

Standard-Dokumentation Metainformationen

(Definitionen, Erläuterungen, Methoden, Qualität)

zu

EU-SILC 2007

Diese Dokumentation gilt für den Berichtszeitraum:
2006/2007

Diese Statistik war Gegenstand eines [Feedback-Gesprächs zur Qualität](#) am 02.06.2005

Bearbeitungsstand: **30.06.2010**



STATISTIK AUSTRIA
Bundesanstalt Statistik Österreich
A-1110 Wien, Guglgasse 13
Tel.: +43-1-71128-0
www.statistik.at

Inhaltsverzeichnis

1. Zusammenfassung, wichtige Hinweise	3
Erhebungsgegenstand	3
Erhebungsdesign	3
Durchführung der Erhebung	4
Publikation	4
2. Allgemeine Informationen.....	4
Statistiktyp	4
Fachgebiet	4
Verantwortliche Organisationseinheit, Kontakt.....	4
Ziel und Zweck, Geschichte	4
Periodizität	5
Auftraggeber	5
Nutzer	5
Rechtsgrundlage(n).....	6
3. Statistische Konzepte, Methodik.....	6
Gegenstand der Statistik.....	6
Beobachtungs-/Erhebungs-/Darstellungseinheiten	6
Datenquellen	6
Meldeeinheit/Respondenten	6
Erhebungsform	6
Charakteristika der Stichprobe.....	6
Erhebungstechnik/Datenübermittlung	7
Erhebungsbogen (inkl. Erläuterungen).....	7
Teilnahme an der Erhebung.....	7
Erhebungs- und Darstellungsmerkmale, Maßzahlen; inkl. Definition	8
Verwendete Klassifikationen	9
Regionale Gliederung der Ergebnisse	10
4. Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen... 10	10
Datenerfassung.....	10
Signierung (Codierung)	10
Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen.....	10
Imputation (bei Antwortausfällen bzw. unvollständigen Datenbeständen)	11
Hochrechnung (Gewichtung)	15
Erstellung des Datenkörpers, (weitere) verwendete Rechenmodelle, statistische Schätzmethoden	18
Sonstige qualitätssichernde Maßnahmen	19
5. Publikation (Zugänglichkeit).....	19
Vorläufige Ergebnisse	19
Endgültige Ergebnisse	19
Revisionen	20
Publiziert in:	20
Behandlung vertraulicher Daten.....	20
6. Qualität	21
6.1. Relevanz.....	21
6.2. Genauigkeit.....	22
6.2.1. Stichprobenbedingte Effekte, Repräsentativität	22
6.2.2. Nicht-stichprobenbedingte Effekte.....	22
Qualität der verwendeten Datenquellen	22
Abdeckung (Fehlklassifikationen, Unter-/Übererfassung)	22
Antwortausfall (Unit-Non Response, Item-Non Response)	23
Messfehler (Erfassungsfehler)	25
Aufarbeitungsfehler.....	26
Modellbedingte Effekte.....	26
6.3. Rechtzeitigkeit und Aktualität.....	26
6.4. Vergleichbarkeit	27
6.5. Kohärenz	27
Hinweis auf ergänzende Dokumentationen/Publikationen	33

1. Zusammenfassung, wichtige Hinweise

EU-SILC (Gemeinschaftsstatistik über Einkommen und Lebensbedingungen; SILC steht für **S**tatistics on **I**ncome and **L**iving **C**onditions) ist eine europaweite Erhebung und bildet die Grundlage für vergleichende Einkommensstatistiken sowie für Daten zu Armut und Soziale Ausgrenzung.

Erhebungsgegenstand

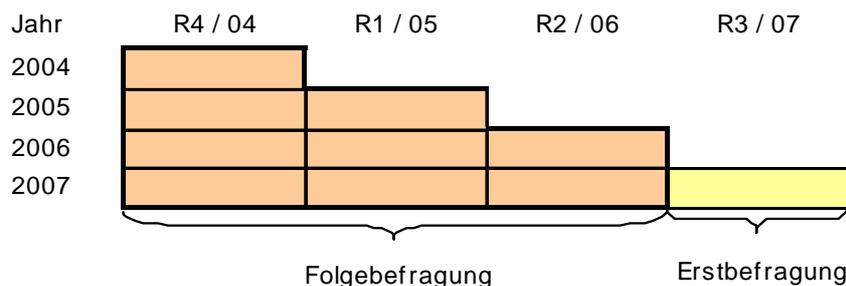
EU-SILC dient der Erfassung von Einkommen und Lebensbedingungen. Der Referenzzeitraum für das Einkommen der Haushalte und Personen und den Beschäftigungskalender ist das der Erhebung vorangegangene Kalenderjahr, 2007 also das Einkommen 2006. Alle anderen Merkmale beziehen sich auf den Erhebungszeitpunkt. Ab dem Jahr 2005 wird jährlich ein Modul in die Befragung integriert, welches jeweils ein vertiefendes Thema zur sozialen Lage und zur Lebenssituation der Bevölkerung behandelt. 2007 wurden dabei Fragen zur „Wohnsituation“ integriert.

Erhebungsdesign

Grundgesamtheit der Erhebung EU-SILC sind Personen in Privathaushalten in Österreich. Durch EU-Verordnung ist eine Mindestgröße des Samples von 4.500 Haushalten im Querschnitt und 3.250 Haushalten im Längsschnitt vorgeschrieben. Die Teilnahme an der Erhebung ist freiwillig. Nachdem EU-SILC im Jahr 2003 als reine Querschnittserhebung durchgeführt wurde, startete die Erhebung im Jahr 2004 als integrierte Längsschnittserhebung. 2007 ist somit das vierte Jahr von EU-SILC als integrierte Längsschnitterhebung

Integriertes Quer- und Längsschnittdesign heißt, dass die Stichprobe in vier etwa gleich große Teile (Rotationsgruppen) aufgeteilt wird und jedes Viertel in vier aufeinander folgenden Jahren interviewt wird. In einem Jahr werden also drei Viertel der Stichprobe wiederbefragt (Folgebefragung), und ein Viertel scheidet aus der Erhebung aus bzw. ein neues Stichprobenviertel wird in die Stichprobe aufgenommen (Erstbefragung). 2007 wird somit erstmalig eine Rotationsgruppe bereits zum vierten Mal befragt (R4). Das Grundprinzip der Rotationsstichprobe wird in untenstehender Grafik dargestellt.

Abbildung 1: Rotationsstichprobe in EU-SILC



Der Datensatz von EU-SILC 2007 enthält somit Haushalte bzw. Personen, die 2004 zum ersten Mal befragt wurden, Haushalte bzw. Personen, die 2005 oder 2006 zum ersten Mal befragt wurden und Haushalte bzw. Personen, die 2007 zum ersten Mal befragt wurden. Mit dem Jahr 2007 stehen somit auch Längsschnittdaten zur Verfügung, die 4 Jahre erfassen.

Durchführung der Erhebung

Die konkrete Art der Datenerhebung bleibt den einzelnen Ländern vorbehalten. In Österreich wird EU-SILC bislang als CAPI-Erhebung (Computer Assisted Personal Interviewing) durchgeführt.¹ Die Feldarbeit von EU-SILC 2007 wurde von zwei unabhängigen Erhebungsinstituten durchgeführt: Neben dem bereits 2005 und 2006 beauftragten Institut Spectra Marktforschungs GesmbH wurde ein Teil der Erhebung von der hauseigenen Erhebungsinfrastruktur der Bundesanstalt Statistik Austria durchgeführt. Somit gab es zwei Feldphasen: Die Erhebungsinfrastruktur der Bundesanstalt Statistik Austria interviewte von Mitte April bis Mitte September Erstbefragungshaushalte und Folgebefragungshaushalte, die im Rahmen eines Versuchs telefonisch erfasst wurden. Die Feldarbeit des Erhebungsinstituts Spectra, von dem die verbleibenden Folgehaushalte befragt wurden, dauerte von Mitte März bis Ende September.

Publikation

Eine umfassende Ergebnispublikation wurde im März 2009 veröffentlicht, ausgewählte Hauptergebnisse sind auch im Internet in Tabellenform verfügbar.

2. Allgemeine Informationen

Statistiktyp

Primärstatistische Erhebung.

Fachgebiet

Einkommen und Lebensbedingungen.

Verantwortliche Organisationseinheit, Kontakt

Bereich Soziales und Wohnen; Direktion Bevölkerung

Mag. Richard Heuberger

Tel. +43 (1) 71128- 8285, e-mail: richard.heuberger@statistik.gv.at

Mag. Matthias Till

Tel. +43 (1) 71128- 7106, e-mail: matthias.till@statistik.gv.at

Elisabeth Kafka

Tel. +43 (1) 71128- 7278, e-mail: elisabeth.kafka@statistik.gv.at

Ziel und Zweck, Geschichte

Ziel von EU-SILC ist es für alle Staaten der Europäischen Union vergleichende Daten über die soziale Lage und die Einkommen der Bevölkerung zu erhalten. Auf EU-Ebene ist EU-SILC die Grundlage für Statistiken zur Einkommensverteilung, Armut und sozialen Eingliederung. Die Ergebnisse des Projekts fließen beispielsweise in die von der Kommission jährlich erstellten Berichte zur sozialen Lage der Bevölkerung in Europa ein und bilden die Grundlage für einige der Strukturindikatoren und Laeken-Indikatoren.

Unmittelbares Vorläuferprojekt von EU-SILC war die Statistik ECHP (European Community Household Panel). Dieses Projekt diente ebenfalls als Datengrundlage für vergleichende Daten zu Einkommen, Armutsgefährdung und Lebensbedingungen auf Ebene der Europäischen Union. Das ECHP war als reine Panelerhebung konzipiert. In Österreich wurde das ECHP seit dem Jahr des EU-Beitritts 1995 durchgeführt, in allen anderen Staaten – bis auf Finnland, wel-

¹ Für die kommenden Jahre ist weiterhin die Verwendung der CATI-Technik für Folgebefragungsinterviews geplant.

ches 1996 mit dem Projekt startete – begann das ECHP im Jahr 1994. Die letzte und nach nationaler Zählung siebente Welle des ECHP wurde im Jahr 2001 erhoben.²

Nachdem es in Österreich nach dem Auslaufen des ECHP keine alternativen Datenquellen gab, um für das Jahr 2002 die auf europäischen Räten (insbesondere dem Europäischen Rat von Laeken 2001) vereinbarten Statistiken und Indikatoren über den sozialen Zusammenhalt liefern zu können, wurde im Jahr 2003 auf Basis bilateraler Verträge mit Eurostat einerseits und dem Bundesministerium für soziale Angelegenheiten und Generationen (BMSG) andererseits eine einmalige EU-SILC Querschnitterhebung durchgeführt werden. 2003 führten auch Belgien, Dänemark, Griechenland, Irland und Luxemburg, sowie Norwegen eine EU-SILC-Erhebung durch.

Ab dem Jahr 2004 wird EU-SILC in Österreich als integrierte Quer- und Längsschnitterhebung durchgeführt. Auch alle anderen EU-15-Staaten, mit Ausnahme von Deutschland, den Niederlanden und Großbritannien, begannen 2004, ebenso wie Estland, Norwegen und Island. Ab 2005 wird EU-SILC in allen bislang 25 EU-Staaten sowie in Norwegen und Island durchgeführt. Im Jahr 2006 starteten Bulgarien, Rumänien mit der Erhebung EU-SILC und 2007 wurde die Erhebung auch in der Schweiz durchgeführt. In der Türkei ist die Durchführung der Erhebung ebenfalls geplant.

2007 ist das vierte Jahr, in dem EU-SILC in Österreich als integrierte Quer- und Längsschnitterhebung durchgeführt wird. Das heißt ein Teil der Haushalte wurde nunmehr zum vierten Mal wiederbefragt, einer zum dritten, und ein Teil zum zweiten Mal.

Periodizität

EU-SILC wird jedes Jahr durchgeführt. Ab 2004 wird EU-SILC als integrierte Quer- und Längsschnitterhebung durchgeführt. Das Sample ist in vier Rotationsgruppen geteilt. Jede Rotationsgruppe verbleibt vier Jahre in der Erhebung.

Auftraggeber

Seit dem Jahr 2004 ist die Republik Österreich gemäß der EU-Verordnung (VO 1177/2003 des Europäischen Parlaments und des Rates der Europäischen Union) zur Durchführung von EU-SILC unmittelbar verpflichtet. Zur Finanzierung sind allerdings weitere Verträge notwendig. Finanziert wird EU-SILC einerseits von Eurostat und andererseits dem Bundesministerium für Arbeit, Soziales und Konsumentenschutz (BMASK). Als Auftraggeber fungieren also einerseits Eurostat und das Sozialministerium. 2008 läuft der Finanzierungsbeitrag von Eurostat aus und ab diesem Zeitpunkt ist die Erhebung komplett national zu finanzieren. Ab 2008 wird die Finanzierung somit allein durch das BMASK getragen.

Nutzer

- Europäische Union/Eurostat (v.a. zur Berechnung der Laeken- und Strukturindikatoren)
- BMASK
- Bundesländer (Landesstatistik)
- Armuts- und Sozialforschung an Universitäten, nicht-universitäre Forschungseinrichtungen und Interessensvertretungen
- Öffentlichkeit

² Für weitere Informationen über das ECHP in Österreich vgl. die Projekthomepage <http://www.iccr-international.org/echp>; für Informationen auf europäischer Ebene vgl. <http://circa.europa.eu/irc/dsis/echpanel/info/data/information.html>

Rechtsgrundlage(n)

EU-SILC wird auf der Grundlage der Verordnung des Europäischen Parlaments und des Rates der [EU VO \(EG\) Nr. 1177/2003](#) durchgeführt. Diese legt fest, dass EU-SILC durchgeführt werden soll, um die Kommission über die Einkommensverteilung, Armut und soziale Ausgrenzung in den Mitgliedstaaten zu informieren.

Weiters regeln fünf ausführende Kommissionsverordnungen, sowie jeweils Verordnungen zu den einzelnen Modulen, die Durchführung der Statistik.

Die Verordnungen können auch auf der [Homepage der Statistik Austria](#) aufgerufen werden.

3. Statistische Konzepte, Methodik

Gegenstand der Statistik

Gegenstand der Statistik ist das Einkommen und die Lebensbedingungen der Bewohnerinnen und Bewohner von Privathaushalten in Österreich. Bezugszeitraum der Einkommensvariablen ist (außer für das aktuelle Einkommen) das der Erhebung vorangegangene Kalenderjahr; alle anderen Variablen beziehen sich auf den Erhebungszeitpunkt.

Beobachtungs-/Erhebungs-/Darstellungseinheiten

Grundgesamtheit der Erhebung EU-SILC 2007 sind Personen in Privathaushalten in Österreich. Ausgenommen sind daher Personen in Anstaltshaushalten bzw. Gemeinschaftsunterkünften und Personen ohne festen Wohnsitz. Erhebungseinheiten sind private Haushalte und deren Bewohnerinnen und Bewohner, wenn in diesem Privathaushalt mindestens eine Person wohnt, die (1) ihren Hauptwohnsitz in dieser Wohnung hat und die (2) mindestens 16 Jahre alt ist.

