

Dipl.-Mathematikerin Birgit Horneffer, Dipl.-Soziologin Birgit Kuchler

Drei Jahre Panelerhebung EU-SILC

Erfahrungen und methodische Weiterentwicklungen

Mit der Einführung der Gemeinschaftsstatistik über Einkommen und Lebensbedingungen (EU-SILC) im Jahr 2004 entwickelte das Statistische Amt der Europäischen Gemeinschaften (Eurostat) die Idee einer jährlichen europäischen Sozialberichterstattung weiter. Neben Informationen zur Einkommensarmut und Einkommensungleichheit in den Mitgliedstaaten der Europäischen Union (EU) stehen mit EU-SILC nun auch zahlreiche Informationen über die subjektiven Einschätzungen der Menschen zu ihren Lebensbedingungen zur Verfügung.

In Deutschland werden die Daten für diese Gemeinschaftsstatistik mit einer eigenen Haushaltsbefragung: *LEBEN IN EUROPA* gewonnen. Die Erhebung startete im Jahr 2005, jährlich werden schriftlich etwa 14 000 Haushalte befragt, die Teilnahme ist freiwillig. Derzeit liegen Mikrodaten für die Erhebungsjahre 2005 und 2006 vor.

In einem Beitrag von Körner u. a. (2005)¹⁾ wurden bereits ausführlich die Zielsetzung, die Rechtsgrundlagen der Erhebung und die Outputharmonisierung beschrieben. Nach einem Überblick über die Erhebungsorganisation geht es in diesem Beitrag um bisher durchgeführte Maßnahmen zur Erhöhung der Datenqualität im Bereich der Erhebungsorganisation und der Hochrechnung.

Der Beitrag schließt mit einem Ausblick auf die derzeit auf nationaler wie auch auf europäischer Ebene diskutierten methodischen Weiterentwicklungen.

Rechtsgrundlagen der Gemeinschaftsstatistik über Einkommen und Lebensbedingungen

Rechtsgrundlagen für die deutsche EU-SILC-Erhebung *LEBEN IN EUROPA* sind eine europäische Rahmenverordnung²⁾ und verschiedene europäische Durchführungsverordnungen, in denen das Erhebungsdesign, die Behandlung von Nonresponse, der Inhalt der an Eurostat zu liefernden Mikrodaten (Zielvariablenliste), das Hochrechnungsverfahren und die Dokumentation der Datenproduktion festgelegt sind. Die Datenproduktion erfolgt outputharmonisiert: Die Mitgliedstaaten können selbst entscheiden, in welcher Form sie die Daten produzieren – sofern diese den in den Verordnungen festgelegten Qualitätsanforderungen entsprechen.

Die jährlich an Eurostat zu übermittelnden Mikrodaten bestehen aus einer Querschnitts- und einer Längsschnittkomponente und sind in vier Zieldateien gespeichert:

- Haushaltsmetadaten – Bruttostichprobe
- Personenmetadaten – Nettostichprobe
- Daten aus dem Haushaltsfragebogen – Nettostichprobe
- Daten aus dem Personenfragebogen – Nettostichprobe

1) Siehe Körner, T./Meyer, I./Minkel, H./Timm, U.: „LEBEN IN EUROPA – Die neue Statistik über Einkommen und Lebensbedingungen“ in WiSta 11/2005, S. 1137 ff.

2) Siehe 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 (Amtsbl. der EU Nr. L 255, S. 6).

Die Zielvariablenliste enthält verschiedene Standardmodule³⁾ und ein sich jährlich änderndes Schwerpunktmodul⁴⁾. Für einen Teil der Zielvariablen gibt es genau festgelegte Definitionen und Beziehungen zu anderen Zielvariablen, für andere dagegen lautet die Vorgabe, auf Definitionen oder Zuordnungshinweise zu verzichten, um die Selbsteinschätzung der Befragten nicht zu beeinflussen (z. B. bei der Selbsteinschätzung der Befragten zur aktuellen Lebens-/Erwerbssituation). Der Handlungsspielraum für die Gestaltung der Erhebungsunterlagen ist zum großen Teil durch festliegende logische Beziehungen zwischen den Zielvariablen vorgegeben oder aufgrund von Empfehlungen strukturiert. Informationen zur Feldphase und zur Panelmortalität (Auszüge, Einzüge, Geburten, Ausfallgründe, Ergebnisse der Kontaktierung) sind in Form von Zielvariablen bereitzustellen. Für jedes Erhebungsjahr sind zwei Qualitätsberichte⁵⁾ anzufertigen, deren Inhalte in einer Durchführungsverordnung festgelegt sind.

Die an Eurostat übermittelten anonymisierten Mikrodaten werden geprüft und anschließend zu einer Gemeinschaftsdatenbasis EU-SILC zusammengeführt.⁶⁾ Die Ergebnisse von EU-SILC fließen in zahlreiche europäische Sozialberichte (z. B. Joint Report on Social Protection and Social Inclusion 2008) ein und stellen die zentrale Informationsquelle für die europäische Sozialberichterstattung dar. Auf nationaler Ebene fanden die Daten der deutschen EU-SILC-Erhebung im 3. Armuts- und Reichtumsbericht der Bundesregierung Verwendung.

Das Stichprobendesign der deutschen EU-SILC-Erhebung

Die deutsche EU-SILC-Erhebung ist als vierjähriges Rotationspanel angelegt und setzt sich aus vier Teilstichproben (Rotationsvierteln) zusammen. Grundlage für die jährliche Ziehung des neuen Zufalls-Rotationsviertels ist die Dauerstichprobe befragungsbereiter Haushalte⁷⁾. Im ersten Erhebungsjahr von EU-SILC (2005) befand sich die Dauerstichprobe befragungsbereiter Haushalte noch im Aufbau und es wurde lediglich ein Rotationsviertel aus der Dauerstichprobe befragungsbereiter Haushalte gezogen. Die drei anderen Rotationsviertel im ersten Erhebungsjahr wurden nach einem Quotenverfahren bestimmt. Der Aufbau der Dauerstichprobe befragungsbereiter Haushalte und die Zusammensetzung der EU-SILC-Gesamtstichprobe in den Folgejahren sind in Schaubild 1 auf S. 652 illustriert.

Sowohl die Dauerstichprobe befragungsbereiter Haushalte als auch die deutsche EU-SILC-Erhebung sind dezentral orga-

nisiert. In Tabelle 1 ist für jedes Bundesland ausgewiesen, in welchen Erhebungsjahren eine Stichprobenziehung für ein Rotationsviertel auf Basis der jeweiligen landesspezifischen Datenbank der Dauerstichprobe befragungsbereiter Haushalte möglich war (S) und in welchen die landesspezifische Datenbank der Dauerstichprobe befragungsbereiter Haushalte noch zu wenig Haushalte enthielt, um eine Ziehung durchzuführen (A). Im zweiten Fall wurden alle in der Paneldatenbank der Dauerstichprobe vorhandenen Haushalte für den Aufbau des neuen Rotationsviertels herangezogen. Wie Tabelle 1 zeigt, war in allen Erhebungsjahren für eine deutliche Mehrheit der Statistischen Ämter der Länder eine Stichprobenziehung aus der Dauerstichprobe befragungsbereiter Haushalte möglich. Allerdings zeigt das Ergebnis auch, dass der Aufbau der landesspezifischen Datenbanken der Dauerstichprobe als Ziehungsbasis für die deutsche EU-SILC-Erhebung im Jahr 2008 noch nicht abgeschlossen ist.

Tabelle 1: Ziehungen der Haushalte für EU-SILC aus der Dauerstichprobe befragungsbereiter Haushalte in den einzelnen Erhebungsjahren nach Bundesländern Erhebungsjahre

Land	2005	2006	2007	2008
Schleswig-Holstein	A	S	S	S
Hamburg	A	A	A	S
Niedersachsen	A	A	A	S
Bremen	A	A	A	A
Nordrhein-Westfalen	S	S	S	S
Hessen	S	S	S	S
Rheinland-Pfalz	S	S	S	S
Baden-Württemberg	S	S	S	S
Bayern	S	S	S	S
Saarland	S	S	S	S
Berlin	S	S	S	A
Brandenburg	S	S	S	S
Mecklenburg-Vorpommern ..	S	S	S	S
Sachsen	S	S	S	S
Sachsen-Anhalt	A	S	A	A
Thüringen	S	S	S	S
Stichprobenziehungen insgesamt	11	13	12	13

S = Stichprobenziehung; A = Stichprobenziehung nicht möglich.

