

SOEP Survey Papers

Series B - Survey Reports (Methodenberichte)

SOEP 2002 – Pretestbericht Feasibility Study zum Befragungsjahr 2002 (Welle 19) des Sozio-oekonomischen Panels (3. Bericht)

Running since 1984, the German Socio-Economic Panel Study (SOEP) is a wide-ranging representative longitudinal study of private households, located at the German Institute for Economic Research, DIW Berlin.

The aim of the SOEP Survey Papers Series is to thoroughly document the survey's data collection and data processing.

The SOEP Survey Papers is comprised of the following series:

Series A – Survey Instruments (Erhebungsinstrumente)

Series B – Survey Reports (Methodenberichte)

Series C – Data Documentation (Datendokumentationen)

Series D – Variable Descriptions and Coding

Series E – SOEPmonitors

Series F – SOEP Newsletters

Series G – General Issues and Teaching Materials

The SOEP Survey Papers are available at
<http://www.diw.de/soepsurveypapers>

Editors:

Prof. Dr. Gert G. Wagner, DIW Berlin and Technische Universität Berlin

PD Dr. Joachim R. Frick, DIW Berlin

Prof. Dr. Jürgen Schupp, DIW Berlin and Freie Universität Berlin

Please cite this paper as follows:

Infratest Sozialforschung. 2011. SOEP 2002 – Pretestbericht Feasibility Study zum Befragungsjahr 2002 (Welle 19) des Sozio-oekonomischen Panels (3. Bericht). SOEP Survey Papers 42: Series B. Berlin: DIW/SOEP

ISSN: 2193-5580 (online)

Contact: DIW Berlin
SOEP
Mohrenstr. 58
10117 Berlin

Email: soepapers@diw.de

INFRATEST SOZIALFORSCHUNG

SOEP 2002 – PRETESTBERICHT FEASIBILITY STUDY ZUM BEFRA- GUNGSJAHR 2002 (WELLE 19) DES SOZIO-OEKO- NOMISCHEN PANELS (3. BERICHT)

München, 2001

Reprint 2011

Inhalt	Seite
Vorbemerkung	3
1. Ziele und Ergebnisse der Testerhebung im Jahr 2001	4
1.1 Ziele	4
1.2 Die Stichprobenbasis	5
1.3 Durchführung der Befragung	6
1.4 Validierung der Stichprobe	7
2. Optionen für die Haupterhebung im Jahr 2002	9
2.1 Ziele	9
2.2 Die Stichprobenbasis	10
2.3 Zwei alternative Modelle für die Haupterhebung	16

Vorbemerkung

Der vorliegende Bericht schließt an folgende bereits vorgelegten Berichte zur Feasibility Study an:

1. Zwischenbericht

Testerhebung 2001: Die Stichprobenbasis.
Infratest Sozialforschung, Juli 2001

2. Zwischenbericht

Testerhebung 2001: Feldbericht und Hinweise zum Fragebogen.
Infratest Sozialforschung, September 2001

Wesentliche Punkte aus diesen Berichten werden im folgenden zusammenfassend dargestellt. Für Einzelheiten wird auf die genannten Berichte verwiesen.

1. Ziele und Ergebnisse der Testerhebung 2001

1.1 Ziele

Das Sozio-ökonomische Panel (SOEP) wird seit dem Jahr 2000 erweitert durch das vom BMBF geförderte Ergänzungsprojekt „Erprobung innovativer Erhebungskonzepte für Haushalts-Panel-Studien“. Die Feasibility Study „Einkommensstarke Haushalte“, über die hier berichtet wird, ist Teil dieser Erprobungen innovativer Ansätze.

Die Fragestellung ist, ob und mit welchem Erhebungsdesign die repräsentative Bevölkerungstichprobe des SOEP im Bevölkerungssegment der „Einkommensstarken Haushalte“ aufgestockt werden kann, um die Analysemöglichkeiten im oberen Bereich der Einkommensschichtung zu erweitern.

Der Wunsch nach einer derartigen *Stichprobenaufstockung* steht dabei in unmittelbarem Zusammenhang mit der Planung des *Fragenprogramms* für das SOEP 2002, das u.a. den Themenschwerpunkt „Vermögen“ behandeln soll. Wie bekannt, konzentrieren sich Vermögensbestände und die Vermögensbildung relativ stark auf die oberen Einkommensgruppen.¹ Wenn es gelänge, die Stichprobenbasis des SOEP für diesen Teil der Bevölkerung gezielt zu verbreitern, wäre damit der Wert des angestrebten Datenbestands zum Thema „Vermögen“ erheblich vergrößert.

Disproportionale Stichprobenansätze zur Überrepräsentation bestimmter Bevölkerungsteile werden im SOEP bereits praktiziert und haben seine Nutzungsmöglichkeiten nachhaltig gesteigert:

- Die *Ausländer-Stichprobe B* wurde 1984 gebildet, um die fünf zahlenmäßig stärksten „Gastarbeiter“-Nationalitäten mit analysefähigen Fallzahlen zu repräsentieren.
- Die *Zuwanderer-Stichprobe D* wurde 1995/96 gebildet, um die Zuwanderung im zurückliegenden Jahrzehnt seit 1984 im Stichprobensystem des SOEP abzubilden.