Von der Stichprobenziehung der Ersterhebung ausgeschlossen wurden Haushalte, die bereits an der Erhebung EU-SILC 2003 bzw. EU-SILC 2004 - 2006 teilgenommen haben.

Datenquellen

Keine. EU-SILC ist eine Primärerhebung.

Meldeeinheit/Respondenten

Die Informationen des Haushaltsregisters sind durch die Erhebungspersonen einzutragen. Die Angaben für den Haushaltsfragebogen und das Personenregister werden bei einem Mitglied des Haushalts erfragt. Mittels Personenfragebögen werden alle Personen des Haushalts ab 16 Jahren befragt. Fremdauskünfte – so genannte Proxy-Interviews – sind nur in Ausnahmefällen gestattet. Es wurden etwa 20% der Personen (2.650 von insgesamt 13.391 Personen) durch Proxy-Interviews erfasst (vgl. das Kapitel zu Messfehlern), bei 59 Personen wurde das fehlende Personeninterview vollständig imputiert.

Erhebungsform

Stichprobenerhebung.

Charakteristika der Stichprobe

Grundgesamtheit und Auswahlrahmen

Die Grundgesamtheit besteht aus privaten Haushalten, das sind Wohnungen mit wenigstens einer mindestens 16-jährigen Person, die ihren Hauptwohnsitz in dieser Wohnung hat. Auswahlrahmen war das zentrale Melderegister (ZMR) mit Stand vom 31.12.2006. Die Haushalte der Stichprobe EU-SILC 2003 und EU-SILC 2004 -2006 wurden aus dem Auswahlrahmen für die Ersterhebung ausgeschlossen.

Samplingverfahren

Für die Ersterhebung wurden die 3.380 Adressen mittels Zufallsauswahl aus dem ZMR gezogen. Die Stichprobe wurde nach Interviewersprengeln geschichtet. Diese Vorgehensweise soll einerseits die regionale Streuung des Samples, andererseits eine ausgewogene Zuteilung von Adressen an die Erhebungspersonen gewährleisten.

Die gezogene Stichprobengröße leitet sich einerseits aus den durch EU-Verordnung geforderten Mindestgrößen der Stichprobe ab, andererseits aus dem zu erwartenden Panelausfall (Panelattrition) der nächsten Jahre, um zu gewährleisten, dass die Rotationsgruppen auch in den nächsten Jahren ausreichend groß sind. Die maßgebliche EU-Verordnung (VO Nr. 1177/2003 des Europäischen Parlamentes und des Rates) schreibt für Österreich vor, dass die Nettostichprobe im Querschnitt mindestens 4.500 Haushalte und im Längsschnitt mindestens 3.250 Haushalte enthalten muss.

Erhebungstechnik/Datenübermittlung

Wurde die Erhebung 2006 (für Details zur Feldarbeit siehe den Bericht über die [Feldarbeit](#) an das BMSK) noch ausschließlich von der Spectra Marktforschungs GesmbH durchgeführt, hat Statistik Austria beginnend mit EU-SILC 2007 einen Teil der Datenerhebung übernommen. Vorerst handelte es sich hierbei um die Ersterhebungshaushalte, sowie einen kleinen Teil der Folgehaushalte bei dem die Durchführung von CATI-Interviews in der Folgeerhebung getestet wurde. Inklusive Split-Haushalten wurden 780 Haushalte im Rahmen dieses CATI-Tests bearbeitet. Der Großteil der Folgehaushalte wurde wie bereits in den Vorjahren durch die Spectra Marktforschungs GesmbH mittels CAPI-Technik bearbeitet. Ab 2008 soll ein möglichst großen Anteil der Haushalte der Folgeerhebung telefonisch befragt werden. Langfristig ist mit dem Einstieg von Statistik Austria die Qualität der Daten wesentlich besser gewährleistet, es ist eine größere Transparenz der Erhebung gegeben und durch die Möglichkeit von kontrollierten Telefoninterviews sollen kostenmäßig Einsparungspotentiale geschaffen werden.

Die Feldarbeit dauerte von März (Beginn von Spectra mit der Folgeerhebung, bei Statistik Austria wurde im April mit der Ersterhebung und im Juni mit der Folgeerhebung begonnen) bis September 2007.

Dabei werden in persönlichen oder telefonischen Interviews die Antworten der Respondentinnen und Respondenten direkt in einen Laptop-PC eingegeben. Statistik Austria setzte dafür 138 Erhebungspersonen im Face to Face Bereich und 15 im Telefonstudio ein. Für Spectra führten 76 Erhebungspersonen Interviews durch.

Die Rohdaten wurden von den Interviewerinnen und Interviewern via Internet laufend übermittelt und aufbereitet. Sowohl von Spectra, als auch von der hauseigenen Erhebungsinfrastruktur wurden 3 Zwischendatenlieferungen an den Fachbereich übermittelt. Auf Grund technischer Probleme verzögerte sich die erste Datenlieferung von Statistik Austria.

Erhebungsbogen (inkl. Erläuterungen)

Nachdem EU-SILC eine CAPI/CATI-Erhebung ist, sind Erhebungsunterlagen nur insoweit vorhanden, als sie gemeinsam mit den Codebüchern und den vorgegebenen Plausibilitätschecks als Grundlage für CAPI/CATI-Programmierung dienen.

Teilnahme an der Erhebung

Freiwillig.

Erhebungs- und Darstellungsmerkmale, Maßzahlen; inkl. Definition

Erhebungsmerkmale auf Haushaltsebene: Einkommenskomponenten auf Haushaltsebene (brutto und netto), Fragen zur Wohnung, zur Ausstattung mit Konsumgütern und zur finanziellen Lage, nicht-monetäre Deprivationsindikatoren, Kinderbetreuung. 2007 wurden detaillierte Fragen zur Wohnsituation gestellt.

Erhebungsmerkmale auf Personenebene: Einkommenskomponenten auf Personenebene (brutto und netto), demographische Daten, Fragen zur Erwerbstätigkeit, zur Bildung, zur Gesundheit und zur Zufriedenheit.

Einkommenskomponenten in EU-SILC

Diese Komponenten werden sowohl brutto als auch netto erhoben. Das Nettoeinkommen ist das Bruttoeinkommen abzüglich der Beiträge zur gesetzlichen Sozialversicherung sowie der Einkommen- und Vermögenssteuern. Fallen bei einer Komponente keine derartige Abzüge an (so wie in Österreich z.B. bei der Familienbeihilfe), so ist das Bruttoeinkommen gleich dem Nettoeinkommen.

Referenzzeitraum für die Einkommenskomponenten ist in EU-SILC das dem Erhebungsjahr vorangegangene Kalenderjahr (2006).

Komponenten auf Haushaltsebene

Staatliche Beihilfen zu Wohnkosten

Familienbeihilfe

Sonstige Familienleistungen (ausgenommen Wochengeld/Kinderbetreuungsgeld)

Einkommen von Personen unter 16 Jahren

Komponenten auf Personenebene

Einkommen aus unselbständiger Erwerbstätigkeit (inkl. privat genutzter Firmen-Pkw und anteilige Abfertigungen)

Einkommen aus selbständiger Erwerbstätigkeit (inkl. Vermietung/Verpachtung)

Vermögenseinkommen

Einkommen aus Präsenz-/Zivildienst

Arbeitslosenleistungen

Pensionen und Unfallrenten

Pflegegeld

Krankengeld, Entgeltfortzahlung im Krankheitsfall, sonstige Krankheitsleistungen

Staatliche Beihilfen bei Ausbildung (z.B. Stipendien)

Sozialhilfe

Wochengeld, Kinderbetreuungsgeld (inkl. Karenzgeld), Unterhaltsvorschüsse

Sonstige staatliche Beihilfen

Unterhaltszahlungen (erhalten wie geleistet)

Sonstige private Transfers (erhalten wie geleistet)

Steuernachzahlungen bzw. -rückerstattungen

Sonstiges Einkommen

Die angeführten Komponenten wurden entweder auf Monats- oder auf Jahresbasis erfragt. Konnten oder wollten Respondenten oder Respondentinnen einen Betrag nicht nennen, so wurden sie gebeten, auf einer vorgelegten Karte eine Stufe (Einkommensintervall) zu nennen.

Einen Sonderfall stellt das Einkommen „Familienbeihilfe“ dar. Nachdem die Sätze für diese Leistung bekannt sind, wurde diese Einkommenskomponente betragsmäßig nicht erhoben, sondern berechnet.

Das Haushaltseinkommen entspricht der Summe aller dieser Einkommenskomponenten im jeweiligen Haushalt.

2007 wurden erstmalig imputierte Mieten, Zinsszahlungen für Hypotheken sowie Arbeitgeberbeiträge zur Sozialversicherung als zusätzliche Einkommenskomponenten berechnet. Imputierte Mieten wurden dabei auf der Basis des Mikrozensus für 67% aller Haushalte berechnet, um finanzielle Vorteile von Haus- und Wohnungseigentümern oder mietfrei wohnenden Personen zu quantifizieren und in Zukunft für das Haushaltseinkommen zu berücksichtigen. Durch die Berechnung von Zinsszahlungen für Hypotheken für imputierte Mieten konnten Kosten zur Wohnraumbeschaffung berücksichtigt und vom Haushaltseinkommen abgezogen werden. Die Arbeitgeberbeiträge zur Sozialversicherung wurden als Prozentsatz des Unselbständigeneinkommens berechnet, werden aber derzeit wie die imputierten Mieten noch nicht in die Berechnung des Haushaltseinkommens einbezogen.

Darstellungsmerkmal Äquivalenzeinkommen

Die meisten EU-Indikatoren, die im Rahmen von EU-SILC erhoben werden, bauen auf dem Äquivalenzeinkommen auf. Darunter zu verstehen ist ein auf einen Einpersonenhaushalt standardisiertes Haushaltseinkommen. Verwendet werden die Gewichtungsfaktoren der EU-Skala (modifizierte OECD-Skala):

Erste erwachsene Person	1,0
Jede weitere Person ab 14 Jahren	0,5
Kinder von 0 bis 13 Jahren	0,3

Das Äquivalenzeinkommen wird berechnet, indem man das Haushaltseinkommen durch die Summe der Gewichte im Haushalt dividiert. Es fällt daher in Einpersonenhaushalten gleich hoch, ansonsten immer niedriger aus als das Haushaltseinkommen. Mit Hilfe dieser Größe sollen unterschiedlich große und verschieden strukturierte Haushalte vergleichbar gemacht werden.

Definition von Armutsgefährdung (Einkommensarmut)

Personen gelten, entsprechend der von EU/Eurostat verwendeten Definition, als armutsgefährdet, wenn sie in einem Haushalt leben, dessen äquivalisiertes Einkommen unterhalb von 60% des Medians des Netto-Haushaltsäquivalenzeinkommens liegt.

Es sei hier noch einmal darauf verwiesen, dass mit „Einkommen“, sofern nicht anders angeführt, stets das Jahreseinkommen in dem Erhebungsjahr vorangegangenen Kalenderjahr gemeint ist, also bei der Erhebung 2007 das Einkommen im Jahr 2006.

Definition der Armutsgefährdungslücke

Darunter ist der Median der prozentuellen Abweichung des Einkommens armutsgefährdeter Haushalte von der Armutsgefährdungsschwelle zu verstehen.

Verwendete Klassifikationen

Die Adressen im Haushaltsregister waren gemäß EU-Vorgaben regional nach der Besiedlungsdichte³ sowie nach [NUTS 2](#) (in Österreich identisch mit den Bundesländern) zu klassifizieren.

In der Personenbefragung wurden, gemäß den Selbsteinschätzungen der Respondenten und Respondentinnen, die berufliche Tätigkeit nach der [Berufssystematik ISCO-88](#) (2-Steller) und

³ Hohe Besiedlungsdichte: eine Gruppe aneinander grenzender Gemeinden mit jeweils mehr als 500 Einwohner pro km² und insgesamt mindestens 50.000 Einwohner.

Mittlere Besiedlungsdichte: eine Gruppe aneinander grenzender Gemeinden mit 101-500 Einwohner pro km² und insgesamt mindestens 50.000 Einwohner.

Geringe Besiedlungsdichte: alle übrigen Gemeinden.

der Wirtschaftszweig der Betriebsstätte bzw. Dienststelle nach der [NACE](#) (2-Steller) vercodet. Die Bildungsebene wurde analog zum Mikrozensus-Merkmal „höchste abgeschlossene Schulbildung“ erhoben und anschließend nach der [ISCED](#) (1-Steller) vercodet. Die berufliche Funktion war erhebungsseitig ebenfalls an ein Mikrozensus-Merkmal angelehnt, nämlich an die „Stellung im Beruf“.

Regionale Gliederung der Ergebnisse

Die auf Basis von EU-SILC 2007 berechneten Indikatoren werden nicht regional gegliedert. In der Publikation wird eine Differenzierung nach Gemeindegrößenklassen⁴ und nach Bundesländern verwendet.⁵ Dabei ist auf eine größere Schwankungsbreite der Ergebnisse auf Bundesländerebene aufgrund der kleineren Fallzahlen hinzuweisen. Ergebnisse auf Bundesländerebene weisen daher eine geringere Präzision auf.

4. Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen

Datenerfassung

CAPI (Computer Assisted Personal Interviewing)

CATI (Computer Assisted Telephone Interviewing)

Signierung (Codierung)

Keine.

Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen

Es zählt zu den großen Vorteilen einer CAPI/CATI-Erhebung, dass Erfassungsfehler (z.B. bei Betragsangaben eine Null zu viel oder zu wenig) sowie technisch unplausible Antwortmuster (z.B. wechselseitig nicht übereinstimmende Angaben zur Verwandtschaft) zum Teil bereits während der Befragung korrigiert werden können. Dies ist möglich durch die elektronische Datenerfassung und die integrierte Durchführung von Haushalts- und Personeninterviews. Mehr als 150 derartige Checks wurden von Statistik Austria entwickelt und von den Erhebungsinstituten implementiert.

Zu diesen Plausibilitätschecks im CAPI/CATI-Programm gehören Errors und Warnings⁶. Die beiden Begriffe bezeichnen zwei Arten von Fehler. Errors treten auf, wenn die gemachten Angaben realiter nicht möglich ist (etwa wenn mehr einkommensbeziehende Kinder im Haushalt angegeben werden, als es Kinder im Haushalt gibt) – hier ist eine Eingabe der fehlerhaften Angabe nicht möglich. Warnings treten auf, wenn die gemachte Angabe zwar nicht gänzlich unmöglich ist, aber doch recht unwahrscheinlich ist (etwa der Bezug von Arbeitslosengeld von 3000 Euro im Monat) – hier ist eine Eingabe möglich, die Interviewerin bzw. der Interviewer wird aber darauf hingewiesen, dass die Richtigkeit der Angabe eher unwahrscheinlich ist.