Auswahlplan

Für jedes Erhebungsjahr wird ein Auswahlplan für das neu hinzukommende Rotationsviertel in EU-SILC erstellt. Die EU-Rahmenverordnung sieht für Deutschland bei einfacher Zufallsauswahl für die Querschnittskomponente einen Stichprobenumfang von 8.250 Haushalten vor. Da die EU-SILC-Stichprobe aus der Dauerstichprobe befragungsbereiter Haushalte gezogen wird, für die Haushalte ausschließlich aus dem Mikrozensus angeworben werden, können

3) Siehe Lenuweit, B.: „LEBEN IN EUROPA 2005“ in WiSta 1/2007, S. 31 ff. Die Haushaltserhebung LEBEN IN EUROPA umfasst Themenbereiche wie Einkommen, Wohnverhältnisse und Wohnkosten, Kinderbetreuung, Gesundheit, soziale Ausgrenzung sowie soziodemografische Informationen. Sie dient als Grundlage der Berechnung der sog. „Laeken-Indikatoren“. Zudem gibt sie einen Einblick in den Bereich nicht-monetärer Deprivation und ermöglicht vergleichende Analysen der allgemeinen Lebensbedingungen der Bevölkerung. Die Erhebung spielt damit eine wichtige Rolle im Rahmen der offenen Koordinierung der europäischen Sozialpolitik.

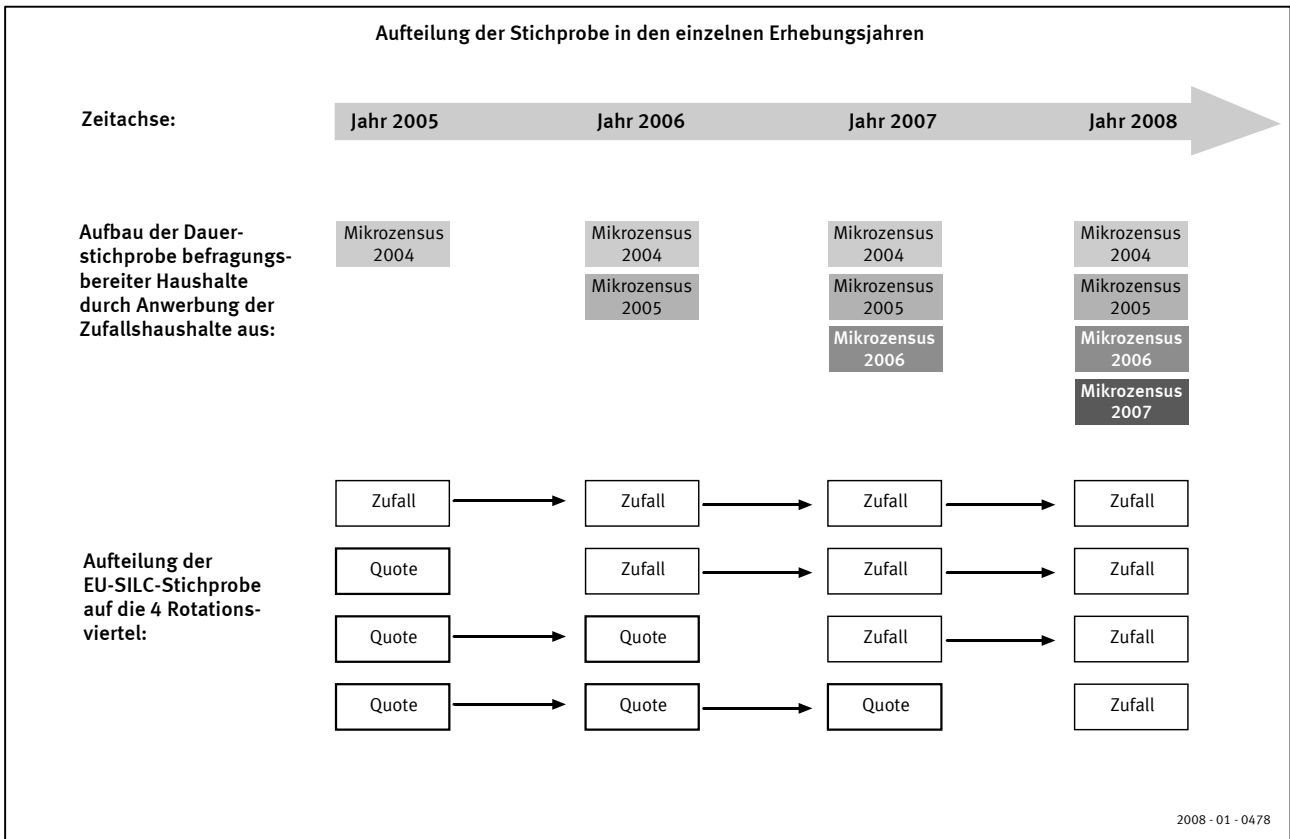
4) Der Inhalt eines Schwerpunktmoduls wird in einer eigenen europäischen Durchführungsverordnung festgelegt. Folgende Schwerpunktmodule gab es bisher: 2005: Finanzielle Situation im Elternhaus während der eigenen Jugend; 2006: Freizeitaktivitäten; 2007: Subjektive Einschätzungen zum Wohnumfeld; 2008: Verschuldung. Die Umsetzung der Schwerpunktmodule sowie die methodische Weiterentwicklung der Datenproduktion erfolgt in verschiedenen Unterarbeitsgruppen (Task Forces) bei Eurostat.

5) Im Folgejahr der Erhebung ist ein Zwischenqualitätsbericht anzufertigen, in dem die Datenqualität der Querschnittskomponente dokumentiert ist. Zwei Jahre nach der Erhebung bzw. nach der Lieferung der Längsschnittkomponente ist ein Abschlussqualitätsbericht zu erstellen. Eurostat stellt die nationalen Qualitätsberichte zu einem europäischen Qualitätsbericht für jedes Erhebungsjahr zusammen. Der europäische Abschlussbericht für ein Erhebungsjahr wird veröffentlicht.

6) Für die wissenschaftliche Nutzung der Daten gibt Eurostat ferner eine anonymisierte Gemeinschaftsdatenbasis (Scientific-Use-File) heraus. Die anonymisierten Mikrodaten der deutschen EU-SILC-Erhebung stehen beim Forschungsdatenzentrum des Statistischen Bundesamtes für Auswertungen zur Verfügung.

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

Schaubild 1



damit auch Klumpeneffekte aus dem Mikrozensus fortwirken. Das ist beim Stichprobenplan zu berücksichtigen, wodurch die Anzahl der zu befragenden Haushalte steigt. Ferner wird eine Panelmortalität von 10% angenommen. Insgesamt wird daher ein jährlicher Nettostichprobenumfang von 14 100 teilnehmenden Haushalten angestrebt. Eine Beschreibung des Auswahlplans erfolgte bereits in dem Beitrag von Körner u. a. (siehe Fußnote 1).

Aufteilung des Stichprobenumfangs auf die Zufalls- bzw. Quotenstichprobe

Der Gesamtstichprobenumfang von 14 100 Haushalten wird proportional zur Zahl der dort vorhandenen Haushalte auf die Bundesländer aufgeteilt, um möglichst präzise Bundesergebnisse zu erhalten. Für das erste Erhebungsjahr 2005 war ein Nettostichprobenumfang von 4 100 Haushalten für die Zufallsstichprobe und 10 000 Haushalten für die Quotenstichprobe geplant. Ausgehend von einer Ausschöpfungsquote von 65% müssen jährlich mindestens 6 300 Haushalte für die Zufallsstichprobe angesprochen werden, um diesen Nettostichprobenumfang zu realisieren.

Bei allen vier Teilstichproben wurden die Schichten bzw. die Quotierungszellen definitorisch auf dieselbe Art abgegrenzt – abgesehen von eventuell bei der Zufallsstichprobe wegen der geringeren Fallzahl erforderlichen Schichtzusammenfassungen. Auch bei der Aufteilung des Stichprobenumfangs

auf die Schichten bzw. Quotierungszellen wurden keine Unterschiede zwischen den Teilstichproben gemacht. Im Folgenden ist der Einfachheit halber nur von Schichtungsmerkmalen und Schichten die Rede.⁸⁾ Bei der Zufallsstichprobe sind die Strukturen auf die Nettostichprobe bezo-

Tabelle 2: Aufteilung des Stichprobenumfangs nach Bundesländern für das Erhebungsjahr 2005

Bundesland	Gesamtstichprobenumfang (netto)	Zufallsstichprobenumfang (netto)	Zufallsstichprobenumfang (brutto)	Quotenstichprobenumfang (Erhebungssoll)
Schleswig-Holstein	482	140	215	342
Hamburg	337	98	151	239
Niedersachsen	1 358	395	608	963
Bremen	131	38	58	93
Nordrhein-Westfalen	3 086	897	1 380	2 189
Hessen	1 028	299	460	729
Rheinland-Pfalz	673	196	302	477
Baden-Württemberg	1 745	507	780	1 238
Bayern	2 024	589	906	1 435
Saarland	188	55	85	133
Berlin-West	420	122	188	298
Berlin-Ost	258	75	115	183
Brandenburg	436	127	195	309
Mecklenburg-Vorpommern	304	88	135	216
Sachsen	777	226	348	551
Sachsen-Anhalt	446	130	200	316
Thüringen	407	118	182	289
Deutschland ...	14 100	4 100	6 308	10 000

⁸⁾ Im Fall der Quotenstichprobe sind damit die Quotierungsmerkmale bzw. die Quotierungszellen gemeint.

gen. Aus einem umgerechneten Bruttostichprobenumfang je Schicht und der Schichtbesetzungszahl in der Dauerstichprobe befragungsbereiter Haushalte werden dann die Auswahlätze (bzw. die Auswahlabstände) für die Ziehung berechnet.

