

Standard-Dokumentation Metainformationen

(Definitionen, Erläuterungen, Methoden, Qualität)

zu

EU-SILC 2021

Diese Dokumentation gilt für folgende Berichtszeiträume:
2021

Die Statistik war Gegenstand eines Feedback-Gesprächs zur Qualität am 24.04.2014.

Bearbeitungsstand: **30.11.2022**



Die Informationsmanager

STATISTIK AUSTRIA
Bundesanstalt Statistik Österreich
A-1110 Wien, Guglgasse 13
Tel.: +43 1 711 28-0
www.statistik.at

**Direktion Bevölkerung
Bereich Soziales und Lebensbedingungen**

Ansprechperson:
MMag. Thomas Glaser
Tel.: +43 1 711 28-7039
E-Mail: thomas.glaser@statistik.gv.at

Ansprechperson:
Mag. Richard Heuberger
Tel. +43 1 711 28-8285
E-Mail: richard.heuberger@statistik.gv.at

Inhaltsverzeichnis

Executive Summary	4
1 Allgemeine Informationen	8
1.1 Ziel und Zweck, Geschichte.....	8
1.2 Auftraggeber:innen.....	9
1.3 Nutzer:innen.....	9
1.4 Rechtsgrundlage(n).....	10
2 Konzeption und Erstellung	12
2.1 Statistische Konzepte, Methodik	12
2.1.1 Gegenstand der Statistik.....	12
2.1.2 Beobachtungs-/Erhebungs-/Darstellungseinheiten	12
2.1.3 Datenquellen, Abdeckung	12
2.1.4 Meldeeinheit/Respondent:innen	13
2.1.5 Erhebungsform	13
2.1.6 Charakteristika der Stichprobe	14
2.1.7 Erhebungstechnik/Datenübermittlung.....	24
2.1.8 Erhebungsbogen (inkl. Erläuterungen).....	26
2.1.9 Teilnahme an der Erhebung.....	26
2.1.10 Erhebungs- und Darstellungsmerkmale, Maßzahlen; inkl. Definition	27
2.1.11 Verwendete Klassifikationen	34
2.1.12 Regionale Gliederung.....	35
2.2 Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen	35
2.2.1 Datenerfassung.....	35
2.2.2 Signierung (Codierung)	35
2.2.3 Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen.....	35
2.2.4 Imputation (bei Antwortausfällen bzw. unvollständigen Datenbeständen)	37
2.2.5 Hochrechnung (Gewichtung).....	42
2.2.6 Erstellung des Datenkörpers, (weitere) verwendete Rechenmodelle, statistische Schätzmethoden	56
2.2.7 Sonstige qualitätssichernde Maßnahmen	56
2.3 Publikation (Zugänglichkeit)	57
2.3.1 Vorläufige Ergebnisse	57
2.3.2 Endgültige Ergebnisse.....	57
2.3.3 Mikrodaten	57
2.3.4 Revisionen.....	57
2.3.5 Publikationsmedien	57
2.3.6 Behandlung vertraulicher Daten	58
3 Qualität.....	60

3.1	Relevanz	60
3.2	Genauigkeit	61
3.2.1	Stichprobenbedingte Effekte, Repräsentativität	61
3.2.2	Nicht-stichprobenbedingte Effekte	63
3.3	Aktualität und Rechtzeitigkeit	70
3.4	Vergleichbarkeit	71
3.4.1	Zeitliche Vergleichbarkeit/Veränderungen gegenüber dem Vorjahr	71
3.4.2	Internationale und regionale Vergleichbarkeit	71
3.5	Kohärenz	71
4	Ausblick	79
5	Glossar	79
6	Abkürzungsverzeichnis	79
7	Hinweis auf ergänzende Dokumentationen/Publikationen	80

Executive Summary

EU-SILC (Statistics on Income and Living Conditions) ist eine auf europäischer Ebene harmonisierte Statistik über Einkommen und Lebensbedingungen von Privathaushalten in Europa und 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, die es erlauben, die Lebenssituation von Menschen in Privathaushalten abzubilden.

EU-SILC ist die zentrale Quelle zur Erhebung der vom Europäischen Rat verabschiedeten Indikatoren zur sozialen Eingliederung zur Messung von Armut und sozialer Ausgrenzung. Das waren zuvor die vom Europäischen Rat in Lissabon (2000) gesteckten Ziele, bis 2010 die Reduktion von Armut deutlich voranzutreiben, das Verständnis von Armut und sozialer Ausgrenzung im europäischen Rahmen zu verbessern sowie den Erfahrungsaustausch zwischen den Mitgliedsländern zu fördern. Die in der Nachfolge dieser so genannten Lissabon-Strategie im Jahr 2010 beschlossene Strategie „Europa 2020“ definiert als Ziel die EU-weite Reduktion der Zahl an armuts- oder ausgrenzungsgefährdeten Personen um 20 Millionen. In Nachfolge der Europa 2020-Strategie wurde in der Europäischen Säule sozialer Rechte im zugehörigen Aktionsplan die Reduktion der Zahl der von Armut betroffenen Personen als wichtiges Ziel näher definiert. Im Zuge dessen wurde in den beiden Indikatorengruppen von SPC (Social Protection Committee, Sozialschutzausschuss) und EMCO (Employment Committee, Beschäftigungsausschuss) an der Überarbeitung des Indikatorensets zum Social Scoreboard gearbeitet. Die überarbeiteten Indikatoren, die sich stark an den bisher verwendeten Indikatoren orientieren, wurden für EU-SILC 2021 von Statistik bereits ausgewertet.

Die Verwendung von EU-SILC für die Überprüfung der Zielsetzungen auf europäischer Ebene unterstreicht die Bedeutung von EU-SILC für die europäische Sozialstatistik. Darüber hinaus ist anzumerken, dass EU-SILC eine der am meisten nachgefragten Datenquellen für sozialwissenschaftliche Forschungsprojekte auf nationaler und internationaler Ebene darstellt.

Seit dem Jahr 2005 umfasst die Erhebung EU-SILC zusätzlich jährlich wechselnde Sondermodule, mit denen bestimmte Bereiche der Lebensbedingungen genauer betrachtet werden sollen. Im Jahr 2021 wurden zusätzliche Fragen zu den Themen Verschuldung, Konsum, Vermögen und Beschäftigung gestellt.

In Österreich wurde EU-SILC erstmals 2003 als einmalige Querschnitterhebung von Statistik Austria durchgeführt. Mit 2004 begann die integrierte Längs- und Querschnitterhebung – das heißt jeweils rund drei Viertel der Haushalte werden auch im Folgejahr wieder befragt, ein Viertel der Stichprobe kommt jährlich neu hinzu. Die Stichprobe wird aus dem Zentralen Melderegister gezogen. Die Ausschöpfung der Erstbefragungsadressen betrug 2021 48,6 % der für die Befragung verfügbaren Adressen und lag damit etwas über dem Niveau der Ausschöpfung der Erstbefragung im Jahr 2020 (45,2 %). Die Ausschöpfung der Folgebefragungsadressen lag bei 84,6 %. Im Jahr 2021 umfasste die realisierte Stichprobe 12 274 Personen (davon 1 865 Personen jünger als 16 Jahre) in 6 021 Haushalten.

Alle Personen eines Haushaltes ab 16 Jahren werden persönlich befragt. Die Teilnahme an der Erhebung ist freiwillig und erfolgt mittels CAPI- (Computer Assisted Personal Interviewing) bzw. CATI- (Computer

Assisted Telephone Interviewing) Technik. Zusätzlich werden grundlegende Informationen zu Kindern im Haushalt erhoben. Rechtliche Grundlage der Erhebung sind EU-Verordnungen und die Einkommens- und Lebensbedingungen-Statistikverordnung (ELStV) des Sozialministeriums¹. Ab dem Jahr 2022 wird die europäische Rechtsgrundlage von EU-SILC mit der gemeinsamen Rahmenverordnung für die Sozialstatistik ("Integrated European Social Statistics – IESS") auf eine neue Basis gestellt. Damit bildet ab dem Erhebungsjahr 2022 die Verordnung (EU) Nr. 256/2021 den gemeinsamen Rahmen für europäische Statistiken über Personen und Haushalte – es wird dazu in der Standard-Dokumentation für EU-SILC 2022 berichtet werden. Die ELStV regelt unter anderem die Durchführung der Erhebung und die Möglichkeit der Verknüpfung mit Daten aus Verwaltungsregistern.

Wie auch in den Vorjahren werden mit dem Jahr 2021 wesentliche Komponenten des Haushaltseinkommens aus Verwaltungsdatenquellen berechnet. Etwa 87 % des Volumens des Gesamteinkommens wird aus Verwaltungsdaten ermittelt.

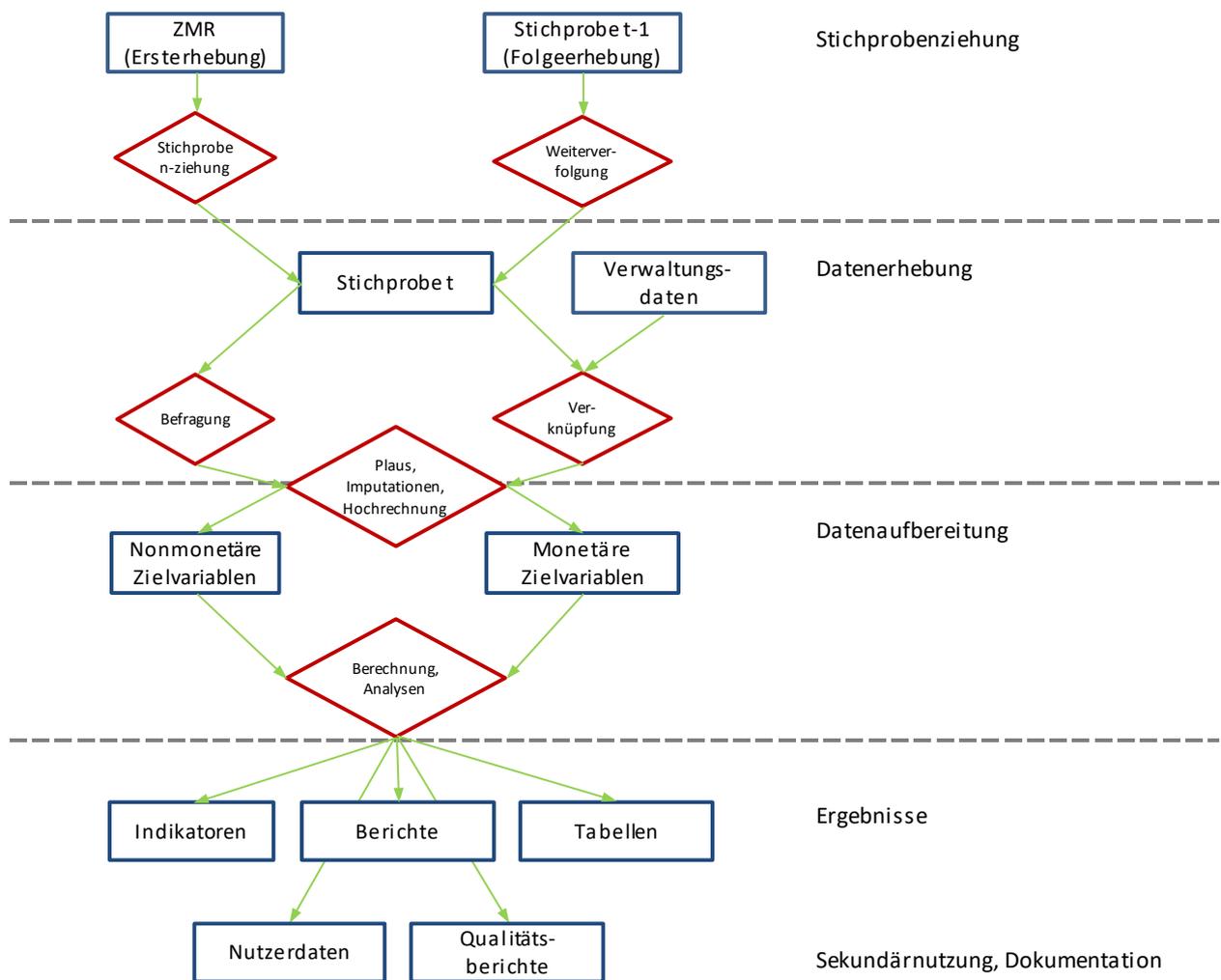
Die Feldarbeit des Jahres 2021 war durch die Bedingungen der Corona-Pandemie geprägt. Für die Erhebung hieß das mit den besonderen Bedingungen wie Kontaktbeschränkungen und Quarantänen in einzelnen Gemeinden umzugehen und Befragungszeitpunkt und -modus darauf abzustimmen.

Ergebnisse werden jährlich auf der Internetseite von Statistik Austria und in Form eines Tabellenbandes publiziert. Die anonymisierten Mikrodaten der österreichischen Datensätze werden kostenfrei für wissenschaftliche Zwecke von Statistik Austria zur Verfügung gestellt. Die europäischen Datensätze aller teilnehmenden Länder sind über Eurostat verfügbar.

Abbildung 1 beschreibt den Datenproduktionsprozess von EU-SILC von der Stichprobenziehung bis zum Vorliegen der Ergebnisse. In der vorliegenden Standarddokumentation werden die einzelnen Arbeitsschritte genauer erläutert.

¹ Verordnung des Bundesministers für Arbeit, Soziales und Konsumentenschutz über die Statistik der Einkommen und Lebensbedingungen, BGBl. II Nr. 277/2010.

Abbildung 1: Der Datenproduktionsprozess in EU-SILC



Q: STATISTIK AUSTRIA.

EU-SILC 2021 – Wichtigste Eckpunkte

Gegenstand der Statistik	Erfassung von Haushaltseinkommen und Lebensbedingungen nach Eurostat-Vorgaben
Grundgesamtheit	Wohnbevölkerung in Privathaushalten in Österreich
Statistiktyp	Primärstatistische Erhebung (CAPI- und CATI-Interviews) unter Einbeziehung von Verwaltungsdaten
Datenquellen/Erhebungsform	Stichprobenerhebung bei 6 018 Haushalten CATI- und CAPI-Interviews; integrierte Quer- und Längsschnitterhebung
Berichtszeitraum bzw. Stichtag	2021, Bezugsjahr der Jahreseinkommen: 2020
Periodizität	Jährlich
Teilnahme an der Erhebung (Primärstatistik)	Freiwillig
Zentrale Rechtsgrundlagen	IESS (VO (EU)2019/1700 und fünf ausführenden Verordnungen (drei Verordnungen zu allen IESS-Erhebungen, zwei EU-SILC spezifisch) Verordnung des Bundesministers für Arbeit, Soziales und Konsumentenschutz über die Statistik der Einkommen und Lebensbedingungen (Einkommens- und Lebensbedingungen-Statistikverordnung – ELStV) BGBl. II Nr 277/2010
Tiefste regionale Gliederung	Österreich – eingeschränkt Bundesländer (NUTS 2)
Verfügbarkeit der Ergebnisse	Endgültige Ergebnisse: Erhebungsjahr t + 4 Monate
Sonstiges	EU-SILC ist die Grundlage der europäischen Indikatoren zu Armut und sozialer Eingliederung EU-SILC ist Grundlage der nationalen Eingliederungsindikatoren

1 Allgemeine Informationen

1.1 Ziel und Zweck, Geschichte

In EU-SILC werden Informationen zur Einkommenssituation der Haushalte und aller Haushaltsmitglieder und deren Lebenssituation und Lebensbedingungen erfasst. Grundgesamtheit der Erhebung sind somit Personen in Privathaushalten in Österreich.

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 Strukturindikatoren und Indikatoren der sozialen Eingliederung.

Unmittelbares Vorläuferprojekt von EU-SILC war die Erhebung ECHP (European Community Household Panel). Dieses Projekt diente ebenfalls als Datengrundlage für vergleichende Daten zu Einkommen, Armutsgefährdung und Lebensbedingungen auf Ebene der Europäischen Union. Das ECHP war als reine Panelerhebung konzipiert. In Österreich wurde das ECHP seit 1995 (dem Jahr des EU-Beitritts) 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 des ECHP wurde im Jahr 2001 erhoben.

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. 2003 führten auch Belgien, Dänemark, Griechenland, Irland und Luxemburg sowie Norwegen eine EU-SILC-Erhebung durch.

Seit dem Jahr 2004 wird EU-SILC in Österreich als integrierte Quer- und Längsschnitterhebung durchgeführt. Auch alle anderen EU-15-Staaten, mit Ausnahme von Deutschland, den Niederlanden und Großbritannien, begannen 2004, ebenso Estland, Norwegen und Island. Seit 2005 wird EU-SILC in allen bis dahin 25 EU-Staaten, sowie in Island, durchgeführt. Im Jahr 2006 startete Bulgarien und 2007 Rumänien mit EU-SILC. Ab 2006 wurde die Erhebung auch in der Türkei und ab 2007 in der Schweiz durchgeführt, ab 2010 auch in Kroatien. Serbien lieferte für das Jahr 2015 erstmals SILC-Daten an Eurostat. Im Jahr 2010 startete auch Mazedonien mit der probeweisen Durchführung der Erhebung.

2021 ist das 17. Jahr, in dem EU-SILC in Österreich als integrierte Quer- und Längsschnitterhebung durchgeführt wird. Nachdem bereits 2011 zum ersten Mal Verwaltungsdaten für die Berechnung eines Teils der Einkommenszielvariablen herangezogen wurden, erfolgt seit 2012 die Berechnung eines Großteils des gesamten Haushaltseinkommens auf Basis von Verwaltungsdaten. Eine Beschreibung zur Verwendung von Verwaltungsdaten findet sich im Kapitel zur Vergleichbarkeit (3.4.1).

Seit 2018 erfolgt die Erhebung in STATsurv, dem neu entwickelten Tool für die Datenerhebung aller Haushalts- und Personenerhebungen von Statistik Austria. Die Umsetzung des Workflows (Ablauf der Avisierung und Erhebungstätigkeit, Zuordnung zu Modes, Umzüge, Anlage von Splithaushalten usw.) sowie die Umstellung des Telefonstudios auf eine neue Telefonanlage und Integration in STATsurv stellten die größten Herausforderungen dar. Da die aus dem neuen System generierten Daten hinsichtlich ihrer Struktur gegenüber den früheren Daten abwichen, mussten die Reports und Rohdaten-Aufbereitungsschritte angepasst werden.

1.2 Auftraggeber:innen

Seit der Erhebung EU-SILC 2008 wird die Finanzierung allein durch das Sozialministerium getragen, davor wurde das Projekt von Eurostat kofinanziert.

1.3 Nutzer:innen

Nachfolgend wird eine Auswahl an regelmäßig die EU-SILC Daten nutzenden Institutionen angeführt.

Nationale Institutionen

- Bundeskanzleramt
- Bundesministerien
- Politische Institutionen (Nationalrat, Bundesrat, Landtage etc.)
- Interessenvertretungen (z. B. Sozialpartner, Kammern, Standesvertretungen etc.)
- Oesterreichische Nationalbank
- Österreichischer Rechnungshof
- Gebietskörperschaften (Bund, Länder, Gemeinden)
- Statistik Austria (interne Nutzer:innen)
- Wirtschaftsforschungsinstitute
- Bundesanstalt für Agrarwirtschaft und Bergbauernfragen (BAB)
- Agentur für Gesundheit und Ernährungssicherheit (AGES)
- Tourismusverbände
- Umweltbundesamt
- Österreichische Energieagentur
- E-Control
- Kuratorium für Verkehrssicherheit
- Bundesanstalt für Verkehr (BAV)

Internationale Institutionen

- Europäische Kommission
- Europäischer Rechnungshof
- Europäische Zentralbank
- OECD
- UNO bzw. Suborganisationen
- IWF

- FAO
- Non-Profit-Organisationen
- Internationale Energieagentur
- La Conference Européenne des Ministres des Transports (CEMT)

Sonstige Nutzer:innen

- Medien
- Bildungseinrichtungen
- Forschungseinrichtungen
- Gesundheitseinrichtungen
- Unternehmen
- Gemeinnützige Organisationen bzw. Non-Profit-Organisationen
- Allgemeine Öffentlichkeit

1.4 Rechtsgrundlage(n)

EU-SILC wird seit dem Jahr 2021 auf der Grundlage der Rahmenverordnung für die Erstellung Europäischer Statistiken für Personen und Haushalte (Integrierte Europäische Sozialstatistiken – IESS) durchgeführt. Diese regelt die Europäischen Erhebungen von Personen und Haushalte zu den Bereichen Arbeitsmarkt, Einkommen und Lebensbedingungen, Gesundheit, Bildung, die Nutzung von Informationstechnologien, die Zeitverwendung und den Verbrauch (Vo (EU) 2019/1700).

Die jährlichen Modulfragen werden in eigenen Verordnungen geregelt; für EU-SILC 2021 war dies die [Verordnung \(EG\) Nr. 414/2019](#) der Kommission, sie regelte die sekundären Zielvariablen zu den Themen Verschuldung, Konsum, Vermögen und Beschäftigung. Eine Übersicht über die Verordnungen, die auf europäischer Ebene EU-SILC regeln, findet sich auf der Webseite von Eurostat.²

Auf nationaler Ebene wurde am 31.8.2010 die Verordnung des Bundesministers für Arbeit, Soziales und Konsumentenschutz über die Statistik der Einkommen und Lebensbedingungen Einkommens- und Lebensbedingungen-Statistikverordnung – ELStV; ([BGBl. II Nr 277/2010](#)) erlassen, die die Erhebung und die Verknüpfung mit Verwaltungsdatensätzen regelt. Diese Verordnung wurde im Jahr 2013 ([BGBl. II, Nr 230/2013](#)), im Jahr 2018 ([BGBl. II, Nr. 313/2018](#)), im Jahr 2019 ([BGBl. II Nr.319/2019](#)) sowie im Jahr 2022 ([BGBl. II Nr.38/2022](#)) geändert. Die aktuell geltende Fassung kann unter <https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20006884> abgerufen werden.

Die derzeit gültigen Rechtsgrundlagen können auch über die [Internetseite von Statistik Austria](#) abgerufen werden. Ab 2022 wird die Rechtsgrundlage von EU-SILC mit der gemeinsamen Rahmenverordnung für die Sozialstatistik ("Integrated European Social Statistics – IESS") auf eine neue Basis gestellt.

² <https://ec.europa.eu/eurostat/web/income-and-living-conditions/legislation>

2 Konzeption und Erstellung

2.1 Statistische Konzepte, Methodik

2.1.1 Gegenstand der Statistik

EU-SILC dient der Erstellung von Statistiken zu Einkommen und Lebensbedingungen der Wohnbevölkerung in Privathaushalten in Österreich. Eingeschlossen sind damit alle Personen, die an Adressen von Privathaushalten leben, in denen mindestens eine Person laut Zentralem Melderegister (ZMR) ihren Hauptwohnsitz hat. Ausgenommen sind daher Personen in Anstaltshaushalten bzw. Gemeinschaftsunterkünften und Personen ohne festen Wohnsitz. Im Jahr 2021 lebten ca. 8,76 Mio. Personen in Privathaushalten in Österreich.

Persönlich befragt werden in EU-SILC Personen, die mindestens 16 Jahre alt sind. Von den unter 16-Jährigen werden nur wenige Personenmerkmale erhoben. Es werden alle Personen eines Haushalts erfasst, d.h. auch jene, die nicht ihren Hauptwohnsitz laut ZMR im für die Erhebung ausgewählten Privathaushalt haben.

Bezugszeitraum der Einkommensvariablen ist (außer für das aktuelle Einkommen) das der Erhebung vorangegangene Kalenderjahr, für EU-SILC 2021 also das Jahr 2020. Die meisten anderen Variablen beziehen sich auf den Erhebungszeitpunkt (Ausnahmen sind z.B. Beschäftigungskalender im Einkommensbezugsjahr, Frage nach der Veränderung des Haushaltseinkommens in den letzten 12 Monaten).

2.1.2 Beobachtungs-/Erhebungs-/Darstellungseinheiten

Grundgesamtheit der Erhebung EU-SILC sind Personen in Privathaushalten in Österreich. Erhebungseinheiten sind private Haushalte und deren Bewohner:innen, wenn in diesem Privathaushalt mindestens eine Person wohnt, die (1) ihren Hauptwohnsitz an der Adresse des Haushalts hat und die (2) mindestens 16 Jahre alt ist. Haushalte sind definiert als wirtschaftliche Einheiten, welche aus einer oder mehreren Personen bestehen. Befragt werden alle Personen eines für die Erhebung ausgewählten Haushalts.

