportionale Ziehung, ein sogenanntes *Oversampling*, erfüllt zwei wichtige Aufgaben: Es ermöglicht genauere Schätzungen des Vermögens,

Tabelle 1

Internationaler Vergleich von Vermögensverteilungen

	Kanada SFS 1999	Finnland HWS 1998	Deutschland ¹ SOEP 2002	Italien SHIW 2002	Schweden HINK 2002	Vereinigtes Königreich BHPS 2000	USA PSID 2001	USA SCF 2001
Vermögensanteile in %								
Nettovermögensgruppen								
Top 10%	53	45	54	42	58	45	64	71
Top 5%	37	31	36	29	41	30	49	58
Top 1%	15	13	14	11	18	10	25	33
Ginikoeffizient	0,75	0,68	0,78	0,61	0,89	0,66	0,81	0,84
Relation Quantil/Median								
10. Perzentil	-17	-6	0	0	-84	0	-11	-15
25. Perzentil	0	1	0	8	-1	2	0	0
75. Perzentil	350	218	886	209	447	238	378	368
90. Perzentil	708	390	1.818	359	972	482	925	980

Quelle: Schürz (2008, Tabelle 1).

¹ Schwellenwert von 2.500 für Finanzaktiva und -passiva.

² Statistisch gesehen steht die Präzision eines Schätzers für den Mittelwert (etwa das durchschnittliche Haushaltsbruttovermögen) in einem inversen Zusammenhang zum Standardmessfehler, der proportional mit der Wurzel der Stichprobengröße sinkt.

insbesondere hinsichtlich bestimmter Finanzaktiva, und es erleichtert die Behandlung des Problems fehlender Antworten (in Bezug auf Unit- sowie Item-Non-Response³). Grundsätzlich existieren zwei Möglichkeiten des Oversampling. Entweder kann in derselben Stichprobe die Wahrscheinlichkeit für vermögendere Haushalte, in die Stichprobe aufgenommen zu werden, erhöht werden. Oder es wird eine zusätzliche Stichprobe gezogen, die besonders vermögende Haushalte beinhaltet. Im Nachhinein können diese dann, entsprechend gewichtet nach ihrem tatsächlichen Bevölkerungsanteil, mit der Grundstichprobe kombiniert werden. Wieder hängt die Wahl der Methode stark von der Datenverfügbarkeit in den jeweiligen Ländern ab. Sowohl das deutsche sozioökonomische Panel (SOEP) als auch der US-SCF erheben zusätzliche Stichproben, die dann mit der Grundstichprobe kombiniert werden. Im Fall des SOEP wird die Grundgesamtheit der Zusatzstichprobe durch eine zusätzliche telefonische Befragung zum Einkommen erstellt. Im Fall des US-SCF werden Daten des Finanzministeriums verwendet. Im deutschen Pre-Test des HFCS kam die erste Variante zum Einsatz. Dabei wurde jenen Haushalten, die in besseren Wohngebieten (auf Basis von Kfz-Zulassungs- und Immobiliendaten identifiziert) liegen, eine höhere Ziehwahrscheinlichkeit zugeordnet. Grundsätzlich ist die Verwendung von Steuerdaten aber die zuverlässigste Methode. In Kapitel 3 wird detailliert auf Bei-

spiele aus Deutschland, den USA und Spanien eingegangen.

2.2 Problem der Teilnahme- und Antwortverweigerung

Eine zuverlässige Erfassung von Vermögen im Rahmen von freiwilligen Haushaltsbefragungen ist schwierig. Insbesondere ergeben sich Probleme bei der Erfassung der Ränder der Verteilungen, das heißt bei der Erhebung von sehr niedrigen und sehr hohen Vermögensbeständen. Bei Vermögenden kommt es häufiger zu einer Teilnahmeverweigerung. Und auch wenn die Vermögenden teilnehmen, stellen sich Probleme. Vermögendere Haushalte haben komplexere Finanzprodukte und vielfältigere Anlagen. Interviews dauern daher zwangsläufig länger. Auch die Kontaktaufnahme gestaltet sich schwieriger. Es kommt zu häufigeren Terminverschiebungen und -absagen. Da Haushaltsgewichte nur auf Basis der in die Stichprobe gezogenen Haushalte korrigieren können, kann das Problem der systematischen Unit-Non-Responses nicht gelöst werden.

