

# Standard-Dokumentation Metainformationen

(Definitionen, Erläuterungen, Methoden, Qualität)

zu

## EU-SILC 2006

Diese Dokumentation gilt für den Berichtszeitraum:  
**2005/2006**

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

Bearbeitungsstand: **19.05.2008**



STATISTIK AUSTRIA  
Bundesanstalt Statistik Österreich  
A-1110 Wien, Guglgasse 13  
Tel.: +43-1-71128-0  
[www.statistik.at](http://www.statistik.at)

# Inhaltsverzeichnis

<b>1. Zusammenfassung, wichtige Hinweise .....</b>	<b>3</b>
<b>2. Allgemeine Informationen.....</b>	<b>4</b>
Statistiktyp.....	4
Fachgebiet .....	4
Verantwortliche Organisationseinheit, Kontakt.....	4
Ziel und Zweck, Geschichte .....	4
Periodizität .....	5
Auftraggeber .....	5
Nutzer .....	5
Rechtsgrundlage(n).....	5
<b>3. Statistische Konzepte, Methodik.....</b>	<b>5</b>
Gegenstand der Statistik.....	5
Beobachtungs-/Erhebungs-/Darstellungseinheiten .....	6
Datenquellen.....	6
Meldeeinheit/Respondenten .....	6
Erhebungsform .....	6
Charakteristika der Stichprobe.....	6
Erhebungstechnik/Datenübermittlung .....	6
Erhebungsbogen (inkl. Erläuterungen).....	7
Teilnahme an der Erhebung.....	7
Erhebungs- und Darstellungsmerkmale, Maßzahlen; inkl. Definition .....	7
Verwendete Klassifikationen .....	9
Regionale Gliederung der Ergebnisse .....	9
<b>4. Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen.....</b>	<b>9</b>
Datenerfassung.....	9
Signierung (Codierung) .....	9
Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen.....	9
Imputation (bei Antwortausfällen bzw. unvollständigen Datenbeständen) .....	10
Hochrechnung (Gewichtung) .....	14
Erstellung des Datenkörpers, (weitere) verwendete Rechenmodelle, statistische Schätzmethoden ....	17
Sonstige qualitätssichernde Maßnahmen .....	18
<b>5. Publikation (Zugänglichkeit).....</b>	<b>18</b>
Vorläufige Ergebnisse .....	18
Endgültige Ergebnisse .....	18
Revisionen .....	18
Publiziert in: .....	18
Behandlung vertraulicher Daten.....	19
<b>6. Qualität .....</b>	<b>19</b>
<b>6.1. Relevanz.....</b>	<b>19</b>
<b>6.2. Genauigkeit.....</b>	<b>20</b>
<b>6.2.1. Stichprobenbedingte Effekte, Repräsentativität .....</b>	<b>20</b>
<b>6.2.2. Nicht-stichprobenbedingte Effekte .....</b>	<b>20</b>
Qualität der verwendeten Datenquellen .....	20
Abdeckung (Fehlklassifikationen, Unter-/Übererfassung) .....	21
Antwortausfall (Unit-Non Response, Item-Non Response) .....	21
Messfehler (Erfassungsfehler) .....	23
Aufarbeitungsfehler.....	24
Modellbedingte Effekte.....	24
<b>6.3. Rechtzeitigkeit und Aktualität.....</b>	<b>24</b>
<b>6.4. Vergleichbarkeit .....</b>	<b>25</b>
<b>6.5. Kohärenz.....</b>	<b>27</b>
<b>Hinweis auf ergänzende Dokumentationen/Publikationen .....</b>	<b>33</b>

# 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. Von Eurostat werden Zielvariablen vorgegeben, die von den Mitgliedstaaten für eine Mindestzahl von Haushalten geliefert werden müssen. Die konkrete Art der Datenerhebung bleibt den einzelnen Ländern vorbehalten. In Österreich wird EU-SILC bislang als CAPI-Erhebung (Computer Assisted Personal Interviewing) durchgeführt.<sup>1</sup>

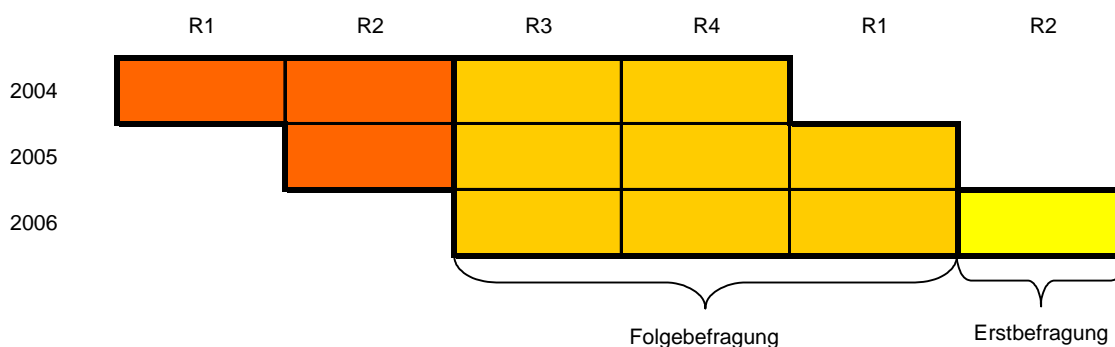
EU-SILC dient der Erfassung von Einkommen und Lebensbedingungen. Der Referenzzeitraum für das Einkommen der Haushalte und Personen und den Beschäftigungskalender ist das der Erhebung vorangegangene Kalenderjahr, 2006 also das Einkommen 2005. Alle anderen Merkmale beziehen sich auf den Erhebungszeitpunkt.

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 Querschnitterhebung durchgeführt wurde, startete die Erhebung im Jahr 2004 als integrierte Längsschnitterhebung. 2006 ist somit das dritte 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). Das Grundprinzip der Rotationsstichprobe wird in untenstehender Grafik dargestellt.

**Abbildung 1: Rotationsstichprobe in EU-SILC**



Der Datensatz von EU-SILC 2006 enthält somit Haushalte bzw. Personen, die 2004 zum ersten Mal befragt wurden, Haushalte bzw. Personen, die 2005 zum ersten Mal befragt wurden und Haushalte bzw. Personen, die 2006 zum ersten Mal befragt wurden. Mit dem Jahr 2006 werden somit auch Längsschnittdaten zur Verfügung stehen, die 2 Jahre erfassen.

Die Feldarbeit von EU-SILC 2006 fand von April bis September 2006 statt und wurde – nach der Neuausschreibung der Erhebung für die Jahre 2005 bis 2007 - durch das Erhebungsinstitut Spectra durchgeführt.

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. 2006 wurden dabei Fragen zur „sozialen und kulturellen Teilhabe“ integriert.

<sup>1</sup> Für die kommenden Jahre ist die Verwendung der CATI-Technik für Folgebefragungsinterviews geplant.

## 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](mailto:richard.heuberger@statistik.gv.at)

Mag. Matthias Till

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

Elisabeth Kafka

Tel. +43 (1) 71128- 7278, e-mail: [elisabeth.kafka@statistik.gv.at](mailto: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).<sup>2</sup> 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, welches 1996 mit dem Projekt startete – begann das ECHP im Jahr 1994. Die letzte und nach nationaler Zählung siebente Welle<sup>3</sup> des ECHP wurde im Jahr 2001 erhoben.<sup>4</sup>

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 in allen anderen EU-15-Staaten, mit Ausnahme von Deutschland, den Niederlanden und Großbritannien, begannen 2004, ebenso wie Estland, Norwegen und Island.

---

<sup>2</sup> Diese Statistik wurde, ebenso wie EU-SILC, in den meisten teilnehmenden Staaten durch eine Erhebung erstellt, und in manchen Staaten vorrangig durch Registerdaten befüllt.

<sup>3</sup> Nachdem das Projekt auf europäischer Ebene 1994 gestartet wurde, handelte es sich nach europäischer Zählung um die achte Welle.

<sup>4</sup> 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>

Ab 2005 wird EU-SILC in allen bislang 25 EU-Staaten sowie in Norwegen und Island durchgeführt. Im Jahr 2006 starteten Bulgarien, Rumänien und die Türkei mit der Erhebung EU-SILC, und ab dem Jahr 2007 wird die Erhebung auch in der Schweiz durchgeführt.