2.1.3 Datenquellen, Abdeckung

EU-SILC ist eine Primärerhebung. Auswahlrahmen und Datengrundlage für die Stichprobenziehung ist das ZMR. Das ZMR wurde nach der letzten Volkszählung 2001 durch die Zusammenführung der Gemeindemelderegister erstmalig befüllt und wird auf Basis der Wohnsitzmeldungen der Gemeinden laufend aktualisiert. Es enthält für alle in Österreich gemeldeten Personen die jeweiligen Adressdaten der gemeldeten Haupt- und Nebenwohnsitze. Bei der Zusammenführung der Adressdaten von verschiedenen Personen eines Haushalts kann es dabei durch unterschiedliche Schreibweisen der Adresse zu nicht erkannten Wohnungszusammenhängen kommen. In der Regel befindet sich an einer bestimmten Adresse genau ein Haushalt. In seltenen Fällen kann es jedoch sein, dass es an einer

Adresse mehrere Haushalte, verstanden als wirtschaftliche Einheiten, gibt. Ob eine Adresse mehrere Haushalte enthält, kann erst im Zuge der Datenerhebung eindeutig festgestellt werden. Des Weiteren ist zu berücksichtigen, dass die so genannte Melderealität nicht immer mit der Lebensrealität von Personen übereinstimmt, d.h. mitunter unterscheidet sich die tatsächliche, beim Interview erhobene Haushaltszusammensetzung von jener im ZMR.

2.1.4 Meldeeinheit/Respondent:innen

Die Befragung setzt sich aus fünf Teilen zusammen: dem Haushalts- und Personenregister sowie dem Haushalts-, Personen- und Kinderfragebogen. Das Haushaltsregister enthält grundlegende Informationen zu den Haushalten der Bruttostichprobe. Die Angaben für das Personenregister (enthält Informationen zu allen Personen der befragten Haushalte), der Haushaltsfragebogen und der Kinderfragebogen werden von einem Mitglied des Haushalts erfragt. Zusätzlich werden mittels Personenfragebögen alle Personen des Haushalts ab 16 Jahren einzeln befragt. Fremdauskünfte – so genannte Proxy-Interviews – sind nur in Ausnahmefällen gestattet (vgl. Kapitel 2.1.6.1 und Kapitel 3.2.2.4).

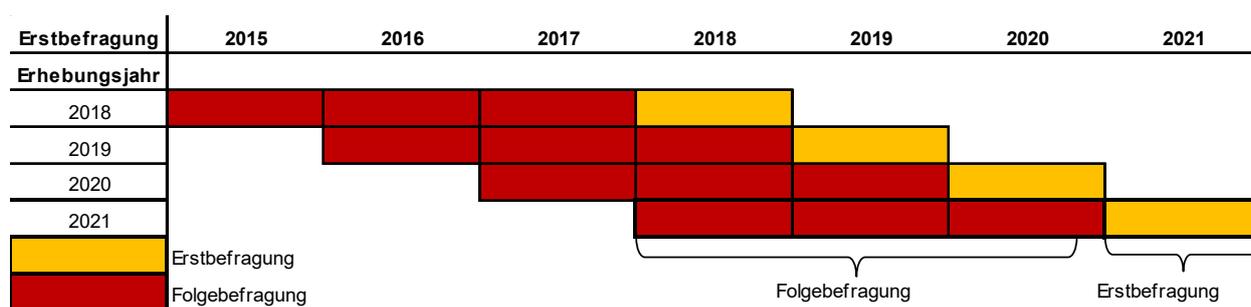
2.1.5 Erhebungsform

Seit dem Jahr 2004 wird EU-SILC mit einem integrierten Quer- und Längsschnittdesign als rotierende Panelerhebung durchgeführt (Abbildung 2). Das bedeutet, dass die Stichprobe in etwa vier gleich große Rotationsgruppen geteilt wurde und jedes Jahr eine Rotationsgruppe aus der Stichprobe ausscheidet, zugleich kommt eine neue Rotationsgruppe zur Stichprobe hinzu. Jedes Viertel der Stichprobe verbleibt somit für vier Jahre im Panel.

2021 ist das 17. Jahr von EU-SILC als integrierte Quer- und Längsschnitterhebung. Haushalte der Erhebungen vor 2018 befinden sich nicht mehr in der Stichprobe.

Die Erhebung EU-SILC 2021 wurde wie auch in den Vorjahren vollständig von Statistik Austria durchgeführt. Die Interviews wurden durch persönliche (CAPI) und telefonische (CATI) Befragungen erhoben. Zusätzlich wurde im Falle von COVID-19-bedingten Kontaktbeschränkungen und Quarantänen Telefoninterviews durch Erhebungspersonen des CAPI-Stabes zugelassen – siehe Kapitel 2.1.7.

Abbildung 2: Zusammensetzung der Stichprobe EU-SILC 2021



Q: STATISTIK AUSTRIA.

2.1.6 Charakteristika der Stichprobe

2.1.6.1 Zusammensetzung der Querschnittstichprobe

Die in der europäischen EU-SILC-Verordnung (EG) 1177/2003 vorgegebene effektive Stichprobengröße für Österreich umfasst 4 500 Haushalte. Die dafür tatsächlich benötigte Anzahl an Befragungen hängt vom Designeffekt der Armutsgefährdungsquote ab. Dieser berücksichtigt Genauigkeitsverluste, die aus Abweichungen von einer einfachen Zufallsauswahl (simple random sampling) entstehen.³ Ist der Wert des Designeffekts größer als 1, dann müssen mehr Haushalte befragt werden, um dieselbe Genauigkeit zu erzielen, die eine einfache Zufallsauswahl von 4 500 Haushalten bieten würde.

Um den Indikator „Armutsgefährdungsquote“ mit derselben Genauigkeit wie bei einer einfachen Zufallsauswahl schätzen zu können, musste beim Stichprobendesign 2021 davon ausgegangen werden, dass um rund ein Drittel mehr Haushalte – also in Summe ca. 6 000 – befragt werden müssen, um die effektive Zielgröße zu erreichen.⁴ Die Nettostichprobengröße von EU-SILC 2021 umfasst letztlich 6 018 Haushalte und erfüllt somit die Vorgaben an die effektive Stichprobengröße.

Wie in jedem Erhebungsjahr war auch 2021 eine neue Rotationstichprobe erforderlich. Die dafür notwendigen Adressen für die Ersterhebung wurden mittels einer einstufigen, stratifizierten Wahrscheinlichkeitsstichprobe mit disproportionaler Allokation aus einem Abzug des ZMR gezogen (Stichtag 30. September 2020). Um doppelte Ziehungen innerhalb des Vierjahrespanels zu vermeiden und zu verhindern, dass ein Haushalt für zwei aufeinanderfolgende Vierjahrespanels gezogen wird, werden Personen, die in einer der Erst-Stichproben von EU-SILC 2017 bis 2020 oder Personen aus vergangenen beziehungsweise derzeit laufenden Erhebungen automatisch ausgeschlossen, deren Erhebungszeit sich mit der Erhebungszeit von EU-SILC um bis zu plus/minus 3 Monate überschneidet.

Da im Stichprobenrahmen eine Vielzahl an Einkommenskomponenten des EU-SILC Haushaltseinkommens aus Verwaltungsdaten vorhanden sind, konnte für die Erststichproben 2016 und 2017 vor der Stichprobenziehung eine Annäherung des Haushaltseinkommens für jede Adresse berechnet werden. Dafür wurden die Nettobeträge aller über das bereichsspezifische Personenkennzeichen (bPK) an einer Adresse verknüpfbaren Einkommensinformationen auf Haushaltsebene aggregiert. Ab EU-SILC 2018 wurde dieser Ansatz durch die Verwendung des neu verfügbaren „Richframe“ erweitert. Es handelt sich dabei um einen vierteljährlich generierten Rahmen der gesamten österreichischen Bevölkerung, der auf mehreren verschiedenen Registern basiert. Soziodemographische Variablen in Kombination mit den verfügbaren Einkommensinformationen wurden verwendet, um eine Schätzung des Merkmals „Armuts- oder Ausgrenzungsgefährdung“ (AROPE) mittels eines Machine Learning Algorithmus (Random Forest) für den gesamten Rahmen zu „trainieren“. Dieses für den Stichprobenrahmen vorhergesagte AROPE_{ML} wurde schließlich als

³ Vgl. Statistik Austria (2006), Kap. „Variance estimation methodology“, S. 80f.

⁴ Der Designeffekt auf Basis aus EU-SILC 2021 betrug nach Berechnung von Statistik Austria 1,30. Im Vorjahr lag der Wert ebenfalls bei 1,30.

Substratifikationsmerkmal innerhalb jedes Bundeslandes verwendet. Die Erststichprobe EU-SILC 2021 hat somit $9 \times 2 = 18$ Strata. Bundesländer sind aus erhebungstechnischen Gründen wiederum regional aufgeteilt in Interviewsprengel. Eine Substratifizierung der Interviewsprengel nach dem geschätzten AROPE_{ML} würde zu kleine Schichten bilden, in denen es in der Nettostichprobe erwartungsgemäß leere Sprengel gegeben hätte. Deshalb beschränkt sich die Stratifizierung wie oben beschrieben auf die Merkmale Bundesländer \times AROPE_{ML}. Um trotzdem eine weitgehend gleichmäßige Auslastung der Interviewer pro Sprengel zu erreichen, wurden innerhalb jedes Stratums mittels „Systematic Sampling“⁵ Adressen gezogen.

Das Stichprobendesign für die Erstbefragung EU-SILC 2021 wurde in Bezug auf die Allokation pro Bundesland im Vergleich zum Vorjahr adaptiert. Die zuvor verwendete Methodik der Schätzung der Ausschöpfungsquote hat sich aufgrund der Erfahrung der vergangenen drei Jahre als schwierig erwiesen, insbesondere, weil die Ausschöpfung von Jahr zu Jahr pro Bundesland mitunter stark schwankt und immer schwerer antizipierbar wird. Allokationskriterium ist seit 2021 die Präzision der Armut- oder Ausgrenzungsgefährdung (AROPE) im Querschnitt (basierend auf den zuletzt verfügbaren Ergebnissen aus EU-SILC 2020). Ziel dieses Stichprobendesigns ist es eine möglichst präzise Stichprobe (im Sinne des Standardfehlers von AROPE) zu erhalten. Deshalb wurde die Allokation der Stichprobe für SILC 2021 pro Bundesland in Bezug auf die Präzisionsvorgaben von Eurostat angepasst. Dabei wurde auf die von Eurostat definierte Präzision der Stichprobe in Bezug auf den Standardfehler (SE) von AROPE pro Bundesland⁶ eingegangen. Pro Bundesland wurde errechnet wieviel größer die Stichprobe sein müsste, um den Präzisionsanforderungen zu genügen⁷. Die daraus resultierende Verteilung der zu ziehenden Adressen wurde danach auf die benötigte Bruttostichprobe skaliert (da das Ergebnis der SE Abschätzung eine zu große Stichprobe erfordern würde).

Das endgültige Stichprobendesign verwendet somit eine disproportionale Allokation pro Bundesland. So wurden beispielsweise in Wien verhältnismäßig mehr Haushalte als bei proportionaler Stratifizierung gezogen, da in Wien die Ausschöpfungsquote erfahrungsgemäß geringer ist als in anderen Bundesländern. Das heißt, bei Bundesländern mit vergleichsweise geringer Ausschöpfungsquote wurde ein Oversample angewendet. Die Anzahl der zu ziehenden Haushalte weicht bis auf Wien allerdings nicht sehr stark von einer proportionalen Stratifizierung ab. Die Erstbefragungsstichprobe umfasste nach der Ziehung 3 866 Adressen. Um auf den Fall einer zu geringen Zahl an vollständigen Interviews in der ersten Erhebungswelle 2021 vorbereitet zu sein, wurde wie in jedem Jahr zur Erststichprobe vorsorglich eine Reserve im Ausmaß von einem Drittel der aktiven Erststichprobe gezogen. Für EU-SILC 2021 umfasste diese Reservestichprobe 1.289 Adressen. Im Laufe der Erhebung stellte sich schließlich heraus, dass die Anzahl der abgeschlossenen Interviews unter den Erwartungen blieb. Deshalb wurden 439 Adressen aus der Reserve zufällig, stratifiziert nach Bundesland, ausgewählt und Teil der Erhebung. Somit umfasste die Brutto-Stichprobe der ersten Welle von EU-SILC 2021 insgesamt 4 305 Adressen.

⁵ Vgl. Särndal et al. (2003), S. 73ff.

⁶ Ausgenommen Bundesländer unter 500 000 Einwohner (Burgenland, Vorarlberg)

⁷ Vereinfachte Standardfehlerabschätzung aufgrund des SE von simple random sampling

Das Stichprobendesign der Erststichprobe EU-SILC 2021 beruht auf einer stratifizierten Wahrscheinlichkeitsstichprobe mit disproportionaler Allokation und Systematic Sampling als Auswahlmechanismus. Die Stratifizierung beinhaltet insgesamt 18 Schichten, wobei pro Bundesland jeweils zwei durch AROPE_{ML} definierte Schichten vorhanden sind.

Wie bereits in Kapitel 2.1.5 erläutert, ist EU-SILC eine rotierende Panelerhebung mit einer Paneldauer von vier Jahren. Deswegen müssen in jedem Erhebungsjahr neben der Erststichprobe auch die Stichproben der drei Folgebefragungen bestimmt werden. Diese ergeben sich aus den Personen der bereits im Vorjahr befragten Haushalte. Im Unterschied zur Erststichprobe handelt es sich bei den Folgestichproben also nicht um Haushalts-, sondern um Personenstichproben.

Im zweiten Erhebungsjahr befinden sich all jene Personen in der Folgestichprobe, die im ersten Erhebungsjahr 14 Jahre oder älter waren und sich in einem Haushalt befanden, der an der Erhebung teilnahm. Diese Personen sind „Stichprobenpersonen“. Jeder Haushalt, der zumindest eine Stichprobenperson enthält, wird befragt. Zieht eine Stichprobenperson in einen anderen Privathaushalt innerhalb Österreichs, so wird sie weiterverfolgt und im neuen Haushalt befragt. Wenn es in Folgewellen mehrere Haushalte gibt, in denen Stichprobenpersonen des Ursprungshaushalts leben, werden alle diese Haushalte (Ursprungshaushalte und so genannte Splithaushalte) befragt. Zieht eine Nichtstichprobenperson in einen Haushalt, in dem eine oder mehrere Stichprobenpersonen leben, so wird auch diese zugezogene Person im Haushalt erfasst (bzw. persönlich befragt, sofern die Person mindestens 16 Jahre alt ist).⁸

Die Stichprobe der Folgebefragung von EU-SILC 2021 ergibt sich aus den Stichprobenpersonen erfolgreich befragter Haushalte im Jahr 2020. Diese Haushalte befinden sich im Erhebungsjahr 2021 in den Folgestichproben mit den Ersterhebungsjahren 2018, 2019 und 2020 (Abbildung 2).

2.1.6.2 Ausschöpfung der Querschnittstichprobe EU-SILC 2021

Die Stichprobe von EU-SILC 2021 umfasste brutto 9 145 Adressen, 4 305 davon für die Erstbefragung, 4 840 für die Folgebefragung. Die Stichprobe von EU-SILC 2021 umfasste brutto 9 145 Adressen, 4 305 davon für die Erstbefragung, 4 840 für die Folgebefragung⁹. Davon stellten sich insgesamt 257 Adressen als nicht existent heraus, drei Adressen konnten nicht bearbeitet werden. Die verbleibenden 8888 Adressen bilden die gültigen Adressen. Von diesen wurden 8 867 Adressen erreicht. 6 045 Haushalte wurden erfolgreich befragt. Bei 550 Haushalten war während der Feldarbeitszeit keine Person anwesend, 2 006 Haushalte verweigerten das Interview. Die Interviews von 27 Haushalten mussten schließlich aufgrund von Qualitätsmängeln ausgeschlossen werden, sodass insgesamt 6 021 Haushaltsinterviews für die Analyse verblieben (Tabelle 1).

⁸ Ausführliche Informationen hierzu finden sich in der [Verordnung \(EG\) Nr. 1982/2003 der Kommission](#).

⁹ Die Folgebefragungsadressen umfassen auch bereits die während der Erhebung erfassten sogenannten Splithaushalte. Das sind Haushalte, die durch den Auszug von Samplepersonen aus dem Ursprungshaushalt entstehen.

Tabelle 1: Stichprobe EU-SILC 2021

	Total		Erstbefragungsadressen		Folgebefragungsadressen	
	N	in %	N	in %	N	in %
Bruttostichprobe EU-SILC 2021*	9 145	100,0	4 305	100,0	4 840	100,0
Adresse bearbeitet	9 145	100,0	4 305	100,0	4 840	100,0
Adresse nicht bearbeitet	0	100,0	0	0,0	4 840	100,0
Bearbeitete Adressen	9 145					
Adresse existiert	8 888	97,2	4 051	94,1	4 837	99,9
Adresse existiert nicht**	257	2,8	254	5,9	3	0,1
Gültige Adressen	8 888	100,0	4 051	100,0	4 837	100,0
Adresse erreicht	8 867	99,8	4 046	99,9	4 821	99,7
Adresse nicht erreicht	21	0,2	5	0,1	16	0,3
Erreichte Adressen	8 867	100,0	4 046	100,0	4 821	100,0
Haushaltsinterview abgeschlossen	6 045	68,2	1 842	45,5	4 203	87,2
Mitarbeit verweigert	2 006	22,6	1 584	39,1	422	8,8
Niemand anwesend	550	6,2	484	12,0	66	1,4
Keine auskunftsfähige Person	87	1,0	55	1,4	32	0,7
Sonstige Ausfallsgründe	179	2,0	81	2,0	98	2,0
Haushaltsinterview durchgeführt	6 045	100,0	1 842	100,0	4 203	100,0
Interview aufgenommen in die Datenbank	6 018	99,6	1 834	99,6	4 184	99,5
Aus Qualitätsgründen abgelehnte Interviews	27	0,4	8	0,4	19	0,5

Q: STATISTIK AUSTRIA, EU-SILC 2021. – * Inklusive Splithaushalte bei Folgebefragungsadressen. – ** Bei Folgebefragungsadressen handelt es sich um umgezogene Haushalte, deren neue Adresse nicht gefunden wurde.

Tabelle 3 unterscheidet qualitätsneutrale und nicht qualitätsneutrale Ausfälle. Erstere beeinflussen die Qualität der Stichprobenerhebung nicht. Dabei handelt es sich etwa um Adressen, an denen kein Gebäude mehr existiert oder Haushalte, bei denen alle Haushaltsmitglieder ins Ausland verzogen sind. Die qualitätsneutralen Ausfälle setzen sich aus den nichtexistierenden und den qualitätsneutralen nicht erreichten Haushalten (Tabelle 2) zusammen. Nicht qualitäts-neutrale Ausfälle beeinträchtigen hingegen die Qualität der Stichprobe, da diese Verzerrungen in der Ausschöpfung verursachen. Dabei handelt es sich um Ausfälle, die etwa durch die Teilnahmeverweigerung des Haushalts verursacht werden bzw. um Haushalte, bei denen niemand anwesend war.

Tabelle 2: Ausfälle und Interviews auf Haushaltsebene EU-SILC 2021

Folge Erhebungsjahr	Insgesamt		Erstbefragung 2021		1. Folgebefragung 2020		2. Folgebefragung 2019		3. Folgebefragung 2018	
	N	in %	N	in %	N	in %	N	in %	N	in %
Bruttoadressen	9 145	100,0	4 305	100,0	2 021	100,0	1 498	100,0	1 321	100,0
Qualitätsneutrale Ausfälle	335	3,7	259	6,0	30	1,5	26	1,7	20	1,5
Nicht qualitätsneutrale Ausfälle	2 765	30,2	2 204	51,2	353	17,5	140	9,3	68	5,1
Erfolgreiche Interviews	6 045	66,1	1 842	42,8	1 638	81,0	1 332	88,9	1 233	93,3
Aus Qualitätsgründen abgelehnte Interviews	27	0,3	8	0,2	14	0,7	3	0,2	2	0,2
akzeptierte Interviews	6 018	65,8	1 834	42,6	1 624	80,4	1 329	88,7	1 231	93,2

Q: STATISTIK AUSTRIA, EU-SILC 2021.

In den 6 018 befragten Haushalten lebten insgesamt 12 340 Personen. Davon waren 1 905 Personen unter 16 Jahre alt und 10 435 mindestens 16 Jahre alt und kamen daher für ein persönliches Interview in Frage. Von den 10 435 Personen wurden 5 109 durch CAPI-Interviews befragt, 4 425 telefonisch mittels CATI-Interviews. Insgesamt wurden 852 Proxyinterviews durchgeführt, 362 mittels CAPI- und 490 mittels CATI-Technik. Die Proxyrate lag insgesamt bei 8,2 % und war mit 10,0 % bei CATI-Interviews höher als bei CAPI-Interviews (6,6 %). 50 Personeninterviews fehlten und mussten vollständig imputiert werden.

In Tabelle 4 ist die Anzahl der Haushalte, Personen und Personeninterviews in EU-SILC 2021 auch für die Bundesländer angegeben.

Tabelle 3: Haushalte, Personen und Personeninterviews in EU-SILC 2021 nach Bundesland Teil 1

	Befragte Haushalte	Personen		
		Insgesamt	Personen 16 Jahre und älter	Personen unter 16 Jahre
Österreich	6 018	12 340	10 435	1 905
Burgenland	227	483	421	62
Kärnten	410	879	734	145
Niederösterreich	1 144	2 462	2 052	410
Oberösterreich	876	1 818	1 559	259
Salzburg	326	676	571	105
Steiermark	898	1 864	1 604	260
Tirol	480	986	828	158
Vorarlberg	268	602	469	133
Wien	1 389	2 570	2 197	373

Tabelle 4: Haushalte, Personen und Personeninterviews in EU-SILC 2021 nach Bundesland Teil 2

	Personeninterviews					Vollständig imputiert
	Insgesamt	CAPI*		CATI**		
		Persönliches Interview	Proxy- Interview	Persönliches Interview	Proxy- Interview	
Österreich	10 435	5 109	362	4 424	490	50
Burgenland	421	258	12	129	22	0
Kärnten	734	429	50	221	31	3
Niederösterreich	2 052	899	68	954	117	14
Oberösterreich	1 559	781	34	683	58	3
Salzburg	571	286	29	226	27	3
Steiermark	1 604	820	59	640	78	7
Tirol	828	475	37	274	39	3
Vorarlberg	469	218	16	217	16	0
Wien	2 197	943	57	1 080	102	15

Q: STATISTIK AUSTRIA, EU-SILC 2021. – * Computerunterstütztes persönliches Interview. –

** Computerunterstütztes telefonisches Interview.

Tabelle 5 präsentiert die Ausschöpfung der Erhebung EU-SILC 2021. Die Ausschöpfung der Erstbefragungsadressen betrug 45,3 % und lag damit etwas niedriger als die Ausschöpfung der Erstbefragung im Jahr 2020 (48,6 %). Die Ausschöpfung der Folgebefragungsadressen lag bei durchschnittlich 86,6 %. Die höchste Ausschöpfung wies wie in den Vorjahren die Rotationsgruppe mit der längsten Verweildauer im Panel auf (3. Folgebefragung).

Tabelle 5: Ausschöpfung EU-SILC 2021 nach Rotationen (mit Splithaushalten)

Folge Erhebungsjahr	Insgesamt	Erst 2021	1. Folge 2020	2. Folge 2019	3. Folge 2018
Bruttostichprobe	9 145	4 305	2 021	1 498	1 321
Gültige Adressen	8 888	4 051	2 019	1 498	1 320
Erfolgreiche Interviews	6 018	1 834	1 624	1 329	1 231
Ausschöpfung in % ^{1,2}		45,3	80,4	88,7	93,2

Q: STATISTIK AUSTRIA, EU-SILC 2021.–

1) Für die Erstbefragung: Anzahl der erfolgreich befragten Haushalte durch die Anzahl der gültigen Adressen. – 2) Für die Folgebefragung: Anzahl der erfolgreich befragten Haushalte durch die Bruttostichprobe; Folgebefragungen enthalten auch Splithaushalte. Eine Gesamtausschöpfungsquote lässt sich hierbei nicht berechnen.

