

Dipl.-Sozialwissenschaftlerin Urszula Sikorski, Dipl.-Mathematikerin Birgit Horneffer,
Dipl.-Soziologin Birgit Kuchler

Die Längsschnittdaten von EU-SILC – Datenstruktur und Hochrechnungsverfahren

Im Jahr 2004 führte Eurostat, das Statistische Amt der Europäischen Gemeinschaften, die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) ein, in der neben jährlichen Querschnittsinformationen zu Einkommensarmut, Einkommensungleichheit und Lebensbedingungen auch Längsschnittinformationen über Veränderungen in den Privathaushalten erhoben werden. Deutschland beteiligt sich seit dem Jahr 2005 an dieser Gemeinschaftsstatistik und erhebt die Daten schriftlich in einer eigenen Haushaltsbefragung: *LEBEN IN EUROPA*. Die deutsche EU-SILC-Erhebung umfasst etwa 14 000 Haushalte, die über mehrere Jahre hinweg teilnehmen; die Teilnahme ist freiwillig. EU-SILC ist die einzige Haushaltserhebung der amtlichen Statistik, mit der soziodemografische und sozioökonomische Veränderungen in den Privathaushalten über einen längeren Zeitraum (4 Jahre) untersucht werden können.

Inhalt (Zielvariablen) und Datenstruktur der zu erhebenden Längsschnittinformationen sind in einer europäischen Rahmenverordnung festgelegt. Inzwischen liegen die Längsschnittdaten für den Zeitraum 2005 bis 2007 vor. Die Längsschnittdaten für den Zeitraum 2005 bis 2008 werden derzeit aufbereitet und sind im März 2010 an Eurostat zu liefern.

Im ersten Teil dieses Beitrags wird die Datenstruktur der Längsschnittdaten vorgestellt. Anschließend werden das Paneldesign der deutschen EU-SILC-Erhebung und die verschiedenen Maßnahmen zur Qualitätssicherung erläutert.

Ein Überblick über das Hochrechnungsverfahren wird im dritten Teil des Beitrags gegeben. Der Beitrag schließt mit empirischen Ergebnissen zur Ausschöpfung und zur Panelmortalität für die Erhebungsjahre 2005 bis 2008. Danach nahm knapp die Hälfte der im Jahr 2005 aus der Dauerstichprobe befragungsbereiter Haushalte gezogenen Haushalte bis zum Jahr 2008 kontinuierlich an der EU-SILC-Erhebung teil.

Rechtsgrundlagen

Mit den Längsschnittdaten von EU-SILC soll die Armutsgefährdung in den Mitgliedstaaten der Europäischen Union (EU) dauerhaft gemessen werden: „Zur Untersuchung zeitlicher Veränderungen bei Einzelpersonen und zur Berechnung des Indikators für den sozialen Zusammenhalt ‚Quote der dauerhaften Armutsgefährdung‘ werden alle in ihrem privaten Haushalt verbleibenden oder in einen anderen privaten Haushalt innerhalb des nationalen Erhebungsgebiets umziehenden Stichprobenpersonen für einen Zeitraum („Paneldauer“) von mindestens vier Jahren in der EU-SILC-Stichprobe belassen.“¹⁾

Rechtsgrundlage für die Längsschnittdaten ist die europäische Rahmenverordnung für die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC).²⁾ Der Begriff Längsschnittdaten ist darin wie folgt definiert:

1) Verordnung (EG) Nr. 1982/2003 der Kommission vom 21. Oktober 2003 zur Durchführung der Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates für die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) im Hinblick auf die Regeln für die Stichprobenauswahl und die Weiterbefragung (Amtsbl. der EU Nr. L 298, S. 29, Anhang, Abschnitt 6).

2) Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates vom 16. Juni 2003 für die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) (Amtsbl. der EU Nr. L 165, S. 1), geändert durch die Verordnung (EG) Nr. 1553/2005 des Europäischen Parlaments und des Rates vom 7. September 2005 (Amtsblatt der EU Nr. L 255, S. 6).

„Längsschnittdaten“ sind einschlägige Daten auf der Ebene von Einzelpersonen, die sich mit der Zeit verändern und regelmäßig über einen bestimmten Zeitraum beobachtet werden. Die Längsschnittdaten können entweder aus einer Querschnitterhebung mit Rotationsstichproben stammen, bei der einmal ausgewählte Personen weiterbefragt werden, oder aus einer reinen Panelerhebung; sie können mit Registerdaten kombiniert werden.“³⁾

Die zur Rahmenverordnung gehörenden Durchführungsverordnungen legen weitere Details zum Inhalt (Zielvariablen) und zur Datenstruktur der Längsschnittdaten fest. Darüber hinaus enthalten die Verordnungen methodische Vorgaben

zur Erhebungsorganisation und Erhebungsdokumentation, beispielsweise

- zur Stichprobengröße der Längsschnittdaten,
- wie Haushalte und Befragungspersonen weiterverfolgt werden sollen,
- wie soziodemografische Veränderungen in den Haushalten zu erfassen sind,
- wie Abspaltungen von Haushalten zu behandeln sind,
- zum Aufbau der Identifikationsschlüssel,

Übersicht 1: Vergleich der erfassten Bereiche in der Querschnittskomponente und in der Längsschnittkomponente

Einheit (Personen oder Haushalte)	Erfassungsmodus	Sachgebiete	Bereiche	Erfasst im Querschnitts- (X) und/oder Längsschnittbereich (L)
1. Haushaltsdaten (d-file/h-file)				
Haushalt	Datenerhebung bei einem Haushaltsmitglied ab 16 Jahren	Grunddaten	Grunddaten des Haushalts einschließlich Grad der Verstärkerung	X, L
		Einkommen	Haushaltseinkommen insgesamt, Bruttoeinkommenskomponenten auf Haushaltsebene	X, L
		Soziale Ausgrenzung	Wohnungsbezogene und sonstige Zahlungsrückstände	X, L
			Nichtmonetäre haushaltsbezogene Mangelindikatoren einschließlich finanzieller Engpässe, der Schuldenhöhe und der erzwungenen Unterversorgung in Bezug auf grundlegende Bedürfnisse	X, L
			Physisches und soziales Umfeld	X
		Daten über Erwerbstätigkeit	Kinderbetreuung	X
		Wohnverhältnisse	Wohnungstyp, Besitzverhältnisse und Wohnbedingungen	X, L
Ausstattung der Wohnung	X, L			
Kosten der Wohnung	X			
2. Personendaten (r-file/p-file)				
Alle Personen unter 16 Jahren	Datenerhebung bei einem Haushaltsmitglied ab 16 Jahren	Grunddaten	Demografische Daten	X, L
Ehemalige Haushaltsmitglieder			Demografische Daten	X
Alle Personen ab 16 Jahren im Haushalt	Datenerhebung bei sämtlichen Haushaltsmitgliedern ab 16 Jahren (ausnahmsweise Proxyinterview für vorübergehend abwesende oder kranke Personen)	Einkommen	Persönliches Bruttoeinkommen und Bruttoeinkommenskomponenten auf persönlicher Ebene	X, L
		Grunddaten	Persönliche Grunddaten	X, L
			Demografische Daten	X, L
		Bildung	Bildung einschließlich höchster erreichter ISCED-Stufe	X, L
		Daten über Erwerbstätigkeit	Grunddaten über den derzeitigen Beschäftigungsstatus und die derzeitige Haupttätigkeit; bei Arbeitslosen Daten über die letzte Haupttätigkeit	X, L
Grunddaten über den Beschäftigungsstatus im Einkommensbezugszeitraum	X			
Gesamtzahl der im Rahmen einer Zweit-/Drittätigkeit oder weiterer Tätigkeiten geleisteten Arbeitsstunden	X			
Mindestens ein Haushaltsmitglied ab 16 Jahren (Stichprobenperson)	Erhebung persönlicher Daten bei der/den Person(en) (ausnahmsweise Proxyinterview)	Gesundheit	Gesundheit einschließlich Gesundheitszustand und chronischer Krankheiten	X, L
			Zugang zum Gesundheitswesen	X
		Daten über Erwerbstätigkeit	Detaillierte Daten zur Erwerbstätigkeit	X, L
			Bisherige Erwerbstätigkeit	L
		Kalender der Erwerbstätigkeit	L	

3) Siehe Fußnote 2, Artikel 2, Abschnitt h.