2006 ist das dritte Jahr, in dem EU-SILC in Österreich als integrierte Quer- und Längsschnitterhebung durchgeführt wird. Das heißt ein Teil der Haushalte wurde nunmehr zum dritten Mal wiederbefragt 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 Soziale Sicherheit, Generationen und Konsumentenschutz. 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 BMSK getragen.

### **Nutzer**

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

### **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 2006 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 oder 2005 teilgenommen haben.

## **Datenquellen**

Keine. EU-SILC ist eine Primärerhebung.

## **Meldeeinheit/Respondenten**

Die Informationen des Haushaltsregisters sind durch die Interviewerinnen und Interviewer 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 (2.352 von insgesamt 12.010 Personen) durch Proxy-Interviews erfasst (vgl. das Kapitel zu Messfehlern).

## **Erhebungsform**

Stichprobenerhebung

## **Charakteristika der Stichprobe**

### Grundgesamtheit und Auswahlrahmen

Grundgesamtheit sind private Haushalte, das sind Wohnungen mit zumindest einer mindestens 16-jährigen Person, die ihren Hauptwohnsitz in dieser Wohnung hat. Auswahlrahmen war das zentrale Melderegister (ZMR) mit Stand vom 31.12.2005. Die Haushalte der Stichprobe EU-SILC 2003, EU-SILC 2004 und EU-SILC 2005 wurden aus dem Auswahlrahmen für die Ersterhebung ausgeschlossen.

### Samplingverfahren

Für die Ersterhebung wurden die ursprünglich 3.588 Adressen mittels einfacher Zufallsauswahl aus dem ZMR gezogen. 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. Die Ausschöpfung der Erststichprobe muss darüber hinaus 60% betragen, zusätzlich wird eine Rate von 5% neutraler Ausfälle angenommen.

## **Erhebungstechnik/Datenübermittlung**

Wie bereits im Vorjahr wurde die Feldarbeit an die Spectra MarktforschungsgesmbH ausgelagert. Die Erhebung wurde vom 6. April bis zum 24. September mittels CAPI-Technik durchgeführt.

Dabei werden in persönlichen Interviews die Antworten der Respondentinnen und Respondenten direkt in einen Laptop-PC eingegeben. Spectra setzte dafür 106 Interviewende ein. Die Rohdaten wurden von den Interviewerinnen und Interviewern via Internet laufend an Spectra übermittelt und dort aufbereitet.

Spectra übermittelte die Daten in insgesamt sechs Datenlieferungen an die Statistik Austria. Die letzte Datenlieferung war die Enddatenlieferung.

### **Erhebungsbogen (inkl. Erläuterungen)**

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

Erhebungsmerkmale auf Personenebene: Einkommenskomponenten auf Personenebene (brutto und netto), demographische Daten, Fragen zur Erwerbstätigkeit, zur Bildung, zur Gesundheit und zur Zufriedenheit. Im Jahr 2006 kommen zusätzlich noch Fragen zur kulturellen und sozialen Teilhabe hinzu.

#### *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 (2005).

#### *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

Wohngeld, 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, 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. Ab 2007 wird das Haushaltseinkommen noch weitere Komponenten umfassen, z.B. unterstellte Mieten (*imputed rents*).

#### 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 2006 das Einkommen im Jahr 2005.

#### 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 Besiedlungsdichte<sup>5</sup> 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 [ISCO-88-COM - Berufsklassifikation](#) (2-Steller) und 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 2006 berechneten Indikatoren werden nicht regional gegliedert. In der Publikation wird eine Differenzierung nach Gemeindegrößenklassen<sup>6</sup> und nach Bundesländern verwendet.<sup>7</sup> 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)

### Signierung (Codierung)

Keine

### Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen

Es zählt zu den großen Vorteilen einer CAPI-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 Spectra implementiert.

Zu diesen Plausibilitätschecks im CAPI-Programm gehören Errors und Warnings<sup>8</sup>. 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 angege-

---

<sup>5</sup> Hohe Besiedlungsdichte: eine Gruppe aneinander grenzender Gemeinden mit jeweils mehr als 500 Einwohner pro km<sup>2</sup> und insgesamt mindestens 50.000 Einwohner.

Mittlere Besiedlungsdichte: eine Gruppe aneinander grenzender Gemeinden mit 101-500 Einwohner pro km<sup>2</sup> und insgesamt mindestens 50.000 Einwohner.

Geringe Besiedlungsdichte: alle übrigen Gemeinden.

<sup>6</sup> 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).

<sup>7</sup> Vergleiche hierzu unten den Abschnitt ‚Stichprobenbedingte Effekte, Repräsentativität‘ sowie die Publikation *Einkommen, Armut und Lebensbedingungen. Ergebnisse aus EU-SILC 2006*, Statistik Austria 2007 ([http://www.statistik.at/wcm/idc/idcplg?ldcService=GET\\_PDF\\_FILE&RevisionSelectionMethod=LatestReleased&dDocName=030260](http://www.statistik.at/wcm/idc/idcplg?ldcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=030260)) und die darin angegebenen Hinweise zur Schwankungsbreite.

<sup>8</sup> 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.

ben 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 Angabe eher unwahrscheinlich ist.

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. Dazu erhielten sie regelmäßig Rückmeldung von Statistik Austria, wo die Daten dann eingehend plausibilisiert wurden.

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. Dazu ist anzumerken, dass oftmals nicht klar ist, wie viele dieser fragwürdigen Antworten tatsächlich vom Respondenten bzw. der Respondentin kamen und in wie vielen Fällen es sich um Fehler des Interviewers bzw. der Interviewerin handelt. 2006 wurden die Möglichkeiten der Längsschnittplausibilisierung, also die Kontrolle und der Vergleich der Angaben mit den Angaben des Vorjahres weiterentwickelt.

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 2006 angewandten Verfahren der Imputation. Diese basieren auf den bereits für EU-SILC 2005 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. Die Distanzfunktion verwendet zur Bestimmung der Ähnlichkeit bestimmte 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 2006 zum ersten Mal interviewt worden sind, unterscheidet sich: kann für diese nur die Information aus dem aktuellen Jahr zur Bestimmung von ähnlichen Fällen verwendet werden, kann für jene 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 2006 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 36 Personeninterviews imputiert; im Vorjahr mussten 55 Personeninterviews ersetzt werden. In 19 Fällen stand dabei Information aus dem Vorjahr zur Verfügung, für 17 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 Längsschnittsimputation 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.