2.1.6.3 Die Entwicklung der Stichprobe seit EU-SILC 2018

Tabelle 6: Ausschöpfung der Haushalte 2018 - 2021

Erhebungsjahr		Erstbefragung				
		Insgesamt	2018	2019	2020	2021
2018	Bruttostichprobe	3 684	3 684			
	Gültige Adressen	3 459	3 459			
	Erfolgreiche Interviews	1 958	1 958			
	Ausschöpfung in %		56,6			
2019	Bruttostichprobe	6 261	1 990	4 271		
	Gültige Adressen	5 993	1 923	4 070		
	Erfolgreiche Interviews	3 326	1 486	1 840		
	Ausschöpfung in %		74,7	45,2		
2020	Bruttostichprobe	7 685	1 509	1 882	4 294	
	Gültige Adressen	7 466	1 504	1 880	4 082	
	Erfolgreiche Interviews	4 744	1 300	1 462	1 982	
	Ausschöpfung in %		86,1	77,7	48,6	
2021	Bruttostichprobe	9 145	1 321	1 498	2 021	4 305
	Gültige Adressen	8 888	1 320	1 498	2 019	4 051
	Erfolgreiche Interviews	6 018	1 231	1 329	1 624	1 834
	Ausschöpfung in %		93,2	88,7	80,4	45,3

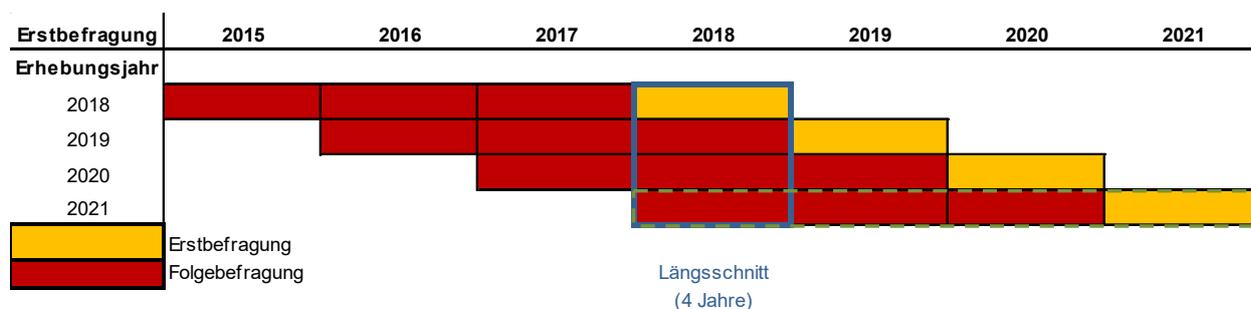
Q: STATISTIK AUSTRIA, EU-SILC 2018 - 2021. – Anmerkungen: (1) Ausschöpfung der Erstbefragung: Anzahl der erfolgreich befragten Haushalte dividiert durch die Anzahl der gültigen Adressen – (2) Ausschöpfung der Folgebefragungen: Anzahl der erfolgreich befragten Haushalte dividiert durch die Bruttostichprobe. – (3) Folgebefragungshaushalte enthalten auch Splithaushalte.

Tabelle 6 gibt einen Überblick der Entwicklung der Haushaltsstichprobe in Bezug auf Ausschöpfung und Größe seit dem Jahr 2018. Mit zunehmender Dauer einer Rotation steigt der Tendenz nach auch der Anteil der erfolgreichen Interviews an den ausgegebenen Adressen (Folgebefragungen). Haushalte, welche wiederholt befragt werden, zeigen in der Regel mit zunehmender Dauer der Erhebung eine hohe Bereitschaft an der Erhebung teilzunehmen, der Ausfall wird über die Zeit geringer.

2.1.6.4 Die Längsschnittstichprobe 2018-2021

Das integrierte Quer- und Längsschnittdesign von EU SILC erlaubt zwei analytische Perspektiven. Einerseits werden im Querschnitt alle Rotationen eines Jahres betrachtet. Andererseits ist im vierjährigen Längsschnitt nur eine Rotation von Interesse, wobei für diese jedoch Daten über vier Jahre vorhanden sind. Abbildung 3 stellt den Unterschied von Quer- und Längsschnitt anhand der Querschnittstichprobe 2021 (gestrichelte Linie) und der Längsschnittstichprobe 2018-2021 (blaue Umrandung) dar.

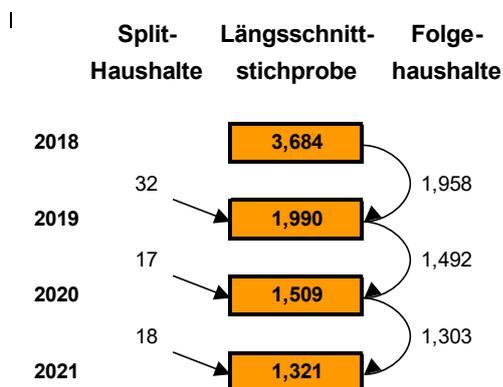
Abbildung 3: Querschnitt EU-SILC 2021 und Längsschnitt EU-SILC 2018-2021



Q: STATISTIK AUSTRIA.

Die Berechnung des Indikators „dauerhafte Armutsgefährdung“ basiert auf den Daten der Erhebungsjahre 2018-2021 (Erstbefragung 2018). Diese Rotation erlaubt im Rahmen des EU SILC Paneldesigns eine Perspektive im vierjährigen Längsschnitt. Abbildung 4 gibt einen Überblick über die Entwicklung dieses vierjährigen Längsschnitts.

Abbildung 4: Entwicklung des vierjährigen Längsschnitts 2018-2021 (Bruttostichprobe Haushalte)



Q: STATISTIK AUSTRIA.

Die Nettostichprobengröße des zusammengesetzten Vier-Jahres-Panels beträgt 2 307 Personen. Diese Zahl umfasst all jene Personen, die vier Jahre in Folge befragt wurden und für die somit in jedem Erhebungsjahr Daten zur Armutsgefährdung vorliegen. Da die zusammengesetzte vierjährige Längsschnittstichprobe nur aus einer Rotation besteht und nur Personen enthält, die in allen vier Jahren

teilnahmen, ist sie deutlich kleiner und beträgt auf Haushaltsebene nur 20,2 % der Querschnittstichprobe 2021.¹⁰

2.1.7 Erhebungstechnik/Datenübermittlung

Die Feldarbeit der Erhebung von EU-SILC 2021 wurde von durch Statistik Austria beauftragten Erhebungspersonen (Werknehmer:innen) sowie einen externen Telefonstudiobetreiber durchgeführt.

3 162 Haushalte wurden von Erhebungspersonen face-to-face befragt (CAPI). In 2 856 Haushalten wurde die Befragung durch das Telefonstudio durchgeführt (CATI). Anders als im Vorjahr wurden 2021 keine Interviews mittels TAPI befragt (Telefonische Interviews durch CAPI-Erhebungspersonen). Für 10 385 Personen, die mindestens 16 Jahre alt waren, wurde ein Personenfragebogen ausgefüllt. Davon wurden 8,2 % mittels Proxy-Interview (durch Fremdauskünfte) erhoben. 50 Interviews mussten vollständig imputiert werden.

CAPI-Interviews wurden von durch Statistik Austria beauftragten Erhebungspersonen (Werknehmer:innen) durchgeführt. Die Bearbeitung der CATI-Fälle wurde durch einen externen Betreiber abgewickelt. Dessen Interviewer:innen führten die Interviews in einem Telefonstudio in den Räumlichkeiten von Statistik Austria durch.

Eingeschränkt auf die 4 184 Folgebefragungsinterviews betrug der Anteil der CATI-Interviews 67,5 % (2640 Interviews).

Tabelle 7: Interviewmodus nach Rotationsgruppen 2021 (Haushaltsinterviews)

Folge Erhebungsjahr	Insgesamt		Erstbefragung 2021		1. Folgebefragung 2019		2. Folgebefragung 2018		3. Folgebefragung 2017	
	N	in %	N	in %	N	in %	N	in %	N	in %
CAPI	3 162	52,5	1 801	98,2	561	34,5	424	31,9	376	30,5
CATI	2 856	47,5	0	0,0	1 063	65,5	905	68,1	855	69,5
Insgesamt	6 018	100,0	1 834	100,0	1 624	100,0	1 329	100,0	1 231	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2021.

Auf Personenebene wurden insgesamt 5 471 Personen mittels CAPI befragt (52,7 %), 4 914 Personen mit CATI (47,3 %). Davon wurden 362 Personen in CAPI mit Proxyinterviews erfasst (6,6 % aller CAPI -

¹⁰ Der Prozentsatz bezeichnet den Anteil der 1.213 Haushalte, welche im Jahr 2021 die Personen aus dem zusammengesetzten Längsschnitt 2018-2021 enthalten, bezogen auf die Nettostichprobe von 6.018 Haushalten im Querschnitt 2021.

Interviews) und für 490 Personen konnten mit CATI-Interviews Proxyangaben ermittelt werden (10,0 % aller CATI-Interviews). Die Proxyrate insgesamt betrug damit 8,2 %. 50 Personeninterviews wurden vollständig imputiert und scheinen daher nicht in Tabelle 8 auf.

Tabelle 8: Interviewmodus 2021 (Personeninterviews)

	Insgesamt		CAPI		CATI	
	N	%	N	%	N	%
Persönliches Interview	9 533	91,8	5 109	93,4	4 424	90,0
Proxyinterview	852	8,2	362	6,6	490	10,0
Insgesamt	10 385	100,0	5 471	100,0	4 914	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2020.

Die Feldphase dauerte von Februar bis Juli, mit einzelnen Nachbearbeitungen im August 2021 (siehe Tabelle 9). Dies inkludiert eine Verlängerung der Feldzeit um rund 3 Wochen.

Tabelle 9: Interviews in der Feldarbeitszeit 2021

	Insgesamt		Erstbefragung		Folgebefragung	
	Interviews	%	Interviews	%	Interviews	%
Insgesamt	6 018	100,0	1 834	100,0	4 184	100,0
Februar	742	12,3	98	5,3	644	15,4
März	1 925	32,0	600	32,7	1 325	31,7
April	1 117	18,6	294	16,0	823	19,7
Mai	960	16,0	267	14,6	693	16,6
Juni	844	14,0	325	17,7	519	12,4
Juli	411	6,8	238	13,0	173	4,1
August	19	0,3	12	0,7	7	0,2

Q: STATISTIK AUSTRIA, EU-SILC 2021.

Bei den persönlichen oder telefonischen Interviews wurden die Antworten der Respondent:innen direkt in einen Laptop eingegeben. Statistik Austria setzte rund 160 Erhebungspersonen für die CAPI-Befragung ein. Die Rohdaten wurden von den Interviewer:innen via Datenverbindung laufend übermittelt.

2.1.8 Erhebungsbogen (inkl. Erläuterungen)

Nachdem EU-SILC eine elektronische Erhebung im CAPI/CATI-Modus ist, sind Erhebungsunterlagen nur insoweit vorhanden, als sie gemeinsam mit den Codebüchern und den vorgegebenen Plausibilitätschecks als Grundlage für die Programmierung dienen. Der Fragebogen wurde in der neuen Erhebungsumgebung STATSurv programmiert, und steht auf den PCs des Telefonstudios und den Laptops der Feldinterviewer:innen zur Verfügung. Eine Printversion des Fragebogens steht auf der Homepage von Statistik Austria zur Verfügung.

2.1.9 Teilnahme an der Erhebung

Die Teilnahme an EU-SILC ist freiwillig.

Zur Avisierung der Erhebung direkt vor Kontaktaufnahme wurden auf die Erhebungswelle abgestimmte Briefe versandt. Beigelegt waren für die Haushalte der Ersterhebung die Broschüre „Zahlen-Daten-Fakten“ und ein Informationsfolder und für Folgehaushalte ein SILC-Infowürfel und ein Pre-Incentive (2-Euromünze). Jeder Haushalt erhielt für die Teilnahme nach erfolgreichem Interview einen Einkaufsgutschein im Wert von 15 €.

2.1.10 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.

Für das Modul 2021 waren Fragen zu Lebensformen und Lebensbedingungen von Kindern in getrennt lebenden und Patchwork-Familien¹¹ zu stellen.

EU-SILC ist eine überwiegend outputharmonisierte Erhebung, d.h. Eurostat gibt die zu erhebenden Variablen und deren Operationalisierung vor. Die in EU-SILC in Österreich erhobenen Merkmale sind somit weitgehend vorgegeben. Einige Variablen sind hinsichtlich ihrer Fragestellung harmonisiert (Inputharmonisierung), z.B. Fragen zur Leistbarkeit bestimmter Tätigkeiten und Güter.

Einkommenskomponenten in EU-SILC

Diese Komponenten wurden sowohl brutto als auch netto erhoben. Das Nettoeinkommen auf Personenebene ist das Bruttoeinkommen abzüglich der Beiträge zur gesetzlichen Sozialversicherung sowie der Einkommensteuern. Fallen bei einer Komponente keine derartigen Abzüge an (so wie in Österreich z.B. bei der bedarfsorientierten Mindestsicherung), so ist das Bruttoeinkommen gleich dem Nettoeinkommen. Referenzzeitraum für die Einkommenskomponenten ist in EU-SILC das dem Erhebungsjahr vorangegangene Kalenderjahr (d.h. 2020 für EU-SILC 2021).

¹¹ "Living arrangements and conditions of children in separated and blended families"

Tabelle 10: Einkommenskomponenten in EU-SILC auf Haushalts- und Personenebene aus der Befragung

Erhebungseinheit	Einkommenskomponente	Eurostat Zielvariable*	Für Berechnung des Haushaltseinkommens
Haushalt	Staatliche Beihilfe zu Wohnkosten	HY070G/N	Ja
	Sonstige Familienleistungen	HY050G/N	Ja
	Sozialhilfe/BMS	HY060G/N	Ja
	Sonstige Bildungsleistungen	PY140G/N	Ja
	Imputierte Miete	HY030G/N	Nein
	Zinszahlungen	HY100G/N	Nein
	Unterhaltszahlungen	HY080G/N	Ja
Personen	Einkommen aus selbständiger Tätigkeit	PY050G/N	Ja
	Vermögenseinkommen	HY090G/N	Ja
	Private Pensionen	PY080G/N	Ja
	Sonstige private Transfers (erhalten und geleistet)	HY080G/N und HY130G/N	Ja
	Arbeitgeberbeiträge zur Sozialversicherung	PY030G/N	Nein

* Vgl. Eurostat 2021.

Tabelle 10 gibt einen Überblick der auf Haushalts- und Personenebene mittels Fragebogen zu erhebenden Einkommenskomponenten. Nicht alle Einkommenskomponenten bzw. Eurostat-Zielvariablen werden für die Berechnung des Haushaltseinkommens herangezogen (vgl. Tabelle 12). Die in Tabelle 10 angeführten Komponenten wurden entweder auf Monats- oder auf Jahresbasis erfragt. Konnten oder wollten Respondent:innen einen Betrag nicht nennen, so wurden sie gebeten, eine Stufe (Einkommensintervall) zu nennen.

Einen Überblick über die Einkommenszielvariablen, die vorrangig durch Verwaltungsdaten befüllt werden, gibt Tabelle 11. Vorrangig meint, dass einige Einkommenszielvariablen sich aus mehreren Einkommenskomponenten zusammensetzen, von denen der überwiegende Teil aus Verwaltungsdaten stammt. Dennoch ist es möglich, dass diese Zielvariablen Einkommensinformation aus der Erhebung enthalten. Beispielsweise umfasst die Variable PY140 Verwaltungsdaten zu Studienbeihilfe und der Schüler- bzw. der Heimbeihilfe, aber auch sonstige Bildungsleistungen von Ländern und Gemeinden aus der Erhebung.

Tabelle 11: Einkommenskomponenten in EU-SILC auf Haushalts- und Personenebene aus Verwaltungsdaten

Erhebungseinheit	Einkommenskomponente	Eurostat Zielvariable*	Für Berechnung des Haushaltseinkommens
Haushalt	Familienleistungen	HY050G/N	Ja
	Einkommen von Personen unter 16 Jahren	HY110G/N	Ja
	Steuernachzahlungen bzw. -rückerstattungen	HY145G/N	Ja
Personen	Einkommen aus unselbständiger Erwerbstätigkeit	PY010G/N	Ja
	Einkommen aus Präsenz-/Zivildienst	PY010G/N	Ja
	Arbeitslosenleistungen	PY090G/N	Ja
	Altersleistungen	PY100G/N	Ja
	Hinterbliebenenleistungen	PY110G/N	Ja
	Krankenleistungen	PY120G/N	Ja
	Invaliditätsleistungen	PY130G/N	Ja
	Ausbildungsleistungen	PY140G/N	Ja

* Vgl. Eurostat 2021.

Tabelle 12 gibt einen Überblick über die Größe der jeweiligen Einkommenskomponente und die Zusammensetzung des Haushaltseinkommens nach Verwaltungs- und Befragungsdaten.

Tabelle 12: Haushaltseinkommensmodell EU-SILC 2021

			Summe in Mio. Euro	Anteil an HY020 in %
	PY010	Unselbständigeneinkommen	101 579	55,0
+	PY050	Selbständigeneinkommen	14 289	7,7
+	PY090	Arbeitsloseneinkommen	5 510	3,0
+	PY100	Altersleistungen	38 925	21,1
+	PY110	Hinterbliebenenleistungen	4 772	2,6
+	PY120	Krankenleistungen	1 149	0,6
+	PY130	Invaliditätsleistungen	2 854	1,5
+	PY140	Ausbildungsleistungen	509	0,3
+	PY080	Renten aus privaten Systemen	1 218	0,7
=		Summe der Einkommen auf Personenebene	170 804	92,6
+	HY040	Einkommen aus Vermietung und Verpachtung	2 830	1,5
+	HY050	Familienleistungen	7 356	4,0
+	HY060	Sonstige Leistungen gegen soziale Ausgrenzung	1 334	0,7
+	HY070	Wohnungsbeihilfen	295	0,2
+	HY080	Erhaltene Transfers zwischen privaten Haushalten	1 681	0,9
+	HY090	Zinsen und Dividenden	1 662	0,9
+	HY110	Einkommen von Personen unter 16 Jahren	136	0,1
=		Summe der Einkommen auf Haushaltsebene	15 295	8,3
-	HY130	Geleistete Transfers zwischen privaten Haushalten	2 079	1,1
-	HY145	Einkommensteuernachzahlungen/-erstattungen	-501	-0,3
=	HY020	Verfügbares Haushaltseinkommen	184 522	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2021.

Eine Übersicht über die verwendeten Verwaltungsdatenquellen gibt Tabelle 13. Einige dieser Datensätze wurden bereits bislang von Statistik Austria zur Erstellung von (jährlichen) Statistiken verwendet, vier (Datensätze zu Kinderbetreuungsgeld, Unfallrenten, Schülerbeihilfe, Studienbeihilfe) wurden für die Erstellung der Statistik bei den zuständigen Institutionen angefordert.

Tabelle 13: Verwaltungsdatenquellen für EU-SILC 2021

Datensätze	Beschreibung
Lohnsteuerdatensatz	Enthält alle Einkommen aus unselbständiger Beschäftigung und Pensionen, unabhängig davon, ob dafür Lohnsteuer bezahlt wurde oder nicht, aber auch Informationen zum Pflegegeldbezug und zum Wochengeld. Vor der Weiterverwendung der Daten kommen umfangreiche Plausibilisierungsverfahren zur Anwendung. Die Lohnzettel werden pro Person zusammengeführt und die jeweiligen Einkommen summiert.
Pensions-Jahresdatensatz	Enthält Informationen zu allen versicherungsrechtlichen Pensionen (z.B. Alterspensionen, Invaliditäts- und Hinterbliebenenleistungen). Dieser Datensatz des Dachverbands der österreichischen Sozialversicherungsträger beinhaltet umfassende Informationen zu allen in einem Kalenderjahr aufrechten Pensionen, Neuzugängen und Abgängen und ermöglicht es, Pensionen hinsichtlich unterschiedlicher individueller Merkmale wie z.B. der Pensionsart zu klassifizieren.
DV-Qualifikationen	Enthält die sozialversicherungsrechtlichen Meldungen beim Dachverband österreichischer Sozialversicherungsträger. Dieser Datensatz enthält keine Einkommensinformation, aber vielfältige Informationen zum jeweiligen sozialversicherungsrechtlichen Status, z.B. Arbeitslosenmeldungen.
Transferdatensatz	Enthält Informationen zu Leistungen des Arbeitsmarktservice, d.h. vor allem Auszahlungen aus der Arbeitslosenversicherung wie Arbeitslosengeld oder Notstandshilfe.
Arbeitnehmerveranlagungsdatensatz	Dieser Datensatz enthält die Rückzahlungen bzw. Nachzahlungen der Lohnsteuer aus der Arbeitnehmerveranlagung.
Familienbeihilfedatensatz	Enthält die Daten der ausbezahlten Familienbeihilfe.
Studienbeihilfedatensatz	Enthält die Informationen zu den im Kalenderjahr ausbezahlten Studienbeihilfen der Studienbeihilfenbehörde. Weiters sind auch etwaige Rückzahlung an die Behörde erfasst.
Schülerbeihilfedatensatz	Enthält den zusammengeführten Datensatz der Auszahlungen der Schülerbeihilfe der österreichischen Schülerbeihilfenbehörden.
Kinderbetreuungsgeld-datensatz	Enthält die Auszahlung des Kinderbetreuungsgelds und des Familienzeitbonus, die von der Österreichischen Gesundheitskassa (ÖGK) Niederösterreich verwaltet wird. Der Datensatz enthält auch etwaige Rückzahlungen an die Behörde.
Unfallrentendatensatz	Enthält die ausbezahlten Unfallrenten und Hinterbliebenenleistungen der Unfallversicherung.

Das Haushaltseinkommen entspricht der Summe aller in Tabelle 10 und Tabelle 11 angegebenen Einkommenskomponenten im jeweiligen Haushalt (bis auf HY030, HY100, PY030, PY020). Eine detaillierte Beschreibung der Eurostat-Bezeichnungen der Einkommenskomponenten befindet sich im Eurostat-Dokument EU SILC 065 (2021 operation).¹²

¹² Grundsätzlich steht die Information auch auf der Webseite von Eurostat zur Verfügung: <http://ec.europa.eu/eurostat/web/income-and-living-conditions/methodology/>

Die im Jahr 2007 erstmals erfassten Komponenten imputierte Mieten, Zinszahlungen für Hypotheken und Arbeitgeberbeiträge zur Sozialversicherung werden entsprechend der Vorgaben von Eurostat nicht in die Berechnung des Haushaltseinkommens einbezogen.

Darstellungsmerkmal Äquivalenzeinkommen

Die meisten EU-Indikatoren, die im Rahmen von EU-SILC erhoben werden, bauen auf dem Äquivalenzeinkommen auf. Darunter zu verstehen ist ein auf einen Einpersonenhaushalt standardisiertes Haushaltseinkommen. Verwendet werden die Gewichtungsfaktoren der EU-Skala (modifizierte OECD-Skala): für die erste erwachsene Person des Haushalts wird ein Gewicht von 1,0 angenommen, für jede weitere Person ab 14 Jahren ein Gewicht von 0,5 und für Kinder von 0 bis 13 Jahren ein Gewicht von 0,3. Die Summe der Äquivalenzgewichte des Haushalts bildet den Nenner, um das Äquivalenzeinkommen der Haushaltsmitglieder aus dem Haushaltseinkommen zu berechnen. Die untenstehende Tabelle zeigt die Äquivalenzgewichte unterschiedlicher Haushaltstypen, eine alleinlebende Person wird dabei als Referenz herangezogen.

Tabelle 14: Beispiele für die Berechnung der Haushaltsäquivalente

	Fixbedarf des Haushalts	Bedarf für Erwachsene	Bedarf für Kinder	Gesamtbedarf
Einpersonenhaushalt	0,5	0,5	0,0	1,0
1 Erwachsener + 1 Kind	0,5	0,5	0,3	1,3
2 Erwachsene	0,5	1,0	0,0	1,5
2 Erwachsene + 1 Kind	0,5	1,0	0,3	1,8
2 Erwachsene + 2 Kinder	0,5	1,0	0,6	2,1
2 Erwachsene + 3 Kinder	0,5	1,0	0,9	2,4

Q: STATISTIK AUSTRIA, EU-SILC 2021. – Kind = Personen unter 14 Jahren.

Das Äquivalenzeinkommen, auch „äquivalisiertes Haushaltseinkommen“ genannt, wird berechnet, indem man das verfügbare Nettojahreshaushaltseinkommen durch die Summe der Gewichte des Haushalts 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 der Armutsgefährdungsschwelle

Der Betrag des äquivalisierten Haushaltseinkommens, der die Grenze für Armutsgefährdung bildet. Bei äquivalisierten Haushaltseinkommen unter diesem Schwellenwert wird Armutsgefährdung angenommen. Wenn nicht anders ausgewiesen, handelt es sich um die normative Festlegung der Armutsgefährdungsschwelle nach EU-Definition bei 60 % des Medians des äquivalisierten Haushaltseinkommens (verwendet werden für analytische Zwecke auch Schwellen von 40 %, 50 % und 70 % des Medians). Der Median teilt die Verteilung der Einkommen in genau zwei Hälften, sodass 50 %

der Personen ein Einkommen unter und 50 % der Personen ein Einkommen über diesem Wert zur Verfügung haben. Der Median des äquivalisierten Haushaltseinkommens beträgt 2021 27 428 Euro im Jahr oder 2 286 Euro im Monat. Der Betrag der Armutsgefährdungsschwelle liegt 2021 bei einem äquivalisierten Haushaltseinkommen von 1 371 Euro pro Monat (12 Mal für einen Einpersonenhaushalt gerechnet) oder bei 16 457 Euro pro Jahr.