⁴ Dabei werden folgende Ausprägungen unterschieden: Wien, Gemeinden mit über 100.000 Einwohner, Gemeinden mit über 10.000 Einwohner, sonstige Gemeinden (d.s. alle Gemeinden mit maximal 10.000 Einwohner).

⁵ Vergleiche hierzu unten den Abschnitt ‚Stichprobenbedingte Effekte, Repräsentativität‘ sowie die Publikation *Einkommen, Armut und Lebensbedingungen. Ergebnisse aus EU-SILC 2007*, Statistik Austria 2008 (http://www.statistik.at/wcm/idc/idcplg?ldcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=035744) und die darin angegebenen Hinweise zur Schwankungsbreite.

⁶ So heißen diese beiden Typen von Fehlern im von Spectra verwendeten CAPI-Programm Compass. In anderen CAPI-Programmen haben diese Fehlertypen andere Bezeichnungen. Bei Blaise entsprechen diese Fehlertypen den Signals (Hard) Checks und (Active) Signals.

Die Mitarbeiterinnen und Mitarbeiter von Spectra führten ihrerseits Datenchecks durch, um die Umwandlung der Daten vom Erhebungsformat ins ASCII-Format und die Qualität der Interviews zu prüfen. Auch in der hauseigenen Erhebungsabteilung wurden erste Datenchecks durchgeführt. Der Fachbereich gab nach eingehender Plausibilisierung der Daten regelmäßig Rückmeldungen.

In der Vorplaus wurde zunächst die Vollständigkeit der übermittelten Datenfiles kontrolliert. Anschließend wurden die verschiedenen Interviews in einem Haushalt auf ihre Konsistenz geprüft und weitere technische Prüfungen vorgenommen.

Für die Mikroplaus wurden eigene SPSS-Programme geschrieben, wobei sich zeigte, dass diese inhaltlichen Datenprüfungen sehr aufwändig waren und sind. Insbesondere die Bearbeitung der häufig auftretenden unplausiblen oder fehlenden Antworten auf Einkommensfragen erwies sich als sehr zeitaufwändig. Fragwürdigen Antworten können vom Respondenten bzw. der Respondentin selbst kommen oder durch Fehleingaben der Erhebungspersonen entstehen. Diese werden mittels Nachrecherchen oder Plausibilisierung korrigiert oder als fehlende Werte definiert. 2007 wurden sowohl Querschnittschecks als auch Längsschnittchecks, die auf einem Vergleich der aktuellen mit den Vorjahresdaten (und falls vorhanden auch älteren Daten von 2004-2006) basieren, durchgeführt.

Die Makroplaus erfolgte in erster Linie über Häufigkeitsauszählungen und Kohärenzprüfungen mit externen Datenquellen. Auch hierbei erwies sich das Vorhandensein der Daten aus den vorangegangenen EU-SILC Erhebungen und aus Steuerstatistiken als sehr hilfreich.

Nach dem Vorliegen des authentischen internen Datenbestandes konnten die Zielvariablen gebildet werden. Um diese zu überprüfen hatte Eurostat eigene SAS-Programme entwickelt, die auch den Mitgliedstaaten zur Verfügung gestellt worden waren.

Imputation (bei Antwortausfällen bzw. unvollständigen Datenbeständen)

Der folgende Abschnitt beschreibt die in EU-SILC 2007 angewandten Verfahren der Imputation. Diese basieren auf den bereits in den Vorjahren entwickelten Verfahren.

Verfahren der Imputation meint alle Verfahren, um fehlende Werte zu schätzen und zu ersetzen. Im Allgemeinen beziehen sich Imputationen auf Schätzverfahren, um entweder fehlende Information von ganzen Erhebungseinheiten (unit non-response) oder einzelne Werte von Erhebungseinheiten (item non-response) zu ermitteln.

Das Hauptinteresse der Erhebung EU-SILC liegt in der Erfassung des Haushaltseinkommens der befragten Haushalte. Bei fehlenden Werten auch nur einer Variablen bei nur einem Haushaltsmitglied kann kein korrektes Haushaltseinkommen berechnet werden. Deshalb werden in EU-SILC fehlende Einkommenswerte imputiert. Dasselbe gilt für fehlende Personeninterviews: fehlt ein Personeninterview eines Haushalts, so müssen die Angaben dieser Person geschätzt werden, um für diesen Haushalt das Einkommen berechnen zu können.

Verfahren für den Umgang mit unit non-response

In EU-SILC werden fehlende Interviews von Personen, die aufgrund von Verweigerungen, vorübergehender Abwesenheit, oder aus anderen Gründen nicht interviewt werden konnten, ersetzt. Dabei wird mit einer Distanzfunktion ein ähnlicher Fall gesucht, dessen Interviewergebnis auf den Fall mit den fehlenden Werten übertragen wird (Hot-Deck verfahren). Die Distanzfunktion verwendet zur Bestimmung der Ähnlichkeit bestimmter Variablen. Die Fälle werden nach ihrer Ähnlichkeit sortiert, und der nächstliegende Fall wird als Spenderwert verwendet, unter der Voraussetzung, dass folgende Anforderungen erfüllt werden:

- Der Spenderfall und der Fall mit fehlenden Werten haben das gleiche Geschlecht.
- Das Interview des Spenderfalls ist kein Proxy-Interview.
- Der Spenderfall hat denselben Beschäftigungsstatus wie der Fall mit den fehlenden Werten.

Die Imputation von Fällen, von denen es bereits ein Interview aus dem Vorjahr gibt, und von Fällen, die 2007 zum ersten Mal interviewt worden sind, unterscheidet sich: Während für letztere nur die Information aus dem aktuellen Jahr zur Bestimmung von ähnlichen Fällen verwendet werden, kann für erstere auch die Vorjahresinformation zur Bestimmung eines Spenderfalles herangezogen werden. Die Distanzfunktion für Folgebefragungsinterviews verwendet nun das Vorjahresinterview zur Bestimmung eines Spenderfalls, und die Information aus dem Interview des Jahres 2007 des Spenderfalls wird für den Fall mit fehlender Information herangezogen. Hierfür wurden die folgenden Variablen verwendet:

- Geschlecht
- Alter
- Gegenwärtige Beschäftigungssituation
- Haushaltsgröße
- Anzahl der Kinder unter 18 Jahre im Haushalt
- Anzahl der Personen über 60 Jahre im Haushalt
- Bundesland
- Höchste abgeschlossene Schulausbildung
- Chronische Krankheit/Beeinträchtigung bei alltäglichen Arbeiten durch Behinderung und sonstige gesundheitliche Beeinträchtigung
- Haushaltseinkommen
- Anzahl der Monate mit bezahlter Beschäftigung
- Anzahl der Monate mit selbständiger Beschäftigung

War nur Information aus dem aktuellen Erhebungsjahr verfügbar (Erstbefragung oder kein Interview im Vorjahr), konnte nur die Haushalts- bzw. Personenregisterinformation zur Berechnung der Distanzfunktion herangezogen werden. Die Liste der verwendeten Variablen ist somit kürzer, und beinhaltet die folgenden Variablen:

- Bevölkerungsdichte
- Geschlecht
- Alter
- Haushaltsgröße
- Beschäftigungsstatus
- Bundesland
- Anzahl der Kinder unter 18 Jahre im Haushalt
- Anzahl der Personen über 60 Jahre im Haushalt
- Haushaltseinkommen

Insgesamt wurden 59 Personeninterviews imputiert; im Vorjahr mussten 36 Personeninterviews imputiert werden. In 10 Fällen stand dabei Information aus dem Vorjahr zur Verfügung, für 49 Personen konnte nur auf Informationen aus dem aktuellen Jahr zurückgegriffen werden.

Verfahren für den Umgang mit item non-response

Im Fragebogen von EU-SILC finden sich zahlreiche detaillierte Fragen zum Einkommen der Befragten. Dabei kommt es immer wieder zu „fehlenden Werten“; sei es, dass die Befragten die Antwort verweigern, sei es, dass es zu falschen Eintragungen von Seiten der Interviewerinnen und Interviewer kommt, oder sei es, dass bestimmte Werte aufgrund von Plausibilitätsprüfungen gelöscht werden müssen. Das Problem ist nun, dass bei fehlenden Werten auch nur einer

Variablen bei nur einem Haushaltsmitglied kein korrektes Haushaltseinkommen berechnet werden kann. Deshalb werden in EU-SILC fehlende Einkommensvariablen imputiert, wobei nur Netto-Einkommensvariablen imputiert werden. Fehlende Bruttowerte werden mittels Netto-Brutto-Konversion berechnet.

Fehlende Werte in den Einkommensvariablen von EU-SILC können auf drei verschiedene Arten entstehen. Entweder es fehlt die Angabe darüber, ob die befragte Person eine bestimmte Einkommenskomponente bezogen hat oder nicht, oder es fehlt die Angabe darüber, wie oft eine bestimmte Einkommenskomponente bezogen wurde, oder es fehlt die Angabe darüber, wie hoch der Betrag einer bestimmten Einkommenskomponente ist.

Fehlt die Angabe darüber, ob eine bestimmte Einkommenskomponente bezogen wurde, so wird von Seiten der Statistik Austria versucht, diese Angabe aus anderen Variablen abzuleiten. Dabei wird insbesondere auf den Aktivitätenkalender zurückgegriffen. Kann die Information, ob eine Einkommenskomponente bezogen worden ist oder nicht, nicht abgeleitet werden, wird davon ausgegangen, dass diese Einkommenskomponente nicht bezogen wurde.

Eine ähnliche Behandlung finden fehlende Angaben zu der Anzahl der Monate, die eine Einkommenskomponente bezogen wurde. Zunächst wird versucht, die Angabe über die Monatsanzahl aus anderen Variablen des Datensatzes herzuleiten. Gelingt dies nicht, wird nach dem Muster der empirischen Verteilung ein Zufallswert imputiert.

Beim Auftreten eines fehlenden Betrags einer Einkommenskomponente ist die Vorgehensweise ein wenig komplexer. Grundsätzlich haben die Befragten mehrere Möglichkeiten Angaben zur Höhe einer Einkommenskomponente zu machen: entweder der/die Befragte nennt sowohl den Brutto-Betrag als auch den Netto-Betrag der Einkommenskomponente, oder es wird die Angabe nur zu entweder dem Brutto- oder dem Nettobetrag gemacht, oder der/die Befragte gibt eine Einkommensstufe an. Die Möglichkeit der Auswahl einer Einkommensstufe – anstelle der Angabe eines konkreten Betrages – war in der Befragung vorgesehen worden, um die Befragten in ihrer Erinnerung der Einkommenshöhe zu unterstützen, oder, im Falle einer drohenden Verweigerung, zumindest die ungefähre Einkommenshöhe erfragen zu können.

Ist bei Einkommensvariablen entweder der Bruttobetrag oder der Nettobetrag im Datensatz vorhanden, so wird der korrespondierende Wert durch die Brutto-Netto- bzw. die Netto-Brutto-Konversion errechnet. Diese Konversion erfolgt im Falle von unselbständigen Einkommen und Pensionseinkommen auf der Grundlage der Steuerwerte der Lohnsteuerdaten, im Falle von Selbständigeneinkommen auf der Basis der empirischen Werte des Datensatzes. Geben die Befragten anstelle eines exakten Wertes eine Einkommenskategorie an, so wird ein Zufallswert innerhalb dieser Kategorie berechnet.

Fehlt jedwede Angabe zur Höhe der bezogenen Einkommenskomponente, so wird zunächst versucht, die Höhe des Einkommens entweder aus anderen Variablen des Datensatzes oder aus gesetzlichen Regelungen zu ermitteln. Fehlt etwa die Angabe zur genauen Höhe des bezogenen Wochengeldes, so kann unter bestimmten Voraussetzungen der Wochengeldbezug aufgrund der gesetzlichen Vorgaben geschätzt werden.

Kann auch vermittels dieser Methoden kein Wert ermittelt werden, so stehen zwei Ansätze zur Berechnung eines Schätzwertes zur Verfügung: Längsschnitts- und Querschnittsverfahren. Das Längsschnittsverfahren kann dann verwendet werden, wenn die Person mit fehlenden Werten bei einer Variablen im Vorjahr einen Wert für diese Variable angegeben hat. Für alle anderen Fälle können nur Querschnittsverfahren verwendet werden.

Das Verfahren zur Längsschnittimputation basiert auf der *row-and-column*-Methode von Little und Su. Dieses Verfahren verwendet Reihen und Spalteneffekte, um einen geeigneten Spenderwert zu identifizieren. Der Reiheneffekt quantifiziert die Veränderung der Variable zwischen den Wellen, während der Spalteneffekt das Verhältnis eines Falles zu jeweils allen anderen Fällen bestimmt. Beide Effekte gemeinsam führen zu einem Gesamteffektwert, mit dem der Datensatz sortiert werden kann. Der nächstliegende Nachbarwert wird dann als Spenderwert verwendet.

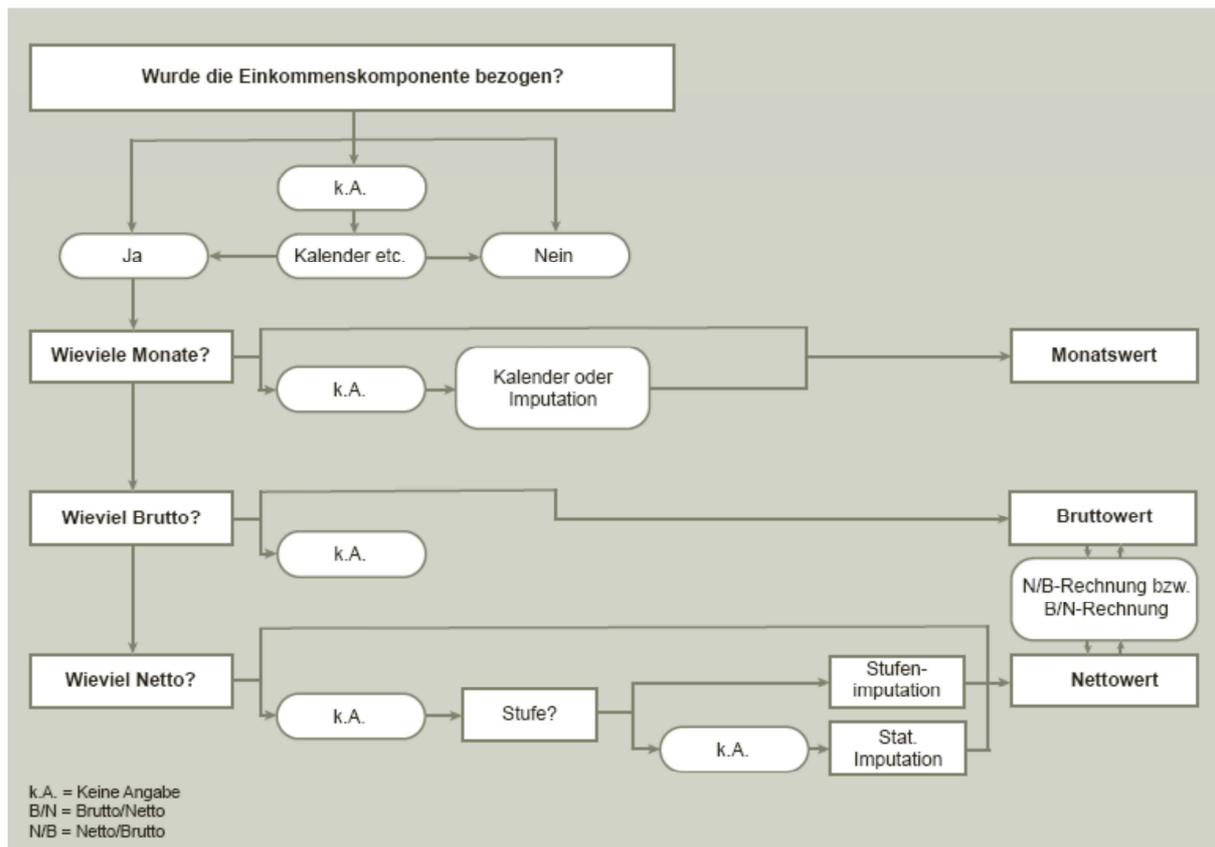
Als Querschnittsverfahren werden in EU-SILC 2007 Regressionsmodelle zur Berechnung von Schätzwerten verwendet. Für jede Einkommenskomponente wurden dabei mehrere Modelle spezifiziert, um sicherzustellen, dass auch beim Vorliegen von fehlenden Werten in den Prädiktorvariablen Schätzwerte berechnet werden können.