Oversampling kann einen Beitrag zu einer Lösung des Unit-Non-Response-Problems und der damit verbundenen niedrigen Präzision leisten. Die Präzision kann entsprechend der gewählten Stichprobengröße, der zu analysierenden Forschungsfrage und dem Grad der Vermögenskonzentration beliebig erhöht werden. Zudem steigt die Wahrscheinlichkeit, die tatsächliche Variabilität abzudecken, da sich auch die Wahrscheinlichkeit erhöht,

³ *Unit-Non-Response bezeichnet die Nichtteilnahme in die Stichprobe gezogener Untersuchungseinheiten an einer Erhebung. Dadurch kann die sogenannte Ausschöpfungsquote (Response Rate) stark gesenkt werden. Gründe sind meist Verweigerung, Nichterreichbarkeit oder Nichtteilnahmefähigkeit, die durch Krankheit, Alter oder Sprachbarrieren gegeben sein kann.*

Item-Non-Response bezeichnet einen partiellen Ausfall, wenn Untersuchungseinheiten einzelne Fragen nicht beantworten. Gründe dafür sind meist Verweigerung oder Nichtwissen. Speziell Verweigerung tritt häufig bei sensitiven Fragen auf.

Haushalte am äußersten rechten Rand der Verteilung, etwa im Promillebereich, in die Stichprobe einzubeziehen. Die massive Ungleichheit innerhalb der Gruppe der Vermögendsten kann zu deutlichen Unterschieden für die Schätzer führen. Aber auch das Oversampling ist nur ein Element der Problemlösung. Zwar reduziert es durch den erhöhten Umfang der Stichprobe eine zufällige Unterschätzung⁴ des Mittelwerts der Gruppe und erhöht die Präzision desselben. Wenn aber kein Haushalt ab einer bestimmten Vermögenshöhe zu einem Interview bereit ist, kann auch das Oversampling dieses Problem nicht lösen.

Neben diesen Vorteilen hinsichtlich Unit-Non-Response wird Oversampling vor allem bei der Lösung von Item-Non-Response-Problemen hilfreich sein. Item-Non-Response bedeutet, dass ein Haushalt an der Erhebung teilnimmt, aber einige Fragen nicht beantworten kann oder will. Ein Item-Non-Response-Problem bei Vermögensfragen ergibt sich insbesondere bei bestimmten Vermögensarten (Aktien, Anleihen usw.), die zwar gehalten werden, deren (Beträge) Werte aber in der Befragung nicht angegeben werden. Sind diese fehlenden Werte nicht zufällig über die Haushalte der Stichprobe verteilt, führt dies zu verzerrten Schätzern. Im Allgemeinen treten Verweigerungen zu Einkommens- oder Vermögensfragen besonders bei einkommens- und vermögensstarken Haushalten auf.

Wird nicht für diese fehlenden Werte imputiert, werden entsprechende Durchschnitte bzw. Vermögensaggregate unterschätzt. In der Literatur wird dies als *Listwise Deletion* bezeichnet (Durrant, 2005; Little und Rubin, 2002). Wenn die fehlenden Werte mit bestimmten Haushaltscharakteristika korreliert sind, insbesondere mit Variablen, die mit hohem Vermögen positiv korrelieren, wie z. B. hohe Einkommen, hohe Bildung oder teure Wohngegend, ist es sinnvoll, diese Werte zu imputieren, um zumindest teilweise für die Verzerrung korrigieren zu können. Durch die größere Anzahl an Beobachtungen, besonders für nicht weit verbreitete Finanzaktiva, werden Imputationen erleichtert und deren Qualität verbessert. Rezentere statistische Imputationsmethoden (Multiple Imputation) erlauben zudem eine korrekte Schätzung der Varianz, da sie miteinbeziehen, dass auch der imputierte Wert nicht dem wahren Wert entspricht, sondern einer gewissen Unsicherheit unterworfen ist. Für die Immobilienvermögensbefragung der OeNB wird ein multiples Imputationsverfahren angewandt (Albacete et al., 2009). Tabelle 2 zeigt exemplarisch einige Item-Non-Response-Raten aus der OeNB-Immobilienvermögenserhebung 2008, wobei „Angabe eines Betrags“ eine genaue Angabe des Werts und „Angabe von Bändern“ die Angabe von Intervallen, in denen sich der Wert befindet, das heißt eine ungefähre Angabe, bedeutet. Die vorgegebenen Bänder sind in der Regel recht schmal.