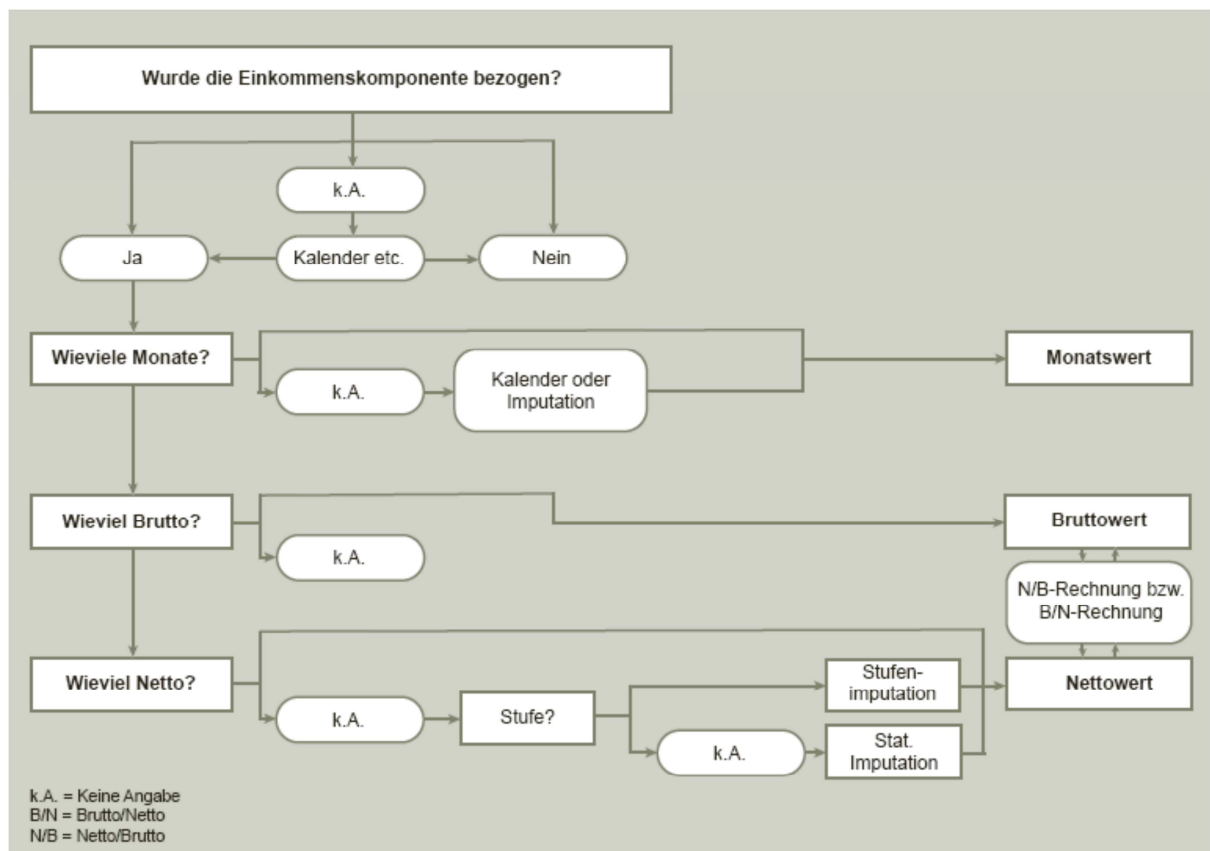
Als Querschnittsverfahren werden in EU-SILC 2006 Regressionsmodelle zur Berechnung von Schätzwerten verwendet. Für jede Einkommenskomponente wurden dabei mehrere Modelle spezifiziert, um sicherzustellen, dass auch beim Vorliegen von fehlenden Werten in den Prädiktorvariablen Schätzwerte berechnet werden 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, beim Arbeitslosengeld und bei unselbständigem Einkommen mehr als 90% der Angaben direkt aus der Befragung, so sind es bei Einkommen aus Land- und Forstwirtschaft 79% und beim Krankengeld nur etwas mehr als 72%. Im Vergleich zur Erhebung EU-SILC 2005 mussten etwas weniger Werte imputiert werden.

**Tabelle 1: Imputationen ausgewählter Nettoeinkommenskomponenten**

	Insgesamt	Betrag laut Befragung	Betrag aus Stufe	Brutto-Netto-Rechnung	Längsschnitt-Imputation	Querschnitt-Imputation	Sonstige Korrekturen
<i>Absolut</i>							
Unselbständigen Einkommen	6.186	5.916	0	8	103	118	41
Land-/forstwirtschaftlicher Betrieb	354	281	24	11	22	16	0
Arbeitslosengeld	530	498	9	0	6	16	1
Alterspension	2.992	2.667	83	19	104	61	58
Krankengeld	116	84	7	8	4	9	4
<i>Relativ</i>							
Unselbständigen Einkommen	100,0	95,6	0,0	0,1	1,7	1,9	0,7
Land-/forstwirtschaftlicher Betrieb	100,0	79,4	6,8	3,1	6,2	4,5	0,0
Arbeitslosengeld	100,0	94,0	1,7	0,0	1,1	3,0	0,2
Alterspension	100,0	89,1	2,8	0,6	3,5	2,0	1,9
Krankengeld	100,0	72,4	6,0	6,9	3,4	7,8	3,4

Quelle: Statistik Austria, EU-SILC 2006

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**

	Insgesamt	Betrag laut Befragung	Brutto-Netto-Rechnung	Längsschnitt-Imputation	Querschnitt-Imputation	Sonstige Korrekturen
Anzahl	6.185	5.916	7	103	118	41
In %	100,0	95,7	0,1	1,7	1,9	0,7
Arithmetisches Mittel	17.482	17.519	13.606	19.382	15.900	12.529
Median	16.800	16.800	11.036	18.567	15.654	7.000
Minimum	100	100	8.614	189	309	250
Maximum	134.630	134.630	20.895	72.089	38.707	88.400
Standardabweichung	11.153	11.147	5.081	11.785	8.725	15.241

Quelle: Statistik Austria, EU-SILC 2006

## 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 Gewichte getrennt für die 3 Teilstichproben, jeweils nach dem Jahr der Ersterhebung berechnet. Gegenüber dem letzten Jahr wurde die Gewichtung daher geändert und verbessert.

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

- a) 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.
- b) 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 mehreren Arbeitsschritten.

#### Designgewicht

Das Designgewicht wird berechnet, um Effekte des Designs der Stichprobe bei der Gewichtung zu berücksichtigen, denn das Stichprobendesign entscheidet über die Auswahlwahrscheinlichkeit eines Elements der Grundgesamtheit. Das Designgewicht wird als die Inverse der Auswahlwahrscheinlichkeit berechnet. Im Falle einer reinen Zufallsauswahl, wie bei den Erstbefragungshaushalten in EU-SILC, hat jedes Element der jeweiligen Teilstichprobe dasselbe Designgewicht. Das Designgewicht der drei Teilstichproben unterscheidet sich demnach durch die unterschiedliche Größe der Stichprobe, aber nicht durch die prinzipielle Art der Berechnung.

#### Non-Response-Gewichtung

Ziel der Non-Response-Gewichtung ist die Reduktion des Bias, der durch den Antwortausfall von Haushalten entsteht (Verweigerungen etc.). Die Korrektur dieser Verzerrung erfordert allerdings die Kenntnis der Antwortwahrscheinlichkeit für jeden Haushalt. Diese Antwortwahrscheinlichkeit ist in der Regel nicht bekannt und muss daher geschätzt werden. Grundlage für die Schätzung der Antwortwahrscheinlichkeit sind die empirischen Antwortraten für verschiedene Gruppen von Haushalten.

Für die Erstbefragungshaushalte wurde im Vorjahr nur die Antwortquote nach Bundesland und Urbanisierungsgrad berücksichtigt. In diesem Jahr wurden noch zusätzliche Informationen aus dem Auswahlrahmen (ZMR) herangezogen: die Anzahl Alter der Personen im Haushalt, sowie , die Staatsbürgerschaft und das Geschlecht.<sup>9</sup> 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, schließlich kann sich die Zusammensetzung des Haushalts zwischen den Wellen verändern. Personen, die auf Basis der natürlichen Bevölkerungsbewegung das Sample verlassen (Sterbefälle und Migration) reflektieren dabei 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 beide Teilsamples der Folgebefragung (Personen mit Ersterhebungsjahr 2004 bzw. 2005) zwei logistische Regressionsmodelle entwickelt. Bei der Entwicklung ergaben sich dabei zwei Herausforderungen: erstens die Auswahl der in das Model einbezogenen Variablen und zweitens, für die Teilstichprobe, die mit der Erhebung 2004 begonnen hat, der Umgang mit Personen, die nicht an der Erhebung EU-SILC 2005, aber wieder an der Erhebung 2006 teilgenommen haben.

---

<sup>9</sup> Diese Informationen über die Zusammensetzung der Personen im Haushalt 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 bzw. 2005 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 Ausfalls maximiert, und die Zahl der benötigten erklärenden Kategorien (Freiheitsgrade) minimiert. Die Modellierung folgt somit ausschließlich technischen Kriterien. Eine Beurteilung der substanzwissenschaftlichen Relevanz der erklärenden Variablen ist lediglich für die Auswahl der in die Optimierung berücksichtigten Kategorien maßgeblich. Die Gewichtungsfaktoren sind daher unabhängig von den üblicherweise notwendigen Entscheidungen für eine bestimmte Modellierung.