Würden man den aus der linearen Regression ermittelten Schätzwert imputieren, würde sich die Varianz der Variablen reduzieren. Um diesen Effekt möglichst gering zu halten, wurde den mittels linearer Regression ermittelten Werten ein stochastischer Störterm beigefügt, d.h. der ermittelte erwartete Wert wurde mit einem fiktiven Residuum addiert. Dieser Störterm entsprach in seiner Verteilung der Verteilung der geschätzten Residuen der erwarteten Werte. Dies führt dazu, dass die bei reinen, d.h. ohne Störterm durchgeführten Imputationen auf der Basis linearer Regression zu erwartende Reduktion der Varianz verringert werden konnte.

Die Prädiktorvariablen wurden nach ihrer Vorhersagekraft ausgewählt, das heißt, technisch ausgedrückt, nach der Variation des R^2 der Regressionsgleichung und nach der inhaltlichen Plausibilität der Prädiktoren. Konnte für eine Variable kein angemessenes Regressionsmodell spezifiziert werden, so wurde ein Schätzwert auf der Basis des Mittelwerts bzw. des Medians berechnet, welcher wiederum mit einem stochastischen Störterm addiert wurde.

Die folgende Grafik fasst die Vorgehensweise des Umgangs mit fehlender Information bei Einkommensvariablen zusammen.

Abbildung 2: Vorgehen bei fehlender Einkommensinformation in EU-SILC



Untenstehende Tabelle zeigt am Beispiel einiger Variablen die Verteilung der verschiedenen Arten der Ersetzung fehlender Werte an. Die ausgewählten Variablen sind die nach der Fallzahl größten Einkommenskomponenten der Einkommenskategorien unselbständiges und selbständiges Einkommen, Pensions-, Arbeitslosen- und Krankenleistungen. Die Anteile der Ersetzungsarten variiert dabei zwischen den verschiedenen Einkommenskomponenten: Stammen bei unselbständigem Einkommen und beim Arbeitslosengeld mehr als 90% der Angaben direkt aus der Befragung, so sind es bei Alterspensionen 89%, bei Einkommen aus Land- und Forstwirtschaft 83% und beim Krankengeld nur etwas mehr als 77%. Wie bereits in den bisherigen Erhebungen müssen nur wenige Werte imputiert werden. Bei CATI-Interviews ist insbesondere bei Einkommenskomponenten ein etwas höherer Antwortausfall festzustellen.

Tabelle 1: Imputationen ausgewählter Nettoeinkommenskomponenten

	Gesamt	Betrag laut Befragung	Betrag aus Stufe	Brutto-Netto- Rechnung	Längsschnitt- Imputation	Querschnitt- Imputation	Sonstige Korrekturen
<i>Absolut</i>							
Unselbständigen Einkommen	6.980	6.770	0	48	51	64	47
Land-/forstwirtschaftlicher Betrieb	392	324	30	0	24	14	0
Arbeitslosengeld	652	614	17	0	8	12	1
Alterspension	3.393	3.028	95	29	144	50	47
Krankengeld	131	101	16	3	1	7	3
<i>in %</i>							
Unselbständigen Einkommen	100,0	97,0	0,0	0,7	0,7	0,9	0,7
Land-/forstwirtschaftlicher Betrieb	100,0	82,7	7,7	0,0	6,1	3,6	0,0
Arbeitslosengeld	100,0	94,2	2,6	0,0	1,2	1,8	0,2
Alterspension	100,0	89,2	2,8	0,9	4,2	1,5	1,4
Krankengeld	100,0	77,1	12,2	2,3	0,8	5,3	2,3

Quelle: Statistik Austria, EU-SILC 2007

Am Beispiel der unselbständigen Einkommen zeigt folgende Tabelle, wie sich die Verteilung der Variable durch die Ersetzung der fehlenden Werte verändert, bzw. welche Werte durch die verschiedenen Ersetzungsarten geschätzt werden. Bei fast allen Ersetzungsarten ist der Median und das arithmetische Mittel höher als die durch die Befragung erfassten Werte. Aufgrund der geringen Anzahl der ersetzten Werte verändert dies aber den Median und den Mittelwert nicht bzw. nur geringfügig.

Tabelle 2: Verteilung des unselbständigen Einkommens nach Ersetzungsart der fehlenden Werte

	Betrag laut Befragung	Brutto-Netto- Rechnung	Betrag aus Plausibilitäten	Längsschnitts- imputation	Querschnitts- imputation	Sonstige Korrekturen
Anzahl	6.770	48	21	51	64	26
In %	97,0	0,7	0,3	0,7	0,9	0,4
Arithmetisches Mittel	17.909	19.600	5.882	13.742	18.007	23.961
Median	16.800	17.270	4.200	12.250	17.539	22.261
Minimum	40	302	270	401	2.437	2.300
Maximum	273.600	91.496	30.000	65.872	41.031	76.000
Standardabweichung	12.885	15.336	7.409	12.052	9.307	16.410

Quelle: Statistik Austria, EU-SILC 2007

Hochrechnung (Gewichtung)

Die Berechnung der Gewichte zur Hochrechnung der Stichprobenergebnisse auf die Gesamtbevölkerung basiert auf Empfehlungen von Eurostat (EU-SILC Doc. 65/05.1). Für die Berechnung werden wie bereits 2006 Gewichte getrennt für die 4 Teilstichproben, jeweils nach dem Jahr der Ersterhebung berechnet.

Ziel der Gewichtung ist die Beseitigung von bekannten Verzerrungen (Bias) in den Ergebnissen und die Schätzung von Populationswerten. Es gibt prinzipiell zwei Gewichtungsfaktoren:

- das sogenannte Basisgewicht, das im Falle eines Panels nur für jene Personen definiert ist, die bereits bei der ersten Befragung im Haushalt gelebt haben oder neu geboren wurden. Das Basisgewicht ist in der Regel nur bei der Erstbefragung für alle Personen in einem Haushalt identisch.
- Das sogenannte Querschnittsgewicht, das auch für jene Personen definiert ist die nach dem ersten Befragungsjahr in einen Stichprobenhaushalt eingezogen sind. Es ist für alle Personen im Haushalt identisch.

Die Berechnung der Gewichte erfolgt in drei Arbeitsschritten. Siehe dazu auch den für Eurostat erstellten [Qualitätsbericht](#) (in englischer Sprache).

Designgewicht

Das Designgewicht wird berechnet, um Effekte des Designs der Stichprobe bei der Gewichtung zu berücksichtigen, da das Stichprobendesign über die Auswahlwahrscheinlichkeit eines Elements der Grundgesamtheit entscheidet. Das Designgewicht wird als die Inverse der Auswahlwahrscheinlichkeit berechnet, wobei sich die Auswahlwahrscheinlichkeit aus der Division der Anzahl der Haushalte in der Stichprobe durch die Anzahl der Haushalte in der Grundgesamtheit ergibt. Im Falle einer reinen Zufallsauswahl, wie bei den Erstbefragungshaushalten 2004-2006⁷ in EU-SILC, hat jedes Element dasselbe Designgewicht innerhalb der jeweiligen Teilstichprobe. Bei der Erstbefragung 2007⁸ handelt es sich allerdings nicht wie in den Vorjahren um eine einfache, sondern um eine nach Interviewersprengeln geschichtete Zufallsauswahl. Die Auswahlwahrscheinlichkeit im geschichteten Stichprobendesign wird im Grunde wie bei einer einfachen Zufallsauswahl berechnet, allerdings für jeden Sprengel extra. Die zu erzielende Stichprobengröße pro Sprengel ergibt sich aus dem Anteil der Haushalte im Sprengel an der Gesamtbevölkerung. Deshalb unterscheidet sich das Designgewicht in der Erstbefragung 2007 je nach Interviewersprengel, ist aber innerhalb eines solchen Sprengels für alle Haushalte gleich.

Non-Response-Gewichtung

Ziel der Non-Response-Gewichtung ist die Reduktion des Bias, der durch den Antwortausfall von Haushalten entsteht (Verweigerungen etc.)⁹. Dabei wird zwischen den Erstbefragungshaushalten 2007 und den Folgebefragungshaushalten mit dem Erhebungsjahr 2004, 2005 und 2006 unterschieden. Die Korrektur dieser Verzerrung würde die Kenntnis der Antwortwahrscheinlichkeit für jeden Haushalt erfordern. Da diese in der Regel nicht bekannt ist, muss sie unter Verwendung bekannter Variablen geschätzt werden. Grundlage für die Schätzung der Antwortwahrscheinlichkeit sind die empirischen Antwortraten für verschiedene Gruppen von Haushalten.

Für die Erstbefragungshaushalte wurden wie im Vorjahr die Antwortquote nach Bundesland und Urbanisierungsgrad sowie folgende zusätzliche Informationen aus dem Auswahlrahmen (ZMR) berücksichtigt: die Anzahl Alter der Personen im Haushalt, sowie die Staatsbürgerschaft und das Geschlecht.¹⁰ Auf Basis dieser Informationen wurde ein logistisches Regressionsmodell entwickelt, um die Antwortwahrscheinlichkeit zu schätzen.

Für die Folgebefragungshaushalte steht im Vergleich zu den Erstbefragungshaushalten ungleich mehr Information zur Verfügung. Die Berechnung der Antwortwahrscheinlichkeit bezieht sich dabei allerdings auf die Individuen und nicht auf die Haushalte, da sich die Zusammensetzung des Haushalts zwischen den Wellen verändern kann. Personen, die auf Basis der natürlichen Bevölkerungsbewegung das Sample verlassen (Sterbefälle und Migration) reflektieren Veränderungen der Zusammensetzung der Grundgesamtheit und stellen damit kein Problem für die Gewichtung dar, da die Veränderung des Samples tatsächliche Veränderungen der Grundgesamtheit reflektieren. Problematisch sind Personen, die das Sample verlassen, aber nicht die Grundgesamtheit, also die Selektivität der Teilnahmebereitschaft. Um damit umzugehen, wird das Basisgewicht der das Panel verlassenden Personen auf ähnliche Personen aufgeteilt.

Zur Schätzung der Antwortwahrscheinlichkeiten wurden für die drei Teilsamples der Folgebefragung (Personen mit Ersterhebungsjahr 2004, 2005 oder 2006) drei logistische Regressionsmodelle entwickelt. Jeweils unterschiedliche Variablen erwiesen sich als relevant für den Ausfall.

⁷ Rotationen 1/05, 2/06 & 4/04, siehe Abbildung 1

⁸ Rotation 3/07, siehe Abbildung 1

⁹ Ein Bias aufgrund von Antwortausfall tritt dann auf, wenn bestimmte Personengruppen, z.B. junge Erwachsene, eine verhältnismäßig höhere Non-Response Rate haben. Dies führt dazu, dass bestimmte Gruppen unterrepräsentiert sind.

¹⁰ Diese Informationen über die Zusammensetzung der Haushalte stammen aus dem ZMR und müssen nicht mit der beim Interview erfassten Zusammensetzung der Personen im Haushalt identisch sein.

Für Personen, die seit den Erhebungen 2004, 2005 oder 2006 an EU-SILC teilnahmen, stehen aus den Vorjahren zahlreiche Variablen zur Berechnung der Antwortbereitschaft zur Verfügung. Das Modell zur Schätzung von Ausfallswahrscheinlichkeiten wird mechanisch nach einem Optimierungsalgorithmus ermittelt. Dabei wird die Vorhersagegenauigkeit für einen Ausfall maximiert und die Zahl der benötigten erklärenden Kategorien (Freiheitsgrade) minimiert. Die Modellierung folgt somit ausschließlich technischen Kriterien. Dies führt dazu, dass eine inhaltliche Interpretation der Variablen welche unterschiedliche Antwortwahrscheinlichkeiten erklären soll nur schwer möglich ist. Eine Beurteilung der substanzwissenschaftlichen Relevanz der erklärenden Variablen ist also lediglich für die Auswahl der in der Optimierung berücksichtigten Kategorien maßgeblich. Die Verwendung des Optimierungsalgorithmus soll lediglich die Große Zahl an erklärenden Variablen mittels eines automatisierten Prozesses reduzieren.

Beim Einfügen kategorialer Variablen in ein logistisches Regressionsmodell werden die einzelnen Kategorien in Dummyvariablen umgewandelt. Im Falle von bundesländerspezifischen Ausfällen erhalten beispielsweise eben nur jene Bundesländer spezifische Non-Response-Gewichte, die auch tatsächlich signifikante Unterschiede aufweisen.

Personen die im Laufe des drei- bzw. vierjährigen Panels¹¹ ein- oder zweimal nicht an der Erhebung teilgenommen haben, weisen für das Jahr 2005 und/oder 2006 kein Gewicht auf, weil ihre Gewichte im Zuge der Vorjahresgewichtung an andere Haushalte verteilt wurde. Demnach fehlt für die Anpassung ein entsprechendes Basisgewicht. Eurostat sieht hierzu vor, dass das Gewicht des Haushalts, in den diese Personen einziehen, aufgeteilt werden soll. Allerdings wurden die meisten diesen in die Erhebung zurückkehrenden Personen in Haushalten interviewt, in denen sie entweder alleine oder mit Personen, die bislang nicht an der Erhebung teilgenommen haben, leben. Somit gibt es für die meisten dieser Personen auch kein Haushaltsgewicht, das aufgeteilt werden könnte. Deshalb wurde im Modell für Personen, die 2004 mit der Erhebung begonnen haben, die Antwortwahrscheinlichkeit direkt zwischen den Jahren 2004 und 2007, für Personen die 2005 begonnen haben, direkt zwischen 2005 und 2007, geschätzt.