Aufteilung des Stichprobenumfangs auf die Schichtungsmerkmale

Die Aufteilung wird für die drei Schichtungsmerkmale (Haushaltstyp, soziale Stellung des Haupteinkommensbeziehers, Haushaltsnettoeinkommen) getrennt durchgeführt, sodass man zunächst nur die Stichprobenumfänge für die Randsummen erhält.

Als Schichtumfänge der Grundgesamtheit wurden die Ergebnisse des Mikrozensus herangezogen.

a) Aufteilung nach Haushaltstyp

Die Aufteilung der Stichprobe nach Haushaltstyp erfolgte disproportional: Schwach besetzte Haushaltstypen [z. B. Paare mit Kind(ern)] sind mit einem höheren Auswahlatz als stärker besetzte Haushaltstypen (Eiersonenhaushalte, Paare ohne Kind) zu erfassen. Der Grad der Disproportionalität wird durch einen Exponenten gesteuert, der für diese Erhebung auf 0,4 gesetzt wurde. Gleichzeitig wird für jede Ausprägung j ($j=1, \dots, 5$) des Merkmals Haushaltstyp ein Gewicht w_j eingeführt, sodass Paare mit Kind(ern) stärker erfasst und die Stichprobenumfänge für Eiersonenhaushalte und Paare ohne Kind reduziert werden können:

$$n_{L,k}'' = \frac{W_k \cdot Z_k^{0,4}}{\sum_k W_k \cdot Z_k^{0,4}} \cdot n_L$$

mit

n_L : Stichprobenumfang des Landes L

Z_k : Zahl der Haushalte des Haushaltstyps k im früheren Bundesgebiet bzw. in den neuen Ländern in der Grundgesamtheit

$n_{L,k}''$: Stichprobenumfang für Haushaltstyp k im Land L

$w_1 = 0,9$ (Eiersonenhaushalte)

$w_2 = 0,9$ (Paare ohne Kind)

$w_4 = 1,2$ [Paare mit Kind(ern)]

$w_j = 1$ für $j=3,5$

b) Aufteilung nach sozialer Stellung des Haupteinkommensbeziehers

Die Aufteilung nach dem Merkmal „Soziale Stellung“ erfolgte nach dem Verfahren von Neyman-Tschuprow mit

dem Ziel, ein möglichst präzises Gesamtergebnis über alle Schichten zu erhalten.

c) Aufteilung nach Haushaltsnettoeinkommen

Die Aufteilung nach dem Haushaltseinkommen erfolgte nach dem Verfahren von Neyman-Tschuprow mit dem Ziel, ein möglichst präzises Gesamtergebnis über alle Schichten zu erhalten.

Im nächsten Schritt wurden für jedes Land getrennt die vorher berechneten Stichprobenumfänge auf die einzelnen Schichten (Kombination der Ausprägungen der drei Schichtungsmerkmale) verteilt, und zwar so, dass einerseits die Stichprobenumfänge für die Randsummen eingehalten werden und sich andererseits die Verteilung der Stichprobenhaushalte auf die einzelnen Schichten möglichst wenig von der Verteilung in der Grundgesamtheit unterscheidet. Dieses Aufteilungsziel wird durch das sogenannte iterative proportionale Anpassungsverfahren (‘‘Iterative Proportional Fitting’’⁹⁾ erreicht. Mit der Fertigstellung der landesspezifischen Auswahlpläne für die Ziehung des neuen Rotationsviertels kann die jährliche EU-SILC-Erhebung beginnen.

Erhebungsorganisation

Die Erhebung wird – wie die anderen freiwilligen Haushaltsbefragungen der amtlichen Statistik in Deutschland – schriftlich¹⁰⁾ und dezentral durchgeführt. Während die Ziehung des neuen Rotationsviertels auf der Grundlage des zuvor vorgestellten Auswahlplans aus methodischen Gründen zentral im Statistischen Bundesamt erfolgt, finden die Feldphase und die Erfassung der zurückgesandten Erhebungsunterlagen dezentral in den Statistischen Ämtern der Länder statt.¹¹⁾ Die Feldphase beginnt im April und endet im September eines Erhebungsjahres.

Jeder Haushalt erhält einen Personenfragebogen für jedes Haushaltsmitglied über 16 Jahre (befragungsfähiges Alter) sowie einen Haushaltsfragebogen. Der Haushaltsfragebogen soll von der Person im Haushalt beantwortet werden, die sich am besten mit den Belangen des Haushalts auskennt. Nach den europäischen Rechtsgrundlagen dürfen für EU-SILC nur solche Haushalte erfasst werden, für die mindestens ein Personenfragebogen und ein Haushaltsfragebogen zurückgesandt wurden und in denen mindestens eine Stichprobenperson wohnt (Erfassungsbedingung). Erfüllt ein Haushalt diese Erfassungsbedingung nicht, so ist er als Ausfall zu registrieren. Für ausführlichere Informationen zur Erhebungsorganisation und -methodik siehe den Beitrag von Körner u. a. (Fußnote 1).

Die Paneldatenbank

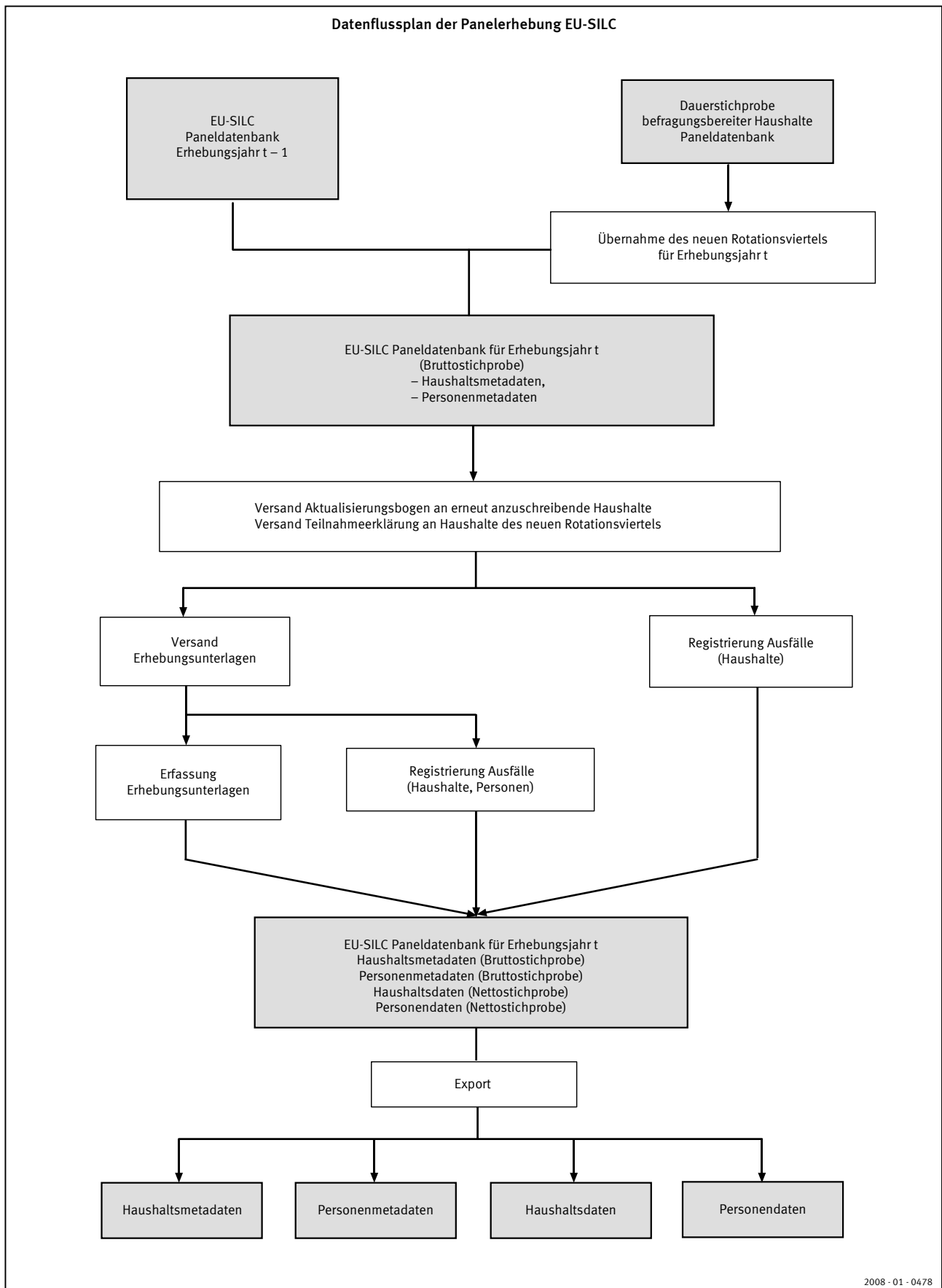
Jedes Statistische Landesamt verfügt über eine dezentrale SILC-Paneldatenbank, um die einzelnen Arbeitsschritte durchzuführen. In Schaubild 2 ist der Datenfluss für eine

9) Siehe z. B. Bishop, Y. M. M./Fienberg, S. E. /Holland, P. W.: ‘‘Discrete Multivariate Analysis: Theory and Practice’’, Cambridge 1980, S. 85 f.