- zum Hochrechnungsverfahren der Längsschnittinformationen und
- zum Inhalt der Abschlussqualitätsberichte.

Danach muss Deutschland bei den Längsschnittdaten eine effektive Mindeststichprobengröße von 6 000 Haushalten und 10 500 Personen im Alter ab 16 Jahren erreichen. Termin für die jährliche Lieferung der Längsschnittdaten an Eurostat ist der 31. März des zweiten auf das Erhebungsjahr folgenden Jahres (N + 2).

Aufbau der Längsschnittdaten

Die Längsschnittdaten werden aus den jährlichen EU-SILC-Daten erstellt und entsprechen in ihrem Aufbau und in ihrer Funktion den Querschnittsdaten⁴⁾. Danach sind vier Dateien zu erstellen:

- Haushaltsmetadaten mit Informationen zur Teilnahme und zu Veränderungen des Haushalts,
- Haushaltsdaten,
- Personenmetadaten mit stichprobenrelevanten und Teilnahmeinformationen aus den vorhergehenden Erhebungsjahren der befragten Personen sowie
- Personendaten.

Im Gegensatz zu den Querschnittsdaten bleiben die Identifikationsschlüssel für die Haushalte und Personen zeitlich unverändert. Damit wird es möglich, die Lebensbedingungen ein- und derselben Haushalte über vier Erhebungsjahre hinweg zu beobachten.

Ein Teil der EU-SILC-Längsschnittdaten wird auf der Basis der jährlich erhobenen Querschnittsdaten gebildet.⁵⁾ Im Gegensatz zu den Querschnittsdaten enthalten die Längsschnittdaten jedoch grundsätzlich nur primäre Zielvariablen und keine sekundären Zielvariablen.⁶⁾ Übersicht 1 gibt einen Überblick über die gemeinsame Schnittmenge von primären Zielvariablen in Querschnitts- und Längsschnittdaten.

Haushaltsmetadaten und Haushaltsdaten im Längsschnitt

Die Metadaten der Haushalte enthalten Informationen zur Teilnahmebereitschaft in den jährlichen Bruttostichproben. Zusätzlich werden die Ausfallgründe für bestehende Haushalte und die Herkunft (Abspaltung von einem Ursprungshaushalt usw.) von neuen Haushalten in der Variablen Haushaltsstatus erfasst. Auf diese Weise kann für jede Erhebungswelle die Panelmortalität untersucht werden:

Die Variable Haushaltsstatus hat folgende Ausprägungen:

- (1) Haushalt befindet sich an der gleichen Adresse wie im Vorjahr.
- (2) Haushalt ist umgezogen.
- (3) Haushalt ist nicht mehr als Privathaushalt anzusehen (z. B. weil er in eine Anstalt verzogen ist).
- (4) Haushalt ist ins Ausland verzogen.
- (5) Haushalt existiert aufgrund von Tod der Haushaltsmitglieder nicht mehr.
- (6) Haushalt enthält keine Stichprobenperson mehr.
- (7) Haushalt war nicht erreichbar.
- (8) Haushalt ist ein von einem Ursprungshaushalt abgespaltener Haushalt.
- (9) Haushalt nimmt das erste Mal an EU-SILC teil (Erhebungsjahr).
- (10) Haushalt kam aufgrund einer Fusion von bereits bestehenden EU-SILC-Haushalten zustande.
- (11) Haushalt war im Erhebungsjahr nicht mehr aufzufinden (unbekannt verzogen).

Nach den methodischen Vorgaben von Eurostat werden nicht alle Haushaltsmitglieder bei Auszug aus einem Ursprungshaushalt (Stichprobenhaushalt der ersten Erhebungswelle) weiterverfolgt, sondern nur die Stichprobenpersonen.⁷⁾ Stichprobenpersonen sind alle Mitglieder des Ursprungshaushalts, die in der ersten Erhebungswelle mindestens 14 Jahre alt waren. Die anderen Haushaltsmitglieder – auch die zu einem späteren Zeitpunkt neu hinzukommenden Personen – gelten als sonstige Personen. Je nach Art der Haushaltsveränderungen im Zeitverlauf (Tod, Auszug von Stichprobenpersonen) muss daher für jeden Haushalt geprüft werden, ob der Haushalt noch zur Erhebungsgesamtheit gehört. Ferner können ausgezogene Stichprobenpersonen die Weiterbefragung an ihrer neuen Adresse verweigern. Diese Fälle sind im engen Sinne keine Abspaltungen von einem Ursprungshaushalt, sondern werden als klassische Ausfälle aufgrund der Verweigerung registriert. Haushalte, in denen mindestens eine Stichprobenperson lebt, werden weiter befragt. Lebt keine Stichprobenperson mehr in einem Haushalt, so wird der Haushalt in der nächsten Erhebungswelle nicht weiter befragt. Auch Haushalte, die nicht aufzufinden oder zu erreichen waren oder die Teilnahme verweigerten, werden in den Folgejahren nicht mehr befragt.

4) Siehe hierzu auch Sikorski, U.: „Europäische Panelerhebung über Einkommen und Lebensbedingungen“ in WiSta 8/2009, S. 795 ff.

5) Zielvariablen, die sowohl in der Querschnittsdatenlieferung als auch in der Längsschnittdatenlieferung vorkommen, weisen die gleiche Bezeichnung auf.

6) Sekundäre Zielvariablen sind Zielvariablen, die im Rahmen der sich jährlich ändernden Schwerpunktmodule gebildet werden.

7) Siehe Verordnung (EG) Nr. 1981/2003 der Kommission vom 21. Oktober 2003 zur Durchführung der Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates für die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) im Hinblick auf die Aspekte der Feldarbeit und die Imputationsverfahren (Amtsbl. der EU Nr. L 298, S. 23, Anhang, Punkt 1c).

In Übersicht 2 sind die in den Rechtsgrundlagen⁸⁾ festgelegten Weiterbefragungsregeln auf Haushaltsebene aufgeführt:

Übersicht 2: Weiterbefragungsregeln auf Haushaltsebene

Kontaktresultat	Regel
Haushalt verweigert Teilnahme.	Haushalt scheidet aus.
Haushalt ist nicht auffindbar.	Haushalt scheidet aus.
Haushalt ist nicht erreichbar.	Haushalt scheidet aus.
Im Haushalt lebt keine Stichprobenperson mehr.	Haushalt scheidet aus.
Im Haushalt lebt mindestens eine Stichprobenperson.	Haushalt ist weiter zu befragen.