Werden kategoriale Variablen in ein logistisches Regressionsmodell integriert, werden die einzelnen Kategorien in Dummyvariablen umgewandelt. Diese kategorialen Variablen werden „blockweise“ verarbeitet, d.h. die zugehörigen Dummyvariablen werden entweder alle in das Modell integriert oder aus dem Modell ausgeschieden. Deshalb wurden bei kategorialen Variablen die einzelnen Kategorien in Dummyvariablen umgewandelt, bevor diese in das Modell integriert wurden. Im Falle von bundesländerspezifischen Ausfällen erhalten beispielsweise eben nur jene Bundesländer spezifische Non-Response-Gewichte, die auch tatsächlich signifikante Unterschiede aufweisen.

Personen die an der Erhebung 2004 und 2006, aber nicht an der Erhebung 2005 teilgenommen haben, weisen für das Jahr 2005 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 für das Modell für Personen, die 2004 mit der Erhebung begonnen haben, die Antwortwahrscheinlichkeit direkt zwischen den Jahren 2004 und 2006 geschätzt, das Jahr 2005 bleibt dabei unberücksichtigt.

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 oder 2006 in die Grundgesamtheit zugewandert sind, wurden entsprechend den Vorgaben von Eurostat die Querschnittsgewichte um einen fixen Faktor (3 bzw. 3/2) erhöht. Externe Randverteilungen aus der Wanderungsstatistik wurden nicht berücksichtigt. Für EU-SILC 2007 ist zukünftig eine Anpassung der Gewichte an die Anzahl der in Privathaushalte zugezogenen Personen (aus dem POPREG) beabsichtigt.

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.



### 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 2006 und für die Anzahl der Bezieher und Bezieherinnen von Arbeitslosenleistungen aus Verwaltungsdaten.

Auf Haushaltsebene wird an die Haushaltsgröße (1,2,3 und 4 und mehr Personen), den Mietstatus (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 heuer erstmalig mit dem von INSEE entwickelten und von Eurostat empfohlenen SAS-Makro CALMAR durchgeführt und stellt sicher, dass die genannten Verteilungen von Haushalten und Personen mit externen Eckzahlen übereinstimmen.

Die Gewichtung von EU-SILC 2006 – insbesondere die Einbeziehung der Anzahl der Bezieherinnen und Bezieher von Leistungen der Arbeitslosenversicherung und die Anzahl von Personen außerhalb von EU/EFTA – hat sich gegenüber dem Vorjahr 2005 nicht unbeträchtlich verändert. Im Abschnitt zur Vergleichbarkeit wird auf diesen Vergleich noch eingegangen.

### **Erstellung des Datenkörpers, (weitere) verwendete Rechenmodelle, statistische Schätzmethode**

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 2005, 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“ w. o. zu finden.

## Sonstige qualitätssichernde Maßnahmen

Um den Einfluss einer selektiven Adressenbearbeitung auf die Qualität der Ergebnisse einzuschränken, wurden in die Ausschreibung folgende Bestimmungen aufgenommen: Es sollten pro Haushalt zumindest drei Kontaktversuche erfolgen, bevor eine Verweigerung akzeptiert werden konnte, und es mussten regionale Antwortquoten erfüllt werden. Diese sind für Ersterhebung und Folgeerhebung unterschiedlich festgelegt: bei der Ersterhebung müssen in jeder Region mindestens 50% der Haushalte erfolgreich befragt werden (insgesamt ist die geforderte Ausschöpfungsquote in der Ersterhebung 60%), bei der Folgeerhebung müssen mindestens 75% aller ausgewählten Haushalte einer Region erfolgreich befragt werden (bundesweit 93%). Die geforderten regionalen Mindestausschöpfungsquoten führten zu einem hohen Einsatz des Erhebungsinstituts, auch als „schwierig“ empfundene Haushalte (v.a. im städtischen Bereich) zur Mitarbeit zu bewegen.

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 Interviewerinnen und Interviewer rückgemeldet werden und das Feldinstitut konnte notwendig gewordene Rückrufe bei den Interviewten machen.

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

## 5. Publikation (Zugänglichkeit)

### Vorläufige Ergebnisse

Keine

### Endgültige Ergebnisse

Siehe Standardpublikation und Internet

### Revisionen

Trifft nicht zu.

### Publiziert in:

#### Standardpublikation

[Einkommen, Armut und Lebensbedingungen. Ergebnisse aus EU-SILC 2006](#). Herausgegeben von Statistik Austria (inkl. CD-ROM); Wien 2007 (ISBN 978-3-902587-44-2).

#### 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 vom Erhebungsinstitut 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 lag aus der Stichprobe die Adresse 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 wurden Name, Adresse und Geburtstag aus den Datensätzen gelöscht. Im Datensatz verblieben 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 und 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: den Geschlechtsspezifischen Lohnunterschied (Gender Pay Gap), die Armutsgefährdungsquote vor und nach sozialen Transfers, das Verteilungsquintil und die Quote der dauerhaften Armutsgefährdung.

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

Darüber hinaus stehen auf der Website der Statistik Austria Mikrodaten von EU-SILC 2004 ([Standardisierte Datensätze SDS](#)) mit einer Zufallsstichprobe im halben Stichprobenumfang (2.260 Haushalte mit 5.751 Personen) zur Verfügung. 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 2006 19.673 Euro, der Schätzwert für den Standardfehler beträgt 174,7 Euro.<sup>10</sup> 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. 2005 beträgt die Obergrenze des Konfidenzintervalls 20.016 Euro und die Untergrenze 19.332 Euro.<sup>11</sup>

Der Designeffekt ist ein Maß für das Verhältnis der Varianz einer Schätzfunktion einer gegebenen Stichprobe zur Varianz der Schätzfunktion bei einer einfachen Zufallsauswahl. Für EU-SILC 2006 wurden die [Designeffekte für die Laeken-Indikatoren](#) nach einem Linearisierungsverfahren berechnet.

Die Berechnung der Konfidenzintervalle für die [Laeken-Indikatoren](#) wurde nach den beiden Methoden Linearisierungsverfahren und Bootstrapping-Verfahren durchgeführt. Für die Schwankungsbreiten der Armutsgefährdungsquoten der Bundesländer wurde nur das Linearisierungsverfahren verwendet.

**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		
<b>Österreich</b>	<b>12,6</b>	<b>1.027</b>	<b>11,6</b>	<b>952</b>	<b>13,5</b>	<b>1.102</b>	<b>6.028</b>	<b>14.883</b>
Burgenland	12,4	32	6,5	17	18,3	47	226	549
Kärnten	14,6	83	10,5	59	18,8	106	427	1.081
Niederösterreich	12,1	189	10,2	159	14,1	219	1.139	2.820
Oberösterreich	10,1	144	8,2	116	12,1	172	1.073	2.817
Salzburg	11,8	63	8,5	45	15,2	81	361	924
Steiermark	11,6	136	9,3	109	14,0	163	920	2.305
Tirol	10,0	70	7,0	49	13,0	91	498	1.320
Vorarlberg	9,2	35	4,1	16	14,2	54	270	733
Wien	17,2	276	14,7	236	19,7	315	1.114	2.334

Quelle: Statistik Austria, EU-SILC 2006

### 6.2.2. Nicht-stichprobenbedingte Effekte

#### Qualität der verwendeten Datenquellen

Nicht zutreffend, da eigene Datenerhebung.

<sup>10</sup> Dieser Wert resultiert aus der Varianzschätzung mit dem Verfahren der Linearisierung; wird die Varianzschätzung nach einem Bootstrapping-Verfahren durchgeführt, beträgt der Schätzwert für den Standardfehler 204 Euro.

<sup>11</sup> Diese Werte beziehen sich auf die Berechnung nach dem Linearisierungsverfahren. Berechnet wird das Konfidenzintervall nach der Formel:  $KI_{95} = 19.673 \pm 1,96 * 174,7$ . Berechnet nach dem Bootstrapping-Verfahren beträgt die Obergrenze 19.963 Euro und die Untergrenze 19.385 Euro.

## **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.

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 2006 8.450 Adressen verwendet. 75 dieser Adressen erwies sich als nicht existent. Von den verbleibenden 8.375 Adressen wurden 8.338 erfolgreich kontaktiert. Von diesen Adressen wurden schließlich 6.064 Haushalte erfolgreich interviewt, die restlichen 2.274 Haushalte verweigerten entweder die Auskunft, konnten nicht angetroffen werden, oder nahmen aus sonstigen Gründen nicht an der Erhebung teil. Allerdings stellten sich 36 dieser Haushaltsinterviews aufgrund von Qualitätsmängeln als nicht verwendbar heraus, sodass letztlich 6.028 Haushalte im Datensatz verblieben.