Definition von Armutsgefährdung (Einkommensarmut)

Alle Personen, deren äquivalisiertes Haushaltseinkommen unterhalb eines festgelegten Schwellenwertes (Armutsgefährdungsschwelle = 60 % des Medians) liegt, gelten als armutsgefährdet.

Definition der Armutsgefährdungslücke

Die Armutsgefährdungslücke ist ein Maß für die Intensität der Armutsgefährdung und ist definiert als Median der individuellen relativen Abweichungen der Äquivalenzeinkommen der Armutsgefährdeten von der Armutsgefährdungsschwelle in Prozent dieser Schwelle. 2021 beträgt die Armutsgefährdungslücke 22,2 %. Dies bedeutet, dass die Äquivalenzeinkommen der Armutsgefährdeten im Mittel um 22,2 % unter der Armutsgefährdungsschwelle liegen. In absoluten Zahlen beläuft sich die Armutsgefährdungslücke auf 3 656 Euro pro Jahr bzw. 305 Euro pro Monat.

Dauerhafte Armutsgefährdung

Armutsgefährdung am Ende des Beobachtungszeitraumes, das heißt im letzten Jahr der Erhebung, und in mindestens zwei von drei vorhergehenden Jahren. Die dauerhafte Armutsgefährdung kann per Definition nur von Personen, die vier Jahre in Folge befragt wurden, ermittelt werden.

Materielle und soziale Deprivation

Von erheblicher materieller Deprivation sind Personen betroffen, die in Haushalten leben, auf die fünf oder mehr der folgenden 13 Merkmale zutreffen:

Für den Haushalt ist Folgendes finanziell nicht leistbar:

- Unerwartete Ausgaben in der Höhe von 1 290 Euro zu tätigen
- Einmal im Jahr auf Urlaub zu fahren
- Miete, Betriebskosten oder Kredite pünktlich zu bezahlen
- Jeden zweiten Tag Fleisch, Fisch oder eine vergleichbare vegetarische Speise zu essen
- Die Wohnung angemessen warm zu halten
- Abgenützte Möbel zu ersetzen
- Ein Auto zu besitzen

Personen ab 16 Jahren ist Folgendes finanziell nicht möglich:

- Eine zufriedenstellende Internetverbindung zu haben
- Abgenützte Kleidung zu ersetzen
- Zwei Paar passende Schuhe zu besitzen

- Jede Woche einen kleinen Betrag für sich selbst auszugeben
- Regelmäßig kostenpflichtige Freizeitaktivitäten auszuüben
- Einmal im Monat Freund:innen oder Familie zum Essen/Trinken zu treffen

Haushalt mit keiner oder sehr niedriger Erwerbsintensität

Dies sind jene Haushalte, in denen Personen im Erwerbsalter (zwischen 18 und 59 Jahren, außer Studierende) in geringem Ausmaß erwerbstätig sind (weniger als 20 % ihres Erwerbspotentials innerhalb eines Jahres ausschöpfen). Als „betroffen“ werden hier nur Personen unter 60 Jahren gewertet.

Armuts- oder Ausgrenzungsgefährdung

Als armuts- oder ausgrenzungsgefährdet werden Personen dann gezählt, wenn sie entweder von Armutsgefährdung oder von materieller Deprivation betroffen sind, oder in einem Haushalt mit keiner oder sehr niedriger Erwerbsintensität leben.

Weitere Definitionen finden sich im Tabellenband von EU-SILC.

2.1.11 Verwendete Klassifikationen

Die Adressen im Haushaltsregister waren gemäß EU-Vorgaben regional nach der Besiedlungsdichte¹³ sowie nach NUTS2 (in Österreich identisch mit den Bundesländern)¹⁴ zu klassifizieren.

In der Personenbefragung wurden die berufliche Tätigkeit nach der Berufssystematik ISCO-88 (2-Steller)¹⁵ und der Wirtschaftszweig der Betriebsstätte bzw. Dienststelle nach der NACE (2-Steller)¹⁶

¹³ Diese Klassifikation (DEGURBA – Degree of Urbanisation) klassifiziert alle Gemeinden nach ihrer Besiedlungsdichte: hohe, mittlere und geringe Besiedlungsdichte. Die Definition dieser Klassifikation wurde im Jahr 2011 revidiert.

Hohe Besiedlungsdichte: eine Gruppe aneinandergrenzender Gemeinden mit jeweils mehr als 500 Einwohner:innen pro km² und insgesamt mindestens 50.000 Einwohner:innen.

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

Geringe Besiedlungsdichte: alle übrigen Gemeinden.

Informationen zur Revision sind auf folgender Webseite verfügbar: http://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA (Stand Mai 2022).

¹⁴ Vgl. <https://www.statistik.at/services/tools/services/regionales/regionale-gliederungen> (Stand August 2022).

¹⁵ Vgl. in der Klassifikationsdatenbank von Statistik Austria http://www.statistik.at/KDBWeb/kdb_Einstieg.do (Stand Mai 2022).

¹⁶ Vgl. in der Klassifikationsdatenbank von Statistik Austria http://www.statistik.at/KDBWeb/kdb_Einstieg.do (Stand Mai 2022) und <https://statistik.at/ueber-uns/erhebungen/unternehmen/oenace-2008> (Stand August 2022).

vercodet. Die Bildungsebene wurde analog zum Mikrozensus-Merkmal „höchste abgeschlossene Schulbildung“ erhoben und anschließend nach der ISCED (1-Steller)¹⁷ vercodet. Die berufliche Funktion war erhebungsseitig ebenfalls an ein Mikrozensus-Merkmal angelehnt, nämlich an die „Stellung im Beruf“.

2.1.12 Regionale Gliederung

Die veröffentlichten Indikatoren EU-SILC 2021¹⁸ werden nach Gemeindegrößenklassen differenziert.¹⁹

EU-SILC weist auch ausgewählte Ergebnisse nach Bundesländern (NUTS2) aus – seit dem Tabellenband 2018 beruhen diese auf Dreijahresdurchschnitten, da die Stichprobe für EU-SILC repräsentativ für Österreich gezogen wurde und daher Auswertungen auf Bundesländerebene mit einer höheren Schwankungsbreite der Ergebnisse behaftet sind.

2.2 Erstellung der Statistik, Datenaufarbeitung, qualitätssichernde Maßnahmen

2.2.1 Datenerfassung

Die im Zuge der CAPI-Interviews erhobenen Daten wurden durch die Erhebungspersonen mittels elektronischer Datenübertragung an Statistik Austria übermittelt. Die Übertragung der Daten aus den CAPI-Fragebögen erfolgt ebenso elektronisch, durch die CAPI-Erhebungspersonen, die die Datensätze an Statistik Austria übermitteln.

2.2.2 Signierung (Codierung)

Freitexteinträge werden automatisch in so genannten Look-up-Listen den entsprechenden Codes und Zuordnungslisten zugeordnet.

2.2.3 Plausibilitätsprüfung, Prüfung der verwendeten Datenquellen

Es zählt zu den großen Vorteilen einer computerunterstützten Erhebung, dass Erfassungsfehler (z.B. bei Angaben zum Betrag eine Null zu viel oder zu wenig) sowie technisch unplausible Antwortmuster (z.B. wechselseitig nicht übereinstimmende Angaben zur Verwandtschaft) zum Teil bereits während der

¹⁷ Vgl. http://www.statistik.at/web_de/klassifikationen/klassifikationsdatenbank/weitere_klassifikationen/bildungsklassifikation/index.html (Stand August 2022).

¹⁸ Vgl. https://www.statistik.at/fileadmin/pages/338/Tabellenband_EUSILC_2021.pdf, S. 175 ff. (Stand August 2022).

¹⁹ Dabei werden folgende Ausprägungen unterschieden: Wien, andere Gemeinden mit über 100.000 Einwohnerinnen und Einwohnern, Gemeinden mit über 10.000 und max. 100.000 Einwohner:innen, sonstige Gemeinden (d.s. alle Gemeinden mit maximal 10.000 Einwohner:innen).

Befragung korrigiert werden können. Dies ist möglich durch die elektronische Datenerfassung und die integrierte Durchführung von Haushalts- und Personeninterviews. Mehr als 300 derartige Checks wurden von Statistik Austria entwickelt und im Zuge der Feldarbeit implementiert.

Zu diesen Plausibilitätschecks im CAPI/CATI-Programm gehören so genannte Warnings und Errors. Die beiden Begriffe bezeichnen zwei Arten von Fehlern. Errors treten auf, wenn die gemachten Angaben realiter nicht möglich sind – hier ist eine Eingabe der fehlerhaften Angabe überhaupt nicht möglich, sie muss zwingend korrigiert werden. Warnings treten auf, wenn die gemachte Angabe zwar nicht gänzlich unmöglich, aber doch recht unwahrscheinlich ist (etwa der Bezug einer bedarfsorientierten Mindestsicherung von 3 000 Euro im Monat) – hier ist die entsprechende Eingabe möglich, die Interviewerin bzw. der Interviewer wird aber darauf hingewiesen, dass die Richtigkeit der Angabe eher unwahrscheinlich ist.

In der Vorplausibilisierung 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.

Die Mikroplausibilisierung erfolgt durch SPSS-Programme, die alljährlich angepasst werden. Dabei werden in standardisierten Verfahren die Datensätze kontrolliert und bearbeitet und Probleme bei Einzelfällen identifiziert. Soweit möglich, werden auch bei der Plausibilisierung standardisierte Verfahren verwendet, da die Prüfung auf Einzeldatenebene sehr aufwendig ist. Die verbleibenden Einzelfälle von unplausiblen und fehlenden Antworten bei Einkommensfragen werden mithilfe der offenen Anmerkungsfelder geprüft und die Bearbeitung in die Programmierung einbezogen und dokumentiert.

Fragwürdige Antworten können von der Respondentin bzw. dem Respondenten selbst stammen oder durch Fehleingaben durch die Erhebungspersonen entstehen. Diese werden mittels Nachrecherchen oder Plausibilisierung korrigiert oder als fehlende Werte definiert. 2021 wurden sowohl Querschnittchecks als auch Längsschnittchecks, die auf einem Vergleich der aktuellen mit den Vorjahresdaten (und falls vorhanden auch älteren Daten von 2018 und 2019) basieren, durchgeführt.

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

Nach dem Vorliegen des authentischen internen Datenbestandes können die Eurostat-Zielvariablen gebildet werden. Um diese zu überprüfen, wurden von Eurostat eigene SAS-Programme entwickelt, die auch den Mitgliedsstaaten zur Verfügung gestellt wurden.

2.2.4 Imputation (bei Antwortausfällen bzw. unvollständigen Datenbeständen)

Der folgende Abschnitt beschreibt die in EU-SILC angewandten Verfahren der Imputation. Diese basieren weitgehend auf den bereits in den Vorjahren entwickelten Prozeduren.²⁰

Imputation meint alle Verfahren, um fehlende Werte zu schätzen und zu ersetzen. Im Allgemeinen sind dies Schätzverfahren, um entweder fehlende Informationen von ganzen Erhebungseinheiten (unit nonresponse) oder einzelne Werte von Erhebungseinheiten (item nonresponse) zu ermitteln. Unit Nonresponse bezieht sich hier nur auf den Ausfall von Personeninterviews in interviewten Haushalten (within unit nonresponse), nicht auf gesamte nichtinterviewte Haushalte.

Das Hauptinteresse der Erhebung EU-SILC liegt in der Erfassung des Haushaltseinkommens der befragten Haushalte. Bei fehlenden Werten auch nur einer einkommensbezogenen Variable 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 Nonresponse

In EU-SILC werden fehlende Interviews von Personen, die aufgrund von vorübergehender Abwesenheit, Krankheit, sprachlichen Problemen oder aus anderen Gründen²¹ nicht interviewt werden konnten, ersetzt. Dabei wird mit einer Distanzfunktion ein ähnlicher Fall gesucht, dessen Interviewergebnis auf den Fall mit den fehlenden Werten übertragen wird (Hot-Deck-Verfahren). Die Distanzfunktion verwendet zur Bestimmung der Ähnlichkeit bestimmte Variablen (siehe Tabelle 15). 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.

Insgesamt wurden in EU-SILC 2021 50 von insgesamt 10 409 Personeninterviews vollständig imputiert (vgl.). Für 29 Personen standen Informationen aus der vorangegangenen Erhebung zur Verfügung, für 21 Personen waren nur Informationen aus dem Personenregister verfügbar.

Je nachdem, welche Informationen über fehlende Fälle zur Verfügung standen, wurden unterschiedliche Variablen in die Berechnung der Distanzfunktion einbezogen. Die folgende Tabelle gibt einen Überblick über die berücksichtigten Variablen:

²⁰ Die Vorgehensweise bei Imputationen im Datensatz von EU-SILC in Österreich ist genauer beschrieben in Heuberger/Lamei 2006, und in BMASK 2010.

²¹ Aufgrund einer Verweigerung nur in Ausnahmefällen

Tabelle 15: Variablen für die Distanzfunktion imputierter Personeninterviews

Imputation fehlender Personeninterviews	
Information aus dem Vorjahr (N = 29)	Information aus dem Register (N = 21)
Geschlecht	Geschlecht
Alter	Alter
Gegenwärtige Beschäftigungssituation	Bundesland
Bundesland	Haushaltsgröße
Haushaltsgröße	Anzahl Personen jünger 18 im Haushalt
Anzahl Personen jünger 18 im Haushalt	Anzahl Personen älter 60 im Haushalt
Anzahl Personen älter 60 im Haushalt	Haushaltseinkommen
Höchste abgeschlossene Schulausbildung	
Beeinträchtigung durch Krankheit/Behinderung	
Haushaltseinkommen	
Anzahl der Monate in Beschäftigung	
Anzahl der Monate mit selbständiger Tätigkeit	

Q: STATISTIK AUSTRIA, EU-SILC 2021.

Verfahren für den Umgang mit Item Nonresponse

Item Nonresponse bezeichnet das Fehlen von Werten in den entsprechenden Variablen. Diese fehlenden Werte entstehen durch die Verweigerung einer Antwort oder durch Nicht-Wissen seitens der Respondent:innen, durch Eingabefehler von Interviewer:innen, durch Fehler beim Editieren des Datensatzes oder das Nichtvorhandensein in einem Verwaltungsdatensatz. Den meisten Ursachen von Item Nonresponse kann durch geeignete Maßnahmen (Gestaltung der Fragen, Schulung der Interviewer:innen etc.) wirksam entgegengesteuert werden, dennoch ist ein gewisser Anteil fehlender Werte unvermeidbar.

In EU-SILC werden grundsätzlich fehlende Angaben zu Einkommensfragen durch geeignete Imputationsverfahren ersetzt. Durch die weitgehende Verwendung von Verwaltungsdaten für die Berechnung der Einkommensvariablen hat sich der Anteil der fehlenden Werte deutlich reduziert: Sind Werte in Verwaltungsdaten vorhanden, kann es keine fehlenden Werte geben (außer durch Fehler beim Editieren der Datensätze). Item Nonresponse kann somit nur dann entstehen, wenn davon ausgegangen werden muss, dass es einen entsprechenden Wert in Verwaltungsdaten geben müsste, der betreffende Fall aber nicht mit dem jeweiligen Verwaltungsdatensatz verknüpft werden konnte. D.h. der Grund für Item Nonresponse bei der Verwendung von Verwaltungsdaten liegt in fehlenden Verknüpfungen von Befragungs- und Verwaltungsdaten, oder genauer ausgedrückt, im Fehlen des Verknüpfungsschlüssels in einem der beiden Datensätze (fehlendes bereichsspezifisches Personenkennzeichen, kurz bPK). Leider

ist es aber grundsätzlich nicht möglich, zu unterscheiden, ob eine Person nicht im Verwaltungsdatensatz vorkommt, weil eine Verknüpfung fehlt, oder weil diese Person tatsächlich über kein entsprechendes Einkommen verfügt. Daher ist es bei der Verwendung von Verwaltungsdaten nur in einer eingeschränkten Anzahl von Fällen möglich, Item Nonresponse festzustellen (etwa bei Grenzgänger:innen, die per Definition nicht im Lohnsteuerdatensatz auffindbar sein können, aber dennoch über ein unselbständiges Einkommen verfügen). Für Einkommen, die weiterhin im Fragebogen erfragt werden, kommt es nach wie vor zu fehlenden Werten (etwa beim Einkommen aus selbständiger Erwerbstätigkeit). Die beschriebenen Imputationsverfahren bzw. der Umgang mit fehlenden Einkommenswerten in EU-SILC bezieht sich somit vorrangig auf jene Einkommen, die in der Erhebung erfragt werden.

In der Erhebung EU-SILC werden nur Netto-Einkommensvariablen imputiert. Fehlende Bruttowerte werden mittels Netto-Brutto-Konversion berechnet.

Fehlende Werte bei der Befragung von Einkommensvariablen in 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 versucht, diese Angabe aus anderen Variablen abzuleiten. Dabei wird insbesondere auf den Aktivitätenkalender (Hauptbeschäftigung in den Monaten des Vorjahres) zurückgegriffen. Kann die Information, ob eine Einkommenskomponente bezogen worden ist oder nicht, nicht abgeleitet werden, wird davon ausgegangen, dass diese Einkommenskomponente nicht bezogen wurde.

Ähnlich behandelt werden fehlende Angaben über die Anzahl der Monate, in denen eine Einkommenskomponente bezogen wurde. Zunächst wird versucht, die Angabe über die Monatsanzahl aus anderen Variablen des Datensatzes herzuleiten. Gelingt dies nicht, wird ausgehend von der empirischen Verteilung der Monatsangaben 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 die/der Befragte nennt sowohl den Bruttobetrag als auch den Nettobetrag der Einkommenskomponente, oder es wird die Angabe nur zu entweder dem Brutto- oder dem Nettobetrag gemacht, oder die/der Befragte gibt eine Einkommensstufe an. Die Möglichkeit der Auswahl einer Einkommensstufe – anstelle der Angabe eines konkreten Betrages – ist 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 auf der Grundlage der empirischen Verteilung der Einkommenswerte ein Schätzwert ermittelt.

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 mit Hilfe dieser Methoden kein Wert ermittelt werden, so stehen zwei Ansätze zur Berechnung eines Schätzwertes zur Verfügung: Längsschnitt- und Querschnittverfahren. Das Längsschnittverfahren kann dann verwendet werden, wenn die Person mit fehlenden Werten in den Vorjahren einen Wert für diese Variable angegeben hat. Für alle anderen Fälle können nur Querschnittverfahren verwendet werden.

Das Verfahren zur Längsschnittimputation basiert auf der Row-and-Column-Methode von Little und Su²². 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 Querschnittverfahren wurden in EU-SILC 2021 lineare Regressionsmodelle zur Berechnung von Schätzwerten verwendet.²³ Für jede Einkommenskomponente wurden dabei mehrere Modelle spezifiziert, um sicherzustellen, dass auch beim Vorliegen von fehlenden Werten in den Prädiktorvariablen Schätzwerte berechnet werden können.

Würde man den aus der linearen Regression ermittelten Schätzwert als Imputationswert verwenden, so würde sich die Varianz der Variablen reduzieren und die durch die Regressionsgleichung modellierten Zusammenhänge verstärken. Um diese Effekte möglichst gering zu halten, wurde den mittels linearer Regression ermittelten Werten ein stochastischer Störterm beigefügt, d.h. zum ermittelten erwarteten Wert wurde ein fiktives Residuum addiert. Dieser Störterm entsprach in seiner Verteilung der Verteilung der geschätzten Residuen der erwarteten Werte. Dies führte dazu, dass der Varianzreduktion entgegengewirkt werden konnte.

Die Prädiktorvariablen wurden nach ihrer Vorhersagekraft ausgewählt (also nach der Signifikanz der Regressionskoeffizienten), und nach der inhaltlichen Plausibilität der Prädiktoren. Konnte für eine

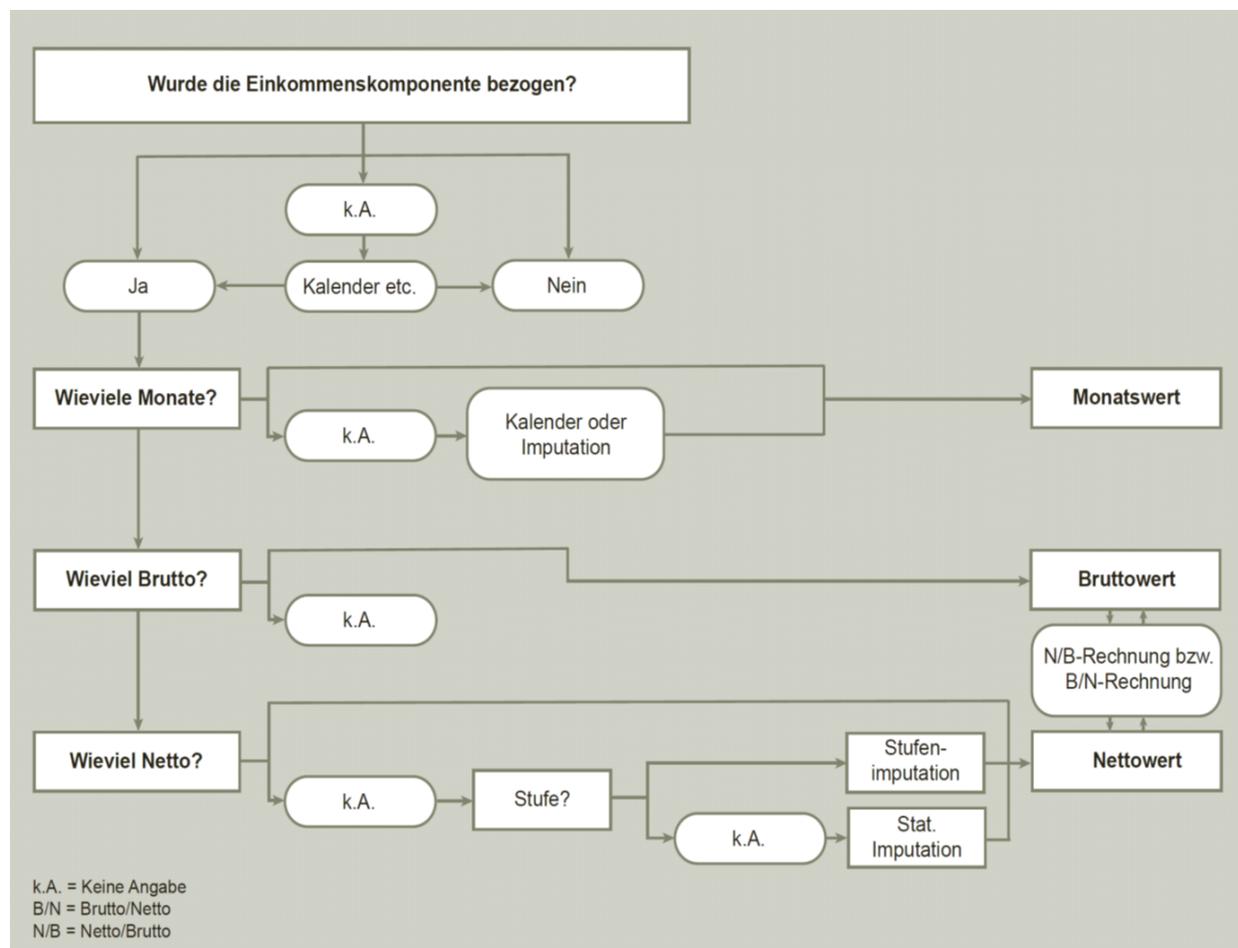
²² Little, Roderick J.A. / Su, Hong-Jin (1989). Item Non-response in Panel Surveys. In: Kasprzyk, D./Duncan, G./Kalton, G./Singh, M. Panel Surveys. New York; Wiley, S.400-425.

²³ Die Regressionskoeffizienten wurden dabei mit dem OLS-Verfahren ermittelt.

Variable kein angemessenes Regressionsmodell spezifiziert werden, so wurde ein Schätzwert auf der Basis des Mittelwerts bzw. des Medians berechnet, zu welchem wiederum ein stochastischer Störterm addiert wurde.

Abbildung 5 fasst den Umgang mit fehlender Information bei Einkommensvariablen zusammen.

