

Poverty Measurement

Guide to Data Disaggregation



UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

Poverty Measurement Guide to Data Disaggregation



United Nations
New York and Geneva, 2020

© 2020 United Nations

This work is available open access by complying with the Creative Commons license created for inter-governmental organizations, available at <http://creativecommons.org/licenses/by/3.0/igo/>

Publishers must remove the UN emblem from their edition and create a new cover design. Translations must bear the following disclaimer: “The present work is an unofficial translation for which the publisher accepts full responsibility.” Publishers should email the file of their edition to permissions@un.org.

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries.

Photocopies and reproductions of excerpts are allowed with proper credits.

This publication is issued in English and Russian.

ECE/CES/STAT/2020/4

ISBN: 978-92-1-117260-7
eISBN: 978-92-1-005352-5
Sales no.: E.20.II.E.43
ISSN: 0069-8458

United Nations publication issued by the United Nations Economic Commission for Europe.

Preface

In these times of growing inequalities, poverty is increasingly recognized as a global challenge. The pledge of the 2030 Agenda for Sustainable Development is to reach the furthest behind first and leave no one behind. However, the most vulnerable or disadvantaged groups are particularly hard to identify or capture in statistical surveys. Greater data granularity is therefore needed to understand who is at risk of poverty and where support is needed the most.

To take further the methodological work and develop recommendations to countries on data disaggregation with particular focus to poverty measurement, the Conference of European Statisticians established a task force, which worked through 2018 and 2019 to develop the present Guide. The task force consisted of 30 statistical experts from UNECE member countries, other countries participating in the work of the Conference of European Statisticians, international organizations and academia. It builds on the UNECE Guide on Poverty Measurement published in 2017.

The Guide's recommendations aim to improve national poverty statistics by ensuring better availability of disaggregated data for measuring poverty in alignment with international standards. The publication mainly targets national statistical authorities and provides useful information for policymakers, researchers and other users of poverty data. The Guide was endorsed by the 68th plenary session of the Conference of European Statisticians in 2020.

UNECE is grateful to all the experts who were involved in the preparation of this Guide.

Acknowledgements

This Guide has been prepared by the UNECE Task Force on Disaggregated Poverty Measures, which consisted of the following members:

Matthias Till (Statistics Austria) – Chair of the Task Force

Nadja Lamei, Magdalena Skina-Tabue Fotso and Martin Bauer (Statistics Austria)

Andrew Heisz, Cilanne Boulet and Samir Djidel (Statistics Canada)

Jakub Hrkal (Institute of Health Information and Statistics of the Czech Republic)

Federico Polidoro (Italian National Institute of Statistics)

Edgar Vielma and Óscar Ramirez (INEGI, Mexico)

Robert Vlacuha and Michaela Glaser-Opitzova (Statistical Office of the Slovak Republic)

Marina Malincheva, Natalia Ignatova and Tatiana Velikanova (Rosstat, Russian Federation)

Martina Guggisberg (Swiss Federal Statistical Office)

Richard Tonkin and Dominic Webber (United Kingdom Office for National Statistics)

Trudi Renwick (United States Census Bureau)

Valentina Bryseva (CIS-STAT)

Didier Dupré, Agata Kaczmarek-Firth, Emilio Di Meglio and Sigita Grundiza (Eurostat)

Ursula Till-Tentschert (EU Agency for Fundamental Rights)

Marco Mira d'Ercole and Carlotta Balestra (OECD)

Elena Danilova-Cross (UNDP Istanbul Regional Hub)

Andres Vikat and Vania Etropolska (UNECE)

Joanne Bosworth and Enrique Delamonica (UNICEF)

Ginette Azcona (UN Women)

Nobuo Yoshida, Paul Andres Corral Rodas and Monica Robayo-Abril (World Bank)

Gianni Betti (University of Siena)

Adriana Conconi and Ricardo Nogales (Oxford Poverty and Human Development Initiative)

Dave Gordon (University of Bristol)

Teresa Munzi (Luxembourg Income Study)

The chapters of the Guide have been discussed and agreed by the entire Task Force. Some members took primary responsibility for drafting certain chapters, as follows:

Chapter 1 by Matthias Till (Austria) and Vania Etropolska (UNECE)

Chapter 2 by Jakub Hrkal (Institute of Health Information and Statistics of the Czech Republic)

Chapter 3 by Ursula Till-Tentschert (EU Agency for Fundamental Rights) and Dominic Webber (United Kingdom Office for National Statistics)

Chapter 4 by Gianni Betti (University of Siena), Cilanne Boulet (Statistics Canada) and Matthias Till (Austria)

Chapter 5 by Trudi Renwick (United States Census Bureau)

Chapter 6 by Matthias Till (Austria)

The Task Force wishes to recognize the contribution of the late Professor Vijay Verma who inspired the work on this Guide, especially on chapters 3 and 4.

UNECE acknowledges financial support through the United Nations Development Account (10th tranche) project “Data and Statistics” for the work related to Chapter 4.

Table of Contents

Preface	III
List of Figures	VIII
List of Tables	X
List of Boxes	XI
Acronyms / Abbreviations	XII
1 Introduction	1
1.1 Why this Guide?	1
1.2 Background.....	3
1.3 Outline of the Guide	5
1.4 Summary of recommendations.....	7
1.4.1 Chapter 2: Standard core variables for disaggregation.....	7
1.4.2 Chapter 3: Including hard to reach groups in poverty measurement.....	7
1.4.3 Chapter 4: Assessing and improving survey methods.....	8
1.4.4 Chapter 5: Supplemental or experimental poverty measures	9
2 Standard Core Variables for Disaggregation	13
2.1 Policy relevant target groups for poverty disaggregation.....	13
2.2 Defining variables for poverty disaggregation	21
2.2.1 Sex	21
2.2.2 Age.....	22
2.2.3 Disability status	24
2.2.4 Migratory status	29
2.2.5 Ethnicity.....	32
2.2.6 Household type	35
2.2.7 Current employment status	37
2.2.8 Tenure status of the household	38
2.2.9 Receipt of cash (or near cash) social transfers.....	40
2.2.10 Degree of urbanisation.....	42
2.2.11 Educational attainment level	44
2.3 Examples of applications to the poverty disaggregation	46
2.3.1 Example 1. Gender and poverty measurement: approaches and challenges.....	46
2.3.2 Example 2: Disability and more appropriate poverty measures (United Kingdom)	49
2.3.3 Example 3: Disaggregation and the use of alternative poverty measures (United States).....	50
2.3.4 Example 4: Joint disaggregation of child poverty.....	52
2.3.5 Example 5: Presentation of poverty across population groups	58
2.3.6 Example 6: Poverty status and receipt of social transfers	60
2.3.7 Example 7: Interactive presentation of poverty indicators.....	60
2.3.8 Example 8: Poverty and housing tenure (United Kingdom).....	62
2.4 Recommendations.....	63
3 Including Hard to Reach Groups in Poverty Measurement	64
3.1 Introduction.....	64
3.2 Improving inclusiveness of poverty surveys.....	65

3.2.1	Fieldwork approaches for hard to reach populations	65
3.2.2	Principles to facilitate surveying vulnerable groups	65
3.2.3	Interviewing hard to reach groups	66
3.2.4	Using simplified proxy indicators	67
3.2.5	Training and field work materials.....	68
3.3	Dedicated survey programmes for hard to reach groups	69
3.3.1	Design modifications to include hard to reach groups in standard surveys	70
3.3.2	Special survey programmes.....	72
3.3.3	Sampling target groups	74
3.4	Including qualitative information in quantitative data collections through micronarratives	82
3.5	Poverty among children	86
3.5.1	Street children	86
3.5.2	Children with disabilities	87
3.5.3	Foster children.....	87
3.6	Homelessness and housing difficulties in the European Union	88
3.6.1	Policy justification.....	89
3.6.2	How to define homelessness	90
3.6.3	Sample Issues	90
3.6.4	Implementation in EU-SILC.....	91
3.6.5	Plans for the revised EU-SILC.....	93
3.6.6	Conclusion	93
3.7	Surveying Institutional Households.....	94
3.7.1	Case Study: United Kingdom - including non-private household populations (NPHP) .	97
3.8	Recommendations.....	101
4	Assessing and Improving Survey Methods.....	103
4.1	Survey errors and quality	103
4.1.1	A typology of survey errors	103
4.1.2	Multidimensional quality frameworks	107
4.1.3	Relationship between different aspects of quality	111
4.2	Improving quality in each survey step.....	112
4.2.1	Survey design.....	112
4.2.2	Data collection.....	120
4.2.3	Data processing and imputation	124
4.2.4	Weighting	128
4.2.5	Variance estimation.....	135
4.2.6	Dissemination	137
4.3	Other methodological issues related to measuring poverty.....	140
4.3.1	Pooling.....	140
4.3.2	Small Area Estimation.....	142
4.3.3	Rapid estimates	145
4.4	Conclusion and recommendations.....	145
5	Supplemental or Experimental Poverty Measures	147
5.1	Adjustments to poverty thresholds.....	147
5.1.1	Spatial differences within countries with regard to consumption and income poverty	148
5.1.2	Household size and age of household members	154

5.1.3	Disability status/High medical expenditures	158
5.2	Adjustments to resources	159
5.2.1	Social transfers in kind (STiK)	159
5.2.2	Accounting for housing wealth.....	170
5.2.3	Asset poverty: Why assets should be considered in the measurement of poverty and how it could be done	174
5.2.4	What is the role of unequal sharing of resources within households?	180
5.2.5	Going beyond income: The role of multidimensional poverty measures	187
6	Future Work – measuring what matters	199
6.1	Pending problems in measurement of resources	199
6.2	Pending problems in measurement of needs	201
6.3	Microsimulation and impact of transfers.....	201
6.4	Dedicated surveys on hard to reach populations	202
6.5	Use of registers.....	203
6.6	Link to macro aggregates	203
6.7	Dynamics of poverty and exclusion (panel data)	205
6.8	Privacy and access to microdata	206
6.9	Comparative quality reporting	207
	References	208

List of Figures

Figure 2.1 United States example of measuring race and ethnicity.....	34
Figure 2.2 Poverty rate for older persons (age 65 and over) by sex, in all households and in one-person households, European countries, 2017	48
Figure 2.3 Proportion of families by material deprivation and disability status	50
Figure 2.4 Percentage of people in poverty by different poverty measures, 2017.....	51
Figure 2.5 Change in number of people in poverty after including each element, 2017 (in millions)	52
Figure 2.6 At risk-of-poverty rates for the total population, households with dependent children and households with three or more children	53
Figure 2.7 Number of children in Kosovo deprived as percentage, by dimension for children age 15-17 years, according to the gender of the child	54
Figure 2.8 Children in the EU Member States aged 0-17 years at risk of poverty, citizenship of their parents, 2016 (in %)	55
Figure 2.9 Venn diagram of deprivation overlaps between nutrition, water, housing, by urban and rural areas, children aged 0-4 years (Tajikistan)	56
Figure 2.10 Venn diagram of deprivation overlaps between education, water, housing, by urban and rural areas, children aged 5-10 years (Tajikistan)	56
Figure 2.11 Venn diagram of deprivation overlaps between water, sanitation, housing, by urban and rural areas, children aged 5-10 years (Tajikistan)	57
Figure 2.12 Venn diagram of deprivation overlaps between education, water, housing, by urban and rural areas, children aged 11-17 years (Tajikistan)	57
Figure 2.13 Venn diagram of deprivation overlaps between housing, sanitation, water, by urban and rural areas, children aged 11-17 years (Tajikistan)	58
Figure 2.14 Poverty among the basic sectors in the Philippines	59
Figure 2.15 Program participation status of household (per cent of population) by poverty status of people in 2017	60
Figure 2.16 At risk of poverty or social exclusion, Austria, 2014.....	61
Figure 2.17 Percentage of people (all ages) in relative low income, by housing tenure, United Kingdom, 2016-17	62
Figure 2.18 Percentage of children in relative low income, by housing tenure, United Kingdom, 2016-17	62
Figure 3.1 Micronarrative framework.....	84
Figure 3.2 Self-coded reasons to go abroad in Western Balkans (n=1136 micro-narratives)	84
Figure 3.3 Self-coded reasons to go abroad in Bosnia and Herzegovina (n=203 micro-narratives).....	85
Figure 3.4 Pattern of responses for unfavourable treatment based on gender, ethnicity and age, n=510	85
Figure 4.1 Increasing sampling error/bias ratios for disaggregations	106
Figure 4.2 EU-SILC standardised rotating scheme	119
Figure 4.3 Distribution of response Modes in the American Community Survey 2017	122
Figure 5.1 Impact of geographic adjustments on poverty rates, 2018.....	149
Figure 5.2 Regional poverty rates using subsistence minimum (RF SM) and regional subsistence minimum (Reg. SM) of the Russian Federation for three climatic regions	151
Figure 5.3 Relationship of deprivation index and income among two-person households	153
Figure 5.4 Poverty Rate with and without Care Allowances for Austria, 2008.....	159
Figure 5.5 Social transfers in kind by income quintile group, United Kingdom, 2016/17	162

Figure 5.6 Original, disposable, post-tax and final income by quintile groups, all households, financial year ending 2016/17	163
Figure 5.7 Relative at-risk-of-poverty rates, United Kingdom, 2011/12	164
Figure 5.8 GINI coefficients for income (SNA scale) with and without educational and health STiK.....	165
Figure 5.9 Low-income rates by age of household head, Canada 2016	172
Figure 5.10 Impact of Differential Thresholds on the U.S. Supplemental Poverty Measure, 2017.....	173
Figure 5.11 Illustration of the two-dimensional poverty index	175
Figure 5.12 Share of income poor that do not have sufficient liquid assets, by poverty line and reference period, in percent of the population (individuals), with 95%-confidence interval.....	176
Figure 5.13 Share of income poor that do not have sufficient liquid assets for three months, by subgroups, in percent of the population (individuals), with 95%-confidence interval	177
Figure 5.14 Share of population indicating material or subjective difficulties, by poverty status, in percent of the population (individuals), with 95%-confidence interval	178
Figure 5.15 The standard household income model – exemplary household	180
Figure 5.16 Dimensions of deprivation for households.....	191
Figure 5.17 Relationship of Monetary Poverty (below 60% of Median) and Material Deprivation among children below 16 years	195
Figure 6.1 Difference between national accounts aggregates and income micro data	205

List of Tables

Table 2.1 Educational attainment levels used in Mexico's National Household Income and Expenditure Survey.....	44
Table 3.1 Proportion of population with household income below 60% of national median	72
Table 3.2 Summary of sampling approaches in previous Roma surveys.....	79
Table 3.3 EU-SILC variables on past episodes of homelessness and housing difficulties.....	92
Table 3.4 Estimated size of each including non-private household sub-populations in the United Kingdom	98
Table 3.5 Proposed approaches to capture data for the different categories of including non-private household populations	99
Table 4.1 Concepts and categories used by different organizations to identify dimensions of quality...	107
Table 4.2 Use of administrative data and registers for each domain covered by EU-SILC (2012)	114
Table 4.3 Two overlapping phases of the American Community Survey	121
Table 4.4 Demographic control totals used by the Canadian Income Survey	132
Table 5.1 Proposal of new equivalence scales for Tajikistan.....	157
Table 5.2 At-risk-of-poverty by sex.....	184

List of Boxes

Box 2.1 Approaches to disaggregation of SDG indicators in a global framework	16
Box 2.2 National responses to improving disaggregation of SDG indicators in the United Kingdom	17
Box 2.3 Regional approaches for harmonizing variables for disaggregation across statistical domains – Eurostat’s project on standardization of social variables	18
Box 2.4 OECD survey of current practices to measure ethnicity, race and indigenous identity	19
Box 2.5 Detailed age disaggregation of poverty measures for children.....	23
Box 2.6 The Washington Group Short Set on Functioning (WG-SS)	27
Box 2.7 United States’ experience with measuring race and ethnicity	34
Box 3.1 Practical challenges in surveying respondents in areas controlled by organised crime in Mexico.....	68
Box 3.2 Surveying marginalized Roma communities in Slovakia	70
Box 3.3 Surveying the Roma, Sinti and Caminanti (RSC) populations in Italy.....	72
Box 3.4 Generating data on hard-to-reach populations at risk of violation of their rights in Bulgaria	73
Box 3.5 The first nationwide Survey on persons with disabilities in Germany.....	74
Box 3.6 Previous results and future development of statistical information about homelessness in Italy.....	77
Box 3.7 UNDP – Sampling approaches in previous Roma Surveys	79
Box 3.8 Special surveys on the Roma population in the Multiple Cluster Survey (MICS)	81
Box 3.9 UNDP’s collection of micronarratives from Roma in the Western Balkans.....	83
Box 3.10 Measuring poverty for the institutionalized population in the United States.....	95
Box 3.11 Data gaps on persons with disabilities in the European Union	96
Box 4.1 Framework to assess errors in poverty measurements (Verma et al, 2010).....	106
Box 4.2 European Union: content of quality reports as required by the law	108
Box 4.3 Use of administrative tax data at Statistics Canada.....	113
Box 4.4 Use of registers in the European Union	114
Box 4.5 Record linkage through statistical identifiers in Austria	115
Box 4.6 Design of the American Community Survey in the United States	117
Box 4.7 Design of the Canadian Income Survey	118
Box 4.8 EU-SILC Sample Design.....	119
Box 4.9 Data collection for the American Community Survey	120
Box 4.10 Imputation of Income in Albanian Survey of Income and Living Conditions.....	126
Box 4.11 Calibration in the Canadian Income Survey.....	132
Box 4.12 Calibration in Household Living Condition Survey 2009 in Ukraine.....	134
Box 4.13 Calculating variance for EU-SILC in Austria	137
Box 4.14 Suppression rules in the United States	139
Box 4.15 Small Area Income and Poverty Estimates in the United States	144
Box 5.1 Assessment of the 60% of median threshold for the elderly in Austria	152
Box 5.2 Equivalence scale based on a demand model in Tajikistan	157
Box 5.3 Approaches to the estimation of imputed rent.....	170
Box 5.4 The national Multidimensional Poverty Index (MPI) of Armenia	190
Box 5.5 Definition of the material and social deprivation measure in the European Union.....	193
Box 5.6 Definition of the child deprivation measure in the European Union	194
Box 5.7 Methodology for determining deprivation measures for the European Union	195

Acronyms / Abbreviations

ABS	Australian Bureau of Statistics
ACS	American Community Survey
ASEC	Annual Social and Economic
BEA	Bureau of Economic Analysis
BHPS	British Household Panel Survey
BPA	Beijing Platform for Action
BRR	Balanced Repeated Replication
CES	Conference of European Statisticians
CIS-STAT	Interstate Statistical Committee of the Commonwealth of Independent States
CPS ASEC	Current Population Survey Annual Social and Economic Supplement
CRC	Convention on the Rights of the Child
CRPD	Convention on the Rights of Persons with Disabilities
DIW	German Institute for Economic Research (<i>Deutsches Institut für Wirtschaftsforschung</i>)
DWP	Department for Work and Pensions
DSQ	Disability Screening Questions
ESCAP	Economic and Social Commission for Asia and the Pacific
ESS	European Statistical System
Eurostat	Statistical Office of the European Union
ECHP	European Community Household Panel Survey
EHIS	European Health Interview Survey
EU-SILC	European Union Statistics on Income and Living Conditions
FEANTSA	European Federation of National Organisations Working with the Homeless
FRA	European Union Agency for Fundamental Rights
GPSDD	Global Partnership for Sustainable Development Data
GSOEP	German Socio-Economic Panel
GQ	Group quarters
HBAI	Households below average income
HBS	Household budget survey
HRBAD	Human rights-based approach to data
IAEG-GS	Inter-Agency and Expert Group on Gender Statistics
IAEG-SDGs	Inter-Agency and Expert Group on SDG Indicators
ICF	International Classification of Functioning, Disability and Health
ICP	International Comparison Program
ILO	International Labour Organisation
ISCED	International Standard Classification of Education
ISG	Indicators Subgroup
MBM	Market Basket Measure
MEHM	Minimum European Health Module
MICS	Multiple Indicator Cluster Surveys
MODA	Multiple Overlapping Deprivation Analysis
NPHP	Non-private household populations
NSOs	National Statistical Offices
LFS	Labour Force Surveys

LAU2	Local Administrative Unit 2 (formerly NUTS 5)
LSMS	Living Standards Measurement Study
OECD	Organisation for Economic Co-operation and Development
ONS	Office for National Statistics
OHCHR	Office of the High Commissioner for Human Rights
OPHI	Oxford Poverty and Human Development Initiative
PPPs	Purchasing Power Parities
RDS	Respondent Driven Sampling
Reg. SM	Regional subsistence minimum
RELAIS	Record Linkage at ISTAT
RPPs	Regional price parities
RF SM	Russian Federation's subsistence minimum
SDG	Sustainable Development Goals
SIPP	Survey on Income and Programme Participation
STiK	Social Transfers in Kind
SPM	Supplemental Poverty Measure
UBN	Unmet basic needs
UK GSS	United Kingdom Government Statistical Service
UN	United Nations
UNDP	United Nations Development Programme
UNECE	United Nations Economic Commission for Europe
UNFPA	United Nations Population Fund
UNICEF	United Nations Children's Fund
UNSC	United Nations Statistical Commission

1 Introduction

1.1 Why this Guide?

1. The Guide on Poverty Measurement (UNECE, 2017a) laid out important standards for the measurement of poverty in the UNECE region with a view of SDG 1 “End poverty in all its forms everywhere”. The present Guide intends to provide further guidance to consolidate the measurement of poverty as well as inspiration for the practice of statistical offices.
2. First and foremost, this requires that all population groups are counted. People living in poverty are increasingly missed by household surveys, particularly if they belong to ethnic or other minorities whose life circumstances differ from the general population. Therefore, the Commission on Global poverty, chaired by Tony Atkinson recommended that: “There should be an investigation of the extent to which people are “missing” from the global poverty count, and proposals made for adjustments where appropriate at the national level for survey underrepresentation and noncoverage by surveys; ...” (World Bank, 2017, recommendation 3: 33).
3. Governments and stakeholders may have quite different views on how to eradicate poverty. Different policy approaches are typically connected to different measurement approaches, such as consumption or income-based measurement (Rio Group, 2006: 140f). This is a challenge to the first of the United Nation’s Fundamental Principles of Official Statistics which demands relevance, impartiality and equal access (UNSD, 2014). The measurement of poverty by independent statistical offices does not only serve governments. It has to consider all stakeholders involved in the political debates, including the public at large. Statistical offices provide reliable data infrastructures. They have to be neutral as to whether poverty should be reduced by redistribution through cash transfers, economic growth, and active labour market policies or addressing unmet needs directly through non-cash transfers for certain groups among the poor. But governments may formulate their targets effectively and be held accountable for it based on these data.
4. Regardless of how accurate the national measures of poverty are, their usefulness will be limited if they mask existing disparities within societies. For example, it is still a true description that some countries are poorer than others. However, it would be misleading to donors of foreign aid to assume that all people in a poor country are poor. Income distributions are overlapping between rich and poor countries. Everyone who travels can see the same cars, clothes, and food in almost any country – albeit such signs of material wealth are clearly not available for everyone. Relative poverty may be on the rise when certain segments of the population take more advantage from rapidly increasing levels of prosperity. With ongoing progress, it becomes more important to identify fault lines within countries.
5. The pledge of the 2030 Agenda for Sustainable Development is to reach the furthest behind first and leave no one behind (UN, 2015). Evidence based policies need to know who is poor and where support is needed the most. The General Assembly Resolution “Transforming our world: the 2030 Agenda for Sustainable Development” therefore calls for “high-quality, timely and reliable data disaggregated by

income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts" (A/RES/70/1).

6. The measurement and disaggregation of poverty measures provides a particular description of the social world. Amartya Sen understood description as choice (Sen, 1980). He did however reject the idea of Orshansky that deliberate decisions of poverty measurement are the result of purely subjective value judgements, resting solely in the eye of a particular "beholder" (Orshansky, 1969). It may be helpful to rather think of measuring poverty like constructing a map to be potentially used for reaching different destinations or using different paths. Human rights principles are essential for that. It must be guaranteed that vulnerable populations can participate, are protected from any harm and can benefit from the results. Further, privacy and self-identification must be respected and transparent processes must be followed (OHCHR, 2018).

7. For example, the European Union has established a transparent legal framework for the measurement of poverty in the European Union Statistics on Income and Living Conditions (EU-SILC), including access to micro data and quality reports. The framework reflects (rather than imposes) different ideas about economic growth and employment policy as much as about redistribution or social and human rights. Indicators on poverty (and their disaggregation) hence provide a common ground for rational policies with potentially opposed objectives. Its function is comparable to that of the consumer price index in a collective bargaining process which ensures that negotiations maintain a minimum of shared empirical foundation. The choices made in disaggregations have a similar function and hence are far from arbitrary.

8. Disaggregation of poverty measures may imply demanding sample size requirements and coverage (including e.g. rare, institutionalised or homeless populations). The provision of sufficient means and statistical capacity is therefore crucial. The Guide aims to inspire their best possible use but cannot substitute for the relevant legal and policy processes. This document should provide orientation for the whole UNECE region and may also be discussed by Working Groups on EU-level.

9. Disaggregation has practical consequences. Measurement of poverty may be called good if it supports the effective allocation of various resources such as food banks, medicines, teachers or law enforcement. If for example household income inadequately reflects children's needs for such resources, it is probably not the ideal measure for that group.

10. International standards provide an important point of reference for government and non-government stakeholders, especially when no specific national policy objectives are articulated (Till and Keindl, 2015). These standards need to be internationally comparable and often help to assess the need and impact of foreign aid and development partnerships. If priorities are set accordingly, limited resources can be spent efficiently. Harmonisation of disaggregated poverty measures also allows countries to benefit from each other's experience in tackling poverty and reaching further those in need. Joining efforts across countries and regions would help to identify similar patterns that would permit policymakers to search for already existing solutions and to address together challenges that trigger inequalities. Such challenges are for example related to women who tend to earn less than men, persons with no post-compulsory training who often have less employment opportunities, elderly who by and large live in poorer housing conditions, or children in rural areas who are less educated than those living in the cities.

11. The guidance offered on the definitions and operationalization of the disaggregations aims to serve as an international standard for poverty disaggregation while respecting differences in national realities or in the existence of international standards. This guidance can further be refined when results of other methodological developments become available.

1.2 Background

12. The first United Nations handbook on poverty measurement was published by the United Nations Statistical Division in 2005 (UNSD, 2005). It responded to the increased need for measuring poverty in the context of the Millennium Development Goals. The diversity of approaches around the globe was addressed in four regional workshops and was supported by a survey among statistical offices which was answered by 91 countries. Globally, measures based on consumption were found to dominate, while in the UNECE region, most responding countries were measuring poverty on the basis of income. Only a few countries reported that they measured poverty on the basis of unmet basic needs (UBN), and disaggregation was mentioned in this handbook only with reference to geographic disparities.

13. An Expert Group on Poverty Statistics (Rio Group) which involved 22 countries and 18 regional or international organisations was active between 1996 and 2006. It had found that a universally applicable handbook was difficult to accomplish given the heterogeneity of methodologies, degrees of sophistication and policy purposes. Instead, a compendium of different practices was published, which included monetary poverty as well as non-monetary measures of deprivation. Apart from geographical disaggregation, the compendium addressed the measurement of poverty among children. It particularly acknowledged that the needs of children differ in degree and kind from the adult population and thus require specific approaches to measurement.

14. Within the narrower focus of monetary poverty, the World Bank published a streamlined handbook on the poverty and inequality (Haughton and Khandker, 2009) and maintains the ADePT software to obtain standard disaggregation tables from micro data (Foster et al, 2013). The Commission on Global Poverty led by Sir Tony Atkinson raised a number of caveats concerning the monetary approach which were summarized in 21 recommendations. The recommendations include the use of a multidimensional poverty indicator as a complement to the \$1.90/day measure (World Bank, 2017). Such a global multidimensional indicator was developed by the Oxford Poverty and Human Development Initiative (OPHI) and UNDP. It is regularly published since the first publication of the Multidimensional Poverty Index (MPI) of the Human Development Report in 2010 (e.g., UNDP, 2019). A handbook on how to implement national versions of the global MPI has recently been published (UNDP and OPHI, 2019).

15. Against the background of these mainly global initiatives, the present Guide focuses on the practical approaches currently followed in statistical offices of countries participating in the Conference of European Statisticians (CES)¹. Most UNECE countries (34) already participate in the collection of EU-

¹ The Conference of European Statisticians is composed of national statistical organizations in the UNECE region (for UNECE member countries, see http://www.unece.org/oes/nutshell/member_states_representatives.html)

SILC. This statistical instrument has been designed explicitly to provide harmonised indicators for the coordination of EU policy processes. Therefore, the European Union has also set up a so-called Indicators Subgroup (ISG) which is a special committee of statistical and social policy experts for the continuous development of indicators to guide policies on social inclusion.²

16. EU-SILC is based on a common legal framework and reflects fundamental methodological decisions for the measurement of poverty which were originally proposed by Tony Atkinson et al (Atkinson et al, 2002). Eurostat updates comprehensive guidelines (Doc65) for the implementation of every new round of the EU-SILC data collection.³ Quality reports which include information on how closely these implementation guidelines were followed are published regularly both by Eurostat as well as the Member States.⁴ The continuous evolution of indicators and methodologies are also documented in two important Eurostat publications (Atkinson and Marlier, 2011; Atkinson et al, 2017).

17. The UNECE Statistical Division has been involved in poverty statistics since 2012. The activities started with an in-depth review of this topic prepared by the State Statistics Service of Ukraine and Eurostat for the Bureau of the Conference of European Statisticians (CES Bureau). Since then, expert meetings and capacity-development workshops on poverty measurement were held regularly on an annual basis.

18. During these meetings, the Interstate Statistical Committee of the Commonwealth of Independent States (CIS-STAT) - among other participants - regularly presented updated information about the progress made in the 11 CIS countries. The latest of these reports noted that national SDG platforms with roadmaps, lists of indicators as well as data and metadata became increasingly available for certain indicators (CIS-STAT, 2019: 2). Nonetheless, common regional standards to obtain comparable measures for monetary or multidimensional poverty are yet to be established (CIS-STAT, 2019: 5). Neither the global poverty line nor the global MPI seem appropriate for the context of rapidly developing middle income countries. Most CIS countries already publish disaggregated poverty measures (CIS-STAT, 2019: 6) but obtaining sufficiently large samples for representative data and especially the definition of persons with disability remains problematic (CIS-STAT, 2019: 8).

19. In addition, to the expert meetings, task forces mandated by the CES and composed of national statistical offices and international organizations have been working on poverty-related issues. A first such task force produced and published in 2017 the Guide on Poverty Measurement (UNECE, 2017a). The Guide

and includes in addition Australia, Brazil, Chile, Colombia, Japan, Mexico, Mongolia, New Zealand and Republic of Korea. The major international organizations active in statistics in the UNECE region also participate in the work, such as the statistical office of the European Commission (Eurostat), the Organization for Economic Cooperation and Development (OECD), the Interstate Statistical Committee of the Commonwealth of the Independent States (CIS-STAT), the International Monetary Fund (IMF) and the World Bank.

² <https://ec.europa.eu/social/main.jsp?catId=830&langId=en>

³

https://circabc.europa.eu/faces/jsp/extension/wai/navigation/container.jsp?FormPrincipal: idcl=FormPrincipal: id1&FormPrincipal_SUBMIT=1&id=334d943f-6f71-4f4b-9c7e-a6767a3fe164&javax.faces.ViewState=JwW05I73o2Dvufk86FDq%2BI3s4dGhWgljenexRTXh12jKnzZh01ftwoaJz0VG7IDMf6D417EEa00cQbME8b6E1QYt3Tc3T3fGOMsc0h1TS7UyeOOeZALTRVdnwwwkzWTfRms%2FUBEdRfo1wdTcxQgqZGp%2BVU%3D

⁴ <https://ec.europa.eu/eurostat/web/income-and-living-conditions/quality/eu-and-national-quality-reports>

provided an overview of definitions, data requirements and measurement issues related to various approaches for poverty measurement and made recommendations for improving the international comparability of poverty statistics. The Guide also highlighted methodological challenges such as measurement of resource sharing within households and coverage of hard-to-reach populations and the population living in institutions. It was stressed that comparability of poverty measures across population groups will depend on other factors taken into consideration that go beyond pure income or consumption measures at the national level, for example on the inclusion of assets and housing wealth or social transfers in kind. The Guide advised additional work to be conducted on several topics, where no clear-cut recommendations were made due to insufficient evidence from the current practices, such as hard-to-reach populations, imputed rent and housing cost, individual level poverty measures, wealth, spatial differences with regard to income and consumption poverty, subjective poverty and comparability of welfare aggregates.

20. In February 2017, the CES Bureau established the Steering Group on Measuring Poverty and Inequality to provide direction to the CES work on poverty and inequality statistics. Based on the challenges identified in the Guide on Poverty Measurement and in consultation with countries' experts, the Steering Group recommended establishing a task force to take further the methodological work and develop recommendations to countries on data disaggregation for poverty measurement. The CES Bureau approved the task force's terms of reference in 2017. The task force consisted of 30 statistical experts from UNECE member countries, other developed countries participating in the work of the Conference of European Statisticians, international organizations and academia. The experts in poverty measurement worked through 2018-2019 to develop the present Guide on data disaggregation.

21. The Guide does not recommend a particular definition of poverty. The objective of the Guide is to consolidate current and emerging good practices in disaggregating poverty indicators and in assessing their robustness. The Guide provides methodological and practical guidance on data disaggregation to improve poverty-related measures at a national level and further enhancing international harmonisation. The Guide takes into consideration the SDG reporting needs and already available international guidance on data disaggregation. Drawing on the experiences of countries in the UNECE region, the Guide includes specific recommendations to national statistical offices and sets minimum standards on disaggregation of poverty measures that could serve as guidance.

1.3 Outline of the Guide

22. The ethical and practical purpose of disaggregating poverty measures has at least four implications that define the structure of this Guide. Firstly, there should be clarity about the definition of the groups to be considered. Secondly, every effort must be taken to ensure coverage of groups that are most vulnerable in data collection. Thirdly, it is paramount to continuously assess and document the quality of the poverty measures. Fourthly, poverty profiles obtained from conventional measures should be compared to profiles that would be obtained from supplemental measures to assess their robustness.

Where this appears necessary, complementary measures need to be published together with conventional measures.

23. **Chapter 2** provides methodological and practical guidance on standard set of core variables for poverty disaggregation with the purpose to enhance further international harmonisation. The chapter offers recommendations to improve the quality, relevance, and use of data, consistent with international human rights norms and principles for the identification of policy-relevant groups. It suggests a set of essential variables with background information, concepts and definitions of the variable and related categories, and implementation guidelines. Reference questions are provided to illustrate the implementation of data disaggregation in practice.

24. Some vulnerable or disadvantaged groups are particularly hard to count in general population surveys – either because they are hard to reach, such as the homeless, undocumented migrants, members of ethnic minorities or older people living in institutions, or because they are hard to identify, such as gender minorities. **Chapter 3** looks at how to adapt data collection to fulfil the 'no-one left behind' commitment and reach beyond the traditional and established survey methodologies to capture those groups, that are most exposed to the risk of poverty.

25. To establish trust in poverty measurement and prevent misguided policies, statistical offices have to regularly assess and continuously improve the quality of their processes and accuracy of their data. Quality reports which describe the quality criteria and explain any instances in which these criteria could not meet, or statistical concepts could not be correctly applied will not only assist the correct interpretation but can also provide the basis for future improvements. **Chapter 4** provides analysis of how response rates and sampling precision may be improved in the measurement of poverty among relevant social groups. The chapter contains practical recommendations for national statistical offices on controlling sampling and non-sampling errors for small domains, weighting and improving coverage for hard-to-reach population groups.

26. Any methodology for measuring poverty relies on several assumptions and a multitude of decisions about how to set the poverty threshold, how to define resources and how to implement any particular methodology. While it is impossible to avoid these assumptions or “arbitrary” decisions, countries are developing supplemental or experimental poverty measures to assess their impact and improve the accuracy of the measurement. **Chapter 5** describes individual country experiences with adjustments for difference in cost of living across regions and household circumstances and measurement issues such as social transfers in kind, disability cost or high medical expenditures, housing wealth, imputed rent, assets poverty and unequal sharing of resources within households as well as multi-dimensional poverty measures.

27. An outlook on pending issues for future work is presented in **Chapter 6**.

1.4 Summary of recommendations

1.4.1 Chapter 2: Standard core variables for disaggregation

28. **Recommendation 1:** In line with the guidance provided by the Office of the United Nations High Commissioner for Human Rights, development and use of measures for disaggregation of poverty should follow the Human rights-based approach to data with emphasis on principles of self-identification and participation.

29. **Recommendation 2:**

- a) In line with the Guide on Poverty Measurement and considering other emerging needs, especially in the context of SDGs, groups which are defined by the following variables are suggested as a priority for poverty disaggregation, wherever relevant and permitted by national legislation: age, sex, disability status, migratory status and ethnicity.
- b) In addition, the following disaggregation variables are suggested to better understand socio-economic processes and effectiveness of policies: household type, employment status, tenure status of the household, receipt of social transfers, educational attainment and degree of urbanisation. These minimum suggestions should be complemented by additional or more detailed disaggregations as required by the national context.

1.4.2 Chapter 3: Including hard to reach groups in poverty measurement

30. **Recommendation 3:** Countries are encouraged to first identify, and then quantify, parts of the population not sufficiently covered in their poverty statistics. This assessment should especially address ethnic minorities, persons who are homeless, live in institutional households or have disabilities. They should further explore the methodological and empirical causes and challenges of why these or other populations may be underrepresented or not covered.

31. **Recommendation 4:** Countries should develop suitable methods to facilitate participation of hard-to-reach households in their survey programmes. The principle of self-identification or the use of survey instruments in different languages are simple examples for some basic measures, which aim to establish trust and improve accessibility.

32. **Recommendation 5:** If it is not feasible to include hard-to-reach groups on existing survey programmes, countries should develop targeted surveys to collect data on poverty and social exclusion specific to the groups, to be run at least every 5 years.

33. **Recommendation 6:** Censuses and large-scale surveys should include questions on self-identification of ethnicity, which allow for multiple identities. It helps to establish sampling frames which are needed to target ethnic minorities and hard to reach groups. This requirement is critical for the

objective of leaving no one behind and to disaggregate survey data and poverty statistics for vulnerable groups.

34. **Recommendation 7:** Interviewer training for surveying minority groups should address cultural and group-specific aspects. Persons of the target populations should be involved in the set-up and development of fieldwork materials.

35. **Recommendation 8:** Further research should be directed at creating supplementary measures of poverty to reflect that the living circumstances of target groups do not always fit with standard household definitions.

36. **Recommendation 9:** Further research is needed to explore the possibilities for alternative sampling approaches, such as non-probability designs including online surveys, to measure poverty in hard-to-reach populations.

1.4.3 Chapter 4: Assessing and improving survey methods

37. **Recommendation 10:** Document coverage, precision and any departure from international standards in quality reports to be issued with each data release.

38. **Recommendation 11:** Strictly adhere to probability sampling and refrain from substitution. Alternative practices should be reserved exclusively for the purpose of representing specific population groups which otherwise could not be reached with capacities normally available for a developed statistical system.

39. **Recommendation 12:** Adjust the sample design and allocation for the required disaggregation. As a minimum, the sample size for each group should never be below 50 units. Ideally, estimation domains will be used as strata so that their size may be controlled.

40. **Recommendation 13:** Where available, consider the use of administrative or register data since it can limit response burden and increase accuracy.

41. **Recommendation 14:** Establish a clear collection protocol and monitor field work to ensure that it is respected. This should include sufficient interviewer training, ensuring a sufficient number of contact attempts and procedures to address barriers to survey response for at-risk groups (for instance, translations to address language barriers).

42. **Recommendation 15:** Ensure that vulnerable populations participate in the design of the study, are protected from any harm and have access to its results.

43. **Recommendation 16:** Minimize the impact of item non-response through imputation. Because poverty status is usually derived from several variables that have to be observed simultaneously; it is essential that group representation be preserved by adequate and transparent imputation rules.

44. **Recommendation 17:** Compensate for unit non-response and other sources of variation by applying a non-response adjustment and calibration. Ensure that these adjustments remain reasonable

by imposing a maximum and minimum factor by which weights are multiplied (for instance between 1/3 and 3).

45. **Recommendation 18:** Carefully assess the variance of all key estimates. Publish confidence intervals or coefficients of variation with the survey estimates and in quality reports so their precision can be compared to other surveys and over time.

46. **Recommendation 19:** Establish filtering rules for publications that eliminate unreliable estimates. Do not publish results based on fewer than 20 observations. Similarly, do not publish results with a coefficient of variation that is too high or a confidence interval that is too wide.

1.4.4 Chapter 5: Supplemental or experimental poverty measures

47. **Recommendation 20:** As a general rule, measurement of monetary poverty has to ensure equivalence of standards of living for all regions and groups within a country, notably with regard to the needs of children/ persons with health impairments or disabilities. When disaggregating poverty estimates for smaller geographies, methodologies that take into account spatial differences in the cost of living are important, particularly in large countries with heterogeneous cultural norms and price levels. Relative thresholds such as those used in the European Union are a convenient practical alternative.

48. **Recommendation 21:**

- a) Equivalence of measurement should be assessed empirically. Sensitivity analysis can compare poverty profiles of official poverty measures with supplemental poverty measures which may be either monetary (based on income, expenditure or assets) or nonmonetary.
- b) The sensitivity of monetary poverty measures should be regularly assessed using alternative equivalence scales. While simple equivalence scales based on the square root of family size are useful for international comparability, countries should explore the empirical development of equivalence scales that go beyond family size to include differences in needs based on age, disability status and health.
- c) If Social Transfers in Kind (STiK) are included in resources, special consideration needs to be given to appropriate equivalence scales.

49. **Recommendation 22:** Further research on the special needs of persons with disabilities is necessary to ensure these are taken into account in monetary poverty measures. These adjustments may be done on the thresholds, perhaps through equivalence scales, or on the resource-side of the measure.

50. **Recommendation 23:**

- a) While the measurement of Social Transfers in Kind (STiK) continues to pose serious challenges, it is important to develop a mechanism to take them into account when estimating poverty and the impact of these transfers on poverty estimates. Supplemental or alternative poverty measures are important tools for illustrating the impact of these transfers on economic well-being. STiK can

be particularly relevant for comparisons between different welfare systems, where STiK are more important than cash transfers in one country (or group) than another.