Gemäß den Vorgaben von Eurostat wurden die Gewichte getrimmt, um die Varianz der Gewichte zu reduzieren bzw. um extreme Gewichte zu vermeiden.

Kombination der Teilstichproben

Eine Gesamtstichprobe von EU-SILC besteht nach dem von Eurostat empfohlenen, integrierten Quer- und Längsschnittdesign aus drei Teilstichproben mit Panels von unterschiedlicher Dauer sowie einer Stichprobe von erstmalig befragten Haushalten. Streng genommen ist nur die jeweils neue Erstbefragungsstichprobe repräsentativ für alle Teile der Bevölkerung. Insbesondere jene Personen die vor nach dem Stichtag der jeweiligen Erstbefragung in die Population zugewandert sind können in den Folgebefragungsstichproben nicht erfasst werden. Die Kombination der Teilstichproben erfordert daher eine gesonderte Gewichtung.

Bei der "IN-Population", d.h. jene Personen die 2005, 2006 oder 2007 in die Grundgesamtheit zugewandert sind, wurden entsprechend den Vorgaben von Eurostat die Querschnittsgewichte um einen fixen Faktor erhöht¹².

Vor der abschließenden Anpassungsgewichtung werden die Gewichte der Stichprobenpersonen in jedem Haushalt auf alle darin lebenden Personen aufgeteilt in dem die Summe der um Antwortverweigerungen bereinigten Basisgewichte durch die Gesamtzahl der Haushaltsmitglieder dividiert wird. So erhalten auch Nicht-Stichprobenpersonen, deren Basisgewicht Null ist, einen Anteil an den Gewichten der anderen Personen im Haushalt, ohne dass sich die Summe der Gewichte des Haushaltes ändert. Mitunter kann dieser Anpassungsschritt eine deutliche Erhöhung der Streuung der Gewichte bewirken. Etwa dann, wenn eine einzelne Stichprobenperson in einen großen Haushalt einzieht.

¹¹ Rotation 1/05 bzw. Rotation 4/04

¹² Für Rotation 1/07 beträgt der IN-Populationsfaktor 4, für Rotation 1/05 beträgt er 4/3 und für die Rotation 2/06 beträgt er 4/2. Für die Rotation 4/04 gibt es keine In-Population.

Anpassungsgewichtung

Das Anpassungsgewicht wird berechnet, damit die Schätzwerte auf Basis der erhobenen Daten so weit wie möglich dem „wahren Wert“ der Grundgesamtheit entsprechen. Dazu werden Gewichtsummen für bestimmte Personengruppen an externe Randverteilungen angepasst. Externe Datenquelle für EU-SILC 2006 ist der Jahresdurchschnitt des Mikrozensus 2007, für die Anzahl der Bezieher und Bezieherinnen von Arbeitslosenleistungen wurden Administrativdaten verwendet.

Auf Haushaltsebene wird an die Haushaltsgröße (1, 2, 3 bzw. 4 und mehr Personen), das Rechtsverhältnis an der Wohnung/dem Haus (Miete/Eigentum), und das Bundesland angepasst. Auf Personenebene wird an die Merkmale Geschlecht und Alter, an die Anzahl der Personen von außerhalb des EU/EFTA-Raums und die Anzahl der Bezieherinnen und Bezieher von Leistungen aus der Arbeitslosenversicherung angepasst. Diese Anpassung wurde wie bereits im Vorjahr mittels des von INSEE entwickelten SAS-Makro CALMAR (Logit-Methode) simultan auf Haushalts- und Personenebene durchgeführt.

Erstellung des Datenkörpers, (weitere) verwendete Rechenmodelle, statistische Schätzmethoden

Die Berechnung von Familienbeihilfe/Kinderabsetzbetrag erfolgte entsprechend den gesetzlich festgelegten Sätzen. Auch das Kinderbetreuungsgeld wird aus den gesetzlichen Regelungen eingesetzt.

Ein weiteres Rechenmodell betraf die Zuordnung von Einmalzahlungen wie Abfertigungen oder Urlaubsentschädigungen. Die Antwort auf die Frage, inwieweit solche Zahlungen Teil des laufenden Einkommens sind, ist international umstritten.

Nachdem diese Frage auch von Eurostat nicht eindeutig geregelt wurde, wurden die Einmalzahlungen nach einer nationalen Regelung zugeordnet. Die Einmalzahlungen werden anteilig entweder als Unselbständigen-, Arbeitslosen- oder Pensionseinkommen klassifiziert, je nach aktuellem Erwerbsstatus der interviewten Person.

Neben den Einmalzahlungen waren zwei der österreichischen Erhebungsvariablen den Eurostat-Zielvariablen nicht eindeutig zuordenbar. Das Einkommen aus Präsenz- und Zivildienst wurde als Einkommen aus unselbständiger Erwerbstätigkeit klassifiziert. Das „sonstige Einkommen“ wurde je nach Plausibilität dem Unselbständigen-, Selbständigen- oder Pensionseinkommen zugerechnet.

Bei Pensionisten und Pensionistinnen war nach dem Monatseinkommen gefragt worden. Um die Nettowerte auf Jahresbasis bringen zu können, musste die Steuerbegünstigung von Sonderzahlungen rechnerisch inkludiert werden. Dies geschah mit der Berechnung eines entsprechenden Faktors auf Basis der Lohnsteuerdaten 2006.

Ein weiteres angewandtes Rechenmodell war die Brutto-Netto-Rechnung bzw. Netto-Brutto-Rechnung. Diese wurde notwendig, wenn eine Zielperson entweder nur das Brutto- oder – was häufiger der Fall war – nur das Nettoeinkommen angegeben hatte. Die Umrechnung erfolgte bei Unselbständigen und Pensionistinnen und Pensionisten auf Basis der Lohnsteuerdaten 2006, also auf Basis realer Faktoren. Für selbständig Erwerbstätige sind solche Informationen nicht vorhanden; hier wurde die Brutto-Netto- bzw. Netto-Brutto-Relation auf der Basis von Regressionsmodellen anhand der erhobenen Daten ermittelt.

Informationen zur Imputation von Antwortausfällen sind im Kapitel „Imputation bei Antwortausfällen“ zu finden.

Sonstige qualitätssichernde Maßnahmen

Um den Einfluss einer selektiven Adressenbearbeitung auf die Qualität der Ergebnisse einzuschränken, wurden auch 2007 folgende Bestimmungen vereinbart: Erstens sollten pro Haushalt zumindest drei Kontaktversuche erfolgen, bevor eine Verweigerung akzeptiert werden konnte. Zweitens sollten regionale Antwortquoten erfüllt werden: Bei der Ersterhebung sollten in jeder Region mindestens 50% der Haushalte erfolgreich befragt werden (insgesamt lag die geforderte Ausschöpfungsquote in der Ersterhebung bei 60%), bei der Folgerhebung sollten mindestens 75% aller ausgewählten Haushalte einer Region erfolgreich befragt werden (bundesweit 85%). Diese Ausschöpfungsquoten wurden in der Erstbefragung erreicht, in der Folgebefragung etwas unterschritten. Obwohl die Ausschöpfung der Erhebung 2007 deutlich verbessert werden konnte, wurde der ursprüngliche Zielwert von 93% in der Folgebefragung deutlich unterschritten.

Die Vereinbarung von zweiwöchigen Feldberichten und Zwischendatenlieferungen ermöglichte von Seiten der Statistik Austria die Kontrolle der Daten noch während der Feldarbeitszeit. Somit konnten eventuell auftretende Fehler und Fehlentwicklungen an das Feldinstitut bzw. an die Erhebungspersonen rückgemeldet werden und die Feldinstitute konnten notwendig gewordene Rückrufe bei den Interviewten machen.

Als inhaltliche Qualitätssicherungsmaßnahme seien schließlich die Zwischendatenlieferungen angeführt. Die Erhebungsinstitute erhielt zu jeder Datenlieferung umfassende Rückmeldung, sodass Informationen über die Qualität der Erhebung noch während der Feldarbeit bekannt waren. Dies führte in Einzelfällen auch zur Nachbefragung einzelner Haushalte, wodurch die Qualität der endgültigen Daten weiter gesteigert werden konnte.

Der SILC CATI Test

Im Rahmen der Erhebung EU-SILC 2007 wurde erstmals der Einsatz von CATI (Computer assisted Telephone Interview) getestet. 750 Haushalte (+ 30 Splithaushalte) wurden dafür ausgewählt, 365 der ersten und 415 Haushalte der zweiten Rotationsgruppe. Falls Haushalte nicht am Telefon Auskunft geben wollten, wurden sie von einer Erhebungsperson aufgesucht und wie bisher mittels CAPI Technik befragt. 541 der 780 Haushalte konnten von CATI Erhebungspersonen kontaktiert werden, 441 dieser Haushalte wurden erfolgreich mittels CATI und 120 HH mittels CAPI befragt.

Der CATI Test zeigt, dass die Bereitschaft, telefonisch Auskunft zu geben mit der Erwerbsintensität, dem Bildungsniveau und der beruflichen Stellung steigt. CATI Interviews weisen höhere Item non-response Raten auf, vor allem bei Einkommenskomponenten. Fehlende Werte resultieren meist aus Angaben von „weiß nicht“ oder „keine Angabe“. Durch die nachweisbar geringere Dauer der CATI Interviews tragen diese zur Entlastung der Respondenten und Respondentinnen bei.

Nach dem positiven Test Ergebnis soll der Einsatz der CATI Technik in der SILC Erhebung unter Voraussetzung ständiger Qualitätskontrollen in den nächsten Jahren ausgeweitet werden.

5. Publikation (Zugänglichkeit)

Vorläufige Ergebnisse

Keine.

Endgültige Ergebnisse

Siehe Standardpublikation und Internet.

Revisionen

Im Herbst 2009 wurden die Daten von EU-SILC 2007 revidiert. Die Datenrevision betrifft die Einkommen aus unselbständiger Erwerbstätigkeit. Dabei wurde erstens die Einbeziehung von Sachleistungen in die Berechnung des Haushaltseinkommens geändert und zweitens Probleme bei der Berechnung der Brutto-Netto-Konversion korrigiert.

Der Marktwert von Sachleistungen unselbständig Erwerbstätiger ist in EU-SILC als eigene Einkommenskomponente konzipiert. Seit Anlaufen von EU-SILC bis zur Erhebung 2007 war nur die private Nutzung von Firmen-PKWs in das Haushaltseinkommen aufzunehmen und verpflichtend an Eurostat zu liefern. Die übrigen Sachleistungen waren ab EU-SILC 2007 zu erheben und umfassen etwa freies Essen, Essensgutscheine, Brennstoffe, Strom oder Kinderbetreuung. Die Sachleistung von freiem oder vergünstigtem Wohnraum ist davon ausgeschlossen, da der entsprechende Wert konzeptionell von der seit 2007 berechneten imputierten Miete abgedeckt wird.

Ursprünglich war die Einberechnung der zusätzlichen Sachleistungen in das Haushaltseinkommen ab EU-SILC 2007 vorgesehen. Auf europäischer Ebene wurde jedoch beschlossen, erst die Auswirkung einer geänderten Einkommenszusammensetzung auf die mit EU-SILC befüllten Indikatoren zu analysieren und die Komponenten des Haushaltseinkommens wie in den Vorjahren zu belassen. Die Daten von EU-SILC 2007 wurden dementsprechend revidiert, sodass wie bislang im Haushaltseinkommen nur die private Nutzung von Firmen-PKWs als Sachbezug enthalten ist - der entsprechende Betrag ist im Unselbständigeneinkommen enthalten. Die Daten von EU-SILC 2008 sind ebenfalls so aufgebaut. Neu ist hingegen, dass der Wert der ab 2007 zusätzlich erhobenen Sachbezüge in den revidierten Daten von EU-SILC 2007 als eigene Komponente ausgewiesen wird.

Nach der Veröffentlichung der Daten entdeckte Probleme bei der Brutto-Netto-Rechnung führten zu überhöhten Bruttobeträgen in den Datensätzen von EU-SILC 2007. Diese Probleme wurden im Zuge der Datenrevision beseitigt.

Publiziert in:

Standardpublikation

[Einkommen, Armut und Lebensbedingungen. Ergebnisse aus EU-SILC 2007](#). Herausgegeben von Statistik Austria (inkl. CD-ROM); Wien 2008 (ISBN 978-3-902587-93-0)

Internet

Auf der Website der Statistik Austria sind im Bereich Statistiken, Soziales unter dem Punkt [Armut und soziale Eingliederung](#) die Ergebnisse von EU-SILC hinsichtlich Armutsgefährdung und soziale Eingliederung abrufbar. Unter dem Punkt Statistiken, Soziales, Haushaltseinkommen sind die Ergebnisse von EU-SILC bezüglich des [Haushaltseinkommens](#) abrufbar.

Eurostat

Tabellen mit den Daten für alle Mitgliedstaaten der EU stehen auf der Website von Eurostat kostenlos zur Verfügung:

Behandlung vertraulicher Daten

Um die von den Erhebungsinstituten gelieferten Daten auf Vollständigkeit und Richtigkeit zu prüfen, und um Haushalte in den kommenden Jahren erneut befragen zu können, sind die Familien- und Vorname der befragten Personen sowie das vollständige Geburtsdatum zu erheben. Weiters liegen die Adressen aus der Stichprobe vor. Diese Merkmale waren Teil des Registers in der Erhebungsphase, sind also in einem anderen Datensatz gespeichert als die inhaltlichen Erhebungsdaten wie Einkommen, Gesundheit usw. Nach Abschluss der Plausibilitätsprüfungen werden Name, Adresse und Geburtstag aus den Datensätzen gelöscht. Im Datensatz verbleiben lediglich Geburtsmonat, Geburtsjahr und das Bundesland (NUTS2), da es sich hierbei um Eurostat-Zielvariablen handelt.

Die Adressinformation bzw. Namen und Geburtsdatum der an der Studie teilnehmenden Personen und Haushalte sind, um eine Weiterverfolgung der Panelhaushalte/Personen zu ermöglichen, in einer Access-Datenbank eingelagert. Diese ist nur für die mit dem Projekt EU-SILC betrauten Personen zugänglich.

Im Eurostat-Datenfile selbst sind die Haushalte durch eine Laufnummer (5-stellig) gekennzeichnet. Die Identifikation einer Person erfolgt über eine 7-stellige Personennummer, wobei es sich bei den ersten 5 Stellen um die Laufnummer des Haushalts, bei den letzten 2 Stellen um eine Personen-Laufnummer innerhalb des Haushalts handelt.

Die Geheimhaltungsbestimmungen für Daten, die im Bundesstatistikgesetz 2003 konsolidierte Fassung §19 (2) und (3) geregelt sind, werden strikt eingehalten.