In beiden Fällen war die Entwicklung eines geeigneten Stichproben- und Erhebungsdesigns eine eigenständige, keineswegs einfache Aufgabe. Während für Stichprobe B Auswahlgrundlagen für eine Stichprobenziehung aus Einwohnermelderegistern vorhanden waren, traf dies für Stichprobe D nicht zu. Hier wurde der Weg über ein gezieltes Adressensammeln („Screening“) im Zusammenhang mit laufenden Infratest-Bevölkerungsumfragen gewählt, um die gewünschte (Über-)Repräsentation der Zuwanderer zu erreichen.

Für die nunmehr angestrebte Aufstockung der SOEP-Stichprobe im oberen Einkommensbereich ist die Problemlage ähnlich:

¹ Vgl. Lebenslagen in Deutschland: Der erste Armuts- und Reichtumsbericht der Bundesregierung. Berlin 2001, S. 53 f.

- Man benötigt für die Stichprobenbildung eine Vorinformation über bestimmte Merkmale der potentiellen Zielhaushalte.
- Es gibt aber keine Auswahlgrundlage, um aus der Grundgesamtheit der einkommensstarken Haushalte – wie immer sie definiert sind – mit einer direkten Zufallsauswahl eine Stichprobe zu bilden.

DIW und Infratest haben in einer längeren Vorphase verschiedene konzeptionelle Ansätze diskutiert. Als aussichtsreichster Weg, der auch mit den methodischen Standards des SOEP gut vereinbar ist, wurde das Stichproben- und Erhebungsdesign gewählt, das im Rahmen dieser Feasibility Study mit einer Testerhebung erprobt wurde.

1.2 Die Stichprobenbasis

Eine repräsentative Stichprobe der einkommensstarken Haushalte kann nicht direkt gebildet werden, sondern nur auf dem (Um-)Weg einer repräsentativen Bevölkerungsstichprobe, in der per Befragung festgestellt wird, welche Personen bzw. Haushalte zum oberen Einkommenssegment gehören. Diese Personen bilden dann die Ausgangsstichprobe für die eigentliche Befragung im Rahmen des SOEP.

Wenn ein solches Verfahren erfolgreich umgesetzt werden soll, müssen eine Reihe von Voraussetzungen gegeben sein:

(1) *Eine ausreichend große Basisstichprobe der Bevölkerung*

Infratest kann hierfür auf die kumulierte Jahresstichprobe der telefonischen Mehrthemenumfrage *InfraScope* zurückgreifen. Für die Testerhebung bei einkommensstarken Haushalten im Sommer 2001 wurde die Jahresstichprobe 2000 verwendet, die 87.000 befragte Personen umfasst. Es handelt sich um eine Zufallsstichprobe von Personen ab 14 Jahren, die in Privathaushalten mit Telefonanschluss (Festnetz) leben. Die Methodik der Stichprobenbildung ist im *1. Zwischenbericht* zu dieser Feasibility Study (S. 5 ff.) genauer beschrieben.

(2) *Eine Einkommensinformation in der zugrundeliegenden Befragung*

Die Standarddemographie der *InfraScope*-Interviews enthält die Frage nach dem *Haushaltsnettoeinkommen*. Vorgegeben sind Einkommensklassen mit der oberen Kategorie „7.500 DM und mehr“. 81% der Befragten haben zur Einkommensfrage eine Angabe gemacht. Von diesen liegen 6% in der oberen Einkommensklasse.

(3) *Ein datenschutzrechtlich einwandfreies Verfahren der Wiederbefragung*

Die Telefonnummer der Befragten darf nur mit der Einwilligung der Befragten gespeichert werden. Um eventuelle Wiederbefragungen durchführen zu können, holt Infratest am Ende

jedes Interviews diese Einwilligung ein. Von den 87.000 Personen, die im Jahr 2000 im Rahmen von InfraScope befragt wurden, gaben 75.000 ihre Einwilligung zur Speicherung der Telefonnummer.

Die Abgrenzung der interessierenden Teilpopulation – hier: der oberen Einkommensgruppe – für eine erneute Befragung erfolgt im anonymisierten Datenbestand. Darüber hinaus wurde von der Vorabinformation zum Einkommen in keiner Weise Gebrauch gemacht – weder den Interviewern noch den Befragten gegenüber. Die erneute Ansprache der Befragungshaushalte erfolgt völlig neutral, wie bei jedem anderen Verfahren einer Zufallsauswahl von Haushalten. Wie beim ursprünglichen Interview, so erhalten die Befragten auch beim erneuten Interview eine datenschutzrechtliche Information, in der auf die Befragungsthemen und auf die Freiwilligkeit der Teilnahme hingewiesen wird.

(4) *Ein Verfahren zur Ermittlung der Haushaltsadressen*

Aus der ursprünglichen Telefonbefragung liegt nur die Telefonnummer der Befragten vor. Wenn die erneute Befragung mündlich-persönlich erfolgen soll, wie es im Rahmen des SOEP geschieht, müssen zuvor Name und Adresse der Befragungshaushalte ermittelt werden.

Im vorliegenden Fall erfolgte dies über ein telefonisches Kontaktinterview, in dem die Befragten eine erste Information über die angestrebte Befragung erhalten und um Angabe ihrer Adresse gebeten werden, damit ihnen genaueres Informationsmaterial zugesandt werden kann. Nur wenn die angesprochenen Personen durch Angabe ihrer Adresse grundsätzliches Interesse signalisieren, werden sie in die weiteren Verfahrensschritte einbezogen.