- b) Figures on total STiK should be presented together with poverty measures wherever possible as a useful indicator in its own right.
- c) STiK should be included in the measurement of poverty if their value can be empirically estimated on household or individual level with sufficient precision. Particularly relevant for poverty measurement are STiKs for food, shelter, clothing, and utilities. Some countries also make provisions for health care and education. If STiK are included in the resource measure this may affect the equivalence scale (see Recommendation 21 above).
- d) If poverty headcounts of relevant groups would change by 10% after STiK some consideration in the poverty measure is highly advisable. If however measurement is very poor or its effect on poverty profiles is within the margin of sampling error, STiK should not be included in poverty measures.
- e) Given the unavoidable and essentially arbitrary methodological choices regarding valuation and distribution of STiK, these need to be made fully transparent in regularly updated quality reports. In any case, users should be given the possibility to assess poverty measures with and without adjustments for STiK.
- f) STiK can be valued at equivalent insurance cost or actual consumption or as a mix. Its total value and estimated number of recipients need to be assessed against administrative data on the total public cost on STiK.
- g) Caution is needed when analysing STiK and their potential distributive consequences. If STiK are included in the resource measure, its value needs to be capped and by all means should not exceed the poverty threshold.
- h) If the value of STiK received is too difficult to obtain, the deduction of out-of-pocket expenses from the resource measure are a viable alternative. In such a situation however some poor individuals who have already curtailed certain expenditure may eventually appear as non-poor.

51. **Recommendation 24:**

- a) Given that the choice of method for obtaining monetary poverty measures may especially depend on the available data and may contain essentially arbitrary elements, methodological choices need to be made fully transparent in regularly updated quality reports. In any case, users should be given the possibility to assess poverty measures with and without adjustments for imputed rent.
- b) As an alternative to imputing rent, residual income after actual housing cost may be considered as a resource measure which reflects housing equity.
- c) Surveys should consider adding questions to measure housing costs and property values and other characteristics of the dwelling, which facilitate the calculation of imputed rent.

52. **Recommendation 25:** Assets are an important resource for determining well-being and countries should continue to experiment with ways to measure both income and asset poverty. At this point, the two-dimensional approach with a separate asset-based poverty rate calculated in addition to income poverty seems to be the easiest to communicate to the public.

53. **Recommendation 26:**

- a) In the immediate term, official poverty rates, collected at the household level, should be routinely disaggregated by sex and age of each household member. This shall provide estimates for the poverty rate of prime age women who have children and/or live without a partner (e.g. in lone parent households).
- b) More research is needed to determine how best to address intra-household resource sharing in order to develop sub-household measures of poverty status. This is particularly important for the disaggregation of poverty estimates by sex and age. Surveys should continue to experiment with questions designed to determine individual control of resources and to measure material deprivation at the person level.
- c) Where possible, official poverty rates for men, women, children or other socio-demographic subgroups should be accompanied by results that consider unequal sharing of resources. At a minimum, national statistical offices should carry out sensitivity analysis for poverty profiles contrasting the conventional full pooling assumption with partial pooling and full separation of resources.
- d) To validate assumptions on within-household distribution of economic resources, these should be considered in combination with material living standards, wherever possible. Questions on sharing of personal economic resources and/or personal material living standard have been used in EU-SILC. Before such questions are adapted it should however be ensured by pre-tests that sharing of resources is not considered as sensitive by respondents.

54. **Recommendation 27:** Following the SDGs, countries should implement multidimensional poverty measures to complement existing monetary measures of poverty. These multidimensional measures should be tailored to the national context and policy priorities, and be tracked over time.

55. **Recommendation 28:**

- a) Deprivation measures need to be based upon a clear and explicit theory or normative definition of poverty in order to ensure that each indicator is a valid measure, i.e. that it measures poverty and not some other related (or unrelated) concept such as wellbeing or happiness.
- b) The validity of each indicator should be demonstrated, i.e. the amount of systematic error should be formally assessed and indicators should be dropped if they have a low validity.
- c) The reliability of each indicator should be determined, i.e. the amount of random error should be formally assessed and indicators should be dropped if they have a low reliability.

- d) Normally, a deprivation or poverty index should count each item with an equal weight. Differential weights should only ever be used when this reduces measurement error (i.e. validity and/or reliability of the index are improved).

56. **Recommendation 29:**

- a) Countries should use available datasets, such as Multiple Indicator Cluster Surveys (MICS) or household surveys to develop child-specific and life-cycle adapted multidimensional poverty measures that reflect the needs of children at different stages of development and allow for identification of intra-household differences between children. This type of measure can be performed at intervals of 3-5 years to complement more frequent disaggregated national measures, as this will give greater insight into childhood and adolescent poverty.
- b) In order to enhance availability and use of child poverty data, countries should collect data on all key dimensions related to children's rights, including health and nutrition, and introduce lifecycle appropriate indicators to measure the situation of each child in the household. Countries should consider introducing innovative ways to collect, monitor and report on child poverty data, including ways to encourage child participation in the monitoring and discussion of child poverty data and potential policy responses.
- c) In accordance with national definitions of monetary and multidimensional poverty, countries should revise and adopt survey tools to best serve their national needs for poverty measurements. Both household budget surveys (HBS) and MICS are flexible and can be adapted to reflect a national context, but without compromising cross-country comparability. MICS offers the potential to obtain data on a broader list of child-focused indicators that can be used to measure multidimensional child poverty.
- d) Statistical data is an important source for evidence-based decision making by policy makers, not only at national but also at regional and international level. Therefore, it is important to make anonymised statistical data openly available for all users. To meet data protection and confidentiality concerns, efforts must be taken to ensure that the identity of respondents is not disclosed or can be inferred. Hence countries should make all poverty-related data, including micro-data, publicly available and easily accessible for scientific research and production purposes. This would enhance research, policy design and policy innovation in this field, which is of utmost importance for devising policies for poverty reduction.

2 Standard Core Variables for Disaggregation

57. This section provides methodological and practical guidance on standard core variables to enhance further international harmonisation of disaggregation of poverty-related measures. Section 2.1 provides background for identification of potential policy-relevant groups for which disaggregation of poverty measures is needed and an overview of respective international and national initiatives to serve as basis for defining standard core variables. Section 2.2 guides definitions and operationalization of standard core variables recommended for purposes of disaggregation. Section 2.3 provides practical examples of applications of disaggregation of poverty measures and refers also to methodological and practical challenges.

2.1 Policy relevant target groups for poverty disaggregation

58. The main purpose of the disaggregation of poverty measures is to enable the identification and monitoring of policy-relevant target groups: *“Understanding which groups experience the highest levels of poverty is important for targeting policy interventions effectively.”* (UNECE, 2017a)

59. The selection of target groups for poverty disaggregation should be consistent with international human rights norms and principles as highlighted in the general principle of *“Leaving No One Behind”* in promoting and monitoring global development under the Sustainable Development Goals (UN, 2015). Specific guidance on implementing this principle in statistical work is provided in **a human rights-based approach to data** (OHCHR, 2018). It focuses, for example, on issues such as how to improve the quality, relevance, and use of data and statistics consistently with international human rights norms and principles to support monitoring of progress towards the SDGs. According to this approach *“data disaggregation and collection which allow for comparison of different population groups is central and forms part of States’ human rights obligations”*. It further recommends *“focusing on the most disadvantaged or marginalized and on inequalities among the population”* and gives examples of relevant variables for data collection and disaggregation: *“sex, age, ethnicity, migration or displacement status, disability, religion, civil status, income, sexual orientation, and gender identity.”* For improving the quality of data on target groups and ensure that data is collected ethically appropriate, the principles of self-identification (freedom to self-identify) and participation (involvement of all relevant stakeholders) are arguably the most important. Self-identification means that statistical offices respect the freedom of respondents to choose the ethnicity for which they wish to be counted. This holds especially if this may perhaps imply an undercount against methods which would use “objective” criteria. Likewise, participation means for example that collection of ethnicity data is coordinated with ethnic communities, in particular concerning the names by which ethnic communities wish to be referred to in a questionnaire. Both examples also influence the level of trust and cooperation that is essential for any successful fieldwork.

60. With the selection of relevant population groups, those characterised by *“features exogenous to the individual effort (ethnicity, for example) have special normative significance.... poverty differences across such groupings capture structural unfairness of economic processes in a society”* (World Bank,

2017). To better characterise these processes, it is essential that disaggregation of poverty measures based on grounds of discrimination are complemented with more detailed disaggregation by socioeconomic strata and geographic areas.

61. Guidance on target groups and variables for poverty disaggregation presented in this section builds primarily upon international recommendations for disaggregation of income and poverty-related indicators from the following sources:

- **UNECE Guide on Poverty Measurement**, which provides in Recommendation 4 a minimum set of variables for disaggregation: age, sex, employment status, household type, disability status, and urban/rural population (UNECE, 2017: 33-34);
- **World Bank's Monitoring Global Poverty report**, which states in Recommendation 13: *"The global poverty figure, and the counterpart national figures, should be accompanied by the numbers of women, children, and young adults living in households with consumption below the International Poverty Line, as well as the number of female-headed households below the International Poverty Line"* (World Bank, 2017). The report also discusses possibilities in sub-national poverty measurement regarding rural/urban and regional areas, and socioeconomic groupings such as formal or informal sector, ethnicity, religion, and caste.
- **Canberra Group Handbook on Household Income Statistics**, which suggests categorizing households for purposes of income distribution statistics by: *"a. household size and composition based on characteristics such as age and sex, marital status, number of dependent children, dependency ratio, number of income earners, the main source of income; b. housing status, e.g., ownership of a home, access to subsidised housing, market renters; c. physical location, e.g., urban, non-urban, region."* (UNECE, 2011). Individual characteristics such as gender or employment status are also mentioned.

62. In addition to the income and poverty-related recommendations, the specific disaggregation variables defined in the SDG global indicator framework (income, sex, age, race, ethnicity, migratory status, disability and geographic location) and additional requests as defined in the terms of reference of the Task Force on Disaggregated Poverty Measures (housing tenure and receipt of social transfers) determine the scope of the specific guidance provided further in this section.

63. To describe target groups covered in this Guide, the following disaggregation variables should be considered for all members of a household:

- **Sex** (target group of women and girls);
- **Age** (target groups of children, youth and older people);
- **Disability status** (target group of persons with disabilities);
- **Migratory status** (target group of migrant population);
- **Ethnicity** (target groups of ethnic minorities);

64. In addition, variables referring to socioeconomic and geographic strata may be considered as follows:

- **Household type** (characteristics of household composition);
- **Educational attainment level** (characteristics of qualification and social status);
- **Employment status** (characteristics of labour force participation);
- **Tenure status of the household** (characteristics of an arrangement of occupancy of housing unit by a private household);
- **Receipt of social transfers** (characteristics of income composition);
- **Degree of urbanisation** (characteristics related to urban/rural composition).

65. For analytic purposes, combinations of two or more of these variables (especially age and sex) are often helpful to identify drivers and specific vulnerabilities across the life cycle. The list of variables is also consistent with the minimum set of dimensions for disaggregation used by the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs) in their background paper to the fiftieth session of the UN Statistical Commission (UN, 2019a).⁵

66. It should be stressed that the list of selected variables for disaggregation of poverty measures presented in this chapter is not exhaustive. Examples of target groups that are not covered are homeless people, drug users, sex workers, refugees or undocumented immigrants. As these groups are usually not well covered by regular data collections, they are focused on in other parts of this Guide, such as in Chapter 3.

67. The basis for defining standard core variables for poverty disaggregation is currently available guidance. Examples of international and national initiatives which aim to improve standards for data disaggregation from different perspectives are:

- **Global SDG indicators framework:** The work of the IAEG-SDGs⁶ (see Box 2.1); and a United Kingdom example of national initiatives related to improving the disaggregation in SDG framework (see Box 2.2);
- **Censuses:** Guidance for population and housing censuses on a global or regional level (UNFPA, 2014; UN, 2017a; UNECE, 2015a);
- **Sample surveys:** The Intersecretariat Working Group on Household Surveys⁷ established by the United Nations Statistical Commission (UNSC), the International Household Survey Network⁸, or Eurostat project on standardization of social variables (see Box 2.3);

⁵ <https://unstats.un.org/sdgs/iaeg-sdgs/disaggregation>

⁶ <https://unstats.un.org/sdgs/iaeg-sdgs/>

⁷ <https://unstats.un.org/iswghs/>

⁸ <http://www.ihsn.org/>

- International **sample survey programmes or data collections with an important income or poverty component**: European Statistics on Income and Living Conditions (EU-SILC)⁹ or Living Standards Measurement Study (LSMS);¹⁰
- On-going work of various **UN** bodies related to disaggregation **in specific domains**, such as in:
 - **Ageing**: UNECE recommendations on ageing-related statistics (UNECE, 2016) and UNECE Task Force on Measuring Old-age Population in Institutions;¹¹ and a newly established Titchfield City Group on Ageing and Age-disaggregated data (UN, 2017b);
 - **Disability**: Activities of the UN Statistics Division, the World Health Organization and the Washington Group on Disability Statistics (UN, 2017c);
 - **Migration**: On-going activities of the Expert Group on Refugee and Internally Displaced Persons Statistics (UN, 2017d), and the United Nations Expert Group on Migration Statistics (UN, 2017e);
 - **Gender**: The Global Gender Statistics Programme implemented by the UN Statistics Division and coordinated by the Inter-Agency and Expert Group on Gender Statistics (IAEG-GS),¹² and the work of the UNECE on gender statistics;¹³
- On-going work of other international organisations: OECD work related to measuring selected aspects of diversity (see Box 2.4).

Box 2.1

Approaches to disaggregation of SDG indicators in a global framework

Resolution 70/1 on the 2030 Agenda for Sustainable Development adopted by the General Assembly in 2015 (UN, 2017f) emphasized the need for “*quality, accessible, timely and reliable disaggregated data to help with the measurement of progress and to ensure that no one is left behind.*” Such data is key to decision-making. Furthermore, the Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs) was tasked to develop and implement the global indicator framework for the goals and targets of the 2030 Agenda.

Disaggregation of SDG indicators is addressed in various ways:

- By defining specific breakdowns for most indicators;
- By suggesting a list of breakdowns for general use: “*The Sustainable Development Goal indicators should be disaggregated, where relevant, by income, sex, age, race, ethnicity, migratory status, disability, and geographic location, or other characteristics, in accordance with the Fundamental Principles of Official Statistics.*” (General Assembly Resolution 68/261);

⁹ <http://ec.europa.eu/eurostat/web/income-and-living-conditions/overview>

¹⁰ <http://surveys.worldbank.org/lsm>

¹¹ <http://www.unece.org/statistics/networks-of-experts/task-force-on-measuring-old-age-population-in-institutions.html>

¹² <https://unstats.un.org/unsd/demographic-social/gender/>

¹³ <http://www.unece.org/statistics/areas-of-work/statssoc/gender-statistics.html>

- In target 18 under Goal 17 on data, monitoring and accountability: *“By 2020, enhance capacity-building support to developing countries, including for least developed countries and small island developing States, to increase significantly the availability of high-quality, timely and reliable data disaggregated by income, gender, age, race, ethnicity, migratory status, disability, geographic location and other characteristics relevant in national contexts”*.

With respect to poverty-related SDG indicators under Goal 1. *End poverty in all its forms everywhere* and Goal 10. *Reduce inequality within and among countries*, breakdowns on a global level are or will be defined.¹ The Guide on Poverty Measurement further provides recommendations for disaggregation of SDG monetary poverty indicators for ECE countries (UNECE, 2017a: Section 3.5.2).

As part of the work the global indicator framework, the IAEG-SDGs has been developing further guidance on data disaggregation via a work stream on data disaggregation. A respective Expert Group developed a document containing a stocktaking analysis of the availability of data disaggregation in the global SDGs database and an overview of existing and possible standards of disaggregation. The Group also identified key challenges on which to focus.

The IAEG-SDGs was asked by the UNSC to clearly determine the dimensions and categories of data disaggregation required for the indicators in the framework and prepared a background document on data disaggregation for consideration by the UNSC at its fiftieth session in 2019 (UN, 2019). Also, IAEG-SDG's Working Group on Geospatial Information is mandated with tasks related to data disaggregation, i.e. *“providing granularity and disaggregation of the indicators where appropriate”* and *“proposing strategies for undertaking methodological work on specific areas for improving disaggregation by geographic location concepts for national and sub-national reporting”* (UN, 2016b).

¹<https://unstats.un.org/sdgs/metadata/>

Box 2.2

National responses to improving disaggregation of SDG indicators in the United Kingdom

As part of the United Kingdom's commitment to the “Leaving No One Behind” agenda, the Office for National Statistics (ONS) is supporting the United Kingdom's Department for International Development to develop a new global **Inclusive Data Charter**¹ together with the Global Partnership for Sustainable Development Data (GPSDD) and their partners. The Inclusive Data Charter was launched at the High-Level Political Forum in July 2018 and then at the World Data Forum in October 2018. This multi-stakeholder charter recognised the need to build data disaggregation into country systems including civil and vital registration systems, administrative data systems and censuses.

In addition, the SDG and Migration Statistics teams at the ONS are working in collaboration to develop alternative data sources to improve the evidence base for international migration. This work is part of an ambitious programme, across the United Kingdom Government Statistical Service (UK GSS), looking to improving **international migration data**, particularly around the impact of international migration on society and the economy, at national and local levels. These developments will assist in the disaggregation of some of the SDGs.

In an open consultation,² users responded that ONS should pay special attention to **geographic breakdowns**. Therefore, where possible, it is an ambition of ONS to go down to small area statistical geographies as well as breaking down by geographic classifications such as rural or urban or the local authority classification. ONS are collaborating with geography experts to develop new data sources or

methods to allow them to fill geographic data gaps – using earth observation, satellite imagery, and geospatial data.

¹ <http://www.data4sdgs.org/initiatives/inclusive-data-charter>

² <https://consultations.ons.gov.uk/sustainable-development-goals/ons-approach-to-measuring-reporting-sdgs-in-the-uk/>

Box 2.3

Regional approaches for harmonizing variables for disaggregation across statistical domains – Eurostat’s project on standardization of social variables

The current Eurostat project on social variables’ standardisation follows-up on previous efforts on promoting the use of a number of standardised variables in all social surveys conducted within the European Statistical System (ESS) (Eurostat, 2011; Eurostat, 2018c). This project is closely related to activities on the modernisation of European social statistics and is expected to contribute to efficiency gains in statistical processes such as streamlining of surveys, data matching, and improved capacity to cross tabulate data.

The project aims at developing standard descriptions for the key variables which are present in at least two of the European social micro-data collections concerning households/persons¹ and related definitions (e.g. definitions of a private household, institutional households, concepts of ‘sharing household expenses’ versus ‘house-related expenses’ to distinguish between a multi-person household as oppose to a dwelling shared by housemates/flatmates). The list includes 38 variables, 18 of which core variables – to be implemented in all the European social surveys – and 20 additional variables that are common to at least two (but not all) social surveys.

The total list of 38 variables divided by topics is as follows:

- **Demographic characteristics:** Sex, Age in completed years, Partners living in the same household, Household grid, Household size, Household type;
- **Geographic characteristics:** Country of residence, Region of residence, Degree of urbanization;
- **Migration characteristics:** Country of birth, Country of main citizenship, Country of birth of the father, Country of birth of the mother, Duration of stay in the country of residence in completed years;
- **Labour market characteristics:** Main activity status (self-defined), Full- or part-time main job (self-defined), Permanency of main job, Status in employment in main job, Economic activity of the local unit for main job, Occupation in main job, Existence of previous employment experience, Size of the local unit for main job, Supervisory responsibilities in main job, Year in which the person started working for current employer or as self-employed in main job;
- **Education and training characteristics:** Educational attainment level, Participation in formal education and training (student or apprentice) in <reference period>, Level of the current/most recent formal education or training activity, Year when the highest level of education was successfully completed, Field of the highest level of education successfully completed;
- **Income and living conditions characteristics:** Net current monthly household income, Tenure status of the household;

- **Health characteristics:** Self-perceived general health, Long-standing health problem, Limitation in activities because of health problems;
- **Technical survey variables:** Interviewing mode used, Nature of participation in the survey, Stratum, Primary Sampling Unit.

For each variable the following common information is provided: name of the variable, scope (micro-data collection/survey concerned), variable definition (reporting unit, filter, and concept), category concept, categories for the variable, implementation guidelines, and reference question (optional).

¹ Statistics on Income and Living Conditions, Labour Force Survey, Household Budget Survey, Adult Education Survey, European Health Interview Survey, Harmonised European Time Use Survey, and Survey on Information and Communication Technology usage in households.

Box 2.4

OECD survey of current practices to measure ethnicity, race and indigenous identity

On November 2017, the OECD Statistics and Data Directorate conducted a survey of national statistical offices (NSOs) in OECD and key partner countries to assess current diversity data collection practices in the areas of race and ethnicity, indigenous identity, as well as migration status (Balestra and Fleischer, 2018). All OECD statistical offices, except for Estonia, France, Italy, Republic of Korea, and the United States, have participated, and completed questionnaires have been received from Bulgaria, Colombia, Costa Rica, Romania and the Russian Federation. The issues covered in the survey include relevant legal frameworks that underpin diversity data collection, available data sources, ways of identification, and detailed question wording and respondent instructions. The main findings of the survey are as follows:

- Virtually all countries which participated in the survey collected information on imperfect diversity proxies such as country of birth;
- A small majority, mostly Eastern European countries as well as the United Kingdom and Ireland, gather additional information on race and ethnicity;
- Only a handful of countries in the Americas and Oceania also collect data on indigenous identity.

Migration statistics are collected in all countries except Japan. Out of the countries for which information is available, 37 gather data on country of birth, 30 on the country of birth of the respondent's parents, 16 on the year of arrival in the country, and 16 on the mother tongue or main languages spoken by the person surveyed in at least one data source. Information on migrant status is generally self-reported via sample surveys or censuses.

Among all countries participating in the OECD survey, 17 collect official statistics on ethnicity and 8 on race, using different measurement approaches. Most of the countries that collect data on both race and ethnicity use separate questions. In data sources on race, most countries collect information via both the census (except Chile and Mexico) and sample surveys (except Ireland). The situation is somewhat different for ethnicity, where 6 out of 17 countries only include a question on ethnicity in the census.

Most countries that only collect information on ethnicity but not on race are in Eastern Europe and the Baltics and share a concept of ethnicity as related to nationality and ancestry (but not current citizenship). Poland and Slovakia include religion-related groups such as Jewish/Yiddish as ethnic categories alongside nationalities, and the Israeli measure of ethnicity is a mix of information on country of birth from the population register and a self-reported question on religion. The Romani, as

a distinct ethnic group, is only listed among the response options in Hungary, Poland, Slovakia, and Romania. Ireland is currently testing the addition of a Romani category in its Pilot Census 2018.

International comparability of questions on both race and ethnicity is low and question wording differs across clusters of countries. Even when the concept (e.g. ethnicity as nationality) is similar, pre-coded response categories vary significantly, as they naturally reflect a country's racial and ethnic makeup.

Since most countries adjust over time the content and format of their ethnicity questions, comparability over time might be an issue (OECD, 2017b), while internal consistency of questions across current instruments is usually guaranteed. Some countries have developed specific statistical standards in this regard (e.g., the Ethnicity Standard in New Zealand or the harmonised groups in the United Kingdom). Only Hungary, Slovakia, and Costa Rica use slightly different diversity categories for some surveys.

All countries participating in the OECD survey use self-reporting as an identification method, except Israel, which draws partly on nationality data from its population registry, and Latvia, which has a register-based system but did not provide further information. Mexico, interestingly, is the only country with an additional item on reflected diversity for its race and ethnicity questions, asking the respondent for an evaluation of the personal characteristics that a third person would likely attribute to him/her.

Only 3 out of 8 countries that collect data on race (United Kingdom and the United States) or visible minority populations (Canada) enable individuals to declare more than one category. 12 out of 17 countries allow multiple responses for ethnicity, except Israel, Latvia, Slovakia, and Slovenia.

Statistics Canada does not define race or produce racial statistics. Relevant questions relate to the ethnic or cultural origin or origins of a person's ancestors. For persons of East Indian or South Asian origins, they would report a specific origin or origins. For example, they would not report "Indian"; instead they would indicate the specific group, such as "Punjabi" or "Tamil" (Statistics Canada, 2017b).

The vast majority of countries use a mix of pre-coded response options with an "other" category that usually allows respondents to specify the race/ethnic affiliation in a write-in box. Only Chile, Mexico, Colombia and Costa Rica, for their race data collection, do not make use of such an open-ended option. Czechia, Romania, and the Russian Federation are the only countries to not have pre-coded ethnicity categories at all.

Seven out of all countries surveyed (Australia, Canada, Chile, Mexico, New Zealand, the United States, Colombia, and Costa Rica) collect data on their indigenous populations. Markedly, almost as many OECD countries in which indigenous people live (Denmark, Finland, France-New Caledonia, Japan, Norway, Sweden) do not gather any such information, which significantly limits policy options to improve the well-being of these communities (OECD, 2017a; OECD, 2018). For those countries that engage in gathering information, data collection is relatively frequent as both the census and sample surveys include relevant questions.

Determination of indigenous status is frequently carried out in several steps, with an initial binary question asking about whether a person identifies as indigenous, and follow-up questions propping for the respondent's specific group/tribe and/or indigenous dialects spoken. These follow-up questions are often either completely open-ended or use a pre-coded higher-level category of indigenous communities with additional instruction to fill in the exact tribe.

Even more so than with race and ethnicity data collection, international comparability for indigenous identity is very limited given the strong local identity of communities. Countries perform well in current internal comparability of their items. However, comparability over time can be limited due to changes in response options between census waves.

2.2 Defining variables for poverty disaggregation

68. This part provides methodological and practical guidance on the implementation of core variables recommended for disaggregation of poverty measures. For each variable, background information, concepts and definitions of the variable and related categories, and implementation guidelines are provided. Reference questions are also proposed to illustrate a possible way of implementation. In practice, the mode of collection could also affect the formulation of questions and related interviewing guidelines and may determine the accessibility of certain population groups, such as persons with disabilities. Measures implemented to protect personal data, i.e. guaranteeing the confidentiality of data and their exclusive use for statistical purposes can have an important impact on the participation of some vulnerable groups.

69. Most of the proposed guidance was adapted from the CES Recommendations for the 2020 Censuses of Population and Housing (UNECE, 2015a) and the Eurostat project on standardisation of social variables (Eurostat, 2018c).¹⁴ The guidance follows the breakdowns as presented in the Guide on Poverty Measurement while at the same time it proposes some further adjustments in terminology or alternatives for their measurement. It is preferable to disaggregate poverty rates by individual characteristics, which are collected from questionnaires filled for each individual in a household (instead of only using characteristics of a reference person for the household). If no individual questionnaires are used, it is necessary to count the number of household members for each of the categories listed in the following sections in order to perform these disaggregations.

2.2.1 Sex¹⁵

2.2.1.1 Background

70. Sex is the most basic type of demographic information collected about individuals and a fundamental requirement for gender statistics. Data on two categories, male and female, are usually collected in most countries. Canada and the United Kingdom have recently conducted an in-depth review on gender identity (UNECE, 2019) in which they examine the emerging need for collection of two different variables, “sex” (at birth) and “gender”, (with possibilities for responses other than only male and female), in recognition of the fact that sex and gender¹⁶ are not necessarily the same in all cases - for example for those who self-identify with a gender that is not the same as the sex assigned to them at birth, or those who do not identify themselves within the binary categories of male and female. While this is clearly an

¹⁴ The project is ongoing, and the draft document used is subject to further changes to be considered for purposes of poverty disaggregation once available.

¹⁵ See more in UNECE, 2015b: 93-94 and Eurostat, 2017a: 6.

¹⁶ Gender is generally defined as a social construction relating to behaviours and attributes based on labels of masculinity and femininity; gender identity is a personal, internal perception of oneself and so the gender category someone identifies with may not match the sex they were assigned at birth. For a discussion of these issues in the context of the SDGs, see Talland and Evans, 2019.

emerging field, data quality and confidentiality/sensitivity issues may not be sufficiently resolved yet for recommending a common standard for disaggregation of poverty measures by non-binary categories of sex or gender. More information on the importance of gender statistics is available in the UNECE publications on “Indicators of Gender Equality” (UNECE, 2015b) and “Developing Gender Statistics: A Practical Tool” (UNECE, 2010).

2.2.1.2 Variable definition

71. The concept of the variable sex refers to the biological and physiological characteristics that define a person to be either male or female. It will typically refer to sex as registered at birth.

72. The reporting units are all individuals.

2.2.1.3 Categories definition

- Male
- Female

2.2.1.4 Implementation guidelines

73. Due to the importance of sex characteristics, it is fundamental that information is as complete and accurate as possible. In the absence of this information, or when respondents cannot identify with either of the two categories, for the purpose of disaggregation it may simply be imputed into the data record by attributing male and female interchangeably.

2.2.1.5 Reference question

74. Depending on the data collection mode or information being available from administrative sources, it might usually not be necessary to ask the respondents directly. In cases when this information needs to be asked directly to the respondents the recommended question is: *“What is your sex?”*

2.2.2 Age¹⁷

2.2.2.1 Background

75. Age is the most basic type of demographic information collected about individuals and essential to provide data on important target groups such as children, youth or elderly.

¹⁷ See more in UNECE, 2015: 94 and Eurostat, 2017a: 7.

2.2.2.2 Variable definition

76. The suggested concept is “age in completed years” which is the age at the last birthday before the reference date of the data collection, i.e. the interval of time between the date of birth and the reference date, expressed in completed years.

77. The reporting units are all individuals.

2.2.2.3 Categories definition

78. As a minimum and in line with the Guide on Poverty Measurement, the following categories used for poverty disaggregation are recommended:

- 0-17 (children)¹⁸
- 18-24¹⁹
- 25-49
- 50-64
- 65 and over (older persons).

79. This recommendation defines the minimum number of categories for which poverty rates should be calculated. If sample sizes do not allow for a separate disaggregation of youth in the age group 18-24, poverty rates should alternatively be presented for broader categories, (e.g. combining youth and children). For the purpose of calculating consumption units and equivalised incomes it is advisable for each household to record the number of children at different age groups (e.g. 0-13 and 14-17) during data collection. More detailed age disaggregation would also be of high policy relevance as shown for children and youth in Box 2.5.

Box 2.5

Detailed age disaggregation of poverty measures for children

The child population could further be disaggregated into smaller age groups as there are often significant differences in poverty rates between these age groups. Different rationale can be applied to this, for example, related to policy objectives (schooling status: pre-school, school age, secondary school); or age groups (0-4, 5-9, 10-14, and 15-17).

In terms of multidimensional poverty, using life-cycle related age categories makes it possible to distinguish between the needs of children at different stages, for example, early childhood, school age, and adolescence. Some indicators of multidimensional poverty (e.g., stunting) would usually only be defined for children under 5 years of age.

¹⁸ The current UN standard for children is 0-15 years. However, the UN Convention on the Rights of the Child defines a child as a person under the age of 18 years, unless the national age of majority is lower, which is the case in only a small number of countries.

¹⁹ The age group 18-24 can serve for purposes of reporting on youth, however the UN defines youth as those persons between the ages of 15 and 24 years. For statistical reporting on the EU's youth strategy, the definition of young people as those aged 15 to 29 years is used (European Commission, 2011).

2.2.2.4 Implementation guidelines

80. The variable age can be collected directly but due to its importance, it is generally recommended to collect information on date of birth. This yields more precise information and allows for use in two ways: by year of birth and by age. However, it brings legal issues to data confidentiality and is not allowed in all countries. Therefore, suggested categories for data collection are the following:

- Year of birth (4 digits)
- Passing of birthday (yes/no)
- Reference date (DD/MM/YYYY)

81. The integer number of the age in completed years is then deduced from this information.

82. Due to the importance of age characteristic, it is fundamental that information on it is as complete and accurate as possible. In the absence of this information, it should be imputed into the data record by attributing the most plausible value. It is also important to pay special attention to validation of data on age and eliminate possible errors by checking age together with family relationships (for example children older than their parents), possible misreporting of infant's age in months and quality assuring information for the elderly.

2.2.2.5 Reference question

83. When asking about the passing of birthday at the time of the interview, the following questions are recommended: *"What is your year of birth? Have you already had your birthday this year?"* When the reference date is not the time of the interview, the question has to be adapted, clearly explaining which reference date the passing of birthday should refer to (e.g. the end of an income reference period).

At the data collection level, the exact date of birth might be asked directly to the respondents. In this case, the recommended question is: *"What is your date of birth?"* (DD/MM/YYYY).

2.2.3 Disability status²⁰

2.2.3.1 Background

84. Disability status characterises the population as those with and without a disability. The UN Convention on the Rights of Persons with Disabilities defines: *"Persons with disabilities include those who have long-term physical, mental, intellectual or sensory impairments which in interaction with various barriers may hinder their full and effective participation in society on an equal basis with others"* (OHCHR, 2008). Such persons are considered to be at greater risk for limitations in activities and/or restrictions in participation than the general population even if the necessary adaptations have been made at the person or environmental levels to compensate for their impairments. This implies that a social model of disability is to be used.

²⁰ See more in UNECE, 2015a: 154-161.

2.2.3.2 Variable definition

85. A conceptual framework for and classification of disability is provided in the International Classification of Functioning, Disability and Health (ICF).²¹ According to ICF: *“Disability is an umbrella term for impairments, activity limitations, and participation restrictions. It denotes the negative aspects of the interaction between an individual (with a health condition) and that individual’s contextual factors (environmental and personal factors).”* ICF defines four major dimensions: body structure and function, activities, participation and environmental factors.

86. The ICF definition of disability may only partly derived from an individual’s condition, because the same condition may affect activities and participation in very different ways when the environment is adjusted to special needs. Also, the definition of disability will not necessarily coincide with government’s definitions which may be used for welfare entitlements.

87. The reporting units are all individuals.

2.2.3.3 Categories definition

- A person with disabilities (disabled)
- A person without disabilities (non-disabled)

2.2.3.4 Implementation guidelines

88. The Guide on Poverty Measurement (UNECE, 2017a) recommends using a short set of six questions proposed by the Washington Group for the disaggregation. These questions were originally developed to be used in a census context to identify persons at risk of disability. They have been thoroughly tested for their comparability and are included in the Conference of European Statisticians Recommendations for the 2020 Censuses of Population and Housing. The short list has been implemented in over 80 countries including comparative surveys such as the Demographic and Health Surveys (DHS), the Multiple Indicator Cluster Surveys (MICS) and the Living Standards Measurement Studies (LSMS). The Washington Group short set is also promoted by international aid programmes (DEFID, United Kingdom and the DFAT, Australia) as well as United Nations agencies and NGOs (UN 2020, p. 15).

89. The Washington Groups’ short list is certainly useful for the disaggregation of SDG indicators. It distinguishes between several specific functional limitations. A full representation of a social model of disability requires the collection of additional information. This may not always be feasible in practice in already lengthy survey questionnaires which need to cover a broad range of topics.

90. For the European Union, questions on functional limitations are not included in the regulations that determine the content of EU-SILC. Instead, the Global Activity Limitation Instrument (GALI) is used as a straightforward disaggregation variable among a set of questions that define the Minimum European Health Module (MEHM). This instrument requires careful implementation and translation. It has been shown to yield valid and comparable overall measures (Van Oyen et al. 2018). When aiming for

²¹ <http://www.who.int/classifications/icf/en/>

comparability within the UNECE region it should be acknowledged that for the EU Member States, the disaggregation of poverty measures will rather rely on the two GALI questions below than the short set of functional limitations proposed by the Washington Group.

2.2.3.5 Reference question

91. For the purpose of disaggregating poverty measurement, it is recommended to use a two-question instrument (GALI) as follows:

- QUESTION_1: "Are you limited because of a health problem in activities people usually do? Would you say you are... severely limited, limited but not severely, or not limited at all?"

92. If answer to QUESTION_1 is "severely limited" or "limited but not severely" ask QUESTION_2:

- QUESTION_2: "Have you been limited for at least the past 6 months? Yes, No"

93. The following text presents approaches of three international entities that could be used for purposes of more detailed disaggregation in the context of specific disability surveys:

- a) The comprehensive scale used by the World Health Organization (Model Disability Survey),
- b) The questions proposed by the Washington Group on Disability Statistics, and
- c) The ESS, which includes the most pragmatic approach reflected in the above reference question.

94. Additional country examples from Canada and Mexico further illustrate the broad variety of approaches, which are currently used in practice.

World Health Organization - Model Disability Survey

95. The Model Disability Survey²² was developed by the World Health Organization and the World Bank as a general population survey to provide detailed information on the life of persons (2 years and over) with and without disabilities and on the difficulties they encounter, regardless of any underlying health condition or impairment. Its brief version²³ consisting of 40 questions is suitable as a module to be integrated into existing household surveys and used for disaggregation. Disability is characterized by limitations in functioning in multiple life areas depending on how impairments and health conditions play out in a person's real-life environment. Disability is conceived as a continuum, but the methodology also establishes a cut-off point to determine the group of persons with disabilities. The model disability survey has been mainly implemented in dedicated disability surveys in Chile, Sri Lanka, the Philippines and Qatar as well as surveys conducted in Cameroon (Adamawa), Pakistan (Balochistan) and the United Arab Emirates (Dubai) (UN 2020).

The Washington Group on Disability Statistics

96. The Washington Group on Disability Statistics constituted as a city group under the UNSC developed several modules for measuring disability in censuses and surveys.²⁴ The Short Set of Disability

²² <http://www.who.int/disabilities/data/mds/en/>

²³ Idem.

²⁴ <http://www.washingtongroup-disability.com/washington-group-question-sets/>

Questions²⁵ is a tool that enables comprehensive measurement of disability in population 5 years and over. The concept of disability refers to the presence of difficulties in an unaccommodated environment in six core functional domains: walking, seeing, hearing, cognition (essential domains), and self-care and communication (additional domains).

Box 2.6

The Washington Group Short Set on Functioning (WG-SS)

Introduction: The next questions ask about difficulties you may have doing certain activities because of a HEALTH PROBLEM.

1. Do you have difficulty seeing, even if wearing glasses?
2. Do you have difficulty hearing, even if using a hearing aid?
3. Do you have difficulty walking or climbing steps?
4. Do you have difficulty remembering or concentrating?
5. Do you have difficulty (with self-care such as) washing all over or dressing?
- 6 Using your usual language, do you have difficulty communicating, (for example understanding or being understood by others)?

Each question has four response categories, which are read after each question: 1/ No, no difficulty, 2/ Yes, some difficulty, 3/ Yes, a lot of difficulty and 4/ Cannot do it at all.

Disability is determined, according to the WG-SS, as anyone having at least 'a lot of difficulty' on at least one of the six questions.

Source: <http://www.washingtongroup-disability.com/wp-content/uploads/2016/12/Disability-Measurement-and-Monitoring-Using-the-WG-Disability-Questions-July-2018.pdf>

97. It is important to understand that these questions capture only specific impairments. They should never be used in isolation to define disability. They were designed to be combined with additional questions to reflect the social model. The main strength of this approach is that it allows to clearly distinguish between prevalence of impairments and the social barriers to inclusion that define the social model of disability.

98. In order to address measuring children's disability in a more comprehensive manner, a module on Child Functioning was developed together with UNICEF (The Washington Group and UNICEF, 2018). Disability is measured for children between 2 and 17 years of age via difficulties in functional domains such as hearing, seeing, communication/comprehension, learning, mobility, and emotions. The questionnaire for children 2-4 years old consists of 16 questions whereas the questionnaire for children 5-17 consists of 24 questions. Apart from child functioning, there is also a special module on inclusive education among children.

²⁵ Idem.

99. Finally, the Washington Group has developed together with the ILO a special module for the LFS to capture functional limitations and barriers for disability and employment (UN 2020).²⁶

The European Statistical System and the European Health Interview Survey

100. The European Health Interview Survey enables measuring disability using an activity-oriented approach in four domains of functioning: seeing, hearing, mobility and cognition (Eurostat, 2018: 33-42). The instrument was proposed for population 15 years and over²⁷ using eight questions with response categories: “no difficulty”, “some difficulty”, “a lot of difficulty”, “cannot do at all” and “unable to do”. Additional questions in domains of self-care and domestic life are collected for persons 55 years and over.

101. The Minimum European Health Module (Eurostat, 2018: 17-22) is used as a standard tool for measuring health and disability in various surveys within the ESS. Three concepts are used: general health, chronic morbidity and, activity limitation. For measuring disability, a concept of activity limitation is used referring to the severity of participation restriction through long-standing limitation (6 months or more) in activities that people usually do because of health problems. It is measured in population 15 years and over using three response categories: “severely limited”, “limited but not severely” and “not limited at all”. A person with a disability is defined as a person who is severely limited or limited (but not severely) in usual activities.

102. Additional steps were undertaken to measure the health of children within EU Statistics on Income and Living conditions (Eurostat, 2019a). In 2017, data collection on children’s health took place for the first time in an ad hoc EU-SILC module, which included 6 variables: general health, limitation in activities because of health problems, unmet need for medical examination or treatment, main reason for unmet need for medical examination or treatment, unmet need for dental examination or treatment, and main reason for unmet need for dental examination or treatment. Two concepts are used: general health and activity limitation (for measuring disability the concept of activity limitation is used). It is a subjective assessment of whether the child (0-15) is limited by any ongoing physical, mental or emotional health problem, including disease or impairment compared with children of the same age. Consequences of injuries/accidents, congenital conditions, birth defects, etc. are all included while temporary or short-term limitations are excluded. The period of at least the past 6 months is strictly related to the duration of the activity limitation and not to the duration of the health problem. A child is considered to be disabled if it is severely limited or limited but not severely in activities children of its age group usually do. The module is planned to be collected regularly in the revised EU-SILC.

Canada’s Disability Screening Questions

103. Canada’s Disability Screening Questions (DSQ) are used in the Canadian Income Survey (the main source of poverty information in Canada) to identify persons with a disability. The DSQ are based on the social model of disability, which defines disability as the relationship between body function and structure, daily activities and social participation while recognizing the role of environmental factors.