6. Qualität

6.1. Relevanz

Das Projekt EU-SILC wird seit dem Jahr 2004 auf der Basis der Verordnung 1177/2003 des Europäischen Parlaments und des Rates der Europäischen Union durchgeführt. Bis zum Jahr 2008 wird EU-SILC einerseits durch Eurostat und andererseits durch das Sozialministerium finanziert. Diese beiden Institutionen fungieren somit als Auftraggeber, mit denen während der Projektarbeit eine intensive Kommunikation stattfand. Durch die fristgerechte und einwandfreie Erfüllung der aus den Projektverträgen entstehenden Verpflichtungen (Datenlieferung, Berichtspflichten usw.) werden die Interessen der Auftraggeber erfüllt.

EU-SILC bildet eine wichtige Grundlage für die Europäische Sozialstatistik. Zentrale Themen sind Einkommen, Beschäftigung, Wohnen und viele andere Bereiche einschließlich subjektiver Fragen zu Gesundheit und finanzieller Lage. Diese Fragen erlauben es, die Lebenssituation von Menschen in Privathaushalten abzubilden. EU-SILC ist auch die zentrale Quelle zur Erhebung der vom Europäischen Rat verabschiedeten Laeken-Indikatoren zur Messung von Armut und sozialer Ausgrenzung. Diese Indikatoren ermöglichen den Mitgliedstaaten und der Europäischen Kommission, die Fortschritte bei der Erreichung des vom Europäischen Rat von Lissabon gesteckten Zieles zu messen, bis 2010 bei der Ausmerzung von Armut deutlich weiterzukommen, das Verständnis von Armut und sozialer Ausgrenzung im europäischen Rahmen zu verbessern sowie den Erfahrungsaustausch zwischen den Mitgliedsländern zu fördern.

EU-SILC ist auch die Basis für den nationalen Bericht zur sozialen Eingliederung und für den Frühjahrsbericht der Europäischen Kommission. Weiters ist EU-SILC die Datenquelle für mehrere Strukturindikatoren: die Armutsgefährdungsquote vor und nach sozialen Transfers, das Verteilungsquintil und die Quote der dauerhaften Armutsgefährdung.

Für wissenschaftlich und administrativ Interessierte wurde eine so genannte „[Nutzer- und Nutzerinnen-CD](#)“ erstellt, welche Erhebungsdaten im EU-Zielvariablen-Format enthält. Diese kann um 250 € pro Erhebungsjahr käuflich erworben werden.

Darüber hinaus stehen auf der Website der Statistik Austria Mikrodaten von EU-SILC 2004, 2005 und 2006 ([Standardisierte Datensätze SDS](#)) mit einer Zufallsstichprobe im halben Stichprobenumfang (2.260 Haushalte mit 5.751 Personen) zur Verfügung. In Kürze werden auch Mikrodaten von EU-SILC 2007 verfügbar sein. Diese können nach einer Registrierung und der Akzeptierung der Nutzungsbedingungen kostenfrei heruntergeladen werden.

6.2. Genauigkeit

6.2.1. Stichprobenbedingte Effekte, Repräsentativität

Das arithmetische Mittel des äquivalisierten Netto-Haushaltseinkommen beträgt in EU-SILC 2007 20.399 Euro, der Schätzwert für den Standardfehler beträgt 146 Euro.¹³ Mit diesem Wert lässt sich das Konfidenzintervall errechnen, und somit die Grenzen angeben, innerhalb derer der wahre Wert des arithmetischen Mittels mit 95%-Wahrscheinlichkeit liegt. 2007 beträgt die Obergrenze des Konfidenzintervalls 20.685 Euro und die Untergrenze 20.114 Euro.¹⁴

Die Berechnung der Konfidenzintervalle für die [Laeken-Indikatoren](#) wurde nach dem Linearisierungs¹⁵verfahren durchgeführt. Eine kurze Beschreibung dessen befindet sich in der Standard-Dokumentation EU-SILC 2006.

Der Designeffekt beziffert die Genauigkeit des Erhebungsdesigns im Vergleich zu einer einfachen Zufallsstichprobe. In der Verordnung (EG) 1177/2003 wird er definiert als Quotient der Stichprobenvarianz der Armutsgefährdungsquote der realisierten Stichprobe im Verhältnis zur Stichprobenvarianz der Armutsgefährdungsquote in der Modellstichprobe. Mit dem von Eurostat für die Berechnung des Stichprobenfehlers empfohlenen Varianzschätzungsverfahren wurde der Designeffekt in Österreich auf 1,33 geschätzt.

Tabelle 3: Armutsgefährdungsquote und deren Schwankungsbreite für Österreich und die Bundesländer

	Armutsgefährdungsquote		Konfidenzintervall 95%				befragte Haushalte	Personen in befragten Haushalten
			untere Grenze		obere Grenze			
	in %	in 1.000	in %	in 1.000	in %	in 1.000		
Österreich	12,0	989	11,2	917	12,9	1.060	6.806	16.684
Burgenland	13,7	37	9,0	24036	18,5	49.290	247	606
Kärnten	10,8	59	6,5	35557	15,0	81.667	475	1.156
Niederösterreich	10,4	167	8,6	137843	12,3	196.716	1.287	3.212
Oberösterreich	8,1	114	6,3	88265	9,9	140.253	1.233	3.207
Salzburg	10,1	55	7,1	38696	13,1	70.904	440	1.125
Steiermark	13,6	160	11,2	132031	16,0	188.265	1.014	2.547
Tirol	10,1	71	7,5	52184	12,8	89.535	584	1.493
Vorarlberg	13,0	50	9,4	35984	16,6	63.428	307	832
Wien	17,4	276	15,0	238439	19,8	314.202	1.219	2.506

Q: STATISTIK AUSTRIA, EU-SILC 2007

6.2.2. Nicht-stichprobenbedingte Effekte

Qualität der verwendeten Datenquellen

Nicht zutreffend, da eigene Datenerhebung.

Abdeckung (Fehlklassifikationen, Unter-/Übererfassung)

Die Stichprobe stammt aus dem Zentralen Melderegister (ZMR). Dieses war nach der letzten Volkszählung durch die Zusammenführung der Gemeinde-Melderegister erstmals initialisiert worden und wird seither, auf Basis von Wohnsitzmeldungen, laufend aktualisiert. Die Administration obliegt dabei dem Bundesministerium für Inneres, das ZMR umfasst Haupt- und Nebenwohnsitze.

¹³ Dieser Wert resultiert aus der Varianzschätzung mit dem Verfahren der Linearisierung; Bootstrapping-Verfahren wurden 2007 nicht angewandt.

¹⁴ Diese Werte beziehen sich auf die Berechnung nach dem Linearisierungsverfahren. Berechnet wird das Konfidenzintervall nach der Formel: $KI_{95} = 20.399 \pm 1,96 * 146$.

¹⁵ Eine kurze Beschreibung des Verfahrens findet sich in der Standard-Dokumentation EU-SILC 2006.

Leider war bei der Erstbefüllung des ZMR nicht auf den Eintrag eines eindeutigen Wohnungsidentifikators geachtet worden, sodass ein Wohnungszusammenhang nur über die Adresse herstellbar ist. Auch variierte die Qualität der Einträge in den einzelnen Gemeinden erheblich, weshalb in der ZMR-Aufbauphase umfassende Cleanings und Harmonisierungen notwendig wurden.

Die Herstellung eines Wohnungszusammenhangs (und damit eines Haushaltszusammenhangs) über die Adresse funktioniert nicht immer einwandfrei (fehlende Türnummern, unterschiedliche Schreibweisen usw.). Als Abdeckungsfehler sind daher das Nicht-Erkennen von Wohnungszusammenhängen und ein daraus resultierender verzerrter Auswahlrahmen zu nennen. Das wird dadurch belegt, dass die Zahl der Einpersonenhaushalte im ZMR höher ausfällt als z.B. im Mikrozensus. Weniger wahrscheinlich ist ein irrtümliches Zusammenfassen von Personen, die nicht in einer Wohnung leben.

Weiters können, bedingt durch die Konstruktion des ZMR auf Basis des Meldewesens, klassifikatorische Fehler auftreten, wenn die Hauptwohnsitzmeldung einer Person nicht dem tatsächlichen Mittelpunkt der Lebensbeziehungen entspricht und reale Hauptwohnsitze im ZMR nur als Nebenwohnsitze oder gar nicht aufscheinen.

Daneben muss der Zeitabstand zwischen der Stichprobenziehung und der Feldarbeit berücksichtigt werden: Nachdem das ZMR kontinuierlich aktualisiert wird, entspricht der Auswahlrahmen zum Zeitpunkt der Stichprobenziehung zumindest formal zwar tatsächlich der korrespondierenden Population. Da aber bis zur konkreten Datenerhebung mitunter einige Monate vergehen, ist auch hier mit geringfügigen Verzerrungen zu rechnen (Gebäudeabrisse, Neubauten, Umzüge usw.).

Gegenüber anderen Fehlern haben Abdeckungsfehler den Vorteil, dass sie während der Feldphase zumindest teilweise korrigiert werden können. So wurden etwa nicht (mehr) existente Adressen als qualitätsneutrale Ausfälle verbucht. Auch konnte der Haushaltszusammenhang der einzelnen Mitglieder eines Haushalts durch die Interviewerinnen und Interviewer in der Regel korrekt erfasst werden.

Antwortausfall (Unit-Non Response, Item-Non Response)

Unit non-response

Insgesamt wurden für die Erhebung EU-SILC 2007 8.922 Adressen verwendet. 118 dieser Adressen erwiesen sich als nicht existent. Von den verbleibenden 8.804 Adressen wurden 8.710 erfolgreich kontaktiert. Von diesen Adressen wurden schließlich 6.862 Haushalte erfolgreich interviewt, die restlichen 1.848 Haushalte verweigerten entweder die Auskunft, konnten nicht angetroffen werden, oder nahmen aus sonstigen Gründen nicht an der Erhebung teil. Allerdings stellten sich 56 dieser Haushaltsinterviews aufgrund von Qualitätsmängeln als nicht verwendbar heraus, sodass letztlich 6.806 Haushalte im Datensatz verblieben und für die Analyse zur Verfügung standen.

Tabelle 4: Stichprobe EU-SILC 2007

	Total		Erstbefragungsadressen		Folgebefragungsadressen	
	N	%	N	%	N	%
Verwendete Adressen	8.922	100,0	3.380	100,0	5.542	100,0
Adresse existiert	8.804	98,7	3.274	96,9	5.530	99,8
Adresse existiert nicht	118	1,3	106	3,2	12	0,2
Bruttosample	8.804	100,0	3.274	100,0	5.530	100,0
Adresse erreicht	8.710	98,9	3.263	99,7	5.447	98,5
Adresse nicht erreicht	94	1,1	11	0,3	83	1,5
Erreichte Adressen	8.710	100,0	3.263	100,0	5.447	100,0
Haushaltsinterview abgeschlossen	6.862	78,8	2.167	66,4	4.695	86,2
Niemand anwesend	286	3,3	145	4,4	141	2,6
Mitarbeit verwweigert	1.324	15,2	876	26,8	448	8,2
Keine auskunftsfähige Person	28	0,3	26	0,8	2	0,0
Sonstige Ausfallgründe	210	2,4	49	1,5	161	3,0
Erfolgreiche Haushaltsinterviews	6.862	100,0	2.167	100,0	4.695	100,0
Interview aufgenommen in die Datenbank	6.806	99,2	2.124	98,0	4.682	99,7
Interview abgelehnt	56	0,8	43	2,0	13	0,3

Quelle: Statistik Austria, EU-SILC 2007

In diesen 6.806 Haushalten lebten zum Zeitpunkt des Interviews 16.684 Personen. 3.293 dieser Personen waren unter 16 Jahre alt, die übrigen 13.391 Personen waren mindestens 16 Jahre alt oder älter. Von diesen 13.391 Personen stehen Personeninterviews zur Verfügung; 10.682 dieser Interviews wurden dabei durch ein persönliches Interview oder durch Proxy-Interviews erfasst, 642 davon durch telefonische Interviews. Die restlichen 59 Personeninterviews wurden imputiert. Von den 13.332 nicht imputierten Personeninterviews wurden 19,8% (2.650 Interviews) durch Proxy-Interviews erfasst. Bei der Erhebung EU-SILC 2005 betrug die Proxyrate 19,6%.

Tabelle 5: Haushalte, Personen und Personeninterviews in EU-SILC 2007 nach Bundesland

	befragte Haushalte	Insgesamt	Personen		Insgesamt	Personeninterviews			
			Personen 16 Jahre und älter	Personen unter 16 Jahre		CAPI*	CATI**	Proxy-Interview	vollständig imputiert
Österreich	6.806	16.684	13.391	3.293	13.391	10.040	642	2.650	59
Burgenland	247	606	520	86	520	370	30	120	0
Kärnten	475	1.156	933	223	933	700	44	182	7
Niederösterreich	1.287	3.212	2.623	589	2.623	2.010	127	474	12
Oberösterreich	1.233	3.207	2.538	669	2.538	1.932	133	472	1
Salzburg	440	1.125	889	236	889	672	32	183	2
Steiermark	1.014	2.547	2.060	487	2.060	1.558	95	390	17
Tirol	584	1.493	1.173	320	1.173	769	42	357	5
Vorarlberg	307	832	617	215	617	410	28	176	3
Wien	1.219	2.506	2.038	468	2.038	1.619	111	296	12

Quelle: Statistik Austria, EU-SILC 2007

* Computerunterstütztes persönliches Interview

** Computerunterstütztes telefonisches Interview

Item non-response

Von Relevanz ist Item non-response vor allem bei den Einkommensvariablen von EU-SILC. Die Auswertungen der Statistik erfolgt zumeist auf Basis der Einkommenszielvariablen. Diese werden nicht direkt durch die Befragung erfasst, sondern setzen sich aus mehreren Einzelfragen zusammen. Deshalb wird in der folgenden Tabelle, die die Imputationen bei den Einkommenszielvariablen darstellt, zwischen vollständiger, partiell fehlender und vollständig fehlender Information unterschieden.