⁴ Theoretisch ist auch eine Überschätzung möglich, aber eine Unterschätzung ist aufgrund der Schiefe der Verteilung weitaus wahrscheinlicher.

Exemplarische Item-Non-Response-Raten aus der Immobilienvermögenserhebung der OeNB

	Anteil der Haushalte, die zur Frage gelangen	Angabe eines Betrags	Angabe von Bändern	Nichtwissen	Antwortverweigerung
	in %				
Haushaltsnettoeinkommen	100,0	67,3	22,6	0,4	9,6
Geschätzter Verkaufspreis der Immobilie	52,1	73,4	14,8	7,3	4,4
Höhe des Kaufpreises der Immobilie	39,8	65,7	20,9	8,1	5,3
Kredithöhe (für den ersten Kredit)	29,6	84,4	6,9	4,0	4,7
Geschätzter Wert der geerbten Immobilie	20,1	61,5	0,0	6,9	31,6

Quelle: OeNB-Immobilienvermögenserhebung 2008.

2.3 Problem der Überprüfung der Qualität der Mikrodaten

Bei der Erhebung von Vermögen in Haushaltsbefragungen tritt das Problem auf, dass die Vermögensaggregate, die sich aus den Haushaltsstichproben ergeben, kleiner sind als die in den Volkswirtschaftlichen Gesamtrechnungen (VGR) ausgewiesenen Bestände. Abweichungen können sich aus Abgrenzungsunterschieden bezüglich der relevanten Untersuchungspopulation und den spezifischen Vermögenskomponenten ergeben. So werden große Vermögensbestände teils gar nicht von privaten Haushalten, sondern von Privaten Organisationen ohne Erwerbszweck (POoE) oder von Privatstiftungen gehalten. Durch den Vergleich mit den VGR-Daten können Differenzen zwar identifiziert werden, aber die Frage der Korrektur bleibt ungeklärt. Denn nicht immer muss der Fehler bei den Mikrodaten liegen.

Betrachtet man einzelne Unterpositionen des Vermögens, dann zeigt sich die Problematik einer unzureichenden Erfassung insbesondere bei den Finanzaktiva. Bei den Nicht-Finanzaktiva kommen Erhebungen hingegen auf einen Deckungsgrad von über 90%

(Tabelle 3). Die Nachweisquote bei den Finanzaktiva liegt hingegen nur zwischen 16% und 38%. Basis dieser Daten ist die Luxembourg Wealth Study (LWS), die eine Ex-post-Harmonisierung von zehn nationalen Datensets für Vermögen privater Haushalte vornahm. Gerade bei Pro-Kopf-Vergleichen muss aber bedacht werden, dass bei der Berechnung der Pro-Kopf-Werte auf Makroebene das gesamte Vermögen von POoE, selbstständig Erwerbstätigen oder gar Privatstiftungen fälschlich auf die Bevölkerung verteilt wird, auch wenn sich diese Vermögen definitiv nicht direkt im Eigentum von privaten Haushalten (wie sie als Zielpopulation der Mikroerhebungen definiert sind) befinden. Dies spielt vor allem bei den Finanzaktiva eine Rolle, die für private Haushalte im Vergleich zum realen Vermögen eine untergeordnete Rolle spielen, aber für Privatstiftungen besonders relevant sind; das heißt, sogar im Idealfall, wenn die Haushaltsbefragung auch den rechten Rand der Verteilung abbilden würde, wären beträchtliche Unterschiede zwischen Mikro- und Makrodaten zu erwarten.