²⁶ The disability module can be downloaded in CSPro format from <https://ilostat.ilo.org/resources/lfs-resources/>

²⁷ Some activity limitation questions were also adapted for children and were subject to pretesting at the time of drafting this paper.

104. The DSQ comprehensively identify ten distinct disability types. The questions were drafted following an extensive review of existing disability indicators used in Canada and internationally. Development involved several rounds of qualitative testing conducted to ensure the validity of each question. This was followed by two major quantitative tests to assess the reliability of the DSQ on surveys with different contexts.

105. Using the DSQ, disability is defined to include anyone who reported being "sometimes," "often" or "always" limited in their daily activities due to a long-term condition or health-problem lasting or expected to last six months or more, as well as anyone who reported being "rarely" limited if they were also unable to do certain tasks or could only do them with a lot of difficulty (Statistics Canada, 2016).

Mexico's National Household Income and Expenditures Survey

106. Within Mexico's National Household Income and Expenditure Survey, this theme of disability statistics is addressed by two questions asked to the head of the household or a household member who is informed about the household members. Responses to these questions facilitate the presentation of information about income and expenses items disaggregated by disability status.

107. The first question identifies up to seven different types of disability for each member of the household and literally says: *"In your daily life, (the member of the household in question) has difficulty to..."*

- 1) Walk, move, climb or descend;
- 2) See, even when wearing glasses;
- 3) Talk, communicate or converse;
- 4) Hear, even using a hearing aid;
- 5) Dressing, bathing or eating;
- 6) Pay attention or learn simple things;
- 7) Any mental limitations?"

108. When any difficulty is reported, a second question investigates the cause for each difficulty with the following options: birth, disease, accident, advanced age or another reason. For the 2020 edition, this Survey will have greater attachment to the set of the Washington Group.

2.2.4 Migratory status²⁸

2.2.4.1 Background

109. The proposal presented here is based on recommendations of 2017 UN Expert Group Meeting on Improving Migration Data in the Context of the 2030 Agenda to take a stepwise approach in defining a

²⁸ See more in UNECE, 2015: 134-147 and Eurostat, 2017a: 36, 40.

migratory status for disaggregation of SDG indicators (UN, 2017e). It suggests as a first step, that migratory status can be classified as native-born and foreign-born persons, or citizens and non-citizens (including stateless persons). As a second step, data can be further disaggregated by country of birth of the parents, duration of stay in the country, and reason for migration. Finally, internal migrants and internally displaced persons could also be considered if there is an interest in population mobility within the country. For purposes of disaggregation of poverty-related indicators by stock migrant population, the same approach is recommended and variables *country of birth* together with *country of main citizenship* are suggested as a minimum.²⁹

2.2.4.2 Variable definition

110. The **country of birth** of an individual is defined as the country of usual residence (in its current boundaries) of the individual's mother at the time of delivery.

111. The **country of main citizenship** reports on the country of the person's main citizenship. "Citizenship" is defined as the particular legal bond between an individual and his/her State, acquired by birth or naturalisation, whether by declaration, choice, marriage or other means according to national legislation. A person with two or more citizenships where neither of the countries is the reporting country shall be allocated to only one country of citizenship based on his/her choice; in situations when one of the citizens is of the reporting country the preference should be given to the reporting country.

112. The reporting units for both variables are all individuals.³⁰ Both variables should be complemented by information on the number of years of residence in or citizenship of the country.

2.2.4.3 Categories definitions

113. Country of birth:

- Foreign-born (born in a country other than the reporting country) population/residents;
- Native-born (born in the reporting country) population/residents.

114. Country of main citizenship:

- Foreigners/non-national citizens (individuals having the citizenship of a country other than the reporting country, including stateless persons);
- National citizens (individuals having the citizenship of the reporting country).

115. If data enable, joint use of both variables is recommended by defining the following migrant population groups: Native-born nationals, Native-born foreigners, Foreign-born nationals, Foreign-born foreigners. Even further disaggregation is possible, if information on the year of arrival and/or on parents' country of birth is available. 'First-generation immigrants' are people born in a country other than their

²⁹ These variables are usually complemented by variable Country of residence and Country of birth of father/mother.

³⁰ Specific filters (e.g., aged 15+) may apply depending on the survey target population concerning the variable.

country of residence and whose residence period in the host country is, or is expected to be, at least 12 months. 'Second-generation immigrants' are native-born persons with at least one foreign-born parent.

2.2.4.4 Implementation guidelines

116. The categories used for collection of data are the following:

- Country of birth: Country code;
- Country of main citizenship: Country code, Stateless.

117. Country codes are recorded according to ISO 3166³¹ three-letter code (alpha-3).

118. The category "stateless" (country of main citizenship) corresponds to a person without recognized citizenship of a state.

119. Information on the country of birth should be obtained according to the current national boundaries and not according to the boundaries in place at the time of birth. Further guidance concerns cases when boundaries of a country have changed, and in particular, cases when previously existing countries have split to form two or more new countries, as well as cases when the information on the place of usual residence of the mother at the time of the birth is not available.

120. Information on the country of main citizenship should be obtained in accordance with the current administrative status/legal situation of the individual. Further guidance concerns cases of "recognised non-citizens".

2.2.4.5 Reference questions

121. In case the information is available from administrative sources it might not be necessary to ask the respondents directly. When this information is required from the respondents, the recommended question is, *"In which country were you born?"*

122. The country of birth of a person, who was born during the mother's short-term visit to a country other than her country of usual residence, should be the country where the mother had her place of usual residence. In situations (e.g. specific countries or regions) where this reference question may not capture appropriately the information on the place of the usual residence of the individual's mother at the time of delivery, the following question should be asked: *"Which was the country of usual residence of your mother at the time when you were born?"*

123. Countries with boundaries that have changed, or which split to form two or more new countries, may need to ask additional questions e.g. city/town of birth, in order to determine the country of birth. Particular care is needed in cases where national boundaries have changed and/or where previously existing countries have split to form two or more new countries. As noted in the CES Recommendations for the 2020 Censuses of Population and Housing (UNECE, 2015a: par. 651), a person should not be

³¹ <https://www.iso.org/iso-3166-country-codes.html>

regarded as foreign-born (i.e. recorded as born in a country other than the reporting country) simply because the national boundaries of the country of birth have changed.³²

124. When the information on main citizenship is available from administrative sources it might not be necessary to ask the respondents directly. When this information needs to be supplied by the respondents, the recommended question is: "*What is your citizenship?*"

2.2.5 Ethnicity³³

2.2.5.1 Background

125. Statistical information on ethnicity (or ethno-cultural characteristics in general) is of increasing relevance in some countries, particularly in the context of developing migration, integration, equality and minority policies. Such information is needed to rectify any discrimination and unequal treatment and the need for data may even be driven by legal requirements associated with evaluating anti-discrimination policies.

126. However, the collection of such data may be prohibited by law or its narrow interpretation in some countries. In the European Union data on ethnicity is not commonly collected. It might also require the implementation of special monitoring mechanisms to guarantee the free declaration of the respondents and the protection of the information recorded as it may be viewed as sensitive and potentially misused. In addition, there are no international standards or classification relating to ethnicity.

127. Therefore, the following text presents only general guidance on defining and measuring ethnicity and refers to national examples. Countries are encouraged to develop their national definitions and measurement tools which comply with national legislation, ensure the protection of personal information and enhance the relevance and accuracy of the data.

128. More specific guidance might be available in the future as a result of ongoing activities relating to defining ethnicity in the context of SDGs or other initiatives such as the recent OECD work related to national practices of measuring selected aspects of diversity (see Box 2.4 above).

2.2.5.2 Variable definition

129. According to a broad definition, ethnicity is based on a shared understanding of the history and territorial origins (regional, national) of an ethnic group or community as well as on particular cultural characteristics, such as language and/or religion and/or specific customs and ways of life.

³² The following important exception to the general rule of considering the current borders might exist: a person whose mother's place of usual residence was, at the time of his/her birth, part of the person's actual country of origin (e.g., as indicated by his/her citizenship or current place of usual residence) but is not any more due to changed borders. In this case the country of birth can exceptionally be enumerated at the boundaries at the time of birth.

³³ See more in UNECE, 2015a: 148-153.

130. Ethnic identity can be measured using a variety of concepts, including ethnic ancestry or origin, ethnic group, cultural origins, nationality, race, colour, minority status, tribe, language, religion or various combinations of these concepts. Particular reference should be made, when relevant, to indigenous peoples. In line with CES Recommendations for the 2020 Censuses of Population and Housing (UNECE, 2015a), it is proposed that an affiliation with a certain ethnic group is distinct from an affiliation with a language and/or religious group.

131. National definition of ethnicity and concepts used will depend on national conditions and needs.

2.2.5.3 Implementation guidelines

132. Some general recommendations regarding measuring and dissemination of statistics based on ethnicity are provided here:

- **Ensuring free and open declaration of ethnicity** (self-identification): The subjective and multidimensional nature of the topic requires special attention when designing respective tools. The following guidance should facilitate the data collection:
 - Option for mixed categories or multiple choice of ethnicity;
 - Option allowing respondents to classify themselves as belonging to “other” unspecified ethnic groups and to write in their chosen description;
 - The ethnic group questions are asked in a way that allows the respondent to see all possible response options before making their decision;
 - To guarantee the free self-declaration of ethnicity, respondents should be allowed to indicate “None” or “Not declared”.
- **Involvement of representatives of ethnic communities** (participation): to undertake consultations with representatives of ethnic communities in defining classifications, developing the measurement tool, and designing statistical outputs; this would promote transparency, equality of treatment, better understanding of the reasons for collecting such information, and full participation of the population.
- **Protection of personal data during collection and dissemination**: to implement special monitoring mechanisms to guarantee the free declaration of the respondents and the protection of the information recorded and implementing appropriate disclosure control measures for dissemination purposes by finding a balance between disclosure risks and enough detail in reporting on minorities.
- Statistical categories should reflect demographic changes as well as evolutions in the understanding of ethnic identities, while remaining grounded in sufficient stability to allow comparability over time and across sources (Balestra and Fleischer, 2018).

133. Concrete examples of measuring ethno-cultural characteristics on a national level:

- The United Kingdom: *Measuring equality: a guide for the collection and classification of ethnic group, national identity and religion data in the United Kingdom*³⁴ and *Harmonised Concepts and Questions for Social Data Sources - Ethnic Group (UK GSS, 2017)*;
- The United States: *Questions Planned for the 2020 Census and American Community Survey* (U.S. Census Bureau, 2018). More information is provided in Box 2..

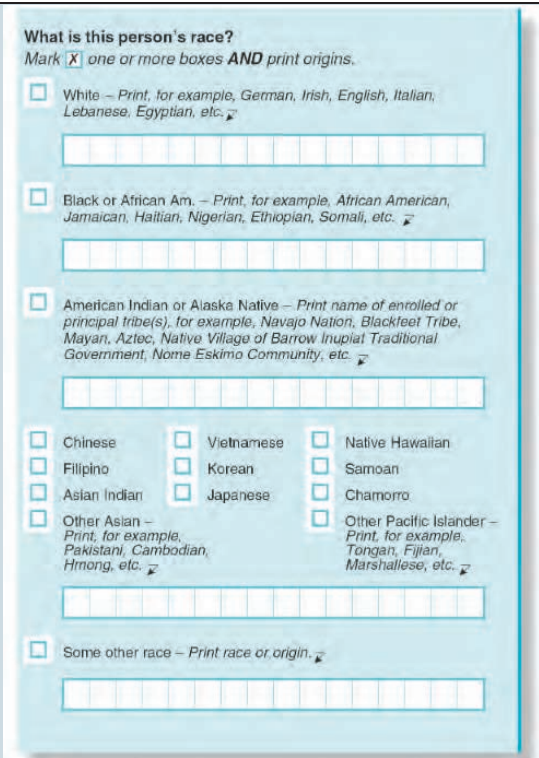
Box 2.7

United States’ experience with measuring race and ethnicity

The U.S. has a long tradition of collecting and using data on race and ethnicity dating back to the first census in 1790 (Pratt et al, 2015). Since the 1970s the U.S. Census Bureau has been developing and testing questions used to measure race and ethnicity.¹ The content and form of collected data have reflected changes in society influenced by social, political and economic factors including emancipation, immigration, and civil rights.

Figure 2.1

United States example of measuring race and ethnicity



The image shows a portion of the 2020 US Census form titled "What is this person's race?". It includes instructions to "Mark ☒ one or more boxes AND print origins." and several checkboxes for race categories: White, Black or African Am., American Indian or Alaska Native, Chinese, Filipino, Asian Indian, Other Asian, Vietnamese, Korean, Japanese, Native Hawaiian, Samoan, Chamorro, Other Pacific Islander, and Some other race. Each checkbox is followed by a list of examples and a write-in area.

A major issue in measuring ethnical background in the last two decades was the growing number of people who did not identify with any of the official race categories. This was primarily due to reporting by Hispanics, but also segments of other populations, such as Afro-Caribbean and Middle Eastern or North African populations. Additional research to tackle the issues on measuring race and ethnicity began in 2008 and consisted mainly of the 2010 Census Alternative Questionnaire Experiment (Stokes et al, 2012) and the 2015 National Content Test.² The goals were to evaluate and compare different versions of questions and to test different contact strategies for optimizing self-response (with focus on the use of the Internet).

The 2020 Census and the American Community Survey will include separate questions on race and Hispanic origin. For each race group, respondents will have the opportunity to indicate “origin” in either a write-in area or a checkbox.

Source: Questions planned for 2020 Census and American Community Survey (U.S. Census Bureau, 2018: 11).

¹ <https://www.census.gov/about/our-research/race-ethnicity.html>

² <https://www.census.gov/programs-surveys/decennial-census/2020-census/research-testing/testing-activities/2015-census-tests/national-content-test.html>

³⁴ <http://webarchive.nationalarchives.gov.uk/20160106185805/http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/index.html>

2.2.5.4 Reference question

134. As collecting information on ethnicity and race differs greatly and is also country specific, no harmonised question is proposed.

2.2.6 Household type³⁵

2.2.6.1 Background

135. From policy perspective, family composition and especially the number of (dependant) children is of main interest and is therefore embedded in the proposed disaggregation according to household type. The Guide on poverty measurement (UNECE 2017a, p. 27) made an appropriate reference to the housekeeping concept as inherent in the definition of a household in the Canberra Group Handbook (UNECE 2011, p. 64).

2.2.6.2 Variable definition

136. The variable describes the household composition of the private household according to the housekeeping concept (UNECE, 2015a: 162; Eurostat, 2017a: 15-20).

137. The reporting units are private households.

2.2.6.3 Categories definition

138. The following categories are recommended for disaggregation:³⁶

- One-person household;
- Couple without any children;
- Couple with one child;
- Couple with two or more children;
- Lone parent with children;
- Other type of household.

³⁵ See more in UNECE, 2015a: 176-177 and Eurostat, 2017a: 16-18

³⁶ These categories reflect a simplification of census recommendations (UNECE, 2015a: 176-177) in a few aspects: cohabiting and registered relationships are not distinguished. Couple households with other persons (multi-family households) are here considered as other type of household. The Guide of poverty measurement (UNECE 2017a) refer to the following categories: One-person households; Two adult household without children; Two adult household with one child under 18; Two adult household with two or more children under 18; One adult households with children under 18; Other. Some categories were slightly changed to stress the reference to the concept of family households by identifying couples, since the type of relationship between the adults is relevant for poverty measurement. However, when data on couples are not available, countries could use the number of adults as an alternative.

139. A “*couple*” is defined as a pair of individuals considered as partners in terms of their actual living arrangements within the household, regardless of whether the relationship with the partner is legally registered (e.g., marriage or civil union) or a de facto relationship.

140. The term “*child(ren)*” refers to the presence of son(s) or daughter(s) in the household where:

- The child is under 25.
- Natural/adopted son/daughter or stepson/stepdaughter refers to a natural (biological), adopted or step member of the family (regardless of age or partnership/relationship status) who has usual residence in the household of at least one of the parents;
- “Adoption” means taking and treating a biological child of other parents as one's own in so far as provided by the laws of the country (by means of a judicial process, whether related or not to the adopter, the adopted child acquires the rights and status of a biological child born to the adopting parents);
- The term “stepson/stepdaughter” refers to a situation in which a step-parent treats the child of his/her partner as one's own in so far as provided by the laws of the country, without adopting the child;
- Foster children and children-in-law are not covered by this category.

141. A “*lone parent*” is a parent not living with a (legal or de facto) partner in the same household who has most of the day-to-day responsibilities for raising a child or children.

2.2.6.4 Implementation guidelines

142. A person is considered to be a member of the household when having his or her usual residence in the household.

143. The categories used for the variable “household type” describing different types of household compositions refer only to one-generation (one-person household; couple without any children) or two-generation (lone parent with children; couple with children) households. Multigenerational households (those consisting of more than two generations) should be classified as “other type of household”.

144. One-person households and lone parent households can be further disaggregated by sex of the single person or lone parent to report estimates such as poverty rates of women or men living alone, single fathers, or single mothers.³⁷

145. The concept of “age (in completed years)” applies to all children in the household.

146. The requested categories would usually be derived from more detailed information on the household structure collected, for example via:

³⁷ The United States regularly reports separate poverty estimates for female householders and male householders (See Fontenot et al, 2018: Table 4).

- The composition of private households and the intra-household relationships between household members using a matrix containing the relationships of each member to each other member of the household (household grid) (Eurostat, 2017a: 9-12);
- Variables on the relationships to a “reference person” and the ones on spouse or cohabiting partner, father, and mother (Eurostat, 2018b: 9-13, 15-16).

2.2.7 Current employment status³⁸

2.2.7.1 Background

147. Classification of persons by their employment (or labour force) status provides important information about their relation to the labour market, in particular, to work performed for pay or profit and is, therefore, an important factor in analysing poverty. Main categories of labour force status divide population being “in labour force” (“employed” or “unemployed”) or “outside the labour force” (economically inactive). In general, two approaches can be envisaged:

- (Current) labour market status as derived by applying criteria according to the definition of the International Labour Organisation (ILO). This approach is usually applied in labour force surveys (LFS).
- Self-defined main activity status. This approach is usually implemented in other than labour force surveys and is therefore further referred to in these guidelines.

2.2.7.2 Variable definition

148. A concept of self-defined main activity status refers to a person’s own perception of his/her main activity status (i.e. the most important one for the respondent) in the respective reference period.

149. In order to keep coherence with poverty measurement, it may be preferable that the employment status refers to the most frequent status during the income/consumption reference period (usually 12 months preceding the data collection). The most frequent activity status is the status that a person declares him/herself to have occupied for more than half of the period for which information on any status is available. In practice, however, it mostly refers to a single question on the current situation (that is the situation at the moment of the data collection).

150. The reporting units are individuals of working age (usually population 15 years and over).

2.2.7.3 Categories definitions

151. According to the Guide on Poverty Measurement, the following categories used for disaggregation are recommended:

³⁸ See more in UNECE, 2015: 103-111.

- Employed;
- Unemployed;
- Retired;
- Other outside the labour force.

152. The main activity status is self-assessed by the respondent and the chosen category should appropriately describe how a person mainly perceives him/herself. The categories are not mutually exclusive and do not have a hierarchical order.

2.2.7.4 Implementation guidelines

153. Variables based on the concept of the most frequent activity status (in the income reference period) can be collected using a set of questions, each referring to the number of months spent in specific activity status³⁹. For implementation of this variable and where relevant, guidelines for using the concept of self-defined (current) main activity status can be used (Eurostat, 2017a: 21-23). The results for both variables are very sensitive to their implementation, especially with regard to the wording and order of the questions and categories, therefore the guidelines should be followed strictly. For purposes of data collection, the following categories of main activity status are proposed: Employed / Unemployed / Retired / Unable to work due to long-standing health problems / Student, pupil / Fulfilling domestic tasks / Compulsory military or civilian service / Other⁴⁰.

2.2.7.5 Reference question

154. The information should be collected, preferably with a single question from the respondent, rather than inferred from other responses.

155. The following question could be used:

156. What is your main current labour status (based on the main job)?

2.2.8 Tenure status of the household⁴¹

2.2.8.1 Background

157. Housing tenure and relating housing costs can be modified by social policies and can also to some extent determine poverty status of households. It is therefore valuable to include tenure status of households in the analysis of poverty.

³⁹ See the definition of derived variable Activity Status (ACTSTA) (Eurostat, 2019a)

⁴⁰ The last five categories compose the disaggregation category 'Other outside the labour force'.

⁴¹ See more in UNECE, 2015: 179-180 and Eurostat, 2017a: 19-20.

2.2.8.2 Variable definition

158. The variable refers to the arrangements under which a private household occupies all or part of a housing unit.

159. The reporting units are private households living in conventional dwellings (UNECE, 2015a: 187-188).

2.2.8.3 Categories definitions

160. The categories recommended for purposes of poverty disaggregation are the following:

- Owner without an outstanding mortgage;
- Owner with an outstanding mortgage;
- Tenant, rent at market price;
- Tenant, rent at a reduced price or rent free.

Owner with an outstanding mortgage/owner without an outstanding mortgage

161. A household is an owner household when at least one of its members is the owner of the housing unit in which the household lives. A person is an *owner* if he/she possesses a title deed, independently of whether the house is fully paid or not. A reversionary owner should be considered as the owner. If the housing unit is owned by a relative (not living in the household) of one household member (e.g., by a parent to a child living in the household) the household member (e.g., child) does not qualify as an owner.

162. The categories “owner with (without) outstanding mortgage” apply to the situations where the owner has to pay off at least one outstanding mortgage (“with outstanding mortgage”), respectively no mortgage (“without outstanding mortgage”), taken to buy this housing unit. Payoffs for mortgages and/or housing loans for any other housing unit (e.g., for a second dwelling) and/or for repairs, renovation, maintenance, or any non-housing purposes do not qualify. If the owner has already fully paid the principal of the mortgage and only the interest remains outstanding, the owner shall be considered as an owner with an outstanding mortgage.

Tenant, rent at market price/tenant, rent at reduced price

163. Households of which at least one member is the tenant (housing unit directly rented from the owner (i.e. “landlady” or “landlord”) or subtenant (housing unit rented from someone who himself is a tenant) of the housing unit in which the household lives, and where no household member is the owner of that housing unit. Households where at least one member pays partly rent and partly mortgage for the housing unit in which the household lives shall be classified under “owner with an outstanding mortgage” if the household member possesses a title deed.

164. *Tenant, rent at market price* applies to households of which at least one member is the tenant or subtenant who pays rent at prevailing or market prices. The category also applies where rent at market price is paid but is partly or wholly recovered from housing benefits or other sources, including public, charitable, or private sources.

165. *Tenant, rent at reduced price* covers households living in housing units at a reduced price, i.e. a price is below the market price (but not rent-free), and includes cases where the price reduction is granted (a) by law, (b) as a result of a social housing scheme, (c) for private reasons or (d) by an employer.

166. *Tenant, rent free* covers households living in housing units rent-free, i.e. where no rent is paid, and includes cases where the free rent is granted (a) by law, (b) as a result of a social housing scheme, (c) for private reasons or (d) by an employer.

2.2.8.4 Implementation guidelines

167. Categories Tenant, rent at reduced price and Tenant, rent free should be collected as separate categories. It is also suggested to collect data for categories covering all private households that do not live in conventional dwellings.

168. Countries may face practical difficulties in collecting data on subsidized housing due to misreporting as demonstrated by experience in the United States (Renwick and Mitchell, 2015). For the Current Population Survey Annual Social and Economic Supplement, one of the U.S.'s major household surveys, respondents are asked two separate questions about their receipt of housing assistance: (1) "Is this public housing, that is, owned by a local housing authority or other public agency?", (2) "Are you paying lower rent because the Federal, State or local government is paying part of the cost?". Administrative data from the United States Department of Housing and Urban Development were matched to survey responses. Of those matched, more than half had misclassified their type of assistance in the survey. It is therefore recommended to use administrative sources when there are difficulties in identifying housing as subsidized.

2.2.8.5 Reference question

169. The recommended question is:

170. *"Does your household own this housing unit or does it rent it?"*, followed by: *"Do you still have to repay money from an outstanding loan or mortgage for this accommodation?"* when the household is the owner; or by: *"Does your household rent this housing unit at market price, below market price, or use it free of charge?"* when the household is renting the housing unit.

2.2.9 Receipt of cash (or near cash) social transfers

2.2.9.1 Background

171. Social transfers/benefits serve to compensate for lack of income and prevent persons and their households from falling into poverty. The information on whether a person/household receives social transfers and its amount or share in disposable income can provide more insight into effectiveness of social policies.

172. There is no international recommendation on how to construct a respective variable. Several methodological elements need to be considered:

1. **Reporting unit.** The reporting unit could be the individual as suggested for example for “Main source of livelihood” (UNECE, 2015a: 120-121). However, certain social transfers are often provided on household level (e.g., child or housing allowances) and thus the household seems more appropriate in the context of poverty statistics.
2. **Scope of social transfers.** Social transfers in kind (STiK) are proposed to be excluded in line with mainstream international practices in poverty measurement. Cash or near cash social transfers (whether contributory or means tested) are defined as old-age (retirement) pensions and survivors’ (widows’ and widowers’, children, grandchildren, parents and other relatives of the deceased person) benefits; unemployment benefits; family/children-related allowances; sickness and disability benefits; education-related benefits; housing allowances; social transfers not elsewhere classified (Eurostat, 2020).
3. **Significance/Importance of social transfers in total income.** It can be based on various approaches:
 - Households classified as those *receiving any social transfers and households not receiving any social transfers*. However, this approach is weaker in classifying households with a very small share of social transfers in total disposable income (e.g. 1%). This could be an issue for example in countries using universal social benefits for children (i.e. all children of certain age are entitled to certain benefits).
 - Households classified as those *“with significant” or “without significant” share of social transfers in total income*. A minimum threshold to define significant share of social transfers would need to be determined (for example 10% of total disposable income). The challenge is the selection of the minimum threshold;
 - Households classified according to categories of *range of share of social benefits* in total disposable income (Shelton Alison, 2015). Categories of the share of social benefits could, for example, be: <10%, 10-50%, 50-90%, 90%<. It provides more detailed view on the importance of social transfers in household income than the previous approach but the question of defining cut-off points remains.
 - Households classified as those *with social benefits as the main source of income* and those *with the main source of income other than social benefits*. This approach is weaker in classifying households with a similar share of income from different sources (e.g. 54 % from social transfers and 46 % from work).

173. Various approaches could have different relevance in different countries and due to lack of any analysis on a global scale, a simple disaggregation is proposed to show poverty rates for persons/households who receive and those who don’t receive social transfers. As a possible extension, countries are encouraged to further disaggregate recipients of social transfers by the share of social transfers in total disposable income.

174. If social transfers are measured separately in an income survey, this can also be used for social policy impact assessment. For example, Eurostat regularly calculates the number of persons who would fall below the poverty line without the social transfers they receive.

2.2.9.2 Variable definition

175. The variable refers to the receipt of any cash social transfers/benefits.

176. Social transfers refer to all cash social transfers, i.e. not including STiK.

177. The variable is proposed to be constructed on an aggregated level, i.e. receiving/not receiving any social transfers but can also be constructed separately for each type of social transfers.

178. The reporting units are private households.

2.2.9.3 Categories definitions

179. The following classification of receipt of social transfers is proposed:

- Household receiving cash social transfers;
- Household not receiving any cash social transfers.

2.2.9.4 Implementation guidelines

180. The variable is derived from variables on receipt of specific individual-level and household-level social benefits.

2.2.10 Degree of urbanisation⁴²

2.2.10.1 Background

181. Spatial patterns of poverty can be analysed using classifications based on urban and rural concepts. In line with the recommendation of the Guide on Poverty Measurement, the variable degree of urbanisation is suggested for poverty disaggregation. The most recent classification was developed by the OECD and the European Commission and more background information and guidance are provided in a working paper “A harmonised definition of cities and rural areas: the new degree of urbanisation” (Dijkstra and Poelman, 2014) and on Eurostat dedicated section on methodology for degree of urbanisation (Eurostat, 2018d).

⁴² For more details, see UNECE, 2015: 88-89 and Eurostat, 2017a: 51-52.

2.2.10.2 Variable definition

182. The variable reports on the degree of urbanisation of the area where the person/household has his/her usual residence. The degree of urbanisation classifies local administrative units/municipalities⁴³ based on the share of local population living in urban clusters and in urban centres identified by population size and population density.

183. The reporting units are all individuals or households.

2.2.10.3 Categories definition

184. The following three types of area are classified and recommended for poverty disaggregation:

- Densely populated areas (cities);
- Intermediate density areas (towns and suburbs);
- Thinly populated areas (rural areas).

185. *Densely populated area (city)* is a territorial unit where at least 50% of the population lives in an urban centre (high-density clusters).

186. *Intermediate density area (town or suburb)* is a territorial unit where at least 50% of the population live in urban clusters but is not “city”.

187. *Thinly populated area (rural area)* is a territorial unit where more than 50% of the population lives in rural grid cells.

188. This classification is based on a combination of criteria of geographical contiguity and minimum population threshold applied to 1 km² population grid cells.

2.2.10.4 Implementation guidelines

189. The information on degree of urbanisation should be constructed according to the suggested methodology and therefore should not be collected from respondents.

190. For EU and some other countries, the information on the degree of urbanisation is to be extracted from the LAU2 list including the degree of urbanisation, which is available on the RAMON server.⁴⁴

⁴³ For EU and some other countries, it corresponds to LAU2 or communes, see <http://ec.europa.eu/eurostat/web/nuts/local-administrative-units>

⁴⁴ RAMON Eurostat's metadata server: http://ec.europa.eu/eurostat/ramon/miscellaneous/index.cfm?TargetUrl=DSP_DEGURBA

2.2.11 Educational attainment level

2.2.11.1 Background

191. The educational attainment level of an individual is the highest International Standard Classification of Education (ISCED) level successfully completed, the successful completion of an education programme being validated by a recognised qualification, i.e. a qualification officially recognised by the relevant national education authorities or recognised as equivalent to another qualification of formal education. In countries where education programmes, in particular those belonging to ISCED levels 1 and 2, do not lead to qualification the criterion of full attendance of the programme and normally gaining access to a higher level of education may have to be used instead. When determining the highest level, both general and vocational education should be taken into consideration.

2.2.11.2 Variable definition

192. The educational attainment level is defined according to the ISCED 2011.⁴⁵

2.2.11.3 Categories definitions

193. As a minimum, the following categories used for poverty disaggregation are recommended:

- **Low:** Less than primary, primary and lower secondary education (levels 0-2);
- **Medium:** Upper secondary and post-secondary non-tertiary education (levels 3- 4);
- **High:** Tertiary education (levels 5-8).

194. More detailed education attainment disaggregation following ISCED could be used depending on the purpose of the survey.

195. For example, Table 2.1 shows the detailed educational attainment levels used in Mexico.

Table 2.1

Educational attainment levels used in Mexico's National Household Income and Expenditure Survey

Level of attendance/Level approved	Instructions
0 None	Persons who do not have approved any degree of studies in the school. It includes those who only took literacy classes.
1 Preschool	Persons who only attended some degree in preschool or kindergarten. (Between 3 and 6 years old)
2 Primary school	Persons who have approved six years in this level (1 st to 6 th grade). (Between 6 to 12 years old)
3 Secondary education	Persons who have approved three years in this level. (Between 12 to 15 years old)
4 High school	Persons who have approved three years in high school. (Between 15 to 18 years old)
5 Normal	Persons who passed between one and six degrees that covers this level. Teachers are trained for preschool, primary or secondary education. (18 years and up)

⁴⁵ <http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx> – ISCED-A for levels of educational attainment.

Level of attendance/Level approved	Instructions
6 Technical or commercial career	<ul style="list-style-type: none"> Persons who passed between one and four degrees in technical or commercial careers. The answer of the respondent can be: secretary, computer, accounting assistant, electrician technician, dental technician, senior nutrition technician, senior tourism technician, etc.; Technical or commercial careers can be considered after having finished: <ul style="list-style-type: none"> Primary; Secondary; or High school or baccalaureate; Higher technicians are considered at this level. Although their requirement or antecedent is high school or high school, they do not grant a bachelor's degree.
7 Bachelor	Persons with undergraduate studies carried out in universities, technological universities, polytechnics and other institutions of higher education, public or private, whose requirement to enter is high school or baccalaureate. (1 to 5 years). (18 years and up)
8 Master	<ul style="list-style-type: none"> Persons who passed one or two years of this level, after completing the degree. The specialty of medic and nurses, whose duration is at least two years, is equivalent to a master's degree (after bachelor's degree).
9 PhD	<ul style="list-style-type: none"> Persons who passed from one to four years of this level, which are done after the completion of the master's degree. The subspecialty of medic is equivalent to doctorate (after master's degree).

2.2.11.4 Implementation guidelines

196. Joint Eurostat-OECD guidelines on the measurement of educational attainment in household surveys (Eurostat/OECD, 2014a; Eurostat/OECD 2014b) are available.

197. Coding of the variable should be based on the ISCED integrated mapping which is elaborated in each country. An ISCED integrated mapping is a table including information on national educational programmes and qualifications, their main characteristics and coding in ISCED.

198. When determining the highest educational level, both general and vocational education should be taken into consideration. If a person has successfully completed more than one programme at the same ISCED level, the most recent qualification should be reported (ISCED, 2011: 87).

199. Persons who have attended but not successfully completed a formal education programme should be coded according to the highest level of formal programme that they have (previously) successfully completed.

2.2.11.5 Reference question

200. The “diploma approach”, where respondents are asked about their diplomas instead of the level of education completed, is highly recommended. For details on the “diploma approach”, see the joint Eurostat-OECD guidelines (Eurostat/OECD, 2014).

201. Alternatively, the question(s) for this variable should be phrased by countries in a way that the concept of educational attainment level (qualification officially recognised by the relevant national education authorities or recognised as equivalent to another qualification of formal education) is

described as fully as possible. This can be achieved by listing the formal education programmes/qualifications (or categories thereof).

2.3 Examples of applications to the poverty disaggregation

202. This part provides examples of methodological approaches and challenges, and graphical presentations related to disaggregation of poverty measures. It is based on a selection of good practices and recommendations and of small-format electronic products for presenting poverty measures such as press releases, infographics and factsheets.

203. The review of just a few of these products revealed that all variables defined in this section are commonly used for analytical and dissemination purposes; age, sex, household/family type, employment status and migratory status (especially nativity) being the most used. Other commonly used disaggregations are for example region/residence, characteristics of job (full- or part-time) and status in employment.⁴⁶

2.3.1 Example 1. Gender and poverty measurement: approaches and challenges⁴⁷

204. The Beijing Platform for Action (BPA) (UN, 1995) and, more recently, the UNECE publication on indicators of gender equality (UNECE, 2015b) consider poverty among the most relevant indicators to assess inequality between men and women. Nonetheless, poverty remains one of the most problematic areas of gender statistics.

205. The difficulty of estimating monetary poverty by sex stems from the use of household level instruments to collect poverty data. Poverty has been conventionally measured using a monetary measure based on income or consumption and collected at the household level. The household-level estimate of poverty is attributed to all individuals living there, which means all household members are classified the same way—as either in or out of poverty. Intra-household disparities in access to income, consumption and other entitlements are not captured. As a result, the extent to which women and men experience poverty differently on an individual level remains largely opaque and does not feature prominently in poverty statistics (UN Women and World Bank 2018).

206. Some researchers have attempted to side-step this issue by looking at headship and comparing male versus female headed households in poverty. There are many questions however about the reliability and significance of estimates differentiated by ‘headship’, given the ambiguities in how it is defined and understood by both enumerators and respondents. For example, social norms and views

⁴⁶ For additional examples on poverty risk groups see https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_poverty_and_social_exclusion#Which_groups_are_at_greater_risk_of_poverty_or_social_exclusion.3F

⁴⁷ Adapted from publications: UN 2015 and UN 2016a.

about who is understood as the head of the household may vary across countries or across income groups and may privilege one sex over the other. Moreover, important questions also exist regarding the usefulness of a category such as ‘female-headed household’ that covers many different sub-groups of households that may not be comparable (e.g., female headed household where the husband is absent due to migration versus female headed households with no access to outside income) (UN Women 2019).

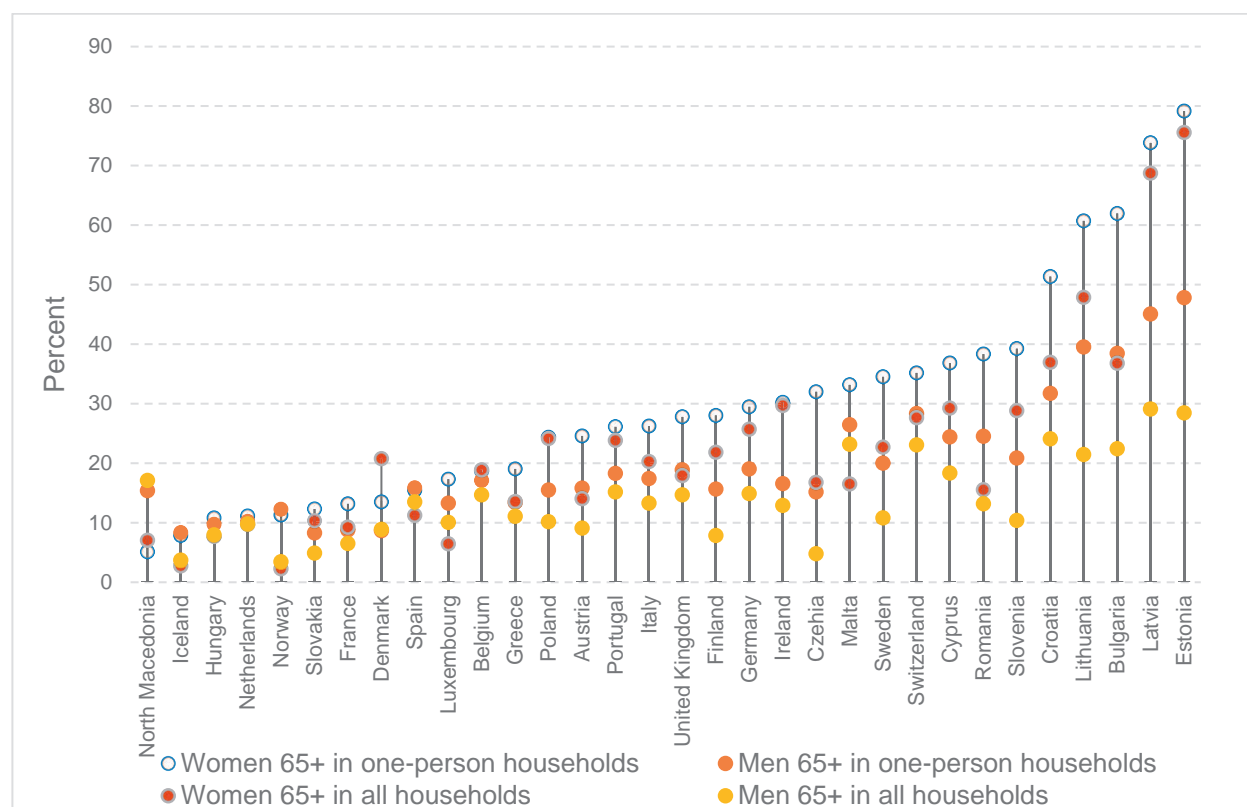
207. Even though it is difficult to capture the distribution of resources to household members through nationally representative surveys, it is possible to use existing household-level data to get a better—though imperfect—understanding of the gender dimensions of poverty. The findings of a joint UN Women and World Bank study show that a life-cycle approach can help to reveal meaningful differences in the way women, men, girls and boys experience poverty. Based on an analysis of data from 91 developing countries, the study shows that there are specific pressure points in women’s life course that place them at greater risk of poverty compared with men (UN Women and World Bank 2018). Focused on the percentage of women and men who live in poor households, the study finds poverty rates for women between 25 and 34 years of age are two percentage points higher than for men in the same age cohort. The difference in extreme poverty rates among women and men between 25 and 34 years of age coincides with differences in men’s and women’s life courses. Women marry and have children earlier in life than men. The responsibility for caring for young dependents earlier in life deprives women of opportunities to develop their capabilities and diminishes their chances of gaining access to decent work (UN 2019b).

208. To have a better understanding of the poverty profile across countries, more detailed disaggregations are needed, including by sex, age, and importantly by sex and age simultaneously, as well as other demographic and social characteristics of all household members, and by types of households (or living arrangements), taking into account the composition of those households.

209. In some countries poverty rates may vary among the elderly, for example indicating highest poverty rates among very old persons. The degree to which such differences are observed will also depend on the extent to which this population can be surveyed within private households (see section Surveying Institutional Households 3.7). But even within private households there are often considerable differences in poverty rates within the elderly population. The following example shows how at-risk-of poverty rates of older women and men differ when type of household is also considered. As shown in Figure 2.2, poverty rates among older persons (age 65+) are higher for women than for men in most European countries. On average, 23% of older women and 14% of older men were poor in that region in 2017. Living in one-person households, in particular, increases the risk of poverty for both older women and men, and in three quarters of the countries, it is more so for women than for men. The average poverty rate for older persons living in one-person households in European countries in 2017 was 30% for women and 20% for men. Furthermore, women living in one-person households are in most of the countries at higher vulnerability to poverty relative to when living with another person.

Figure 2.2

Poverty rate for older persons (age 65 and over) by sex, in all households and in one-person households, European countries, 2017



Source: Eurostat, 2019a.

Note: The data for Iceland is for 2016.

210. As described above, one major limitation of using household-level poverty data from a gender perspective is the lack of information on inequality in the consumption of goods and services among various household members. The household-level approach assumes that all individual incomes are pooled together, the resources are shared equitably, and all household members enjoy the same level of well-being. Existing data on intrahousehold sharing of resources suggest that income is most often pooled together within the household, but not always (Eurostat, 2013a), and the allocation of expenditures may reflect a gender dimension. Systematic national statistics are missing on this topic, and some of the statistical evidence on discrimination against women and girls is inconclusive (Duflo, 2012) and dependent on the statistical methods used (Zimmermann, 2012). Research has shown, however, that in some specific settings, particularly in the context of limited economic resources, inequality in the distribution of resources among girls and boys is evident, especially when it comes to private education, time devoted to child care, and access to health services.

211. Standard household-level measures of poverty do not take into account inequality within the household because it is difficult to know how household income/expenditure is distributed to each household member, particularly when it comes to common goods such as food, housing, water supply or sanitation. Besides, when different patterns of consumption are observed, it is not always clear if they are

related to different levels of individual biological need, to different preferences or the unequal distribution of resources.