In der Testerhebung konnten auf diese Weise für 63% der eingesetzten Telefonnummern Namen und Adresse ermittelt werden.² Erst damit ist der Punkt erreicht, an dem eine Stichprobenbasis für die angestrebte Befragung im Rahmen des SOEP gegeben ist.

1.3 **Durchführung der Befragung**

Die Befragungshaushalte der Sonderstichprobe „Einkommensstarke Haushalte“ (High Income) erhalten dasselbe Fragenprogramm wie die der anderen SOEP-Stichproben. In der Testerhebung im Sommer 2001 wurde die vorläufige Fassung des Haushalts- und des Personenfragebogens für 2002 verwendet.³ Anforderungen und Regeln entsprechen im übrigen dem SOEP-Standard.

² Vgl. 1. *Zwischenbericht*, a.a.O., S. 14.

³ Auf diese Weise konnte die Testerhebung zugleich als Pretest für das SOEP 2002 dienen, was insbesondere für die Fragen zum Thema „Vermögen“ wichtig und von großem Nutzen war. Schlussfolgerungen zur Gestaltung des Fragebogens sind im 2. *Zwischenbericht*, a.a.O., S. 8 ff. dargestellt.

Als Bruttostichprobe wurden 312 Haushaltsadressen eingesetzt, die nach dem oben beschriebenen Verfahren ermittelt waren. Die Interviewer konnten im Rahmen einer Feldzeit von etwa 8 Wochen 66% dieser Haushalte zur Teilnahme gewinnen. Die folgende Datenprüfung zeigte allerdings, dass das aktuelle Haushalts-Nettoeinkommen, wie es in der jetzigen Befragung angegeben wurde, nicht immer den angestrebten Schwellenwert von 7.500 DM erreichte. Haushalte unterhalb dieses Schwellenwerts wurden als ungültig behandelt und nicht in die Nettostichprobe übernommen (16%). Der an das DIW übergebene Datenbestand über „einkommensstarke Haushalte“ betrug 157 Haushalte, das sind 50% der eingesetzten Bruttostichprobe.

Dieses Ergebnis ist durchaus zufriedenstellend und entspricht etwa dem erreichten Feldergebnis in einer allgemeinen Bevölkerungsstichprobe. Ein Teil der nicht befragungsbereiten Haushalte ist allerdings bereits in der vorangegangenen Stufe der telefonischen Kontaktinterviews herausgefallen. Wie oben dargestellt, konnten dort 63% der eingesetzten Telefonnummern erfolgreich bearbeitet werden. Beide Verfahrensstufen zusammengenommen beträgt die in der Testerhebung erreichte Ausschöpfung der Ausgangsstichprobe 36%.

1.4 Validierung der Stichprobe

Das gewählte Verfahren der Stichprobenbildung ermöglicht weitreichende Nonresponse-Analysen zur Validierung der Stichprobenqualität. Damit kann überprüft werden, ob die Ausfälle in den verschiedenen Stufen der Befragung zu systematischen Verzerrungen der Stichprobenstruktur führen oder ob die Ausfälle weitgehend zufällig verteilt und damit ergebnisneutral sind.

Die Validierung wird dadurch möglich, dass für jeden Fall der Bruttostichprobe Vorinformationen aus der Ursprungsbefragung vorliegen, nämlich die Standarddemographie aus dem InfraScope-Telefoninterview. Auf dem „Bruttoband“, das Infratest dem DIW zusätzlich zum „Nettoband“ übergibt, sind die Merkmale der Standarddemographie für jeden einzelnen Fall enthalten.

Das Bruttoband der Testerhebung umfasst diejenigen 700 Fälle, die – wie oben beschrieben – die Ausgangsbasis für die Befragung darstellten. Tabelle 1 zeigt die Merkmalsverteilungen in der eingesetzten Bruttostichprobe (n = 700) im Vergleich zur realisierten Nettostichprobe (n = 157). Beide Verteilungen beziehen sich auf die im ursprünglichen Telefoninterview gemachten Angaben der Befragten.

Die Abweichungen in der Struktur der Nettostichprobe gegenüber der Bruttostichprobe sind erfreulich gering. So bleibt als Ergebnis der Testerhebung festzuhalten:

- Einkommensstarke Haushalte aus einer Zufallsstichprobe der Bevölkerung können zu gut einem Drittel für die Teilnahme am SOEP gewonnen werden.
- Die Ausfälle führen zu keinen Stichprobenverzerrungen und sind somit unter Qualitätsgesichtspunkten unbedenklich.
- Das gewählte Stichproben- und Erhebungsdesign hat sich demnach bewährt und kann im Grundsatz für eine Haupterhebung im Jahr 2002 zugrunde gelegt werden.

Tabelle 1
Einkommensstarke Haushalte:
Strukturvergleich zwischen Bruttostichprobe und realisierter Nettostichprobe

Merkmale	Brutto (n = 700) %	Netto (n = 157) %
<i>Geschlecht</i>		
Männer	62	62
Frauen	38	38
<i>Alter</i>		
bis 20 Jahre	6	3
21 – 30 Jahre	6	4
31 – 40 Jahre	21	19
41 – 50 Jahre	28	34
51 – 60 Jahre	26	27
61 – 70 Jahre	10	8
71 Jahre und älter	3	6
<i>Schulbildung</i>		
einfach	11	10
mittel	28	30
hoch	61	60
<i>Erwerbsstatus</i>		
berufstätig	75	78
in Ausbildung	7	4
Hausfrau	7	5
Rentner	9	11
<i>Stellung im Beruf 1)</i>		
Arbeiter	2	2
einfache/qualifizierte Angestellte	15	15
hochqualifizierte/leitende Angestellte	28	25
einfache/mittlere/gehobene Beamte	4	7
höhere Beamte	4	6
Landwirte	0	0
Freiberufler	6	6
Selbständige	14	15
<i>Haushaltsgröße</i>		
1 Person	7	6
2 Personen	40	40
3 Personen	21	19
4 Personen	23	22
5 Personen und mehr	9	13