10) In anderen Mitgliedstaaten wird die Erhebung telefonisch oder in Form von Face-to-face-Interviews durchgeführt.

11) Die Statistischen Ämter der Länder machen einen Auszug der für die Ziehung benötigten Daten aus ihrer landesspezifischen Datenbank der Dauerstichprobe befragungsbereiter Haushalte, den sie an das Statistische Bundesamt senden. Dort findet die Ziehung des neuen Rotationsviertels statt. Anschließend werden die Statistischen Ämter der Länder informiert, welche Haushalte gezogen wurden.

Schaubild 2



EU-SILC-Erhebung grob zusammengefasst. Die erneut anzuschreibenden Haushalte aus dem Vorjahr und die Haushalte aus dem neu gezogenen Rotationsviertel werden in die Paneldatenbank eingespeist. Sie bilden die Bruttostichprobe eines Erhebungsjahres. Die erneut anzuschreibenden Haushalte erhalten einen Aktualisierungsbogen, die neu gezogenen Haushalte dagegen eine Teilnahmeerklärung. In beiden Unterlagen wird die aktuelle Haushaltszusammensetzung für die Aktualisierung der Paneldatenbank zu Beginn der Erhebung erfragt. Dieser Arbeitsschritt ist vergleichbar mit der Pflege des Adressprotokolls durch den Interviewer in Face-to-face-Interviews.

Den Kern der Paneldatenbank bilden die Haushaltsmetadatei und die Personenmetadatei, in denen die Teilnahmebereitschaft und die Haushaltszusammensetzung seit Eintritt des Haushalts in die Panelerhebung registriert sind. Die Pflege der Metadateien ist über verschiedene Eingabemaschinen der Software möglich und basiert auf der manuellen Sichtkontrolle der zurückgesandten Aktualisierungsbogen/Teilnahmeerklärungen und der Erhebungsunterlagen (Haushalts- und Personenfragebogen). In der Haushalts- und Personendatei werden ausschließlich die erfassten Angaben aus den Erhebungsunterlagen (Haushaltsfragebogen und Personenfragebogen) gespeichert. Nach dem Abschluss der Erfassung in einem Statistischen Landesamt stehen vier Exportdateien für den Versand an das Statistische Bundesamt bereit. Im Folgejahr dienen die Metadateien zum Aufbau der EU-SILC-Paneldatenbank für die jährliche Erhebung.

Aufbereitung

Die landesspezifischen Exportdateien werden im Statistischen Bundesamt zu bundesweiten Dateien zusammengespielt und in das für Eurostat erforderliche Datenformat transformiert. Dabei werden:

- die Erhebungsmerkmale outputharmonisiert aufbereitet,
- fehlende Angaben bei den Einkommensinformationen geschätzt (Imputationsverfahren) sowie
- die Zielvariablen generiert, die sich aus mehreren Erhebungsmerkmalen zusammensetzen bzw. nicht direkt erhoben werden können.

Hochrechnung

Nach der Fertigstellung der Mikrodaten für Eurostat erfolgt die Hochrechnung nach den methodischen Vorgaben und Empfehlungen von Eurostat. Im Jahr 2005 wurde bei der Hochrechnung der Querschnittsdaten allerdings noch ein Verfahren angewandt, das nicht den methodischen Vorgaben von Eurostat entsprach. Für das Jahr 2006 wurde das Hochrechnungsverfahren für die deutsche EU-SILC-Erhebung weiterentwickelt und anschließend rückwirkend auf die Mikrodaten des ersten Erhebungsjahres 2005 ange-

wandt, um die Interpretation von Zeitreihen insbesondere zu Armut und Einkommen auf der Grundlage eines einheitlichen und den methodischen Vorgaben von Eurostat entsprechenden Hochrechnungsverfahrens zu ermöglichen. Dies führte zu einer Revision des Haushaltsgewichtungsfaktors (DB090) und des Personengewichtungsfaktors für alle Personen (RB050) in den Mikrodaten des Erhebungsjahres 2005. Das neue Hochrechnungsverfahren für die Querschnittsgewichte wird im Folgenden vorgestellt.

Gewichtungsfaktoren

Das neue Hochrechnungsverfahren basiert auf einer Doppelkalibrierung: Alle Personen eines Haushalts erhalten das gleiche Querschnittsgewicht (RB050) und dieses Gewicht wird auch für Analysen auf Haushaltsebene (DB090) angewendet. Für jeden Haushalt gibt es folglich nur ein Gewicht, das sowohl für Haushalts- als auch für Personenanalysen (Gesamtbevölkerung) zum Einsatz kommt (RB050 = DB090). Für spezifische Analysen von Personen im Alter von 16 Jahren und mehr und von Kindern unter 12 Jahren gibt es zusätzlich noch weitere Gewichtungsfaktoren (PB040, RL070). In Übersicht 1 sind alle in den EU-SILC-Mikrodaten enthaltenen Gewichtungsfaktoren noch einmal aufgeführt:

Übersicht 1: Querschnittsgewichte in EU-SILC

Analyseebene	Gewicht	Bezugspopulation für das Gewicht
Haushalte Personen	DB090	Haushalte
	RB050	Alle Haushaltsmitglieder
	RL070	Kinder bis 12 Jahre
	PB040	Personen über 16 Jahre

Hochrechnungsrahmen

Den Hochrechnungsrahmen für sämtliche Anpassungen der Stichprobe im jeweiligen Erhebungsjahr bildeten die Ergebnisse der Mikrozensus-Erhebung aus diesem Jahr.¹²⁾ Die Auswahl und Kombination der Hochrechnungsmerkmale erfolgte – soweit verfahrensbedingt möglich – vor dem Hintergrund des Verwendungszwecks und der potenziell möglichen Analysetiefe. Für die Erzeugung des oben vorgestellten Haushalts- und Personengewichtungsfaktors (RB050, DB090) waren aufgrund der erforderlichen Doppelkalibrierung die mögliche Gesamtzahl und damit auch die Kombinationsmöglichkeiten der Hochrechnungsmerkmale stark beschränkt. Eine Anpassung an die Haushaltsebene durfte nicht zu Lasten der Anpassung an die Personenebene gehen und umgekehrt. Es musste für beide Analyseebenen sichergestellt werden, dass dieses Gewicht die soziodemografischen Strukturen sowohl von Haushalten als auch von der Gesamtbevölkerung in Deutschland möglichst angemessen abbildet. Ferner wurde auf die Einbeziehung der Merkmale „Bundesland“ und „Nationalität“ verzichtet, weil bei beiden Merkmalen nur stark begrenzte Analysemöglichkeiten bestehen und ihre Einbeziehung weder die Analyse mög-

12) Beim alten Hochrechnungsverfahren für die EU-SILC-Daten des Erhebungsjahres 2005 wurden die Ergebnisse der Mikrozensus-Erhebung 2004 herangezogen; bei der rückwirkenden Anwendung des neuen Verfahrens (Revision) auf die Daten des Jahres 2005 dagegen die Ergebnisse der Mikrozensus-Erhebung 2005. Da sich die Revision nur auf die Gewichtungsfaktoren DB090 und RB050 bezieht und die Gewichtungsfaktoren PB040 und RL070 im Jahr 2005 unverändert blieben, haben die ersten beiden Gewichtungsfaktoren den Mikrozensus 2005 als Hochrechnungsrahmen und die letzten beiden Gewichtungsfaktoren den Mikrozensus 2004 als Hochrechnungsrahmen.

lichkeiten nach diesen Merkmalen deutlich erhöht noch die Hochrechnung insgesamt verbessert hätte.

Die Ermittlung des Gewichtungsfaktors für Personen im Alter von 16 Jahren und mehr (PBO40) erfolgte auf der Grundlage einer eigenen Hochrechnung, um die soziodemografische Struktur dieser Population (Familienstand, soziale Stellung der Personen usw.) möglichst optimal abzubilden. Dies wäre im Rahmen einer Doppelkalibrierung nicht möglich gewesen bzw. hätte für diese Analysepopulation zu Qualitätsverlusten der Ergebnisse geführt. Hier wurde deshalb das gleiche Hochrechnungsverfahren wie im Jahr 2005 angewendet. Ab dem Jahr 2006 wurde der Hochrechnungsrahmen für dieses Gewicht lediglich um die Merkmale Bildungsabschluss und Haushaltstypzugehörigkeit erweitert, um die Datenqualität zu erhöhen.