212. Currently, there is no single straightforward measure of individual poverty from a gender perspective, and no single internationally agreed-upon indicator that can give more meaningful poverty counts for women and men (aside from total number living in poor households). That would require taking into account intra-household inequality, including through the use of some individual level indicators on selected dimensions of poverty. Nevertheless, recent methodological developments suggest a shift in thinking on poverty and gender from a perspective focused on the household as an economic unit, to women and men with an individual agency (capacity for individualized choice or action) and specific constraints, needs, and preferences. This would include the measurement, at the individual level, of asset ownership (UNSD, 2017); individual experience of food insecurity; and individual access to formal financial services.

213. The need to measure poverty and deprivation from a gender perspective has been recognised by the ESS. Consequently, Eurostat introduced a revised measure of deprivation which has the weighted proportion of people lacking/not being able to do at least 5 of 13 items.⁴⁸ The majority of the items used in the measure of new deprivation is collected at the personal level and, moreover, puts the focus on both material deprivation and social deprivation. Therefore, the new data collection opens many analysis possibilities e.g. cluster analysis with two categories of deprivation items “social inclusion items” and “financial items” as well as analysis of deprivation rate by sex (see also 3).

2.3.2 Example 2: Disability and more appropriate poverty measures (United Kingdom)

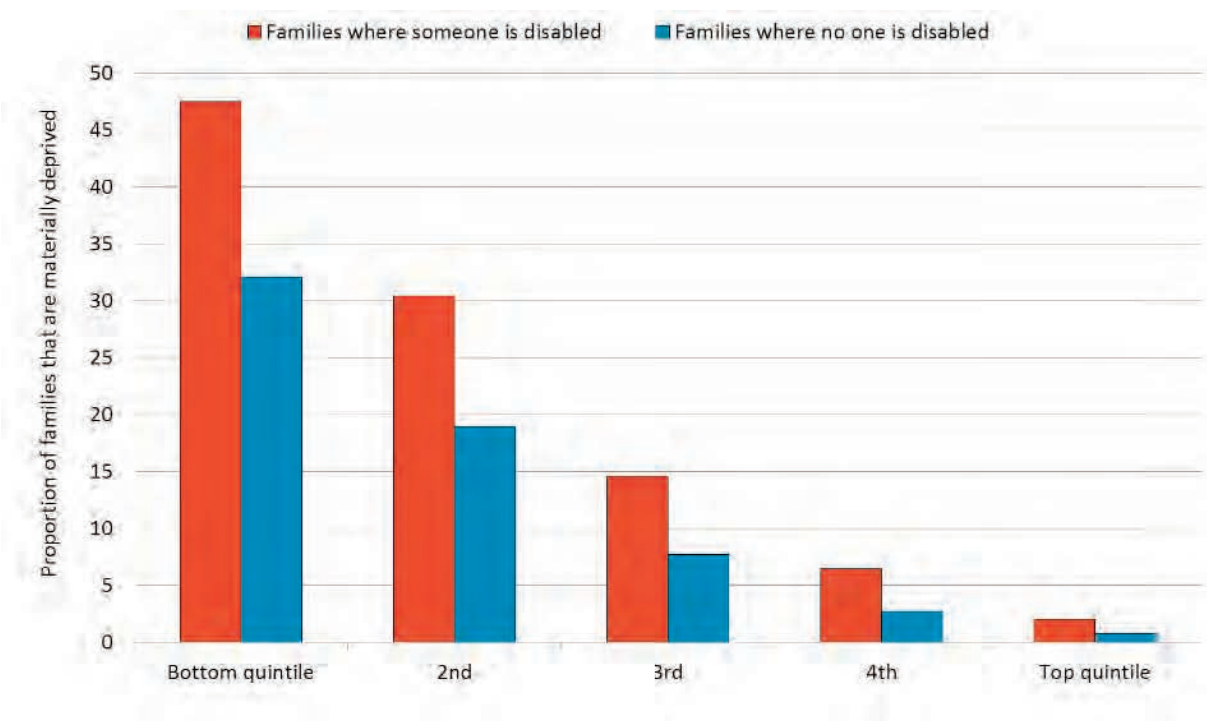
214. Some disaggregations need to be interpreted with caution as the needs of different subgroups may not be fully reflected in standard poverty lines. For example, the material needs of people with disabilities are often greater, due to both additional costs as a result of goods and services needed due to their disability as well as higher costs for some other items compared with individuals without disabilities (UNECE, 2017a: 24).

215. In 2014, a study on disability, long term conditions, and poverty (MacInnes et al, 2014) examined the degree to which poverty estimates are underestimated among disabled people. It finds that households with a disabled family member are more likely to be impoverished. The authors also note the inadequacy of headline poverty measures, recognising that disability and poor health are typically associated with higher costs to meet basic needs. This was exemplified by the fact that 47% of households at the bottom income quintile, in which someone is disabled, are materially deprived compared with 32% for households where no one is disabled (see Figure 2.3).

⁴⁸ Face unexpected expenses; one-week annual holiday; avoid arrears; meal with meat; chicken or fish every second day; home adequately warm; a car; replace worn-out furniture; replace worn-out clothes; two pairs of shoes; 'pocket money'; regular leisure activities; get together with friends/family for a drink/meal; have an internet connection.

216. The authors discuss the extra costs associated with disability but stopped short of adjusting household income due to the wide range and drivers of these costs. However, they do adjust household income to account for disability-related benefits arguing that, for instance, £100 of disability benefit merely makes up for the extra £100 of needs required and should not be considered as extra income. Adjusting income in this manner increases the proportion of people with a disability benefit who are in poverty from 17% to 33%.

Figure 2.3
Proportion of families by material deprivation and disability status



Source: Family Resources Survey, UK Department for Work and Pensions.
Note: The data is an average for 2009-10 and 2010-11.

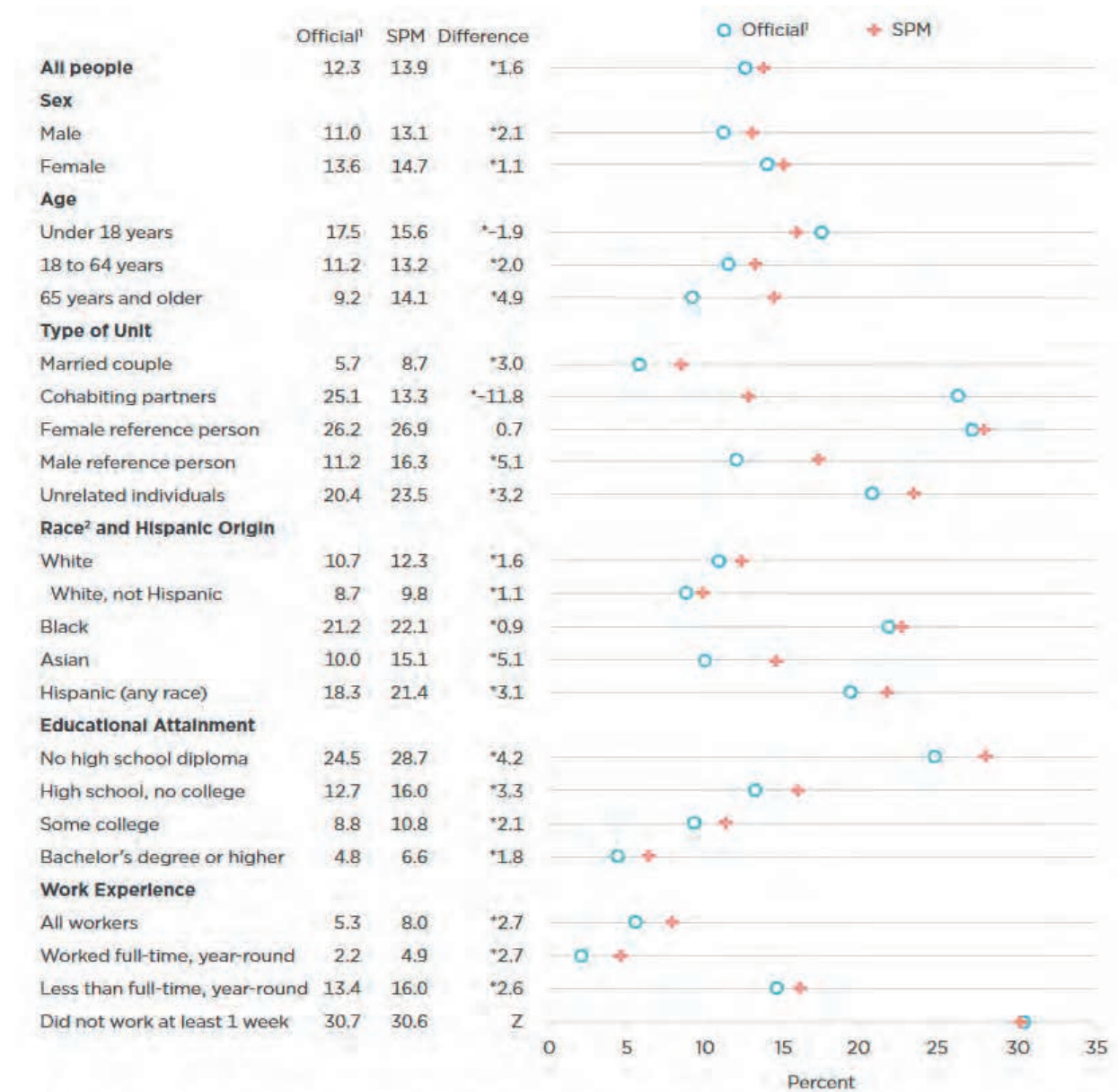
2.3.3 Example 3: Disaggregation and the use of alternative poverty measures (United States)

217. The impact of the use of an alternative poverty measure on different population groups is shown in an example from the United States (Fox, 2018). The U.S. Census Bureau publishes the so-called Supplemental Poverty Measure (SPM), which extends the official poverty measure by taking account of many of the government programs designed to assist low-income families and individuals that are not included in the official poverty measure. Figure 2.4 shows the different impact of using SPM on various population groups. While for most groups, SPM rates were higher than official poverty rates, the SPM shows lower poverty rates for children and individuals living in cohabiting partner units. The higher

poverty rates under the SPM for those aged 65 and over partially reflects that the official thresholds are set lower for units with householders in this age group, while the SPM thresholds do not vary by age.

Figure 2.4

Percentage of people in poverty by different poverty measures, 2017



* An asterisk preceding an estimate indicates change is statistically different from zero at the 90 percent confidence level.

Z Represents or rounds to zero.

¹ Includes unrelated individuals under the age of 15.

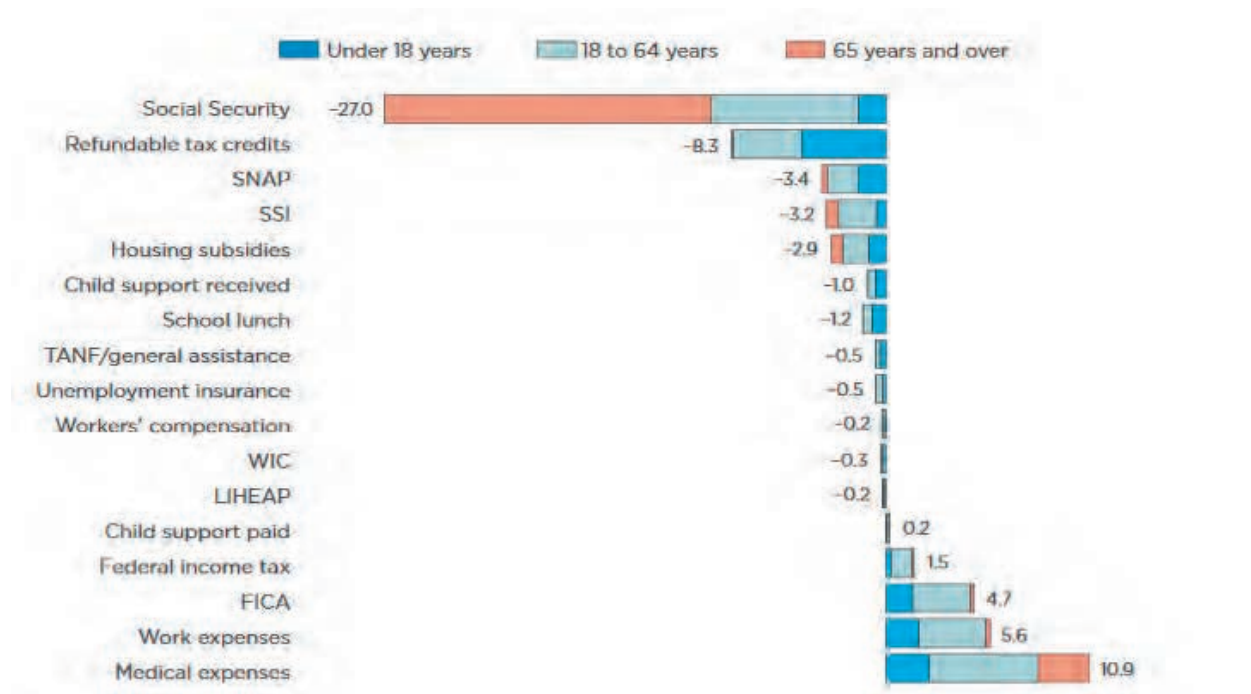
² Federal surveys give respondents the option of reporting more than one race. Therefore, two basic ways of defining a race group are possible. A group such as Asian may be defined as those who reported Asian and no other race (the race-alone or single-race concept) or as those who reported Asian regardless of whether they also reported another race (the race-alone-or-in-combination concept). This table shows data using the first approach (race alone). The use of the single-race population does not imply that it is the preferred method of presenting or analyzing data. The Census Bureau uses a variety of approaches. Information on people who reported more than one race, such as White and American Indian and Alaska Native or Asian and Black or African American, is available from the 2010 Census through American FactFinder. About 2.9 percent of people reported more than one race in the 2010 Census. Data for American Indians and Alaska Natives, Native Hawaiians and Other Pacific Islanders, and those reporting two or more races are not shown separately.

Source: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement. Fox (2008).

218. Another advantage of SPM is that it shows the impact on poverty estimates of not only cash benefits from the government but also the impact of non-cash transfers, taxes and other nondiscretionary expenses. Figure 2.5 shows the effect of various government programmes on the number of people who would have been considered poor in 2017. It shows the social security transfers and refundable tax credits had the largest impacts, preventing 27.0 million and 8.3 million individuals, respectively, from falling into poverty. Medical expenses were the largest contributor to increasing the number of individuals in poverty.

Figure 2.5

Change in number of people in poverty after including each element, 2017 (in millions)



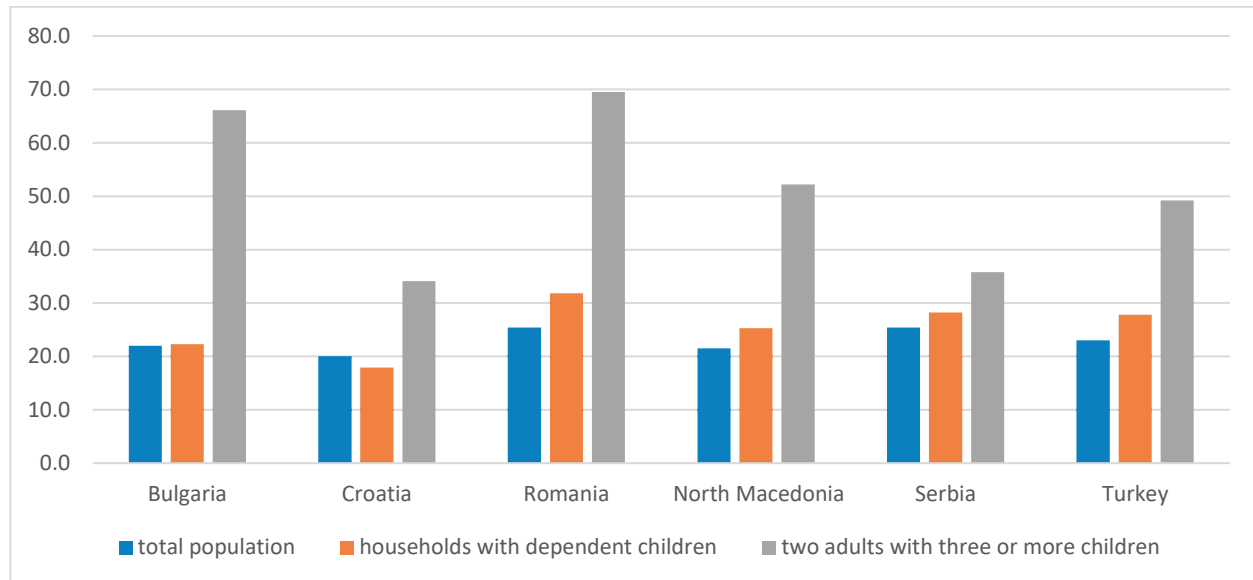
Source: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement. Fox (2008).

2.3.4 Example 4: Joint disaggregation of child poverty

2.3.4.1 Examples of relevant disaggregation of children poverty

219. Child poverty is often related to the composition of the household; therefore many statistical agencies disaggregate poverty statistics on the basis of the number of adults or children in the household.

Figure 2.6

At risk-of-poverty rates for the total population, households with dependent children and households with three or more children

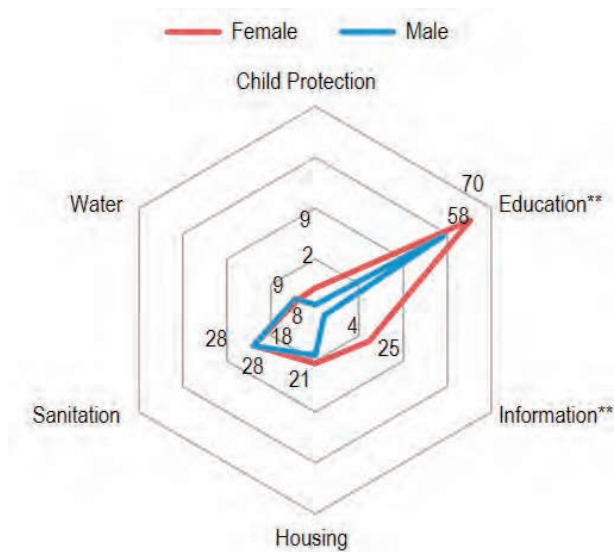
Source: Eurostat, Income and Living Conditions database online.

Note: The data for Bulgaria, Croatia, North Macedonia, Romania and Serbia is for 2015 and for Turkey for 2014.

220. In addition, where sample size allows, it is useful to disaggregate child poverty according to other criteria, notably disability status, ethnicity, or migration status. Because of the significant influence of adult working status (e. g. presence of pensioners, unemployment), and the sex/gender of the household head on child poverty it is particularly useful to also disaggregate according to these categories. Finally, stark differences between rural and urban child poverty are to be expected, which will ideally be reflected in further regional disaggregations, possibly even addressing inequalities within urban areas.

221. Disaggregation of multiple deprivations can be used to identify statistically significant differences even in very specific population groups. Figure 2.7 below shows the percentage of children between the ages of 15-17 defined as deprived by different dimensions of deprivation among Roma, Ashkali and Egyptian populations in Kosovo.

Figure 2.7
Number of children in Kosovo deprived as percentage, by dimension for children age 15-17 years, according to the gender of the child



Source: UNICEF, 2015.

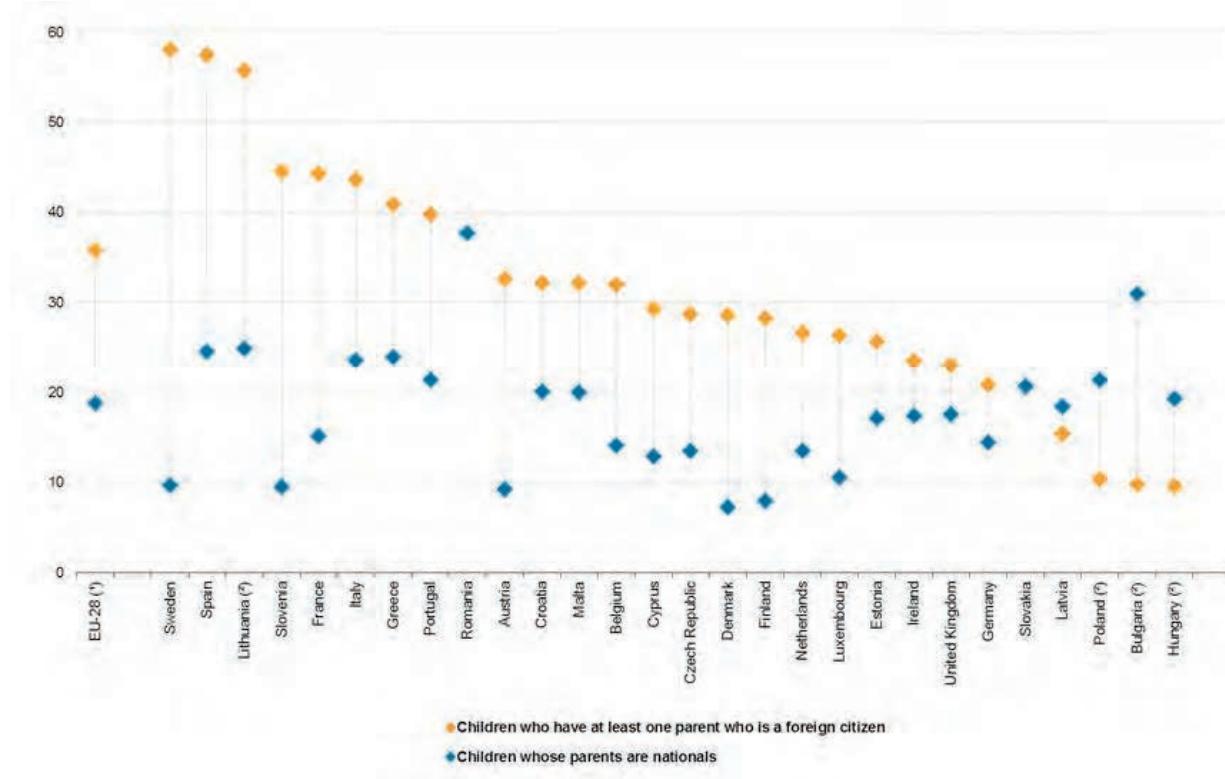
2.3.4.2 Children at risk of poverty according to citizenship of their parents

222. Eurostat provides an example of combining disaggregation of poverty measure for population groups defined by age and migratory status to make international comparisons. Figure 2.8 below shows differences in the risk of poverty for children in the EU Member States who had at least one parent with foreign citizenship compared to children whose parents were both nationals.⁴⁹

⁴⁹ <http://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/DDN-20180319-1>

Figure 2.8

Children in the EU Member States aged 0-17 years at risk of poverty, citizenship of their parents, 2016 (in %)



Source: Eurostat, Migrant integration statistics - at risk of poverty and social exclusion. Available from https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Migrant_integration_statistics_-_at_risk_of_poverty_and_social_exclusion.

Notes: Romania and Slovakia, data not available for children who have at least one parent who is a foreign citizen.

1/ Children with parents who are foreign citizens: estimate

2/ Children with parents who are foreign citizens: low reliability

2.3.4.3 A comparison of deprivation between children living in rural and urban areas

223. An example of an application of joint disaggregation poverty by age and urban/rural residence comes from the Multiple Overlapping Deprivation Analysis (MODA) conducted in Tajikistan. Figures 2.9-2.13 clearly show how rural children are much more likely to be simultaneously deprived in multiple dimensions compared to children in urban areas.

Figure 2.9
Venn diagram of deprivation overlaps between nutrition, water, housing, by urban and rural areas, children aged 0-4 years (Tajikistan)

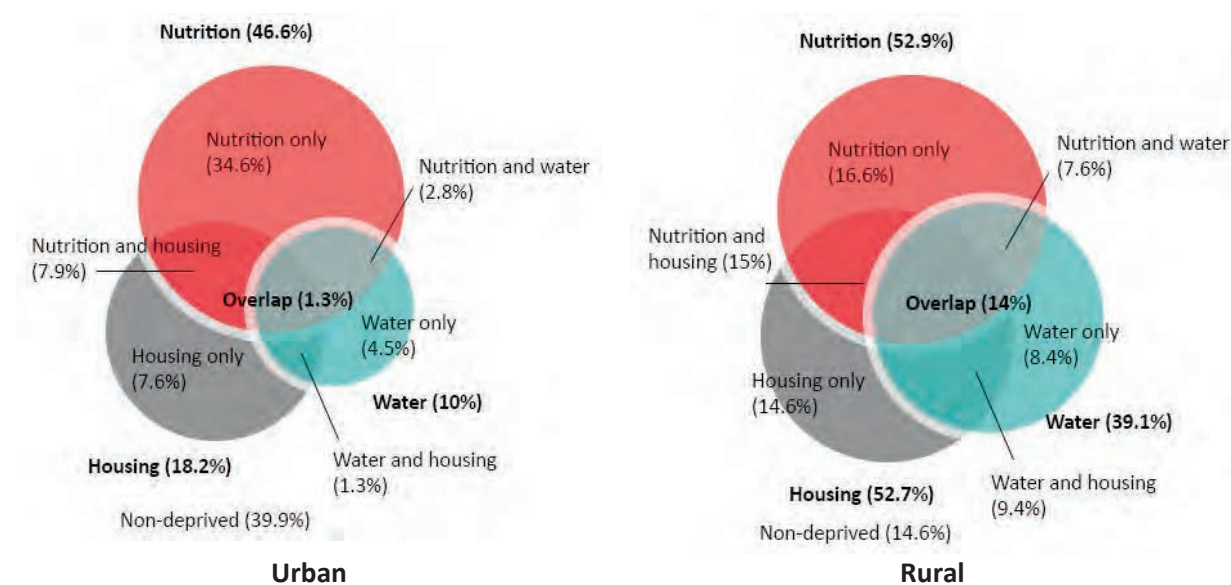


Figure 2.10
Venn diagram of deprivation overlaps between education, water, housing, by urban and rural areas, children aged 5-10 years (Tajikistan)

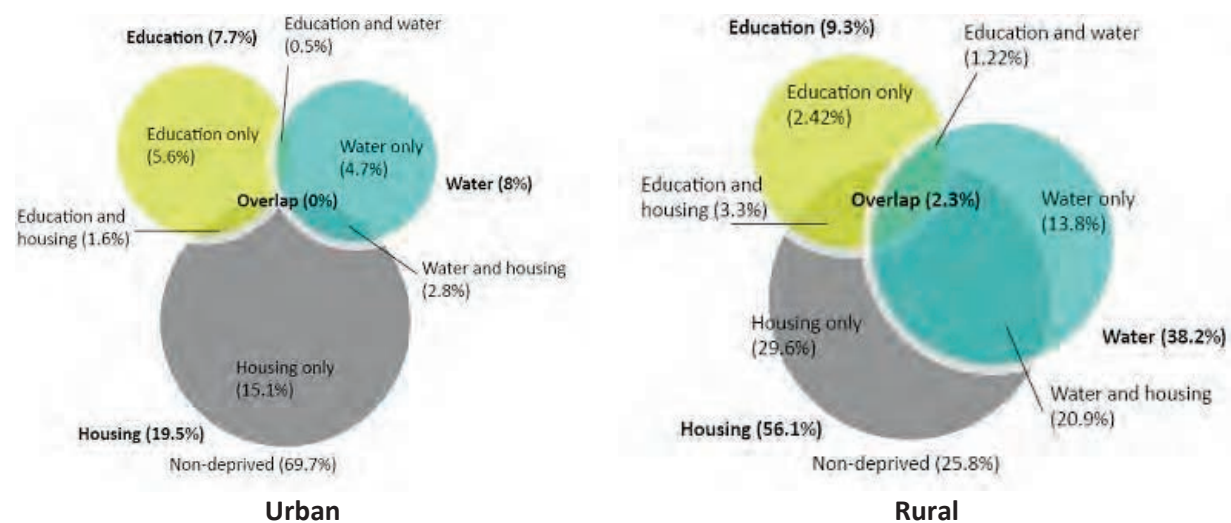


Figure 2.11

Venn diagram of deprivation overlaps between water, sanitation, housing, by urban and rural areas, children aged 5-10 years (Tajikistan)

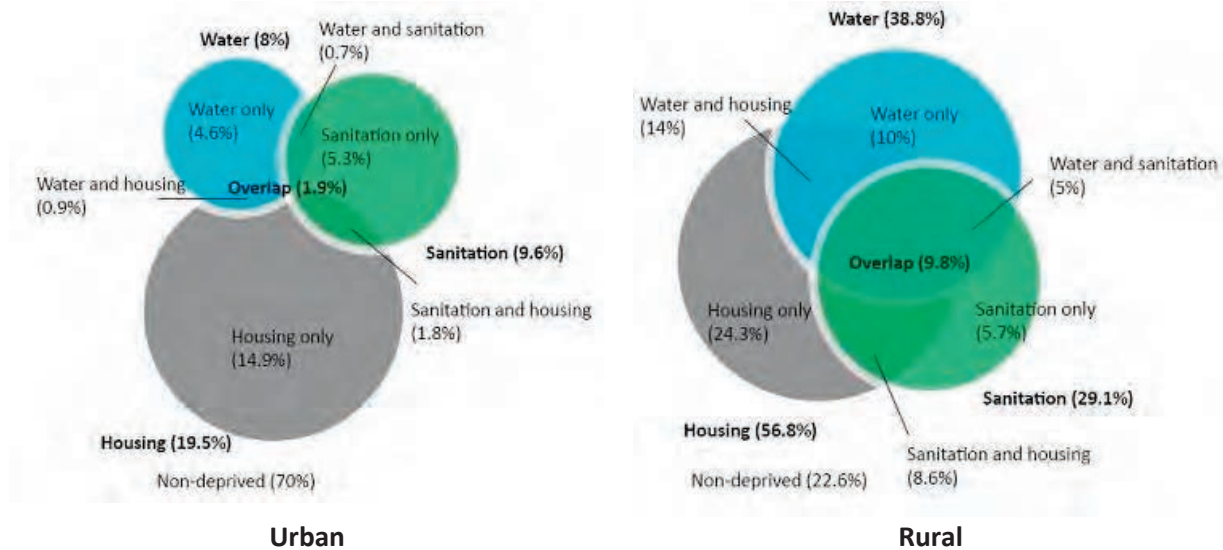


Figure 2.12

Venn diagram of deprivation overlaps between education, water, housing, by urban and rural areas, children aged 11-17 years (Tajikistan)

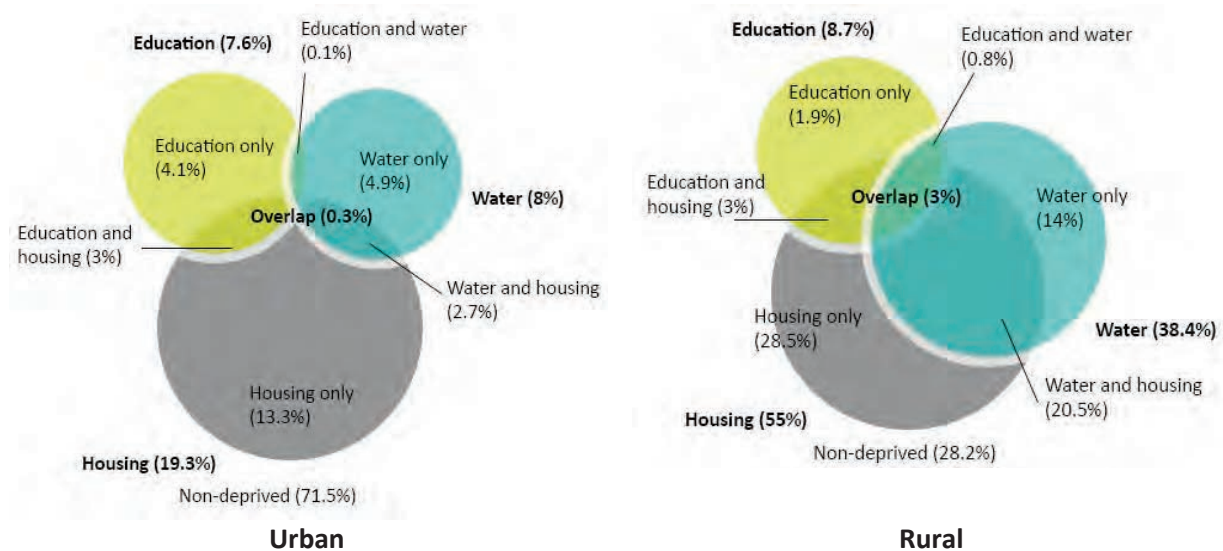
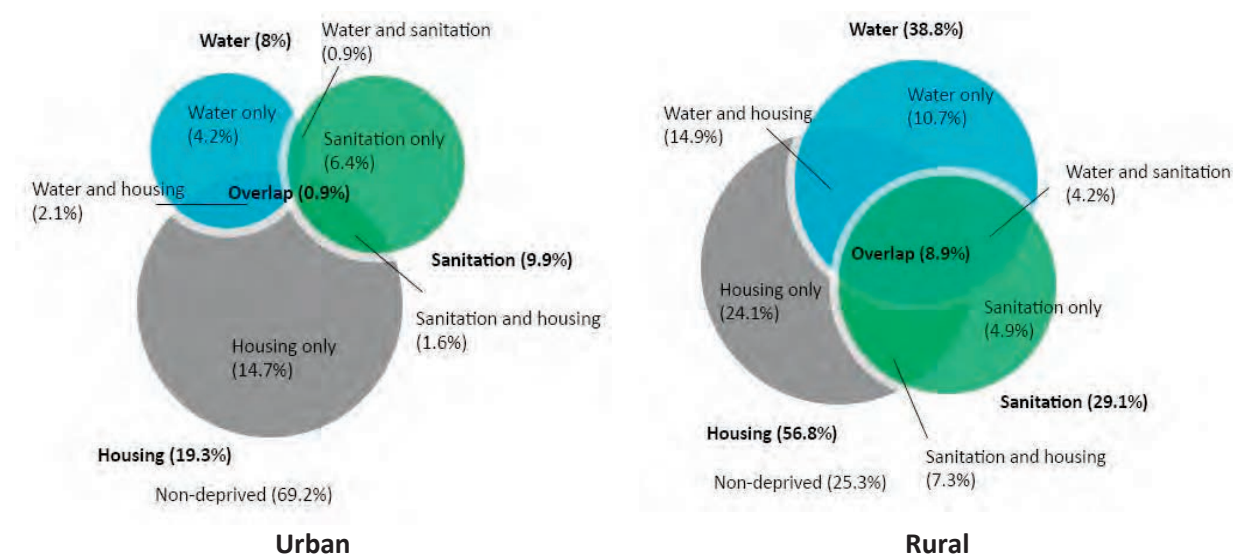


Figure 2.13
Venn diagram of deprivation overlaps between housing, sanitation, water, by urban and rural areas, children aged 11-17 years (Tajikistan)



Source: UNICEF, MODA, Tajikistan.

2.3.5 Example 5: Presentation of poverty across population groups

224. Poverty measures can also be presented across various population groups. In this example, the Philippine Statistics Authority presents poverty incidence for sectors as defined in poverty alleviation legislation.⁵⁰

⁵⁰ <http://psa.gov.ph/poverty-press-releases>

Figure 2.14

Poverty among the basic sectors in the Philippines

Source: Philippine Statistics Authority. Available from:

https://psa.gov.ph/system/files/kmcd/2015%20povstat%20for%20basic%20sectors_final_18Jul17.pdf

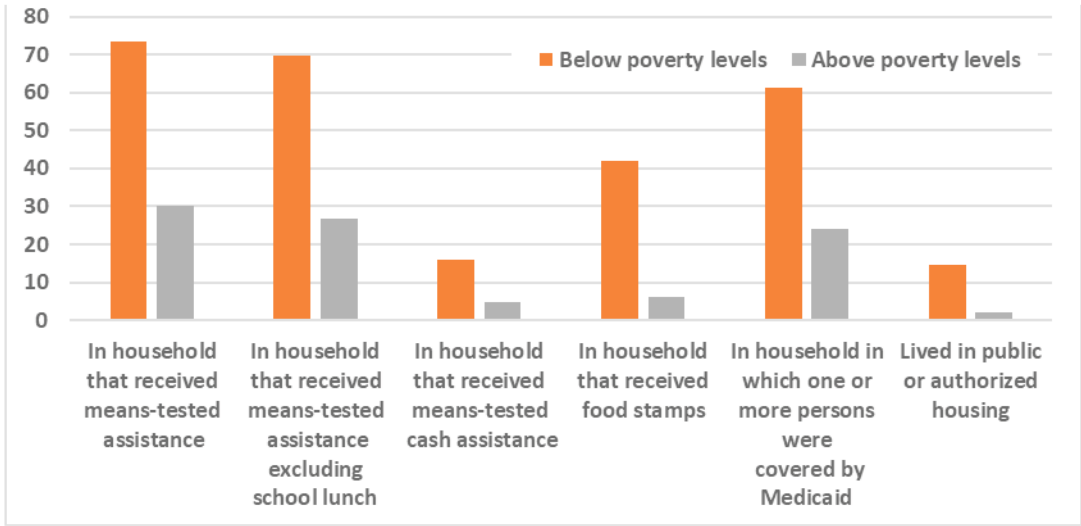
Notes: 1/ Merged 2015 family Income and Expenditure Survey and January 2016 Labour Force Survey

2/ Poverty Incidence is the proportion of families/individuals with per capita income less than the per capita poverty threshold to the total number of families/individuals.

2.3.6 Example 6: Poverty status and receipt of social transfers

225. The example of disaggregation of population by participation in various programmes and poverty status is taken from the U.S. Census Bureau Current Population Survey, Annual Social and Economic Supplement. Figure 2.15 shows the differences in participation rates in selected government programmes between persons below and above poverty line. The data are disseminated for different population groups defined by race and Hispanic origin and further broken down by age, sex, and household relationship.⁵¹

Figure 2.15
Program participation status of household (per cent of population) by poverty status of people in 2017



Source: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement. See <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-26.html>

2.3.7 Example 7: Interactive presentation of poverty indicators

226. Interactive visualisation tools can facilitate the use of poverty statistics. An example of such tool is a Eurostat interactive tool put in place on the occasion of the International Day for the Eradication of Poverty and enabling visualisation of data on risk of poverty or social exclusion for the EU28 and its member states for the total population as well as for selected breakdowns (sex, age, household composition and education level).⁵²

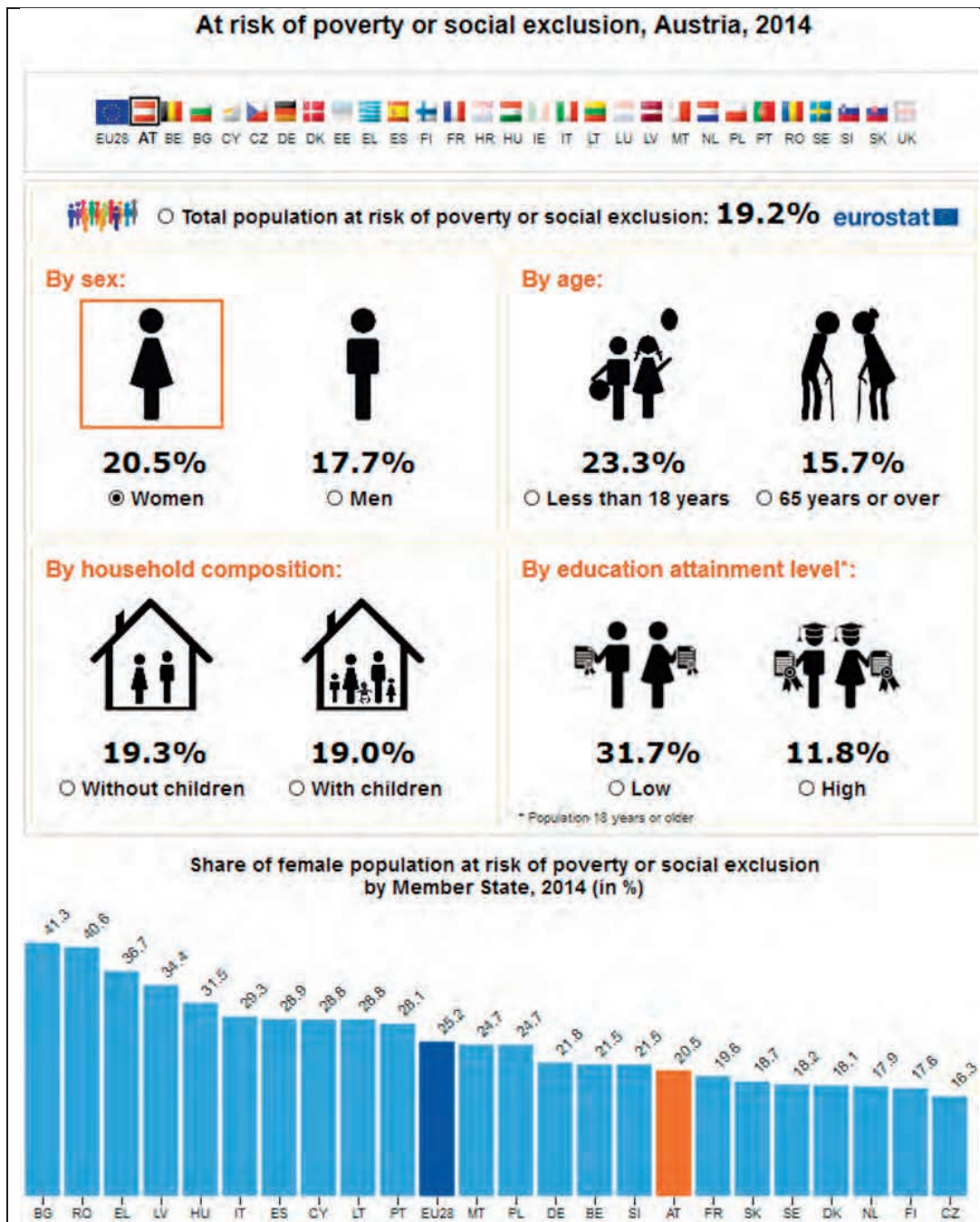
⁵¹ See more at <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-26.html>.