Abbildung 5: Vorgehen bei fehlender Einkommensinformation bei Fragebogenvariablen in EU-SILC



Q: STATISTIK AUSTRIA.

Tabelle 16 zeigt die Anzahl der zu imputierenden Werte für ausgewählte Einkommenskomponenten. In EU-SILC 2021 mussten wie in den letzten Jahren nur wenige Werte imputiert werden. Der Anteil der fehlenden Werte ist bei Pensionen und Unselbständigeneinkommen mit rund 1 % am niedrigsten, beim Selbständigeneinkommen etwas mehr als 17 % am höchsten. Bis auf selbständige Einkommen aus einem land- oder forstwirtschaftlichen Betrieb werden in EU-SILC 2021 alle der in dieser Tabelle dargestellten Variablen durch Verwaltungsdaten befüllt.

Tabelle 16: Imputationen ausgewählter Nettoeinkommenskomponenten

	Gesamt	Betrag laut Befragung	Betrag aus Stufe	Betrag aus Brutto/Netto-Rechnung	Längsschnitt-imputation	Querschnitt-Imputation	Sonstige Korrekturen
Absolut							
Unselbständigeneinkommen ¹	5 625	5 570	0	0	7	48	0
Selbständigeneinkommen aus land-/forstwirtschaftlichem Betrieb	211	174	19	0	3	15	0
Arbeitslosengeld ¹	765	729	0	0	0	36	0
Alterspension ¹	3 014	2 975	0	0	0	39	0
Krankengeld ¹	250	220	0	0	3	27	0
Relativ in %							
Unselbständigeneinkommen	100,0	99,0	0,0	0,0	0,1	0,9	0,0
Selbständigeneinkommen aus land-/forstwirtschaftlichem Betrieb	100,0	82,5	9,0	0,0	1,4	7,1	0,0
Arbeitslosengeld	100,0	95,3	0,0	0,0	0,0	4,7	0,0
Alterspension	100,0	98,7	0,0	0,0	0,0	1,3	0,0
Krankengeld	100,0	88,0	0,0	0,0	1,2	10,8	0,0

Q: STATISTIK AUSTRIA, EU-SILC 2021.

1) Bei Einkommen aus Verwaltungsdaten ist eine Imputation aus einer Stufenangabe nicht möglich.

2.2.5 Hochrechnung (Gewichtung)

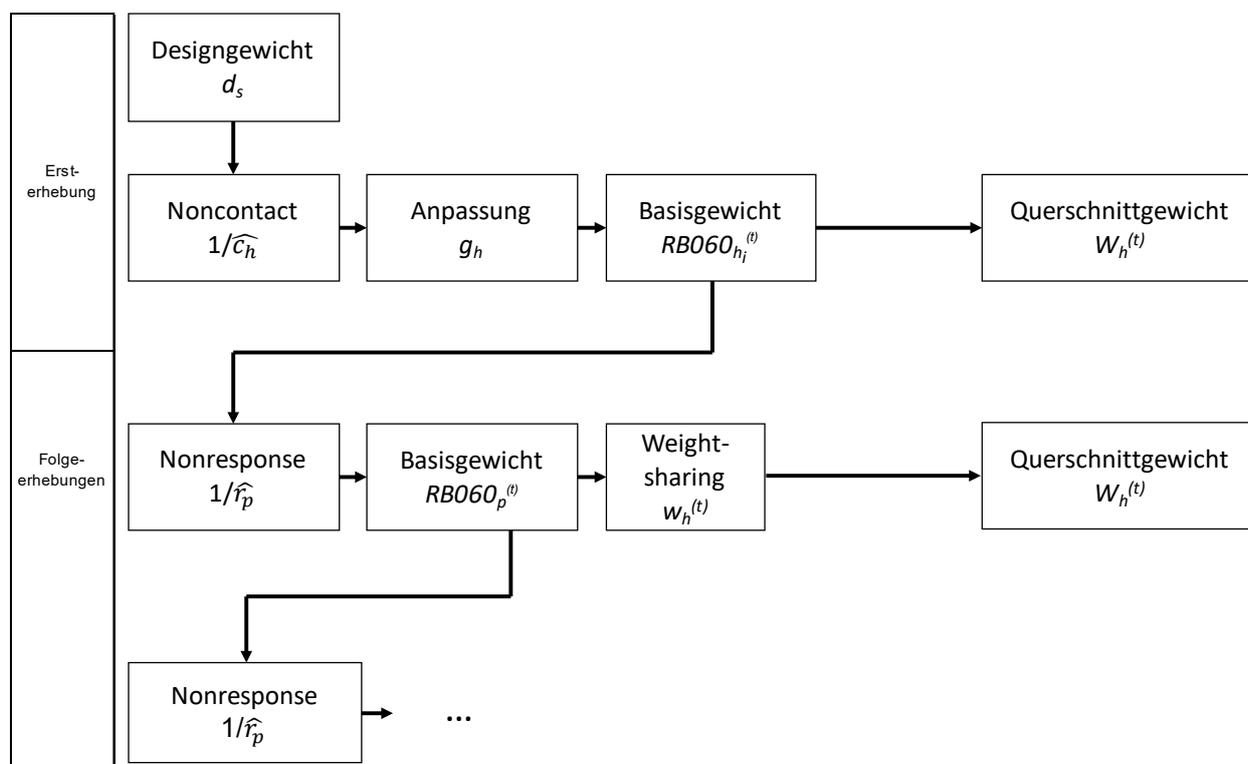
2021 war das 17. Erhebungsjahr, in dem das integrierte Quer- und Längsschnittdesign in EU-SILC implementiert wurde. In ihren Grundzügen erfolgte die Gewichtung nach der bereits seit EU-SILC 2006 angewendeten Methodik. Auf Basis der Empfehlungen von Eurostat wurde die Stichprobe getrennt nach den vier Rotationen gewichtet, wobei jeweils nach dem Jahr der Ersterhebung (2018, 2019, 2020 oder 2021) unterschieden wurde.²⁴

²⁴ Vgl. Eurostat (2021), S. 34ff. Eine detaillierte Beschreibung des EU-SILC Gewichtungsverfahrens 2008 (welches auch die Grundlage für die Hochrechnung von EU-SILC 2021 ist) befindet sich in Glaser/Till 2010.

2.2.5.1 Querschnittgewichtung 2021

Die Durchführung der Querschnittgewichtung gliedert sich im Wesentlichen in fünf Bearbeitungsschritte: Designgewicht (1), Unit Nonresponse Gewicht (2), Basisgewicht (3), Weight-Sharing (4) und Kombination der Teilstichproben inkl. Anpassungsgewicht (5). Beginnend mit dem Designgewicht, welches nur für die Erstbefragung erforderlich ist, werden die Schritte (1) bis (4) auf jede der vier Rotationsgruppen separat angewendet, Schritt (5) betrifft schließlich den gesamten Querschnitt.²⁵ Abbildung 6 stellt den Ablauf des Gewichtungsverfahrens überblicksmäßig dar.

Abbildung 6: Gewichtungsverfahren EU-SILC



Q: STATISTIK AUSTRIA.

Die in Abbildung 6 dargestellten Gewichte bzw. Anpassungsfaktoren beziehen sich auf unterschiedliche Aggregate der Stichprobe. D.h. sie variieren für unterschiedliche Stichproben-Strata bzw. Interviewsprengel (vgl. Kapitel 2.1.6), Haushalte oder Personen. Tabelle 17 beschreibt die wichtigsten verwendeten Formelzeichen und Indizes, welche auch in Abbildung 6 enthalten sind.

²⁵ Jedes Jahr tritt eine neue Rotationsgruppe das erste Mal in das Panel ein und eine Rotationsgruppe, die in den vorangegangenen Jahren schon viermal befragt wurde, verlässt das Panel, vgl. Kapitel 2.1.5.

Tabelle 17: Formelzeichen und Indizes des Gewichtungsverfahrens

Formelzeichen	Bezeichnung	Index	Index-Bezug
d_s	Designgewicht	$s \in \{1, \dots, K\}$	Strata s aus Ziehung der Erststichprobe
$\frac{1}{\hat{c}_h}$	Noncontact-Gewicht	$h \in \{1, \dots, H^{(1)}\}$	Haushalte h , die an der Erstbefragung teilnehmen ($t=1$)
$g_{h^{(1)}}$	Anpassungsfaktor Kalibrierung Erstbefragung	$h \in \{1, \dots, H^{(1)}\}$	Haushalte h , die an der Erstbefragung teilnehmen
g_h	Anpassungsfaktor Kalibrierung Querschnitt	$h \in \{1, \dots, H\}$	Haushalte h , die an der Befragung
$RB060_{h_i}^{(1)}$	Basisgewicht Erstbefragung	$i \in \{1, \dots, m_h\},$ $h \in \{1, \dots, H^{(1)}\}$	Personen i in an der Erstbefragung teilnehmenden Haushalten h ($t=1$)
$W_h^{(1)}$	Haushaltsgewicht Erstbefragung	$h \in \{1, \dots, H^{(1)}\}$	Haushalte h , die an der Erstbefragung teilnehmen ($t=1$)
$\frac{1}{\hat{r}_p}$	Nonresponse-Gewicht	$p \in \{1, \dots, P^{(t)}\}$	Personen p , die an der Befragung im Jahr t teilnehmen ($t \in \{2,3,4\}$)
$RB060_p^{(t)}$	Basisgewicht Folgebefragung	$p \in \{1, \dots, P^{(t)}\}$	Personen p , die an der Befragung im Jahr t teilnehmen ($t \in \{2,3,4\}$)
$W_{h_i}^{(t)}$	Gewicht nach weight-sharing	$h \in \{1, \dots, H^{(t)}\}$	Personen i in an einer Folgebefragung teilnehmenden Haushalten h ($t \in \{2,3,4\}$)
$hgew_{h_i}$	Haushaltsgewicht Querschnitt	$i \in \{1, \dots, m_h\},$ $h \in \{1, \dots, H\}$	Personen i in an der Befragung teilnehmenden Haushalten h

Jede der vier Rotationsgruppen stellt gemäß dem Gewichtungsverfahren ein Abbild der Grundgesamtheit dar, welches die Erstellung von repräsentativen Statistiken ermöglicht. Deshalb ist es für den Abschluss des Gewichtungsverfahrens des Querschnitts notwendig, die Gewichte der vier Teilstichproben so zu kombinieren, dass mit Hilfe des endgültigen Querschnittsgewichts, welches sich auf alle Rotationsgruppen bezieht, repräsentative Statistiken erstellt werden können. Es folgt nun eine detaillierte Beschreibung der einzelnen Gewichtungsschritte.

Schritt 1: Designgewicht

Am Beginn des Gewichtungsverfahrens wurde das Design der Erstbefragungsstichprobe der beschriebenen Stichprobenauswahl berücksichtigt. Da das Stichprobendesign aufgrund der disproportionalen Allokation pro Bundesland unterschiedliche Auswahlwahrscheinlichkeiten beinhaltet, müssen die so entstandenen Unterschiede in der Anzahl der gezogenen Haushalte im Zuge der Gewichtung kompensiert werden. Dieses so genannte „Designgewicht“ beachtet die unterschiedlichen Wahrscheinlichkeiten eines Haushalts, in die Stichprobe aufgenommen zu werden. Um einen Ausgleich zwischen hohen und niedrigen Auswahlwahrscheinlichkeiten zu schaffen, wurde das Designgewicht als Inverse der Auswahlwahrscheinlichkeiten berechnet (siehe Formel (1)). Da die Auswahlwahrscheinlichkeit p_s innerhalb jedes der insgesamt $K=18$ Strata (Interviewsprengel) gleich ist, ist auch das Designgewicht d_s innerhalb jedes Stratums konstant.

$$d_s = \frac{1}{p_s} \quad s \in \{1, \dots, K\} \quad (1)$$

K ... Anzahl der Strata in der Erstbefragungsstichprobe

Jene Adressen, deren Auswahlwahrscheinlichkeit erhöht war, erhielten ein entsprechend reduziertes Designgewicht. Diese Berechnungsart ermöglicht eine, was den Auswahlrahmen betrifft, unverzerrte Hochrechnung mit Hilfe der Designgewichte.²⁶

Schritt 2: Nonresponse

Die Nonresponse-Gewichtung stellt den zweiten Schritt in dem hier beschriebenen Gewichtsungsverfahren dar. Durch die Nicht-Teilnahme von Personen in bestimmten Haushaltsgruppen an der Befragung kann es dazu kommen, dass bestimmte, für die Erhebung wichtige, Eigenschaften über- bzw. unterrepräsentiert sind. Dies führt mitunter zu verzerrten Schätzungen eines Erhebungsmerkmals. Die Nonresponse-Gewichtung soll diesem selektiven Ausfall entgegenwirken. Um vermutete Verzerrungen aufgrund von Ausfällen korrekt vorherzusagen, wäre die Kenntnis der Antwortwahrscheinlichkeiten notwendig. Da diese in der Regel nicht bekannt ist, muss die Antwortwahrscheinlichkeit unter Verwendung bekannter Variablen x_j , wie z.B. Haushaltseinkommen oder berufliche Stellung, geschätzt werden.

Dabei muss hier zwischen der Unit Nonresponse im Rahmen der Erstbefragung und in den Folgebefragungswellen unterschieden werden.

Vorangegangene Simulationen für die Erstbefragung haben gezeigt, dass alle Aspekte der Unit Nonresponse für die erste Erhebungswelle nur schwer zu modellieren sind. Erfolgt jedoch eine Aufteilung der Unit Nonresponse in nicht erfolgten Kontakt (Noncontact) sowie aktive Verweigerung, zeigt sich ein differenzierteres Bild. Auf Basis verschiedener Modellierungen und Simulationen konnte gezeigt werden, dass der Teil der Unit Nonresponse, welcher den nicht erfolgten Kontakt betrifft, gut mittels logistischer Regressionsmodelle geschätzt werden kann. Für die aktive Verweigerung nach erfolgtem Kontakt in der ersten Erhebungswelle ist es aber effizienter diese Form des mitunter selektiven Antwortausfalls im Rahmen der Anpassungsgewichte (Kalibrierung) zu kompensieren²⁷.

Die Gewichtung für Noncontact stellt den zweiten Schritt in dem hier beschriebenen Gewichtsungsverfahren dar. Durch die Nicht-Teilnahme von Personen in bestimmten Haushaltsgruppen an der Befragung kann es dazu kommen, dass bestimmte, für die Erhebung wichtige, Eigenschaften über- bzw. unterrepräsentiert sind. Dies führt mitunter zu verzerrten Schätzungen eines Erhebungsmerkmals. Die

²⁶ Es handelt sich dabei um den Horvitz-Thompson Schätzer. Vgl. Horvitz/Thompson 1952.

²⁷ Im Zuge des Eurostat Grants „Improving Methodology on Sampling, Weighting, Imputation and Variance Estimation in the Austrian EU-SILC with regard to Administrative Data“ konnte gezeigt werden, dass die logistischen Regressionsmodelle zur Schätzung der Kontaktwahrscheinlichkeit einen vergleichsweise besseren Modellfit aufweisen (Pseudo-R² ≈ 12%) als zur Schätzung der gesamten Unit Nonresponse bzw. der Verweigerung nach erfolgtem Kontakt (jeweils Pseudo-R² ≈ 3%).

Noncontact-Gewichtung soll diesem selektiven Ausfall entgegenwirken. Um vermutete Verzerrungen aufgrund von Ausfällen korrekt vorherzusagen, wäre die Kenntnis der Kontaktwahrscheinlichkeiten notwendig. Da diese in der Regel nicht bekannt ist, muss die Kontaktwahrscheinlichkeit unter Verwendung bekannter Variablen x_j , wie z.B. Haushaltseinkommen oder berufliche Stellung, geschätzt werden. Dies geschah bei EU-SILC 2021 mit Hilfe eines logistischen Regressionsmodells, indem die Wahrscheinlichkeit \widehat{c}_h zur Gruppe der Kontaktierten (*Contact*) zu gehören, mit Hilfe der erklärenden Variablen x_j geschätzt wurde (Formel (2)).

$$\widehat{c}_h = P(\text{Contact} = 1 | x_1, \dots, x_J) = \frac{\exp(\widehat{\beta}_0 + \widehat{\beta}_1 X_1 + \dots + \widehat{\beta}_J X_J)}{1 + \exp(\widehat{\beta}_0 + \widehat{\beta}_1 X_1 + \dots + \widehat{\beta}_J X_J)} \quad (2)$$

Bei der Auswahl der erklärenden Variablen musste zwischen den Modellen zur Schätzung der Kontaktwahrscheinlichkeit in den Erstbefragungshaushalten 2021 und der Antwortwahrscheinlichkeit in den Folgebefragungshaushalten mit den Ersterhebungsjahren 2018, 2019 und 2020 unterschieden werden. In der Ersterhebung war nur wenig über die nicht an der Befragung teilnehmenden Haushalte bekannt. Deshalb mussten hier für die Schätzung der Kontaktwahrscheinlichkeiten der Haushalte hauptsächlich aus der Stichprobenziehung zur Verfügung stehende Variablen verwendet werden: Bundesland, Besiedlungsdichte, Anzahl der Personen mit nicht-österreichischer Staatsbürgerschaft, Anzahl der Kinder im Haushalt, Anzahl der Frauen im Haushalt, Anzahl der Männer im Haushalt, Alter der ältesten bzw. jüngsten Person im Haushalt gemäß ZMR. Zusätzlich wurde die Summe aus Unselbständigen- und Pensionseinkommen pro Haushalt (in Dezilen) sowie die aus dem Stichprobenrahmen geschätzte Armut- oder Ausgrenzungsgefährdung (vgl. $AROPE_{ML}$ in Kapitel 2.1.6.1) als erklärende Variablen verwendet.²⁸ Diese Variablen wurden dummykodiert und mit Hilfe eines schrittweisen Ausschlussalgorithmus wurden nur Prädiktoren mit einem bestimmten Signifikanzniveau (maximal zulässige Irrtumswahrscheinlichkeit $\alpha=0,1$) im Modell belassen. Die Dummykodierung wurde dabei angewandt, um nur die für die Schätzung der Antwortwahrscheinlichkeit signifikanten Kategorien der Prädiktoren im Modell zu belassen. Das resultierende Modell hatte ein Pseudo- R^2 von 14,5 %²⁹. Dieser Wert deutet darauf hin, dass die Gründe für nicht erfolgten Kontakt nur zum Teil mit den vorhandenen Prädiktoren geschätzt werden können, nicht erfolgter Kontakt aber auch nicht rein zufällig Auftritt in der Erstbefragung.

Für die Erstbefragungsrotation R1/21 wurden Anpassungsgewichte berechnet, weil diese nach der erfolgten Gewichtung zum Ausgleich von Noncontact auch den Ausfall in Bezug auf Verweigerung ausgleichen sollte. Dies ist auch durch Kalibrierung unter der Verwendung von Anpassungsgewichten möglich. Zusätzlich erlaubt die Kalibrierung Konsistenz mit Randverteilungen der Grundgesamtheit (welche durch den Mikrozensus ermittelt wurden bzw. aus Verwaltungsdaten zur Verfügung standen). Ziel ist es somit für die Erstbefragungsrotation ein Haushalts- und Personengewicht zur Verfügung zu haben, mit dessen Hilfe repräsentative Statistiken erstellt werden können.

²⁸ Selbständig, unselbständig, in Pension, Arbeitslosigkeit gemäß Verwaltungsdaten.

²⁹ Nagelkerkes Pseudo- R^2 .

Dabei erhält vorerst jeder der h Haushalte, die an der Befragung teilnehmen, ein an den Noncontact angepasstes Gewicht b_h .

$$b_h = d_s \cdot \frac{1}{\widehat{c}_h} \quad h \in \{1, \dots, H^{(1)}\} \quad (3)$$

$H^{(1)}$... Anzahl der Haushalte, die an der Erstbefragung teilnahmen

Um zu großen Schwankungen der an den Noncontact angepassten Basisgewichte entgegenzuwirken, wurden diese Gewichte getrimmt. Dabei wurden folgende Schranken für das Gewicht b_h jedes Haushalts h definiert:

$$1/2 \leq \frac{\frac{b_h}{1/H^{(1)} \sum_{i=1}^{H^{(1)}} b_i}}{\frac{d_h}{1/H^{(1)} \sum_{j=1}^{H^{(1)}} d_j}} \leq 2 \quad (4)$$

Das Gewicht b_h , also das Produkt aus Design- und Noncontact-Gewicht, wurde an externe Randverteilungen aus verlässlichen Datenquellen³⁰ angepasst, um die Genauigkeit der Daten zu erhöhen und Kohärenz mit externen Datenquellen zu gewährleisten. Die Kalibrierung bzw. Anpassung im Zuge von EU-SILC 2021 wurde mit Haushaltsmerkmalen und mit auf Haushaltsebene aggregierten Personenmerkmalen durchgeführt. Dabei wurden 2021 folgende Variablen an externe Quellen angepasst:

Haushaltsebene:

- Haushaltsgröße (fünf Kategorien: Haushalte mit ein, zwei, drei, vier, fünf und mehr Haushaltsmitgliedern)
- Rechtsverhältnis an der Wohnung (zwei Kategorien: Eigentum und Nicht-Eigentum)
- Region (neun Kategorien: Bundesländer)
- Urbanisierungsgrad: DEGURBA (drei Kategorien)
- Häufigster (Modalwert) Erwerbsstatus im Haushalt (von Personen ab 16 Jahren): VERWERB (fünf Kategorien)³¹

³⁰ Mikrozensus 2021, Daten des Arbeitsmarktservice 2021, Daten des Dachverbands der Österreichischen Sozialversicherungsträger sowie Lohnsteuerdaten jeweils bezogen auf das Jahr 2020.

³¹ Details zum Verwaltungsdatenerwerbstatus „VERWERB“ vgl.

https://www.statistik.at/fileadmin/pages/54/die_neue_hochrechnung_des_mikrozensus_methodenbeschreibung.pdf (Stand Oktober 2022).

Personenebene:

- Alter (sieben Klassen)
- Geschlecht
- Anzahl ausländischer Staatsbürger:innen (ab 16 Jahren)
- Anzahl Bezieher:innen von Arbeitslosengeld bzw. Notstandshilfe für eine Dauer von mind. 60 Tagen gemäß Daten des Dachverbands der Österreichischen Sozialversicherungsträger (DV)

Die Eckzahlen der Randverteilungen auf Haushaltsebene sowie die soziodemographischen Merkmale auf Personenebene (Alter, Geschlecht, Staatsbürgerschaft) stammen aus dem hochgerechneten Mikrozensus. Die Quartalsgewichte des Mikrozensus werden auf den Quartalsbestand des POPREG³², mit Stichtag zum ersten Tag des Quartals, kalibriert. Wegen der verhältnismäßig frühen Datenlieferung an Eurostat mit Jahresbeginn des an die Erhebung anschließenden Jahres³³ ist es nicht möglich, soziodemographische Eckzahlen für das gesamte Erhebungsjahr zu verwenden. Es stehen demnach nur die Quartale 1 bis 3 zur Verfügung. Allerdings zeigt ein Vergleich, dass zu Gunsten der schnelleren Verfügbarkeit die Unterschiede zwischen der Verwendung der ersten drei Quartale und aller vier Quartale vernachlässigbar sind.

Das Kalibrierungsverfahren liefert das Anpassungsgewicht g_h . In Verbindung mit dem an den Noncontact angepassten Designgewicht ergibt dieser Faktor das Haushaltsgewicht $W_h^{(1)}$ der Erstbefragung welches die Basis für die Berechnung des Querschnittsgewichts in der Ersterhebungsrotation darstellt:

$$W_h^{(1)} = g_h \times b_h \quad (5)$$

Für Folgebefragungshaushalte konnte eine größere Zahl an Prädiktoren in das logistische Regressionsmodell zur Schätzung der Antwortwahrscheinlichkeiten einfließen, weil für jede Person der Folgebefragungsrotationen Informationen aus der Vorjahresbefragung vorhanden waren. Für jede der drei Folgebefragungsrotationen (R2/18, R3/19 und R4/20) wurde ein separates Modell (Formel (6)) auf Personenebene mit einer weit höheren Anzahl an Prädiktoren b_j ($j \in \{1, \dots, J\}$) als im Modell in Formel (2) geschätzt.

$$\hat{r}_p = P(\text{Resp} = 1 | x_1, \dots, x_j) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_j X_j)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 X_1 + \dots + \hat{\beta}_j X_j)} \quad (6)$$

³² Das POPREG ist ein registerbezogenes Datenbanksystem, welches auf dem Zentralen Melderegister (ZMR) basiert. Es liefert die grundlegenden Daten zur Publikation des Bevölkerungsstands. Siehe: <https://www.statistik.at/services/tools/services/publikationen/detail/977?cHash=092f5d7a6cf2016f2c53bf92c3c9bef8> Stand August 2022).