Das Hochrechnungsverfahren für Kinder unter 12 Jahren (RLO70) blieb im Jahr 2006 gegenüber 2005 unverändert.

Anpassung an Auswahlwahrscheinlichkeiten und Antwortausfälle

Die in EU-SILC zu verwendenden Gewichtungsfaktoren sollen eine Anpassung an unterschiedliche Auswahlwahrscheinlichkeiten (hier: Teilnahme- und Bleibewahrscheinlichkeit an/in der Dauerstichprobe befragungsbereiter Haushalte) und Antwortausfälle ermöglichen. Darüber hinaus wird die Stichprobe – wie bereits erwähnt – an einen Hochrechnungsrahmen auf der Basis von externen Datenquellen (Mikrozensus) angepasst (siehe Übersicht 2).

Zur Korrektur der durch Antwortausfälle verursachten Verzerrung (Nonresponse-Bias), die im Zuge der Anwerbung für die Dauerstichprobe befragungsbereiter Haushalte entsteht, erhalten alle Einheiten der Dauerstichprobe befragungsbereiter Haushalte ein Gewicht. Dieses Gewicht wird unter Nutzung der aus dem letzten Mikrozensusinterview übernommenen Daten bzw. der aktuellen Profildaten unter Anwendung einer multivariaten Modellierung (Logitmodell) der individuellen Teilnahme-/Bleibewahrscheinlichkeiten ermittelt. Bei der Gewichtung der Ergebnisse von EU-SILC fließt dieses Gewicht in die Berechnung der Auswahlwahrscheinlichkeiten ein.

Die Berechnung der Teilnahme- und Bleibewahrscheinlichkeiten wurde bereits in dem Beitrag von Körner u. a. (2006)¹³⁾ beschrieben und wird hier deshalb nicht vorgestellt.

Schätzmethode (Gewichtung)

Allgemeines Ziel einer Gewichtung ist es, bestimmte Parameter der Grundgesamtheit (Gesamtwert, Mittelwert, Anteilswert, Varianz) mithilfe geeigneter Schätzfunktionen aus der Stichprobe zu schätzen.

Einen Schätzwert für den unbekanntenen Gesamtwert eines interessierenden Merkmals Y liefert ein verallgemeinerter Regressionsschätzer. Die lineare Schätzfunktion für den Totalwert t_y einer entsprechenden Variablen y lässt sich darstellen als:

$$\begin{aligned} \hat{t}_y &= \hat{t}_{y,HT} + \hat{B}' \cdot (t_x - \hat{t}_{x,HT}) \\ &= \sum_{k=1}^n \left(1 + (t_x - \hat{t}_{x,HT})' \left(\sum_{k=1}^n d_k x_k x_k' \right)^{-1} x_k \right) d_k y_k \\ &= \sum_{k=1}^n g_k d_k y_k = \sum_{k=1}^n W_k y_k \end{aligned}$$

mit

w_k : Gewichtungsfaktor für Haushalt/Person k

g_k : Korrekturfaktor für Haushalt/Person k

y_k : Wert des Erhebungsmerkmals y für Haushalt/Person k

n : (Netto-)Stichprobenumfang

und wobei

$$\hat{t}_{y,HT} = \sum_{k=1}^n \frac{y_k}{\pi_k \hat{\theta}_k} = \sum_{k=1}^n d_k y_k$$

der frei hochgerechnete Totalwert für die Variable Y ist („Horvitz-Thompson-Schätzer“).

Des Weiteren ist:

π_k : die Wahrscheinlichkeit für Haushalt/Person k , für EU-SILC ausgewählt zu werden. Sie ist das Produkt aus der

- Wahrscheinlichkeit, im herausrotierenden Viertel des Mikrozensus zu sein,
- Wahrscheinlichkeit, an der Dauerstichprobe befragungsbereiter Haushalte teilzunehmen (Teilnahmewahrscheinlichkeit an der Dauerstichprobe befragungsbereiter Haushalte),
- Wahrscheinlichkeit, in der Dauerstichprobe befragungsbereiter Haushalte zu verbleiben (Bleibewahrscheinlichkeit in der Dauerstichprobe befragungsbereiter Haushalte) und
- Wahrscheinlichkeit eines in der Dauerstichprobe vertretenen Haushalts, für EU-SILC ausgewählt zu werden.

$\hat{\theta}_k$: Wahrscheinlichkeit eines für EU-SILC ausgewählten Haushalts, an EU-SILC teilzunehmen (Teilnahmewahrscheinlichkeit an der Erhebung)

x_k : Vektor aller Ausprägungen der Hilfsmerkmale für Haushalt/Person k

t_x : Vektor der Totalwerte der Hilfsmerkmale („Eckwerte“)

$$\hat{B} = \left(\sum_{k=1}^n d_k x_k x_k' \right)^{-1} \left(\sum_{k=1}^n d_k x_k y_k \right): \text{Vektor der geschätzten Regressionskoeffizienten}$$

Der Regressionsschätzer ist eine lineare Schätzfunktion und hat die Eigenschaft, dass die Eckwerte getroffen werden, wenn sie aus der Stichprobe hochgerechnet werden, das heißt

$$\hat{t}_x = t_x.$$

13) Siehe Fußnote 7.

Übersicht 2: Hochrechnungsrahmen

Gewicht	Kalibrierungsmerkmale bzw. Hochrechnungsterm
DB090 RB050	Die Kalibrierung erfolgte separat für die Merkmale Haushaltstyp, Haushaltsgröße und Einkommen für die Haushalte und separat für die Merkmale Alter und Geschlecht für die Personen. – Haushaltstyp (5) ¹⁾ – Haushaltsgröße (5) – Einkommen (6) – Alter (7) – Geschlecht (2)
PB040	Die Kalibrierung erfolgte auf Personenebene. – Bundesland (17) – Familienstand (4) – Bildung (3) – soziale Stellung (4) – Haushaltstyp (5) – Alter (7) nach alte/neue Bundesländer – Geschlecht (2) nach alte/neue Bundesländer – Nationalität (2) nach alte/neue Bundesländer
RL070	Die Kalibrierung erfolgte auf Personenebene. – Alter (7) – Geschlecht (2) nach alte/neue Bundesländer

1) Zahl der Ausprägungen in Klammern.

Für die oben bereits erwähnten Gewichte der Querschnittskomponente wurden die in der Übersicht 2 dargestellte Kalibrierungsmodelle verwendet.

Empirische Ergebnisse zur Datenqualität

Strukturvergleich von ausgewählten Ergebnissen aus EU-SILC mit den Ergebnissen des Mikrozensus

Im Folgenden wird anhand ausgewählter Ergebnisse aus EU-SILC und aus dem Mikrozensus gezeigt, inwiefern das verwendete Hochrechnungsmodell die soziodemografische Struktur der Bevölkerung adäquat abbildet (siehe die Tabellen 3 und 4). Für die Ergebnisse zur Gesamtbevölkerung (Tabelle 3) wurde der Gewichtungsfaktor für alle Personen (RB050) verwendet, für die Ergebnisse zu Personen im Alter von 16 Jahren und mehr (Tabelle 4) dagegen der Gewichtungsfaktor für Personen über 16 Jahre (PB040).

Im Vergleich zu den Ergebnissen des Mikrozensus ist beim Merkmal Haushaltstyp der Anteil der Gruppe „sonstige Haushalte“ bei EU-SILC geringer; im Jahr 2006 sogar um 3,3 Prozentpunkte. Bei den Alleinerziehenden und den Paaren mit Kind(ern) liegt der Anteil bei EU-SILC dagegen in beiden Jahren geringfügig höher als im Mikrozensus. Beim Merkmal „Alter“ gibt es in den hier ausgewiesenen Alterskategorien keine Abweichungen im Vergleich zum Mikrozensus.

Die Ergebnisse zum Merkmal „Nationalität“ für Personen über 16 Jahre belegen zwar ähnlich hohe Anteile bei EU-SILC und Mikrozensus, aber tiefer gegliederte Auswertungen nach diesem Gliederungsmerkmal zeigen, dass die Zusammensetzung dieser Gruppe bei EU-SILC stark von der im Mikrozensus abweicht. Es ist daher notwendig, die Repräsentativität in diesem Bereich zu verbessern.