Tabelle 3

Pro-Kopf-Haushaltsvermögen in den LWS-Daten und den VGR-Daten

	Kanada SFS 1999	Zypern SCF 2002	Finnland HWS 1998	Deutsch- land SOEP 2002	Italien SHIW 2002	Nor- wegen IDS 2002	Schwe- den HINK 2002	Verein- igtes König- reich BHPS 2000	USA PSID 2001	USA SCF 2001
in EUR										
LWS-Datenbank										
Nicht-Finanzvermögen	28.237	32.763	31.920	53.507	50.965	14.605	33.132	61.436	63.170	77.686
Finanzvermögen	8.018	6.294	6.181	7.971	8.913	22.066	12.943	11.036	31.332	47.059
Verpflichtungen	9.577	3.719	6.032	11.202	2.590	29.561	16.159	13.572	20.857	26.707
Nettovermögen	26.678	35.339	32.069	50.276	57.288	7.110	29.916	58.901	73.646	98.037
Volkswirtschaftliche Gesamtrechnung										
Nicht-Finanzvermögen	32.492	x	x	69.234	78.417	x	x	67.728		66.679
Finanzvermögen	51.157	38.099	20.317	44.731	48.780	42.268	40.927	87.199		123.768
Verpflichtungen	13.813	15.825	7.147	18.750	7.089	33.629	16.577	20.471		31.003
Nettovermögen	69.836	x	x	95.215	120.108	x	x	134.457		159.444
in %										
Relation Mikro-/Makrodaten										
Nicht-Finanzvermögen	87	x	x	77	65	x	x	91	95	117
Finanzvermögen	16	17	30	18	18	52	32	13	25	38
Verpflichtungen	69	23	84	60	37	88	97	66	67	86
Nettovermögen	38	x	x	53	48	x	x	44	46	61

Quelle: Sierminska et al. (2006, Tabelle 5).

3 Erfahrungen in anderen Ländern**3.1 SOEP – Deutschland**

Das SOEP stellt eine für Deutschland repräsentative Zufallsstichprobe von Haushalten und Personen dar, bei der (zu einem hohen Grad) dieselben Haushalte und Personen jährlich nach demografischen Merkmalen und Einkommen sowie nach vielfältigen Merkmalen ihrer Lebenslage und deren subjektiver Einschätzung befragt werden (Wagner et al., 2006). Der Stichprobenumfang inkludiert mehr als 11.400 Haushalte mit knapp 21.100 befragten Personen im Alter von 16 und mehr Jahren.

Die gesamte SOEP-Stichprobe setzt sich aus acht Teilstichproben zusammen:

- Stichprobe A: Westdeutsche Haushalte;
- Stichprobe B: Ausländische Haushalte, deren Haushaltsvorstand türkischer, italienischer, jugoslawischer, griechischer oder spanischer

Nationalität ist (Ziehungsbeginn: 1984);

- Stichprobe C: Haushalte der ehemaligen DDR (Ziehungsbeginn: 1990);
- Stichprobe D: Zuwandererhaushalte (Ziehungsbeginn: 1994/95);
- Stichprobe E: Auffrischungsstichprobe (Ziehungsbeginn: 1998);
- Stichprobe F: Ergänzungsstichprobe (Ziehungsbeginn: 2000);
- Stichprobe G: Hocheinkommensstichprobe mit Privathaushalten in Deutschland, die mindestens ein verfügbares monatliches Hauhaushaltseinkommen von mehr als 3.835 EUR aufweisen (Ziehungsbeginn: 2002);
- Stichprobe H: Ergänzungsstichprobe (Ziehungsbeginn: 2006).