⁵² <https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/poverty-day>

227. The example below shows that poverty is more gender sensitive in Austria compared to the whole EU but less unequal for age, education or household composition.

Figure 2.16

At risk of poverty or social exclusion, Austria, 2014

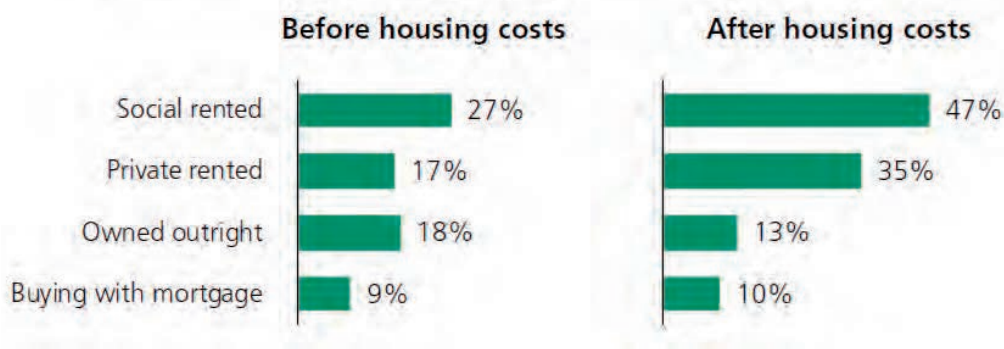


Source: Eurostat. People at Risk of Poverty or Social Exclusion by Age and Sex. Available from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_peps01&lang=en.

2.3.8 Example 8: Poverty and housing tenure (United Kingdom)

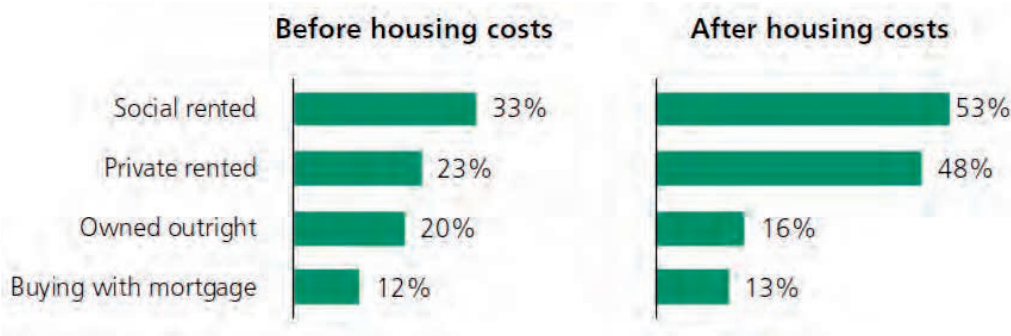
228. The following example of disaggregation of poverty by housing tenure comes from a research briefing “Poverty in the UK: statistics” (McGuinness, 2018) produced by the House of Commons Library as part of their work for members of the Parliament and their staff. The note provides statistics and analytical notes on the levels and rates of poverty in the United Kingdom, including historical trends and forecasts for future years and covers various domains, including poverty and tenure status. Figures 2.17 and 2.18 show that people living in social rented or private rented accommodation are more likely to be in relative low income after housing costs than home owners. The situation is even worse for children living in both the social and private rented sectors. Both figures also show the impact of treatment of housing costs in the calculation on poverty levels for various groups of housing tenure.

Figure 2.17
Percentage of people (all ages) in relative low income, by housing tenure, United Kingdom, 2016-17



Source: McGuinness, 2018.

Figure 2.18
Percentage of children in relative low income, by housing tenure, United Kingdom, 2016-17



Source: McGuinness, 2018.

2.4 Recommendations

Recommendation 1: In line with the guidance provided by the Office of the United Nations High Commissioner for Human Rights, development and use of measures for disaggregation of poverty should follow the Human rights-based approach to data with emphasis on principles of self-identification and participation.

Recommendation 2:

- a) In line with the Guide on Poverty Measurement and considering other emerging needs, especially in the context of SDGs, groups which are defined by the following variables are suggested as a priority for poverty disaggregation, wherever relevant and permitted by national legislation: age, sex, disability status, migratory status and ethnicity.
- b) In addition, the following disaggregation variables are suggested to better understand socio-economic processes and effectiveness of policies: household type, employment status, tenure status of the household, receipt of social transfers, educational attainment and degree of urbanisation. These minimum suggestions should be complemented by additional or more detailed disaggregations as required by the national context.

3 Including Hard to Reach Groups in Poverty Measurement

3.1 Introduction

229. Sometimes, legal frameworks for measuring poverty (including e.g. EU-SILC) define the population living in private households as a reference population, which excludes some of the most vulnerable populations such as the homeless or persons living in institutions. This apparent shortcoming reflects practical difficulties of data collection.

230. There are persons or groups in the population, which cannot be easily captured in general population surveys which are typically used for measuring poverty. Some vulnerable groups such as ethnic minorities may simply be too small or regionally concentrated to be adequately represented with proportional sample allocation. Sample precision may therefore be insufficient for statistical analysis. Moreover, some population groups require special effort beyond traditional and established survey methodologies. Short term labour migrants who perhaps do not speak the national language may be a typical example. Besides small and inaccessible groups, the term hard-to-reach also refers to elusive populations, meaning populations for which – by virtue of their characteristics, or of the lack of suitable sampling frames, or difficulties in obtaining the required information – adequate samples cannot be defined, drawn or implemented using the normal procedures of general population sampling. (Verma, 2013: 4). The UNECE Guide on Poverty Measurement (UNECE, 2017a) mentions coverage problems with other smaller groups, such as street children, drug users, sex workers, people who reside in institutions, including elderly care homes, children's homes, and mental health institutions.

231. There is an increasing demand from policy makers to cover population groups which in many cases have been excluded from general population sample frames as being hard to reach. This is particularly relevant for those countries which are committed to deliver on “Leaving No One Behind” principle of Sustainable Development Goals and Agenda 2030. Dedicated surveys on sub-groups of the population have brought significant methodological advancement and experience which can be extremely useful for improving fieldwork of general population surveys for the measurement of poverty. On the other hand, dedicated surveys often make use of specialised or simplified proxy measures for poverty, which may be added to conventional poverty measurement.

232. One of the main challenges survey researchers face is the absence of sampling frames (or their poor quality) for certain target groups. Alternative sampling methods include time-location sampling (e.g. Baio et al, 2011) or chain referral methods such as Respondent Driven Sampling (RDS, e.g. Heckathorn, 1997). Some of these methods have been developed recently and experiences are still being collected concerning their applicability and reliability, in particular in cross-country research. A recent large-scale example is documented in the technical report of the EU-MIDIS II survey which was coordinated by the EU Agency for Fundamental Rights in all EU-Member States (FRA, 2017a).

233. This chapter examines issues related to poverty measurement due to inadequate coverage for certain hard-to-reach populations. It first considers approaches to improve the inclusiveness of existing surveys to reach isolated populations. It then explores whether specifically dedicated surveys are required and discusses different options available to producers of statistics. Finally, it presents a range of examples that countries have adopted for capturing hard-to-reach populations in their areas.

3.2 Improving inclusiveness of poverty surveys

3.2.1 Fieldwork approaches for hard to reach populations⁵³

234. Tourangeau (Tourangeau et al, 2014) distinguish populations that are hard to sample, those whose members are hard to identify, those that are hard to find or contact, those whose members are hard to persuade to take part, and those whose members are willing to take part but nonetheless hard to interview. Surveys including hard to reach populations vulnerable to poverty face several measurement issues beyond but also related to its sampling and coverage. This concerns in particular the definition of a household, how to approach, how to identify these populations and how to minimise non-response. Fieldwork tools (e.g. respondents might be illiterate), the interview mode and the setting of the fieldwork (e.g. facilitators or recruitment of interviewers of the target population) have to be adjusted to better capture certain target groups. Standard survey tools and approaches might not cover such aspects sufficiently.

3.2.2 Principles to facilitate surveying vulnerable groups

235. The preparation of the survey needs to take into account the particularities of the target population and ensure that ethical and cultural aspects for minorities, and specific needs or vulnerability of the target group are addressed. To reach out to all populations in line with the Agenda 2030 the UN-OHCHR has formulated a set of principles, recommendations and good practices to ensure respect of human rights-based principles in data collection: Participation, self-identification, transparency, privacy and accountability in the design, collection and use of data (OHCHR, 2018).

236. Participation should be considered in all stages of a survey. For instance, to consult with stakeholders and members of the target populations throughout the survey cycle to develop and test survey tools. Stakeholders can help to establish contact with the hard to reach populations, through mediators, by hiring interviewers from the community or identification of locations where the target group congregates. It can also help to overcome prejudices and safety concerns among interviewers. For

⁵³ This section is inspired by examples and experiences from the second European Union Minorities and Discrimination survey (EU-MIDIS II) which was launched by the EU Agency for Fundamental Rights (FRA) in 2015/16. A detailed description of the survey methods used in the survey and described below is available in the technical report (FRA, 2017a).

example, during the course of both rounds of the UNDP-WB-EC Regional Roma Survey (UNDP and Ipsos, 2019) in 2011 and 2017, Roma organisations were engaged in sampling process and fieldwork implementation. Roma NGOs assisted in the process of identification and mapping of Roma settlements. Interviewers from Roma communities were recruited wherever possible to be part of interviewer teams. Roma interviewers were engaged by local Roma NGOs and through survey company's contacts.

237. Some groups are hard to identify or reluctant to self-identify. This might be due to historical persecution (Roma, Jewish) as well as mistrust in authorities or experiences of discrimination and exclusion due to ethnic, racial or religious origin. In some countries legal restrictions prevent data collection on ethnic or racial origin. For example, French law prohibits the collection of data based on race, ethnicity or religion. The principle of self-identification can overcome these barriers; it should always be applied for populations with sensitive personal identity characteristics. It can also help to overcome mistrust and reluctance to participate. The concept suggests that ethnicity, race, religion, gender identity and sexual orientation are fluid concepts and that the size of the total population remains unknown. Target population should be consulted beforehand in which way they would like to be self-identified. For example, when the EU Agency for Fundamental Rights (FRA) launched its dedicated surveys on Roma and Traveller populations, respondents were asked to self-identify selecting all relevant group definitions from country specific show cards. There are several Statistical offices asking for ethnicity and ethnic origin in their surveys and in the census, but questions on ethnic identity can only reflect a person's self-perception on and should allow for multiple identities. As another example, the 2017 UNDP-WB-EC Regional Roma Survey that covered Western Balkans region has used two control questions for respondents to self-identity on their ethnic background. Given post-Yugoslavian context where identifying their main ethnicity could lead to omitting to identify their Romani ethnic background, the respondents were given an opportunity to identify if they feel belonging to another ethnicity in addition to the first identified ethnicity.

238. The concepts of *privacy*, *transparency* and *data protection* are inherent to the principles of statistical offices and are important to overcome mistrust and reluctance to participate.

239. The fifth principle is accountability. The UN OHCHR (OHCHR, 2018) outlines that *as state institutions, national statistical offices are themselves human rights duty-bearers. They have obligations to respect, protect and fulfil human rights in their daily exercise of statistical activities. Independent statistics, free from political interference, are fundamental tools to inform and hold those in power accountable for their policy actions (or inactions).*

3.2.3 Interviewing hard to reach groups

240. Enhancing trust is key to overcoming some of the difficulties in reaching populations on the fringes of society. Schepers et al (Schepers et al, 2017) identified five widely adopted survey strategies in the literature on poverty:

- Doing (community-based) participatory research;

- Recruiting peer researchers or interviewers;
- Providing training, support and supervision for interviewers;
- Elaborating culturally appropriate questionnaires;
- Developing innovative data collection methods.

241. The field work should consider the specificities and living conditions of the target population and avoid intrusive or humiliating questions. Questions should be amended for group specific categories, e.g. when asking for type of housing for travellers to add the category “caravan”. Interviewer need special training and fieldwork tools and interview mode should be fit for any hearing, reading or writing impairments. For instance, FRA uses introductory videos in its Roma and Traveller survey (FRA, 2017c) in case of illiteracy instead of the introductory letter. The Australian Bureau of Statistics has developed cultural training for its interviewers to raise awareness on the impact that conscious or unconscious biases and stereotypical thinking may have when interviewing indigenous population. Pairing interviewers with persons of the communities who facilitate access, gender matching and offering interviews in the main language of the respondent are some of the tools which can help to facilitate fieldwork. Incentives used in general population surveys might have a different effect on minority groups and should be tested beforehand.

242. Interviewers *may assist and help persons with disabilities* – long-term physical, mental, intellectual or sensory impairments – to participate in the survey.

243. Fieldwork should provide a safe and respectful environment to ensure the privacy and dignity of respondents. This may include offering a space to do interviews outside home. For example, in FRA’s Roma and Traveller Survey 2019, interviews for Roma in the Netherlands rented a camper van for fieldwork to ensure privacy for respondents living in overcrowded housing.

244. The *mode of the survey* should be tailored to the needs of the specific population to make it accessible for all groups. Multi-mode surveys might be needed to reach out to hard to reach groups. For example, Persons with low literacy levels may find it difficult to participate in self-administered web or paper and pencil surveys. Sometimes non-probability designs such as opt-in online surveys are the only possibility to reach out to elusive populations (e.g. FRA surveys on LGBTI and Jewish populations)

245. Experiences of poverty and discrimination may cause distress among respondents. Interviewers should be trained for such sensitivities but also be able to provide contact information of support organisations to the respondents if needed.

3.2.4 Using simplified proxy indicators

246. Elaborating culturally and target group specific questionnaires poses specific challenges to comparative and cross-country research. To meet requirements for comparison between different population groups, as well as across countries, proxy indicators could be applied. Collecting full annual disposable income to measure monetary poverty as recommended in the UNECE Guide on Poverty

Measurement and the UN Canberra Group Handbook on Household Income Statistics might be not feasible. Administrative data do not cover hard to reach populations and income surveys do not recognise the different income situations. It might be humiliating to ask homeless persons on income from assets and too challenging to recall a full year of income. Other sources of income such as informal work, begging or donations are usually not covered in standard measures but should be included. A possible solution is to apply proxy measures of material living standards for comparative research.

247. For example, FRA uses current monthly household income in EU-MIDIS II to capture current living standards. In addition, questions on any paid work in the last four weeks as well as questions on child labour in the household are included. The inclusion of income in-kind is particularly relevant for some of the hard-to-reach groups, as it can often be their only form of income. Material deprivation indicators should address the specific living standards of the target population, such as lack of electricity or sewage system and segregation in housing. Other indicators such as arrears for payments of mortgage or credit cards showed high item non-response rates for Roma living in most deprived and segregated housing.

3.2.5 Training and field work materials

248. Interviewer training is key to successfully access some hard-to-reach populations and to reduce non-response. It is recommendable to include members of the community to be surveyed to the trainings. Interviewers, who are not from the target group, might be for the first time in contact with members of the target group and have no knowledge about the target groups apart from stereotypical descriptions. Experts, stakeholders and community members can reduce cultural bias, stereotype thinking and increase sensitivity of the interviewers for the data collection among the hard to reach group. The Australian Bureau of Statistics (ABS) has developed Cultural Protocols for ABS Staff engaging and working with Aboriginal and Torres Strait Islander People. ABS interviewers have to undergo cultural awareness training to improve communication and understanding of indigenous populations. FRA EU-MIDIS II and Roma and Traveller survey training materials address country and target specific context and interviewers get trained to respect human rights in data collections.

249. Field work materials should be easy to read and easy to access and interviewers prepared to explain verbally if reading competencies are not sufficient. Interviews should be provided in the language of the respondent when feasible. If the interview is conducted in the country's main language all field work materials should be offered in the language of the respondent for support.

Box 3.1

Practical challenges in surveying respondents in areas controlled by organised crime in Mexico

In the northern part of Mexico, organized crime poses a serious challenge to data collection for household surveys, particularly in the states of Nuevo Leon, Tamaulipas, Chihuahua and Sinaloa. Access to many roads is controlled by cartels that require travellers to seek permission to use the roads. Field representatives report that it is easy to identify roads that are controlled by organized crime because the cartels put trucks with armed personnel, known as "hawks", at the entrances to the roads. These

hawks report the number of people and vehicles traveling on the road. Field representatives must stop at commercial centres along the road to report what they are doing in the area, including whether or not they are just passing through the area in transit to another area. The Sinaloa cartel requires anyone visiting any town near the mountain to ask permission for access --- specifying the date, time and number of staff for each visit. Some areas are identified as high risk not because of cartel activity in the locations where the sample households reside but rather due to cartel control of the roads and locations through which the interviewers have to travel. In some cases, interview teams have asked municipal police or other authorities to escort them to and from the towns.

250. Fieldwork materials should include information on any barriers or access to certain areas and localities. The ABS census bureau and FRA in its Roma and Traveller Survey developed protocols on how to approach areas or communities which may pose a risk to interviewers or respondents and guidelines on how to proceed in case of restricted access (see forthcoming technical report of the FRA Roma and Traveller Survey 2019).

3.3 Dedicated survey programmes for hard to reach groups

251. For surveying special, hard to reach groups in the population, traditional survey methods need to be adapted for obtaining high quality samples. In some cases, completely different approaches need to be used or even developed. The extent to which more traditional survey methods have to be adapted or which alternative methods should be employed depends on the target group.

252. Before beginning to survey special target groups, detailed information on the group needs to be collected to help make an informed decision on the sampling approaches. This is a crucial step in developing sampling procedures for hard-to-reach groups. Information needs to be collected from all available sources, including experts and members of the target groups, who can provide important insights into the groups and opportunities for sampling. Detailed prior mapping of the target population can increase sampling efficiency or be essential to make probability sampling even feasible. Additionally, reliable background information can be used for weighting and post-stratification.

253. Groups might be hard to reach for different reasons. A group might be small but concentrated in certain geographical areas or dispersed more widely. Additionally, some people might be more mobile, such as travellers or temporary migrants, frequently moving within or across countries. This poses challenges not only to the sampling approaches, but also to the definitions used for target populations in-line with international definitions, such as the place of usual residence. For mobile populations the timing of the fieldwork is relevant.

3.3.1 Design modifications to include hard to reach groups in standard surveys

254. Hard-to-reach groups often are underrepresented or cannot be identified in standard surveys targeting the general population. If the population size is large enough to be covered within a general population survey, the survey design should be adapted.

255. Most importantly, the surveys need to include information which allows identifying the target groups. For example, for ethnic groups, the survey needs to ask respondents about their ethnic group, respecting the principle for self-identification. Respondents should be allowed to identify with more than one ethnicity to ensure the broadest possible coverage and respect multiple identities (Open Society Foundations, 2014). Likewise, information on disability and other characteristics which were presented in Chapter 2 needs to be collected.

256. If the numbers should be too small, strategies to over-sample special groups and tools to improve response rates are appropriate. To oversample certain ethnic groups among migrants, it may be necessary to use country of birth as a proxy for ethnicity. Additional and combined sampling approaches which are described in the next section may also be helpful. However, using proxy information can be challenging for certain groups and needs to be handled with care. For example, nationality-based data might be a very weak indicator for ethnicity due to differences in naturalisation policies, or for indigenous people or autochthonous ethnic groups.

257. Field work materials and approaches can be adapted to better reach out to a certain sub-population, for example by recruiting peer interviewers, special (cultural) training and translation of fieldwork materials. Several national statistics institutes (NSIs) already have developed approaches to cover special groups in their standard surveys. For example, in 2018 and for the first time, the Slovakian statistical office extended its EU-SILC survey to marginalised Roma communities in a separate sample and survey (see Box 3.). In Bulgaria, the national statistical office has been collecting ethnicity in the national SILC surveys since 2012. Also, the Bulgarian national statistics office, in cooperation with FRA, is currently designing a national representative survey of vulnerable communities. The survey will be implemented in 2020. There are efforts by several NSIs to better cover migrants in the Labour Force Survey, through translating survey materials, using multi-lingual or native interviewers and the use of language help lines. Moreover, several NSIs are examining response rates of migrant groups and ways to improve their participation and applying weighting methods (Barnes, 2008).

Box 3.2

Surveying marginalized Roma communities in Slovakia

In February 2017 the Updated Strategy Action Plans for Integration of Roma up to 2020 for years 2016-2018 was approved in Slovakia. One of the objectives of this strategy was to “improve the quality of collection and analysis data about Roma”. This was to be achieved by including “...residents of marginalized Roma community (MRC) in the regular EU-SILC”. Through the collection of this data, the

effectiveness of policies aimed at the MRC will be better assessed and evaluated leading to improved policy targeting.

Based on this directive, the project “Monitoring and evaluation of inclusive policies and their impact on marginalized Roma communities” was established, as a joint cooperation between the Statistical Office of the Slovak Republic and Office of the Plenipotentiary of the Government of the Slovak Republic for Roma communities (OPGSR). The main objective of this project was to ensure monitoring and evaluation of inclusive policies and assessment of their impact on the Roma population, with particular emphasis on the MRC.

The main outcome of the cooperation was the delivery of a specific EU-SILC survey module related to the MRC (hereinafter referred to as “EU-SILC MRC”); the taking place in 2018, and with a second planned for in 2020. The Statistical Office was responsible for implementing EU-SILC MRK and providing data to the OPGSR for further processing. The OPGSR assisted in developing the data collection methodology, and interviewer guidelines. They were also responsible for analysing and reporting the findings, and disseminating through the media and their website, with the aim to increase credibility of survey and facilitate the work of interviewers.

Basic characteristics of the survey:

- The target group are household members and households of MRC.
- The objective of the survey is to measure income and living conditions of the MRC in Slovakia.
- The targeted obtained sample is 1,000 private households. Therefore, the issued sample was 1,500 to account for expected non-response rate.
- The Pen and Paper Interviewing (PAPI) method was used. Survey responses were written by interviewers into paper questionnaires, which then served as entry for electronic processing. There were 3 types of questionnaires: SILC 1-01/A – Household structure; SILC 1-01/B – Household data; and SILC 1-01/C – Personal data for person aged 16 and over. All three types of questionnaires were based on the national EU-SILC, and then adapted for the MRC conditions.
- The Statistical Office tested questionnaires on a sample at least 20 households before the data collection, enabling further refinement of EU-SILC MRK.
- Those working on the project developed manuals for interviewers which included instructions on recording data and other aspects of the fieldwork. In addition, training, and a reward scheme, was organized for interviewers, as well as expenses covered for other logistical and communication costs associated with realization of the survey.
- The Statistical Office created a network of external interviewers for the purpose of collecting data in the field. It also collected management information on interviews including their educational attainment, experience and age. Each interviewer had an identification number, which was recorded in questionnaire.
- The Statistical Office ensured record and contact information about households (e.g., address), which to be used exclusively to create a longitudinal dimension at for the 2020 second survey.
- The micro data will consist of four basic files: file D (household register file), file R (person register file), file H (household data file) and file P (personal data file for all household members aged over 16).
- The Statistical Office will publish a quality report after completion of the survey.

The work started in June 2018 and fieldwork for the first survey was carried out during October and November 2018. The table below presents some initial results from the exercise, comparing the proportions of the MRC and total populations which are below 60% of the national median household income. The MRC are much more likely to be at-risk-of-poverty, with their poverty rate almost 75 percentage points higher than the whole population rate.

Table 3.1

Proportion of population with household income below 60% of national median

	MRC	Whole population	Difference
Total	86.7	12.2	74.5
Single adult household	74.4	17.2	57.2
2 adults, both aged below 65 years old	83.0	9.5	73.5
2 adult household, at least one aged 65+	48.0	3.6	44.4
single parent household with at least 1 dependent child	95.1	36.7	58.4
2 adult household with 1 dependent child	80.0	11.4	68.6
2 adult household with 2 dependent children	83.9	14.5	69.4

Box 3.3

Surveying the Roma, Sinti and Caminanti (RSC) populations in Italy

In 2018 Istat and the National Office against Racial Discrimination (UNAR) signed a new agreement aimed at developing a system of indicators to assess the effectiveness of inclusion policies toward the Roma, Sinti and Caminanti (RSC) populations, with a special focus on housing. The Italian National Statistical Program 2017 - 2019 (in its update 2018 - 2019) has included a "work - sheet" (02741; renewed in the National Statistical Program 2020 - 2022) scheduling survey activities on Roma people.

A web survey (involving UNAR and RSC Associations) was launched in October 2019, in order to assess all housing inclusion projects carried out by municipalities (all municipalities over 15,000 inhabitants; 745 municipalities) and regions. The scope of this background survey was to map all the projects carried out on their own or in cooperation by municipalities, targeting housing exclusion and with the objective to overcome formal and informal settlements. The questionnaire included also a section that will allow updating the picture of settlements in Italy. The further step will focus on the comparison between living conditions of RSC people who left the settlements and those still living there (in formal and informal ones), in the same territories.

Questions will deal with living conditions and in particular with the issues of employment and access to education and health. Information will be also collected on social and demographic variables (age, sex, etc.), access to basic housing facilities (water, electricity, etc.), access to public facilities (bus, trains, hospitals; etc.). The goal of the survey is to build a system of indicators that can monitor regularly the gap in terms of inclusion between RSC in public or private dwellings, and the ones still living in settlements. A mixed qualitative-quantitative research method will be adopted, as the target group are too small to allow only a quantitative approach. The core principles of the survey are: self-identification; voluntary and anonymized data collection, provided with the explicit consent of individuals; involving Roma representatives through the National Roma Associations Platform.

3.3.2 Special survey programmes

258. When the design of standard surveys cannot be sufficiently adjusted to capture certain groups, special survey programmes need to be considered. National Statistical Offices may often lack budget and

capacities to achieve coverage for very small social groups. The development of new methodologies and capacity building is therefore essential. Often this will require joint effort of countries and international organisations. One example is the activities of the EU Agency for Fundamental Rights (FRA). In 2015, FRA launched EU-MIDIS II as a dedicated international survey programme for target groups. It focussed on selected groups of immigrants and ethnic minorities in all 28 EU Member States. The groups included immigrants and descendants of immigrants from specific countries and regions of origin and ethnic minorities. Between one and three groups were selected per country. The results included some simple measures of poverty and deprivation and broad information on discrimination experiences, experiences with violence and harassment, social and living conditions and integration (FRA, 2017b).

Box 3.4

Generating data on hard-to-reach populations at risk of violation of their rights in Bulgaria

The National Statistical Institute of Bulgaria committed to a partnership with the EU Agency for Fundamental Rights to develop and test innovative methods to reach out to specific vulnerable groups at regional and local level. The project aims to collect data for policy purposes and to monitor progress on key SDG indicators between 2018 and 2020 and is supported by the EEA/Norwegian Financial Mechanism.

The project responds to the demand for data disaggregated by various vulnerability criteria (ethnicity, gender, disability, age, region, etc.) needed for informing and monitoring a number of key policy frameworks such as 2030 Agenda for Sustainable Development, poverty reduction, social inclusion, anti-discrimination or hate crime.

Standard statistical surveys tend to focus on selected (sector-specific) dimensions of vulnerability making in-depth analysis of intersectionality or correlating various drivers of vulnerability extremely challenging. The first objective of the project was to come up with a balanced set of statistical indicators relevant to policy frameworks and which can be generated for selected vulnerable groups through a single-source data collection. Based on this policy-driven selection, the team defined the instrument of a big-scale face to face survey. The survey has a net sample size of 15,000 households and is expected to generate representative data on the situation of Roma, children at risk, old-age persons and people with disabilities. Apart from the group coverage, the data would allow in-depth analysis by other characteristics (demographic, territorial, discrimination experience etc.). The questionnaire combines modules of EU-SILC, LFS, European Health Interview Survey (EHIS), UNICEF MICS and the FRA Minorities and Discrimination surveys. Particular attention was given to retaining comparability with data generated through these large-scale surveys.

The sample shall make it possible calculating key indicators for the general population and the key vulnerable groups disaggregated to the level of Bulgaria's 28 regions. Using the information from the survey and the data from 2021 Census, the team will conduct small area estimation or other methods for statistical modelling to calculate key indicators also at the level of municipality. Based on this a number of "vulnerability maps" will be constructed to visualize the territorial distribution of key SDG indicators.

Another important component will be testing methods for more effective usage of data from administrative records combining/matching data from different sources. In view of the methodological, procedural and data protection challenges, the project will investigate at a minimum the feasible options. At maximum, it would come up with a set of tested methods for optimizing the process of

generating policy-relevant statistical data on the situation of people at risk of poverty, social exclusion and violation of their rights and the drivers that put them at such risks.

Box 3.5

The first nationwide Survey on persons with disabilities in Germany

The first nationwide survey among persons with disabilities in Germany is scheduled from 2017 until 2021. The social research company Infas was commissioned by the Federal Ministry of Labour and Social Affairs. The operation involved a large-scale screening survey with more than 300,000 households in 256 sample points. This screening yielded activity limitations among 137,000 individuals which provided the basis for a face to face follow up. Until now more than 19,000 interviews have been conducted with persons with disability and without disability. The target sample comprises 16,000 Individuals with disabilities and 5,000 without.

In parallel, there was a survey among persons with disabilities in institutions which required substantial preparatory work. A sample of 2,900 institutions has been drawn from a total population of approximately 21,000 institutions to conduct a pilot study. This was used for a screening with regard to persons with disabilities. The main sample was drawn out of only those institutions which confirmed the presence of persons from the intended target group. These interviews have just begun in November 2019. By December 2019, 1,300 interviews had been completed. The target number was 5,000 interviews. The whole operation is scheduled to be completed in spring 2020.

The operation included a strong participatory element to identify relevant survey questions on the living conditions of persons with disability. Also, extensive testing of survey instruments was involved. Measures to ensure accessibility included easy reading material and videos (see Harand and Kerstin, 2019). The survey will not produce conventional monetary measures of poverty but is likely to allow for the construction of multidimensional indicators of deprivation.

3.3.3 Sampling target groups

3.3.3.1 Conventional sampling approaches with methods to increase the efficiency of screening

259. In the absence of a sampling frame of the target population, screening methods may be appropriate. These include multi-stage sampling, where first predefined small geographical areas are sampled (Primary Sampling Units, PSUs) and, then screening households for members of the target population. For multistage sampling, a list of PSUs with information on the density of the target population is needed for more efficient sampling. Such lists are not readily available for many groups and often have to be either taken from unofficial sources or through approximation via other variables. For example, associations representing certain groups sometimes have information on where their members live. Sometimes these lists need to be created specifically for a survey with the support of experts. While these lists are important for increasing the efficiency of sampling, they are often based on assumptions and external identification of the area. The data can be verified or updated after fieldwork (for example through comparing eligibility rates with estimates from the area sampling frame) but can only be confirmed through the principle of self-identification at the level of respondent. Due to the sensitivity of

the information even in area sampling frames, the data need to be handled with care considering data protection aspects as well.

260. The multistage sampling approach often needs to be further adapted depending on the resources for the survey and characteristics of the target population – most notably the (estimated) densities of the target group in the PSUs. If screening is feasible in the second stage, it is primarily related to the size of the target population in the respective PSUs.

261. Several methods to increase the efficiency of screening have been employed in the FRA EU-MIDIS II (FRA, 2017a):

- **Coverage reduction**

One approach is to reduce the geographical coverage of the PSUs included based on the density of the target population in the PSUs. In EU-MIDIS II, depending on the feasibility related to the situation in each of the countries and target group, PSUs, where the proportion of the target group falls below a certain threshold, were excluded to reduce the burden of the interviewer of screening. For example, in most countries where Roma have been surveyed, PSUs with less than 5-10% of Roma living there, based on the estimated size of Roma, were excluded. This led to a limited coverage of 6% to 8% of the total Roma population in the nine countries that covered Roma in EU-MIDIS II. With this approach, it needs to be assessed to what extent the coverage minimization impacts on the validity of the sample. Additionally, further adjustments can be made to make sampling more efficient, such as oversampling or clustered sampling, as described below.

- **Oversampling of higher density strata**

In case of availability on the proportion of the target group in the PSUs, those with higher concentrations can be oversampled. This may increase the design effect of the sample and reduce the effective sample size. However, it makes it more realistic to obtain a certain number of interviews.

- **Focused enumeration**

Focused enumeration is a strategy to further reduce the burden of having to screen many households before finding members of the target group. It involves proxy screening through other people. Interviewers have a certain number of core households selected at a PSU; be it via address sampling or other systematic selection, such as random route. At any of the core households selected for screening, the respondent is asked if their neighbours (e.g. the two to the right and the two to the left) are members of the target population. If the respondent indicates that none of the neighbours belongs to the target group, the interviewer continues to the next core household on the list. If the respondent at the core household indicates that at a neighbouring household is a member of the target population or is unsure, the interviewer visits this household and screens the respondents. This approach has the advantage of reducing the screening burden and allows locating respondents much faster, without having to adjust the weights for the additional procedure. The main disadvantage is that it is not feasible and unethical for some population groups. For example, it is not

acceptable to ask people to identify others based on their ethnicity or sexual orientation.⁵⁴ Additionally, if respondents do not know if their neighbours belong to the target population this may introduce bias. This was consequently only used in some countries in EU-MIDIS II, but not with ethnic minorities.

- **Adaptive cluster sampling**

To avoid proxy screening, an alternative is adaptive cluster sampling, which works on the assumption that some target groups are clustered in small geographical areas (e.g. blocks, neighbourhoods). The method involves normal screening in PSUs, and then once a member of the target population is found, the interviewer also screens the neighbouring household to the left and to the right. If a member of the target population is found in the neighbouring household, the interviewer continues to the next neighbour until no member of the target population is found. This method reduces the burden of the interviewer and makes screening more efficient.

This strategy was successfully used in EU-MIDIS II in several countries. However, the method also has drawbacks, as it is complicated to implement for interviewers and potentially error prone, if not explained well. Additional measures need to be taken, such as cutting the number of interviews in clusters of neighbouring households of core households. The weighting needs to adjust the selection probabilities based on the cluster size.

262. Still, for some target groups, multi-stage sampling with screening is not feasible, even if the methods above for boosting the efficiency have been considered. This is particularly the case for very rare and scattered target groups. Many alternative methods for reaching hard-to-reach groups exist depending very much on how the respective group is connected, organised and spread across regions. The next section describes such methods.

3.3.3.2 Alternative sampling approaches

263. As mentioned above, detailed ethno-graphic research is necessary before any alternative sampling method is used. This is because alternative methods rely on assumptions of certain characteristics of the group, which are essential for the success of the method to recruit respondents and weight the samples. Groups can be geographically clustered at specific locations or they can be well connected through personal linkages across members of the target group. The latter situation allows for referral methods, most notably so-called Respondent Driven Sampling (RDS). The former situation means that location sampling methods could be applied. Both methods – with a focus on location sampling – are described briefly.⁵⁵

- **Location sampling**

If a target group is connected in terms of frequently meeting or assembling at a known number of locations, location sampling can be considered. The method, sometimes known as the method of

⁵⁴ Self-identification is one of the principles and recommendations for human rights-based data collection. See OHCHR, 2018.

⁵⁵ For a more comprehensive overview of methods, see for instance Marpsat and Razafindratsima, 2010; Kalton, 2009 or Verma, 2014.

centres of aggregation (Baio et al, 2011) or intercept method (McKenzie and Mistiaen, 2009) is more established in the literature. It is theoretically well developed; however, it has not been applied very often in practice. The method involves collecting information on all possible locations where the target group might congregate. For example, in churches, Non-Government Organizations (NGOs), cultural or ethnic associations, service centres, open places and so forth. There is no restriction in terms of definition of locations, as long as they are clearly geographically defined and can be surveyed. After a list of locations has been compiled, either all or a random sample of locations are visited and surveyed. Respondents are randomly (usually systematically) selected at each location. The crux of the method is to ask every respondent, which of all the locations they attend. This way the probability of selecting respondents can be calculated and weights provided (Baio et al, 2011).

The quality of the sample in terms of representativeness depends on how well the locations cover the entire target population. In practice, there are several topics to keep in mind when implementing the method. Clear guidelines on the how respondents are selected at the different locations need to be prepared to ensure simple random sampling at each location. The quantitative importance of each of the locations in relation to each other needs to be estimated as close as possible. This way a more efficient sample will be obtained. For instance, in locations where not many people go to, fewer interviews should take place; otherwise the weights will increase the impact on the design effect. It is possible that at some locations, interviewers are not allowed to select respondents and carry out fieldwork. This way of non-response influences the opportunities to obtain a representative sample of the target population.

Sample sizes per location should not be very low (e.g. below ten, ideally not below 30), because this limits the opportunities to obtain good estimates of the overlap of locations and hence the calculation of weights. The method is limited to smaller geographical locations (e.g. cities).

Box 3.6

Previous results and future development of statistical information about homelessness in Italy

In 2014, the second survey (the first one was carried out in 2011) on the condition of persons living in extreme poverty was conducted, following an agreement amongst Istat, the Ministry of Labour and Social Policies, Italian Federation of bodies for homeless persons (fio.PSD) and Caritas Italiana.

The sampling base was represented by the benefits provided (meals served and beds) at the types of services taken into consideration (soup kitchens and night shelters). During the survey month (21 November-20 December 2014), 639 centres were visited: 174 soup kitchens at lunch, 89 soup kitchens at supper, and 376 night shelters. It is estimated that 50,724 homeless persons, in the months of November and December 2014, used at least one soup kitchen or night shelter in the 158 Italian municipalities where the survey was carried out. This amount corresponds to 2.43 per thousand of the population regularly registered with the municipalities taken into consideration by the survey, a value higher than three years earlier, when it was 2.31 per thousand (47,648 persons).

National Statistical Program 2020 - 2022 includes a "work - sheet" (02592) aimed at carrying out a new edition of the survey in order to provide policy makers and users with a complete and updated representation of poverty and inequalities in our country.

- **Respondent Driven Sampling (RDS)**

RDS is a method that allows the calculation of selection weights based on referral of respondents. It means that several initial respondents are selected non-randomly (so-called seeds) and each of them is asked to refer one or more other members of the target groups. Based on the assumption that the selection of other respondents is not influenced by previous selection of the respondent, weights can be calculated. This method is tested successfully in some cases, for example in the case of Ukrainian immigrants in Warsaw (Kaczmarczyk, 2013). However, the method has failed to produce samples in other cases – such as in the case of recent immigrants to selected EU countries (Frere-Smith et al, 2014). Therefore, one needs to be sure that there are enough connections in the target group and that respondents are willing and able to refer to other members of the target group.

FRA has tested this methodology in 2012 in an online survey with Jewish respondents in selected EU Member States. It has not produced the desired numbers of interviews, which led to a change in the sampling strategy (open opt-in online survey, see FRA, 2013). Given these challenges, the method was not considered for EU-MIDIS II. It might, however, be well suited for other target groups and surveys.

- **Mixture of conventional with alternative sampling methods**

It is important to note that sampling methods can be combined, which can deliver very promising results. Most notably, if insufficient sampling frames are available, the sampling methods can be extended by combining them with alternative methods.

For example, in the FRA EU-MIDIS II the population register in Poland was considered to not capture well undocumented migrants and hence was combined with location sampling. The sample from the population register was considered as one location. In Cyprus, location sampling was combined with two stage sampling and screening for immigrants and their descendants from Asian countries.

A survey among immigrants from outside the EU was conducted in 2011 in Budapest, where the population register was also considered one location and other locations were included as well. The results of the survey are very consistent with the results from the census 2011 on the respective target group (Reichel and Morales, 2017).

264. An exceptional case in EU-MIDIS II was the sampling method employed in Germany, where immigrants and their descendants from Sub-Saharan African countries and Turkey were surveyed. A sampling frame of the target population was used that identified potential members of the target population based on their names (so-called onomastic sampling). This sampling frame was deemed incomplete for the target population and was consequently extended through referral methods. Respondents were asked to provide contact information on other people of the target group living in the same PSU. Based on the total number of referrals the respondents provided (network size) the selection probabilities could be calculated.

Box 3.7

UNDP – Sampling approaches in previous Roma Surveys

Surveys are understandably better suited to cover easily accessible populations. Capturing those who are difficult to sample, identify, contact, persuade or interview requires a more complicated - and often costly - approach. These challenges must be met with innovative strategies, such as those highlighted in Table 3.2 below, which outlines various sampling approaches and identification techniques in previous Roma surveys.

Table 3.2

Summary of sampling approaches in previous Roma surveys

Survey	Roma identification	First stage sampling	Second stage sampling
2011 UNDP-World Bank-EC Regional Roma Survey ⁵⁶	Implicit endorsement of the external identification Non-rejection of “We are conducting a survey among the Roma population . Would you mind being interviewed?”	Random sample of municipalities on the basis of lists of settlements from Census with average and above share of Roma updated with information from other relevant sources.	Random start and equal random walk
FRA Roma Survey 2011 (FRA, 2014)	Self-identification Screening question “Could you tell me if any Roma live in this household? By ‘household’, I mean all the people who live within this accommodation and with whom you share cooking and meals.”	Selection from lists, which included only localities where the proportion of Roma was higher than the national average, take into account the region, the urbanisation level and the size of the Roma population of each primary sampling unit ⁵⁷ .	Mixed areas (10% to 40% Roma): Simple random walk with focused enumeration. Segregated areas (40% and higher): Simple random walk with a sampling step of five. Camps: Random walk rules in camps were applied to the extent possible
FRA EU MIDIS I Survey 2008 (FRA, 2007)	Self-identification	Primary sampling units were randomly selected in the medium and high density areas where targeted minorities lived in high concentration	Standard random walk procedure Focused enumeration was applied in order to boost the efficacy of the random walk approach.
FRA EU MIDIS II Survey 2016 (FRA, 2017a)	Self-identification Screening question “Is there anyone aged 16 or over living in this household who is Roma?”	Primary sampling units were selected with probability proportional to size of the target population within strata.	Random walk sampling techniques Adaptive Cluster Sampling (ACS) was applied in areas where Roma were estimated to make up less than 25 % of

⁵⁶ <https://www.eurasia.undp.org/content/rbec/en/home/ourwork/sustainable-development/development-planning-and-inclusive-sustainable-growth/roma-in-central-and-southeast-europe/roma-data.html>

			the population (in Bulgaria, Hungary and Romania.)
MICS Roma settlement Surveys (2010-ongoing in several West Balkan Countries)⁵⁸	Self-identification at listing stage	A random sample of enumeration areas (cluster of households) was selected with probabilities proportional to size within each stratum at the first stage.	Random selection from lists Household lists were updated by The Statistical Offices. Based on the orthophoto map, description of the enumeration area and listing, the interviewers' task was to go to the addresses listed and to identify the current Roma households, together with the number of children under five living in the household.
UNDP-WB-EC Regional Roma Survey 2017 (UNDP and Ipsos, 2019)	Self-identification screening question: 'Which ethnicity do you feel you belong to?' and 'Do you feel belonging to another ethnicity in addition to what you stated above?'.	Multistage clustered random probability sample design: Random sample of municipalities on the basis of lists of settlements from Census with average and above share of Roma updated with information from other relevant sources. The PSUs were selected with probabilities proportional to number of Roma in each PSU.	Within each PSU households would be selected randomly via random walk approach.
FRA Roma and Travellers Survey 2018 (in six EU Member States) (FRA, 2017c)	Self-identification	Primary (areas) and secondary sampling units (halting sites) were selected with probability proportional to size of the target population within strata.	Random selection or full census in halting sites in three countries Quota sampling and approaches through mediators if random selection was not possible in four countries.