³³ Im Fall von EU-SILC 2021: Datenlieferung mit Anfang 2022.

Die Prädiktorvariablen wurden wiederum dummykodiert und mittels eines schrittweisen Auswahlverfahrens wurden nur signifikante Variablen ($\alpha=0,1$) im Modell belassen.

Schritt 3: Basisgewicht

Grundlage der folgenden Gewichtungsschritte, sowohl im Quer- als auch im Längsschnitt, ist das sogenannte „Basisgewicht“ mit der Eurostat-Zielvariablen-Bezeichnung RB060. Für die Ersterhebung von EU-SILC 2021 ist das Basisgewicht ident mit dem an externe Randverteilungen angepassten Produkt aus Design-Gewicht und Noncontact-Gewicht. Genauer wird dabei zuerst das Designgewicht d_s mit dem Inversen der geschätzten Kontaktwahrscheinlichkeit \widehat{c}_h multipliziert (vgl. Formel (3)).

Das Basisgewicht auf Personenebene $RB060^{(1)}$ ist für die Erstbefragung mit dem auf Haushaltsebene an die Nonresponse und externe Datenquellen angepassten Gewicht $W_h^{(1)}$ ident, d.h. alle m_h Personen i im Haushalt h erhalten dasselbe Gewicht.

$$RB060_{h_i}^{(1)} = W_h^{(1)} \quad i \in \{1, \dots, m_h\}, h \in \{1, \dots, H^{(1)}\} \quad (7)$$

m_h ... Anzahl der Person im Haushalt h

In den Folgerhebungsrotationen wurde das Basisgewicht als Produkt aus dem Nonresponse-Gewicht und dem Basisgewicht aus dem Vorjahr für alle Personen p , die an der Erhebung 2021 teilnahmen, berechnet. Das Basisgewicht einer Folgerhebungsrotation ist somit das um den Antwortausfall korrigierte Basisgewicht aus dem Vorjahr. Für die Folgerhebungsrotationen von EU-SILC 2021 wurde das Basisgewicht 2021 ermittelt, indem das Basisgewicht aus 2020 mit dem Nonresponse-Gewicht aus 2021 multipliziert wurde. Da das Nonresponse-Gewicht das Inverse der geschätzten Antwortwahrscheinlichkeit \widehat{r}_p ist, kann das Basisgewicht für die Folgerhebungen folgendermaßen berechnet werden:

$$RB060_p^{(t)} = \frac{RB060_p^{(t-1)}}{\widehat{r}_p^{(t)}} \quad t \in \{2,3,4\}, p \in \{1, \dots, P^{(t)}\} \quad (8)$$

$P^{(t)}$... Anzahl der Personen, die an der Befragung im Jahr t teilnahmen

Um zu großen Schwankungen der an die Nonresponse angepassten Basisgewichte entgegenzuwirken, wurden diese Gewichte getrimmt. Dabei wurden folgende Schranken für das Basisgewicht jeder Person p definiert:

$$1/2 \leq \frac{\frac{RB060_p^{(t)}}{1/P^{(t)} \sum_{i=1}^{P^{(t)}} RB060_i^{(t)}}}{\frac{RB060_p^{(t-1)}}{1/P^{(t-1)} \sum_{j=1}^{P^{(t-1)}} RB060_j^{(t-1)}}} \leq 2 \quad (9)$$

In den Folgebefragungen erkennt man zwei Personengruppen, die nicht zur ursprünglichen Erstbefragungsstichprobe der jeweiligen Rotation gehören konnten und deshalb kein Basisgewicht besitzen: Neugeborene und Personen, die nach dem Ersterhebungsjahr in einen befragten Haushalt aus dem Ausland zugezogen sind. Für Neugeborene wird das Basisgewicht der Mutter verwendet.³⁴ Für neu in den Haushalt zugezogene Personen wird das Basisgewicht auf den Wert Null gesetzt, falls diese Personen nicht in der Stichprobe des ersten Erhebungsjahres der betreffenden Rotation vorhanden waren (Co-Residents). Eine Ausnahme bilden Personen, die erst nach dem Erhebungsbeginn in einen Haushalt der Folgebefragungsstichprobe von außerhalb der Grundgesamtheit (Ausland, Nicht-Privathaushalte) eingezogen sind. Diese hatten nicht die Chance in die Ersterhebung zu gelangen und erhielten das mittlere Basisgewicht aller Personen im Haushalt, um den durch den Zuzug verursachten Zuwachs der Grundgesamtheit abzubilden.

Schritt 4: Weight-Sharing

Nach Berechnung des Basisgewichts im dritten Gewichtungsschritt wurde im vierten Bearbeitungsschritt für jeden Haushalt der Folgebefragungsstichprobe das arithmetische Mittel dieser Personengewichte berechnet und anschließend jedem Haushaltsmitglied zugeteilt (Weight-Sharing).

$$W_h^{(t)} = \frac{1}{m_h} \sum_{i=1}^{m_h} RB060_i^{(t)} \quad t \in \{2,3,4\}, i \in \{1, \dots, m_h\} \quad (10)$$

Personen, die zur Grundgesamtheit gehören, aber erst nach dem Erhebungsbeginn in einen Haushalt der Folgebefragungsstichprobe von außerhalb der Stichprobe zugezogen sind (Co-Residents), besaßen auf Personenebene kein Basisgewicht, weil sie nicht zum Panel gehörten. Da sie aber Teil des Haushalts waren, bekamen auch sie das über den gesamten Haushalt gemittelte Querschnittsgewicht. Das gleiche gilt für Splithaushalte. Dabei handelt es sich um Haushalte, die durch den Auszug einer Sampleperson aus dem Ursprungshaushalt entstehen. Alle Nicht-Samplepersonen in dem entstandenen Splithaushalt besitzen kein Basisgewicht, da sie sich nicht im ursprünglich in der Erstbefragung gezogenen Haushalt befanden. Erst durch das Weight-Sharing wurde ihnen ein Gewicht zugeteilt.³⁵ In der Erstbefragungsstichprobe entfällt das Weight-Sharing, da die Basisgewichte auf Haushaltsebene ident sind (vgl. Formel (7)).

Am Ende von Schritt 4 stand für jeden Haushalt der Folgebefragung (und alle darin enthaltenen Personen) ein Gewicht auf Haushaltsebene zur Verfügung, welches Besonderheiten des Stichprobendesigns, selektive Ausfälle und demographische Veränderungen der Panelbevölkerung berücksichtigte. Für die Folgebefragungsrotationen wurde im Zuge von EU-SILC 2021 auf eine separate

³⁴ Falls sich die Mutter nicht im Haushalt befindet, erhalten Neugeborene das mittlere Basisgewicht der Personen im Haushalt.

³⁵ Eine Ausnahme bilden hier Neugeborene und nach dem Ersterhebungsjahr aus dem Ausland Zugezogene, da sie das Basisgewicht einer Sampleperson zugewiesen bekommen.

Kalibrierung verzichtet, da diese am Ende nach Kombination aller Teilstichproben sowieso durchgeführt wird.³⁶

Schritt 5: Kombination von Teilstichproben

Jede Rotation repräsentiert am Ende des in Abbildung 6 dargestellten Gewichtungsverfahrens die Grundgesamtheit. Für die Gesamtstichprobe mussten die Gewichte in geeigneter Weise kombiniert werden. Zuerst wurde die Summe der pro Rotation vorhandenen Gewichte³⁷ auf die Anzahl der Personen in der jeweiligen Rotationsstichprobe skaliert (normiert), was in einem Haushaltsgewicht $W_h^{(1)}$ für den gesamten Querschnitt aus EU-SILC 2021 resultierte.

Die drei Rotationen der Folgebefragungen bezogen sich auf leicht unterschiedliche Grundgesamtheiten, da die Ersterhebungs-Stichproben der Jahre 2018, 2019 und 2020 der betreffenden Rotationen jene Haushalte nicht repräsentieren konnten, deren Bewohner:innen zu dieser Zeit nicht Teil der Grundgesamtheit waren. Beispielsweise konnten in der EU-SILC-Erhebung des Jahres 2021 jene Personen, die im Jahr 2020 zugewandert waren und separate Haushalte bildeten, nur in der Erstbefragungsstichprobe des Jahres 2021 erfasst werden. Haushalte von Personen, die im Jahr 2019 zugewandert waren, konnten immerhin in zwei Teilstichproben, nämlich der aktuellen Erstbefragung und der Rotation, die im Jahr 2020 erstmals befragt wurde (R4/20), enthalten sein. Zuwander:innen aus dem Jahr 2018 waren in drei von insgesamt vier Teilstichproben repräsentiert. Um dieser Verzerrung zu begegnen, wurden die Gewichte W_h der Zuwanderer im Zuge der Kombination der vier Rotationen so hochskaliert, dass ihre Gewichte jene Rotationen, in denen sie nicht repräsentiert sein konnten, kompensierten. Der Wert der entsprechenden Inflationsfaktoren für die Gewichte $W_h^{(1)}$ betrug 3,28, 1,74 oder 1,26, je nachdem, ob die neu Zugewanderten in einer, zwei oder drei Rotation(en) repräsentiert wurden.

Am Ende des Gewichtungsverfahrens für den Querschnitt der EU-SILC-Erhebung 2021 stand noch eine abschließende gebundene Hochrechnung der Gewichte des gesamten Querschnitts um die Präzision der Daten zu erhöhen und Kohärenz mit verlässlichen externen Datenquellen sicherzustellen. Dabei wurde genau das gleiche Verfahren wie in Schritt 2 des Gewichtungsverfahrens für die Erstbefragung beschrieben angewendet, wobei zusätzlich an die Zahl der Bezieher:innen von Unselbständigeneinkommen (Personen ab 15 Jahren) und Pensionseinkommen laut Lohnsteuer kalibriert wurde. Das heißt es wurde an alle in Tabelle 18 bis Tabelle 22 dargestellten Randverteilungen kalibriert.

³⁶ Dies ist auch im Einklang mit den Eurostat-Vorgaben, vgl. Eurostat (2021), S. 34ff.

³⁷ Für die Folgebefragungen handelt es sich um die an die Nonresponse angepassten Gewichte $W_h^{(1)}$ in Formel (10), für die Erstbefragung handelt es sich um das kalibrierte Gewicht $W_h^{(1)}$ aus Formel (5).

Tabelle 18: Randverteilung Kalibrierung Bundesland

Bundesland	Haushaltsebene	Personenebene
Burgenland	128 783	292 262
Kärnten	256 915	553 599
Niederösterreich	742 414	1 672 475
Oberösterreich	650 920	1 474 135
Salzburg	556 852	1 227 235
Steiermark	245 278	551 129
Tirol	335 590	747 442
Vorarlberg	172 332	394 794
Wien	924 550	1 883 829

Q: STATISTIK AUSTRIA. EU-SILC 2021, Mikrozensus 2021.

Tabelle 19: Randverteilung Kalibrierung Haushaltsgröße

Haushaltsgröße	Haushaltsebene	Personenebene
1 Person	1 521 763	1 521 763
2 Personen	1 218 266	2 436 531
3 Personen	582 328	1 746 985
4 Personen	450 327	1 801 308
5+ Personen	240 952	1 290 315

Q: STATISTIK AUSTRIA. EU-SILC 2021, Mikrozensus 2021.

Tabelle 20: Randverteilungen Kalibrierung Haushaltsebene

Modalwert des Erwerbstatus (Personen ab 16 Jahren)		Rechtsverhältnis an der Wohnung x Urbanisierungsgrad		
		Besiedlungsdichte	Eigentümer	Nicht-Eigentümer
Unselbständig (standard)	2 202 746	hoch	314 322	1 038 606
Unselbständig (nicht standard)	111 424	mittel	646 405	586 655
Selbständig	156 485	niedrig	967 380	460 267
Arbeitslos	148 913			
Nicht-Erwerbstätig	1 394 068			

Q: STATISTIK AUSTRIA. – EU-SILC 2021, Mikrozensus 2021.

Tabelle 21: Randverteilungen Personenebene Teil 1

Alter & Geschlecht			Staatsbürgerschaft	
Alter	Männer	Frauen		
0-13	618 288	581 602	Österreich	7 316 402
14-24	509 716	486 909	Nicht-Österreich	1 480 500
25-34	606 792	584 611		
35-44	591 842	585 055		
45-54	641 962	650 016		
55-64	689 016	704 721		
65+	679 321	867 051		

Q: STATISTIK AUSTRIA. EU-SILC 2021, Mikrozensus 2021.

Tabelle 22: Randverteilungen Kalibrierung Personenebene Teil 2

	Bezieher:innen
Arbeitslosengeld oder Notstandshilfe (mind. 60 Tage)	679 253
Unselbständigeneinkommen (Personen ab 15 Jahren)*	4 269 371
Pensionseinkommen*	2 123 125

Q: STATISTIK AUSTRIA. Daten des Dachverbands der Sozialversicherungsträger 2020, Daten des Arbeitsmarktservice 2020, Lohnsteuer 2020. – *Randverteilungen werden nur für die Kalibrierung der Gewichte des gesamten Querschnitts verwendet.

Das Haushaltsgewicht $hgew_{hi}$, welches allen Personen i pro Haushalt h zugewiesen wird, ermöglicht schließlich für die Grundgesamtheit repräsentative Schätzungen auf Basis der EU-SILC Stichprobe.

$$hgew_{hi} = hgew_h = g_h \times W_h \quad i \in \{1, \dots, m_h\}, h \in \{1, \dots, H\} \quad (11)$$

m_h ... Anzahl der Person im Haushalt h

H ... Anzahl der Haushalte des gesamten Querschnitts, die an der Befragung teilnahmen

2.2.5.2 Längsschnittgewichtung 2018-2021

Im Unterschied zum Querschnittsgewicht, welches für alle Personen in einem Haushalt gleich ist, handelt es sich bei den Längsschnittgewichten um Personengewichte, welche für jede Person separat berechnet werden. Grundlage dieser Personengewichte sind jene Basisgewichte, die im dritten Gewichtungsschritt des vorigen Abschnitts beschrieben wurden. Insgesamt gibt es für den Längsschnitt EU-SILC 2018-2021 drei verschiedene Längsschnittgewichte, die jeweils für die Zwei-, Drei- oder Vierjahrespopulation ermittelt wurden. Da ausschließlich die vierjährige Rotation (R2/18) die Berechnung des Indikators „dauerhafte Armutsgefährdung“³⁸ ermöglicht, beziehen sich die folgenden Ausführungen nur auf dieses Panel.

Für die gebundene Hochrechnung der Längsschnittpopulation ist ein eigener Kalibrierungsschritt notwendig. Grundlage sind alle Stichprobenpersonen, die in allen vier Jahren befragt werden konnten. Im vierjährigen Panel nicht erfasst sind Personen, die zwischen 2018 und 2021 geboren, nach Österreich zugezogen oder von einem Anstaltshaushalt in einen Privathaushalt umgezogen³⁹ sind. Auch Personen, die nach 2018 in die befragten Haushalte zugezogen sind, können nicht mit einbezogen werden. Des

³⁸ Nur für Personen, die vier Jahre in Folge befragt wurden. Armutsgefährdung am Ende des Beobachtungszeitraumes, d.h. im aktuellsten Jahr und in mindestens zwei von drei vorhergehenden Jahren. Somit bezieht sich der Indikator dauerhafte Armutsgefährdung auf den Zeitraum 2018-2021.

³⁹ Z.B. Umzug von einem Studentenwohnheim in einen Privathaushalt.

Weiteren reduziert sich die Personenanzahl um jene, die in diesen vier Jahren verstorben, ins Ausland oder in einen Anstaltshaushalt verzogen sind oder für ein Jahr abwesend waren. Personen, die innerhalb des vierjährigen Befragungszeitraums in einen anderen Privathaushalt in Österreich umzogen, wurden im neuen Haushalt weiter befragt, sofern es sich nicht um unter 14-Jährige handelte, die bei einem Wegzug aus dem Ursprungshaushalt in der Erhebung deshalb nicht weiterverfolgt wurden.

Die Stichprobengröße im Vier-Jahres-Panel der Erhebungen EU-SILC 2018-2021 beträgt 2 307 Personen. Diese stammen aus 1 213 Haushalten. Diese Zahl umfasst all jene Personen, die vier Jahre in Folge befragt wurden. Die 2 307 Personen des Panels 2018-2021 entsprechen mit dem Gewicht für den Längsschnitt hochgerechnet rund 8 265 000 Personen, das sind rund 94 % der Bevölkerung im Querschnitt 2021.

Die Bevölkerungsstruktur des Längsschnitts 2018-2021 wurde über eigene Längsschnittgewichte so weit wie möglich an die Querschnittbevölkerung 2021 angepasst. Um Abweichungen der Längsschnittergebnisse zu den Querschnittergebnissen gering zu halten, wurde außerdem an die Armutsgefährdungsquote und den Median der Äquivalenzeinkommen im Querschnitt kalibriert.

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

Für die Berechnung der Einkommen aus unselbständiger Erwerbstätigkeit und aus Altersleistungen müssen etwaige Einmalzahlungen (Abfertigungen) herausgerechnet werden. 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, geschah die Zuordnung der Einmalzahlungen nach einer nationalen Regelung. Die Einmalzahlungen werden anteilig entweder als Unselbständigen-, Arbeitslosen- oder Pensionseinkommen klassifiziert, je nach aktuellem Erwerbsstatus der interviewten Person. Der Anteil der Einmalzahlungen, der die Höhe eines Jahreseinkommens übersteigt, wird nicht berücksichtigt.

Neben den Einmalzahlungen ist eine österreichische Einkommensvariable den Eurostat-Zielvariablen nicht eindeutig zuordenbar. Das Einkommen aus Präsenz- und Zivildienst wurde als Einkommen aus unselbständiger Erwerbstätigkeit klassifiziert.

Ein weiteres angewandtes Rechenmodell war die Brutto-Netto-Rechnung bzw. Netto-Brutto-Rechnung. Diese wurde notwendig, wenn für eine Zielperson entweder nur das Brutto- oder – was häufiger der Fall war – nur das Nettoeinkommen in den Daten vorhanden war. Dies war etwa der Fall, wenn das betreffende Einkommen im Zuge der Bearbeitung des Datensatzes imputiert wurde. Die Umrechnung erfolgte bei Unselbständigen und Pensionist:innen auf Basis der Lohnsteuerdaten 2020, also auf Basis realer Faktoren.

Bei selbständigen Einkommen wurde kein Bruttowert erfragt, d.h. alle Bruttowerte wurden aus den Nettowerten berechnet. Dazu wurden die tatsächlich bezahlten Sozialversicherungsbeiträge und Einkommenssteuerzahlungen erfragt und zum Nettowert hinzugerechnet.

Informationen zur Imputation von Antwortausfällen sind im Kapitel 2.2.4 „Imputation“ zu finden.

2.2.7 Sonstige qualitätssichernde Maßnahmen

Um den Einfluss einer selektiven Adressenbearbeitung auf die Qualität der Ergebnisse einzuschränken, wurden auch 2021 folgende Maßnahmen implementiert: Erstens hatten pro Haushalt zumindest drei Kontaktversuche zu erfolgen, bevor eine Verweigerung akzeptiert werden konnte. Zweitens mussten regionale Antwortquoten erfüllt werden: Bei der Ersterhebung mussten in jeder Region mindestens 55 % der Haushalte erfolgreich befragt werden (insgesamt lag die geforderte Ausschöpfungsquote in der Ersterhebung bei 65%), bei der Folgerhebung mussten mindestens 75% aller ausgewählten Haushalte einer Region erfolgreich befragt werden (bundesweit 82,5%). Diese Ausschöpfungsquoten wurden in der Erstbefragung nicht erreicht, in der Folgebefragung überschritten.

Die Vereinbarung von Feldberichten und Zwischendatenlieferungen ermöglichte vonseiten der Fachabteilung die Kontrolle der Daten noch während der Feldarbeitszeit. Somit konnten eventuell auftretende Fehler und Fehlentwicklungen an die Erhebungspersonen rückgemeldet werden und notwendig gewordene Rückfragen bei den Interviewten und Erhebungspersonen gemacht werden.

Als inhaltliche Qualitätssicherungsmaßnahme seien schließlich die Zwischendatenlieferungen angeführt. Die Erhebungsabteilung erhielt zu jeder Datenlieferung umfassende Rückmeldungen, sodass Informationen über die Qualität der Erhebung noch während der Feldarbeit bekannt waren.

2.3 Publikation (Zugänglichkeit)

2.3.1 Vorläufige Ergebnisse

Keine

2.3.2 Endgültige Ergebnisse

Die EU-Indikatoren zu Armut und sozialer Eingliederung 2021 wurden am 29.4.2022, also fünf Monate nach dem Ende des Erhebungsjahres, auf der Internetseite von Statistik Austria veröffentlicht. Sie sind in Kapitel 5 des [Tabellenband zur Publikation EU-SILC 2021](#) (ebenfalls erschienen am 29.4.2022) auf der Seite von Statistik Austria abrufbar.

[Indikatoren auf Basis von EU-SILC](#) sind auch auf der Website von Eurostat verfügbar.

2.3.3 Mikrodaten

Seit dem Beginn der Erhebung werden von Statistik Austria Mikrodaten des Projekts EU-SILC für wissenschaftliche Forschung und Lehre zur Verfügung gestellt. Die Mikrodaten werden etwa vier Monate nach der Veröffentlichung der Ergebnisse zur Verfügung gestellt.

2.3.4 Revisionen

Keine

2.3.5 Publikationsmedien

Publiziert in:

Statistik Austria (2022). Tabellenband EU-SILC 2021 und Bundesländertabellen mit Dreijahresdurchschnitt EU-SILC 2018 bis 2021. Einkommen, Armut und Lebensbedingungen.

Auf der Website von 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 → Bevölkerung und Soziales → Einkommen und soziale Lage → Haushalts-Einkommen 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.

2.3.6 Behandlung vertraulicher Daten

Jede Person hat ein Grundrecht auf Datenschutz. Dies ist im Datenschutzgesetz 2000 (DSG 2000), welches im Verfassungsrang steht, verankert (§ 1).⁴⁰ Laut Bundesstatistikgesetz 2000 (BStatG 2000) ist es erforderlich, dass „personenbezogene Daten zum frühestens möglichen Zeitpunkt gelöscht und anonymisiert werden müssen (§ 15 Abs. 1 und 5 BStatG) oder dass die Identität von Personen verschlüsselt werden muss, falls die Beibehaltung des Personenbezuges notwendig ist (§ 15 Abs. 2 BStatG)“⁴¹. Ein Personenbezug ist dann gegeben, wenn „die Identität der Betroffenen für das Organ der Bundesstatistik bestimmt oder bestimmbar ist (§ 3 Z 15 BStatG)“.

Die personenbezogene Identifizierung ist gemäß § 5 BStatG nur unter gewissen Gesichtspunkten gestattet.⁴² Für EU-SILC ist die personenbezogene Identifizierung notwendig für die Kontaktaufnahme bei der Erstbefragung und für die Weiterverfolgung von Stichprobenpersonen im Zuge der Folgebefragungen. Die Information der Befragten zum Datenschutz und zur Verwendung ihrer personenbezogenen Daten erfolgt über ein Datenschutzblatt.⁴³

Für EU-SILC sind diese gesetzlichen Richtlinien in der ELStV (BGBl 2010 II/277)⁴⁴ genau festgeschrieben.

Im Zuge der Erstbefragungsstichprobe von EU-SILC 2021 wurden Haushalte aus dem ZMR gezogen. Die innerhalb dieser Haushalte mit Hauptwohnsitz gemeldeten Personen besitzen im ZMR einen anonymisierten 28-stelligen Personenschlüssel, das sogenannte „bereichsspezifische Personenkennzeichen“, kurz „bPK“. Die am 31. August 2010 in Kraft getretene „Einkommens- und

⁴⁰ Vgl. <http://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=bundesnormen&Gesetzesnummer=10001597> (Stand Mai 2022).

⁴¹ Vgl. hausinterne Richtlinie „Statistische Geheimhaltung in Publikationen und bei Weitergabe von Daten“ S. 12: <https://www.statistik.at/fileadmin/pages/547/Geheimhaltung-in-Publikationen.pdf> (Stand August 2022).

⁴² <https://www.statistik.at/ueber-uns/aufgaben-und-grundsaeetze/rechtsgrundlagen/bundesstatistikgesetz-und-verordnung-ueber-europaeische-statistiken> (Stand August 2022).

⁴³ <https://www.statistik.at/fileadmin/pages/1215/SCDatenschutz.pdf>

⁴⁴ http://www.ris.bka.gv.at/Dokumente/BgblAuth/BGBLA_2010_II_277/BGBLA_2010_II_277.pdf (Stand Mai 2022).