Von den Haushalten, in denen mindestens eine Stichprobenperson lebt, dürfen – nach den methodischen Vorgaben von Eurostat – nur jene Haushalte aus dem Vorjahr auch im Folgejahr an einer EU-SILC-Erhebung teilnehmen, die den Haushaltsfragebogen und mindestens einen Personenfragebogen (von einem Haushaltsmitglied im Alter über 16 Jahren) ausgefüllt haben (erfüllte Erfassungsbedingung).⁹⁾

Die Haushaltsdaten der Längsschnittkomponente enthalten bis auf die Zielvariablen zur subjektiven Einschätzung des Wohnumfelds (wie Lärmbelästigung, Verschmutzung, Kriminalität usw.) und zu den Wohnkosten die gleichen Zielvariablen wie die Haushaltsdaten der Querschnittsdaten.¹⁰⁾

Personenmetadaten und Personendaten im Längsschnitt

In den Personenmetadaten des Längsschnitts sind neben den klassischen soziodemografischen Informationen (Alter, Geschlecht, Partner-Identifikationsnummer) die Rücklaufinformationen zu den einzelnen Befragungspersonen aus allen vorliegenden Erhebungsjahren abgelegt. Zusätzlich sind für jede Person Eintritts- und Austrittsinformationen, wie Monat und Jahr des Ein- bzw. Austritts in den Haushalt und die entsprechenden Gründe (z. B. Geburt, Auszug usw.), festgehalten. Die in der Querschnittskomponente erfassten Daten zur Kinderbetreuung sind in den Personenmetadaten des Längsschnitts nicht enthalten.

Analog zum Haushaltsstatus in den Haushaltsmetadaten gibt es in den Personenmetadaten des Längsschnitts die Zielvariable Mitgliedstatus, in der für jede Person die befragungsrelevante Herkunft bzw. die individuelle Teilnahmebiografie gespeichert ist:¹¹⁾

Die Variable Mitgliedstatus hat folgende Ausprägungen:

- (1) Person lebte bereits bei der letzten Befragung in diesem Haushalt.

- (2) Person lebte bei der letzten Befragung noch in einem anderen EU-SILC-Haushalt, lebt jetzt aber in diesem EU-SILC-Haushalt.
- (3) Person ist seit der letzten Befragung neu in den Haushalt gekommen.
- (4) Person wurde nach der letzten Befragung geboren.
- (5) Person ist nach der letzten Befragung ausgezogen.
- (6) Person ist nach der letzten Befragung gestorben.
- (7) Person lebte während der Einkommensreferenzperiode mindestens drei Monate in dem Haushalt, wurde aber nicht als Haushaltsmitglied registriert.

Jede Abspaltung auf Haushaltsebene, jede Rückkehr, jeder Einzug oder Auszug einzelner Haushaltsmitglieder wird auf diese Weise nachvollziehbar und ermöglicht damit, eine sogenannte Haushaltsbiografie für den gesamten Erhebungszeitraum abzubilden.

Analog zu den im vorangegangenen Abschnitt dargestellten Weiterbefragungsregeln auf Haushaltsebene gelten auf der Personenebene die in Übersicht 3 dargestellten Weiterbefragungsregeln:

Übersicht 3: Weiterbefragungsregeln auf Personenebene

Kontaktresultat	Regel
Stichprobenperson ist in einen anderen Haushalt im nationalen Erhebungsgebiet gezogen.	Stichprobenperson ist weiter zu befragen.
Stichprobenperson ist in einen anderen Haushalt außerhalb des nationalen Erhebungsgebiets gezogen.	Registrierung als Auszug bzw. als „Austritt“
Stichprobenperson ist in einen Anstaltshaushalt gezogen.	Registrierung als Auszug bzw. als „Austritt“
Sonstige Person (Nicht-Stichprobenperson) ist in einen anderen Haushalt gezogen.	Registrierung als Auszug bzw. als „Austritt“
Neu eingezogenes befragungsfähiges (über 16 Jahre) Haushaltsmitglied	Zu befragen und Registrierung als „sonstige Person“
Neu eingezogenes nicht-befragungsfähiges (unter 16 Jahre) Haushaltsmitglied	Registrierung als „sonstige Person“

Während auf der Haushaltsebene ein temporärer Ausfall oder ein Aussetzen in einem Erhebungsjahr nicht möglich ist bzw. zum Ausscheiden des Haushalts führt, ist nach den methodischen Vorgaben von Eurostat ein temporärer Ausfall auf Personenebene erlaubt: Haushaltsmitglieder, die in einem Jahr die Teilnahme verweigern (z. B. aufgrund von Krankheit oder temporärer Abwesenheit), werden im nächsten Erhebungsjahr erneut kontaktiert. Wenn die Teilnahme

8) Siehe Fußnote 1, hier: Tabelle 2.

9) Die jährlich abgelegte Zielvariable Haushaltsinterviewakzeptanz enthält die Information zur Erfassungsbedingung (Code 1 = Erfassungsbedingung erfüllt).

10) Die Datei umfasst alle erfolgreich befragten Haushalte eines Erhebungsjahres – unabhängig davon, ob der Haushalt im nächsten Jahr an der Erhebung ebenfalls erfolgreich teilgenommen hat.

11) Siehe auch Verordnung (EG) Nr. 1983/2003 der Kommission vom 7. November 2003 zur Durchführung der Verordnung (EG) Nr. 1177/2003 des Europäischen Parlaments und des Rates für die Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) im Hinblick auf das Verzeichnis der primären Zielvariablen (Amtsbl. der EU Nr. L 98, S. 34, Anhang, Variable RB 110).

dann wieder verweigert wird, gilt dies als endgültige Verweigerung und wird entsprechend registriert.

Bei den Personendaten ist die gemeinsame Schnittmenge der Zielvariablen von Querschnitts- und Längsschnittkomponente geringer als bei den Haushaltsdaten: In den Bereichen Gesundheit, Bildung und Erwerbstätigkeit wurden nur ausgewählte Zielvariablen aus den Querschnittsdaten mit in die Längsschnittdaten übernommen. Auch auf die Zielvariablen zur Staatsangehörigkeit wurde verzichtet. Dafür enthalten die Längsschnittdaten zusätzliche Informationen zur Erwerbsbiografie der teilnahmebereiten Personen: Neben Informationen zum Arbeitsplatzwechsel (einschl. der Gründe für den Wechsel) sind Änderungen des Erwerbsstatus sowie Angaben zur Dauer der Erwerbstätigkeit oder zum Eintrittsalter in die Erwerbstätigkeit gespeichert.