1) Heutige oder – falls nicht mehr erwerbstätig – frühere Stellung

2. Optionen für die Haupterhebung im Jahr 2002

2.1 Ziele

Das SOEP wurde im Laufe der Zeit mehrfach durch Stichprobenaufstockungen erweitert. Dabei handelte es sich zunächst um Stichproben zur Einbeziehung neuer Grundgesamtheiten:

1990: Stichprobe C / Haushalte Ost

1995: Stichprobe D / Zuwanderer.

Die nächsten Aufstockungen dienen der Erhöhung des Stichprobenumfangs:

1998: Stichprobe E / Querschnitt 1998 (n = ca. 1.000)

2000: Stichprobe F / Querschnitt 2000 (n = ca. 6.000)

Künftig ist angestrebt, den Stichprobenumfang des SOEP durch regelmäßige Aufstockungen längerfristig zu stabilisieren, und zwar bei insgesamt mindestens 12.000 Befragungshaushalten. Dafür sollen im zweijährigen Abstand Aufstockungsstichproben im Umfang von 1.000 bis 2.000 Haushalten hinzukommen. Die jeweilige Grundgesamtheit der Aufstockungsstichproben ist vorab nicht festgelegt: Es können entweder Querschnittsstichproben der Wohnbevölkerung insgesamt sein oder Stichproben einer bestimmten Teilpopulation, die damit überproportional im Stichprobensystem des SOEP repräsentiert werden könnte.

Bestandteil dieser längerfristigen Planungsperspektive ist es, das SOEP im Jahr 2002 um eine neue Stichprobe von 2.000 Haushalten zu erweitern. Die vorgesehene Erweiterung im Jahr 2002 bildet den Rahmen für die gewünschte Erweiterung des SOEP um eine Sonderstichprobe „Einkommensstarke Haushalte“.

Die Vorstellungen zum Stichproben- und Erhebungsdesign bewegen sich dabei zwischen zwei alternativen Modellen:

Modell A:

Die gesamte Stichprobenaufstockung des Jahres 2002 wird für die überproportionale Abbildung der Haushalte im oberen Einkommensbereich verwendet. Wenn möglich, soll die High Income-Stichprobe demnach einen Umfang von 2.000 Haushalten haben.

Modell B:

Die Stichprobenaufstockung des Jahres 2002 wird aufgeteilt auf eine bevölkerungsproportionale Querschnittsstichprobe und eine High Income-Sonderstichprobe. Letztere soll einen Mindestumfang von 700 Haushalten haben; die Bevölkerungsstichprobe hätte in diesem Fall einen Umfang von 1.300 Haushalten.

Ob das Modell A oder das Modell B gewählt wird, ist einerseits eine Frage inhaltlicher Präferenzen. Es ist aber auch eine Frage der Machbarkeit, also der realisierbaren Größe einer High Income-Stichprobe unter gegebenen methodischen Prämissen und Rahmenbedingungen. Diese Frage zu klären, war Aufgabe der Feasibility Study und der in diesem Rahmen durchgeführten Testerhebung.

Die Antworten hängen auch von methodischen und inhaltlichen Anforderungen ab, die an eine High Income-Stichprobe gestellt werden. Die wesentlichen zwei Fragen sind,

- (a) wie hoch die Einkommensschwelle für die Abgrenzung „einkommensstarker Haushalte“ angesetzt wird
- (b) ob eine einheitliche Schwelle für alle Haushalte angestrebt wird oder ob nach verschiedenen Regionen (Ost / West) oder nach Haushaltsgrößen (Äquivalenzeinkommen) differenziert werden soll. Hierzu sind Entscheidungen erforderlich, bei denen einerseits wiederum inhaltliche Kriterien, andererseits aber auch Gesichtspunkte der Machbarkeit zu berücksichtigen sind.

2.2 Die Stichprobenbasis

Wie bei der Testerhebung, so besteht auch bei der geplanten Haupterhebung die **Ausgangsstichprobe** aus der großen Zufallsstichprobe der Bevölkerung ab 14 Jahren, die im Rahmen der telefonischen Mehrthemenbefragung **InfraScope** interviewt wurden. Während für die Testerhebung die kumulierte Jahresstichprobe des Jahres 2000 verwendet wurde, ist für die Haupterhebung die InfraScope-Jahresstichprobe 2001 die Basis. Bei Bedarf spricht allerdings nichts dagegen, auch noch auf Teile der Vorjahresstichprobe 2000 zurückzugreifen.

Wir stellen im Folgenden wesentliche Punkte des Mengengerüsts und der Strukturen in der InfraScope-Stichprobe 2001 dar. Die Auswertungen stützen sich auf den bisher verfügbaren Bestand von Interviews in den Monaten Januar bis Juli 2001, das ist eine Fallzahl von rd. 58.000 Befragten. Für Gesamtzahlen des Jahres 2001 wird auf dieser Basis hochgerechnet.