Lebensbedingungen-Statistikverordnung – ELStV“ ([BGBl. 2010 II/277](#); Änderungen in den Jahren 2013 [BGBl. II, Nr 230/2013](#) und im Jahr 2018 [BGBl. II, Nr 313/2018](#)) erlaubt es, nach der Stichprobenziehung personenbezogene Daten (Vor- und Nachname, Geburtsdatum, Geschlecht und Adresse) der Personen in den gezogenen Haushalten über das Bundesministerium für Inneres (BMI) zu beziehen.

Vor- und Nachname, Geburtsdatum und Geschlecht wurden im Laufe der Erhebung erfragt. 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. Die Identität von Personen im EU-SILC Datensatz ist ab diesem Zeitpunkt nicht mehr bestimmbar. Im Datensatz verblieben lediglich Geschlecht, Geburtsmonat, Geburtsjahr und das Bundesland (NUTS2), da es sich hierbei um Eurostat-Zielvariablen handelt und diese Merkmale für die Datenanalyse von Bedeutung sind.

In den Mikrodatensätzen, die für wissenschaftliche Zwecke zur Verfügung gestellt werden, werden weitere Schritte zur Verhinderung der Identifizierung der befragten Haushalte und Personen getroffen. So wird die Angabe des Alters der befragten Personen mit 80 Jahren beschränkt, sodass ältere Personen nicht mehr mit ihrem genauen Alter identifiziert werden können. Bei Fragen zur Staatsbürgerschaft, zum Geburtsland und bei der Frage nach dem Zuzugsjahr nach Österreich werden die Antwortkategorien in größere Kategorien zusammengefasst. Bei Angaben zur Wohnung wird die Größe der Wohnung (Quadratmeter und Anzahl der Zimmer) trunkiert. Diese Maßnahmen sind analog den Maßnahmen, die auch von Eurostat für die Weitergabe von Mikrodaten vorgesehen werden.⁴⁵

⁴⁵ https://ec.europa.eu/eurostat/documents/203647/203704/C_differences/0e43a8b5-6101-0aab-4c4c-bd1b37dc0263 (Stand November 2022)

3 Qualität

3.1 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 einschließlich des Erhebungsjahres 2007 wurde EU-SILC einerseits durch Eurostat und andererseits durch das Sozialministerium finanziert. Diese beiden Institutionen fungierten somit als Auftraggeber, mit denen während der Projektarbeit eine intensive Kommunikation stattfand. Seit EU SILC 2008 finanziert das Sozialministerium alleine das Projekt EU-SILC. Durch die fristgerechte und einwandfreie Erfüllung der aus den Projektverträgen entstehenden Verpflichtungen (Datenlieferung, Berichtspflichten usw.) wurden 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 war auch die zentrale Quelle zur Erhebung der vom Europäischen Rat verabschiedeten Indikatoren zur sozialen Eingliederung zur Messung von Armut und sozialer Ausgrenzung. Das waren zunächst die vom Europäische Rat in Lissabon (2000) gesteckten Ziele, bis 2010 bei der Reduktion von Armut deutlich voranzukommen, das Verständnis von Armut und sozialer Ausgrenzung im europäischen Rahmen zu verbessern sowie den Erfahrungsaustausch zwischen den Mitgliedsländern zu fördern. Die in der Nachfolge dieser so genannten Lissabon-Strategie im Jahr 2010 beschlossene Strategie „Europa 2020“ definierte als Ziel die EU-weite Reduzierung der Zahl an armuts- oder ausgrenzungsgefährdeten Personen um 20 Millionen. Im Zuge der 2021 von der EU-Kommission vorgestellten "Europa 2030-Strategie Europäische Säule sozialer Rechte" wurden als neues Ziel die Zahl der von Armut oder sozialer Ausgrenzung bedrohten Menschen bis 2030 um mindestens 15 Millionen zu reduzieren definiert. Für den Indikator "Armut- oder Ausgrenzungsgefährdung" haben sich in Folge geringfügige Definitionsänderungen in den Teilbereichen der materiellen Benachteiligungen und der geringen Erwerbsintensität ergeben.

EU-SILC, als die österreichweit einzige Erhebung zu Haushaltseinkommen, ist auch die Basis für den nationalen Bericht zur sozialen Eingliederung und für die Berichterstattung der Europäischen Kommission. Weiters ist EU-SILC die Datenquelle für mehrere Strukturindikatoren: die Armutsgefährdungsquote vor und nach sozialen Transfers, das Verteilungsquintil und die Quote der dauerhaften Armutsgefährdung. Im Auftrag des Sozialministeriums werden nationale Eingliederungsindikatoren im Rahmen der "Kennzahlen zu Lebensbedingungen" berechnet. Des Weiteren fließen Ergebnisse aus EU-SILC in zahlreiche regelmäßig erscheinende Publikationen wie zum Beispiel „Wohnen“, „Wie geht's Österreich?“, „Bildung in Zahlen“, „Migration und Integration“ und Indikatorenansammlungen u.a. für die Sustainable Development Goals ein.

Für Interessierte aus Wissenschaft, Administration und Interessensvertretungen werden die Erhebungsdaten für wissenschaftliche Zwecke zur Verfügung gestellt. Verfügbar sind derzeit die Daten aus EU-SILC für den Zeitraum 2003 – 2021. Zur Verfügung gestellt wird selbstverständlich auch die Möglichkeit, die Daten eines Jahres mit Daten aus anderen Wellen zu verknüpfen (Panel). Bislang nutzten über 50 Institutionen aus dem In- und Ausland die Nutzerdaten von EU-SILC. Die Daten von EU-SILC stehen auch im Austrian Microdata Center (AMDC) der Statistik Austria zur Verfügung.⁴⁶

EU-SILC dient auch als Grundlage für Forschungsprojekte auf nationaler und internationaler Ebene. EU-SILC ist Basis für nationale und europäische Mikrosimulationsmodelle (SORESI, EUROMOD) und wird von zahlreichen Forschungsinstitutionen und internationalen Organisationen (z.B. OECD, ILO, LIS) genutzt (siehe Kapitel 1.3).

3.2 Genauigkeit

3.2.1 Stichprobenbedingte Effekte, Repräsentativität

Der Median des äquivalisierten Netto-Jahreshaushaltseinkommens beträgt in EU-SILC 2021 27 428 Euro, der Schätzwert für den Standardfehler des Medians (basierend auf Bootstrapping) beträgt rund 245 Euro. Mit diesem Wert lässt sich das Konfidenzintervall errechnen und die Grenzen angeben, innerhalb derer der wahre Wert des Medians des äquivalisierten Netto-Jahreshaushaltseinkommens mit einer Vertrauenswahrscheinlichkeit von 95 % liegt. 2021 beträgt die Untergrenze des Konfidenzintervalls 26 949 Euro und die Obergrenze 27 908 Euro.

Die Berechnung der Konfidenzintervalle für die Indikatoren zur sozialen Eingliederung sowie für Indikatoren der Europa 2030-Strategie (2021 Indikatoren) werden nach einem vereinfachten Verfahren durchgeführt. Prinzipiell handelt es sich bei der Armutsgefährdungsquote um einen nichtlinearen Schätzer, da die Armutsgefährdungsschwelle vom Median des äquivalisierten Netto-Jahreshaushaltseinkommens abhängt. Die theoretisch auftretende Schwankung des Medians bei wiederholter Durchführung der Stichprobenziehung müsste in die Berechnung des Stichprobenfehlers der Armutsgefährdungsschwelle miteinbezogen werden. Um die Varianzschätzung und somit die Schätzung der Konfidenzintervalle für die große Zahl an Indikatoren zu vereinfachen und zu beschleunigen, wurde für die Armutsgefährdung eine fixe Schwelle als vereinfachende Annahme bestimmt. Das bedeutet, das Verfahren zur Varianzschätzung entspricht in dieser Betrachtungsweise dem eines Mittelwerts.

Die Schätzung des Standardfehlers erfolgte mit Hilfe des von der Stabstelle Methodik der Statistik Austria entwickelten R-Pakets `surveysd` über ein kalibriertes Bootstrap-Verfahren. Dabei konnte mit der Inklusion der Strata in Bezug auf die Armuts- oder Ausgrenzungsgefährdung durch AROPEML im Stichprobendesign (vgl. Kap. 2.1.6.1) ein geringerer Standardfehler erreicht werden. Außerdem erlaubt es `surveysd` die durch die spezielle Rotationsstruktur in EU-SILC vorhandene Abhängigkeit der

⁴⁶ <https://www.statistik.at/services/tools/services/amdc-mikrodaten-fuer-die-wissenschaft>

Stichproben zwischen den Jahren in Form der Kovarianz bei Vergleichen zwischen Jahren zu berücksichtigen. Eine Tabelle mit Konfidenzintervallen für die „EU-Indikatoren zu Armut und sozialer Eingliederung 2021“ befindet sich auf der Webseite von Statistik Austria. Des Weiteren wurden die kalibrierten Replikationsgewichte auch für die Schätzung des Designeffekts (deff) mittels des R-Pakets `surveysd` herangezogen. Für die in Kap. 2.1.6.1 erläuterte effektive Stichprobengröße wurde der Designeffekt der Armutsgefährdungsquote herangezogen. Für EU-SILC 2021 beträgt dieser $deff = 1,27$.

Tabelle 22 zeigt Armutsgefährdungsquoten und 95 %-Konfidenzintervalle für Österreich und die Bundesländer. Deutlich zu erkennen ist, dass die statistischen Schwankungsbreiten in den kleineren Bundesländern größer sind als in Bundesländern mit vergleichsweise großer Einwohnerzahl. So hat das 95 %-Konfidenzintervall der Armutsgefährdungsquote im Burgenland eine Breite von rund $\pm 4,6$ Prozentpunkten [5,3 %; 14,5 %]. In Niederösterreich hingegen beträgt die Schwankungsbreite nur $\pm 3,5$ Prozentpunkte [9,7 %; 16,6 %]. Die Schwankungsbreite hängt allerdings nicht nur von der Stichprobengröße, sondern auch von der Höhe der Armutsgefährdungsquote ab. So hat Wien zwar die meisten Einwohner:innen, aber auch eine vergleichsweise hohe Armutsgefährdungsquote und somit eine hohe Streuung dieses Indikators, welcher im 95 %-Konfidenzintervall von 17,7 % und 25,2 % liegt. Besonders wichtig bei Vergleichen dieser Art ist es, dass aufgrund der statistischen Schwankungsbreiten aus den Bundesländer-Ergebnissen keine Rangfolge abgelesen werden kann.

Tabelle 23: Armutsgefährdungsquote und deren Schwankungsbreite für Österreich und Bundesländer

	Armutsgefährdung		Konfidenzintervall 95 %			
			untere Grenze		obere Grenze	
	in %	in 1 000	in %	in 1 000	in %	in 1 000
Österreich	14,7	1 292	13,3	1 170	16,1	1 413
Burgenland	9,9	29	5,3	16	14,5	42
Kärnten	14,2	79	7,3	41	21,1	117
Niederösterreich	13,2	220	9,7	162	16,6	278
Oberösterreich	10,0	147	7,6	113	12,3	182
Salzburg	12,3	68	4,8	27	19,8	109
Steiermark	13,4	164	9,8	121	17,0	208
Tirol	16,1	121	10,7	80	21,6	162
Vorarlberg	15,2	60	9,9	39	20,5	81
Wien	21,4	404	17,7	333	25,2	475

Q: STATISTIK AUSTRIA. EU-SILC 2021.

3.2.2 Nicht-stichprobenbedingte Effekte

3.2.2.1 Qualität der verwendeten Datenquellen

Die im Zuge von EU-SILC analysierten Daten stammen aus der eigens dafür durchgeführten Datenerhebung. Informationen zu Einkommen stammen vorwiegend aus Verwaltungsdaten.

Für die im Zuge der Hochrechnung durchgeführte Kalibrierung der Gewichte ist es ebenfalls erforderlich auf externe Datenquellen zurückzugreifen. Wie bereits in Kapitel 2.2.5 erwähnt, wurde für die Kalibrierung der Gewichte der Mikrozensus 2021 (Quartale 1 bis 3) verwendet. Die Anzahl der Bezieher:innen von Arbeitslosenleistungen stammt aus einem Datensatz des Dachverbands der Österreichischen Sozialversicherungsträger bezogen auf das Jahr 2020, die Anzahl der Bezieher:innen von Unselbständigeneinkommen (ab 15 Jahren) sowie Pensionseinkommen stammen aus Daten der Lohnsteuer 2020 (vgl. Kapitel 2.1.10).

3.2.2.2 Abdeckung (Fehlklassifikationen, Unter-/Übererfassung)

Da sich die Grundgesamtheit/Population und der Auswahlrahmen, aus dem die Erhebungseinheiten gezogen werden, bei Stichprobenerhebungen nicht immer 1:1 decken, kann es zur Unter- bzw. Übererfassung von Erhebungseinheiten kommen.

Der Stichprobenrahmen der EU-SILC-Erstbefragung beruht auf dem ZMR. Dieses ist nach der letzten Volkszählung 2001 durch die Zusammenführung der Gemeinde-Melderegister 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. Bei der Erstbefüllung des ZMR ist nicht auf den Eintrag eines eindeutigen Wohnungsidentifikators Bedacht genommen worden, sodass ein Wohnungszusammenhang nur über die Adresse herstellbar ist.

Die Herstellung eines Wohnungszusammenhangs (und damit eines Haushaltszusammenhangs) über die Adresse funktioniert nicht immer einwandfrei (fehlende Türnummern, unterschiedliche Schreibweisen usw.). Als Abdeckungsfehler ist 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 (Übererfassung von Einpersonenhaushalten in EU-SILC). 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.).

In den Stichproben der Folgebefragungen werden nur jene Personen weiterverfolgt, die im Ersterhebungsjahr 14 Jahre oder älter waren – sogenannte „Stichprobenpersonen“ (vgl. Kapitel 2.1.6.1). Das bedeutet, es kommt in der Erstellung der Folgebefragungsstichprobe zu einer Untererfassung, nämlich genau dann, wenn eine Person, die im Ersterhebungsjahr jünger als 14 Jahre alt war, alleine in einen Haushalt ohne Stichprobenpersonen umzieht. Da diese Person nicht weiterverfolgt wird, fallen sie und der Haushalt, in den sie einzieht, aus dem Panel.

Gegenüber anderen Fehlern haben stichprobenbedingte 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 Interviewer:innen in der Regel korrekt erfasst werden.

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

Unit Nonresponse

Siehe Kapitel 2.1.6.2 Ausschöpfung der Stichprobe.

Item Nonresponse

Von Relevanz ist Item Nonresponse vor allem bei den Einkommensvariablen von EU-SILC. Die Auswertungen der Statistik erfolgen 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, welche die Imputationen bei den Einkommenszielvariablen darstellt, zwischen vollständiger, partiell fehlender und vollständig fehlender Information unterschieden.

Nachdem seit EU-SILC 2012 vor allem Verwaltungsdaten für die Berechnung von Einkommensvariablen verwendet werden, ist der Anteil der Item Nonresponse geringer als in den Vorjahren. Allerdings wurden auch in Einkommensvariablen, für die Verwaltungsdaten verwendet wurden, einzelne Fälle mit fehlenden Werten imputiert, sei es, dass etwa angenommen wurde, dass eine Verknüpfungsmöglichkeit fehlt, oder dass keine Verknüpfung stattfinden konnte (z.B. Grenzgänger:innen bei unselbständigen Einkommen).

Als einzige Variable in dieser Auflistung erfasst die Variable PY200G ein monatliches Einkommen (unselbständig Beschäftigter). Diese Einkommen gehen nicht in die Berechnung des Haushaltseinkommens ein und stammen nicht aus Verwaltungsdaten.

Einkommenskomponenten wie imputierte Miete (HY030) und Zinszahlungen für Hypotheken (HY100) werden nicht in der Tabelle angegeben, weil diese nicht direkt von den Respondent:innen erfragt werden. Die imputierte Miete wird auf Basis von Angaben zur Wohnung basierend auf Mikrozensusdaten geschätzt und die Zinszahlungen für Hypotheken werden auf Basis der Angaben zu den Krediten berechnet.

Tabelle 24: Imputationen der Einkommenszielvariablen

		Haushalte/ Personen	Davon:		
			Vollständige Information	Partiell fehlende Informationen	Vollständig fehlende Information
		N	N	N	N
Netto-Einkommenskomponenten auf Haushaltsebene					
HY010	Gesamtes Bruttohaushaltseinkommen	6 008	4 747	1 170	91
HY020	Gesamtes verfügbares Haushaltseinkommen	6 008	5 402	570	36
HY040N	Einkommen aus Vermietung und Verpachtung	397	387	2	8
HY050N	Familienleistungen/Kindergeld	1 544	1 542	2	0
HY060N	Sonst. Leistungen gegen soziale Ausgrenzung	230	218	1	11
HY070N	Wohnungsbeihilfen	197	191	0	6
HY080N	Erhaltene Transfers zwischen Privathaushalten	479	476	1	2
HY090N	Zinsen, Dividenden	3 777	3 130	182	465
HY110N	Einkommen von Personen unter 16	50	50	0	0
HY130N	Geleistete Transfers zwischen privaten Haushalten	407	407	0	0
HY145N	Einkommenssteuernachzahlungen/- erstattungen	1 539	1 539	0	0

Tabelle 25: Imputationen der Einkommenszielvariablen

		Haushalte/ Personen	Davon:		
			Vollständige Information	Partiell fehlende Informationen	Vollständig fehlende Information
		N	N	N	N
Netto-Einkommenskomponenten auf Personenebene					
py010N	Unselbständige Erwerbstätigkeit	5 688	5 581	24	83
py035N	Beiträge aus privaten Rentensystemen	2 300	2 187	0	113
py050N	Selbständige Tätigkeit	986	50	272	664
py080N	Privatpensionen	481	465	3	13
py090N	Arbeitslosenleistungen	1 040	997	34	9
py100N	Alterspensionen	3 231	3 178	1	52
py110N	Hinterbliebenenleistungen	680	674	0	6
py120N	Krankenleistungen	418	386	28	4
py130N	Invaliditätsleistungen	260	252	1	7
py140N	Ausbildungsleistungen	151	149	0	2

Q: STATISTIK AUSTRIA, EU-SILC 2021.

3.2.2.4 Messfehler (Erfassungsfehler)

Als Messfehler wird die Differenz zwischen dem erhobenen Wert einer Variable und dem wahren – aber unbekanntem – Wert der Variable 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 der Interviewenden auf die zu Befragenden und Erfassungsfehler durch die Interviewende bzw. den Interviewenden),
- die/der Befragte (Effekt der Respondent:innen etwa auf die Interpretation der Fragen).

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

Wichtiges Instrument der Fehleranalyse ist die Nachbefragung der Erhebungspersonen (Debriefing). Die Ergebnisse des Debriefings der Interviewer:innen sowie andere Rückmeldungen der Erhebungsinfrastruktur zum Erhebungsinstrument sind in die CAPI/CATI-Gestaltung der Erhebung EU-SILC 2021 eingeflossen.

Wesentlichste Maßnahme zur Vermeidung von Effekten der Befragenden sind Projektbriefings der Interviewer:innen, in denen sichergestellt wird, dass alle Befragten unter möglichst denselben Bedingungen interviewt werden.

Laufend erstellte Feldberichte und Zwischendatenabzüge ermöglichten die Kontrolle der Daten noch während der Feldarbeitszeit. Somit konnten eventuell auftretende Fehler an die Erhebungspersonen rückgemeldet und bei den Haushalten recherchiert werden.

Proxyinterviews sind in EU-SILC nur in Ausnahmefällen erlaubt, zum Beispiel weil ein Mitglied zum Zeitpunkt der Befragung bzw. für die Dauer der Feldarbeitszeit außerhalb des Haushalts lebt, dauerhaft krank ist oder aufgrund von Sprachproblemen nicht selbst befragt werden kann. Eine Übersicht zu den Proxyinterviews zeigt Tabelle 26.

Tabelle 26: Interviewmodus und Proxyinterviews in EU-SILC 2021

	N	%
CAPI-Interviews	5 109	49,2
CATI- Interviews	4 424	42,6
Proxyinterviews	852	8,2
Insgesamt	10 385	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2021.

Tabelle 27 stellt den Proxyanteil⁴⁷ unterschieden nach dem Befragungsmodus dar.

⁴⁷ Proxyinterviews für Einpersonenhaushalte können dann auftreten, wenn eine nicht dem Haushalt angehörende Person Auskunft über die alleinlebende Person gibt (etwa bei Sachwalterschaften). Dies war in EU-SILC 2021 bei keinem Haushalt der Fall.

Tabelle 27: Proxyinterviews nach Modus – Alle Personen

	Insgesamt					
	Insgesamt		CAPI		CATI	
	N	%	N	%	N	%
Persönliches Interview	9 533	91,8	5 109	93,4	4 424	90,0
Proxyinterview	852	8,2	362	6,6	490	10,0
Insgesamt	10 385	100,0	5 471	100,0	4 914	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2021.

In der untenstehenden Tabelle werden die Proxyinterviews nach Befragungsart und Erst- und Folgebefragung unterschieden und nach der Hauptbeschäftigung dargestellt (in der Erstbefragung werden nahezu alle Interviews mittels CAPI durchgeführt). In den bisherigen Erhebungen konnte festgestellt werden, dass insbesondere Erwerbstätige eine höhere Proxyrate haben, wohingegen arbeitslose Personen und Pensionist:innen einen höheren Anteil bei persönlichen Interviews aufweisen. Dies gilt in der Erhebung EU-SILC 2021 (wie bereits in den Vorjahren) nur bedingt bzw. nur für Erwerbstätige, die mittels CAPI in der Folgebefragung interviewt wurden.

3.2.2.5 Aufarbeitungsfehler

Die Interviewer:innen übermittelten die durchgeführten Interviews über eine gesicherte Internetverbindung. Die Daten wurden dann aufbereitet und in das passende Datenformat transferiert. Anschließend fanden Plausibilitätsprüfungen, die Datenaufbereitung und Fehlerkorrektur sowie die Imputationen statt. Danach wurden die Daten in das Eurostat-Zielvariablenformat transferiert. Diese Zielvariablen wurden schließlich, nach Vorliegen des authentischen Datenbestandes, ins EXCEL-CSV-Format gebracht und an Eurostat übermittelt.

3.2.2.6 Modellbedingte Effekte

Querschnittimputationen (von Item Nonresponse bei Einkommensfragen) haben einen Einfluss auf die Genauigkeit der Ergebnisse, weil bei Imputationen mittels linearer Regression die Varianz des interessierenden Merkmals unterschätzt wird. Dies deshalb, weil 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 entspricht.

3.3 Aktualität und Rechtzeitigkeit

Die Verordnung von EU-SILC sieht eine vorläufige Datenlieferung neun Monate nach dem Erhebungsjahr ($t + 9$ Monate) und eine Lieferung der endgültigen Daten 12 Monate nach dem Erhebungsjahr vor ($t + 12$ Monate). Dies wurde durch die Lieferung der Daten am 23. Februar 2022 ($t + 3$ Monate) gewährleistet bzw. übertroffen.

3.4 Vergleichbarkeit

3.4.1 Zeitliche Vergleichbarkeit/Veränderungen gegenüber dem Vorjahr

Nutzer:innen unserer Daten haben uns auf einen Fehler bei der Berechnung der Variable Privattransfers (HY130) in EU-SILC 2021 aufmerksam gemacht. Dies führt dazu, dass bei etwa 3% aller Personen in der Stichprobe die geleisteten Privattransfers unterschätzt werden. Da diese Komponenten von den Einkommen aller Haushaltsmitglieder abgezogen wird, wird sohin das Haushaltseinkommen in diesen Fällen geringfügig überschätzt.

Adaptierungen des Befragungsinstruments

Insgesamt änderte sich das Befragungsinstrument nicht wesentlich gegenüber dem Vorjahr. Wesentlichste Änderung war die Integration der Modulfragen zu den Themen Überschuldung, Konsum, Vermögen und Beschäftigung.

Der jeweils aktuelle Fragebogen ist auf der Seite der Statistik Austria verfügbar:

<https://www.statistik.at/ueber-uns/erhebungen/personen-und-haushaltserhebungen/eu-silc-einkommen-und-lebensbedingungen>

3.4.2 Internationale und regionale Vergleichbarkeit

Die Erhebung EU-SILC ist outputharmonisiert, daher können die Daten der an EU-SILC teilnehmenden Länder miteinander verglichen werden. Outputharmonisierung bedeutet, dass von EUROSTAT ein Katalog an Zielvariablen vorgegeben wird, aber kein Fragebogen, wie diese Zielvariablen befüllt werden sollen. Im Gegensatz dazu werden bei input-harmonisierten Erhebungen die Fragebögen und Fragewortlaute vorgegeben. Vorteil der Outputharmonisierung ist die größere Flexibilität und Anpassungsfähigkeit an die jeweiligen nationalen Gegebenheiten, Nachteil ist eine mögliche Einschränkung der Vergleichbarkeit.