Das Paneldesign der deutschen EU-SILC-Stichprobe

Die deutsche EU-SILC-Stichprobe wird seit dem Jahr 2005 erhoben und ist in vier Unterstichproben – sogenannte Rotationsgruppen – unterteilt, deren Zusammensetzung jeweils repräsentativ für die Gesamtbevölkerung ist. Jedes Jahr scheidet eine Rotationsgruppe aus (Vierjahres-Rotationspanel).¹²⁾ Dieses Paneldesign folgt den Empfehlungen von Eurostat und führt für die jährliche Gesamtstichprobe zu folgenden vier Querschnitts-Rotationsgruppen (Q-R):

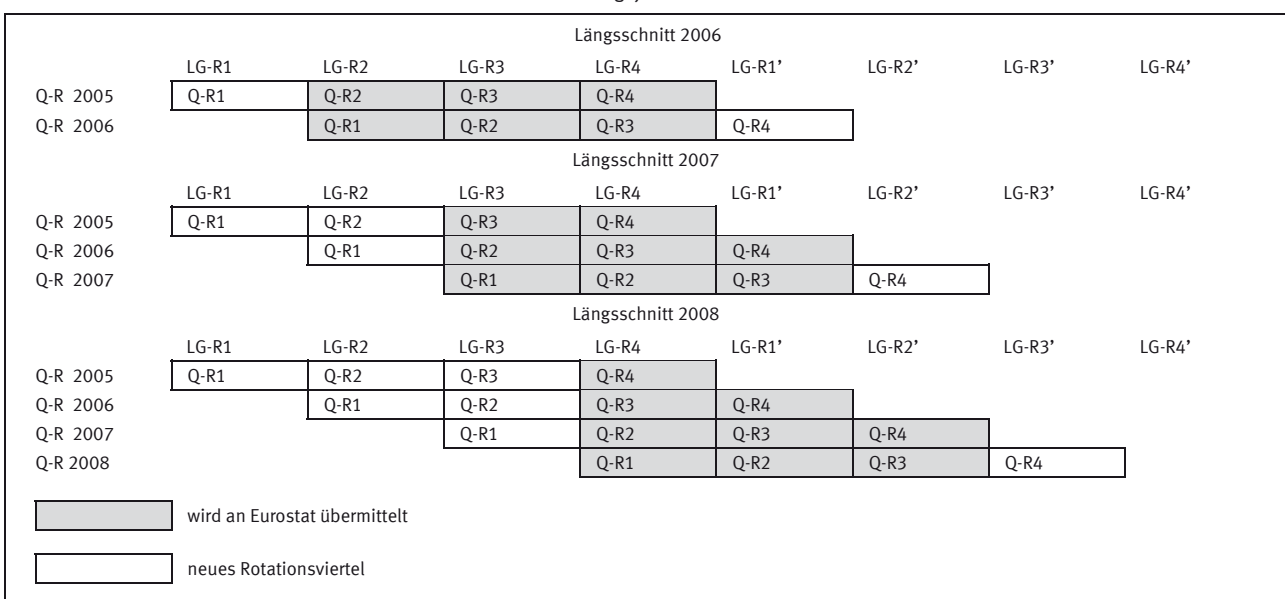
- Rotationsgruppe Q-R1: Haushalt ist noch ein Jahr zu befragen
- Rotationsgruppe Q-R2: Haushalt ist noch zwei Jahre zu befragen

- Rotationsgruppe Q-R3: Haushalt ist noch drei Jahre zu befragen
- Rotationsgruppe Q-R4: Haushalt ist noch vier Jahre zu befragen

In der Längsschnittperspektive verbleibt also ein Haushalt maximal vier Jahre lang in der EU-SILC-Stichprobe und jeder Haushalt gehört einer bestimmten Rotationsgruppe an. Die aus den jährlichen EU-SILC-Erhebungen nach Rotationsgruppenzugehörigkeit zusammengeführten Einzeldaten bilden verschieden lange Längsschnitt-Rotationsgruppen (LG-R). Übersicht 4 illustriert den Aufbau der jährlichen Gesamtstichprobe aus vier Rotationsgruppen und die deutsche Längsschnittgesamtstichprobe (Längsschnittdaten) für die Erhebungsjahre 2005, 2006, 2007 und 2008. Da im ersten Erhebungsjahr von EU-SILC die Dauerstichprobe befragungsbereiter Haushalte noch aufgebaut wurde, setzte sich die deutsche Gesamtstichprobe von EU-SILC im Jahr 2005 noch zu drei Vierteln aus nach einem Quotenverfahren bestimmten Haushalten und nur zu einem Viertel aus Haushalten aus der Dauerstichprobe befragungsbereiter Haushalte zusammen. Ab 2008 besteht die Gesamtstichprobe von EU-SILC ausschließlich aus Rotationsgruppen, die aus der Dauerstichprobe befragungsbereiter Haushalte gezogen wurden. Das bedeutet für die Längsschnittdaten, dass sich auch erst der Längsschnitt für die Erhebungsjahre 2005 bis 2008 (im unteren Teil der Übersicht 4) ausschließlich aus Rotationsgruppen zusammensetzt, die aus der Dauerstichprobe befragungsbereiter Haushalte gezogen wurde.

Die Längsschnittdaten enthalten ausschließlich Haushalte, für die Daten für mindestens zwei Erhebungsjahre vorliegen. Die Längsschnittdatenstichprobe eines Erhebungsjahres ist folglich nur ein Teil der Querschnittsdatenstichprobe.

Übersicht 4: Aufbau der jährlichen Gesamtstichprobe aus 4 Rotationsgruppen und Längsschnitt-Gesamtstichprobe für die Erhebungsjahre 2005 bis 2008



12) Für ausführliche Informationen zum Stichprobendesign der deutschen EU-SILC-Erhebung siehe Horneffer, B./Kuchler, B.: „Drei Jahre Panelerhebung EU-SILC“ in WiSta 8/2008, S. 650 ff.

Die Längsschnitt-identifikationsschlüssel

Eine Grundvoraussetzung für die Generierung von Längsschnittdaten ist, dass Haushalts- und Personenidentifikationsschlüssel unverändert bleiben, damit die Informationen von ein- und denselben Haushalten und Personen in den verschiedenen Erhebungswellen zusammengespielt werden können. Für die erneut zu befragenden Rotationsgruppen werden deshalb zentrale soziodemografische Informationen zur Haushaltsstruktur sowie die Identifikationsschlüssel aus dem Vorjahr in die EU-SILC-Verwaltungssoftware für das neue Erhebungsjahr eingespielt. Einmal vergebene Haushalts- oder Personenidentifikationsschlüssel werden nicht wieder neu vergeben, damit in der Längsschnittperspektive konsistente Informationen vorliegen. Splithaushalte, die sich aus einem bereits bestehenden EU-SILC-Haushalt abgespalten haben, erhalten in der EU-SILC-Verwaltungssoftware eine neue Haushaltsnummer. Diese Splitnummer wird aus der Haushaltsnummer des Ursprungshaushalts abgeleitet. Die EU-SILC-Verwaltungssoftware stellt dabei sicher, dass im Fall von mehreren (zeitlich versetzten) Abspaltungen jede Splitnummer nur einmal vorhanden ist. Die in den Splithaushalt gezogenen Stichprobenpersonen behalten dagegen ihren Personenidentifikationsschlüssel. Kommen im Rahmen einer Abspaltung neue Personen in den Splithaushalten hinzu, erhalten sie einen Personenidentifikationsschlüssel, der sich aus der Splitnummer ableitet. Dadurch ist für jeden Splithaushalt erkennbar, welche Person aus einem Ursprungshaushalt kommt und welche im Rahmen der Abspaltung neu in EU-SILC aufgenommen wurde. Die Gestaltung der Identifikationsschlüssel ist in den Rechtsgrundlagen zu EU-SILC festgelegt (siehe Fußnote 7).