Tabelle 6: Imputationen der Einkommenszielvariablen

	Haushalte/Personen		Vollständige Information		Partiell fehlende Information		Vollständig fehlende Informationen	
	N	%	N	%	N	%	N	%
Gesamtes Bruttohaushaltseinkommen	6.806	100,00	0	0,0	6.806	100,0	0	0,0
Gesamtes verfügbares Haushaltseinkommen	6.806	100,00	4.688	68,9	2.076	30,5	42	0,6
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>								
Imputierte Mieten	4789	70,36	0	0,0	4.789	100,0	0	0,0
Einkommen aus Vermietung und Verpachtung	280	4,11	272	97,1	0	0,0	8	2,9
Familienleistungen/Kindergeld	2.437	35,81	2.425	99,5	12	0,5	0	0,0
Sonst. Leistungen gegen soziale Ausgrenzung	185	2,72	180	97,3	1	0,5	4	2,2
Wohnungsbeihilfen	267	3,92	256	95,9	6	2,2	5	1,9
Erhaltene Transfers zwischen Privathaushalten	518	7,61	495	95,6	3	0,6	20	3,9
Zinsen, Dividenden	4.547	66,81	3.347	73,6	223	4,9	977	21,5
Zinsen für Hypotheken	1.792	26,33	0	0,0	1.792	100,0	0	0,0
Einkommen von Personen unter 16	65	0,96	58	89,2	0	0,0	7	10,8
Geleistete Transfers zwischen privaten Haushalten	468	6,88	436	93,2	10	2,1	22	4,7
Einkommenssteuernachzahlungen/-erstattungen	2.964	43,55	2.899	97,8	26	0,9	39	1,3
<i>Personenebene</i>								
Unselbständige Erwerbstätigkeit	7.012	52,4	6.300	89,8	443	6,3	269	3,8
Beiträge an private Rentensysteme	3.036	22,7	2.822	93,0	1	0,0	213	7,0
Selbständige Tätigkeit	1.297	9,7	1.158	89,3	19	1,5	120	9,3
Eigenproduktion	411	3,1	376	91,5	0	0,0	35	8,5
Renten aus privaten Systemen	41	0,3	39	95,1	0	0,0	2	4,9
Arbeitslosenleistungen	901	6,7	836	92,8	42	4,7	23	2,6
Altersleistungen	3.447	25,7	3.110	90,2	203	5,9	134	3,9
Hinterbliebenenleistungen	122	0,9	112	91,8	1	0,8	9	7,4
Krankenleistungen	241	1,8	219	90,9	4	1,7	18	7,5
Invaliditätsleistungen	377	2,8	365	96,8	6	1,6	6	1,6
Ausbildungsleistungen	179	1,3	167	93,3	3	1,7	9	5,0

Quelle: Statistik Austria, EU-SILC 2007

Messfehler (Erfassungsfehler)

Als Messfehler wird die Differenz zwischen dem Wert einer Variablen und dem wahren – aber unbekanntem – Wert der Variablen bezeichnet. Es lassen sich in einer Erhebung vier Quellen für solche Messfehler identifizieren:

- der Fragebogen (Effekte durch das Design, den Inhalt oder die Wortwahl),
- die Datenerhebungsmethode,
- die befragende Person (Effekte des Interviewers, der Interviewerin auf die zu Befragenden und Erfassungsfehler durch den bzw. die Interviewende),

der/die Befragte (Effekt der Respondenten und Respondentinnen auf die Interpretation der Fragen).

Das Auftreten von Messfehlern ist (zumindest in sozialwissenschaftlichen Erhebungen) nahezu unvermeidbar. Maßnahmen, wie Datenkontrollen während der Feldarbeit und die CAPI/CATI-Checks helfen diese Fehler zu reduzieren. Weiters werden jährlich im Zuge von Fehleranalysen Verbesserungen vor allem des Fragebogens, der Erhebungsunterlagen (Listen und Karten) und der CAPI/CATI-Programmierung durchgeführt.

Wichtiges Instrument der Fehleranalyse ist die Nachbefragung der Erhebungspersonen (Debriefing). Die Ergebnisse des Debriefings sowie andere Rückmeldungen des Erhebungsinstituts zum Erhebungsinstrument sind in die CAPI/CATI-Gestaltung der Erhebung EU-SILC 2007 eingeflossen.

Wesentlichste Maßnahme zur Vermeidung von Effekten der Befragenden sind die Schulungen, in denen sichergestellt wird, dass alle Befragten unter möglichst denselben Bedingungen interviewt werden. An diesen Schulungen nahmen auch Mitarbeiterinnen und Mitarbeiter der Statistik Austria teil.

Befragteneffekte wurden von den Erhebungsinstituten bzw. von Statistik Austria vorrangig bei der korrekten Terminologisierung öffentlicher Leistungen festgestellt. Befragte verwechselten beispielsweise die korrekten Bezeichnungen bei manchen Familien- oder Pensionsleistungen. Einige dieser Fehler wurden identifiziert und konnten mittels telefonischer Rückfragen korrigiert werden.

Die Vereinbarung von zweiwöchigen Feldberichten und Zwischendatenlieferungen ermöglichte von Seiten der Statistik Austria die Kontrolle der Daten noch während der Feldarbeitszeit. Somit konnten eventuell auftretende Fehler an die Erhebungspersonen rückgemeldet und von den Feldinstituten bei den Haushalten recherchiert werden.

Proxy-Interviews sind in EU-SILC in Ausnahmefällen gestattet. Der Anteil blieb wie bereits im Vorjahr auch 2007 unter 20%.

Aufarbeitungsfehler

Die Interviewerinnen und Interviewer übermittelten die durchgeführten Interviews über das Internet an das jeweilige Feldinstitut. Dort wurden die Daten aufbereitet und in das passende Datenformat transferiert. Nach der Übermittlung der Daten an den Fachbereich wurden die Daten in das SPSS-Datenformat gebracht. Hier fanden die mit dem Programm SPSS durchgeführten Plausibilitätsprüfungen, die Datenaufbereitung und Fehlerkorrektur, die Imputationen statt, bevor die Daten in das Eurostat-Zielvariablenformat transferiert wurden. Diese Zielvariablen werden, nach Vorliegen des authentischen Datenbestandes, ins EXCEL-CSV-Format gebracht und an Eurostat übermittelt.

Durch die Teilung der Feldarbeit in zwei Erhebungseinheiten und die Trennung in CAPI- und CATI-Erhebungen waren im Vergleich zu den Vorjahren jedoch massive Mehrarbeiten notwendig – es musste die komplette Infrastruktur sowie der CAPI- und CATI-Fragebogen neu aufgebaut und die Betreuung und Kontrolle von zwei Erhebungseinheiten gewährleistet werden. Das Feldinstitut Spectra hat langjährige Erfahrung mit CAPI-Studien. Somit gab es wenige Probleme mit der CAPI-Programmierung und der Erfassung einzelner Fragen. Bei den Daten der Statistik Austria wurde die erste Zwischendatenlieferung hingegen mehrmals verschoben, da die technische Umsetzung des Datenbankabzugs nicht wie geplant klappte.

Modellbedingte Effekte

Zum Einfluss der Querschnittimputationen (von item non-response bei Einkommensfragen) auf die Genauigkeit der Ergebnisse ist anzuführen, dass bei Imputation mittels linearer Regression die Varianz des interessierenden Merkmals unterschätzt wird. Dies folgt daraus, dass bei einer Regression im Grunde genommen Mittelwerte imputiert werden (verallgemeinerte Form der *group mean imputation*), d.h. es wird zwar die Streuung zwischen verschiedenen Gruppen, nicht jedoch die Streuung innerhalb einer Gruppe berücksichtigt. Dieser Fehler fällt umso geringer aus, je größer der Anteil der Gesamtstreuung ist, der auf die Streuung zwischen den Gruppen entfällt.

Um die Varianz des interessierenden Merkmals (des Einkommens) dennoch valide abbilden zu können, wurde bei der Imputation ein stochastischer Störterm inkludiert, d.h. die Prognosewerte aus dem Regressionsmodell wurden um ein fiktives Residuum erweitert, das in seiner Verteilung der Verteilung der geschätzten Residuen bei den beobachteten Werten entsprach.

6.3. Rechtzeitigkeit und Aktualität

Die Daten wurden am 31. Juli 2008 in einer ersten Version via eDamis an Eurostat übermittelt. Eine revidierte Version der Daten wurde am 23. September 2008 übermittelt.

Der standardisierte [Qualitätsbericht](#) wurde am 10. Dezember 2008 an Eurostat übermittelt.

6.4. Vergleichbarkeit

2007 wurde EU-SILC zum fünften Mal in Österreich durchgeführt, und zum vierten Mal als integrierte Längsschnitterhebung. Das heißt, die Haushalte der Erhebung wurden zum vierten, dritten, zweiten oder ersten Mal für EU-SILC befragt.

Gegenüber dem Vorjahr wurden wenige Änderungen der Fragebögen bei der Erfassung von Merkmalen vorgenommen. Etwaige Änderungen beziehen sich vorrangig auf das Routing des Fragebogens bzw. auf integrierte Kontrollen des Fragebogens.

6.5. Kohärenz

Kohärenz meint die Validierung der Daten von EU-SILC mit externen Datenquellen. Als Qualitätskriterium meint Kohärenz die Vergleichbarkeit der Ergebnisse der Erhebung mit den Ergebnissen anderer Datenquellen. Im Folgenden werden die Daten von EU-SILC 2007 mit externen Datenquellen verglichen. Einkommensbezugsjahr ist in EU-SILC das der Erhebung vorangegangene Jahr, d.h. im Jahr 2007 das Jahr 2006.

Folgende Datenquellen werden für den Vergleich mit den Daten von EU-SILC 2007 herangezogen: die Vorjahreserhebung EU-SILC 2006, die Lohnsteuerdaten 2006 und die Schätzungen auf der Basis der Sektorkonten der Volkswirtschaftlichen Gesamtrechnungen (VGR) des Jahres 2006.

Vergleich mit EU-SILC 2006

EU-SILC 2006 war die dritte Erhebung des Frageprogramms EU-SILC im integrierten Quer- und Längsschnittdesign, was bedeutet, dass 2006 zum zweiten Mal Haushalte zum zweiten Mal befragt wurden.¹⁶ Datensätze der Erhebung EU-SILC eignen sich aufgrund der gleichen Konzepte und Definitionen in den beiden verglichenen Datensätzen im besonderen Maße für einen Vergleich, allerdings reicht dies nicht für eine Überprüfung der Kohärenz aus. Fehler, die sich eben aus den Konzepten und Definition ergäben, blieben unerkannt.

Die quantitativ bedeutendsten Variablen verändern sich im Vergleich von EU-SILC 2006 und EU-SILC 2007 nicht dramatisch. Zwar gibt es etwa Veränderungen beim verfügbaren Haushaltseinkommen (der Median steigt um 2%, das Volumen um 6%), die auf diesen Werten beruhenden Indikatoren wie die Armutsgefährdungsquote, das arithmetische Mittel der äquivalisierten Einkommen und die Armutsgefährdungsschwelle bleiben aber zwischen EU-SILC 2006 und EU-SILC 2007 relativ konstant. Im Vergleich zum Vorjahr sinkt die Armutsgefährdungsquote leicht von 12,6% auf 12,0%, das mittlere äquivalisierte Haushaltseinkommen steigt von 19.674 auf 20.399 Euro und die Armutsgefährdungsschwelle steigt ebenfalls von 10.711 Euro auf 10.945 Euro. Die Veränderungen in Vergleich zum Vorjahr liegen innerhalb der zufälligen Schwankungsbreite.

¹⁶ Für eine genauere Beschreibung der Erhebung EU-SILC 2006 vgl. „Standard-Dokumentation Metainformation zu EU-SILC 2006“

Tabelle 7: Vergleich der Einkommenszielvariablen von EU-SILC 2006 und EU-SILC 2007

	Median			Haushalte/Personen			Summe (in Millionen Euro)		
	EU-SILC 2006	EU-SILC 2007	Veränderung %	EU-SILC 2006	EU-SILC 2007	Veränderung %	EU-SILC 2006	EU-SILC 2007	Veränderung %
Gesamtes Bruttohaushaltseinkommen	35.210	35.583	1,1	3.508.442	3.537.022	0,8	146.357	154.452	5,5
Gesamtes verfügbares Haushaltseinkommen	27.371	27.971	2,2	3.508.442	3.537.022	0,8	110.635	115.653	4,5
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>									
Imputierte Mieten	-	2.015	-	-	3.537.022	-	-	8.650	-
Einkommen aus Vermietung oder Verpachtung	3.960	3.600	-9,1	123.239	137.859	11,9	1.185	1.161	-2,0
Familienleistungen/Kindergeld	4.171	4.080	-2,2	1.082.567	1.118.834	3,4	5.091	5.166	1,5
Sonst. Leistungen gegen Soziale Ausgrenzung	1.456	1.200	-17,6	78.423	106.477	35,8	241	284	17,6
Wohnungsbeihilfen	1.356	1.296	-4,4	134.686	150.109	11,5	198	218	10,3
Erhaltene Transfers zwischen privaten Haushalten	3.240	3.360	3,7	232.620	255.824	10,0	1.094	1.201	9,8
Zinsen, Dividenden	67	120	78,5	2.589.627	2.286.276	-11,7	875	1.296	48,2
Zinsen für Hypotheken	-	910	-	-	871.897	-	-	1.325	-
Einkommen von Personen unter 16 Geleistete Transfers zwischen privaten Haushalten	1.400	1.700	21,4	28.309	27.259	-3,7	47	89	88,5
Einkommenssteuernachzahlungen/-erstattungen	3.000	3.000	0,0	239.306	247.530	3,4	897	938	4,6
	-260	-300	15,4	1.398.084	1.461.988	4,6	-334	-453	35,8
<i>Netto-Einkommenskomponenten auf Personenebene</i>									
Unselbständige Erwerbstätigkeit	16.603	16.510	-0,6	3.590.363	3.589.351	0,0	61.797	63.226	2,3
Beiträge an private Rentensysteme	840	800	-4,8	1.511.467	1.489.789	-1,4	1.647	1.597	-3,1
Selbständige Tätigkeit	10.800	10.800	0,0	598.153	659.999	10,3	8.659	10.087	16,5
Einkommen aus Eigenproduktion	150	200	33,3	131.744	191.260	45,2	31	119	277,3
Renten aus privaten Systemen	1.800	3.360	86,7	16.950	25.912	52,9	62	226	263,7
Arbeitslosenleistungen	3.500	3.360	-4,0	636.837	614.841	-3,5	2.874	2.879	0,2
Altersleistungen	14.026	14.296	1,9	1.657.060	1.706.017	3,0	25.494	27.273	7,0
Hinterbliebenenleistungen	8.120	6.370	-21,6	54.858	57.344	4,5	465	382	-17,8
Krankenleistungen	1.493	1.440	-3,6	125.305	139.777	11,5	347	405	16,8
Invalitätsleistungen	12.600	11.200	-11,1	199.876	186.143	-6,9	2.508	2.172	-13,4
Ausbildungsleistungen	1.800	2.040	13,3	104.969	104.096	-0,8	323	298	-7,7
Bruttomonatsverdienste von Arbeitnehmern	1.700	1.800	5,9	3.252.714	3.232.045	-0,6	6.242	6.425	2,9

Quelle: Statistik Austria, EU-SILC 2006 und EU-SILC 2007

Für jene Einkommenskomponenten mit dem größten Einfluss auf das Haushaltseinkommen, Erwerbseinkommen und Pensionsleistungen¹⁷, kann von 2006 auf 2007 keine Veränderung festgestellt werden. Das Volumen dieser Variablen erhöht sich dennoch um 2% bzw. 7%.

Deutliche Veränderungen wurden in Bezug auf Privatpensionen und die Höhe der Ersparnisse durch Eigenproduktion festgestellt. Beide Einkommenskomponenten werden nicht für die Berechnung des Haushaltseinkommens berücksichtigt. Hier sollen vertiefende Analysen Aufschluss über die Bedeutung der veränderten Werte in EU-SILC 2007 geben.

Zu den lediglich geringen Veränderungen tragen einerseits die annähernd gleichbleibenden Definitionen und Fragen in EU-SILC 2006 und 2007 bei, andererseits auch die gestiegene Ausschöpfung insbesondere der Folgebefragungen.