Tabelle 3: Strukturvergleich – alle Personen¹⁾

Gegenstand der Nachweisung	Prozent			
	Mikrozensus 2005	EU-SILC 2005	Mikrozensus 2006	EU-SILC 2006
Haushaltstyp ²⁾				
Einpersonenhaushalt	17,5	17,5	18,1	18,2
Paare ohne Kind	27,0	27,0	26,9	27,0
Alleinerziehende	4,5	4,8	4,6	5,5
Paare mit Kind(ern)	33,6	34,0	32,4	34,7
Sonstige Haushalte	17,4	16,7	17,9	14,6
Insgesamt ...	100	100	100	100
Alter				
bis 15 Jahre	15,9	15,9	15,5	15,5
16 – 24 Jahre	10,8	10,8	10,8	10,8
25 – 39 Jahre	12,2	12,2	12,0	12,0
40 – 54 Jahre	17,0	17,0	16,9	16,9
55 – 69 Jahre	14,2	14,2	14,4	14,4
70 – 79 Jahre	12,0	12,0	11,8	11,8
80 Jahre und älter	17,9	17,9	18,6	18,6
Insgesamt ...	100	100	100	100
Geschlecht				
Männlich	49,1	49,1	49,1	49,0
Weiblich	50,9	50,9	50,9	51,0
Insgesamt ...	100	100	100	100

1) Gewichtungsfaktor RB050. – 2) Paare ohne Kind = 2 Erwachsene (Partner lebt im Haushalt); Alleinerziehende = 1 Erwachsener mit Kind(ern) unter 27 Jahren, von denen mindestens ein Kind unter 18 Jahre alt ist.

Bei den Merkmalen „Erwerbsstatus“ und „Familienstand“ weichen die Ergebnisse aus EU-SILC für das Erhebungsjahr 2005 etwas von denen des Mikrozensus ab. Dies liegt möglicherweise an den für das Gewicht (PB040) im Jahr 2005 als Hochrechnungsrahmen verwendeten Ergebnissen aus dem Mikrozensus 2004 (siehe Erläuterungen in Fußnote 12).

Tabelle 4: Strukturvergleich – Personen über 16 Jahre¹⁾

Gegenstand der Nachweisung	Prozent			
	Mikrozensus 2005	EU-SILC 2005	Mikrozensus 2006	EU-SILC 2006
Nationalität				
Deutsch	91,3	91,6	91,2	91,2
Nicht-Deutsch	8,7	8,4	8,8	8,8
Insgesamt ...	100	100	100	100
Erwerbsstatus ²⁾				
Erwerbstätig	52,8	51,1	53,8	53,8
Arbeitslos	5,9	5,2	5,8	5,8
Im Ruhestand	24,9	26,4	25,0	25,1
Sonstig nicht erwerbstätig ..	16,3	17,4	15,4	15,4
Insgesamt ...	100	100	100	100
Familienstand				
Ledig	27,8	26,7	28,3	28,4
Verheiratet ³⁾	57,0	57,9	56,2	56,1
Verwitwet	8,3	8,7	8,3	8,4
Geschieden	6,9	6,7	7,1	7,1
Insgesamt ...	100	100	100	100
höchster Bildungsabschluss ⁴⁾				
Niedrig (ISCED 0, 1, 2)	26,5	17,7	26,2	26,3
Mittel (ISCED 3, 4)	53,8	48,9	54,5	54,5
Hoch (ISCED 5, 6)	19,7	33,5	19,2	19,2
Insgesamt ...	100	100	100	100

1) Gewichtungsfaktor PB040. – 2) Erwerbsstatus nach Selbsteinschätzung. – 3) Verheiratet zusammen lebend und verheiratet getrennt lebend. – 4) International Standard Classification of Education (ISCED).

Tabelle 5: Rücklauf von den Haushalten nach Rotationsvierteln

Gegenstand der Nachweisung	Insgesamt	Quotenhaushalte			Dauerstichprobe befragungsbereiter Haushalte	
		Rotationsviertel				
		1	2	3	4	5
2005						
Bruttostichprobe	15 231	–	–	–	6 110	–
Nettostichprobe	13 106	2 528	3 066	3 527	3 985	–
Rücklaufquote in %	–	–	–	–	65,2	–
2006						
Bruttostichprobe	16 927	–	3 141 ¹⁾	3 589 ¹⁾	4 029 ¹⁾	6 168
Nettostichprobe	13 799	–	2 894	3 316	3 530	4 059
Rücklaufquote in %	–	–	92,1	92,4	87,6	65,8

1) Einschl. Splithaushalten bzw. Personen aus Ursprungshaushalten, die aufgrund eines Auszugs als Abspaltung vom Haushalt (Splithaushalt) registriert sind.

Beim Merkmal „Bildungsabschluss“ belegen die Ergebnisse einerseits eine bessere Datenqualität aufgrund der Einbeziehung dieses Merkmals in die Hochrechnung ab dem Erhebungsjahr 2006, andererseits aber auch die Notwendigkeit dieser Einbeziehung.¹⁴⁾

Teilnahmebereitschaft

Durch die Einbeziehung von Quotenhaushalten in den ersten drei Jahren der EU-SILC-Erhebung ist eine Bewertung des Rücklaufs schwierig. Ein Urteil darüber ist frühestens im Jahr 2008 möglich, da sich die EU-Stichprobe in diesem Jahr zum ersten Mal ausschließlich aus nach dem Zufallsprinzip ermittelten Haushalten der Dauerstichprobe befragungsbereiter Haushalte zusammensetzen wird.

Als erfasste Haushalte (Nettostichprobe) werden in die Analyse des Rücklaufs nur solche Haushalte einbezogen, die erfolgreich die in den Rechtsgrundlagen von EU-SILC definierten Erfassungsbedingungen für EU-SILC erfüllt haben. Für das Rotationsviertel 4, welches die erste Teil-Zufallsstichprobe von Haushalten der Dauerstichprobe befragungsbereiter Haushalte in der EU-SILC repräsentierte, be-

trug die Rücklaufquote im Erhebungsjahr 2005 rund 65 % (siehe Tabelle 5). Ein ähnliches Ergebnis ist auch für die zweite Teil-Zufallsstichprobe von Haushalten der Dauerstichprobe befragungsbereiter Haushalte im Erhebungsjahr 2006 (Rotationsviertel 5) festzustellen (knapp 66%). Im Hinblick auf die Teilnahmebereitschaft der Haushalte im zweiten Erhebungsjahr von EU-SILC war der Rücklauf von den Haushalten aus den Rotationsvierteln 2 und 3 geringfügig höher als der beim Rotationsviertel 4 der Dauerstichprobe befragungsbereiter Haushalte.¹⁵⁾ Er betrug bei den Haushalten dieser drei 2006 erneut zu befragenden Rotationsviertel rund 90 %, wobei die inzwischen entstandenen Abspaltungen (Splithaushalte) von diesen Haushalten mit berücksichtigt sind.

Der Rücklauf von den befragungsfähigen Personen aus Haushalten mit erfüllter Erfassungsbedingung (Nettostichprobe) betrug in allen Rotationsvierteln deutlich über 95 % (siehe Tabelle 6). Das bedeutet, dass eine deutliche Mehrheit der Haushalte in beiden Erhebungsjahren und unabhängig von der Form der Stichprobenziehung bereit war, alle zugesandten Personenfragebogen auch auszufüllen und zurückzusenden. Der Anteil von fehlenden Einkommensinformationen (partieller Unit-Nonresponse) aufgrund fehlender Personen-

Tabelle 6: Rücklauf von den befragungsfähigen Personen nach Rotationsvierteln

Gegenstand der Nachweisung	Insgesamt	Quotenhaushalte			Dauerstichprobe befragungsbereiter Haushalte	
		Rotationsviertel				
		1	2	3	4	5
2005						
Bruttostichprobe	25 116	5 010	6 010	6 826	7 270	–
Nettostichprobe	24 982	4 980	5 987	6 808	7 207	–
Rücklaufquote in %	–	99,4	99,6	99,7	99,1	–
Selbstausfüller	22 012	–	–	–	–	–
Anteil in der Nettostichprobe in %	–	86,6	87,5	88,3	89,4	–
2006						
Bruttostichprobe	26 103	–	5 652 ¹⁾	6 452 ¹⁾	6 456 ¹⁾	7 543
Nettostichprobe	25 979	–	5 634	6 433	6 422	7 490
Rücklaufquote in %	–	–	99,7	99,7	99,5	99,2
Selbstausfüller	20 618	–	–	–	–	–
Anteil in der Nettostichprobe in %	–	–	77,1	78,1	79,5	82,0

1) Einschl. Splithaushalten bzw. Personen aus Ursprungshaushalten, die aufgrund eines Auszugs als Abspaltung vom Haushalt (Splithaushalt) registriert sind.

14) Die Verteilung der deutschen Bildungsabschlüsse auf der Grundlage ungewichteter Ergebnisse belegt in beiden Jahren einen ähnlich hohen Bias zuungunsten von Personen mit niedrigem Bildungsabschluss.