Maßnahmen zur Qualitätssicherung im Statistischen Bundesamt

Eine weitere Voraussetzung für die Generierung von Längsschnittdaten sind zeitlich konsistente und plausible Daten. Da es sich bei der deutschen EU-SILC-Erhebung um eine schriftliche Erhebung handelt, lassen sich Konsistenzprüfungen – gerade auch im Hinblick auf vorangegangene Befragungen dieser Haushalte oder Personen –, wie sie im Rahmen eines mündlichen oder telefonischen Interviews möglich sind, nicht unmittelbar vornehmen. Stattdessen müssen die zurückgesandten Erhebungsunterlagen einer sorgfältigen Sichtprüfung unterzogen werden.¹³⁾ Bereits bei der Erfassung der Daten wird versucht, durch einen Vergleich mit den Werten des jeweiligen Vorjahres eine möglichst umfassende Konsistenz herzustellen. Zusätzlich werden für jede Rotationsgruppe die Informationen des aktuellen Erhebungsjahres mit den Informationen aus vorangegangenen Erhebungswellen abgeglichen. So stimmten zum Beispiel 0,3 % der Angaben zum Geburtsjahr, Geburtsmonat und Geschlecht im Erhebungsjahr 2007 nicht mit den im Erhebungsjahr 2006 gemachten Angaben überein, das

heißt von insgesamt 31 709 Personen mussten beim Generieren der Querschnittsdaten etwa 80 Fälle korrigiert werden. Insgesamt 0,4 % der wiederholt befragten Personen hatten – nach Vorliegen der Daten mehrerer Erhebungsjahre – un plausible Angaben zum Familienstand gemacht.

Auch Veränderungen in der Haushaltszusammensetzung (z. B. durch Tod, Trennung oder Wiederverheiratung von Ehepartnern oder Wegzug von Kindern) sowie Veränderungen, die den gesamten Haushalt betreffen – wie Umzug oder Auflösung des Haushalts –, werden Konsistenzprüfungen unterzogen.

Bisherige Erfahrungen zeigen, dass inkonsistente Angaben im Zeitverlauf unter anderem zustande kommen durch:

- Proxy-Interviews (ein Haushaltsmitglied füllt den Personenfragebogen für ein anderes Haushaltsmitglied aus),
- sich über mehrere Monate hinziehende Haushaltsveränderungen (z. B. Trennung),
- Wechsel der Person, die den Haushaltsfragebogen ausfüllt.

Die Konsistenzprüfungen erfolgen auf drei Ebenen im Zeitverlauf:

- Prüfung der individuellen Angaben,
- Prüfung der haushaltsbezogenen Angaben,
- Prüfung der Beziehungen im Haushalt, insbesondere der verwandtschaftlichen Beziehungen zwischen den Haushaltsmitgliedern und der Partnerangaben.

Hochrechnung

Sobald die Aufbereitung der Längsschnittdaten abgeschlossen ist, werden die Längsschnittgewichte für die EU-SILC-Erhebung ermittelt. Die Berechnung erfolgt nach den methodischen Vorgaben von Eurostat und wird hier kurz zusammengefasst:

Die EU-SILC-Stichprobe wird nach den Rotationsgruppen differenziert ausgewertet und führt zu folgenden verschiedenen Längsschnittpopulationen (Panels), wobei Y das betrachtete Erhebungsjahr ist:

1. 2-Jahres-Panel: Betrachtet wird dabei die Datenmenge
 - vom Jahr $Y-1$ zum Jahr Y und
 - vom Jahr $Y-2$ zum Jahr $Y-1$ und
 - vom Jahr $Y-3$ zum Jahr $Y-2$.
2. 3-Jahres-Panel: Betrachtet wird dabei die Datenmenge
 - vom Jahr $Y-2$ zum Jahr Y und
 - vom Jahr $Y-3$ zum Jahr $Y-1$.

¹³⁾ Z. B. kann für jeden Haushalt in der EU-SILC-Verwaltungssoftware die Haushaltsbiografie (Haushaltskomposition über mehrere Jahre) und die aktuelle Haushaltsstruktur eingeblendet werden, um die zurückgesandten Erhebungsunterlagen noch vor der Erfassung mit der Paneldatenbankinformation abzugleichen.

3. 4-Jahres-Panel: Betrachtet wird dabei die Datenmenge

- vom Jahr $Y-3$ zum Jahr Y .

Nach den methodischen Vorgaben von Eurostat sind für die 2- und 3-Jahres-Panel nur die Daten auf Personenebene von Interesse. Daher werden für diese beiden Panels nur Personenlängsschnittgewichte gebildet.

Für das 4-Jahres-Panel werden dagegen drei Längsschnittgewichte ermittelt: ein Haushaltsgewicht, ein Personengewicht für alle Personen und ein Personengewicht für die Personen ab 16 Jahren.

Die Personenlängsschnittgewichte, die für die 2-Jahres- und 3-Jahres-Panel-Analysen verwendet werden, ergeben sich aus dem Personengewicht für das 4-Jahres-Panel. Für alle zugezogenen, verstorbenen und neugeborenen Personen wird das Längsschnittgewicht auf Null gesetzt. Die Summe der Gewichte für das 2-Jahres-Panel über alle Rotationsgruppen hinweg stellt die Gesamtbevölkerung auf Personenebene dar. Gleiches gilt für die Gewichte für das 3-Jahres-Panel.

Die Gewichte für den 4-Jahres-Längsschnitt werden in Anlehnung an die Hochrechnung im Querschnitt bestimmt: Die Bestimmung des Haushalts- und der beiden Personengewichte erfolgt analog zur Gewichtsbestimmung im Querschnitt. Zur Hochrechnung wird ein verallgemeinerter Regressionsschätzer verwendet und alle Personen eines Haushalts erhalten das gleiche Gewicht.¹⁴⁾

Das endgültige Längsschnittgewicht ergibt sich durch die Berücksichtigung von Veränderungen¹⁵⁾ wie:

- Vergrößerung des Haushalts, weil Personen zum Haushalt zugezogen sind,
- Vergrößerung des Haushalts, weil Personen im Haushalt geboren wurden,
- Verkleinerung des Haushalts, weil die Personen des Haushalts sich getrennt haben,
- Verkleinerung des Haushalts, weil Personen des Haushalts ausgezogen sind,
- Verkleinerung des Haushalts, weil Personen des Haushalts gestorben sind,
- vorzeitiges Ausscheiden des Haushalts aus der Erhebung, weil alle Personen des Haushalts verstorben sind.

Diese Veränderungen sind in verschiedenen Hilfsvariablen der Längsschnittdaten hinterlegt. Dabei ist zu beachten, dass bei ein- und demselben Haushalt mehrere Veränderungen zeitgleich stattfinden können.

Das Haushaltslängsschnittgewicht wird für alle im Längsschnitt befindlichen Haushalte je Erhebungsjahr und Rotationsgruppe berechnet. Die Summe der Haushaltsgewichte je Erhebungsjahr stimmt mit der Anzahl der Haushalte in der Gesamtbevölkerung überein.