Da hohes Vermögen positiv mit hohem Einkommen korreliert, wurde im Rahmen des sozioökonomischen Panels in Deutschland eine Telefonbefragung durchgeführt, um die Hocheinkommensbezieher in der deutschen Bevöl-

kerung zu identifizieren. Da es kein Verzeichnis der Haushalte im oberen Einkommensbereich gibt, wurde eine Zufallsstichprobe aus Telefonadressen in der Größenordnung von 100.000 Haushalten durchgeführt.

Im Jahr 2002 wurde erstmals eine Hocheinkommensstichprobe gezogen. Diese Stichprobe war vergleichsweise klein. Im Einkommensbereich unterhalb der Hocheinkommensschwelle von rund 41.500 EUR lagen hochgerechnet 97,5% der Beobachtungen. Durch die Stichprobe G wurden sehr hohe Einkommen überrepräsentiert. Für die obersten 2,5% der Nettoäquivalenzeinkommen stehen im SOEP jetzt rund 600 Haushalte bzw. 1.400 Personen, die in diesen Haushalten leben, zur Verfügung. 2002 lag das Maximaleinkommen (Vorjahres-Haushaltsnettoeinkommen) der Altstichproben (A–F) bei rund 550.000 EUR, aber im Sample G bei über 1 Mio EUR.

3.2 SCF – USA

Der SCF setzt sich aus einer AP-Stichprobe und aus einer Spezialstichprobe zusammen (Kennickell, 2006a). Auf Basis von Daten des Finanzministeriums zu den Vermögenseinkommen wird eine Spezialstichprobe erhoben. Die Informationen beziehen sich auf die Einkünfte aus Vermögen und nicht auf das Vermögen selbst. Der Besteuerung hinterzogene Beträge müssen unberücksichtigt bleiben. Zudem handelt es sich um Angaben zu Individuen, während im SCF Haushalte untersucht werden.

Da in den USA die Top-1% der Verteilung einen Anteil von einem Drittel des gesamten Vermögens halten (Kennickell, 2006b), ist es wichtig, diese kleine Gruppe in der Stichprobe angemessen abzubilden. Die Vermögenden leben zudem in einem überdurchschnittlichen Maß in städtischen

Gebieten und wären daher in der AP-Stichprobe nicht angemessen erfasst.

Etwa 98% der SCF-Haushalte mit einem Nettovermögen von mindestens 5 Mio USD im Jahr 2004, mehr als 85% der Haushalte mit mindestens 1 Mio USD Nettovermögen und rund 75% der Haushalte mit zumindest 500.000 USD kommen aus dieser Spezialstichprobe. Dies veranschaulicht prima facie, dass die Spezialstichprobe von großer Bedeutung ist. Die Stichprobe schließt jene Personen aus, die auf der Forbes-Liste der 400 Reichsten aufscheinen.

Für den SCF liegt Evidenz vor (Kennickell, 2006a), dass die Antwortraten mit steigenden Vermögenseinkommen zurückgehen und mit dem Alter und wohltätigen Zuwendungen ansteigen. Mehr Freizeit und eine altruistische Gesinnung beeinflussen demnach die Teilnahmebereitschaft. In jenem Stratum der Grundstichprobe, das die Vermögendsten umfassen sollte, liegt die Antwortrate nur bei 10%. Grafik 1 zeigt deutlich die enorme Reduktion der Verzerrung nach unten, die durch die Zusatzstichprobe des SCF erzielt werden kann.