Das Mengengerüst für eine Stichprobenbildung

Tabelle 2 zeigt die Eckwerte, die für eine Stichprobenbildung einkommensstarker Haushalte auf Basis der InfraScope-Befragungen relevant sind, im Vergleich zwischen InfraScope 2000 und InfraScope 2001. Die verwendeten Kennzahlen sind im *1. Zwischenbericht* ausführlicher kommentiert und werden hier lediglich aktualisiert.

Tabelle 2:
Die Ausgangsbasis: Telefonische Bevölkerungstichproben

	InfraScope 2000 Ist	InfraScope 2001 Prognose
Anzahl durchgeführter Telefoninterviews im Verlauf des jeweiligen Jahres	87.000	92.000
RLD-Methode	Nein	Ja
Zustimmung zur Speicherung der Telefon-Nr. zwecks eventueller Wiederbefragung	86%	87%
Haushaltsnettoeinkommen		
• Angaben vorhanden	81%	84%
• Einkommen ab 7.500 DM	5,5%	6,5%
Fallzahl der Haushalte in der oberen Einkommens- kategorie und Erlaubnis zur Wiederbefragung	4.277	5.380
darunter: für Testerhebung verwendet	700	-

Die Ausgangsstichprobe der Bevölkerung aus dem Jahr 2001 wird mit ca. 92.000 Fällen noch etwas größer sein als die des Vorjahres. Die statistische Anlage der Stichprobe ist identisch wie im Vorjahr, das Auswahlverfahren jedoch durch Einbeziehung der RLD-Methode verbessert.⁴ Durch RLD („Randomised Last Digits“) werden auch diejenigen Telefonhaushalte in die Auswahl einbezogen, die ihre Telefonnummern nicht in das Telefonverzeichnis haben eintragen lassen. Der Anteil dieser Haushalte variiert nach Regionen und Bevölkerungsgruppen erheblich; in einzelnen neuen Bundesländern beträgt er über 30%.

Man könnte vermuten, dass es sich bei den nicht im Telefonverzeichnis eingetragenen Anschlüssen um Haushalte handelt, die bei einer eventuellen Telefonbefragung besonders wenig auskunftsbereit sind. Die vorliegenden Daten stützen diese Vermutung jedoch nicht:

- Nach Umstellung auf das RLD-Verfahren ist die Quote der Befragten, die ihre Zustimmung zur Speicherung der Telefonnummer und zu einer eventuellen Wiederbefragung geben, nicht etwa gesunken, sondern um einen Prozentpunkt auf nunmehr 87% gestiegen.
- Der Anteil der Befragten, die eine Angabe zu ihrem Haushaltsnettoeinkommen machen, ist ebenfalls gestiegen und liegt nun bei 84%.

Der Anteil der Haushalte mit einem Nettoeinkommen oberhalb der Schwelle von 7.500 DM ist um einen Prozentpunkt auf nunmehr 6,5% gestiegen.⁵ Darin schlägt sich, so ist zu vermuten, die Einkommensentwicklung vom Jahr 2000 auf das Jahr 2001 nieder.

⁴ Vgl. 1. Zwischenbericht, S. 5 ff. Die RLD-Methode, wie sie in InfraScope seit Anfang 2001 eingesetzt wird, ist dort bereits beschrieben (S. 6)

⁵ Unbereinigter Wert, d.h. die Basis enthält auch die KA-Fälle.

Für die Bildung einer Stichprobe einkommensstarker Haushalte wirken alle genannten Punkte positiv in dem Sinne, dass sie die Fallzahl verfügbarer High-Income-Haushalte in der Ausgangsstichprobe erhöhen. Nach derzeitiger Prognose werden – gegenüber 4.277 Fällen in der Vorjahresstichprobe – aus InfraScope 2001 mehr als 5.300 Fälle vorliegen.

Abbildung hoher Einkommen

Zur Abgrenzung von „einkommensstarken“ Haushalten lag aus der InfraScope-Standarddemographie bisher lediglich die Information vor, wie viele Haushalte ein Nettoeinkommen oberhalb von 7.500 DM haben. Die Verteilung auf differenziertere Einkommensklassen oberhalb dieses Schwellenwerts wird erst seit Mitte März 2001 durch eine Ergänzungsfrage ermittelt. Tabelle 3 zeigt die Verteilung in der InfraScope-Stichprobe 2001 im Vergleich zum SOEP und zum Mikrozensus (jeweils Daten aus dem Jahr 2000).

Die Verteilungen in den drei Datenquellen stimmen erstaunlich gut überein. Die Einkommensschwelle „ab 7.500 DM“ definiert die oberen 7% der Einkommenschichtung. Mit zunehmender Anhebung der Einkommensschwelle nimmt der erfasste Bevölkerungsanteil rasch ab. Der Schwellenwert „ab 10.000 DM“ grenzt in der Einkommenschichtung die oberen 2,5% der Haushalte ab, der Schwellenwert „ab 12.000 DM“ die oberen 1% - 1,5% der Haushalte.