⁵⁸ <http://mics.unicef.org/surveys>

Box 3.8

Special surveys on the Roma population in the Multiple Cluster Survey (MICS)

With a population of roughly 10-12 million – and half of them estimated to be under the age of 18 years – the Roma population constitutes one of the largest ethnic minorities in Europe. To better monitor equity and inclusion of Roma children, Multiple Indicator Cluster Surveys (MICS) have been conducted for Roma Settlements, in parallel with another MICS conducted for the whole country population. This was done between 2011 and 2014 in Montenegro, the Former Yugoslav Republic of Macedonia, Serbia, Bosnia and Herzegovina, and Kosovo (UNSCR 1244).

The MICS Roma sample is specifically designed to allow for comparison of data between the national and Roma population. MICS in Roma Settlements, as well as MICS for non-Roma population, uses census sample frames to identify clusters with high Roma concentration. These are defined as a minimum number of Roma households in each country as a threshold for selection of the enumeration areas. The average sample size for Roma MICS surveys is roughly 1,500 households, and replacement is not applied. Households for interview are selected through random systematic selection procedures once a listing of households has been conducted to update the sampling frame.

Data shows that larger shares of the Roma population, compared with non-Roma, lack access to basic services and knowledge to better prevent and deal with illnesses. Prior to the current round, MICS used a wealth index based on household level assets. MICS6 will allow the calculation of multidimensional poverty, enabling comparisons between Roma and non-Roma populations.

3.3.3.3 Non-probability approaches

265. In some cases, traditional and alternative sampling approaches are not possible. In this case, non-probability samples are the only option left. The application of quotas, systematic and widespread recruitment can support a heterogeneous sample. While no inference to the total target population can be made, the results still help to contextualise and measure the occurrence of certain events. For example, FRA carried out an open opt in online survey with over 90,000 Lesbian, Gay Bisexual, and Transsexual people in the EU. The results were weighted based on assumptions on the proportion within and across countries and provided information on the fundamental rights challenges this population faces in Europe (FRA, 2013b).

3.3.3.4 Weighting

266. The probability methods described allow for the calculation of selection probabilities and design weights. Experienced statisticians should be involved already in the design phase. The calculations of weights for RDS and location sampling require special procedures to adjust for the sampling approach. The latter was developed for the FRA EU-MIDIS II and is based on the information on which locations are attended by the respondents. A detailed description of the weighting approaches of the methods used in EU-MIDIS II can be found in the Technical Report (FRA, 2017a).

267. EU-MIDIS II also used non-response weights based on the respondent (if available) and neighbourhood characteristics. As a third approach, post-stratification weights were used, however, only in rare cases, because of the lack of reliable statistics on the target population in most countries.

268. When different samples are combined within and between countries an adjustment for estimated total population sizes is essential to maintain comparability. Weighting procedures should be harmonised across countries and prevent an excessive range of weights.

269. When weighting is implemented following the sampling approaches outlined above, there needs to be an assessment as to whether the weights can be applied for the non-coverage that is potentially introduced. Additionally, in some cases, the concentration of the target group in the PSUs are based on approximate estimates, sometimes based on the number of individuals, and sometimes the number of households, depending on the availability of estimates. Such estimates can be corrected after fieldwork through better assessments of the concentration of the target group, based on eligibility rates after screening, and of the household sizes to turn estimates of the individuals into household estimates or vice versa.

3.4 Including qualitative information in quantitative data collections through micronarratives

270. Marginalized population segments may not only be underrepresented or lost by definition of the survey scope. An additional set of problems concerns the willingness of individuals from marginalized, hard-to-reach groups to provide information to survey interviewers, especially about the type of sensitive matters that define their marginality (undocumented migrants), or difficulty in being interviewed (low literacy groups, or where there is a language barrier). It can be argued that the same social characteristics and constraints that hinder access to these individuals may also impair their willingness or ability to answer survey questions. These types of bias led to significant underestimation of poverty rates, for example in EU-SILC (Nicaise and Schockaert, 2014). It is therefore also important when reporting estimates to inform the user about the potential sources of bias.

271. Often it will be inevitable to support quantitative studies with qualitative information. This may add further value by allowing not only assessing the situation within target groups, but also gaining insight into their perception and dynamics. To obtain such qualitative data in a structured and participatory way, quantitative studies may use supportive qualitative micro-narratives as a data collection method. UNDP has been using this method to generate, analyse and disseminate evidence on income, and human development outcomes and living conditions for a number of years.⁵⁹ Micronarratives capture important events, subtle changes in attitude behaviours and perceptions. They are a method which provides people a voice, helps understand their attitudes, and contributes to sharing knowledge and making informed decisions. While at present, statistical agencies clearly focus on consolidating their data collection infrastructures, they should be alert about the increasing relevance of qualitative data, including cognitive

⁵⁹ UNDP case studies presents stories told by the Roma are accessible under <http://www.eurasia.undp.org/content/rbec/en/home/library/roma/roma-returnees-to-the-western-balkans.html>; <https://www.eurasia.undp.org/content/rbec/en/home/library/roma/the-position-of-roma-women-and-men-in-the-labour-markets-of-west.html>

interpretations of commonly asked survey questions. It is important to understand that insightful qualitative information does not necessarily imply a specific size or selection of samples.

272. In practice, when micronarratives are used, respondents' memory is triggered to situate them in an experience they have had, seen, or heard about from someone else and that relates to the field of interest. In a second step, respondents are encouraged to share a narrative associated with that experience. The story may be about themselves or someone else they know.

273. After sharing their story, respondents are asked to code (add meaning, self-signify) their narratives by responding to several predefined closed-end questions. Story tellers may be asked to locate aspects of their story within a triangle where each corner represents alternative possible interpretations.

274. In this way categorisation of meaning is made possible without strong interference from interviewers. Storytellers are then asked to respond to several multiple-choice questions, including questions about themselves (demographics) and about their narrative. For data visualisation, demographics and other multiple-choice questions serve as tags and enable colour coding and filtering of data (e.g. by gender, age, or role of other parties in the stories shared) allowing deeper analysis and better understanding of issues and meaning for specific subsets of population.

Box 3.9

UNDP's collection of micronarratives from Roma in the Western Balkans

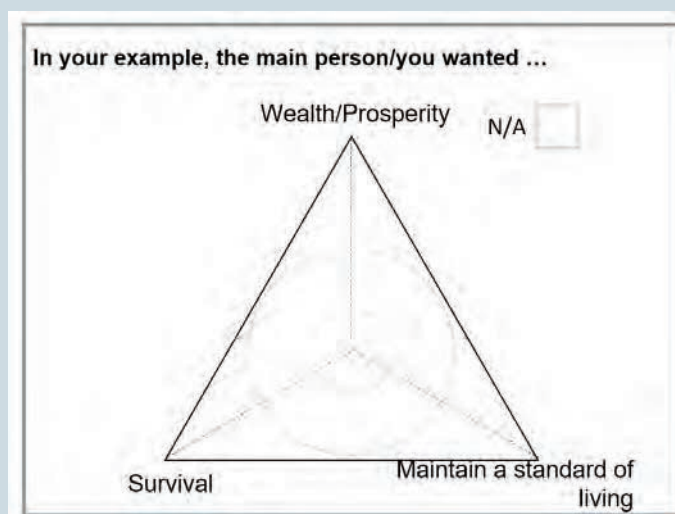
Micronarratives were identified as the most appropriate method for qualitative data collection, as they provide authentic and detailed information from Roma themselves, and also engage them in the process of analysis and intervention design. A team consisting of Roma activists and UNDP Roma focal points was formed to design a micronarratives framework and Roma activists collected a total of 1,175 stories in localities with more compact Roma populations and with already recorded returnees in the Western Balkans. The work started in June 2018 and fieldwork for the first survey was carried out during October and November 2018.

The micronarrative was based on this open-ended question to prompt the story:

"Give us an example (good or bad) of what life is like for you (or someone you know) after coming back from another country. What happened? Tell us about what happened. Say as much or as little as you wish."

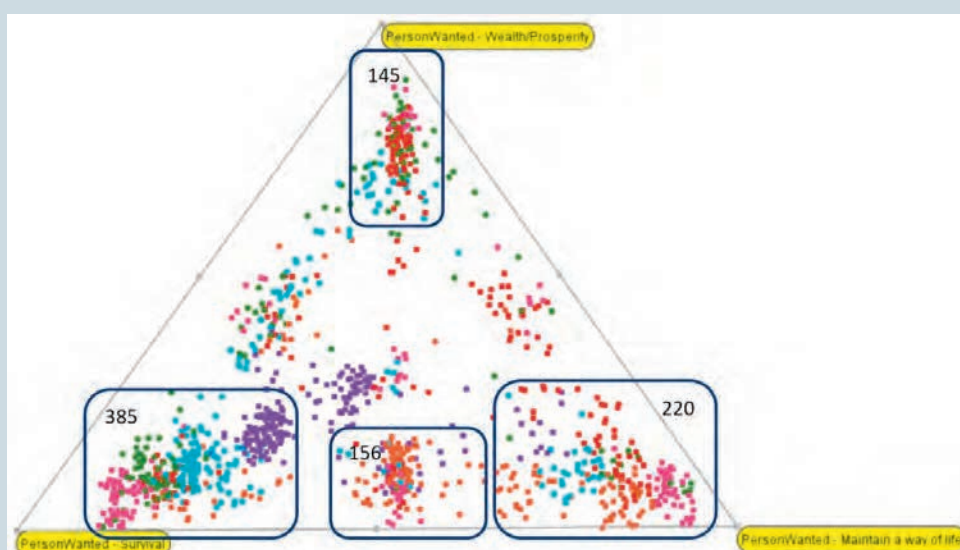
A series of questions about the short story told followed based on developed micronarrative framework and short survey questionnaire. The micronarrative framework included several "triad" questions of the following format so the storyteller could place himself/herself where his/her story fits without external evaluation/coding by a researcher.

Figure 3.1
Micronarrative framework



The self-coding of his/her story allows external researchers to visually observe how stories are getting clustered. For example, when Roma story tellers were asked to tell what had driven the decision of the person described to leave their country. The main reasons in deciding to go abroad were mere survival, maintaining a current standard of life, or a combination of both. Survival is the main theme in a cluster of 385 stories, followed by a cluster of 220 stories about maintaining a standard of living. Another cluster of 156 stories was about the combination of both, which indicates that a number of Roma are balancing between the thin line of survival and maintaining the standard of living.

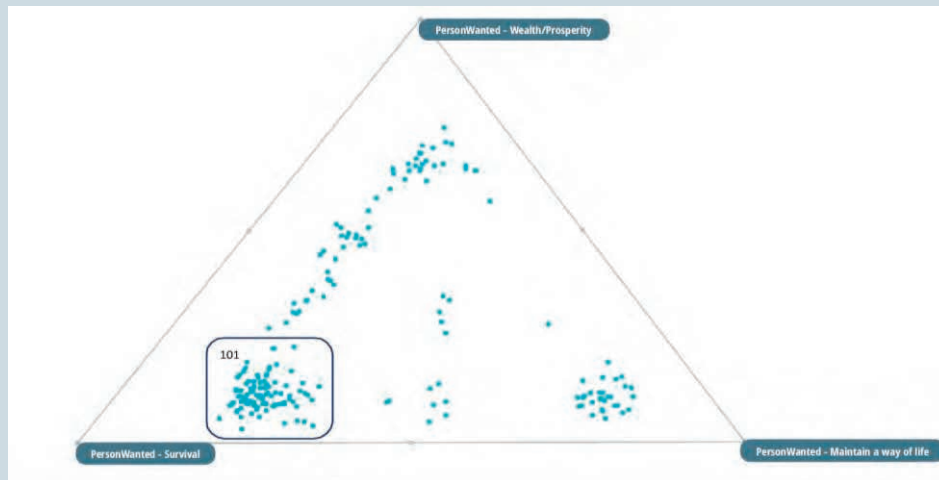
Figure 3.2
Self-coded reasons to go abroad in Western Balkans (n=1136 micro-narratives)



Broken down on country-level (i.e. for Bosnia and Herzegovina), the data shows that half of Roma returnees chose to go abroad due to reasons of survival. The fact that half of Roma storytellers in Bosnia and Herzegovina chose survival as a main reason behind the decision to leave the country indicates higher socio-economic insecurity compared to Roma from other Western Balkan countries.

Figure 3.3

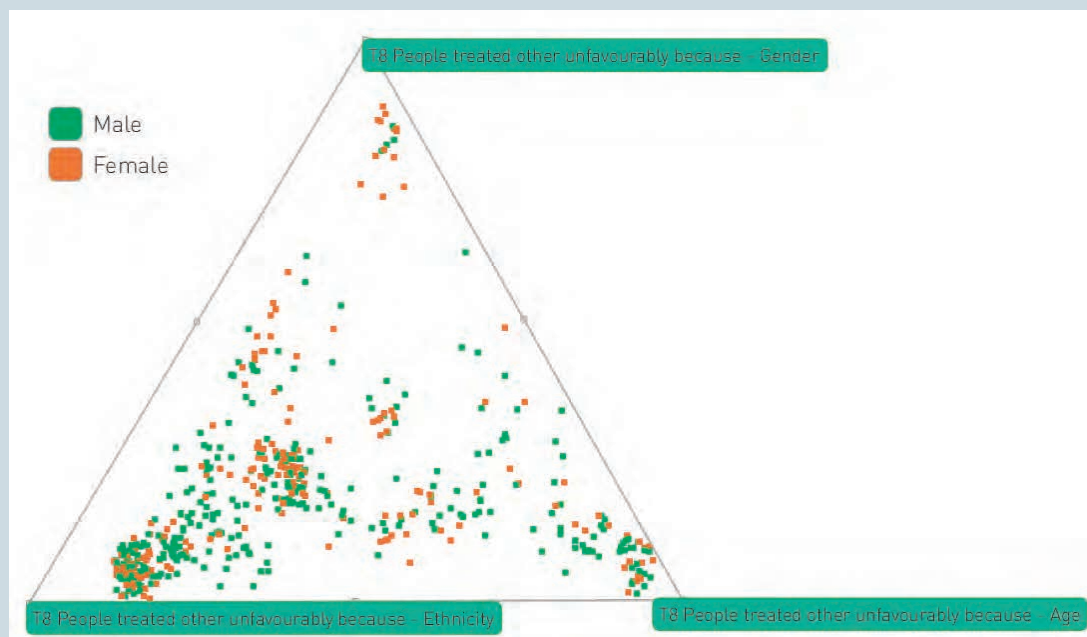
Self-coded reasons to go abroad in Bosnia and Herzegovina (n=203 micro-narratives)



In a different UNDP micronarratives qualitative study on Roma access to employment and experience of discrimination in labour markets, Roma activists collected a total of 1,000 stories in localities with more compact Roma populations. In the question that explored the perceptions of storytellers on the reasons why Roma are treated unfavourably, only half of respondents (510 people out of 1,000) provided answers, indicating perhaps a reluctance to talk about, and the sensitivity surrounding, perceptions of discrimination. Therefore, respondent's choice not to self-code may also inform analysis.

Figure 3.4

Pattern of responses for unfavourable treatment based on gender, ethnicity and age, n=510



A qualitative study was also conducted by the World Bank, in collaboration with the Institute of Ethnography in Serbia, to help understanding the underlying mechanisms behind gaps between Roma

and non-Roma in education and labour markets, with special attention to gender. The aim of the qualitative research was to further unpack the gender gap findings of the survey among the marginalized Roma and neighbouring non-Roma in Serbia and to assess the normative enablers and barriers faced by marginalized Roma and neighbouring non-Roma men and women across three broad outcome themes: human and physical capital, access to livelihoods, and voice. This qualitative research was conducted through direct interviews and focus groups in five Roma settlements (two of which were covered in the pilot research).⁶⁰

3.5 Poverty among children

275. Data on child poverty is essential in fulfilling the United Nations Convention on the Rights of the Child (CRC). Surveys which measure poverty will usually include children as members of private households and collect information through adults. The measurement of child poverty can benefit from dedicated surveys on target groups as there is a severe lack of data on children in vulnerable situations, such as children with ethnic and migrant background. There is rarely data on children with disabilities in many areas, including poverty (FRA, 2018a). Research suggests however that children with disabilities are more likely to live in poverty (Frazer and Marlier, 2017).

276. For example, the second FRA survey on minority discrimination (EU-MIDIS II) revealed that more than 90% of Roma children were found to live at risk of poverty in the nine EU Member States in which the FRA survey including Roma children was carried out. Measuring poverty among minority groups might also require including survey questions specific to certain populations. For example, between 16% (Spain) and 48% (Greece) of Roma children were found to live in households where at least one person went to bed hungry at least once in the month before the data were collected (FRA, 2018a).

277. However, even data collections which target especially vulnerable groups are at risk of missing important aspects of child poverty which will require yet more specialised data collection.

3.5.1 Street children

278. UNICEF (with partners) has some experience of sampling children in street situations in the region, using time-location mapping and capture-recapture techniques to address the elusive and mobile nature of these populations. Examples include Albania's National Study of Children in Street Situations (Di Giglio et al, 2014), and Georgia's Don't Call Me a Street Child (Wargan and Dershem, 2009). Time-location mapping is used to identify locations and time periods with the highest numbers of children on the street in major cities, followed by point-count estimates over a number of days to get a more accurate picture of the numbers of children, their sex and age, to create an estimate of the street child population from which a sample size can be determined. Children are identified for an interview through convenience or

⁶⁰ More details of the qualitative study can be found here:

<https://www.worldbank.org/en/region/eca/publication/breaking-cycle-of-roma-exclusion-in-western-balkans>

snowball sampling methods. These surveys make it possible to obtain quantitative information on a range of living standards and wellbeing related issues.

3.5.2 Children with disabilities⁶¹

279. Within the region, many children with disabilities live in institutions and are not reached by household surveys. However, even those children with disabilities who live with their parents may not be captured in surveys due to stigma, or child disability may not be fully recognized and therefore not captured. It is important to have data on children with disabilities to assist the development of policies and interventions to address the difficulties they face in managing their environment.

280. While data on children with disabilities has been collected in UNICEF's Multiple Indicator Cluster Surveys (MICS) surveys since 2000, the measurement of child disability has recently been improved in line with the World Health Organization's International Classification of Functioning, Disability and Health (ICF). The Multiple Indicator Cluster Surveys, round 6 (MICS6) has the option to include a Washington Group/UNICEF module on child functioning, that covers children between 2 and 17 years of age and assesses functional difficulties in different domains including hearing, vision, communication/comprehension, learning, mobility and emotions. The inclusion of children under 2 is not recommended as the early detection of child disability is highly subjective and the nature of the child development process may lead to a high number of false positives.

281. Under the new module, which has been subjected to rigorous testing and validation, each area of functional difficulty is assessed against a rating scale, making it easier to identify the subpopulation of children who are at greater risk than other children of the same age. Including the module within a larger survey may also mean it is possible to assess multidimensional poverty for this group of children. Most of the estimated eighty MICS6 surveys that will be conducted across the world will include the child functioning module. This will allow understanding the differences and extent of multidimensional poverty between children with functioning difficulties and the rest.

3.5.3 Foster children

282. A first step towards measuring the special conditions of foster children is to ask the household if they are fostering a child. This would potentially allow the disaggregation of poverty numbers by foster children. Without individual poverty measures (see Chapter 5) potential within household inequalities will however remain hidden.

283. The United States Census Bureau's Supplemental Poverty Measure (SPM) includes children in foster care. The Current Population Survey Annual Social and Economic Supplement (CPS ASEC) used for these estimates includes "foster child" as one of the relationship categories. Estimates show that holding

⁶¹ For more information see: UNICEF, 2018.

child and family demographics constant, foster children have a lower risk of poverty than other children (Pac et al, 2017).

284. The Armenia Integrated Living Conditions Survey (ILCS), run each year by the Statistical Service of the Republic of Armenia, is the base for many official statistics including poverty headcounts. In 2013 and 2014, the ILCS included a child-needs module, modelled on a similar one included in EU-SILC in 2009 and 2014 but adapted to national context. Among the child-needs questions, it included questions to adults in the household about the presence of foster children or the absence of children who live in institutions. It included also questions on their opinion on fostering children or orphanages for children with disabilities or from vulnerable families (National Statistical Service of the Republic of Armenia, 2014).

3.6 Homelessness and housing difficulties in the European Union

285. This section illustrates how homelessness is addressed through a special module in EU-SILC. Primary target variables for EU-SILC are collected every year and cover the following domains:

- At household level: basic/core data, income, housing, material deprivation.
- At personal level: basic/demographic data, income, education, labour and health.

286. These variables are complemented by secondary variables, which are collected every five years, or less frequently in the so-called ad-hoc modules. Ad-hoc modules have been included each year since 2005 in order to complement the variables permanently collected in EU-SILC with supplementary variables highlighting unexplored aspects of social exclusion. The modules implemented between 2005 and 2016, and for which data have already been disseminated, covered the following topics: inter-generational transmission of poverty and of disadvantages, over-indebtedness and financial exclusion, housing conditions, material deprivation, intra-household sharing of resources, social and cultural participation and well-being as well as access to services⁶².

287. As part of the modernisation of EU social statistics, it has been agreed that modules will be collected on a regular basis every 3 years or 6 years with the aim to monitor changes over time and also to complement the information collected in the SILC annually. One of the topics which is planned to be collected as part of every 6 years in revised EU-SILC is housing difficulties. This module aims to collect information on past experiences of homelessness.

⁶² <http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/ad-hoc-modules>

3.6.1 Policy justification

288. Adequate, affordable housing is crucial to the enjoyment of fundamental rights, health, well-being and social inclusion. Housing exclusion and homelessness are important themes for the EU's efforts to fight poverty and promote social inclusion.

289. National statistics, albeit non-harmonized across countries, suggest that homelessness is increasing in a majority of Member States (FAP and FEANTSA, 2015). While, designing and implementing measures to reduce homelessness is a (sub) national competence, the European Commission provides Member States with support through policy guidance and through EU funding. Key steps in the emergence of homelessness as a priority on the EU's social agenda include:

- 2008 French Presidency of the EU;
- 2009 Joint Report on Social Protection and Social Inclusion of the Commission and Council;
- European Consensus Conference on Homelessness organized by the Belgian Presidency of the Council in 2010;
- Calls for a European Homelessness Strategy by various EU institutions and bodies, notably the Committee of the Regions, the European Economic and Social Committee, the Council (EPSCO), the European Parliament;
- Publication of the Social Investment Package in 2013, including specific policy guidance on homelessness;
- Inter-ministerial Round Table on Homelessness organized by the Irish Presidency in 2013; and
- Housing exclusion identified as a social trend to watch in the 2015 European Semester.

290. Collecting data on homelessness at EU level has so far been very difficult. However, there has been criticism that the EU's social statistics do not cover more extreme forms of poverty well. Various attempts have been made to improve the data available on homelessness at EU level. In 2007, the European Commission funded a study on measuring homelessness (European Commission, 2007). There was an attempt in the 2011 census to gather comparable data on homelessness. This exercise delivered useful information for some of the Member States, however due to a lack of harmonisation in the definition and methodology, the data available is not comparable across countries.

291. EU-SILC as a survey on private households cannot collect information on persons experiencing homelessness currently. However, the ongoing revision of the EU-SILC legal basis provides an opportunity to monitor past homelessness on a regular basis. For this purpose, a Task Force was set up for preparing variables on homelessness and related guidelines to be implemented as a test in the EU-SILC ad hoc module 2018. In addition to Member States volunteering to participate, FEANTSA (the European Federation of National Organisations Working with the Homeless) provided support to Eurostat and the work of the Task Force.

3.6.2 How to define homelessness

292. The issue of how to define homelessness has been one of the main concerns when developing new variables measuring homelessness in EU SILC. Homelessness is a complex phenomenon and there is no universally accepted definition.

293. FEANTSA developed the European Typology of Homelessness and Housing Exclusion (ETHOS) to aid which consists of four conceptual categories: rooflessness; houselessness; insecure housing and inadequate housing. ETHOS has become a widely accepted frame of reference. For instance, in 2007, a study on measuring homelessness funded by the European Commission used ETHOS as a starting point to propose a harmonized definition of homelessness for data collection purposes. This was a “light” version of the ETHOS, which was simplified for data collection purposes. ETHOS was used also as a starting point when defining homelessness in EU-SILC.

294. The major challenge for EU SILC was to agree on which categories of ETHOS to use in questions for the ad-hoc module. On the one hand, using a broad definition could allow the capture of “hidden” forms of homelessness. On the other hand, there were concerns that when too many situations are considered as homelessness, the usefulness of the data for policy purposes could be reduced. For instance, the inclusion of ETHOS 8.1 (staying temporarily with family/friends) because easily apply to large parts of the population e.g. young people who are studying or saving to buy a first home.

295. Taking into account the above considerations it was agreed that a person should be considered as having past housing difficulties if had no place of their own (either owned or rented) where they could live and therefore were forced to stay with friends/family, stay in an emergency or other temporary accommodation, a place not intended as a permanent home or had to sleep in a public space. Moreover, Eurostat clarified in the guidelines that if respondent was forced to leave their home temporarily due to unforeseen event such as: risk of earthquake; fire in the neighbourhood; evacuation of the neighbourhood because of bomb defuse etc. but in general had place of their own (either owned or rented) to live in then the respondent should not be considered to have had housing difficulties.

296. Furthermore, the category “staying with friends or relatives” – should include only situations when somebody was forced to move (back) to family or friends as they did not have any other place to stay. Visits, staying for limited time (e.g. during refurbishing of own flat) or living with family in order to save money rather than due to an absolute need should not be taken into consideration.

297. Those clarifications in the guidelines enabled a broad definition of homelessness, allowing countries to capture also “hidden” forms of homelessness, while at the same time not include situations which could be considered as normal at certain stages of life.

3.6.3 Sample Issues

298. One of the most important issues when working on the topic of homelessness in EU-SILC was the concern of the occurrence of how many cases of past housing difficulties would be recorded taking into

account the relatively small sample size and the perception that homelessness is a relatively “rare” social issue.

299. Eurostat together with the countries and FEANSTA analysed experiences of other institutions and countries in collecting information on experiences of homelessness (Annex 3). In particular, the following retrospective modules on Homelessness were analysed:

- French Health Survey designed by INSEE, 2013;
- The Survey of English Housing, 1994/95;
- The Scottish Household Survey, 2012;
- Toro et al Survey, 2007;
- Australian General Social Survey, 2010.

300. Based on the experiences from other countries, it could be seen that homelessness in some form was experienced by between 4.5% and 13.5% of the population. The results differed greatly depending on the reference period as well as the definition of homelessness used.

301. There were two strategies Eurostat used to increase the chances of capturing experience of homelessness within EU SILC. The first one was to maximize the number of living situations/ETHOS categories included in the definition. However, a balance needed to be found between larger coverage for sample size reasons and policy relevance as well as its perception by the interviewees (see the discussion on the definition above). The second strategy was to maximize the reference period. During discussions, FEASTA recommended having a reference period of 5 or 6 years. This, however, was viewed as too short and after consultations with Member States it was agreed to extend the reference period to “lifetime”. The downside of using 'lifetime' as a reference period was that there will be no information on when the housing difficulties took place and when the respondent overcame the problem. However, “lifetime” was also used by other surveys and using this reference period allows collecting information from a greater number of respondents.

302. Overall, Eurostat together with Member States agreed that the only way to establish the feasibility of addressing homelessness in the EU SILC sample was to test it as part of an ad hoc module which was implemented in 2018.

3.6.4 Implementation in EU-SILC

303. In developing the ad-hoc module 2018, Eurostat cooperated with FEANSTA to determine what information was most crucial. Selected variables were pre-tested with the use of focus groups and cognitive interviews in English, French and Polish, leading to further refinement. In addition, Eurostat consulted Member States at multiple stages.

304. The feedback was crucial to assist in finalising the list of variables to be collected. It recommended that a filter question asking respondents whether they had any experience of homelessness during their lifetime would be useful to limit the burden for the majority of respondents. As stated earlier, guidelines

for this variable were written in such a way as to preclude people who were not forced to live outside their house (either owned or rented). Following this, it was proposed to have a variable which would measure the duration of housing difficulties.

305. Member States recognised that it is possible for respondents to have experienced difficulties more than once during their lifetime. After discussion, a decision was reached to focus more on the duration of the most recent spell, as opposed to concentrating questions on the longest spell of difficulty the respondent faced. In specific cases when a respondent, during a continuous experience of housing difficulty, changed their place of stay (e.g. from emergency accommodation to a place not intended as a permanent home and then following this was 'sleeping rough') it was agreed that the total duration should be reported.

306. Housing difficulties are a complex, multidimensional issue and often a consequence of multiple events. From a policy perspective, it is crucial to have information regarding which are the main factors which lead to housing difficulties. After discussions with Member States, as well as taking into considerations outcomes of the pre-testing, it was decided to have two variables collecting information on reasons for past housing difficulties.

307. The final selected variable aimed to obtain information on how the respondent resolved his difficulties. Of course, there are likely to be many reasons and causes for resolving housing difficulties. For some, it could be the first event which was a stepping stone leading to escaping housing difficulties, while for others it could be the most recent event. Nonetheless, the respondent is asked to select the most important thing in their opinion. This variable also records whether the respondent is still experiencing housing difficulties, and is therefore currently staying with family or friends (in a household covered by the survey) due to a lack of other housing.

308. The agreed five variables on past housing difficulties are reported in Table 3.3.⁶³ Fifteen countries⁶⁴ implemented the optional housing difficulties variables in the 2018 EU-SILC ad hoc module.

Table 3.3

EU-SILC variables on past episodes of homelessness and housing difficulties

PHD01T Past experience of housing difficulties Yes, staying with friends or relatives temporarily Yes, staying in emergency or other temporary accommodation Yes, staying in a place not intended as a permanent home Yes, 'sleeping rough' or sleeping in a public space No
PHD02T Duration of the most recent experience of housing difficulties Duration

⁶³ Although optional those variables were included in regulation Commission Regulation (EU) 2017/310 of 22 February 2017 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards the list of target secondary variables on material deprivation, well-being and housing difficulties for 2018 (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2017.045.01.0001.01.ENG&toc=OJ:L:2017:045:TOC)

⁶⁴ Albania, Belgium, Bulgaria, Switzerland, Germany, Denmark, Greece, Spain, Hungary, Ireland, Malta, Portugal, Romania, Slovakia and United Kingdom.

PHD03T Main reason for past housing difficulties Relationship or family problems Health problems Unemployment End of rental contract Uninhabitable accommodation Leaving an institution after a long stay and no home to go to Financial problems/ Insufficient income Other
PHD04T Other reason for past housing difficulties Relationship or family problems Health problems Unemployment End of rental contract Uninhabitable accommodation Leaving an institution after a long stay and no home to go to Financial problems/ Insufficient income Other No other reason
PHD05T Exit from housing difficulties Existing, new or renewed relationship with family or partner Addressed health problems Gained employment Moved into social or subsidised private housing Other Still experiencing housing difficulties

3.6.5 Plans for the revised EU-SILC

309. In the revised EU-SILC, a few variables on past housing difficulties are planned to be collected every 6 years. Based on the results of the test of the module in 2018, available in 2019, a decision will be made on the final list of variables to be included in the revised EU-SILC.

3.6.6 Conclusion

310. There is a strong need for comparable information on persons experiencing homelessness. As a household survey, EU SILC is not the appropriate instrument for sampling people who are currently homeless. However, it can provide valuable insights into specific dimensions of past homelessness and housing difficulties. In 2018, information on past housing difficulties was collected for the first time in EU-SILC. The information was collected on a voluntary basis at this testing stage.

311. In the revised EU-SILC, after improving the variables and guidelines based on the outcomes of the 2018 test, information on past housing difficulties is planned to be collected every 6 years. This will

provide information for policy makers on type of housing difficulties, duration, main reasons for housing difficulties and what allowed respondents to exit housing difficulties.

3.7 Surveying Institutional Households

312. The UNECE Guide on Poverty Measurement recommends using the definition of a private household as used in the Conference of European Statisticians (CES) Recommendations for the 2010 Censuses of Population and Housing (UNECE, 2006) and the UN Canberra Handbook on Household Income Statistics (UNECE, 2007: 33):

Either (a) a person living alone in a separate housing unit or who occupies, as a lodger, a separate room (or rooms) of a housing unit but does not join with any of the other occupants of the housing unit to form part of a multi-person household or (b) a group of two or more persons who combine to occupy the whole or part of a housing unit and to provide themselves with food and possibly other essentials for living. The group may be composed of related persons only or of unrelated persons or of a combination of both. The group may also pool their income.

313. The recommendation further defines household membership based on the place of usual residence and specifies rules for temporary absences.

314. It also recognises that such definition does not sufficiently cover all populations of interest and poses practical challenges when applied to some hard-to reach populations. For example, grouped housing for refugees or seasonal worker's homes might neither meet the definition of a private nor institutional household. For travellers who are mobile and live in caravans, the differentiation between family and household needs further specification, likewise the housing unit. A household can encompass several caravans, who share expenses and meals and other essentials of their living.

315. Definitions of what establishes a household need to be adapted for specific groups. For example, FRA used "shared expenses" as a criterion to define a household among persons in grouped and/or precarious housing. The survey included migrants or refugees with irregular or undocumented status. Respondents sometimes share a room without having any relationship, paying rent per bed. In this case each person was treated as a separate household unit. The UNDP-WB Regional Roma Surveys of 2011 and 2017 included students, military and recent migrants.

316. For mobile residents (e.g. travellers) between countries FRA specified usual residence as being more than six months in a country in the last 12 months and the current residence can be a first or second home (e.g. a caravan in summer and a brick and mortar house in winter) at the time of the interview. Similar to location sampling, questions on other residential locations can be used to adjust for overlaps in the sampling frame.

Box 3.10

Measuring poverty for the institutionalized population in the United States

The United States Census Bureau collects data on persons in institutions through the American Community Survey (ACS). In addition to including households in the survey, the ACS collects data for persons in “group quarters” (GQ). Group quarters include correctional facilities, nursing homes, mental hospitals, college dormitories, military barracks, group homes, missions and shelters. Poverty status is determined for residents of a small subset of these group quarters: group homes, missions and shelters. Poverty status is not determined for residents of correctional facilities, nursing homes, mental hospitals, college dormitories or military barracks.

Group quarters facilities are randomly selected each year from a sample list of all group quarters in an area. From a list provided by the GQ contact person of all residents currently staying at the GQ, field representatives randomly select residents to take part in this survey. The interviewer collects data from the selected (sampled) residents of the selected GQ, usually through a computer-assisted personal interview (CAPI) with the resident but sometimes through a paper questionnaire completed by the resident. When collecting data from prison/correctional institutions or when the sampled GQ individual is not capable to complete the questionnaire, the GQ official will complete the questionnaire using administrative records.

The ACS questionnaire used for private households includes two sections – one at the household level (e.g. housing tenure, rental or homeownership costs, housing characteristics) and one for person-level information (e.g. age, marital status, educational attainment). The GQ questionnaire, on the other hand, only includes only the person-level questions, and includes a series of questions at the end of the instrument for use by the interviewer. These questions ascertain whether the answers were given by the sample resident, a proxy respondent, or the interviewer and whether administrative data was used in the response.

All survey respondents 15 years and older, including those in GQ’s, are asked a series of questions about their income over the previous 12 months. These responses are used to calculate total personal income. For individuals in group quarters for whom poverty status is determined, this total personal income is compared to the appropriate poverty threshold to determine poverty status. These thresholds vary by the age of the respondent (lower thresholds are used for respondents ages 65 and older) and by the month of the interview. Since relationship data is not collected for individuals in group quarters, poverty status for individuals in group quarters is based solely on their personal income. Poverty status is not determined for survey respondents in group quarters under 15 years of age.

317. In most countries large scale household surveys providing statistics on poverty and social exclusion cover only persons in private households. This may underestimate the total population at risk of poverty in a country and leave out the very specific vulnerability of persons living in institutional households. Examples of institutional households or collective households are: hospitals, institutions or residential care settings for older people or persons with disabilities, prisons, military barracks, shelters for refugees or homeless persons, boarding houses and workers’ hostels. The proportion of individuals living in these types of accommodation varies between countries significantly and is estimated to be increasing in many countries due to aging populations.⁶⁵ Twenty out of 30 European countries reported a share of collective households higher than 1% of the entire population. Only in seven countries (Bulgaria,

⁶⁵ For definition see the OECD glossary: <https://stats.oecd.org/glossary/detail.asp?ID=1372>

Cyprus, Finland, Italy, Norway, Poland, and Spain) did collective households account for a proportion of less than 0.6% of the entire population (SERISS, 2017). In Germany, 19.4% of persons 65 years or older were living in institutions in 2010, which is expected to increase to 28% by 2035 (Rodrigues et al, 2012). Some national surveys in the EHIS included people living in institutions, and SHARE as a longitudinal survey follows persons even if they move into institutionalised care (FRA, 2018b).

318. The SERISS network⁶⁶ (Synergies for Europe's Research Infrastructures in the Social Sciences) has set up an inventory on surveys covering institutional households. So far 167 European surveys are registered in the survey inventory (+42 surveys from the USA, Canada, Australia, and Israel); 80 of those surveys included at least parts of the institutionalized population at least once (+27 surveys from the USA, Canada, Australia, and Israel).⁶⁷

319. While persons with disabilities in the EU are at a higher risk of poverty (European Commission, 2014), it is only those living in private households that are covered in headline statistics. Around 1.2 million persons with disabilities live in institutions in Europe.⁶⁸ Leaving out their experiences is a significant weakness of the current data. Moreover, the UN Convention on the Rights of Persons with Disabilities (CRPD) is the first UN Human Rights Treaty with a specific article on statistical and data collection. Article 31 obliges State Parties to collect appropriate information, including data on persons with disabilities who live in institutions, to formulate and implement policies to give effect to the Convention.

Box 3.11

Data gaps on persons with disabilities in the European Union

Between 2014 and 2018, the European Union Agency for Fundamental Rights (FRA) implemented a multiannual project to collect and analyse comparable data on the right to independent living of persons with disabilities in the EU. The overall objective was to provide evidence-based assistance and expertise to EU institutions and Member States on how to fulfil the right of persons with disabilities to live independently and be included in the community, as set out in Article 19 of the Convention on the Rights of Persons with Disabilities (CRPD).

The project specifically focused on the local level, giving voice to all main stakeholders, most importantly to people with disabilities themselves, and examined what drivers forward and what holds back the process of deinstitutionalisation across Europe. The results show lack of robust, comparable and timely data on independent living outcomes for persons with disabilities. This impedes evidence-based policymaking and prevents Member States from demonstrating meaningful progress in achieving independent living. The Agency calls to Eurostat and national statistical offices to develop inclusive methodologies for collecting statistical data on the right to independent living. These methodologies

⁶⁶ <https://seriss.eu/about-seriss/project-overview/>

⁶⁷ https://seriss.eu/wp-content/uploads/2017/02/SERISS_WP2_D2.16_Survey-inventory-institutionalized-population.xlsx

⁶⁸ These estimates include Turkey, see https://fra.europa.eu/sites/default/files/fra_uploads/fra-2017-independent-living-part-iii-outcomes_en.pdf.

should facilitate the participation of all persons with disabilities, including those with severe impairments and those living in institutions.⁶⁹

3.7.1 Case Study: United Kingdom - including non-private household populations (NPHP)

320. Following a number of developments which raised concerns about destitution and severe poverty in the United Kingdom, in 2017, the Office for National Statistics and the Joseph Rowntree Foundation, jointly commissioned scoping work to investigate the alternatives for including non-private household populations (NPHP) in estimates of destitution and personal well-being.⁷⁰ The original scope of the project was extended to encompass the wider concepts of living standards and quality of life, to capture not only the extreme cases but also the full range of people's experiences.

321. Currently, United Kingdom estimates of living standards and quality of life derive mainly from household surveys. Unlike many other countries, the United Kingdom does not have a comprehensive population register so household surveys use address-based sampling using products such as the Royal Mail's Postcode Address File. These provide a list of addresses covering most of the population.

322. While these address lists are likely to include some communal establishments, they are not routinely sampled as part of household surveys. This is in part because it is not yet possible to routinely distinguish communal accommodation from private household accommodation and there are practical issues with sampling communal establishments, including the number of residents they contain, the complexity of sampling these residents, issues with accessing some of these types of property and the need for gatekeeper assistance or external support to enable sampling.

323. In addition to the exclusion of communal accommodation from household surveys, those who are temporarily living at an address but are not considered to be a permanent member of the household are also excluded from the majority of surveys. Furthermore, the use of address-based sampling also omits those who do not have a clear or fixed residence.

324. The address-based sampling used in United Kingdom household surveys provides good coverage of the majority of the population, with those not living in a private household estimated to make up less than 2% of the total population. If the characteristics of this part of the population were similar to the wider household population, there would be no issue with omitting them from household surveys. However, there is good reason to expect that the living standards and quality of life of those not living in private households are likely to be very different from the private household population.

⁶⁹ FRA reports on Article 19, CRPD at: <https://fra.europa.eu/en/project/2014/rights-persons-disabilities-right-independent-living/publications>. See, in particular, From institutions to community living - Part III: outcomes for persons with disabilities.

⁷⁰ This work was carried out by Glen Bramley, Filip Sosenko and Jenny Wood from the Institute for Social Policy, Housing and Equalities Research (I-SPHERE) at Heriot-Watt University in association with Joel Williams and Peter Matthews of Kantar Public. The report of this scoping work is available at <https://pureapps2.hw.ac.uk/ws/portalfiles/portal/23481268/NonHouseholdReport.pdf>.