15) Dieser Unterschied wäre vermutlich noch größer, wenn die Zufallsstichprobe direkt aus der Grundgesamtheit gezogen würde, weil die Dauerstichprobe befragungsbereiter Haushalte als Auswahlgrundlage nur grundsätzlich teilnahmebereite Haushalte enthält.

fragebogen ist in EU-SILC folglich sehr gering bzw. besitzt nur eine geringe empirische Bedeutung für die Bewertung der Datenqualität.

Proxy-Interviews

Nach den EU-SILC-Durchführungsverordnungen sind Proxy-Interviews¹⁶⁾ aus Qualitätsgründen weitestgehend zu vermeiden. Die Handlungsmöglichkeiten im Hinblick auf die Vermeidung solcher Interviews hängen von der Erhebungsform ab und sind bei schriftlichen Erhebungen am geringsten. Der Idealfall ist, bei allen Erhebungsformen einen möglichst hohen Anteil von „Selbstaussfüllern“ und möglichst wenig fehlende Angaben in den Erhebungsunterlagen (Item-Nonresponse) zu erhalten. Der geringere Anteil der Selbstaussfüller unter den erfolgreich Befragten im zweiten Erhebungsjahr belegt einen Anstieg der Proxy-Interviews in den Haushalten (siehe Tabelle 6).

Eine weitere Deutung dieses Ergebnisses ist auf der Basis von nur zwei Erhebungswellen noch nicht möglich, es sollte aber in den Folgejahren näher untersucht werden, um mehr Informationen über die Teilnahmebereitschaft im Zeitverlauf und über die Datenqualität der „Proxy-Einkommensinformationen“ zu erhalten.¹⁷⁾

Ferner ist festzustellen, dass sich die teilnahmebereiten Quotenhaushalte bezüglich Rücklauf und Anteil von Proxy-Interviews nicht von den Zufallshaushalten aus der Dauerstichprobe befragungsbereiter Haushalte unterscheiden. Dieses Ergebnis ist bemerkenswert, weil die beiden Teilstichproben aus Haushalten mit unterschiedlichem Erfahrungshintergrund bestehen: Während die Zufallshaushalte aus der Dauerstichprobe befragungsbereiter Haushalte zuvor jahrelang an einer Pflichterhebung (Mikrozensus) mit mündlichen Interviews teilnahmen, wurden die Quotenhaushalte teilweise aus vorangegangenen – für die Haushalte teilweise sehr aufwendigen – schriftlichen Erhebungen auf freiwilliger Basis wie der Einkommens- und Verbrauchsstichprobe 2003 oder den Laufenden Wirtschaftsrechnungen 2004 angeworben.

Bisher durchgeführte Maßnahmen zur Verbesserung der Datenqualität

Überarbeitung der Erhebungsunterlagen

Art und Umfang der Rückfragen im ersten Erhebungsjahr führten zu einer Überarbeitung der Erhebungsunterlagen,

insbesondere der Filterführungen und der Struktur der retrospektiven Einkommensfragen. Die Zuordnung einzelner Einkommenskomponenten sollte dadurch erleichtert werden.

Einführung eines Aktualisierungsbogens

Des Weiteren wurde ab dem Erhebungsjahr 2007 ein Aktualisierungsbogen für die erneut anzuschreibenden Haushalte eingeführt, um Veränderungen in der Haushaltszusammensetzung – insbesondere Zuzüge und Wegzüge von Personen – bereits am Beginn der Erhebung zu erfahren. Zeitgleich mit dem Versand der Teilnahmeerklärung an die Haushalte des neuen Rotationsviertels werden die Aktualisierungsbogen an die erneut zu befragenden Haushalte versandt. Auf diese Weise gewinnt das Statistische Landesamt mehr Zeit und Flexibilität für die Weiterverfolgung¹⁸⁾ von Stichprobenpersonen, die seit dem Ende der letzten Erhebung ausgezogen waren. Ferner erübrigt sich weitestgehend die Nachsendung von Personenfragebogen an neu eingezogene Personen.

Einführung von Umzugsformularen für die Aktualisierung der Datenbanken

Während bei zentral organisierten Haushaltsbefragungen der Umzug eines Haushalts in ein anderes Bundesland lediglich eine Änderung der Adressinformationen in einer Datenbank nach sich zieht, muss bei landesspezifisch organisierten Erhebungen der Umzug physisch in mehreren Datenbanken nachvollzogen werden. Im Herkunftsland muss ein maschineller Export der Metainformationen erfolgen; im Zielland müssen diese Informationen in die Datenbank eingepflegt werden (Import). Zugleich muss dieser Umzug in den landesspezifischen Datenbanken der Dauerstichprobe befragungsbereiter Haushalte berücksichtigt werden.¹⁹⁾ Ab dem Erhebungsjahr 2008 wurden deshalb für die bereits bestehende Schnittstelle zum physischen Umzug der Haushalte zusätzlich ein Umzugsformular für die EU-SILC-Datenbank und ein Umzugsformular für die Datenbank der Dauerstichprobe befragungsbereiter Haushalte eingeführt. Beide Formulare werden vom Herkunftsland ausgefüllt und geben dem Zielland einen Überblick über den Stand der Bearbeitung beim betroffenen Haushalt (z.B. Versand der Erhebungsunterlagen ist bereits erfolgt).

16) Gemeint sind hier Interviews über Dritte bzw. dass ein Haushaltsmitglied für ein anderes Haushaltsmitglied die Erhebungsunterlagen ausfüllt. Im Fall schriftlicher Befragungen bedeutet dies, dass der Personenfragebogen nicht von der Person selbst ausgefüllt wurde. Ein Vergleich solcher Informationen mit denen aus anderen EU-SILC-Ländern ist nur bei gleichen Erhebungsformen sinnvoll, weil die Rahmenbedingungen für die Interviewsituation zu unterschiedlich sind. Z. B. macht es einen Unterschied, ob ein Haushaltsmitglied aus Zeitgründen am Telefon Fragen zum Gesundheitszustand über ein anderes Haushaltsmitglied beantwortet oder die Erhebungsunterlagen zu Hause von einer Person A für eine Person B ausgefüllt werden und sich Person B im Raum befindet.

17) Relevant für diese Betrachtung ist ferner der Anteil der Selbstaussfüller je Haushalt in den verschiedenen Erhebungsjahren, um herauszufinden, ob der Haushalt (a) Proxy-Interviews gezielt aus Zeitgründen nutzt oder (b) sich über eine Zunahme von Proxy-Interviews in einem Haushalt eine „abgestufte“ Verweigerung des gesamten Haushalts abzeichnet. Mit Abstufung ist hier gemeint, dass zuerst ein Teil der Haushaltsmitglieder nicht mehr teilnahmebereit ist und die ursprünglich hohe Teilnahmebereitschaft der verbleibenden Haushaltsmitglieder aufgrund ihrer gestiegenen persönlichen Belastung durch die Befragung (Ausfüllen mehrerer Erhebungsunterlagen) mittelfristig ebenfalls sinkt und/oder der Anteil von Item-Nonresponse signifikant zunimmt. Letzteres kann ebenfalls bei der ersten Gruppe (a) im Laufe der Jahre als negative Erfahrung hinzukommen und somit das ursprüngliche „Zeitargument“ aufheben. Wie hoch der Anteil der beiden Gruppen (a, b) unter den Haushalten mit Proxy-Interviews ist, kann erst auf der Grundlage mehrerer Befragungswellen näher untersucht werden. Bei der Interviewdauer, die in schriftlichen Erhebungen allerdings grundsätzlich nur eine sehr grobe Information darstellt, gibt es keine auffallenden Unterschiede zwischen der Dauer von Proxy-Interviews und der Dauer von Interviews.

18) Sofern es sich bei der neuen Adresse der ausgezogenen Stichprobenperson um einen Privathaushalt handelt.

19) Die Pflege der Datenbanken der EU-SILC und der Dauerstichprobe befragungsbereiter Haushalte liegt nur bei einem Teil der Statistischen Ämter der Länder in einer Hand.

Auswahlplan: Regionale Zusammenlegungen

Für das Erhebungsjahr 2008 wurde getestet, ob die Qualität der Gesamtstichprobe durch Zusammenlegung von Kombinationen der Schichtungsmerkmale Bundesland, Haushaltstyp, soziale Stellung und Haushaltsnettoeinkommen bei der Ziehung des neuen Rotationsviertels erhöht werden kann. Das Ergebnis war nicht zufriedenstellend, da in einigen Ländern der Stichprobenumfang zu gering war und alle Haushalte für den Aufbau des Rotationsviertels herangezogen werden mussten. Für das Erhebungsjahr 2008 wurde deshalb der bisher verwendete Auswahlplan beibehalten und lediglich die Reihenfolge der Schichtungsmerkmale geändert (Haushaltstyp, soziale Stellung, Haushaltsnettoeinkommen, Bundesland). Durch diese Änderung wurde das Merkmal „Bundesland“ im Erhebungsjahr 2008 zu einer Nebenbedingung bei der Schichtung.