Tabelle 3:
Anteil von Haushalten mit hohem Einkommen in verschiedenen Datenquellen

	InfraScope		SOEP	MZ
	2000	2001	2000 ¹⁾	2000
	%	%	%	%
Keine Angabe zur Höhe des Haushaltseinkommens	18,8	16,0	7,2	4,8
Neue Basis: Haushalte mit Einkommensangabe	100	100	100	100
Haushaltsnettoeinkommen				
7.500 DM und höher	6,2	7,7	6,8	7,0
8.000 DM und höher		5,5	5,7	5,6
10.000 DM und höher		2,7	2,4	2,6
12.000 DM und höher		1,3	1,0	1,4

1) Gesamtstichprobe A – F, ungewichtet

Strukturdaten der einkommensstarken Haushalte

Die soziodemographische Struktur der Haushalte im oberen Einkommensbereich weicht von derjenigen der Bevölkerung insgesamt ab. Tabelle 4 gibt eine Zusammenstellung nach verschiedenen Strukturmerkmalen. Dabei wird unterschieden nach einer weiteren Abgrenzung (Schwellenwert 7.500 DM) und nach einer engeren Abgrenzung (Schwellenwert 10.000 DM). Innerhalb der weiter abgegrenzten Gruppe ist zusätzlich danach unterschieden, ob die Einwilligung zu einer Wiederbefragung gegeben wurde oder nicht. Die Ergebnisse lassen sich in zwei Hauptpunkten zusammenfassen:

- Nonresponse Effekte

Unter den Personen mit höherem Haushaltsnettoeinkommen verweigern noch weniger ihre Einwilligung zu einer eventuellen Wiederbefragung als in der Gesamtstichprobe, nämlich nur 10% gegenüber 13% insgesamt. Diese 10% Nonresponse-Fälle unterscheiden sich in ihrer Struktur praktisch nicht von den wiederbefragungsbereiten Personen. Der für eine erneute Befragung einsetzbare Teil der Ausgangsstichprobe – das sind ohnehin 90% – weist damit strukturell keine nennenswerten Verzerrungen auf.

- Struktureffekte der engeren oder weiteren Abgrenzung einkommensstarker Haushalte.

Die enger abgegrenzte High-Income-Bevölkerung (Schwellenwert 10.000 DM) unterscheidet sich von der weiter abgegrenzten High-Income-Bevölkerung (Schwellenwert 7.500 DM) in Bezug auf ihre soziodemographische Zusammensetzung nur graduell und in wenigen Strukturmerkmalen. Einzelne Personenkreise mit höheren Einkommenschancen sind in der engeren Abgrenzung („die oberen 2% - 3%“) noch relativ stärker vertreten als in der weiteren Abgrenzung („die oberen 7%“). Das gilt etwa

- für die Altersgruppe der 51- bis 60-Jährigen
- für Personen mit abgeschlossenem Hochschulstudium
- für Freiberufler und Selbständige
- für Personen in größeren Haushalten („Mehrverdiener-Haushalte“).

Die in der Tabelle ausgewiesenen Strukturen ergeben sich, wenn eine einheitliche Einkommensschwelle für alle Haushalte verwendet wird. Die Merkmalsverteilungen zeigen, welche Struktureffekte diese Abgrenzung „einkommensstarker Haushalte“ hat. Auf dieser Grundlage kann man prüfen und diskutieren, ob die Effekte im Hinblick auf die Analyseziele sinnvoll und erwünscht sind. Das betrifft insbesondere zwei Punkte:

- Anteil der neuen Länder

Die neuen Länder haben in der Stichprobe einkommensstarker Haushalte einen Anteil von lediglich 9%. Das entspricht sicherlich ihrer Position in der gesamtdeutschen Einkommensverteilung. Eine Analyse der Daten speziell für die neuen Länder wird damit aber wegen geringer Fallzahlen rasch an Grenzen stoßen. Zu ändern wäre dies nur, indem die verwendete Einkommensschwelle für die neuen Länder niedriger angesetzt würde als für die alten Länder.

- Anteil der Ein-Personen-Haushalte

Ein-Personen-Haushalte haben in der Stichprobe einkommensstarker Haushalte einen Anteil von lediglich 3%. Das liegt einerseits an der Struktur des Personenkreises, der in Ein-Personen-Haushalten lebt (z.B. Rentnerinnen, junge Singles), andererseits daran, dass es hier nicht mehrere Verdienere geben kann.

Unter Bedarfsgesichtspunkten könnte man argumentieren, dass eine alleinlebende Person mit einem Nettoeinkommen von 6.000 DM besser dasteht als eine Familie mit 8.000 DM. Erstere wäre in der High Income-Stichprobe nicht vertreten, letztere dagegen ja. Will man diesen Effekt vermeiden, müsste man irgendeine Form von *Äquivalenzeinkommen* zugrunde legen.

Unter inhaltlichen Gesichtspunkten würde dies bedeuten, dass man sich nicht an der Einkommensverteilung als solcher orientiert, sondern an einem Lebenslagenkonzept der finanziellen Ressourcen, also einem spiegelbildlichen Ansatz zur Armutsforschung.⁶

Unter Machbarkeitsgesichtspunkten ist zu berücksichtigen, dass die verfügbare Stichprobenbasis eine Berechnung von *Äquivalenzeinkommen* nicht zulässt.⁷ Möglich wäre höchstens, für Ein-Personen-Haushalte eine niedrigere Einkommensschwelle zu verwenden als für Mehr-Personen-Haushalte.

⁶ Die Bildung von Einkommensgruppen anhand von *Äquivalenzeinkommen* kann in der Regel allerdings erst in der Datenanalyse erfolgen, nicht bereits in der Stichprobenbildung. Eine Ausnahme ist das von Infratest durchgeführte *Niedrigeinkommenspanel (NIEP)*, das Bedarfsgesichtspunkte bereits in der Stichprobenbildung berücksichtigte. Ein solches Vorgehen erfordert ein spezielles, aufwändiges „Screening“.