**Tabelle 4: Stichprobe EU-SILC 2006**

	Total		Erstbefragungsadressen		Folgebefragungsadressen	
	N	%	N	%	N	%
<b>Verwendete Adressen</b>	8.450	100,0	3.588	100,0	4.862	100,0
Adresse existiert	8.375	99,1	3.534	98,5	4.841	99,6
Adresse existiert nicht	75	0,9	54	1,5	21	0,4
<b>Bruttosample</b>	8.375	100,0	3.534	100,0	4.841	100,0
Adresse erfolgreich kontaktiert	8.338	99,6	3.515	99,5	4.823	99,6
Adresse nicht erfolgreich kontaktiert	37	0,4	19	0,5	18	0,4
<b>Erfolgreich kontaktiert Adressen</b>	8.338	100,0	3.515	100,0	4.823	100,0
Haushaltsinterview abgeschlossen	6.064	72,7	2.079	59,1	3.985	82,6
Niemand anwesend	615	7,4	397	11,3	218	4,5
Mitarbeit verweigert	1.465	17,6	964	27,4	501	10,4
Keine auskunftsfähige Person	11	0,1	10	0,3	1	0,0
Sonstige Ausfallgründe	183	2,2	65	1,8	118	2,4
<b>Erfolgreiche Haushaltsinterviews</b>	6.064	100,0	2.079	100,0	3.985	100,0
Interview aufgenommen in die Datenbank	6.028	99,4	2.058	99,0	3.970	99,6
Interview abgelehnt	36	0,6	21	1,0	15	0,4

Quelle: Statistik Austria, EU-SILC 2006

In diesen 6.028 Haushalten lebten zum Zeitpunkt des Interviews 14.883 Personen. 2.873 dieser Personen waren unter 16 Jahre alt, die übrigen 12.010 Personen waren mindestens 16 Jahre alt oder älter. Von diesen 12.010 Personen stehen Personeninterviews zur Verfügung; 11.974 dieser Interviews wurden dabei durch ein persönliches Interview oder durch Proxy-Interviews erfasst, die restlichen 36 Personeninterviews wurden imputiert. Von den 11.974 nicht imputierten Personeninterviews wurden 19,6% (2.352 Interviews) durch Proxy-Interviews erfasst. Bei der Erhebung EU-SILC 2005 betrug die Proxyrate 24,4%.

**Tabelle 5: Haushalte, Personen und Personeninterviews in EU-SILC 2006 nach Bundesland**

	Haushalte		Personen		Personeninterviews			Insgesamt
	befragt	Personen unter 16 Jahre	Personen 16 Jahre und älter	Insgesamt	persönliches Interview	Proxy-Interview	vollständig imputiert	
Burgenland	226	81	468	549	348	118	2	468
Kärnten	427	202	879	1.081	718	155	6	879
Niederösterreich	1.139	513	2.307	2.820	1.779	521	7	2.307
Oberösterreich	1.073	571	2.246	2.817	1.875	369	2	2.246
Salzburg	361	181	743	924	649	94		743
Steiermark	920	421	1.884	2.305	1.512	358	14	1.884
Tirol	498	291	1.029	1.320	718	310	1	1.029
Vorarlberg	270	191	542	733	398	142	2	542
Wien	1.114	422	1.912	2.334	1.625	285	2	1.912
<b>Gesamt</b>	<b>6,028</b>	<b>2,873</b>	<b>12,010</b>	<b>14,883</b>	<b>9,622</b>	<b>2,352</b>	<b>36</b>	<b>12,010</b>

Quelle: Statistik Austria, EU-SILC 2006

### 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 Information	
	N	%	N	%	N	%	N	%
gesamtes Bruttohaushaltseinkommen	6.028	100,00	2.045	33,93	3.532	58,59	451	7,48
gesamtes verfügbares Haushaltseinkommen	6.028	100,00	3.688	61,18	2.286	37,92	54	0,90
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>								
Einkommen aus Vermietung und Verpachtung	226	3,75	193	85,40	5	2,21	28	12,39
Familienleistungen/Kindergeld	2.120	35,17	2.110	99,53	9	0,42	1	0,05
Sonst. Leistungen gegen soziale Ausgrenzung	120	1,99	113	94,17	3	2,50	4	3,33
Wohnungsbeihilfen	204	3,38	195	95,59	7	3,43	2	0,98
Erhaltene Transfers zwischen privaten Haushalten	410	6,80	398	97,07	0	0,00	12	2,93
Zinsen, Dividenden	4.588	76,11	3.075	67,02	291	6,34	1.222	26,63
Einkommen von Personen unter 16	53	0,88	40	75,47	0	0,00	13	24,53
Geleistete Transfers zwischen privaten Haushalten	392	6,50	372	94,90	9	2,30	11	2,81
Einkommensteuernachzahlungen/-erstattungen	2.499	41,46	2.437	97,52	27	1,08	35	1,40
<i>Netto-Einkommenskomponenten auf Personenebene</i>								
Unselbständige Erwerbstätigkeit	6.254	52,07	3.768	60,25	362	5,79	2.124	33,96
Beiträge an private Rentensysteme	2.732	22,75	2.537	92,86	2	0,07	193	7,06
Selbständige Tätigkeit	1.098	9,14	524	47,72	69	6,28	505	45,99
Renten aus privaten Systemen	29	0,24	12	41,38	0	0,00	17	58,62
Arbeitslosenleistungen	724	6,03	659	91,02	30	4,14	35	4,83
Altersleistungen	3.045	25,35	1.289	42,33	683	22,43	1.073	35,24
Hinterbliebenenleistungen	105	0,87	41	39,05	27	25,71	37	35,24
Krankenleistungen	181	1,51	63	34,81	29	16,02	89	49,17
Invaliditätsleistungen	366	3,05	213	58,20	43	11,75	110	30,05
Ausbildungsleistungen	178	1,48	154	86,52	1	0,56	23	12,92
Bruttomonatsverdienste von Arbeitnehmern	5.682	47,31	4.141	72,88	43	0,76	1.498	26,36

Quelle: Statistik Austria, EU-SILC 2006

## 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-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-Programmierung durchgeführt.

Wichtiges Instrument der Fehleranalyse ist die Nachbefragung der Interviewerinnen und Interviewer (Debriefing). Die Ergebnisse des Debriefings sowie andere Rückmeldungen des Erhebungsinstituts zum Erhebungsinstrument sind in die CAPI-Gestaltung der Erhebung EU-SILC 2006 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 vom Erhebungsinstitut 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 Interviewerinnen und Interviewer rückgemeldet und vom Feldinstitut bei den Haushalten recherchiert werden.

Proxy-Interviews sind in EU-SILC in Ausnahmefällen gestattet. Nachdem der Anteil der Proxy-Interviews bei etwa 24% gelegen ist, konnte – nicht zu letzt auf Drängen der Statistik Austria – auf etwas unter 20% gesenkt werden.

### **Aufarbeitungsfehler**

Die Interviewerinnen und Interviewer übermittelten die durchgeführten Interviews über das Internet an das Feldinstitut. Dort wurden die Daten aufbereitet und in das passende Datenformat transferiert. Nach der Übermittlung der Daten an Statistik Austria 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.

Das Feldinstitut Spectra hat langjährige Erfahrung mit CAPI-Studien. Somit gab es wenige Probleme mit der CAPI-Programmierung und der Erfassung einzelner Fragen. Aufgrund der Auslagerung der Feldarbeit und der ersten Datenaufbereitung liegen allerdings keine Informationen über konkrete Aufarbeitungsfehler vor. Evidente Aufarbeitungsfehler wurden korrigiert.

### **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 2007 in einer ersten Version via eDamis an Eurostat übermittelt. Eine revidierte Version der Daten wurde am 14. September 2007 übermittelt.

Der standardisierte Qualitätsbericht wurde am 31. Oktober an Eurostat übermittelt.

Am 31. Jänner 2007 erhielt Eurostat den Abschlussbericht.