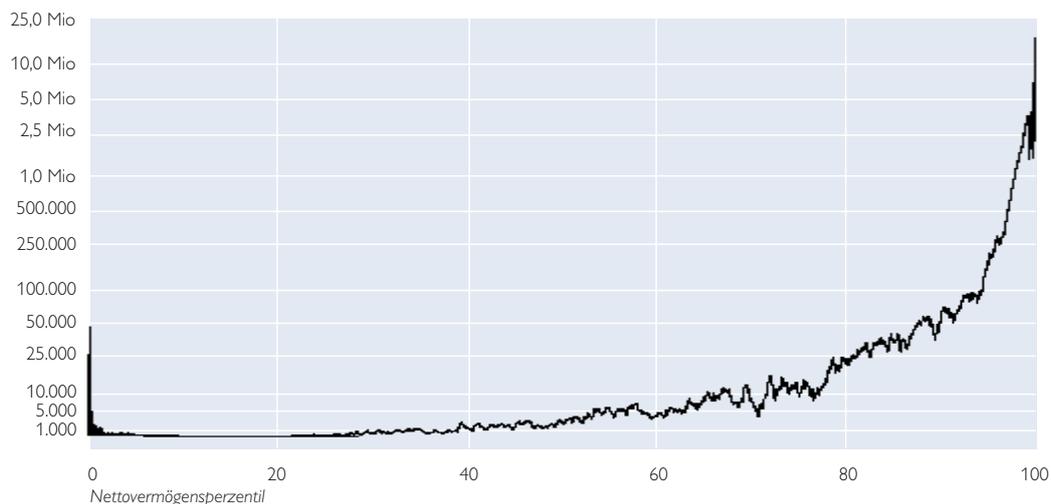
Erkennbar ist die Differenz zwischen den beiden geschätzten Verteilungen zu jedem Punkt der Verteilung des Nettovermögens. Für sehr niedrige Werte macht die Ergänzung durch die Spezialstichprobe die Verteilung negativer. Dies ist aber nur „noise“ und kein systematischer Effekt (Kennickell, 2007). Vom 20. Perzentil aufwärts ist das Niveau des Nettovermögens bei jedem Perzentil in der kombinierten Gesamtstichprobe höher als bei der AP-Stichprobe.

Der geschätzte Anstieg im Niveau liegt bei 250 USD (1,9%) beim 25. Perzentil, bei 3.500 USD (3,9%) beim Median, bei 13.600 USD (4,3%) beim 75. Perzentil, bei 43.600 USD (5,5%)

Grafik 1

Verteilung des Nettovermögens 2004 basierend auf der vollständigen Stichprobe minus der Verteilung basierend auf der geografisch gewichteten Zufallsstichprobe nach dem Perzentil des Nettovermögens

in USD



Quelle: Kennickell (2007, Abbildung 1).

beim 90. Perzentil und bei 2.661.000 USD (74,0%) beim 99. Perzentil.

An der Spitze der Vermögensverteilung erfüllt die Spezialstichprobe demnach eine wichtige Funktion, da es hier kaum Stichprobenfälle gibt. Unter den Top-1% wären es ohne Spezialstichprobe nur 41 Haushalte, mit Oversampling sind es aber 715. Die Spezialstichprobe ersetzt nicht nur das 1% an

der Spitze der Verteilung, sondern hilft bei der Rekalibrierung in den Folgejahren der Erhebung, weil dann für die Untererfassung der Haushalte im obersten Segment korrigiert werden kann. Oversampling ist demnach ein statistisches Unterfangen mit anhaltenden Vorteilen für die Erhebung.

Trotz der großen Differenzen der beiden Stichproben bei den Vermögens-

Tabelle 4

Nettovermögensanteile für verschiedene Perzentile der Verteilung in den USA (2004)

Nettovermögensgruppen	AP-Stichprobe			AP- und Spezialstichprobe kombiniert		
	Anteil der Perzentile	Standardfehler ¹	Anzahl der Beobachtungen	Anteil der Perzentile	Standardfehler	Anzahl der Beobachtungen
	in %			in %		
1% bis 50%	2,9	0,3	1.642	2,5	0,1	1.741
50% bis 90%	30,9	2,8	1.097	27,9	0,9	1.343
90% bis 95%	12,9	1,1	132	12,0	0,7	269
95% bis 99%	19,7	1,4	109	24,1	1,2	454
Top 1%	33,6	5,1	27	33,4	1,2	715
Insgesamt	100,0	0,0	3.007	100,0	0,0	4.522

Quelle: Kennickell (2007, Tabelle 1).

¹ Standardfehler für die AP-Stichprobe reduziert um $\sqrt{3.007/4.522}$ (0,817).

werten an der Spitze der Verteilung liegt der Anteil der Top-1 % am gesamten Vermögen stabil bei etwa einem Drittel. Die Standardfehler unterscheiden sich aber beträchtlich voneinander; sie liegen bei 5,1 % in der AP-Stichprobe und bei nur 1,2 % in der kombinierten Stichprobe.