Die Summe der Personenlängsschnittgewichte für alle Personen einer Rotationsgruppe entspricht der Anzahl der Personen in der Gesamtbevölkerung. Für alle außerhalb der Stichprobe zugezogenen wie auch für alle verstorbenen Personen weist das Gewicht den Wert Null auf. Alle neugeborenen Personen bekommen das Gewicht der Mutter zugeordnet.

Beim dritten Längsschnittgewicht für das 4-Jahres-Panel werden nur die Personen ab 16 Jahren betrachtet. Neugeborene Personen und neu zugezogene Personen unter 16 Jahren werden hier nicht berücksichtigt. Für alle zugezogenen Personen über 16 Jahre wird dieses Gewicht auf Null gesetzt. Die Summe der Personengewichte stimmt je Rotationsgruppe mit der Gesamtbevölkerung ab 16 Jahren überein.

Maßnahmen zur Qualitätssicherung bei Eurostat

Eurostat hat zum Prüfen der Querschnitts- und Längsschnittdaten aus EU-SILC verschiedene Prüfprogramme in SAS (Statistical Analysis System, ein Softwaresystem, das u. a. in der amtlichen Statistik verwendet wird – Anmerkung der Redaktion) entwickelt, mit denen Signierprüfungen und logische Prüfungen für die Zielvariablen und deren Flags (zugehörige Hintergrundinformationen)¹⁶⁾ durchgeführt werden. Mit diesen Programmen kann jeder Mitgliedstaat die eigenen Mikrodaten schon vor dem Versand an Eurostat prüfen und eventuell korrigieren. Liegen Unstimmigkeiten in den Daten vor, gibt das Prüfprogramm Fehlerart und Zahl der betroffenen Fälle aus. Dabei wird unterschieden zwischen „Warnhinweisen“ und „Fehlern“. Die logischen Prüfungen umfassen zum Beispiel:

- die Beziehungen zwischen den Haushaltsmitgliedern [z. B. Altersabstand der Mutter zu(m) Kind(ern) im Haushalt],
- die Konsistenz der verschiedenen Erwerbssinformationen zu einer Person,
- den Zusammenhang zwischen Haushaltsnettoeinkommen und den einzelnen Einkommenskomponenten.

Bei den Längsschnittdaten werden rund 165 Konsistenzprüfungen durchgeführt, wobei zusätzlich zu den umfangreichen Prüfungen der Querschnittsdaten die Längsschnitt-

14) Für ausführliche Informationen zur Hochrechnung der Querschnittsdaten siehe Horneffer, B./Kuchler B., Fußnote 12.

15) (Sozio-)demografische Veränderungen, die die gesamte Bevölkerung betreffen (Bevölkerungsdynamik), wie Zahl der Wegzüge ins Ausland, Zahl der (Lebend-)Geburten, Todesfälle über die Jahre, werden dagegen nicht berücksichtigt.

16) Für weitere Informationen über Flags siehe auch Sikorski, U., Fußnote 4.

informationen geprüft werden, wie zum Beispiel die Entwicklung der Einkommen der Personen und Haushalte und Veränderungen bei der Erwerbstätigkeit der Personen im Zeitverlauf.

Empirische Ergebnisse zur Ausschöpfung und zur Panelmortalität

Aufgrund der Zusammensetzung der EU-SILC-Gesamtstichprobe (Quotenstichprobe und Zufallsstichprobe aus der Dauerstichprobe befragungsbereiter Haushalte)¹⁷⁾ in den Erhebungsjahren 2005 bis 2007 war es nicht möglich, eine Ausschöpfungsquote für die Gesamtstichprobe zu ermitteln. In den Schaubildern 1 und 4 ist deshalb der Rücklauf für alle Erhebungsjahre auf Haushalts- und Personenebene je Rotationsviertel dargestellt.

Danach weisen die im jeweiligen Jahr neu aus der Dauerstichprobe befragungsbereiter Haushalte gezogenen Rotationsgruppen ähnlich hohe Ausschöpfungsquoten auf [2005: 1. Rotationsgruppe (LG-R4) 64,6%; 2006: 2. Rotationsgruppe (LG-R1') 65,8%; 2007: 3. Rotationsgruppe (LG-R2') 65,3% und 2008: 4. Rotationsgruppe (LG-R3') 57,3%]. Die Wiederbefragungsbereitschaft in den verbleibenden Rota-

tionsvierteln ist ebenfalls ähnlich hoch: Deutlich mehr als 80 % der im Vorjahr befragten Haushalte nahmen jeweils erneut an EU-SILC teil.