## 6.4. Vergleichbarkeit

2006 wurde EU-SILC zum vierten Mal in Österreich durchgeführt, und zum dritten Mal als integrierte Längsschnitterhebung. Das heißt, die Haushalte der Erhebung wurden zum 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 de Fragebogens.

Die für die Daten und die Auswertungen bedeutendste Änderung stellt die Neu-Berechnung der Gewichtung, bzw. die Berücksichtigung des Arbeitslosengeldbezugs bzw. der Staatsbürgerschaft für die Gewichtung dar. Im Folgenden soll beispielhaft dargestellt werden, in welcher Weise diese Neugewichtung die Resultate von EU-SILC beeinflusst. Dabei werden nun 4 verschiedene Gewichte miteinander verglichen:

- Die Gewichte unter Berücksichtigung von Staatsbürgerschaft und Arbeitslosenbezug: dies sind die neuen für EU-SILC 2006 ermittelten Gewichte.
- Die Gewichte ohne Berücksichtigung von Staatsbürgerschaft und Arbeitslosenbezug: dies entspricht dem alten Gewichtungsschema adaptiert für die Randverteilungen für EU-SILC 2006.

Um die Effekte der Berücksichtigung der beiden Variablen kontrollieren zu können, wurden zwei zusätzliche Gewichtsvarianten berechnet.

- Die Gewichte ohne Berücksichtigung des Arbeitslosenleistungsbezugs, aber unter Berücksichtigung der Staatsbürgerschaft
- Die Gewichte ohne Berücksichtigung der Staatsbürgerschaft, aber unter Berücksichtigung des Arbeitslosenleistungsbezugs

Die folgende Tabelle zeigt einige Ergebnis des Vergleichs. Bedeutsame Veränderungen ergeben sich dabei vorrangig für die Variablen Einkommen aus unselbständiger Arbeit und selbständiger Arbeit, Einkommen aus Arbeitslosenleistungen, Einkommen aus Krankenleistungen und Einkommen aus Vermögen. Diese Veränderungen ergeben sich vorrangig aus dem erhöhten Gewicht für Personen mit Arbeitslosenbezügen. Feststellbare Veränderungen aufgrund der Berücksichtigung der Staatsbürgerschaft lassen sich nur bei den Einkommenskomponenten Einkommen von Personen unter 16 Jahren und bei Renten aus privaten Systemen feststellen. Für das Gesamtergebnis sind diese allerdings nicht bedeutsam.

Durch die Berücksichtigung von Arbeitslosigkeit für die Gewichtung, wird die hochgerechnete Anzahl der Bezieherinnen und Bezieher von Arbeitslosenleistungen erhöht und damit auch das Gesamtvolumen der Arbeitslosenleistungen. Gleichzeitig verringert sich die Summe der unselbständigen und selbständigen Einkommen und der Einkommen aus Vermögen; dies vor allem durch eine Verringerung der jeweiligen Anzahl der Bezieherinnen und Bezieher. Einen Anstieg der Bezieherinnen und Bezieher und damit auch einen Anstieg des Volumens gibt es auch bei Krankenleistungen zu verzeichnen, ebenso wie bei der Anzahl der Wohnbeihilfen beziehenden Haushalte, wobei letztere Einkommenskomponente für das gesamte Haushaltseinkommen keine wesentliche Rolle spielt.

**Tabelle 7: Einkommenszielvariablen von EU-SILC 2006 mit verschiedenen Gewichtungsvarianten**

	Median				Anzahl				Summe (in Mio. Euro)			
	Gew. Neu	Gew. Alt	Gew. o. AL	Gew. o. ST	Gew. Neu	Gew. Alt	Gew. o. AL	Gew. o. ST	Gew. Neu	Gew. Alt	Gew. o. AL	Gew. o. ST
gesamtes Bruttohaushaltseinkommen	35.210	35.917	35.675	35.343	3.508.442	3.508.442	3.508.442	3.508.442	146.357	149.082	148.530	146.807
gesamtes verfügbares Haushaltseinkommen	27.371	27.660	27.594	27.430	3.508.442	3.508.442	3.508.442	3.508.442	110.635	112.260	111.875	110.953
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>												
Einkommen aus Vermietung und Verpachtung	3.960	4.014	4.014	3.960	123.239	124.807	124.268	123.686	1.185	1.225	1.219	1.189
Familienleistungen/Kindergeld	4.171	4.211	4.225	4.153	1.082.567	1.081.292	1.085.261	1.077.120	5.091	5.099	5.117	5.062
Sonst. Leistungen gegen soziale Ausgrenzung	1.456	1.680	1.680	1.456	78.423	72.822	72.710	78.701	241	234	232	243
Wohnungsbeihilfen	1.356	1.356	1.356	1.356	134.686	120.642	121.269	133.832	198	176	178	196
Erhaltene Transfers zwischen privaten Haushalten	3.240	3.240	3.360	3.240	232.620	225.061	225.728	230.995	1.094	1.071	1.081	1.080
Zinsen, Dividenden	67	70	70	69	2.589.627	2.626.125	2.617.551	2.594.540	875	913	906	880
Einkommen von Personen unter 16	1.400	1.408	1.408	1.400	28.309	25.098	27.686	25.139	47	41	45	42
Geleistete Transfers zwischen privaten Haushalten	3.000	3.000	3.000	3.000	239.306	239.685	239.916	239.285	897	907	907	898
Einkommensteuernachzahlungen/-erstattungen	-260	-260	-260	-260	1.398.084	1.410.303	1.404.190	1.401.352	-334	-332	-329	-336
<i>Netto-Einkommenskomponenten auf Personenebene</i>												
Unselbständige Erwerbstätigkeit	16.603	16.800	16.800	16.660	3.590.363	3.604.738	3.596.716	3.600.129	61.797	63.529	63.291	62.007
Beiträge an private Rentensysteme	840	840	840	836	1.511.467	1.563.851	1.545.474	1.524.028	1.647	1.714	1.695	1.660
Selbständige Tätigkeit	10.800	10.896	10.800	10.800	598.153	619.597	613.023	603.259	8.659	9.163	9.072	8.737
Renten aus privaten Systemen	1.800	1.800	1.800	1.800	16.950	17.266	17.263	16.973	62	62	62	62
Arbeitslosenleistungen	3.500	3.400	3.402	3.480	636.837	427.334	431.202	636.707	2.874	1.945	1.965	2.872
Altersleistungen	14.026	14.026	14.026	14.026	1.657.060	1.673.953	1.668.853	1.661.116	25.494	25.777	25.718	25.537
Hinterbliebenenleistungen	8.120	8.120	8.120	7.700	54.858	55.707	55.585	54.919	465	468	467	464
Krankenleistungen	1.493	1.600	1.560	1.493	125.305	104.601	105.103	125.807	347	308	308	349
Invaliditätsleistungen	12.600	12.600	12.600	12.600	199.876	201.133	198.978	201.498	2.508	2.555	2.526	2.530
Ausbildungsleistungen	1.800	1.920	1.920	1.800	104.969	108.248	107.372	105.970	323	330	327	326
Bruttomonatsverdienste von Arbeitnehmern	1.700	1.750	1.750	1.700	3.252.714	3.281.885	3.275.704	3.260.130	6.242	6.410	6.389	6.260

Quelle: Statistik Austria, EU-SILC 2006

Gew. Neu = Neues Gewichtungsschema

Gew. Alt = Altes Gewichtungsschema

Gew. o AL = Gewichtung ohne Berücksichtigung des Arbeitslosenleistungsbezug, mit Berücksichtigung der Staatsbürgerschaft

Gew. o ST = Gewichtung ohne Berücksichtigung der Staatsbürgerschaft, mit Berücksichtigung des Arbeitslosenleistungsbezug

Insgesamt ergibt sich durch die neue Gewichtung eine Verringerung der Summe der Haushaltseinkommen um etwa 2% verglichen mit der Gewichtung nach dem alten Gewichtungsschema.

Vergleicht man die mit den verschiedenen Gewichten berechneten Armutsgefährdungsquote, so ist diese durch die neue Gewichtungen mit Berücksichtigung des Arbeitslosenleistungsbezugs nur geringfügig höher als bei der Berechnung mit der alten Gewichtung. Allerdings führt die neue Gewichtung zu einer deutlich niedrigeren Armutsgefährdungsschwelle.