⁷ Dies hat zwei Gründe: (a) Die Einkommensinformation liegt nur in groben Größenklassen vor. (b) Die Information zur Haushaltszusammensetzung beschränkt sich auf die Haushaltsgröße, ohne Differenzierung nach Kindern und Erwachsenen.

Tabelle 4
Strukturdaten von Personen in Haushalten mit hohem Einkommen

		Schwellenwert für die Abgrenzung		
		7.500 DM nicht wieder- befragungs- bereit %	7.500 DM wieder- befragungs- bereit %	10.000 DM wieder- befragungs- bereit %
Regionen:	Alte Länder	93	91	91
	Neue Länder	7	9	9
Geschlecht:	Männer	64	63	65
	Frauen	36	37	35
Alter:	bis 20 Jahre	10	8	7
	21 – 30 Jahre	12	12	11
	31 – 40 Jahre	19	17	16
	41 – 50 Jahre	24	26	26
	51 – 60 Jahre	24	24	28
	61 – 70 Jahre	6	9	9
	71 Jahre und älter	5	4	3
Schulbildung:	einfach	11	12	10
	mittel	28	25	22
	hoch	61	63	68
Abgeschlossenes Studium		42	42	46
Stellung im Beruf:	Un-/Angelernte	1	1	0
	Facharbeiter	3	1	1
	Vorarbeiter etc.	0	0	0
	Einfache Angestellte	3	3	2
	Qualifizierte Ang.	10	11	10
	Hochqualifizierte Ang.	8	9	8
	Leitende Angestellte	16	15	15
	Einfache/mittlere Beamte	2	1	0
	Gehobene Beamte	4	5	4
	Höhere Beamte	5	6	6
	Landwirte	0	0	0
	Freiberufler	5	5	9
	Selbständige	15	14	21
Mithelfende	1	0	0	
Haushaltsgröße:	1 Person	5	3	3
	2 Personen	34	33	33
	3 Personen	23	23	20
	4 Personen	23	26	27
	5 Personen und mehr	15	15	17

Quelle: Infratest Sozialforschung, InfraScope 2000/2001, personengewichtet

2.3 Zwei alternative Modelle für die Haupterhebung

In der SOEP-Stichprobe des Jahres 2002 – einschließlich der großen Aufstockungsstichprobe F aus dem Jahr 2000, aber vor einer gezielten Aufstockung um High Income-Haushalte – werden als Erwartungswert etwa 700 – 800 Haushalte mit einem Haushaltsnettoeinkommen von 7.500 DM und mehr zu finden sein. Dies ist für viele Analysen bereits eine recht gute Basis. Haushalte mit einem Nettoeinkommen ab 10.000 DM werden dagegen nur mit einer Fallzahl von rd. 250 Fällen vertreten sein. Dies ist für viele Analysen eine zu schmale Basis.

Von daher konzentriert sich der Wunsch nach einer Stichprobenaufstockung auf die Haushalte im Einkommensbereich ab 10.000 DM. Wir prüfen im Folgenden, wie dieses Ziel am besten zu erreichen ist.

Grundsätzlich sind zwei Modelle denkbar, die auf ihre Effektivität zu prüfen sind.

Modell A:

Es wird eine möglichst große High Income-Stichprobe mit einer relativ niedrigen Einkommensschwelle gebildet.

Konkret prüfen wir die Machbarkeit einer Nettostichprobe von 2.000 Haushalten mit einem Haushaltsnettoeinkommen von mindestens 7.500 DM. Diese Stichprobe würde nach den vorliegenden Verteilungsinformationen (siehe Tabelle 3) zu einem guten Drittel Haushalte mit einem Nettoeinkommen von mindestens 10.000 DM enthalten. Es würden also rd. 700 Haushalte in dieser oberen Einkommensklasse hinzugewonnen.

Zusammen mit den im SOEP bereits vorhandenen Fällen ergäbe das

- eine Fallzahl von fast 1.000 Haushalten oberhalb der Schwelle von 10.000 DM, also im Bereich der oberen 2% - 3% der Einkommensschichtung
- und eine Fallzahl von rd. 2.700 Haushalten oberhalb der Schwelle von 7.500 DM, also im Bereich der oberen 7% der Einkommensschichtung.

Modell B:

Die Aufstockung wird begrenzt auf Haushalte im oberen Einkommensbereich ab 10.000 DM Nettoeinkommen.

Zu prüfen ist, welche Fallzahl bei diesem Vorgehen aus der bestehenden Stichprobenbasis, wie sie oben beschrieben ist, realisiert werden könnte.

Prüfung von Modell A

In der Testerhebung konnten von den Haushalten der Ausgangsstichprobe letztlich 36% für die Teilnahme am SOEP gewonnen und nach Prüfung ihres Haushaltseinkommens in den Datenbestand der High-Income-Stichprobe übernommen werden.

Für die Haupterhebung setzen wir die erwartete Erfolgsquote etwas niedriger an. Grund ist, dass in der Testerhebung wegen der geringen Fallzahl regionale Klumpen gebildet werden konnten, die SOEP-erfahrenen Interviewern eine Bearbeitung mit relativ kurzen Wegstrecken ermöglichten. Dies wird in der Haupterhebung deutlich schwieriger werden. Die Adressen stammen aus der völlig ungeklumpten Telefonstichprobe und sind in ihrer regionalen Verteilung völlig unabhängig von der regionalen Verteilung der SOEP-Interviewer.