Der Vergleich der Querschnittsergebnisse aus EU-SILC 2006 und EU-SILC 2007 ist wichtig zur Beurteilung der Kohärenz der Ergebnisse aus EU-SILC 2007; reale Veränderungen der Einkommen lassen sich klarer betrachten, wenn die Einkommen derjenigen Haushalte und Personen¹⁸ miteinander verglichen werden, die in beiden Jahren an der Erhebung teilgenommen haben. Dieser Vergleich ist in der folgenden Tabelle dargestellt. Dabei wird nur der Median der Einkommenskomponenten dargestellt – die Anzahl der Bezieher und Bezieherinnen und das Volumen dieser Einkommenskomponenten ist nicht sinnvoll vergleichbar.¹⁹

Hier zeigt sich, dass die Höhe des Medians bei der überwiegenden Zahl der Einkommenskomponenten relativ konstant bleibt. Ein Anstieg kann etwa bei den Einkommen aus der Arbeitslosenversicherung beobachtet werden. Insgesamt unterscheidet sich die Entwicklung des Einkommens nicht zwischen dem Vergleich des Gesamtquerschnitts und der Entwicklung der Panelhaushalte bzw. -personen (Ausnahmen sind hier kleinere Einkommenskomponenten wie das Einkommen von unter 16-Jährigen und Ausbildungsleistungen).

¹⁷ Diese beiden Einkommenskomponenten machen mehr als $\frac{3}{4}$ des Haushaltseinkommens aus.

¹⁸ Bei den Haushalten werden nur jene Haushalte für den Vergleich ausgewählt, die sich in ihrer Zusammensetzung nicht verändern, bei den Personen nur diejenigen Personen, die in beiden Jahren mit Personeninterviews erfasst wurden.

¹⁹ Die gewichtete Anzahl der Haushalte und Personen hängt aufgrund der Hochrechnung auf die Gesamtbevölkerung vor allem vom Verhältnis der Größe der in diesem Vergleich nicht berücksichtigten Personen und Haushalte ab. Damit ist auch das Einkommensvolumen dieser Haushalte nicht vergleichbar.

Tabelle 8: Vergleich der Einkommenszielvariablen von EU-SILC 2006 und EU-SILC 2007 (Haushalte/Personen in beiden Erhebungen; Jahresbeträge)

	Median		Veränderung in %
	EU-SILC 2006	EU-SILC 2007	
Gesamtes Bruttohaushaltseinkommen	35.510	36.278	2,2
Gesamtes verfügbares Einkommen	27.676	28.346	2,4
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>			
Imputierte Mieten	-	3.442	-
Einkommen aus Vermietung und Verpachtung	3.960	3.840	-3,0
Familienleistungen/Kindergeld	4.298	4.135	-3,8
Sonst. Leistungen gegen Soziale Ausgrenzung	1.806	2.000	10,7
Wohnungsbeihilfen	1.440	1.440	0,0
Erhaltene Transfers zwischen privaten Haushalten	3.240	3.240	0,0
Zinsen, Dividenden	67	120	78,7
Zinsen für Hypotheken	-	900	-
Einkommen von Personen unter 16 Jahren	1.408	1.657	17,6
Geleistete Transfers zwischen privaten Haushalten	3.360	3.000	-10,7
Einkommenssteuernachzahlungen/-erstattungen	-270	-280	3,7
<i>Netto-Einkommenskomponenten auf Personenebene</i>			
Unselbständige Erw erbstätigkeit	16.603	16.800	1,2
Beiträge an private Rentensysteme	800	800	0,0
Selbständige Tätigkeit	10.564	10.800	2,2
Einkommen aus Eigenproduktion	150	191	27,4
Renten aus privaten Systemen	1.800	2.208	22,7
Arbeitslosenleistungen	3.600	3.600	0,0
Altersleistungen	14.026	14.201	1,2
Hinterbliebenenleistungen	7.560	7.000	-7,4
Krankenleistungen	1.700	1.300	-23,5
Invaliditätsleistungen	12.600	11.760	-6,7
Ausbildungsleistungen	2.040	2.163	6,0
Bruttomonatsverdienste von Arbeitnehmern	1.700	1800	5,9

Quelle: Statistik Austria, EU_SILC 2006 und EU-SILC 2007

Zusammenfassend kann festgestellt werden, dass die Zahlen von EU-SILC 2007 in hohem Maße mit den Daten von EU-SILC 2006 in Einklang zu bringen sind. Die bestehenden Abweichungen sind größtenteils Zufallsschwankungen aufgrund geringer Fallzahlen.

Vergleich mit den Lohnsteuerdaten 2006

Die Lohnsteuerdaten erfassen die Einkommen von unselbständig Beschäftigten und Pensionistinnen und Pensionisten. Sie stellen damit eine wichtige Quelle für Vergleiche der wichtigsten Einkommenskomponente auf Personenebene, dem Einkommen aus unselbständiger Arbeit, dar.²⁰ Allerdings gibt es auch konzeptionelle Unterschiede bei der Erfassung unselbständiger Einkommen bei EU-SILC und den Lohnsteuerdaten, die sich aus den unterschiedlichen Erfassungsbereichen der beiden Datenquellen erklären: In EU-SILC werden keine Personen außerhalb von Privathaushalten erfasst, ebenso wenig wie Personen, die zwischen der Steuerperiode und der Erhebung entweder verstorben oder ins Ausland verzogen sind.

Die Lohnsteuerdaten erfassen etwa 3,70 Millionen unselbständig Beschäftigte, in EU-SILC 2007 sind dies 3,59 Millionen, also etwa 11.600 Beschäftigte weniger. Gründe für diese Differenz liegen in der unterschiedlichen Erfassung unselbständiger Beschäftigter zwischen EU-SILC und der Lohnsteuerdaten und der möglichen Untererfassung von kurzzeitiger Beschäftigung in EU-SILC.

²⁰ Vergleiche der Pensionseinkommen sind aus konzeptionellen Gründen schwieriger und werden aus diesem Grund hier nicht durchgeführt.

Der Vergleich der beiden Statistiken zeigt eine recht gute Deckung der Einkommensverteilungen insbesondere um den Median; Abweichungen zwischen den Lohnsteuerdaten und EU-SILC sind an den Rändern der Verteilungen größer. Das bedeutet, EU-SILC unterschätzt sehr hohe und überschätzt sehr geringe Einkommen. Dies ist mutmaßlich auf Erinnerungsprobleme und inkorrekte Proxy-Angaben in EU-SILC zurückzuführen, aber möglicherweise auch auf Scheinbeschäftigungen in den Lohnsteuerdaten.

Beschäftigungen von nur kurzer Dauer werden, so kann vermutet werden, in einer retrospektiven Befragung schlechter bzw. gar nicht erinnert. Daraus geschlossen werden kann, dass sich der Unterschied zwischen den Lohnsteuerdaten und EU-SILC reduziert, wenn kurzfristige Beschäftigungen (definiert als Beschäftigungen von maximal einem Monat im Jahr) nicht in den Lohnsteuerdaten berücksichtigt werden. Die Lohnsteuerdaten für 2006 ohne Berücksichtigung dieser kurzzeitigen Beschäftigungen sind in untenstehender Tabelle in den mittleren drei Spalten dargestellt. Es zeigt sich eine höhere Übereinstimmung zwischen EU-SILC und diesen eingeschränkten Lohnsteuerdaten. Die Einkommensverteilungen für Männer, Frauen und insgesamt eben so wie die Anzahl der Beschäftigten stimmen hier besser überein als im Vergleich zu den uneingeschränkten Lohnsteuerdaten.

Tabelle 9: Vergleich der Brutto-Jahreseinkommen unselbständig Beschäftigter 2006: Lohnsteuerdaten 2006 und EU-SILC 2007

Perzentile	Lohnsteuerdaten			Bereinigte Lohnsteuerdaten 1)			EU-SILC 2007		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	2.644	3.969	1.960	4.122	6.321	2.930	4.500	7.034	3.492
20% ...	7.903	12.549	5.191	9.566	15.067	6.532	9.650	14.206	7.200
25% ...	10.786	16.735	7.509	12.234	18.548	8.637	12.150	17.775	8.687
30% ...	13.465	19.932	9.594	14.766	21.215	10.573	14.303	19.667	10.383
40% ...	18.342	24.354	13.318	19.368	25.187	14.067	18.200	23.400	13.869
50% ...	22.833	28.100	16.715	23.572	28.767	17.332	22.376	26.600	16.800
60% ...	27.071	32.108	20.405	27.703	32.756	20.981	25.800	30.600	19.600
70% ...	31.821	37.348	24.602	32.414	38.043	25.130	30.400	35.000	23.100
75% ...	34.855	40.815	27.165	35.457	41.530	27.671	32.900	37.800	25.270
80% ...	38.639	45.261	30.184	39.276	45.994	30.661	36.000	42.000	28.396
90% ...	51.100	59.805	39.577	51.767	60.664	40.080	46.816	53.800	36.400
arith. Mittel	26.500	32.478	19.572	27.451	33.645	20.273	25.235	30.275	19.090
Personen	3.704.931	1.988.772	1.716.159	3.571.243	1.917.041	1.654.202	3.589.351	1.972.192	1.617.159

Quelle: Statistik Austria, EU-SILC 2007; Lohnsteuerdaten 2006

1) ohne jene, die max 1 Monat (= 31 Tage) beschäftigt waren.

Es zeigt sich also, dass die Kohärenz zwischen Lohnsteuerdaten und EU-SILC steigt, wenn kurze Beschäftigungen bei den Lohnsteuerdaten nicht berücksichtigt werden. Dies entspricht auch der Vermutung, dass Beschäftigungen und Einkommen in Befragungen umso besser von den Respondentinnen und Respondenten reproduziert werden, je regelmäßiger die jeweilige Beschäftigung war. In der folgenden Tabelle werden ebenfalls die Verteilungen von EU-SILC und Lohnsteuerdaten gegenüber gestellt, allerdings beschränkt auf jene Personen, die im Jahr 2006 mindestens elf Monate beschäftigt waren (hier definiert als ganzjährige Beschäftigung). Auch hier verbessert sich die Kohärenz der Einkommensverteilungen der Lohnsteuerdaten und EU-SILC; insbesondere im unteren Bereich der Verteilung. Die Anzahl der Beschäftigten ist jedoch höher als in den Lohnsteuerdaten.

Tabelle 10: Vergleich der Brutto-Jahreseinkommen unselbständig Beschäftigter 2005: Lohnsteuerdaten 2005 und EU-SILC 2006 (nur ganzjährig Beschäftigte)

	Lohnsteuerdaten			EU-SILC 2007		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	10.368	18.902	6.966	10.800	17.775	7.560
20% ...	16.233	23.793	11.651	16.100	21.600	11.764
25% ...	18.599	25.450	13.398	18.000	22.800	13.300
30% ...	20.829	26.977	15.001	19.600	24.252	14.700
40% ...	24.460	29.928	17.898	22.620	27.385	17.500
50% ...	27.926	33.266	21.001	25.854	30.400	19.600
60% ...	31.651	37.387	24.312	29.400	33.700	22.471
70% ...	36.478	42.954	28.362	33.600	38.549	26.441
75% ...	39.636	46.632	30.802	36.400	42.000	28.985
80% ...	43.642	51.164	33.821	39.600	45.344	31.500
90% ...	56.392	66.562	43.400	50.400	56.994	39.200
arith. Mittel	32.467	40.020	23.922	29.574	35.176	22.464
Personen	2.746.280	1.457.767	1.288.513	2.723.367	1.523.362	1.200.005

Quelle: Statistik Austria, EU-SILC 2007; Lohnsteuerdaten 2006

Vergleich mit den Volkswirtschaftlichen Gesamtrechnungen

Für den Vergleich der Haushaltseinkommen von EU-SILC mit den Sektorkonten der Volkswirtschaftlichen Gesamtrechnungen (VGR) werden die VGR auf die Konten private Haushalte und private Organisationen ohne Erwerbszweck eingeschränkt. Der Ausgangswert der VGR muss dabei angepasst werden. Allerdings gibt es noch weitere Unterschiede zwischen den VGR und EU-SILC, die sich nicht ohne weiteres quantifizieren lassen und die im Folgenden angeführt werden:

- Sachleistungen und einmalige Zahlungen sind in der VGR enthalten aber in EU-SILC 2007 nicht vollständig erfasst.
- Einkommen aus selbständiger Arbeit ist in der VGR eine residuale Kategorie.
- Bei der Berechnung der verfügbaren Einkommen in der VGR wird ein Schätzwert für Spenden für karitative Organisationen und für Mitgliedsbeiträge abgezogen.
- In der VGR sind transnationale Zahlungen mitberücksichtigt.
- Zinszahlungen betragen nach der VGR im Jahr 2006 3,91 Mrd. Euro (D5). Diese Zahlungen (vor allem Hypotheken u.ä.) sind nicht in den Haushaltseinkommen von EU-SILC berücksichtigt.

Wie bereits in den letzten Jahren sind die aggregierten Haushaltseinkommen in EU-SILC deutlich niedriger als die in den VGR geschätzten Werte. Der Unterschied der Schätzwerte verringert sich entscheidend, wenn Vermögenseinkünfte nicht berücksichtigt werden, was die Annahme stützt, dass Vermögenseinkünfte in EU-SILC untererfasst sind.

Tabelle 11: Vergleich des Haushaltseinkommens und des verfügbaren Einkommens VGR 2006 und EU-SILC 2007 (in Mio. Euro)

	Brutto-Haushaltseinkommen		Verfügbares Einkommen
	Insgesamt	Ohne Vermögenseinkommen	
VGR-Ausgangswert aus der Sektorrechnung	202.408	181.953	155.387
Abzug private Organisationen ohne Erwerbszweck ¹⁾	-	-	3.635
Abzug der Bevölkerung außerhalb von Privathaushalten ²⁾	2.368	2.129	1.818
Abzug von Entnahmen ³⁾	1.446	1.446	1.446
Abzug von fiktiven Mieten ⁴⁾	6.788	6.788	6.788
VGR-Schätzwert	191.806	171.590	141.700
EU-SILC Schätzwert	154.452	151.995	115.653
Differenz zw. VGR und EU-SILC	19,47	11,42	18,38

Quelle: Statistik Austria, EU-SILC 2007; VGR 2006

¹⁾ Schätzwert, für das verfügbare Einkommen wird ein gemeinsamer Schätzwert errechnet

²⁾ geschätzt auf der Basis der Bevölkerungsprognose; 1.17% im Jahr 2006

³⁾ geschätzt auf 1,3% der gesamten Konsumausgaben, KE 2004/05

⁴⁾ VGR 2006

Hinweis auf ergänzende Dokumentationen/Publikationen

[EU-Verordnungen](#), auf deren Basis die EU-SILC-Erhebungen ab 2004 erfolgen.

[Strukturindikatoren der Europäischen Union](#) (Beschreibung und Daten).

[Online-Versionen der Publikationen](#).