Die Spezialstichprobe erlaubt zudem eine Analyse von speziellen Finanzportfolios. Im SCF 2004 gaben nur 1,8 % der Haushalte an, direkt Staatsanleihen zu halten. Von den 400 Haushalten der kombinierten Stichprobe, die Anleihen halten, kommen nur 10 % vom AP-Sample.

Tabelle 4 zeigt Schätzungen der Nettovermögensanteile verschiedener Perzentile und die zugehörigen Standardfehler.

3.3 EFF – Spanien

In Spanien gibt es bislang noch eine Vermögensteuer.⁵ Auf dieser Datenbasis kann ein Oversampling vorgenommen werden.

Die Vermögensteuerdaten zeigen, dass 0,4 % der Haushalte 40 % des zu versteuernden Vermögens halten. In einer Stichprobe von 5.000 sollte man

daher nur 20 solcher vermögenden Haushalte erwarten; im EFF sind es aber 500. Die Banco de España definierte acht Vermögensstrata, die mit ansteigenden Raten überrepräsentiert werden (Tabelle 5).

3.4 Andere Möglichkeiten des Oversampling

Neben Vermögen- und Einkommensteuer könnte auch die Grundsteuer als Basis herangezogen werden. Liegen aber keine administrativen Daten vor, ist die Situation noch ungleich schwieriger.

In Zypern wurde – nach Angaben der Projektverantwortlichen recht erfolgreich – der Elektrizitätsverbrauch als Auswahlkriterium für das reiche Bevölkerungssegment herangezogen. Die Besitzer von Zweitwohnsitzen haben einen höheren Elektrizitätsverbrauch als die einkommensschwächeren Mieter.

In Kanada wurde versucht, nach regionalen Merkmalen zu unterscheiden und Personen aus wohlhabenden Gebieten wurden häufiger befragt. Die Ergebnisse waren aber unbefriedigend.

Tabelle 5

Grad des Oversampling in der Stichprobe des EFF

Nettovermögensgruppen	EFF 2002		EFF 2005	
	Anzahl der Beobachtungen	Oversampling Rate ¹	Anzahl der Beobachtungen	Oversampling Rate
1 % bis 50 %	1.878	0,73	2.234	0,75
50 % bis 90 %	1.944	0,94	2.036	0,85
90 % bis 95 %	429	1,67	481	1,61
95 % bis 99 %	524	2,55	675	2,83
Top 1 %	368	7,16	536	8,99

Quelle: Bover (2008, Tabelle 7).

¹ Die Oversampling Rate ist definiert als die Anzahl der Beobachtungen in der Stichprobe in einem bestimmten Nettovermögensbereich der Verteilung dividiert durch die erwartete Anzahl von Beobachtungen in demselben Bereich, wenn die Stichprobe zufällig aus der Population gezogen würde.

⁵ Da die Vermögensteuer in Spanien im Jahr 2009 abgeschafft wird, soll künftig auf Einkommensteuerdaten zugegriffen werden. Für 2011 kann aber, nach Auskunft von O. Bover, noch der Vermögensteuersatz verwendet werden.

In der Haushaltsbefragung der OeNB zum Geldvermögen 2004 konnte kein Oversampling durchgeführt werden. Versuche, ein regionales Muster der Geldvermögensverteilung zu nutzen, blieben erfolglos. Es zeigten sich z. B. in Wien auch nicht die erwarteten Vermögensunterschiede nach Bezirken, wonach der Wohnsitz von Vermögenden eher in der Inneren Stadt, Hietzing und Döbling zu vermuten wäre.

4 Schlussfolgerungen

Die Stichprobenziehung bei Befragungen zu den Haushaltsfinanzen gestaltet sich schwierig. Vermögenskonzentration und Teilnahme- bzw. Antwortverweigerungen setzen freiwilligen Erhebungen enge Grenzen.