Für die Modellrechnungen setzen wir die zu erwartende Ausschöpfung der Ausgangsstichprobe daher niedriger an, und zwar mit 30%.

Für einen Stichprobenumfang von netto 2.000 Haushalten benötigt man dann eine Ausgangsstichprobe von 6.660 Haushalten. Diese Fallzahl ist in der gegebenen Stichprobenbasis vorhanden, wenn neben der InfraScope-Stichprobe von 2001 auch eine gewisse Zahl von Adressen aus der Vorjahresstichprobe 2000 verwendet wird.

Es werden eingesetzt:

- alle 5.300 erwarteten Telefonnummern aus der Jahresstichprobe 2001
- und ergänzend eine Zufallsauswahl von 1.360 Fällen aus der Jahresstichprobe 2000.⁸

Das Modell A ist auf diese Weise realisierbar.

Prüfung des Modells B

Die Zahl der Haushalte mit einem Nettoeinkommen oberhalb von 10.000 DM beträgt etwa ein Drittel der Zahl der Haushalte mit einem Nettoeinkommen oberhalb von 7.500 DM. Dementsprechend geringer ist die Zahl der verfügbaren Adressen bzw. Telefonnummern in der Ausgangsstichprobe.

Erschwerend kommt hinzu, dass die differenzierte Information über höhere Einkommen erst ab Mitte März 2001 erhoben wurde. Für die Zeit davor können Haushalte mit einem Nettoeinkommen ab 10.000 DM in der Ausgangsstichprobe nicht identifiziert werden.

Der Erwartungswert für die Zahl verfügbarer Fälle in der Ausgangsstichprobe nach Modell B beträgt rd. 1.400 Haushalte. Bei der zugrundegelegten erwarteten Erfolgsquote für die Gewinnung dieser Haushalte als Teilnehmer im SOEP beträgt der Erwartungswert für die realisierbare Nettostichprobe rd. 400 – 450 Haushalte.

⁸ Unter Ausschluss der 700 Fälle, die bereits für die Testerhebung verwendet wurden.

Dies ist deutlich weniger als in Modell A, in dem der Erwartungswert für die Fallzahl von neu gewonnenen Haushalten mit einem Nettoeinkommen von mindestens 10.000 DM bei 700 Fällen liegt.

Im Modell B müsste die vorgesehene Aufstockung des SOEP mit insgesamt 2.000 Fällen auf zwei Stichproben verteilt werden:

- (1) die High-Income-Stichprobe mit 400 – 450 Fällen, die auf dem geschilderten Weg gewonnen würde, sowie
- (2) eine Querschnittsstichprobe der Wohnbevölkerung mit 1.600 Fällen, die nach dem Verfahren der Aufstockungsstichprobe F des Jahres 2000 gewonnen würde.

Schlussfolgerung und Empfehlung

- (1) Die Feasibility Study hat gezeigt, dass eine gezielte Aufstockung des SOEP um Haushalte im oberen Einkommensbereich auf der Basis von Random-Verfahren möglich ist.
- (2) Die für das Jahr 2002 vorgesehene Aufstockung des SOEP um 2.000 Haushalte insgesamt könnte entweder ganz auf Haushalte im oberen Einkommensbereich konzentriert werden (Modell A) oder auf zwei Teilstichproben aufgeteilt werden, von denen die kleinere eine High-Income-Stichprobe wäre (Modell B). Im ersten Modell würde die High Income-Stichprobe die oberen 7%, im zweiten Modell die oberen 2% - 3% der Einkommensschichtung abbilden.
- (3) Beide Modelle sind realisierbar. Die Entscheidung kann sich also primär an inhaltlichen Kriterien orientieren. Ein wichtiges inhaltliches Kriterium ist der Bezug zum Themenschwerpunkt „soziale Sicherung und Vermögen“, der im SOEP 2002 im Vordergrund steht.

Die Bedeutung von Vermögen und Vermögensbildung für die soziale Sicherung ist eine Fragestellung, die nicht nur für die „Reichen“ im engeren Sinne bedeutsam ist, sondern für eine breitere Schicht von Haushalten mit guter Einkommenssituation. Unter diesem Gesichtspunkt erscheint Modell A ertragreicher als Modell B.

Die Vermögensbildung in ihrer Funktion als Alterssicherung hat durch die Rentenreform einen neuen sozialpolitischen Stellenwert erhalten. Durch die Reform ist dabei eine Entwicklung angestoßen, die in den kommenden Jahren wirksam werden soll. Insofern ist es auch richtig, wenn die Stichprobenerweiterung im oberen Einkommensbereich nicht auf eine einmalige Querschnittserhebung beschränkt bleibt, sondern als Längsschnittbefragung im Rahmen des SOEP über die kommenden Jahre fortgeführt wird.

- (4) Unter methodischen und praktischen Gesichtspunkten ist ebenfalls dem Modell A der Vorzug zu geben. Im Modell B wäre die vorgesehene Aufstockung auf zwei relativ kleine Teilstichproben verteilt, die methodisch unterschiedlich zu behandeln sind. Das hat erschwerende Auswirkungen
- für die Datenerhebung
 - für den Hochrechnungsrahmen
 - und schließlich auch für die Datenanalyse.

Unter all diesen Gesichtspunkten ist es günstiger, die vorgesehene Aufstockung des SOEP im Jahr 2002 mit einem einheitlichen Stichprobenansatz zu realisieren, wie er in Modell A vorgeschlagen wird.