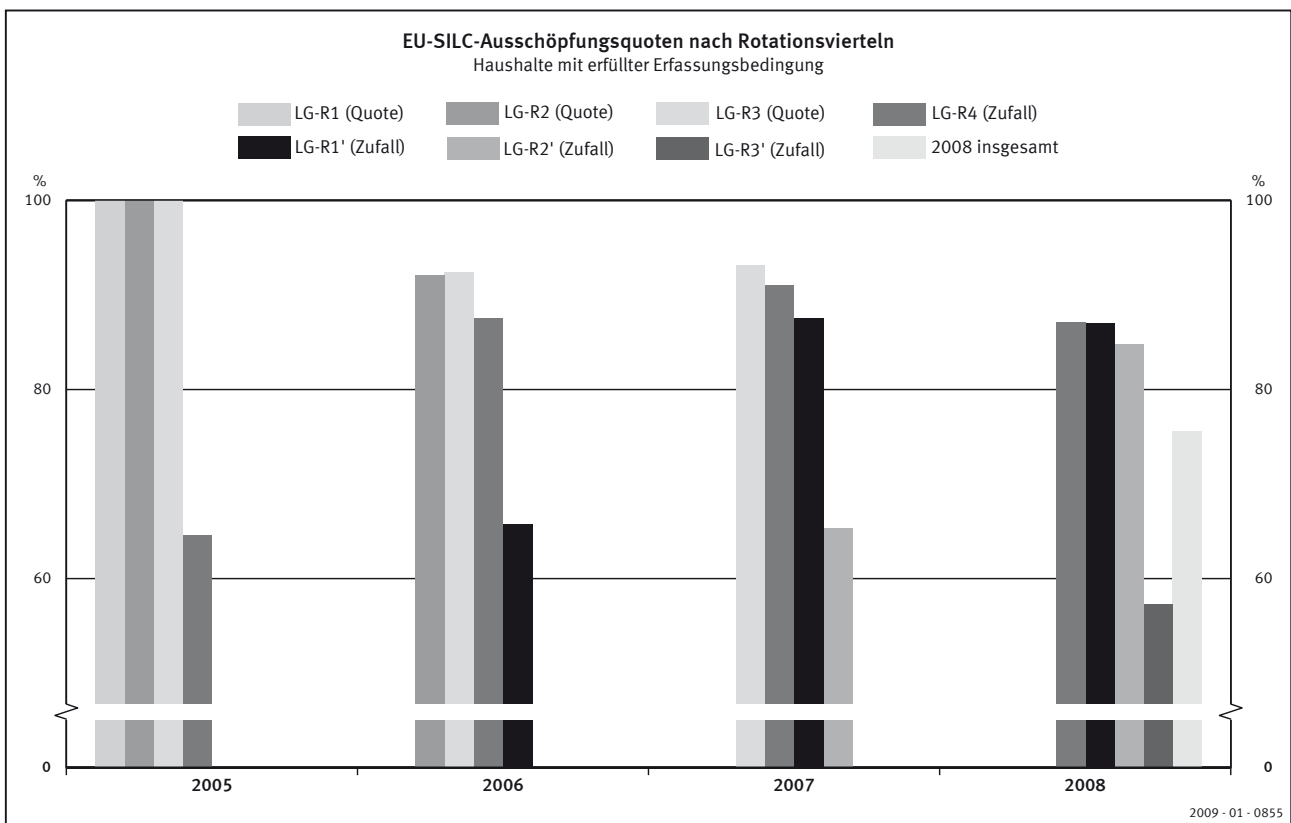
Ausschöpfungsquote der Gesamtstichprobe

Im Jahr 2008 setzte sich die EU-SILC-Gesamtstichprobe erstmals ausschließlich aus Rotationsgruppen zusammen, die zufällig aus der Dauerstichprobe befragungsbereiter Haushalte gezogen wurden. Die Ausschöpfungsquote für die EU-SILC-Gesamtstichprobe betrug in diesem Jahr knapp 76 % (siehe Schaubild 1). Auf Personenebene betrug die Ausschöpfungsquote sogar fast 100%: Von den teilnahmebereiten Haushalten haben fast alle Haushaltsmitglieder auch einen Personenfragebogen ausgefüllt (siehe Schaubild 4).

Teilnahmebereitschaft der ersten Zufallsteilstichprobe in EU-SILC

Die erste Rotationsgruppe (LG-R4), die im Jahr 2005 aus der Dauerstichprobe befragungsbereiter Haushalte zufällig gezogen worden war, hatte im Jahr 2008 ihr letztes Befragungsjahr (siehe Schaubild 2). Aus der Längsschnittpers-

Schaubild 1



17) Für ausführliche Informationen zur Dauerstichprobe befragungsbereiter Haushalte siehe Körner, T./Nimmergut, A./Nökel, J./Rohloff, S.: „Die Dauerstichprobe befragungsbereiter Haushalte“ in WiSta 5/2006, S. 451 ff.

Schaubild 2

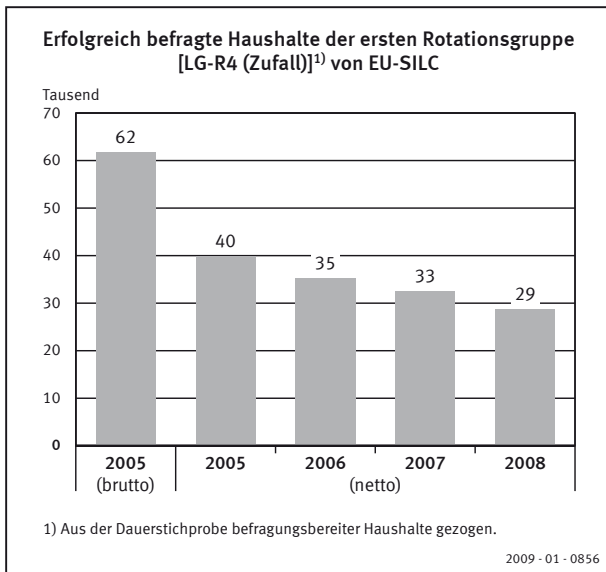
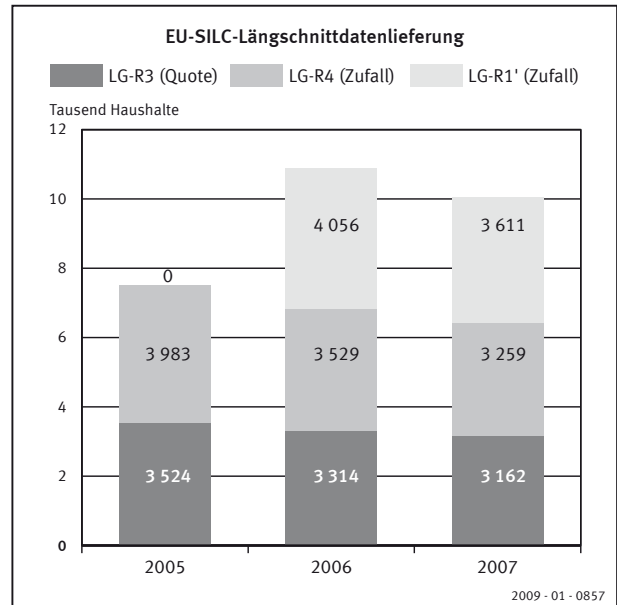


Schaubild 3

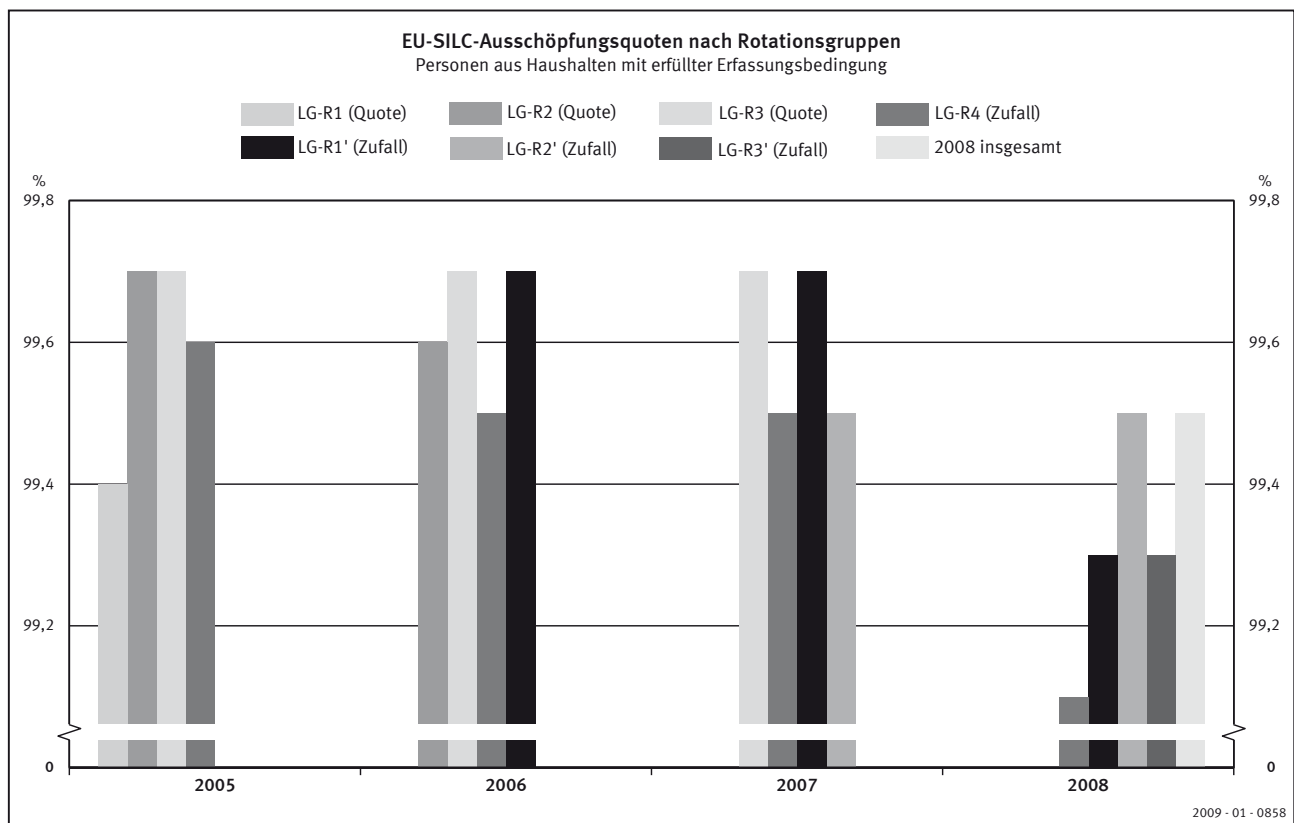


pektive waren von der ersten Rotationsgruppe, die im Jahr 2005 erfolgreich geantwortet hatte (3 985), immerhin 2 878 Haushalte (einschl. Abspaltungen) bereit gewesen, ununterbrochen drei weitere Jahre an EU-SILC teilzunehmen. Das ergibt für diese erste Rotationsgruppe bzw. Zufallsteilstich-