3.7.1.1 Defining the populations of interest

325. The scoping work identified a number of sub-populations that are not routinely captured in household surveys, though it is important to note that these categories are not always distinct from one another or indeed from the wider private household population:

1. **Care home residents:** This includes those living in residential homes and residents of nursing homes. Most of these residents are elderly and many are in poor health.
2. **Long-stay hospital patients:** This category includes those whose stay in hospital is sufficiently long to assume that they would not be captured in a survey at their main place of residence. While, there is no clear threshold for defining this, the scoping work estimate (see below) is based on a stay of more than 4 weeks in hospital.
3. **Military accommodation residents:** this category covers armed forces personnel and their families who live in accommodation provided by the military.
4. **People detained in Immigration Removal Centres:** in most cases, detainees are held for short periods of time, though there are cases where detainees may be kept much longer.
5. **Students living in halls of residence:** although the majority of these establishments are included in the Postcode Address File, they are not routinely sampled, though whether students are sampled at their home address varies for different surveys and at different times of year.
6. **Travellers living in caravans:** the term ‘traveller’ encompasses multiple different communities, including those who adopt a traveller lifestyle and those who identify with a specific ethnic or linguistic group. Only those living in caravans are relevant in the context of the NPHP.
7. **Prisoners:** this category includes large numbers of people with complex needs, particularly mental health issues but whether they are included in the private household population will depend on the length of their sentence.
8. **Those living in hostels/shelters for the homeless:** this category encompasses several different types of accommodation and it can be difficult to define.
9. **Those living in bed and breakfasts, hotels, and other accommodation designed for holidaying:** this category overlaps with those living in hostels or shelters for the homeless and can also be difficult to separate from those using these establishments for tourism.
10. **Homeless people:** this covers people in the previous two categories as well as rough sleepers, those sleeping in cars, tents, public transport, etc., unlicensed squatting or occupation of non-residential buildings and those staying with non-family on a short-term basis in overcrowded conditions, referred to as “sofa surfers”.

Table 3.4

Estimated size of each including non-private household sub-populations in the United Kingdom

Category of establishment	Estimated number
Care home (including Children’s homes)	430,000
Hospital	80,000
Military establishments	115,000
Probation, bail, detention, etc.	3,000
Educational	500,000

Caravan sites	90-120,000
Prisons	92,300
Hostels/shelters for the homeless	46,000
Hotels, guest houses, bed and breakfasts, etc.	20,000

326. Table 3.4 shows the estimated size of each of these sub-populations derived by the authors as part of the scoping work. This highlights that a small number of institutions dominate the wider group in terms of numbers, specifically care homes, educational establishments and military accommodation, though caravan sites may also make up a large proportion of the NPHP.

327. There are currently no reliable estimates of the number of homeless people in the United Kingdom, though the scoping work estimated that there could be up to 200,000 people who could be categorized as homeless in Great Britain alone.

328. Although, collectively referred to as the NPHP, it is important to recognize the highly diverse nature of the different sub-populations both in terms of the expected prevalence of poor living standards and quality of life and in relation to how to collect data for each.

3.7.1.2 Potential ways forward

329. For each of the categories of NPHP identified, there are existing relevant surveys providing at least some information, though the level of detail captured varies for the different sub-populations. None fully capture the information needed to estimate living standards and quality of life in a similar way to that for the private household population and with United Kingdom-wide coverage.

330. Overall, the work identified 3 main ways in which data for the NPHP could be collected:

1. Piggybacking on existing surveys of residents – this would involve adding new questions to existing surveys.
2. New accommodation-based surveys – this would involve developing a new survey based on sampling within different establishments.
3. New service-based surveys – this would involve surveying individuals using particular services.

331. Because of the diversity of the groups being considered, the scoping work proposed different approaches for the different sub-populations, which are summarised in Table 3.5.

Table 3.5

Proposed approaches to capture data for the different categories of including non-private household populations

Category	Preferred approach
Care home residents	Two stage sample: 1. Sampling from lists of care homes 2. Sampling of residents within the selected homes. Computer Assisted Personal Interview (CAPI)
Long-stay medical patients	Analyse existing data on length of hospital day to determine threshold for defining category.

	Two stage sample: 1. Sampling from lists of hospitals 2. Sampling of long-stay patients within selected hospitals. Paper self-completion questionnaire with interviewer assistance where needed
Those in military accommodation	Include relevant questions within existing surveys of armed forces personnel.
Detainees in Immigration Removal Centres	Paper self-completion questionnaire with interviewer assistance where needed of all detainees in each of the 10 Immigration Removal Centres
Students in halls of residence	Bespoke online survey or include relevant questions in existing survey.
Travellers living in caravans	Three stage sample: 1. Sample of Local Authorities (LA) 2. Sample of sites within selected LAs 3. Sample of residents within selected sites Computer Assisted Personal Interview (CAPI)
Prisoners	Self-completion survey
Hostels/shelters for the homeless	Three stage sample: 1. Sample of Local Authorities (LA) 2. Sample of hostels within selected LAs 3. Sample of residents within selected hostels Computer Assisted Personal Interview (CAPI) over 1-2 weeks
Hotels, guest houses, bed and breakfasts, etc.	Three stage sample: 1. Sample of Local Authorities (LA) 2. Sample of establishments within selected LAs 3. Sample of residents within selected establishments Computer Assisted Personal Interview (CAPI) over 1-2 weeks
Sofa surfers	Capture within existing household surveys

332. Although the scoping work did not specifically focus on it, there is also the potential for administrative data to be used, either to provide a sampling frame for new surveys or as a way of providing additional information about the survey respondents.

3.7.1.3 Methodological considerations

333. There are several methodological issues to be considered when surveying the NPHP using the proposed approaches.

334. One of these relates to the potential for counting some individuals both within the NPHP and within a private household, people who are described as “in-betweeners”. Some of the residents in hospitals, prisons, care homes and military accommodation may be classified as usual residents of a

private household but temporarily absent from that household. As such, they may be counted during a household survey, though there will be a lack of information on their living standards and quality of life, since these are not collected by proxy. Given this and that their characteristics are more likely to resemble those of other residents of communal establishments, it is more appropriate to include them in the data collection for the NPHP, however, some adjustment may be needed to ensure that they are not double counted. Similarly, there is a seasonal element in the population of students in halls of residence so there is a risk that they may be captured in household surveys at certain times of the year.

335. Avoiding double counting may also be a consideration in the case of populations that are transitory in nature, particularly sofa surfers, detainees in immigration centres and those making use of homeless shelters and hostels. There is a risk that a service-based survey may also capture these individuals, which may again require some adjustment.

336. Other issues relate to the measurement of the concepts. Typical household survey topics or questions cannot necessarily directly be applied to the NPHP, so it is important to include questions tailored to the specific situation and context. For many of these populations, questions typically used to establish living standards for the private household population will not apply, including those related to the provision of food, accommodation, heating, material goods, etc. Similarly, consideration needs to be given to measuring the quality of life for some of these groups, particularly those in stressful or uncertain situations, such as detainees or those at risk of or experiencing homelessness.

337. Consideration must also be given to the capacity or willingness of residents to answer survey questions, for example, due to poor health, language barriers or trust issues. Some of these may be overcome by providing adequate support or making use of community representatives.

338. The ultimate aim of producing estimates for the NPHP is to enable estimates to be generated for the whole population. This raises issues of how to combine estimates from multiple sources.

3.7.1.4 Next steps

339. The scoping work identified a number of potential ways forward to estimate the living standards and quality of life of different categories of the NPHP. Further feasibility work is needed to establish which of these proposed options might be viable. The Office for National Statistics is considering the recommendations from this scoping work alongside its broader programme of work exploring the potential for non-survey data sources to be used to address existing data gaps.

3.8 Recommendations

Recommendation 3: Countries are encouraged to first identify, and then quantify, parts of the population not sufficiently covered in their poverty statistics. This assessment should especially address ethnic minorities, persons who are homeless, live in institutional households or have disabilities. They should further explore the methodological and empirical causes and challenges of why these or other populations may be underrepresented or not covered.

Recommendation 4: Countries should develop suitable methods to facilitate participation of hard-to-reach households in their survey programmes. The principle of self-identification or the use of survey instruments in different languages are simple examples for some basic measures, which aim to establish trust and improve accessibility.

Recommendation 5: If it is not feasible to include hard-to-reach groups on existing survey programmes, countries should develop targeted surveys to collect data on poverty and social exclusion specific to the groups, to be run at least every 5 years.

Recommendation 6: Censuses and large-scale surveys should include questions on self-identification of ethnicity, which allow for multiple identities. It helps to establish sampling frames which are needed to target ethnic minorities and hard to reach groups. This requirement is critical for the objective of leaving no one behind and to disaggregate survey data and poverty statistics for vulnerable groups.

Recommendation 7: Interviewer training for surveying minority groups should address cultural and group-specific aspects. Persons of the target populations should be involved in the set-up and development of fieldwork materials.

Recommendation 8: Further research should be directed at creating supplementary measures of poverty to reflect that the living circumstances of target groups do not always fit with standard household definitions.

Recommendation 9: Further research is needed to explore the possibilities for alternative sampling approaches, such as non-probability designs including online surveys, to measure poverty in hard-to-reach populations.

4 Assessing and Improving Survey Methods

4.1 Survey errors and quality

340. To establish trust in poverty measurement and prevent misguided policies, Statistical Offices have to regularly assess and continuously improve the quality of their processes and accuracy of their data. Quality reports which describe the quality criteria and explain any instances in which these criteria could not be met, or statistical concepts could not be correctly applied will not only facilitate the correct interpretation but can also provide the basis for future improvements. The World Bank's Report of the Commission on Global Poverty recommended that "The World Bank should make public the principles according to which household survey data are selected for use in the global poverty count; and there should be an assessment at national level of the availability and quality of the required household survey data..." The Commission suggested in particular investigating potential survey underrepresentation and noncoverage. It also recommended that "...poverty estimates should be based on a 'total error' approach, evaluating the possible sources and magnitude of error..." (World Bank 2017 recommendations number 6, 3 and 5 on pages 33, 50 and 59). Certain errors are especially relevant for disaggregation.

4.1.1 A typology of survey errors

341. It is common to broadly separate sampling and non-sampling errors (Assael and Keon, 1982; Kalton, 1983; Kalton et al, 1986). For the assessment of poverty measurement however, Verma et al (Verma et al, 2010) applied a modified framework (Verma, 1981; Hussmanns et al, 1990). Much of this section is drawn directly from Verma et al (Verma et al, 2010) with the permission of the authors.

342. Three broad categories of errors should be distinguished:

(a) Errors in measurement

What is measured on the statistical units enumerated in the survey can be different from the actual (true) values for those units. These errors concern the accuracy of the *substantive content of the survey*: the definition of the survey objectives and questions; the ability and willingness of the respondent to provide the information sought; and the quality of data collection, recording and processing. A typical example of error in measurement would be underreporting of certain income components. This will not only increase uncertainties but can possibly also lead to significant bias in the estimates.

(b) Errors in estimation

The process of extrapolation from individual measurements to the entire study population adds further uncertainties. These result from *sample design and implementation*, notably coverage, sample selection and implementation, and also sampling errors and estimation bias.

(c) Item non-response

For poverty measurement, Verma et al (Verma et al, 2010) especially highlighted *item non-response* as special, mixed category that complements the common distinction between representation and measurement errors (Groves et al, 2004). Item non-response is particularly important in surveys which collect detailed information on components of household and personal income. It is generated in the process of measurement but in its effect, it adds to the existing non-response and thus also amounts to an error of estimation (that may be mitigated by estimation tools such as imputation).

343. Quality reports should describe these broad categories of error in sufficient detail.

4.1.1.1 Errors in measurement

344. It is useful to distinguish conceptual, response (“data collection”) and processing errors. *Conceptual errors* concern the scope, concepts, definitions and classifications adopted in relation to the survey objectives. It is almost impossible to compensate for conceptual errors. *Response errors* concern the process of data collection while *processing errors* concern the subsequent process of transforming the information into a micro database. They result from different survey operations but their effects are similar. Each type of error may further be decomposed into bias and variance components. These distinctions are useful insofar as the components differ in nature and in methods of assessment and control.

4.1.1.1.1 Measurement bias

345. Bias arises from shortcomings which affect the whole survey operation: basic conceptual errors in defining and implementing the survey content; incorrect instructions for interviewers; errors in the coding frame or programs for processing the data; etc. Some errors arise from inherent difficulties in collecting certain types of information given the general social situation and the type of respondents involved. The first step in identifying bias is through logical and substantive analysis of the internal consistency of the data. Beyond that, the assessment requires comparison with more accurate information: Data from external sources or data collected with special, improved methods. When the same collection and processing tools are used for the whole population, most sources of measurement bias will be present across the whole population.

346. Measurement bias which is group-specific can systematically change disaggregated estimates. This will often be related to language and culture of the groups concerned. It matters for example if translations of a questionnaire are available and if the terms which are used have the same meaning. It is a good strategy to ensure equivalence by group translation and participation of representatives of the groups concerned in the questionnaire design (see 3). Other sources of measurement bias are directly related to the resource measures considered in poverty measurement (see, for example, the discussion of different measures of cost of living in Chapter 5.1.).

4.1.1.1.2 Measurement variance

347. Different interviewers (coder etc.) often have a unique influence on measurements due to lack of uniformity and standardisation which can give rise to *correlated response variance*. By contrast, *simple response variance* is random, not correlated with any particular interviewer. Instability of particular items

may indicate problems in the questionnaire's wording (e.g. ambiguous terms). Its measurement requires comparisons between independent repetitions of the survey under the same general conditions. There is no way, in a single survey, to distinguish between variation among the true values of units (which contribute to the sampling error), and the additional variability arising from random factors affecting individual responses. Measurement variance contributes to the uncertainty of estimates and can therefore affect the robustness of disaggregated results.

4.1.1.2 Errors in estimation

4.1.1.2.1 Coverage and related errors

348. Coverage errors arise from discrepancies between the target and the frame populations, and also from errors in the way the sample is selected from the frame. Valid inference is only ensured by probability samples which meet all of these criteria: (a) the survey population is fully and correctly represented in the sampling frame; (b) units from the frame are randomly selected into the sample with known non-zero probabilities for all units; (c) all the units selected into the sample are successfully enumerated. Coverage error concerns primarily (a), but also (b); (c) concerns non-response. Poverty measurement must take any effort to ensure that all vulnerable groups are adequately covered (see also 3).

4.1.1.2.2 Sampling error

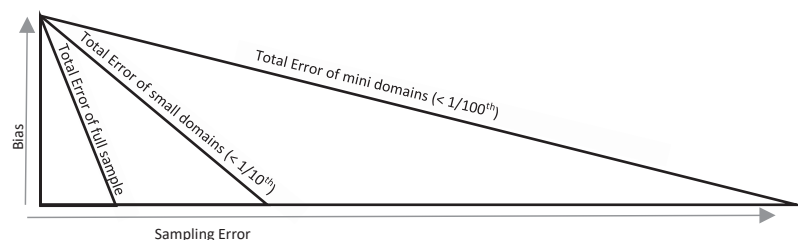
349. Sampling error is the difference between the estimate and the value of the population parameter due to the fact that the estimate is based on measurements taken only on a subset of the population rather than the entire population. The only way to eliminate sampling error from the estimation process is to take a complete census of the population. Each different sample would very likely produce a different estimate creating some variability that would be observed between estimates from different samples drawn using the same sample design as the survey, disregarding any variable errors and biases resulting from the process of measurement and sample implementation. Sampling error represents only one component of the total survey error, the other components being referred to collectively as non-sampling errors. For estimates based on small samples, sampling error is often the dominant one. In other situations, non-sampling errors, in particular coverage, non-response and measurement biases, may be much more important. However, even in these cases, sampling error increases progressively as the estimates are produced for smaller and smaller subgroups of the population, such as for social classes or regions of a country. Therefore, when producing disaggregated poverty statistics, sampling error may well outweigh non-sampling errors.

350. The relative importance of sampling errors for disaggregation is portrayed by the right-angled triangles in Figure 4.1. Total error is often denoted as the root mean squared error and defined by taking the square root of variance and squared bias. It can thus be represented by the hypotenuses of a right-angled triangle (Kish, 1988). The smallest triangle of this figure depicts large survey samples where total error is dominated by the bias. In such situations the precision of poverty measures cannot be increased much by increasing the sample size. Instead it is worthwhile to focus on reducing non-sampling errors. For subpopulations below the national level, the magnitude of bias is, however, often very similar, whereas the sampling error drastically gains in relative importance. Following the classification of Purcell and Kish (Purcell and Kish, 1980), groups which comprise less than $1/10^{\text{th}}$ of the population may be considered as

minor domains. The standard error for estimates for such domains is more than 3 times higher than for the full sample. For these domains, sampling error will overtake bias in many surveys. For mini domains, which Purcell and Kish categorised as groups which comprise between $1/100^{\text{th}}$ and $1/10000^{\text{th}}$ of the population, sampling error will often be the dominant factor and controlling the bias may not substantially improve total error.

Figure 4.1

Increasing sampling error/bias ratios for disaggregations



4.1.1.3 Non-response errors

351. Non-response refers to the partial or complete failure to obtain a measurement on one or more study variables for one or more sample units. More specifically, unit non-response is a type of non-response occurring when no data are collected about a population unit designated for data collection. Item non-response means that a unit is included but information on some items for it is missed. In this typology, item non-response is in an intermediate category between errors in measurement and errors in estimation, whereas unit non-response is considered an error in estimation.

352. Non-response of both types causes an increase in variance due to decreased effective sample size and due to weighting and imputation introduced to control its impact. More importantly, it causes bias in so far as non-respondents are selective with respect to the characteristic being measured. For instance, one might expect persons with high incomes to be more reluctant to give information on their income; similarly, poorer, unemployed and socially excluded persons are more likely to be missed in surveys related to economic well-being.

353. Proposed solutions for item non-response are presented in 2.3section 4.2.3 on imputing missing values, while coverage and related errors and unit non-response are treated in 413 with the introduction of weighting systems.

Box 4.1

Framework to assess errors in poverty measurements (Verma et al, 2010)

Errors in measurement

A) Conceptual errors; these include: i) Errors in basic concepts, definitions and classifications; ii) errors in putting them into practice (questionnaire design, preparation of survey manuals, training and supervision of interviewers and other survey workers).

B) Response (or 'data collection') errors; these include: i) Response bias; ii) simple response variance; iii) correlated response variance.

C) Processing errors; these include: i) Recording, data entry and coding errors; ii) editing errors; iii) errors in constructing target variables; iv) other programming errors.

Mixed category

D) Item non-response; this includes: i) Only approximate or partial information sought in the survey; ii) respondents unable to provide the information sought (“don’t knows”); iii) respondents not willing to provide the information (“refusals”); iv) information suppressed (for confidentiality or whatever reason).

Errors in estimation

E) Coverage and related errors; these include: i) Under-coverage; ii) over-coverage; iii) sample selection errors.

F) Unit non-response; this includes: i) Unit not found or inaccessible; ii) not-at-home; iii) unable to respond; iv) refusal (potentially “convertible”); v) “hard core” refusal.

G) Sampling errors; these include: i) Sampling variance; ii) estimation bias.

Recalling the classical classification into sampling and non-sampling errors, the latter category is comprised of errors of types A) to F) above.

4.1.2 Multidimensional quality frameworks

354. Quality reports are essential tools to assess, improve and communicate the quality of poverty measurement. At a minimum, such reports should describe in sufficient detail all sources of error that limit the accuracy of poverty measures. Overall however, quality should be more broadly defined in terms of user needs, as “fitness to use” for the purpose for which the data were created (Juran and Gryna, 1970).

4.1.2.1 Questions which should be addressed to evaluate the utility of a survey

355. Table 4.1 reports an illustration of overlapping concepts and categories used by different organisations to identify dimensions of quality (taken from Lee and Shon, 2001). For the European Union, the legally required content of quality reports addresses all of these questions which are also reflected in Article 12 of the European Union Statistics Act (Regulation (EC) No 223/2009).

Table 4.1

Concepts and categories used by different organizations to identify dimensions of quality

Canada	Netherlands	R. of Korea	IMF	Eurostat
			Prerequisites of quality	
Relevance	Relevance	Relevance		Relevance
Accuracy	Accuracy	Accuracy	Accuracy and reliability	Accuracy
Timeliness	Timely	Timeliness	Serviceability	Timeliness and Punctuality
Accessibility		Accessibility	Accessibility	Accessibility and clarity

Coherence				Coherence
		Comparability	Methodological soundness	Comparability
Interpretability				
			Integrity	Completeness
	Cost-effectively	Efficiency		
	Without too much a burden			

Source: Lee and Shon, 2001.

Box 4.2

European Union: content of quality reports as required by the law

In October 2019 the European Union established a new framework regulation ((EU) 2019/1700) which integrates all major social surveys in the European Statistical System including the European Union Statistics on Income and Living Conditions (EU-SILC). Member States are thus legally required to meet specified quality criteria and produce regular quality reports. The content of these reports is specified in an implementing regulation which lists the required information as follows:

1. Contacts
2. Statistical Presentation
 - 2.1 Data description
 - 2.2 Classifications
 - 2.3 Sector coverage (main themes)
 - 2.4 Statistical concepts and definitions (including the reference period)
 - 2.5 Statistical units
 - 2.6 Statistical population
 - 2.7 Population(s) not covered
 - 2.8 Reference area
 - 2.9 Time coverage
3. Statistical Processing
 - 3.1 Source data (e.g. interviews, administrative data)
 - 3.2 Sampling frame
 - 3.3 Sample design
 - 3.4 Frequency of data collection
 - 3.5 Data collection (mode such as CAPI, CAWI, CATI, etc., translated questionnaires)
 - 3.6 Data validation (including explanation how it is reflected in the results).
 - 3.7 Data compilation (e.g. data editing, imputation, weighting etc.)
4. Quality Management
 - 4.1 Quality assurance (e.g. EFQM, ISO 9000)
 - 4.2 Quality assessment (main strengths, trade-offs and deficiencies)
5. Relevance
 - 5.1 User needs
 - 5.2 User satisfaction
 - 5.3 Completeness (variables which are not transmitted)
6. Accuracy and Reliability
 - 6.1 Overall accuracy (esp. effect of random and systematic errors for key estimates).

- 6.2 Sampling error (methodology, national and regional standard errors for indicators)
- 6.3 Non-sampling error
- 6.4 Seasonal adjustment (where applicable)
- 6.5 Data revision (policy and practice)
- 7. Timeliness and Punctuality (dates of dissemination and end of fieldwork)
- 8. Coherence and Comparability
 - 8.1 Comparability – geographical
 - 8.2 Comparability – over time
 - 8.3 Coherence – cross domain
 - 8.4 Coherence – National accounts
 - 8.5 Coherence – internal
- 9. Accessibility and Clarity (dissemination formats, documentation)
- 10. Cost and Burden (cost of collection and production, duration of interviews)
- 11. Confidentiality (policy, data treatment)
- 12. Comment (Supplementary descriptive text that can be included in the quality report)

4.1.2.1.1 How relevant is the data?

356. Relevance refers to the capacity of the data to meet users' needs. It implies the identification of users and their needs, and assessment of the extent to which their needs are actually met. The concept also covers the *potential* of the data in meeting the relevant needs. According to Statistics Canada's Survey Methods and Practice (Statistics Canada, 2003): "Assessing relevance is a subjective matter dependent upon the varying needs of users. The statistical agency's challenge is to weigh and balance the conflicting needs of current and potential users to produce a program that goes as far as possible in satisfying the most important needs within given resource constraints". Relevance also depends on the extent to which stakeholders and social groups which are considered for disaggregation were involved.

4.1.2.1.2 How timely and punctual are results available?

357. As defined by Statistics Canada's Survey Methods and Practice (Statistics Canada, 2003): "The *timeliness* of statistical information refers to the delay between the reference point (or the end of the reference period) to which the information pertains, and the date on which the information becomes available. It is typically involved in a trade-off against *accuracy* (see below). The *timeliness* of information will influence its *relevance*."

358. It is important to note that the requirements of timeliness can conflict with those of accessibility and clarity, and above all with those of accuracy. At a minimum, the data must be checked and corrected to a high standard before their public release. Obviously, releasing data or results without adequate editing and correction can be misleading and wasteful. *It can also damage the credibility of the producer organisation.* For instance, Fellegi (2001) identifies credibility as a "survival" issue for a statistical organisation.

359. Punctuality refers to adherence to a pre-established time schedule for the release of statistics. Timeliness is a more objective criterion, assessing how fresh are the data and whether they became available when most needed. Punctuality acquires increased importance in the EU-wide context. The so-called European semester is an annual coordination process for policies in EU Member States. The degree, to which important fiscal and economic decisions may take social conditions into account, depends also

on the timely availability of indicators for many countries simultaneously. The requirements of punctuality have been expressed very strongly in EU-SILC regulations.

4.1.2.1.3 How precise are the results?

360. Data accuracy includes the assessment of survey errors which were discussed in the previous section. It is of such fundamental importance that it has been customary in survey practice to focus on accuracy, sometimes at the expense of – or even to the exclusion of – other dimensions of quality.

361. According to Statistics Canada (Statistics Canada, 2003): “The *accuracy* of statistical information is the degree to which the information correctly describes the phenomena it was designed to measure. It is usually characterized in terms of error in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components. It may also be described in terms of the major sources of error that potentially cause inaccuracy (e.g., sampling, coverage, measurement, nonresponse, and processing).”

362. Ideally all indicators should be published with an indication of their accuracy, including whether there are conceptual differences with regard to international standards. In practice, sometimes sampling errors are presented for selected main indicators only. For disaggregation of poverty measures, it is especially important to identify clearly those results which provide only limited accuracy. Indicators with unacceptable inaccuracy should not be published. More information on best practices for dissemination is discussed in 4.2.6.

4.1.2.1.4 To what degree are findings comparable?

363. Comparability is increasingly considered a central requirement of data quality, especially for measures of poverty. Partnerships for development in the context of the 2030 Agenda for Sustainable Development require comparable measures for poverty. To improve the international comparability and availability of statistics on poverty and the related metadata, the Conference of European Statisticians (CES) established a Task Force in 2014, which worked through 2015 and 2016 to develop a Guide on Poverty Measurement (UNECE, 2017a). The guide states that “many international organizations—the World Bank, OECD, UNDP, Eurostat, just to mention a few—produce poverty data. There have been continuous efforts to improve capacity in statistical offices to develop poverty measures in line with international standards. However, in most cases, these data are not comparable and often cover only a limited number of countries. A lack of comparable data across countries and time impedes effective policy actions. Data produced by countries are not always comparable internationally, largely for two main reasons: i) Country data primarily respond to national needs, which do not always correspond to international standards; and ii) Country data reflect national statistical capacities, which are not always able to meet international standards”.

4.1.2.2 How coherent is the data with other statistics and over time?

364. According to Statistics Canada (Statistics Canada, 2003): “The *coherence* of statistical information reflects the degree to which it can be successfully brought together with other statistical information within a broad analytic framework and over time. The use of standard concepts, classifications and target

populations promotes coherence, as does the use of common methodology across surveys. Coherence does not necessarily imply full numerical consistency.”

365. *Coherence does not necessarily mean identity:* Often there are genuine and inherent differences in the information coming from sources of different types. What it means is whether different sources together lead to a consistent picture, with each making a contribution towards the development of the picture. In the case of surveys on income and poverty, the most relevant sources for external comparison include national household budget and labour force surveys, national accounts, and various administrative and other sources depending on the country. In the European Union all countries conduct EU-SILC while some countries have also well-established official or academic surveys on the same topic. Quality reports should bring to the attention of users any incoherence with external information. Although in practice it will often not be possible to say which source may be more accurate it is important for users to be informed about differences and their possible explanations.

366. In a panel survey and, in fact, in any continuing survey, coherence over time is also a fundamental requirement. Only under this condition can we study trends, aggregate data over time, or construct micro-level longitudinal measures.

4.1.2.3 Are data accessible and clear?

367. These aspects refer to the extent to which the statistical data are available in the form and under conditions which meet users’ requirements, and to how well the data are described and documented for the purpose. Conditions of availability include a whole range of factors such as restrictions on who can or cannot get access to the data, what items of information are suppressed, what restrictions apply on the conditions and purposes of data use, and also the difficulties, delays and the costs involved in gaining access to the micro data and timeliness of accompanying quality reports.

4.1.2.4 What else would users need to know?

368. A variety of other aspects are also covered in various data quality frameworks. Some of these overlap – different terms indicating more or less the same thing, perhaps from a somewhat different point of view or with a somewhat different emphasis. We have, for instance, “completeness” in Eurostat terminology, “integrity” in that of IMF and, along a different line of thought, “interpretability” at Statistics Canada. Cost efficiency and minimisation of respondent burden are other aspects included as quality dimensions, especially in national frameworks (The Netherlands, South Korea in Table 4.1). Surprisingly, not all frameworks explicitly refer to “comparability” as a dimension.

4.1.3 Relationship between different aspects of quality

369. It is safe to assume that no statistical agency is capable of meeting all of the above criteria to the same degree. To a certain extent, the *different dimensions of data quality compete against each other*, an obvious example being the common conflict between timeliness and data accuracy – “quickly released but rough data, versus refined data but much delayed”. Different aspects of data quality can also *mutually*

support and reinforce each other, one often forming a precondition for the other. For instance, it is hardly possible for two data sets to be comparable, when either or both lack statistical accuracy. Perhaps most critically, a survey loses its relevance if it is not timely and accurate enough.

370. Verma (Verma, 1981) and later Verma et al (Verma et al, 2010) propose to think about reduction in data quality in any dimension as a loss in the utility of the information. The loss may be more or less steep depending on the particular context. Often the resources saved by reducing quality in one dimension can be used to improve quality in other dimensions; however, some dimensions can also be linked in such a way that a quality loss in one dimension necessarily implies a loss in the other as well. Beyond a certain point, there is likely to be a critical zone when further reduction in quality along a particular dimension would result in increasing drastically the loss in the overall utility of the data. A certain minimum degree of quality has to be present in *every* dimension for the statistical information to remain useful overall.

4.2 Improving quality in each survey step

371. Once the quality profile of the survey is understood, action for methodological improvements can be taken. Improvements may be considered in each survey step, including design, data collection, edit and imputation, weighting, variance estimation and dissemination (Groves et al, 2004).

4.2.1 Survey design

372. Some survey designs lead to more accurate estimate or can produce more disaggregated statistics. However, global survey design decisions must balance these advantages with costs and other dimensions of quality such as timeliness and comparability.

373. Official measures of poverty that are based on sample surveys must use a probabilistic sample. Beyond this crucial requirement, the following design issues are particularly relevant for disaggregated poverty statistics. Issues specific to hard-to-reach groups are also discussed in 3Including Hard to Reach Groups in Poverty Measurement.

4.2.1.1 Choice of data source

374. The form of data collection can have serious consequences for disaggregation. Income questions can be difficult to answer. When respondents provided income data directly, they will typically provide rounded numbers and may forget to include certain types of income. For instance, a respondent may remember their employment income but forget about or not be able to provide details about investment income.

375. As an alternative, in many countries, income data is taken from administrative sources, such as files created for the administration of income taxes or government programs, or from registers. This

reduces the burden imposed on respondents and this data is often more accurate than respondent data as it is less prone to rounding and recall error. Moreover, administrative tax data may also be classified into more detailed income source categories, depending on the categories that are used for taxation, allowing for the production of statistics that are more disaggregated.

376. This is not to say that measurement errors are non-existent for administrative data. Conceptual errors are an important consideration since the categorization used on the administrative data may not align with the concepts that desired for the income survey. In particular, non-taxable or undeclared types of income may not be found on administrative sources but should be included in income statistics to give a complete picture of an individual's income. Processing errors must always be considered, including for data acquired from a source not controlled by the survey team. Also, the definition of household membership in registers may be quite different from what it is in reality.

Box 4.3

Use of administrative tax data at Statistics Canada

At Statistics Canada, administrative tax data has been the primary source of income data for about 25 years. Using this source has numerous advantages. The Statistics Canada's surveys on income, spending, and consumption all link to the same administrative tax data and process this data in a harmonized way. Additionally, the Canadian Census also uses the same administrative tax sources. This results in increased coherence between these surveys. Using the administrative data led to more precise measurement of the various sources of income, while decreasing the burden place on respondents.

On the Canadian Income Survey, extra questions are asked to respondents to measure income concepts that are not fully included on tax form such as full amounts of support payments (alimony), whether received or sent, and other transfers between households. Collection is also required to obtain variables of interest for disaggregation of poverty statistics such as family composition.

377. Perhaps most importantly, the timeliness of administrative data can be a major drawback to their use for statistical purposes. Since tax data needs to be collected and processed by the tax agency before being provided to the statistical agency, it can take quite a while before it is available. This is a classic example of the conflict between timeliness and accuracy.

378. The choice of data source will have a large impact on what is measurable by the survey. For instance, where administrative data includes details of government transfers such as credits given to families with children, the use of such data can be used to measure the impact of the programs on poverty which may then be different for different types of families. Details about these types of programs could be difficult or even impossible to obtain directly from survey participants as they may not be aware of the details of these transfers or even that a particular program affects them. Appropriate documentation and metadata should be made available and disseminated in order to make the users aware of the specificities of the data coming from countries making use of administrative and register data.

Box 4.4

Use of registers in the European Union

For the European Union the importance of registers for poverty measures is discussed in some detail in a volume by Jäntti, Törmälehto and Marlier (Jäntti et al, 2013). For the year 2012 it was found that 19 out of 28 EU Member States used registers to obtain income information in the year 2012 (Di Meglio and Montaigne, 2013). Apart from the traditional use of registers in the Nordic countries, several countries have seen transitions from interview-based data collection towards use of registers or were planning to do so in 2012. The experience made by these countries appears particularly valuable for countries that intend to use more register information: France, Italy, Latvia, Switzerland, Ireland, Austria and Spain. Overall, it was found to be a good practice to assess carefully the impact and have at least one overlapping measurement from both interviews and registers. For example, in Austria register data was introduced in the year 2011. This led to a decrease in poverty rates by about 2 percentage points for Austria. As register data was accessible also for previous years it was also possible to backcast to earlier waves of the survey including the year 2008 (Statistics Austria, 2017).

Table 4.2

Use of administrative data and registers for each domain covered by EU-SILC (2012)

Using administrative data in the following domains (even partly)	Countries	Number
Demographic/household data	Austria, Bulgaria, Belgium, Denmark, Estonia, Finland, Iceland, Italy, Lithuania, Latvia, Netherlands, Spain, Sweden, Slovenia, and Norway.	15
Education data	Denmark, Finland, Iceland, Spain, and Norway.	5
Labour data	Bulgaria, Denmark, Iceland, Netherlands, Spain and Norway.	6
Housing/dwelling data	Austria, Denmark, and United Kingdom.	3
Income data	Austria, Bulgaria, Belgium, Cyprus, Denmark, Finland, France, Iceland, Ireland, Italy, Lithuania, Latvia, Malta, Netherlands, Spain, Sweden, Switzerland, and Norway.	19
Other	Malta (electricity and water consumption)	1
Not using administrative data	Czechia, Denmark, Greece, Hungary, Luxembourg, Poland, Portugal, and Slovakia.	8

Source: Di Meglio and Montaigne, 2013.

Register-based households are not always composed of the same persons as the survey-based household. In Finland for example, about 10% of households had a different composition according to the register definition than to the information in the conventional housekeeping concept. This discrepancy was found to be only 5% among retired persons and 30% among student households (Jäntti and Törmälehto, 2013). Particular population groups may hence be inappropriately reflected in register data.

Out of the 19 countries that use income information from registers, 9 countries mentioned concerns about the coverage of all groups in the population. For example, in Austria the personal identification number (PIN) required to link survey information with registers is missing for up to 5% of the sampled persons. The information was more often missing for certain groups: persons under the age of 40, persons living in the capital, persons with a foreign citizenship. Because register information is often generated in the context of employment, missing PINs were also more frequent among jobless persons

and persons who are mainly fulfilling domestic tasks (Heuberger et al, 2013). This coverage error of units is potentially an important obstacle to using register data when poverty measures should be disaggregated. If this register under-coverage is known before the fieldwork, survey questions should be adjusted to collect specifically the information on groups which are otherwise missing. For example, for lone parents, transfers between households are often the main source of income. This component of income is usually not included in registers. To get appropriate measures for this group and their children, it is crucial to ask specific questions on alimonies.

379. When administrative data is used, survey data obtained directly from respondents is generally still required to obtain variables for disaggregation. In addition, valid information on certain sources of income such as transfers between households or self-employed income, capital income or hidden economy may only be collected through survey questions. As a result, record linkage is often required to combine the survey and administrative data.

Box 4.5

Record linkage through statistical identifiers in Austria

In Austria, administrative data which are provided by various authorities are entered into registers with a unique personal identifier for statistical purposes. This unique identifier is provided also for each sample unit so that register information can be linked for almost every unit. For privacy concerns it is guaranteed that those identifiers cannot be matched to administrative data held by other authorities. For example, it would be impossible to report back respondent's answers to tax authorities. Also, identifiers have to be kept separate from names and addresses which are used only during fieldwork.

Register households are used as the sampling unit, but household membership is always verified by face to face interviewers. In other words, individuals who are living at the selected address but were not on the register are added to the sample and individuals from the register not (or no longer) living at the selected address are excluded from the sample. This sometimes involves cumbersome enquiries to obtain linkable PIN. This practice also dictates the mode of data collection in the initial contact, leaving telephone or web interviews only as options in the case of follow up panel waves.

4.2.1.2 Level of disaggregation and sample design

380. The production of more disaggregated statistics generally requires larger samples. This objective has to be balanced against budget considerations and clearly communicated to all key stakeholders. For example, the European Union the Directorate General that is responsible for the allocation of regional funds has specified regional precision requirements that EU-SILC must attain. These requirements are partly reflected in Annex II of a newly established framework regulation which applies to EU-SILC (EU, 2019/1700).

381. Where disaggregation variables are available on the survey frame, the sample design may use them as stratification variables. This can help improve accuracy for disaggregations without considerable increases in total sample size. For example, to make inferences on differences in poverty between rural and urban areas, it is useful to consider such variables as strata. The stratification used for a survey will depend primarily on what is available on the frame, though geographic variables are often chosen. This is often done when surveys need to produce estimates for subnational regions. When these regions are of

unequal size, it often translates into different sampling rates by regions in order to achieve a desired level of precision. For example, within the European Union, differences in sample size are much smaller between countries than the differences in population size between countries. This helps make comparisons between Member States.

382. To calculate a minimal sample size for the various domains of interest, the following information is required: the desired precision, or margin of error, for the estimates; the desired confidence level; the size of the domain; the variability of the variable of interest in the domain; and the expected response rate (Eurostat, 2013b). When a complex design such as a clustered sample is used, the effect of the sample design should also be considered. This can be done using the design effect, which is defined as the variance of an estimator under the complex survey design divided by the variance of the estimator had simple random sampling with the same sample size been used. In practice, initial estimates of the required sample sizes are often revised downward when collection costs are calculated. In this case, it is generally necessary to reconsider the precision that is desired.

383. Using different sampling rates to ensure adequate sample sizes for domains of interest may impact negatively on the precision of estimates for the population as a whole. The allocation of the sample to various strata is an important consideration. It will depend on the same factors used to calculate a minimum sample size, such as the variability of the variables of interest by strata, the size of the strata, the expected response rate and the cost of collection. Information about sample size determination and various allocation methods can be found in many of the classical survey sample texts (Kish, 1965; Kish, 1987; Cochran, 1977; Särndal et al, 1992; Lohr, 1999).

384. While the theory is well-established, in practice there will not be one ideal sample allocation for a survey. On the one hand, this occurs because of the multivariate nature of surveys. The allocation that is ideal for one variable of interest may not be for another as they may both have different distributions or variability for example. In a situation where multiple variables are of interest, it is important to remember to choose stratification variables that are related to many of them. Bethel (1989) discusses allocation in the context of multivariate surveys. Additionally, when the goal is to provide statistics at a disaggregated level, the survey will generally be called upon to provide estimates for a whole hierarchy of domains. What is optimal for one class of domains will generally not be for another. In this case there will not be one formula that will give the best allocation. It is good practice to verify the impact of the allocation on a variety of domains that will be used for dissemination. It is also worth verifying that the allocation that is chosen is not too sensitive to small differences in the allocation since the final number of respondents in each stratum will be different from the number selected due to non-response.

385. In all cases, it is of utmost importance that sample design information be appropriately documented and stored as this is an essential condition for sampling variance to be appropriately assessed. If microdata are disseminated, information on stratification and primary sampling units (PSUs) should ideally be provided unless this information could compromise the confidentiality of respondents. In the European Union, the calculation of appropriate sampling errors is made more complicated because this sample design information is not accessible for all countries for Eurostat (Verma et al, 2010; Goedeme

2013; Trinidad and Goedeme, 2016).⁷¹ When sample design information cannot be shared with microdata users, it a minimum they should be given an indication of the design effect which provides a rough estimate of the impact of the sample design on the variance.

Box 4.6

Design of the American Community Survey in the United States

The American Community Survey (ACS) (United States Census Bureau, 2014) has been conducted by the U.S. Census Bureau since 2005. It uses a series of monthly samples to produce annually updated estimates specifically for the small areas (census tracts and block groups). Formerly, these areas were surveyed via the decennial census long-form sample. Five years of samples are required to produce these small-area data. The Census Bureau released its first 5-year estimates in December 2010 and, since then, new small-area statistics are produced annually based on a rolling sample. The ACS includes people living in both housing units (HUs) and group quarters (GQs). The ACS is conducted throughout the United States and in Puerto Rico, where it is called the Puerto Rico Community Survey (PRCS).

In total the ACS sample comprises about 3.54 million addresses per year (approximately 295,000 per month). These addresses are selected independently for each of the 3,143 counties and county equivalents in the U.S., including the District of Columbia, as well as for each of the 78 municipalities in Puerto Rico. Increased sampling rates were used for the smallest sampling entities.

The ACS complements, rather than replaces the monthly Current Population Survey (CPS) which is run jointly by the U.S. Census Bureau and the U.S. Bureau of Labor Statistics and has an annual Social and Economic Supplement which is commonly used for poverty statistics.