**Tabelle 8: Indikatoren von EU-SILC 2006 mit verschiedenen Gewichtungsvarianten**

	Gewichtung Neu	Gewichtung Alt	Gewichtung ohne AL	Gewichtung ohne ST
Median des äquivalisierten verfügbaren Haushaltseinkommens	17.852	18.152	18.050	17.914
Armutsgefährdungsschwelle Single	10.711	10.891	10.830	10.748
Armutsgefährdungsquote	12,6	12,3	12,3	12,5

Quelle: Statistik Austria, EU-SILC 2006

Gewichtung Neu = Neues Gewichtungsschema

Gewichtung Alt = Altes Gewichtungsschema

Gewichtung o AL = Gewichtung ohne Berücksichtigung des Arbeitslosenleistungsbezugs, mit Berücksichtigung der Staatsbürgerschaft

Gewichtung o ST = Gewichtung ohne Berücksichtigung der Staatsbürgerschaft, mit Berücksichtigung des Arbeitslosenleistungsbezugs

## 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 2006 mit externen Datenquellen verglichen. Einkommensbezugsjahr ist in EU-SILC das der Erhebung vorangegangene Jahr, d.h. im Jahr 2006 das Jahr 2005.

Folgende Datenquellen werden für den Vergleich mit den Daten von EU-SILC 2006 herangezogen: die Vorjahreserhebung EU-SILC 2005, die Lohnsteuerdaten 2005 und die Schätzungen auf der Basis der Sektorkonten der Volkswirtschaftlichen Gesamtrechnungen (VGR) des Jahres 2005.

### Vergleich mit EU-SILC 2005

EU-SILC 2005 war die zweite Erhebung des Frageprogramms EU-SILC im integrierten Quer- und Längsschnittdesign, was bedeutet, dass 2005 zum ersten Mal Haushalte zum zweiten Mal befragt wurden.<sup>12</sup> 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 ergeben, blieben unerkannt.

Die nachfolgende Tabelle vergleicht die Ergebnisse der Einkommenszielvariablen der Erhebungen 2005 und 2006. Die quantitativ bedeutendsten Variablen verändern sich in diesem Vergleich nicht dramatisch. Zwar gibt es etwa Veränderungen beim verfügbaren Haushaltseinkommen (der Median und das Volumen dieser Variable sinkt um etwa 2%), die auf diesen Werten beruhenden Indikatoren wie die Armutsgefährdungsquote, das arithmetische Mittel der äquivalisierten Einkommen und die Armutsgefährdungsschwelle bleiben aber zwischen EU-SILC 2005 und EU-SILC 2006 relativ konstant.

<sup>12</sup> Für eine genauere Beschreibung der Erhebung EU-SILC 2005 vgl. „Standard-Dokumentation Metainformation zu EU-SILC 2005“

**Tabelle 9: Vergleich der Einkommenszielvariablen von EU-SILC 2005 und EU-SILC 2006**

	Median			Haushalte/Personen			Summe (in Mio. Euro)		
	EU-SILC 2005	EU-SILC 2006	Veränderung %	EU-SILC 2005	EU-SILC 2006	Veränderung %	EU-SILC 2005	EU-SILC 2006	Veränderung %
gesamtes Bruttohaushaltseinkommen	36.425	35.210	-3,3	3.482.116	3.508.442	0,8	151.821	146.357	-3,6
gesamtes verfügbares Haushaltseinkommen	27.915	27.371	-2,0	3.482.116	3.508.442	0,8	112.519	110.635	-1,7
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>									
Einkommen aus Vermietung und Verpachtung	3.600	3.960	10,0	139.412	123.239	-11,6	1.221	1.185	-2,9
Familienleistungen/Kindergeld	4.080	4.171	2,2	1.117.030	1.082.567	-3,1	5.270	5.091	-3,4
Sonst. Leistungen gegen soziale Ausgrenzung	2.000	1.456	-27,2	73.604	78.423	6,5	202	241	19,6
Wohnungsbeihilfen	1.260	1.356	7,6	138.654	134.686	-2,9	199	198	-0,7
Erhaltene Transfers zwischen privaten Haushalten	3.600	3.240	-10,0	218.222	232.620	6,6	955	1.094	14,6
Zinsen, Dividenden	100	67	-32,8	2.698.868	2.589.627	-4,0	1.023	875	-14,5
Einkommen von Personen unter 16	1.750	1.400	-20,0	16.091	28.309	75,9	43	47	10,4
Geleistete Transfers zwischen privaten Haushalten	3.000	3.000	0,0	288.209	239.306	-17,0	1.114	897	-19,5
Einkommensteuernachzahlungen/-erstattungen	-250	-260	4,0	1.440.500	1.398.084	-2,9	-329	-334	1,4
<i>Netto-Einkommenskomponenten auf Personenebene</i>									
Unselbständige Erwerbstätigkeit	16.144	16.603	2,8	3.619.394	3.590.363	-0,8	62.498	61.797	-1,1
Beiträge an private Rentensysteme	800	840	5,0	1.559.756	1.511.467	-3,1	1.653	1.647	-0,4
Selbständige Tätigkeit	11.856	10.800	-8,9	636.806	598.153	-6,1	10.091	8.659	-14,2
Renten aus privaten Systemen	2.400	1.800	-25,0	29.449	16.950	-42,4	134	62	-53,7
Arbeitslosenleistungen	3.066	3.500	14,2	444.173	636.837	43,4	1.900	2.874	51,3
Altersleistungen	14.107	14.026	-0,6	1.640.795	1.657.060	1,0	26.330	25.494	-3,2
Hinterbliebenenleistungen	7.392	8.120	9,8	60.927	54.858	-10,0	478	465	-2,7
Krankenleistungen	1.500	1.493	-0,5	115.712	125.305	8,3	303	347	14,5
Invalitätsleistungen	12.600	12.600	0,0	194.936	199.876	2,5	2.478	2.508	1,2
Ausbildungsleistungen	1.575	1.800	14,3	111.485	104.969	-5,8	314	323	3,1
Bruttomonatsverdienste von Arbeitnehmern	1.700	1.700	0,0	3.290.013	3.252.714	-1,1	6.272	6.242	-0,5

Quelle: Statistik Austria, EU-SILC 2005, EU-SILC 2006

Hierzu tragen einerseits die geringen Veränderungen der Konzepte, Definitionen und Fragen zwischen 2005 und 2006 bei, andererseits auch die gestiegene Ausschöpfung insbesondere der Folgebefragungen.

Die Einkommenskomponenten auf Personenebene, die einen Großteil zum Haushaltseinkommen (etwa drei Viertel) beitragen, bleiben ebenfalls relativ konstant zwischen 2005 und 2006. Die selbständigen Einkommen gehen zwischen 2005 und 2006 allerdings zurück, und somit auch der Anteil der selbständigen Einkommen am gesamten Haushaltseinkommen (von 9,0% auf 7,8%). Eine Steigerung der Einkommen ist insbesondere bei Arbeitslosenleistungen zu verzeichnen. Dies ist vor allem auf die im Kapitel zur Gewichtung ausgeführte Berücksichtigung der Arbeitslosigkeit bei der Berechnung der Gewichte zurückzuführen.

Starke Veränderungen zeigen sich bei den sonstigen Leistungen gegen soziale Ausgrenzung und den Renten aus privaten Systemen. Die Veränderungen bei sonstigen Leistungen gegen soziale Ausgrenzung lassen sich auf die Veränderung der Gewichtung zurückführen: durch die Verwendung des neuen Gewichtungsschemas sinkt das mediane Einkommen, allerdings steigt die Anzahl der Bezieherinnen und Bezieher deutlich an, sodass sich letztlich das Volumen des Einkommens dieser Komponente im Vergleich zum Vorjahr erhöht. Renten aus privaten Systemen lässt sich keine einfache Erklärung finden. Allerdings wurden in beiden Jahren nur sehr wenige Personen mit Einkommen aus diesen Quellen erfasst, sodass die Ergebnisse für dieses Einkommen starken Zufallsschwankungen unterworfen sein dürfte.