probe von EU-SILC eine Längsschnitt-Ausschöpfungsquote (2005 bis 2008) von etwa 72%.¹⁸⁾

Aus der Perspektive der Bruttostichprobe des ersten Rotationsviertels (6 171) im Jahr 2005 beträgt die Längsschnitt-

Schaubild 4



18) Da die Längsschnittdaten 2005 bis 2008 derzeit aufbereitet werden, basieren die Auswertungen auf den Querschnittsdaten 2005 bis 2008, in denen für jedes Rotationsviertel der jährliche Rücklauf gespeichert ist.

Ausschöpfungsquote etwa 47 %: Die Hälfte der im Jahr 2005 gezogenen Haushalte war folglich bereit, nach vier Jahren Teilnahme am Mikrozensus noch einmal vier Jahre an der EU-SILC-Erhebung teilzunehmen.

Zusammensetzung der jüngsten Längsschnittdatenlieferung (2005 bis 2007)

In Schaubild 3 ist die quantitative Zusammensetzung der EU-SILC-Längsschnittdatenlieferung 2005 bis 2007 dargestellt, die im März 2009 an Eurostat geliefert wurde. Danach sind derzeit 2-Jahres-Längsschnittanalysen für etwa 10 000 und 3-Jahres-Längsschnittanalysen für etwa 6 400 Haushalte möglich.

Teilnahmebereitschaft auf Personenebene

Die Teilnahmebereitschaft der befragungsfähigen Haushaltsmitglieder (im Alter von über 16 Jahren) von erfolgreich befragten Haushalten ist in Schaubild 4 ausgewiesen. Es zeigt sich für alle Jahre und alle Rotationsgruppen, dass trotz der relativ großzügigen Erfassungsbedingung (je Haushalt muss der Haushaltsfragebogen und mindestens ein Personenfragebogen vorliegen) fast alle befragungsfähigen Haushaltsmitglieder auch an EU-SILC teilnehmen. Die Quoten- und Zufallsteilstichproben von EU-SILC unterscheiden sich hierbei nicht.

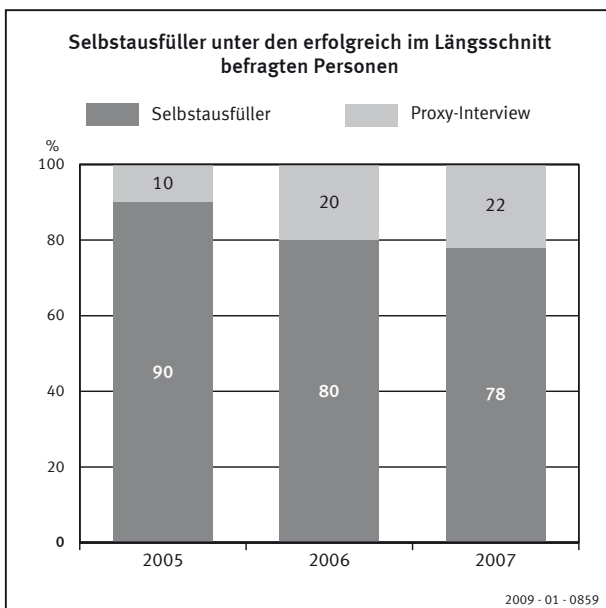
Die Entwicklung von Proxy-Interviews (ein Haushaltsmitglied füllt für ein anderes Haushaltsmitglied den Personenfragebogen aus) unter den in allen Erhebungsjahren erfolgreich befragten Personen der Längsschnittdatenlieferung 2005 bis 2007 ist in Schaubild 5 abgebildet. Danach stieg

der Anteil von Proxy-Interviews in dieser Längsschnittpopulation von 10 % im ersten Erhebungsjahr (2005) auf 22 % im dritten Erhebungsjahr (2007). Die überwiegende Mehrheit der Personen, die in allen drei Erhebungsjahren an EU-SILC teilnahmen, war folglich auch bereit, den Personenfragebogen selbst auszufüllen.

Ausblick

Nach den bisherigen Ergebnissen zum Rücklauf der verschiedenen Zufallsteilstichproben in EU-SILC (Rotationsgruppen) ist die Teilnahmebereitschaft der Haushalte und der befragungsfähigen Haushaltsmitglieder in den Folgejahren sehr hoch. Von der ersten Rotationsgruppe, die aus der Dauerstichprobe befragungsbereiter Haushalte im Jahr 2005 gezogen wurde (Bruttostichprobe), war knapp die Hälfte der Haushalte bereit, an einem kompletten EU-SILC-Befragungszyklus (2005 bis 2008) teilzunehmen. Aus der Perspektive der Nettostichprobe nahmen etwa 72 % (etwa 2 800 Haushalte) der im Jahr 2005 erfolgreich befragten Haushalte aus der ersten Rotationsgruppe bis zum Ende des Befragungszyklus teil. Ausführlichere Analysen zur Panelmortalität – insbesondere im Hinblick auf die Teilnahmebereitschaft von sozialen Gruppen – sind möglich, sobald die Aufbereitung der Längsschnittdaten 2005 bis 2008 abgeschlossen ist. [u](#)

Schaubild 5



Auszug aus Wirtschaft und Statistik

© Statistisches Bundesamt, Wiesbaden 2009

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.

Herausgeber: Statistisches Bundesamt, Wiesbaden

Schriftleitung: Roderich Egeler
Präsident des Statistischen Bundesamtes
Verantwortlich für den Inhalt:
Brigitte Reimann,
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 2086
- E-Mail: wirtschaft-und-statistik@destatis.de

Vertriebspartner: SFG Servicecenter Fachverlage
Part of the Elsevier Group
Postfach 43 43
72774 Reutlingen
Telefon: +49 (0) 70 71/93 53 50
Telefax: +49 (0) 70 71/93 53 35
E-Mail: destatis@s-f-g.com

Erscheinungsfolge: monatlich



Allgemeine Informationen über das Statistische Bundesamt und sein Datenangebot erhalten Sie:

- im Internet: www.destatis.de

oder bei unserem Informationsservice
65180 Wiesbaden

- Telefon: +49 (0) 6 11/75 24 05
- Telefax: +49 (0) 6 11/75 33 30
- www.destatis.de/kontakt