An der Erhebung EU-SILC nehmen alle Mitgliedsstaaten der Europäischen Union sowie zusätzlich Großbritannien, Norwegen, Island, die Türkei, die Schweiz, Albanien, Nordmazedonien, Serbien und Montenegro teil.

3.5 Kohärenz

Im folgenden Abschnitt werden die Daten von EU-SILC 2021 mit anderen österreichischen Datenquellen verglichen. Einkommensbezugsjahr ist in EU-SILC das der Erhebung vorangegangene Jahr, für die Erhebung EU-SILC 2021 also das Jahr 2020.

Für die Vergleiche werden folgende Datensätze herangezogen: die Lohnsteuerdaten 2020, der Mikrozensus-Jahresdurchschnitt 2021 und Schätzungen auf Basis der Sektorkonten der Volkswirtschaftlichen Gesamtrechnungen (VGR) des Jahres 2020.

Einkommen unselbständig Beschäftigter

Die Lohnsteuerdaten erfassen die Einkommen von unselbständig Beschäftigten sowie von Pensionist:innen. Sie stellen damit eine bedeutende Quelle für Vergleiche der wichtigsten Einkommenskomponente auf Personenebene dar, dem Einkommen aus unselbständiger Arbeit.⁴⁸ Nachdem die Einkommen aus unselbständiger Arbeit in EU-SILC 2021 aus den Lohnsteuerdaten selbst befüllt werden, werden somit Einkommen aus ein- und derselben Quelle miteinander verglichen. Allerdings bestehen weiterhin Unterschiede zwischen den in EU-SILC und den in der Lohnsteuer erfassten Einkommen aus unselbständiger Erwerbstätigkeit: Unterschiede liegen im Umfang (Stichprobe und Gesamtdatensatz) der berücksichtigten Bezieher:innen und im Unterschied der berücksichtigten Einkommen. Hinsichtlich der berücksichtigten Bezieher:innen umfassen die beiden Datensätze unterschiedliche Personengruppen: In EU-SILC werden keine Personen außerhalb von Privathaushalten erfasst, ebenso wenig wie Personen, die zwischen dem Einkommensbezug 2020 und der Erhebung 2021 verstorben oder ins Ausland verzogen sind. Allerdings wird die Anzahl der Einkommensbezieher:innen von unselbständigen Einkommen in der Berechnung der Gewichte berücksichtigt (vgl. Kapitel 2.2.5). Hinsichtlich der berücksichtigten Einkommen werden in EU-SILC Abfertigungen (bzw. Zahlungen, die mit festen Sätzen besteuert werden) nur zum Teil für das unselbständige Einkommen berücksichtigt. Außerdem werden auch Einkommen aus Milizübungen und Bezüge politischer Mandatar:innen hinzugezählt. Das Einkommenskonzept von EU-SILC ergibt sich aus der Eurostatvorgabe, das Jahreseinkommen möglichst vollständig zu erfassen, Einmalzahlungen wie Abfertigungen aber nur teilweise in der Berechnung des Jahreseinkommens mitzuberechnen.

Um die Einkommen in EU-SILC und den Lohnsteuerdaten miteinander vergleichen zu können, wurde ein Vergleich der Bruttoeinkommen aus unselbständiger Beschäftigung durchgeführt. Diese Einkommen werden in der Variable KZ210_a der Lohnsteuerdaten erfasst und sind die Grundlage für die Berechnung der Einkommen in EU-SILC. Unterschiede zwischen dem Datensatz von EU-SILC und dem Gesamtdatensatz der Lohnsteuer liegen somit in den erwähnten Unterschieden der erfassten Personengruppe und in der Stichprobe von EU-SILC bzw. der Gewichtung.

Tabelle 28 zeigt den Vergleich der Bruttoeinkommen unselbständig Beschäftigter in EU-SILC und den Lohnsteuerdaten. Die Verteilungen der beiden Datenquellen unterscheiden sich kaum voneinander, Abweichungen sind vor allem an den Rändern der Verteilung festzustellen. Die Werte liegen im unteren und im oberen Bereich der Verteilung unter den Vergleichswerten der Lohnsteuer. Dies ist vorrangig auf fehlende Verknüpfungen zwischen Lohnsteuerdaten und der Stichprobe von EU-SILC zurückzuführen. Sofern für Personen ein Bezug von unselbständigen Einkommen angenommen werden kann und keine Verknüpfung zwischen den Datensätzen möglich ist, werden die Einkommen dieser Personen imputiert. Für die Bruttoeinkommen in EU-SILC werden Abfertigungszahlungen nicht in voller Höhe berücksichtigt.

⁴⁸ Vergleiche der Pensionseinkommen sind aus konzeptionellen Gründen schwierig und werden daher hier nicht durchgeführt.

Tabelle 28: Brutto-Jahreseinkommen unselbständig Beschäftigter 2020: Lohnsteuerdaten 2020 und EU-SILC 2021

Perzentile	Lohnsteuerdaten 2020 (in EUR)			EU-SILC 2021 (in EUR)		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	5 859	7 758	4 633	3 156	3 891	2 725
20% ...	13 308	19 074	10 100	9 136	13 234	7 026
25% ...	16 842	23 970	12 996	13 466	18 498	9 972
30% ...	20 139	27 827	15 569	17 285	23 881	13 169
40% ...	26 160	33 558	20 160	23 951	30 837	18 870
50% ...	31 888	38 341	24 471	29 722	36 360	23 489
60% ...	37 355	43 559	29 400	35 835	41 861	28 164
70% ...	43 539	50 366	35 170	42 215	48 141	34 117
75% ...	47 478	54 895	38 559	45 777	52 765	37 592
80% ...	52 463	60 655	42 546	50 966	57 811	41 913
90% ...	69 008	79 792	55 396	66 939	76 543	55 768
arithm. Mittel	36 839	44 241	28 682	33 897	39 954	27 064
Personen	4 098 256	2 148 604	1 949 652	4 222 973	2 238 728	1 984 245

Q: STATISTIK AUSTRIA, EU-SILC 2021; Lohnsteuerdaten 2020.

Werden nur ganzjährige Beschäftigungen miteinander verglichen, wird der Unterschied zwischen Lohnsteuerdaten und den Daten aus EU-SILC deutlich geringer, ist aber dennoch vorhanden. Betrugen die Einkommen der unselbständig Beschäftigten in EU-SILC im untersten Dezil zunächst nur 60 % der Einkommen der Lohnsteuer, so liegen bei ganzjährig Beschäftigten in EU-SILC im untersten Dezil nur rund 15 % unter den Einkommen der Lohnsteuer (Tabelle 29).

Tabelle 29: Brutto-Jahreseinkommen unselbständig Beschäftigter 2020: Lohnsteuerdaten 2020 und EU-SILC 2021 (ganzjährig Beschäftigte)

	Lohnsteuerdaten 2020 (in EUR)			EU-SILC 2021 (in EUR)		
	Insgesamt	Männer	Frauen	Insgesamt	Männer	Frauen
10% ...	15 152	23 727	11 411	10 879	17 828	6 984
20% ...	22 321	30 835	17 197	19 213	27 491	14 583
25% ...	25 154	33 228	19 409	22 248	30 512	17 118
30% ...	27 872	35 337	21 438	25 444	33 213	19 288
40% ...	32 781	39 322	25 211	30 460	37 570	23 555
50% ...	37 332	43 657	29 376	35 598	41 923	27 584
60% ...	42 220	48 972	33 981	40 867	46 839	32 570
70% ...	48 495	56 117	39 353	46 578	53 797	38 124
75% ...	52 603	60 864	42 652	50 933	57 935	41 684
80% ...	57 782	66 800	46 709	56 068	64 341	45 394
90% ...	74 804	86 566	59 786	71 676	82 938	60 469
arithm. Mittel	43 409	52 171	33 800	40 112	47 817	31 630
Personen	3 209 836	1 678 965	1 530 871	3 362 433	1 761 919	1 600 514

Q: STATISTIK AUSTRIA, EU-SILC 2021; Lohnsteuerdaten 2020.

Aggregiertes Haushaltseinkommen

Im Folgenden wird das in EU-SILC erhobene Haushaltseinkommen mit den Daten der Volkswirtschaftlichen Gesamtrechnungen verglichen. Dazu werden die Sektorkonten „Private Haushalte und private Organisationen ohne Erwerbszweck“ der VGR herangezogen. Die folgenden konzeptionellen Unterschiede zwischen EU-SILC und der VGR müssen bei diesem Vergleich berücksichtigt werden:

- Sachleistungen sind in der VGR enthalten, werden aber in EU-SILC nicht vollständig erfasst
- Einkommen aus selbständiger Arbeit ist in der VGR eine residuale Kategorie, wird also nicht selbst erfasst
- Bei der Berechnung der verfügbaren Einkommen in der VGR wird ein Schätzwert für Spenden an karitative Organisationen und Mitgliedsbeiträge abgezogen
- Transnationale Zahlungen werden in der VGR, aber nicht in EU-SILC erfasst
- Zinszahlungen (insbesondere für Hypotheken) werden in der VGR für die Berechnung des Haushaltseinkommens abgezogen (D4PAY), nicht aber in EU-SILC berücksichtigt
- Erhaltene indirekte Gewinnbeteiligungen sind im Einkommenskonzept der VGR enthalten (D4REC), aber nicht vollständig in EU-SILC

- Erhaltene Versicherungsleistungen bzw. gezahlte Versicherungsprämien werden in der VGR berücksichtigt, nicht aber in EU-SILC

Die in EU-SILC erfassten Einkommen liegen so wie in den Vorjahren unter denen der VGR. Der Unterschied zwischen den Schätzwerten verringert sich auf etwa 4 %, wenn Vermögenseinkünfte nicht berücksichtigt werden. Dies weist auf die Untererfassung von Vermögenseinkommen in EU-SILC hin, die in den Einkommen „Vermietung und Verpachtung“ und „Zinsen, Dividenden, Kapitalerträge“ erfasst werden.

Tabelle 30: Haushaltseinkommen und verfügbares Einkommen: EU-SILC 2021 und VGR 2020

	Brutto-Haushaltseinkommen (in Mio. EUR)		Verfügbares Einkommen (in Mio. EUR)
	Insgesamt	Ohne Vermögens-einkommen	
VGR-Ausgangswert aus der Sektorrechnung	289 270	272 848	220 912
Abzug private Organisationen ohne Erwerbszweck ¹	-	-	5 954
Abzug der Bevölkerung außerhalb von Privathaushalten ²	4 513	4 256	3 446
Abzug von Entnahmen ³	3 471	3 274	2 651
Abzug von fiktiven Mieten ⁴	12 860	12 860	12 860
VGR-Schätzwert	268 427	252 457	196 001
EU-SILC Schätzwert	247 329	241 604	184 522
Differenz zw. VGR und EU-SILC in %	7,9	4,3	5,9

Q: STATISTIK AUSTRIA, EU-SILC 2021; VGR 2020.

1) Schätzwert, für das verfügbare Einkommen wird ein gemeinsamer Schätzwert errechnet. – 2) geschätzt auf der Basis der Bevölkerungsprognose; 1,56 im Jahr 2020. – 3) geschätzt auf 1,3% der gesamten Konsumausgaben, KE 2019/20. – 4) VGR 2020.

Wohnen und Wohnkosten

Für den Vergleich zum Thema Wohnen werden die Daten des Mikrozensus als Vergleichsbasis herangezogen. Der Mikrozensus ist eine der wesentlichen Quellen der österreichischen Wohnstatistik und mit einer im Vergleich zu EU-SILC wesentlich größeren Stichprobe eine sehr verlässliche Datenquelle.

Tabelle 31 vergleicht die Verteilung des Rechtsverhältnisses an der Wohnung im Mikrozensus und EU-SILC. Dieser Vergleich ist insofern von Bedeutung, als der Mikrozensus auch Spenderdatensatz für die Berechnung imputierter Mieten in EU-SILC ist, und die Variable „Rechtsverhältnis an der Wohnung“ eine wesentliche Bedeutung für deren Berechnung hat. Da die Erfassung des Rechtsverhältnisses in beiden

Erhebungen unterschiedlich erfolgt, müssen die Kategorien für diesen Vergleich zusammengefasst werden.⁴⁹

Der Vergleich zeigt im Wesentlichen eine hohe Übereinstimmung zwischen den beiden Erhebungen. So ist der Anteil an Wohneigentum mit etwa 48 % ident. Die Anteile der drei Mietwohnverhältnisse unterscheiden sich leicht, wobei der Anteil der Mietverhältnisse insgesamt mit etwa 43 % wiederum annähernd gleich ist.

Tabelle 31: Rechtsverhältnis an der Wohnung: Mikrozensus 2021 und EU-SILC 2021

	Mikrozensus 2021		EU-SILC 2021	
	in 1 000	in %	in 1 000	in %
Insgesamt	4 020	100,0	4 014	100,0
Hauseigentum	1 482	36,9	1 478	36,8
Wohnungseigentum	444	11,0	450	11,2
Miete: Gemeindewohnung	277	6,9	359	9,0
Miete: Genossenschaftswohnung	662	16,5	625	15,6
Miete: Sonstige Hauptmiete	762	19,0	749	18,7
Untermiete	39	1,0	60	1,5
Mietfreie Wohnung / mietfreies Haus	354	8,8	292	7,3

Q: STATISTIK AUSTRIA, EU-SILC 2021, Mikrozensus 2021.

Auch bei der Erfassung der Wohnkosten unterscheiden sich die beiden Erhebungen EU-SILC und Mikrozensus, und können nicht direkt miteinander verglichen werden⁵⁰. Daher wird nur der Gesamtbetrag für Miete und Betriebskosten miteinander verglichen. Zahlungen für Warmwasser, Heizung und Garagenplätze sind nur dann in diesem Betrag enthalten, wenn diese an die Hausverwaltung gezahlt werden.

Der Median von Miete und Betriebskosten ist in EU-SILC und Mikrozensus mit 550 Euro insgesamt derselbe. Hinsichtlich der Region bzw. dem Rechtsverhältnis zeigen sich größere Unterschiede

⁴⁹ So wird im Mikrozensus nicht zwischen mietfreiem Haus und mietfreier Wohnung unterschieden (hier werden die entsprechenden Kategorien von EU-SILC zusammengefasst), dafür hat der Mikrozensus eine eigene Kategorie für mietfreie Wohnungen von "Verwandten oder Schwägerten des Hauseigentümers" (diese werden mit den sonstigen mietfreien Rechtsverhältnissen zusammengefasst). Die Kategorie "Sonstiges entgeltliches Rechtsverhältnis" wird in EU-SILC nicht erfasst. Diese wird zur Kategorie "Sonstige Hauptmiete" hinzugerechnet. Dadurch ergibt sich im Vergleich zu den Veröffentlichungen des Mikrozensus eine etwas höhere Zahl an Mietrechtsverhältnissen.

⁵⁰ Für den Wohnungsaufwand im Mikrozensus bzw. für die Wohnkosten in EU-SILC vgl. Statistik Austria (2022).

insbesondere bei gering besetzten Zellen, also etwa bei Genossenschaftswohnungen nach der Wohnregion (Tabelle 32).

Tabelle 32: Miete und Betriebskosten nach Wohnnutzfläche und Einwohnerzahl in der Region: Mikrozensus 2021 und EU-SILC 2021

		Mikrozensus 2021				EU-SILC 2021			
		Insgesamt	Ge-meinde	Ge-nossen-schaft	Sonstige Haupt-/Unter-miete	Insgesamt	Ge-meinde	Ge-nossen-schaft	Sonstige Haupt-/Unter-miete
Insgesamt	Median	550	409	540	614	550	460	545	635
Größenklassen									
unter 60 m ²	Median	420	322	380	500	425	380	420	480
60 bis unter 120 m ²	Median	622	498	603	700	630	520	612	730
120 und mehr m ²	Median	1 035	1 048	815	1 141	840	750	900	850
Einwohnerzahl in der Region									
Wien	Median	522	413	570	621	550	445	609	679
> 100 000	Median	566	422	500	645	580	502	540	645
> 10 000	Median	557	400	511	645	533	477	472	650
<= 10 000	Median	559	391	550	590	550	488	548	550

Q: STATISTIK AUSTRIA, EU-SILC 2021, Mikrozensus 2021.

Bildung

Die Frage nach dem höchsten Bildungsabschluss wird in EU-SILC und im Mikrozensus gestellt. Allerdings unterscheiden sich in den beiden Erhebungen die Altersgruppen, auf die sich der höchste Bildungsabschluss bezieht. Im Mikrozensus sind dies Personen ab 15 Jahren, bei EU-SILC werden Daten zum Bildungsabschluss erst für Personen ab 16 Jahren erhoben. Die Daten, welche aus dem Mikrozensus für den Vergleich herangezogen wurden, sind deshalb auf die Bevölkerung ab 16 Jahren eingeschränkt. Tabelle 33 zeigt einen Vergleich der Daten zur höchsten abgeschlossen Bildung zwischen Mikrozensus 2021 und EU-SILC 2021.

Tabelle 33: Höchster Bildungsabschluss: Mikrozensus 2021 und EU-SILC 2021

	Mikrozensus 2021		EU-SILC 2021	
	in 1 000	in %	in 1 000	in %
Pflichtschule/keine Pflichtschule	1 114	20,0	3	2,1
Lehre (Berufsschule)	1 887	33,9	1 386	860,2
Fach- oder Handelsschule	707	12,7	2 404	1492,4
Matura	937	16,8	993	616,7
Abschluss einer Universität, (Fach-)Hochschule	846	15,2	1 219	756,7
Anderer Abschluss nach der Matura	80	1,4	1 231	764,4
Gesamt	5 572	100,0	161	100,0

Q: STATISTIK AUSTRIA, EU-SILC 2021, Mikrozensus 2021. – Personen ab 16 Jahren.

Die Fragestellung unterscheidet sich in den beiden Erhebungen, allerdings wurden die Fragestellungen im Zuge der Erfassung der revidierten ISCED-Kategorien miteinander harmonisiert. Dadurch sind die Abweichungen der beiden Erhebungen gering und sind in jeder Kategorie weniger als zwei Prozentpunkte.

Außerdem ist bei Vergleichen dieser Art zu beachten, dass aufgrund zufallsbedingter Schwankungen in beiden Erhebungen Unterschiede in der Verteilung vergleichbarer Variablen auftreten können, die nicht durch Messfehler bzw. -unterschiede erklärbar sind.

4 Ausblick

Für die Datenerhebung ab **2023** wurde 2021 ein **Eurostat Grant** beantragt und bewilligt. Über diesen wird die Umstellung auf den neuen Modus **CAWI** (Computer Assisted Web Interviewing) **und einen "mobile first"**-tauglichen Fragebogen mitfinanziert. Ziel ist es, den zusätzlichen Interviewmode CAWI neben CAPI und CATI anzubieten, um effizienter und respondentenfreundlicher erheben zu können. Dazu müssen sämtliche (600+) Fragebogenfragen sowie der Fragenfluss, der Workflow und die Kommunikationsstrategie überarbeitet werden. Umfassende Tests (kognitive, Friendly User, Pilottests) begleiten die Neuentwicklung des Frageninstruments. Der Arbeitsaufwand dafür beträgt etwas mehr als 3 Personenjahre und wird zwischen Oktober 2021 und September 2022 geleistet.

5 Glossar

Äquivalenzeinkommen: Das auf einen Einpersonenhaushalt standardisierte Haushaltseinkommen.

Armutsgefährdung: Alle Personen, deren Äquivalenzeinkommen unterhalb eines festgelegten Schwellenwertes (Armutsgefährdungsschwelle = 60 % des Medians) liegt, gelten als armutsgefährdet. Ist von „Armutsgefährdung“ oder „Armutsrisiko“ ohne Zusatz die Rede, ist immer die Armutsgefährdung nach Sozialleistungen gemeint.

Armutsgefährdungslücke: Maß für die Intensität der Armutsgefährdung definiert als Median der individuellen relativen Abweichungen der Äquivalenzeinkommen der Armutsgefährdeten von der Armutsgefährdungsschwelle in Prozent dieser Schwelle.

Längsschnitterhebung: Bei einer Längsschnitterhebung beziehen sich die Variablen auf mehrere Zeitpunkte bzw. -perioden. Bei EU-SILC beziehen sich die Längsschnittdaten der Jahre 2018-2021 auf diese Jahre bzw. auf die Einkommensbezugsperioden 2017-2020.

Panelerhebung: Bei einer Panelerhebung werden dieselben Personen in mehreren Jahren befragt. In EU-SILC werden die Personen etwa eines Viertels der Stichprobe in vier aufeinanderfolgenden Jahren befragt.

Proxy(-interview): Interview, bei dem Dritte über die eigentliche Zielperson Auskunft geben.

Querschnitterhebung: Die Daten einer Querschnitterhebung beziehen sich auf ein und denselben Zeitpunkt bzw. Zeitraum. In der Erhebung EU-SILC beziehen sich die Angaben einer Welle zumeist auf den Befragungszeitpunkt oder das vorangegangene Jahr (Einkommensangaben).

6 Abkürzungsverzeichnis

CAPI

Computer Assisted Personal Interview

CATI	Computer Assisted Telephone Interview
ECHP	European Community Household Panel
ELStV	Einkommens- und Lebensbedingungen Statistikverordnung
EU	Europäische Union
ISCED	International Standard Classification of Education
MZ	Mikrozensus
NACE	Nomenclature statistique des activités économiques dans la Communauté européenne
NUTS	Nomenclature des unités territoriales statistiques
SILC	Statistics on Income and Living Conditions
VGR	Volkswirtschaftliche Gesamtrechnungen
ZMR	Zentrales Melderegister

7 Hinweis auf ergänzende Dokumentationen/Publikationen

Bauer, M./Lamei, N. (2005). EU-SILC – die neue Erhebung zu Einkommen und Lebensbedingungen. In: Statistische Nachrichten 3/2005. Wien. S. 224-231.

BMASK (2010). Armutsgefährdung und Lebensbedingungen in Österreich. Ergebnisse aus EU-SILC 2009. Methoden und Vergleiche zu EU-SILC. Bundesministerium für Arbeit, Soziales und Konsumentenschutz. Wien.

Eurostat (2013). Standard error estimation for the EU-SILC indicators of poverty and social exclusion. Eurostat. Luxemburg.

Eurostat (2021). Methodological Guidelines and Description of EU-SILC Target Variables. DocSILC065 (2021 operation). Eurostat. Luxemburg.

- Geisberger, T./Till, M. (2009). Der neue EU-Strukturindikator „Gender Pay Gap“. In: Statistische Nachrichten 1/2009. Wien. S. 64-70.
- Glaser, T./Till, M. (2010). Gewichtungungsverfahren zur Hochrechnung von EU-SILC-Querschnittergebnissen. In: Statistische Nachrichten 7/2010. Statistik Austria. Wien. S. 566-577.
- Heuberger, R./Lamei, N. (2006). Das Datenmanagement in EU-SILC – von der Befragung zu Sozialindikatoren. In: Statistische Nachrichten 11/2006. Wien. S. 1054-1061.
- Horvitz, D. G./Thompson, D. J. (1952). A Generalization of Sampling Without Replacement From a Finite Universe. In: Journal of the American Statistical Association. 47. S. 663–685.
- Lamei, N./Till-Tentschert, U. (2005). Messung von Armutsgefährdung und Deprivation. In: Statistische Nachrichten 04/2005. Wien. S. 349-359.
- Lamei, N./Till-Tentschert, U. (2009). Dauerhafte Armutsgefährdung in Österreich 2004 bis 2007. In: Statistische Nachrichten 01/2010. Wien. S. 31-40.
- Lamei, N./Skina-Tabue, M. (2011). Lebensstandard aus Geschlechterperspektive. Indikatoren zu Einkommen, Armutsgefährdung und benachteiligten Lebenslagen von Frauen und Männern. In: Statistische Nachrichten 12/2011. Wien. S. 1205-1215.
- Särndal, C.-E./Swensson, B./ Wretman, J. (2003), Model assisted survey sampling, Springer, New York.
- Statistik Austria (2006). Einkommen, Armut und Lebensbedingungen. Ergebnisse aus EU-SILC 2004. Wien.
- Statistik Austria (2021). Wohnen 2021. Zahlen, Daten und Indikatoren der Wohnstatistik. Wien.
- Till, M. (2006). Jahreseinkommen und erwartete Lebensstandardpositionen von Personen in Privathaushalten. In: Statistische Nachrichten 4/2006. Wien. S. 250-260.