Oversampling bei Vermögensbefragungen von privaten Haushalten bietet grundsätzlich einen Ausweg aus dieser

schwierigen Situation. Die Grundlagen für solche Spezialstichproben sind jedoch länderweise verschieden, und die internationalen Erfahrungen unterscheiden sich beträchtlich voneinander.

In Spanien bildet eine Vermögensteuer die Basis, in den USA ein Einkommensteuerdatensatz und das SOEP zieht die Stichprobe auf Basis von Telefonkontakten.

In Österreich ist die statistische Situation schlecht, da für die Stichprobenziehung auf kein Steuerregister zugegriffen werden kann. Ein Fernziel für die regelmäßigen Befragungen des Eurosystems sollte eine statistische Kooperation zwischen Bundesministerium für Finanzen, Statistik Austria und OeNB sein. Dabei ist besonders auf Anonymität und Datenvertraulichkeit zu achten.

Literaturverzeichnis

- Albacete, N., P. Fessler und K. Wagner. 2009.** Multiple Imputation in the Austrian Household Survey on Housing Wealth 2008. Mimeo.
- Bover, O. 2008.** The Spanish Survey of Household Finances (EFF): Description and Methods of the 2005 Wave. Documentos Occasionales 0803.
- Bucks, B. K., A. B. Kennickell und K. B. Moore. 2006.** Recent Changes in Family Finances: Evidence from the 2001 and 2004 Surveys of Consumer Finances. Federal Reserve Bulletin. A1–A38.
- Durrant, B. G. 2005.** Imputation Methods for Handling Item-Nonresponse in the Social Sciences: A Methodological Review. ESRC National Centre for Research Methods and Southampton Statistical Sciences Research Institute. NCRM Methods Review Papers NCRM/002.
- Fessler, P., P. Mooslechner und M. Schürz. 2009.** Herausforderungen der Forschung zur Finanzsituation privater Haushalte im Euroraum. In: Statistiken – Daten & Analysen Q1/09. Wien: OeNB.
- Frankel, M. und A. B. Kennickell. 1995.** Toward the Development of an Optimal Stratification Paradigm for the Survey of Consumer Finances. www.federalreserve.gov/pubs/oss/oss2/papers/frankel.kennickell.asa95.pdf
- Frick, J. R., J. Goebel, M. M. Grabka, O. Groh-Samberg und G. Wagner. 2007.** Zur Erfassung von Einkommen und Vermögen in Haushaltssurveys: Hocheinkommensstichprobe und Vermögensbilanz im SOEP. DIW Data Documentation 19. Berlin.
- Groves et al. 2004.** Survey Methodology. Wiley Series in Methodology. Wiley.
- Kennickell, A. B. 2006a.** How Do We Know if We Aren't Looking? An Investigation of Data Quality in the 2004 SCF. Federal Reserve Board. Paper prepared for the 2006 Annual Meetings of the American Statistical Association. Seattle, Washington D.C.

- Kennickell, A. B. 2006b.** Currents and Countercurrents: Changes in the Distribution of Wealth, 1989–2004.
www.federalreserve.gov/pubs/oss/oss2/papers/concentration.2004.5.pdf
- Kennickell, A. B. 2007.** The Role of Over-sampling of the Wealthy in the Survey of Consumer Finances. ISA 56th Conference.
- Little, R. J. A. und D. B. Rubin. 2002.** Statistical Analysis with Missing Data. Wiley-Interscience.
- Schürz, M. 2008.** Auf der Suche nach dem verschwiegenen Reichtum. *Intervention* 5(1). 62–74.
- Sierminka, E., A. Brandolini und T. Smeeding. 2006.** The Luxembourg Wealth Study – A Cross-Country Comparable Database for Household Wealth Research. In: *Journal of Economic Inequality* 4(3).375–383.
- Wagner, G., J. R. Frick und J. Schupp. 2006.** Enhancing the Power of Household Panel Studies – The Case of the German Socio-Economic Panel Study (SOEP). *Data Documentation* 13.