Überlegungen zu weiteren Qualitätsmaßnahmen

Aktualisierung der Haushalts- zusammensetzung in der Paneldatenbank

Bei der Organisation der Feldphase stellte sich in den ersten Erhebungsjahren heraus, dass eine Aktualisierung der Haushaltszusammensetzung in der Paneldatenbank aufgrund der schriftlichen Erhebungsform schwierig ist und umfangreiche maschinelle Plausibilitätsprüfungen der Angaben von Mitgliedern eines Haushalts erfordert. Zum Beispiel wurden Personen im Haushaltsfragebogen als ausgezogen angegeben (z. B. Partner, erwachsene Kinder), aber es lag zugleich ein Personenfragebogen von dieser Person vor. In anderen Fällen gab eine Person X die Person Y als Partner im Personenfragebogen an, aber die Person Y teilte in ihrem Personenfragebogen mit, dass sie keinen Partner hat. Hier sind schärfere Kombinationsprüfungen zu entwickeln, um die Qualität der erfassten Daten zu erhöhen und den Nachbereitungsaufwand zu verkürzen.

Bildungsabschluss

Für eine schriftliche Panelerhebung zum Thema Armut und Deprivation ist das nicht repräsentative Ergebnis zum Merkmal „Bildung“ nicht überraschend. Dennoch müssen Kompensationsmaßnahmen gefunden werden bzw. muss überlegt werden, inwiefern der Anteil von Personen mit niedrigem Bildungsabschluss in der Stichprobe erhöht werden kann. Die Untersuchungen hierzu sind noch nicht abgeschlossen.

Migranten

Inwiefern die Zahl der Migranten in der Stichprobe erhöht werden kann, ist ebenfalls zu diskutieren. Eine Übersetzung der Erhebungsunterlagen ist hierfür nicht hinreichend, da zur Beantwortung der Fragen auch Kenntnisse über das deutsche Sicherungssystem (Steuern, Sozialabgaben, soziale Transfers) vorliegen müssen. Alternativ könnte

der bereits aus dem Mikrozensus bekannte Interviewer die Befragung mündlich durchführen. Die Sprachkompetenz wäre dann zwar auf beiden Seiten (Interviewer und Befragter) bekannt, aber beide Beteiligten haben keine Erfahrungen mit Einkommensfragen. Insofern müssten die Mikrozensus-Interviewer entsprechend geschult werden.

Überlegungen zu weiteren Qualitäts- maßnahmen auf europäischer Ebene

Bisher durchgeführte Ländervergleiche zeigen, dass das Instrument der Outputharmonisierung im Bereich der subjektiven Fragen nicht zwangsläufig zu einer aussagekräftigen Vergleichbarkeit der Ergebnisse führt. Zum einen können bereits geringfügige Abweichungen bei der Formulierung der subjektiven Fragen zu verschiedenen Frageintentionen führen, zum anderen haben die einzelnen Fragen in den Mitgliedstaaten aufgrund der Unterschiede in den Lebensbedingungen nicht immer die gleiche Relevanz. Letzteres erschwert insbesondere die Berechnung von aussagekräftigen Deprivationsindikatoren für die europäische Sozialberichterstattung. Für das Erhebungsjahr 2009 beschlossen die Mitgliedstaaten der Europäischen Union deshalb, das Schwerpunktmodul für eine Evaluierung der subjektiven Fragen in EU-SILC zu nutzen.

Schlussbetrachtung

Die methodisch bedeutsamste Weiterentwicklung seit Beginn der Erhebung stellt die erstmalige Anwendung des von Eurostat geforderten Hochrechnungsverfahrens im Jahr 2006 und dessen rückwirkende Anwendung für das Erhebungsjahr 2005 dar. Der Strukturvergleich (Mikrozensus, EU-SILC) belegt die zeitliche Robustheit der empirischen Ergebnisse und die erreichte Datenqualität im Hinblick auf die für die Gemeinschaftsstatistik zentralen soziodemografischen Gliederungsmerkmale.

Bei den Merkmalen „Bildung“ und „Nationalität“ zeigen die Ergebnisse, dass die Anteile von Personen mit niedrigem Bildungsstand und von Personen mit Migrationshintergrund in der Stichprobe erhöht werden müssen. Die Untersuchungen hierzu sind noch nicht abgeschlossen. Eine Möglichkeit wäre, diese beiden Gruppen bei der Ziehung des neuen Rotationsviertels stärker zu berücksichtigen. Voraussetzung dafür ist eine hinreichend große Ziehungsbasis. Der Aufbau der Dauerstichprobe befragungsbereiter Haushalte muss also auch im Hinblick auf solche Optionen weiter vorangetrieben werden. Grundsätzlich ist jedoch bei der deutschen EU-SILC-Erhebung zu überlegen, inwiefern die Panelbetreuung (z. B. durch Incentives, Nachbefragung bei temporären Ausfällen) und die Erhebungsform (z. B. durch den Einsatz von mündlichen oder telefonischen Interviews) insgesamt optimiert werden können, um die Teilnahmebereitschaft dieser spezifischen Haushalte zu erhöhen und aufrechtzuerhalten. Die auf diese Weise erreichte Erhöhung der Ausschöpfungsquote würde nicht nur die Repräsentativität der Gesamtstichprobe erhöhen, sondern auch eine effizientere Nutzung der Dauerstichprobe befragungsbereiter Haushalte als Ziehungsbasis bedeuten.

Inwiefern der jährliche Austausch einer Quotenteilstichprobe durch ein neues Rotationsviertel der Dauerstichprobe befragungsbereiter Haushalte zu einer höheren Repräsentativität der Gesamtstichprobe – insbesondere bezüglich der beiden Merkmale Bildung und Nationalität – führen wird, kann frühestens Ende 2009 untersucht werden, wenn die Mikrodaten des Erhebungsjahres 2008 vorliegen bzw. sich die EU-SILC-Gesamtstichprobe erstmalig ausschließlich aus Haushalten der Dauerstichprobe befragungsbereiter Haushalte zusammensetzt.

Im Hinblick auf die allgemeine Teilnahmebereitschaft der Haushalte zeigen die Ergebnisse zum Rücklauf bei den einzelnen Rotationsvierteln die zu erwartenden zeitlichen Verläufe. Bei der Interpretation der Teilnahmebereitschaft der Haushalte im ersten Erhebungsjahr ist zu berücksichtigen, dass die EU-SILC-Erhebung die bisher einzige Haushaltsbefragung ist, deren Ziehungsgrundlage die Dauerstichprobe befragungsbereiter Haushalte ist. Ein Vergleich der Ergebnisse zum Rücklauf mit anderen ähnlich großen Stichproben auf der Grundlage der Dauerstichprobe befragungsbereiter Haushalte ist folglich nicht möglich.

Ein besonderes Problem beim Umgang mit den in einer Erhebungswelle gemachten Erfahrungen ist die zeitlich und methodisch aufwendige Outputharmonisierung der EU-SILC-Mikrodaten für die Gemeinschaftsstatistik. Nur ein geringer Teil der Erhebungsmerkmale – zum Beispiel die subjektiven Fragen zur Einschätzung der Lebensbedingungen – kann direkt in das Zielvariablenformat transformiert werden. Die Auswirkungen von Fehlern bei der Operationalisierung von Erhebungsmerkmalen oder das Fehlen von zusätzlichen Informationen bei der Generierung der Zielvariablen – insbesondere bei neu eingeführten Zielvariablen – werden in der Regel erst zu einem Zeitpunkt festgestellt, in dem die Haushalte bereits die Erhebungsunterlagen der nächsten Befragungswelle (Feldphase) ausfüllen. Erkenntnisse und Weiterentwicklungen können folglich erst in der übernächsten Befragungswelle umgesetzt werden. Im Hinblick auf die Vergleichbarkeit der Ergebnisse kommt eine weitere Verzögerung hinzu, weil Eurostat teilweise erst über einen Ländervergleich der erhobenen Mikrodaten oder bei der Durchsicht der nationalen Qualitätsberichte eigene Qualitätsanalysen durchführen kann. Andererseits wurde dadurch auch der internationale Austausch von Erfahrungen intensiviert, weil Eurostat gemeinsam mit den Mitgliedstaaten in den verschiedenen Arbeitsgruppen über methodische Weiterentwicklungen oder Erhebungsprobleme diskutiert.

In einer der nächsten Ausgaben von Wirtschaft und Statistik werden ausgewählte Ergebnisse aus EU-SILC 2005 und 2006 vorgestellt, u. a. Ergebnisse für die im Rahmen der europäischen Sozialberichterstattung verwendeten Indikatoren. [uu](#)

Auszug aus Wirtschaft und Statistik

© Statistisches Bundesamt, Wiesbaden 2008

Vervielfältigung und Verbreitung, auch auszugsweise, mit Quellenangabe gestattet.

Herausgeber: Statistisches Bundesamt, Wiesbaden

Schriftleitung: Walter Radermacher
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