Der Vergleich der Querschnittsergebnisse von EU-SILC 2005 und EU-SILC 2006 ist wichtig zur Beurteilung der Kohärenz der Ergebnisse von EU-SILC 2006; reale Veränderungen der Einkommen lassen sich klarer betrachten, wenn die Einkommen derjenigen Haushalte und Personen<sup>13</sup> 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.<sup>14</sup>

Hier zeigt sich, dass die Höhe des Medians bei der überwiegenden Zahl der Einkommenskomponenten relativ konstant bleibt. Eine Steigerung ist auch hier vor allem bei den Arbeitslosenleistungen zu verzeichnen. 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).

---

<sup>13</sup> 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.

<sup>14</sup> 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 10: Vergleich der Einkommenszielvariablen von EU-SILC 2005 und EU-SILC 2006 (Haushalte/Personen in beiden Erhebungen; Jahresbeträge)**

	Median		Veränderung %
	EU-SILC 2005	EU-SILC 2006	
gesamtes Bruttohaushaltseinkommen	35.648	35.756	0,3
gesamtes verfügbares Haushaltseinkommen	27.368	27.804	1,6
<i>Netto-Einkommenskomponenten auf Haushaltsebene</i>			
Einkommen aus Vermietung und Verpachtung	3.600	4.200	16,7
Familienleistungen/Kindergeld	4.098	4.080	-0,4
Sonst. Leistungen gegen soziale Ausgrenzung	2.000	1.200	-40,0
Wohnungsbeihilfen	1.464	1.236	-15,6
Erhaltene Transfers zwischen privaten Haushalten	3.480	3.480	0,0
Zinsen, Dividenden	99	80	-18,9
Einkommen von Personen unter 16	1.750	1.408	-19,5
Geleistete Transfers zwischen privaten Haushalten	3.000	3.000	0,0
Einkommensteuernachzahlungen/-erstattungen	-250	-270	8,0
<i>Nettoe-Einkommenskomponenten auf Personenebene</i>			
Unselbständige Erwerbstätigkeit	16.350	16.800	2,8
Beiträge an private Rentensysteme	792	840	6,1
Selbständige Tätigkeit	10.944	10.035	-8,3
Renten aus privaten Systemen	2.400	1.800	-25,0
Arbeitslosenleistungen	3.110	3.280	5,5
Altersleistungen	14.107	14.069	-0,3
Hinterbliebenenleistungen	8.400	7.700	-8,3
Krankenleistungen	1.500	1.560	4,0
Invaliditätsleistungen	12.600	12.600	0,0
Ausbildungsleistungen	1.683	1.400	-16,8
Bruttomonatsverdienste von Arbeitnehmern	1.750	1.750	0,0

Quelle: Statistik Austria, EU-SILC 2005, EU-SILC 2006

Zusammenfassend kann festgestellt werden, dass die Zahlen von EU-SILC 2006 in hohem Maße mit den Daten von EU-SILC 2005 in Einklang zu bringen sind. Die bestehenden Abweichungen lassen sich durch konzeptionelle Veränderungen (Gewichtung) erklären, bzw. sind Zufallsschwankungen aufgrund von geringen Fallzahlen geschuldet.

#### Vergleich mit den Lohnsteuerdaten 2005

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.<sup>15</sup> 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 3,64 Millionen unselbständig Beschäftigte, bei EU-SILC 2006 sind dies 3,59 Millionen, also etwa 54.000 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.

<sup>15</sup> 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 1 Monat im Jahr) nicht in den Lohnsteuerdaten berücksichtigt werden. Die Lohnsteuerdaten für 2005 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 11: Vergleich der Brutto-Jahreseinkommen unselbständig Beschäftigter 2005: Lohnsteuerdaten 2005 und EU-SILC 2006**

	Lohnsteuerdaten			Lohnsteuerdaten 1)			EU-SILC 2006		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	2.567	3.778	1.921	4.165	6.297	3.011	4.800	7.634	3.758
20% ...	7.620	12.050	5.086	9.435	14.817	6.538	10.200	15.167	7.118
25% ...	10.451	16.246	7.326	12.049	18.264	8.578	12.600	18.000	9.000
30% ...	13.093	19.470	9.350	14.516	20.896	10.432	14.700	19.600	10.500
40% ...	17.931	23.800	12.961	19.060	24.673	13.803	18.302	23.100	14.000
50% ...	22.320	27.375	16.296	23.113	28.075	16.994	22.120	26.400	16.802
60% ...	26.394	31.205	19.911	27.060	31.891	20.533	25.200	29.400	19.600
70% ...	30.937	36.180	23.951	31.579	36.904	24.510	29.322	33.960	22.800
75% ...	33.814	39.487	26.407	34.474	40.246	26.945	31.942	36.400	24.500
80% ...	37.441	43.729	29.285	38.115	44.525	29.812	35.000	40.500	27.403
90% ...	49.476	57.914	38.299	50.204	58.813	38.855	44.800	50.400	35.309
arith. Mittel	25.704	31.426	19.005	26.736	32.676	19.776	24.478	29.118	18.815
Personen	3.644.382	1.965.500	1.678.882	3.497.364	1.886.946	1.610.418	3.590.363	1.973.424	1.616.939

Quelle: Statistik Austria, EU-SILC 2006; Sonderauswertung zum Allgemeinen Einkommensbericht 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 2005 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 12: Vergleich der Brutto-Jahreseinkommen unselbständig Beschäftigter 2005: Lohnsteuerdaten 2005 und EU-SILC 2006 (nur ganzjährig Beschäftigte)**

	Lohnsteuerdaten			EU-SILC 2006		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	10.583	19.310	7.303	10.021	15.600	7.000
20% ...	16.295	23.696	11.700	15.358	20.400	10.970
25% ...	18.683	25.254	13.366	17.227	22.000	12.740
30% ...	20.795	26.690	14.932	18.570	23.100	14.070
40% ...	24.257	29.517	17.776	21.700	26.371	16.872
50% ...	27.550	32.719	20.791	24.333	28.710	19.320
60% ...	31.145	36.666	23.972	28.000	32.333	21.944
70% ...	35.775	42.015	27.862	32.200	36.400	24.520
75% ...	38.811	45.591	30.202	34.800	39.372	26.880
80% ...	42.668	50.031	33.073	37.179	42.679	29.400
90% ...	55.123	65.178	42.324	47.100	53.998	37.048
arith. Mittel	31.953	39.317	23.558	27.766	32.826	21.388
Personen	2.616.070	1.393.606	1.222.464	2.907.134	1.621.155	1.285.979

Quelle: Statistik Austria, EU-SILC 2006; Sonderauswertung zum Allgemeinen Einkommensbericht 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 2006 nicht vollständig erfasst.
- Die VGR schätzt zum gesamten Einkommen das Ausmaß an Schwarzarbeit, Trinkgeldern und Eigenleistungen (bei privaten Baustellen u.a.). Dieser Anteil beträgt in etwa 9,4% des Bruttoinlandsprodukts. Dieser Anteil am verfügbaren Einkommen wurde im folgenden Vergleich nicht berücksichtigt, und kann möglicherweise einen Teil der Differenz der Schätzwerte erklären.
- 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 2005 2,32 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 13: Vergleich des Haushaltseinkommens und des verfügbaren Einkommens VGR 2005 und EU-SILC 2006 (in Mio. Euro)**

	Brutto-Haushaltseinkommen		Verfügbares Einkommen
	Insgesamt	Ohne Vermögenseinkommen	
VGR-Ausgangswert aus der Sektorrechnung	195.693	174.505	151.055
Abzug private Organisaationen ohne Erwerbszweck 1)			4.616
Abzug der Bevölkerung außerhalb von Privathaushalten 2)	2.211	1.908	1.706
Abzug von Entnahmen 3)	2.544	2.269	1.964
Abzug von fiktiven Mieten 4)	6.113	6.113	6.113
VGR-Schätzwert	184.825	164.215	136.656
EU-SILC Schätzwert	146.357	145.250	110.635
Differenz zw. VGR und EU-SILC in %	20,81	11,55	19,04

Quelle: Statistik Austria, EU-SILC 2006; VGR 2005

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

2) geschätzt auf der Basis der Bevölkerungsprognose; 1,13% im Jahr 2005

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

4) VGR 2005

## 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](#).