4.2.1.3 Repeated surveys

386. Surveys on poverty and economic resources are generally repeated in order to understand trends over time. It is therefore important to determine whether a cross-sectional or longitudinal design is desired. In a cross-sectional survey, the sample is used once and for the next repetition of the survey a new independent sample is chosen. In a longitudinal survey, data is collected from the selected sample on several occasions, often over many years. In this case, the sample is generally referred to as a panel.

387. For the measurement of poverty, longitudinal surveys can be of particular interest since they allow for the measurement of change at the individual level. As a result, issues such as the persistence of poverty can be measured by a longitudinal survey much better than by a cross-sectional survey. Measuring these types of issues using a cross-sectional survey relies on respondents being able to accurately report on their situation in the past which is known to be definitely less reliable than taking the measurement twice, at two points in time. Additionally, longitudinal surveys reduce the sampling variance for estimates of change ($\hat{Y}_1 - \hat{Y}_2$ where \hat{Y}_1 is the measure at time 1 and \hat{Y}_2 is the measure at time 2).

388. Longitudinal surveys are not without their challenges and disadvantages, however. Collection is complicated by the presence of movers that must be followed to their new address (see Iacovou and Lynn, 2013). The representativity of a longitudinal survey decreases the further you are from the time at which the panel was formed due to non-response, and to changes in the population such as birth, deaths, and

⁷¹ See also the important resources accessible here <https://tingoedeme.com/eu-silc-standard-errors/>

immigration. This is a particular challenge for disaggregation. Without adequate sample refreshments, the population of new migrants – which are often vulnerable to poverty – cannot be represented which can lead to serious bias (Glaser et al, 2015). Non-response due to respondent fatigue is a particular challenge for longitudinal surveys. Since non-response compounds over time, it is harder to measure and treat sufficiently in longitudinal surveys than in cross-sectional surveys. Finally, costs can be a concern for longitudinal surveys since they require budget to be guaranteed over an extended period of time.

389. There exist intermediate solutions between a cross-sectional and longitudinal survey and allow for objectives of both to be balanced. In a rotating panel design a part of the sample is replaced at each iteration of the survey. This leads to more accurate measures of change while limiting the issues relating to representativeness and non-response attrition associated with longitudinal surveys. That being said, many of the collection and definitional issues carry over from a longitudinal design to a rotating panel design. The EU-SILC uses this type of design.

4.2.1.4 Mandatory and voluntary surveys

390. Whether a survey is mandatory is another decision to be made. Mandatory surveys generally lead to higher response rates and lower collection effort being required which can lower costs. However, the decision as to whether a survey can or should be mandatory or voluntary depends on the legal framework under which a statistical agency operates and is often country-specific. Even when the legal framework provides for mandatory surveys, statistical agencies may decide not to make a survey mandatory in order to conform to social expectations and to maintain a good relationship with their constituents. In EU-SILC participation is mandatory in only a few countries. From the perspective of measurement errors, it is not always advisable to make survey participation mandatory (see Glaser et al, 2015).

Box 4.7

Design of the Canadian Income Survey

Statistics Canada has used a series of surveys to provide information on the income and income sources of individuals and families in Canada. Since 2012, this is done using the Canadian Income Survey (CIS), an annual household survey. The CIS is a supplement to the Canadian Labour Force Survey (LFS) which uses a probabilistic sample selected from an area frame using a multi-stage survey design. The CIS inherits its sample from the LFS. From January to June every year, the LFS respondents in their last month of LFS collection are asked to answer the CIS questionnaire immediately following the LFS. Though the LFS has a rotating panel design, by collecting only during the last month of LFS collection, CIS has a purely cross-sectional design. While the LFS is a mandatory survey, CIS is not.

Income data for the CIS is gathered from administrative tax files and record linkage is used to combine it with survey data. Though it has clear advantages from the point of view of accuracy and response burden, the principal disadvantage of using tax data is timeliness. The CIS is disseminated 14 months after the end of its reference year. The principal reason for this is that tax data only becomes available for processing with the survey data nine months after the end of the reference year.

By using administrative data as its income source and combining its collection with that of the LFS, during which personal and household characteristics have already been collected, the CIS can use a relatively short questionnaire that can generally be completed within 10 minutes. The CIS gathers

additional data on labour market activity, school attendance, activity limitation, support payments, inter-household transfers, and characteristics and costs of housing.

Since its sample design and sample size is tied to that of the LFS, CIS can only be used to produce disaggregated statistics to the degree that the LFS design allows. Statistics Canada also builds complementary products produced directly from administrative tax data allowing further geographic disaggregation of income. Though administrative data can be used for producing statistics for small geographies, on its own, it cannot disaggregate along some other variables of interest such as household composition.

Box 4.8

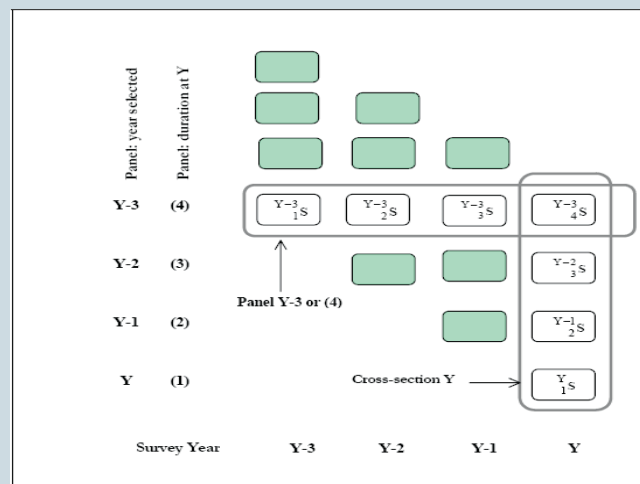
EU-SILC Sample Design

The EU-SILC uses a rotating panel design that allows for the production of both cross-sectional and longitudinal statistics. In most countries⁷², the EU-SILC is composed of four independent samples as rotation groups. In those countries that adopted this scheme, one fourth of the sample, corresponding to one panel, is replaced every year (see Figure 4.2 below). As a result, units in the sample are followed longitudinally over a four-year period. Movers from the original sample are followed-up at their new location for the time their panel remains in the survey. This standard integrated design has been adopted by nearly all the participating countries.

This design yields a cross-sectional sample as well as longitudinal samples of various durations (Verma, 2001; Verma and Betti, 2006). The cross-sectional sample for year Y consists of four subsamples, 1-4, one introduced each year from (Y-3) to Y. A longitudinal sample consists of persons who have remained in the survey since they were first introduced into it. Three overlapping longitudinal samples of different durations are formed: of two-year duration from subsamples (2+3+4), of three-year duration from subsamples (3+4), and of four-year duration from subsample (4).

Figure 4.2

EU-SILC standardised rotating scheme



Source: Verma and Betti, 2006.

⁷² Exceptions are for example the EU-SILC sample in Luxembourg, which is essentially a permanent panel or France and Norway which have a much longer panel duration.

4.2.2 Data collection

391. Research on the Labour Force Survey in Austria (Glaser et al, 2015) has shown that even in official surveys where participation is compulsory and non-response may incur a penalty fine, systematic differences across groups can be observed which may also lead to considerably biased results (Meraner et al, 2016). Moreover, non-response leads to increased variance by decreasing the size of the sample used at estimation. Data collection is the survey step at which these issues can be prevented, or at least minimized.

392. A lower response rate among people at both ends of the income distribution is a common occurrence and impacts the measurement of poverty. Factors potentially related to poverty, for instance lower levels of education or language barriers among recent immigrants, can also make responding to surveys more difficult and lead to higher non-response. Longitudinal surveys are essential for understanding the persistence of poverty but come with the extra challenge of minimizing attrition due to respondent fatigue.

393. In order to encourage high and balanced response rates, collection should be planned so that it minimizes the burden placed on respondents while maximizing their perceivable benefit. The use of administrative data sources to collect income information can greatly reduce the burden imposed on survey participants.

394. For the variables that cannot be obtained from administrative sources, there are many ways to simplify the process for respondents. A survey on income can be conducted at a time of year that is shortly after individuals prepare or review their income tax documents, thereby making it easier to answer questions on income. Letting respondents know ahead of time which type of information is required to answer the survey can help reduce item non-response. Surveys measuring consumption often do so by having respondents fill out a diary of their purchase over a period of time (often a week or two for example) and shorter periods may lead to higher rates of completion of the diary. Proxy interviews, that is, obtaining information for an absent respondent from another knowledgeable person, generally leads to a higher response rate. Though the quality of proxy responses tends to be lower than if the response was obtained directly, depending on the question asked, it can often be adequate especially for members of the same household.

395. Communication strategies can be used to lower unit non-response. Letters sent to the sampled individual before they are contacted by interviewers can increase survey participation. Dillman, Smyth and Christian (Dillman et al, 2014) have demonstrated empirically that when attention is paid to the details of respondent communication, participation can be vastly improved substantially, notably when financial incentives are used to frame participation in the survey as social exchange.

Box 4.9

Data collection for the American Community Survey

Because a high level of self-response is cost critical, the ACS employs multiple mailings to encourage respondents to complete the survey via the Internet or to return a paper questionnaire. ACS materials

for U.S. addresses are printed in English, and Puerto Rico Community Survey (PRCS) materials sent to Puerto Rico are printed in Spanish. U.S. respondents can request Spanish mailing packages, and Puerto Rico respondents can request English mailing packages, via telephone questionnaire assistance (TQA).

For most HUs, the first phase includes a mailed request to respond via Internet, followed later by an option to complete a paper questionnaire and return it by mail. If no response is received, the address may be selected for computer-assisted personal interviewing (CAPI).

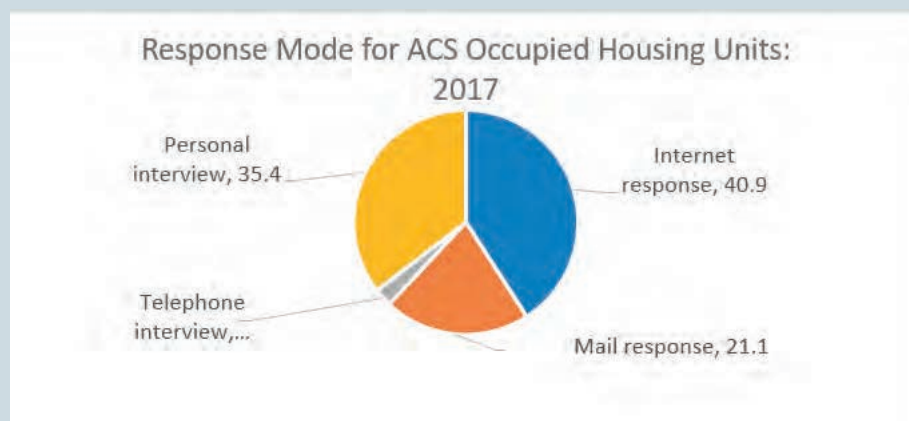
The ACS includes 12 monthly independent samples. Data collection for each sample lasts for three months, with mail and Internet returns accepted during this entire period. This two-phase process operates in continuously overlapping cycles so that, during any given month, two samples are in the mail/Internet phase, one is in the CAPI phase (as shown in Table 4.3).

Table 4.3

Two overlapping phases of the American Community Survey

ACS Sample Panel	Month of Data Collection						
	November	December	January	February	March	April	May
November	Internet / Mail Response (HUs Receive Mailings 1 – 5)		Personal Visit				
December		Internet / Mail Response (HUs Receive Mailings 1 – 5)		Personal Visit			
January			Internet / Mail Response (HUs Receive Mailings 1 – 5)		Personal Visit		
February				Internet / Mail Response (HUs Receive Mailings 1 – 5)		Personal Visit	
March					Internet / Mail Response (HUs Receive Mailings 1 – 5)		Personal Visit
April						Internet / Mail Response (HUs Receive Mailings 1 – 5)	

Figure 4.3

Distribution of response modes in the American Community Survey 2017

Notes: 2017 American Community Survey 1-year estimates; data from group quarters are excluded. The number of households contacted via telephone follow-up interview decreased in 2017. The operation was ended in October 2017.

Source: <https://www.census.gov/library/visualizations/interactive/acs-collection.html>

396. To reduce potential bias due to non-response, response rates should be monitored during collection. Indicators of representativity which are based on the variance of response rates between groups can be calculated. A tool that can be used to evaluate non-response is the R-indicator.⁷³ Introduced by Schouten, Cobben and Bethlehem (Schouten et al, 2009), it measures the degree to which response propensities are constant over subpopulations. Originally, R-indicators had been limited to those variables which are available in the sampling frame only. More recently, the method has been extended so that information which is available only at the population level may be used (Bianchi et al, 2016). If certain areas or domains have lower response rates, collection staff can be reallocated to work on cases in these groups. If variables on the frame identify hard-to-reach groups or groups to which hard-to-reach individuals are more likely to belong, then particular attention should be made to obtaining sufficient response rates in these groups. For longitudinal surveys, data from previous waves is a particularly rich source of data for monitoring collection, and if some groups are trailing in terms of response rates, tracing activities can be concentrated on the groups in question. The R-indicator was recently also used to evaluate the impact of panel attrition of representativeness of the EU-SILC (Luiten and Schouten, 2019).

397. Response rates among groups which are especially hard to contact can also be vastly improved by requesting a specified minimum number of documented visits or calls. Satisfactory contact rates require sufficient time. For groups which can be expected to be away from home such as young single adults, contact modes (phone, text messaging or internet) may be adjusted. Different expected response rates may be considered when the order of interviews or even payment per interview are determined.

398. The practice of sample substitution is not an appropriate measure for managing non-response and should be avoided as it does not reduce the bias introduced by non-response. It can also encourage

⁷³ <https://www.cmi.manchester.ac.uk/research/projects/representative-indicators-for-survey-quality/>

poor collection practices such as not making enough effort to obtain responses from households that do not respond at first contact, which can worsen non-response bias when present. Sample substitution makes it difficult to calculate appropriate response rates and assess the quality of the data which is obtained. In the European Union, guidelines for EU-SILC are clear in this respect, stating that “As a rule, the units enumerated in the survey shall be exactly the same units as those selected for the purpose in accordance with the sampling design, i.e. not substituted for by other units.” (Eurostat, 2014)

399. The second type of error to be addressed during data collection is measurement error. Many of the strategies for reducing response burden mentioned above, also help reduce recall error and lessen measurement error. Yet again, one worth highlighting in the context of poverty measurement is the use of administrative or register data on income. Its use has been shown to reduce measurement error in many countries. Interviewer training is an important way to address measurement issues as well, especially for hard-to-interview populations.

400. For the measurement of poverty, it is still common to use personally assisted modes, rather than self-administered questionnaires. Mostly, computer assisted personal interviews (CAPI) have replaced the conventional paper and pencil mode (PAPI). This gives additional control over measurement errors due to routing mistakes in the questionnaire. The use of automatic checks can ensure that interviewers instantly detect and can probe respondents on potential inconsistencies. For example, if a respondent states that someone in the household receives a pension, this should be aligned with the activity status of that person; gross salaries should always be larger than net salaries; social benefits will usually fall within certain limits; etc. This requires skills and sufficient time in programming the data collection software. With computer assisted data collection data can be processed without any further delay. Logistics with paper questionnaires becomes obsolete as well as provisions for scanning or manual data entry. On the other hand, issues such as data transmission, privacy, and performance of software solutions and reliability of communication technology become more pertinent when computers are used in data collection.

401. An increasingly relevant aspect of data collection comes with the – often simultaneous - use of different data collection modes such as telephone (CATI) and self-administered internet questionnaires (CAWI). In the European Union most countries are about to or have already turned to some kind of mixed mode design. Mixing survey modes is often seen as an opportunity to save survey cost. The potential switches between modes has however implications for the case management software. Furthermore, each mode is likely to exhibit specific selection bias and measurement errors. If data or a sub-population is accessible only by a certain mode, their measurements may reflect the specific methodological effects attached to that mode. When poverty measures are then disaggregated it may then be difficult to distinguish the effect of the different data collection modes from the true differences between groups. Many countries are using BLAISE or have developed their own data collection tools. The current state of mixed mode data collection methods has been documented in the MIMOD project, in which several EU Member States shared their views on survey organisation and software for mixed mode data collection (Signore, 2019).

402. When data is obtained from administrative sources, it is important to consider the details of how it is obtained. When no statistical registers are available to facilitate the association between the

administrative data and the data obtained directly from respondents, software such as Italian RELAIS (Record Linkage at ISTAT) and Statistics Canada's G-Link (Statistics Canada, 2017a) is available to facilitate the process of data linkage to administrative sources. No matter which record linkage software is used, one of the most important steps of record linkage is the pre-processing of linkage variables such as names and addresses. The cleaning and standardization of linkage variables can substantially improve the quality of the subsequent linkage. The record linkage process can be thought of as simply another data collection strategy. Just as the response rate is an important indicator of quality that should be shared with data users, so too is the linkage rate.

4.2.3 Data processing and imputation

403. Raw data that is received from collection is invariably messy and difficult to use directly. Cleaning and processing of the raw data are crucial to transform it into a useable dataset. Some variables will have values that are missing. This can occur for many reasons; for instance, because of refusals on the part of the respondent to answer certain questions, because the respondent did not know the answer, or because the linkage did not find the relevant record. In addition, variables may take on invalid values or values that are incoherent when compared to other variables for the same record. The issues can be mitigated with the use of computer-assisted interviewing but even in this case are still usually present in the data, although to a lesser extent.

404. The first step in the cleaning process is to determine which sampled units provided responses that are sufficiently complete to be considered as respondents. Units that have not provided responses to key questions should be considered to be non-respondents. The number of questions considered crucial for this purpose should be small. When a survey is carried out at the household level, the response status should be defined at the household level, so that either the entire household is determined to be respondent or non-respondent. On the Canadian Income Survey, the entire household is deemed to be respondent if at least one adult household member answers the key questions on the survey. Unit non-response is treated by weighting as described in the next section, while item non-response is treated through imputation which is described further in this section.

405. The second step, editing, involves the correction of values for which it is evident that there has been some measurement error. Generally, very few corrections of this type are made because it is difficult to identify and correct measurement errors with certainty. One instance in which these types of errors can be corrected is when interviewers have left notes identifying problems. As another example, on monetary amounts, it may be possible to detect errors where a decimal point has been placed incorrectly. Extreme values and values that are inconsistent with other variables for the same sampling unit can also be treated at the editing step either by altering the values themselves or by deciding that the value should be imputed.

406. Though editing can be a helpful step and it can improve the overall data quality, it is strongly recommended to use it parsimoniously. On income variables whose distributions can be quite skewed, extreme values are to be expected and applying corrections should only be done rarely since it is likely to

introduce bias. It is also important to remember that collected data will always contain surprising relationships between variables. Survey practitioners should not aim to produce a dataset that is free of relationships that seem inconsistent at first view. Not only would they find the task near impossible, in trying to do so they would be likely to impose preconceived relationships on the data that may turn out to be incorrect.

407. Having cleaned the data, the next step is imputation - that means assigning values to replace erroneous values identified in editing or missing data. There are many imputation methods available. Two examples of imputation for income variables are given below, the first uses a regression-based approach and the second is donor imputation. Regardless of the method chosen, the same principles apply. It is important to find a method that maintains coherence between the variables for the record that is being imputed. At the same time, the process should be automated and objective, avoiding the application of preconceived models on the records to be imputed as manual intervention can. For example, on income surveys, the relationship between labour force status and the types of income received should be taken into account as part of the imputation strategy but not generally imposed as a hard rule. Additionally, imputation rates should be reported as part of the survey's quality report as an important complement to the response rates.

408. One example of an imputation procedure for income variables comes from the EU-SILC and based on the "sequential regression multivariate imputation" (SRMI). This flexible approach was implemented by the University of Michigan in the IVE-ware imputation software which can be used with SAS, STATA, SPSS and R packages or as a standalone in Windows, Linux or Mac OS operating systems. The method proposed by the authors of the software (Raghunathan et al, 2001) constructs the imputed values by fitting a sequence of regression models and drawing values from the corresponding predictive distribution, under the hypothesis of a Missing at Random (MAR) mechanism, infinite sample size and simple random sampling. The MAR mechanism is a standard assumption of imputation procedures and supposes that item non-response depends only on the covariates that have been observed and are used in the modelling.

409. The procedure is a variant of the estimation-maximisation (EM) algorithm and follows a Bayesian paradigm. The sequential multivariate model used is made for more complete imputation of the variables, while at the same time safeguarding their variance and their inter-correlation. A brief outline of the approach may be described as follows:

- Initially, the variables are divided into two types: auxiliary variables used to impute the others, and target variables which are the subject of the imputation. In the initial stages, the auxiliary variables are generally those relating to the demographic characteristics (sex, age) and to labour force characteristics.
- The auxiliary (exogenous) variables are supposed to be available for all cases. If not, some ad hoc procedures are used to perform the necessary imputations. The objective of this is not to impute 'final' values of these variables as such, but to provide a basis for their use in the imputation of the target (income) variables.
- The target variables are arranged in a sequence, starting with those with the smallest proportion of (or with no) missing values. (Alternatively, the ordering can be in terms of decreasing explanatory power of the variables.) Going down in sequence, each target variable is imputed

using all the variables above it, for which all information is available (or has been previously imputed), as auxiliary variables in the multivariate regression.

410. The form of regression depends on the nature of variable being imputed, such as a generalized linear regression for continuous variables (as in the case of income amounts) or a logistical regression for binary variables.

- Once a variable with missing values has been imputed, it is moved from the second set to the first, i.e. used as an auxiliary variable in imputation of the next variable in the list.
- After all variables in the list have been dealt with as above, the process is started again with the first variable in the target set, but this time using all the other variables as predictors, using for each the given or the most recently imputed value is used. The process is performed for each variable in turn, and is repeated iteratively.

Box 4.10

Imputation of Income in Albanian Survey of Income and Living Conditions

The Survey of Income and Living Conditions (SILC) was conducted for the first time in Albania in 2016. Several methods of data verification were implemented before defining the imputation method to be used for missing data. Problems observed include, for example, value incorrectly missing or reported as zero, outliers, and the benefit of the whole enterprise being reported as self-employment income. After these checks, different imputations methods were used

- *Logical/deductive imputation:* Deductive imputation was performed when, given specific values of other fields, and based on a logical or mathematical reasoning, a unique set of values exists causing the imputed record to satisfy all the edits (e.g. when items must sum up to a total and only one item in the sum has to be imputed, then its value is uniquely determined by the values of the other items).⁷⁴
- *Regression (used for imputed rent):*

To derive imputed rent, that is the amount a home owner would have to pay in rent for a similar dwelling as he owns, a regression was used. The regression was defined as $y_{ij} = \beta_0 + \beta_1 x_{ij} + \epsilon_{ij}$ using as auxiliary variables different characteristics of the head or the dwelling: age-group of the household head, number of children, number of adults, education level of the head, gender, household size, characteristics of the dwelling (ownership year of construction, number of rooms, square of space) and also prefecture (Tirana and others).

- *IVE ware imputations.*

For the purpose of imputing the non-response status a logistic regression was defined and used on IVE ware. The following variables were used: income of the previous year, gender, age group and the prefecture.

Also the IVE ware was used especially for imputation of the self-employment income. The first process is to define the missing data and examine the patterns of missing data, i.e. identify variables which are

⁷⁴ <https://ec.europa.eu/eurostat/documents/64157/4374310/30-Recommended+Practices-for-editing-and-imputation-in-cross-sectional-business-surveys-2008.pdf>

correlated with the propensity of missing values. These variables are used to impute the missing self-employment income with a linear regression model. The imputations are performed multiple times. A table with parameter estimates and variance information informs on the impact of the missing data among variables used in the regression model.

411. At Statistics Canada, the linkage, processing and imputation of income variables is done in the same way for each of the three surveys measuring household spending, income, and wealth including the Canadian Income Survey. This is an important factor contributing to producing a more coherent picture of economic well-being by Statistics Canada.

412. On the Canadian Income Survey, around 12% of respondents need to have their income variables imputed. The imputation is carried in phases. In the first phase, the various components of individual market income, such as wages, self-employment income, investment income, and pensions are imputed. Imputing all of these income components is done together to maintain coherence between the various types of income. The next phases impute income variables that are related to family structure, such as alimony and transfers between spouses, and takes into account the vector of variables imputed at the first phase. The final phase imputes income tax amounts, again taking into account variables imputed at earlier phases.

413. At each phase, the income variables are imputed by donor imputation using the Canadian Census Edit and Imputation System (CANCEIS), which is available to users outside of Statistics Canada. This imputation method consists of finding a record without any values needing to be imputed (donor) that resembles the record to be imputed (receiver) on range of auxiliary variables and, for the variables being imputed, using that donor's value for the receiver. The donor selected may be the record that is closest to the receiver (nearest-neighbour hot-deck) or may be selected randomly from among a set of potential donors each of which is close to the receiver (random hot-deck). Depending on the variable being imputed, the receiver may receive the donor's value directly or, when the relationship between variables must be maintained, the imputed value may be calculated by using a ratio to a related auxiliary variable that is available for both.

414. The auxiliary variables used to match donor and receivers include both categorical variables, such as age, sex, labour force status and family characteristics, and numerical values, such as income variables from the previous phases. It is well worth investing the time necessary to select the best auxiliary variables and to determine how much weight to give to each auxiliary variable when calculating the distance between the receiver and potential donors.

415. Additionally, special treatment is sometimes useful for income variables. For example, since income variables generally have a long tail of large values, they are often transformed so that matching between donor and receiver is done on the rank of units by the auxiliary variable rather than on the original values. As well, during the income tax imputation phase, the donor gives his tax rate rather than tax amount to the receiver and, based on this tax rate, the receiver's income tax to be imputed is then calculated. This is an example in which a ratio, the tax rate, is used as part of the imputation.

4.2.4 Weighting

416. The calculation of weights is usually performed by a step-by-step procedure. The main steps common to the production of weights for most surveys include:

1. The calculation of design weights;
2. An adjustment for non-response;
3. Calibration; and
4. The trimming of weight.

417. Implementing these steps can be more complicated for more complex designs. For instance, calculating weights for a cross-sectional survey or for the first wave of a longitudinal survey will generally be simpler than for subsequent waves of a longitudinal survey.

4.2.4.1 Design weights

418. In the first step, design weights are calculated on the basis of the sample design. These weights are of methodological interest since they are the starting point from which the weights will be created, but they are not meant to be used in substantive analysis. A design weight is assigned to all sampled units, not just responding units, and is defined when the survey sample is first selected. This weight will be based on the sampling unit.

419. For example, when a sample of households (or of addresses or other units containing households) is selected, the household design weight, for household i , is computed as

$$\omega^{(HD)} = \frac{1}{\text{probability of selection of the household}}$$

420. The probability of selection is based on the design of the survey and reflects design features such as stratification and multi-stage selection procedures.

4.2.4.2 Non-response adjustment

421. The next step of the weighting process is the non-response adjustment. At this step, the weight of non-responding units is redistributed to responding units. Non-response adjustment procedures aim to redistribute the weight of non-respondents to responding units that have a similar response propensity as this can help minimize the impact of non-response bias. For poverty surveys, the concern being addressed by the non-response adjustment is generally non-response at the household interview stage. Non-response by individuals in the household is often addressed by imputation as described in the previous section.

422. The problem of (unit) non-response can be particularly problematic in some household surveys in some countries; and it occurs for both cross-sectional surveys and longitudinal or panel surveys (in which case it is referred to as *attrition*). In a longitudinal survey, non-response compounds over the waves, with non-respondents at the first wave usually being excluded from subsequent waves and so on. Good,

efficient procedures to re-weight the responding cases are therefore a critical requirement at the first wave. However, the possibilities for non-response adjustment in cross-sectional surveys and at the first wave of longitudinal surveys are often constrained by lack of information since the non-response adjustment has to be based on characteristics which are known for both responding and non-responding households. For the later waves of a longitudinal survey, many variables are available for the non-response adjustment since the first wave data can be used.

423. The non-response adjustment procedure involves estimating response rates or propensities as functions of characteristics available for responding and non-responding households. This includes the use of characteristics of the areas where the households are located. This is also true when a sample of persons has been used. The main difference is that for samples of persons the characteristics of interest that can be used for the adjustment are not only those of households, but also (and perhaps more importantly), personal characteristics of the selected individuals.

424. Generally, it can be useful to apply the adjustment in two steps:

- (i) For non-contact (of households and/or of selected individuals); and
- (ii) For non-response, once a contact with the households or the person concerned has been made.

425. For both steps, especially for (i), area-level characteristics often provide a main part of the auxiliary variables explaining non-response. This is because they are the more easily available variables for both responding and non-responding units.

426. In dealing with the effect of non-response, it is of crucial importance to identify responding and non-responding units correctly. In this context, a “respondent” is not just a collection status, rather it is a unit whose interview is accepted after processing and will be used for estimation. In practice, determining which units are respondents and non-respondents can be complicated because the frames from which units are selected are generally not perfect. Continuing with the household sample example, an address frame will often contain units that do not correspond to a household. This can be because the address is non-existing, corresponds to an unoccupied structure, or is a business rather than a private dwelling. These selected units which turn out to be non-eligible or non-existent must be excluded and not counted as non-responding. Imputation has to be done for units with unknown status, i.e. when it is not clear whether they are non-eligible or non-respondents. Every unit has to be assigned uniquely to one category or the other.

427. Substitutions inevitably reduce control over the survey process and can diminish quality. In surveys where substitution has nonetheless been allowed, such procedures must follow strict rules, such as making best possible use of auxiliary information. In such situations non-responding original units for which successful substitutions have been made are to be considered as ‘responding units’ for the purpose of determining non-response weights.

428. Having done this preliminary step, there are two commonly used procedures for non-response weighting. The first is to modify the design weights by a factor inversely proportional to the response rate within each “weighting cells” (appropriately determined grouping of units). It is common to use sampling strata or other partitions, sometimes geographical, as weighting cells. These classes can also be defined

using classification trees, though this is more relevant if many variables are available as is the case after the first wave of a longitudinal survey. The non-response adjusted weight is:

$$\omega_i^{(HN)} = \frac{\omega_i^{(HD)}}{R_K}$$

where R_K is the response rate in weighting class K , the non-response class to which the household belongs, and is computed with data weighted by the design weights:

$$R_K = \frac{\text{sum of design weights of responding units in cell } K}{\text{sum of design weights of selected units in cell } K}.$$

429. Numerous, very small weighting cells can result in a large variation in R_K values and should be avoided. On the other hand, if only a few broad classes are used, little variation in the response rates across the sample may be captured, making the whole re-weighting process ineffective. On practical ground, cells of average size 100-300 units may be recommended. These cells must also include enough respondents so that the factor applied is not too large. The appropriate maximum factor will depend on the survey's overall response rate. No absolute rule exists but it may be useful to use weighting cells for which the adjustment factor is no more than twice the average adjustment factor, for instance. In other words, if the overall weighted response rate for the survey is 80%, the average adjustment factor would be $1/0.80 = 1.25$ and so the weighting cells used could be defined so that no adjustment over 2.5 would be applied in any of the cells.

430. The other alternative is to use a regression-based approach. Using an appropriate model such as logit regression, response propensities can be estimated as a function of auxiliary variables, which are available for both responding and non-responding cases. When many auxiliary variables are available, this approach is often preferable to simply using sampling strata or a geographic partition.

431. In principle, each responding unit weight is adjusted by the inverse of the estimated response propensity. However, a very important point when using the regression approach is to ensure that weights assigned are confined to be within reasonable limits. This is the case for all non-response adjustments, no matter the modelling approach used. In the case of a regression-based approach, the regression can predict zero or even negative values, which of course must be rejected. The problem is more general than that since extreme values should also not be permitted. To deal with this, a best practice is to classify the units into response homogeneous groups (RHGs) based on the response propensity estimated using the regression. These are defined to be groups of units clustered together because they have similar response propensity. Once these classes are defined, the adjustment can then proceed as in the first method, within the cells:

$$\omega_i^{(HN)} = \frac{\omega_i^{(HD)}}{R_{RHG}}$$

where R_{RHG} is the response rate in the RHG to which unit i belongs. This is known as the score method (Little, 1986; Eltinge and Yanseneh, 1997).

432. Regardless of the non-response adjustment method selected, the choice of variables is fundamental. In fact, the choice of variables will generally have more of an impact on the effectiveness of

the non-response adjustment than the method used. The adjustment will only reduce non-response bias if the variables are related to both the response rate and the estimates being produced. Therefore, variables with a link to income, poverty, or to the variables that will be used to disaggregate the statistics should be prioritized, on the condition that they are also related to non-response of course. When too many variables are included in the model or when the variables are not related to both non-response and the statistics of interest, the non-response adjustment can have the negative effect of increasing the variance of the estimates.

4.2.4.3 Calibration

433. Calibration is a method that adjusts the weights assigned to sample units (individual or household) in order to satisfy (or approximately satisfy) some pre-determined constraints. These are typically based on Census data or other large surveys. The key idea is that estimates formed from the weighted sample should replicate the known values from other sources. The critical requirement in calibration is to ensure that the external control variables are strictly comparable to the corresponding survey variables, the distribution of which is being adjusted.

434. Calibration is used for multiple reasons. The first is to produce results that are coherent with other related surveys and with the Census data that is available. In addition to this, calibration can also improve the accuracy of estimates in two ways. Calibration can serve as a non-response adjustment and can stabilize estimates reducing the variance of the estimates. As a non-response adjustment, calibration is particularly relevant if a control total is available for a variable that was not available for both respondents and non-respondents during the non-response adjustment step and that can explain (at least in part) the observed survey non-response. For income statistics, one example of such a source can be administrative tax data that can give the distribution of some income components for the whole population. Calibrating to match these distributions more closely can be helpful as a way of adjusting for non-response.

435. In household surveys concerning poverty, income and social exclusion, where the household is the sampling unit and both the household and the individual are used as units of analysis, the so-called “integrated” calibration is recommended (Lemaître and Dufour, 1987). This is a calibration which retains the same weights for all members of the same household; characteristics of households and of the total population are controlled. When a separate personal interview sample exists, a further adjustment can be applied to the personal interview sample.

436. Mathematically, calibration is an optimization problem. The goal is to find weights,

$$\omega_i^{(HC)} = \omega_i^{(HN)} g_i ,$$

as close as possible to the non-response adjusted weights that respect the chosen calibration constraints. Different choices of distance functions will result in different estimators. For example, post-stratification and raking ratio estimation can be expressed as special cases of calibration. Numerous software packages have been built to implement calibration. Examples include CALMAR which is widely used in the EU and Statistics Canada’s G-Est. The Slovak Statistical Office provides a freeware Graphical User Interface for the

R-software which makes it extremely easy to implement calibration even without expert knowledge of the R-software and independently from commercial software.⁷⁵

437. There are important conditions to be respected when implementing calibration. First of all, it is important to be selective in terms of the calibration constraints. If the constraints are not related to the statistics to be produced by the survey or to the domains for which they are produced, the calibration can increase the variance of the estimates without a gain in accuracy. Along the same lines, when too many calibration constraints are used, this can also increase the variance rather than stabilize the estimates.

438. The main symptoms of excessive calibration include non-convergence (i.e. no solution being found that satisfies all constraints), the presence of negative weights, and weights that are very close to 0 or that are very large. The calibration factors (ratio of the calibrated weight divided by the pre-calibration/non-response adjusted weight, indicated as g_i above and often referred to as g-factors) should be neither too small nor too large. It therefore a good practice to apply bounds to the calibration adjustment factors. While there is no specific rule for which range of calibration adjustment factors is acceptable, trying to keep these factors between 0.3 and 3 (or between 0.3 and 3 times the average calibration factor) can be preferable. In this range, it is most important to respect the upper bound for the calibration factors.

Box 4.11

Calibration in the Canadian Income Survey

This case study briefly describes the calibration totals used for the Canadian Income Survey. This strategy uses two main sources of control totals: demographic projections derived from Census estimates and administrative tax files providing wage and salary information for all paid employees in Canada. The demography totals used include:

Table 4.4
Demographic control totals used by the Canadian Income Survey

Counts of individuals in 15 sex × age groups	$\left. \begin{matrix} \text{male} \\ \text{female} \end{matrix} \right\} \times \begin{cases} 0 - 6 \text{ (both sexes)} \\ 7 - 17 \\ 18 - 24 \\ 25 - 34 \\ 35 - 44 \\ 45 - 54 \\ 55 - 64 \\ 65 + \end{cases}$
Number of households by household size	sizes 1, 2 and 3+
Number of economic families by family size	sizes 1 and 2
Counts of individuals in 6 Canadian cities	Montreal, Toronto, Winnipeg, Calgary, Edmonton, Vancouver

⁷⁵ <https://github.com/SO-SR/Calif>

The administrative tax data on wages and salaries is used in the following way:

- An administrative tax file provides wage and salary amounts for all paid employees in Canada, not just those who file their taxes. Using this file, the 10th, 25th, 50th, 65th, and 75th percentiles of the wages and salaries amount is calculated, as is the number of employees in each of the six classes defined by these cut-off points (0th – 10th percentile, 10th – 25th percentile, 25th – 50th percentile, 50th – 65th percentile, 65th – 75th percentile, and 75th – 100th percentile).
- On the survey data, a new variable is derived indicating to which of these six classes each survey respondent who has received wages and salaries belongs. If the respondent did not receive wages and salaries, they are put in a seventh class of non-wage earners.
- Six calibration totals are used corresponding to the number of employees in each of the six wage and salary classes.

Using these wages and salaries control totals along with the demographic ones, calibration is carried out separately for each of the ten Canadian provinces.

The wage and salary counts are a particularly important part of the Canadian Income Survey calibration strategy. They provide a way to ensure that the distribution of wages and salaries after calibration matches the distribution coming from administrative tax data. In practice, their effect is to help compensate for higher non-response rates at the top and bottom of the income distribution that cannot be sufficiently corrected during the non-response adjustment.

Along with using a common imputation strategy for income variables, Statistics Canada's surveys on household spending, income, and wealth all use a similar calibration strategy. This has helped make them much more coherent with each other.

4.2.4.4 Treatment of extreme or influential weights

439. Trimming or winsorisation refers to recoding of extreme weights to more acceptable values. The objective of trimming is to avoid excessive increase in variance due to weighting (the so-called Kish effect). It is important to realise that the process will introduce some bias. Even so, the aim is to seek a procedure which reduces the mean squared error. Though treatment of extreme or overly influential weights introduces some bias, the overall error may still be reduced.

440. At each step of the weighting procedure, the distribution of the weight adjustments and the weights should be checked. In principle, the results of every step can be subject to the trimming procedure. This applies to weight adjustments for non-response and calibration as well but, of course, if the adjustment factors are already limited by the non-response or calibrations strategy, this step may not need to be repeated separately.

441. It can also be useful to reduce the weights of units that are influential for certain important variables, such as key income components. Even if the weight of a unit on its own is not too large, the product of the weight and the value of the variable together may make it influential. This can be dealt with by adjusting the value of the variable in question or the value of the weight. When the value of the variable is reasonable but large or if there are relationships between the variables that must be maintained, it can be more practical adjust to the weight.

442. There is no rigorous or absolute procedure for general use for determining the limits for trimming or winsorising and it is very important to use it parsimoniously because it does introduce bias. While sophisticated approaches are possible, it is generally desirable to have a simple and practical approach.

443. The following approach, given as an example for the non-response adjustment, may be quite adequate for the purpose if the permitted limits are wide enough. Where

- $\omega_i^{(HD)}$ is the household design weight;
- $\omega_i^{(HN)}$ is the household weight after non-response adjustment;
- $\bar{\omega}^{(HD)}, \bar{\omega}^{(HN)}$ are their respective mean values; and
- any computed non-response weights outside the following limits are recoded to the boundary of these limits: $1/C \leq \frac{\omega_i^{(HN)} / \bar{\omega}^{(HN)}}{\omega_i^{(HD)} / \bar{\omega}^{(HD)}} \leq C$.

444. A reasonable value for the parameter is $C=3$. A similar process could be used with calibrated weights.

445. As a second example, a unit can be determined to be influential for a statistic in a cell if removing that unit changes the estimate of the statistics in the cell by more than a predetermined percentage. For example, if removing one unit changes the average of wages in a particular age group by more than 10%, it could be deemed influential. The appropriate threshold depends on how common the variable is in the population. As a weight adjustment, units that are influential can have their weight reduced to the point of no longer being influential and have the amount by which their weight was reduced redistributed to other units in the same domain.

446. Since trimming alters the mean and total value of the weights, these types of adjustments may need to be applied iteratively, with the mean re-determined after each cycle. It may also be necessary to iteratively repeat this step with the calibration. In both cases, a very small number of cycles normally suffice.

447. The most important factor to remember when reducing the weight of certain units is to do so very parsimoniously since it will introduce bias. It should only be done for a small number of units that are particularly extreme.

Box 4.12

Calibration in Household Living Condition Survey 2009 in Ukraine

This case study briefly describes the calibration process performed by State Statistics Service of Ukraine for the 2009 round of the Household Living Condition Survey (HLCS). The sample design for this survey consists of a stratified multistage probability sample design with a three-stage sampling procedure for urban area and a two-stage sampling procedure for rural area. The procedures for calculation of final weights for the 11,182 interviewed households were implemented under the generally established steps: i) calculation of design weights; ii) adjustment of design weights for unit non-response; and iii) calibration of weights to external sources. The population characteristic variables used for calibration were quite numerous and disaggregated over 27 geographical regions (25 Oblast plus cities of Kiev and

Sevastopol): the household size; the presence of children in the household; and the number of men and women in the household, classified into four age groups each. The software used at the State Statistics Service of Ukraine, the SPSS g-calib (Statistics Belgium, 2002; Vanderhoeft, 2002), was therefore not able to reach convergence without producing some negative weights and a semi-automatic procedure was used to make these negative weights positive. This procedure leads to a weighting system with the following statistical characteristics:

- Mean weight: 1528.96 (for a total population of 17,096,871 households);
- Standard deviation: 790.23;
- Coefficient of variation: 0.5168;
- 5 minimum value weights: 1.44196, 9.06557, 13.48944, 14.48621, 18.19603; and
- 5 maximum value weights: 7733.00, 7799.99, 7864.44, 8807.92, 9175.46.

Such weights have clearly been calibrated too much, with too many constraints imposed. Although they permit the sample statistics to be unbiased, they introduce extra variability (instability) in such statistics. According to Kish (Kish, 1992), the increase of variance of a generic statistics y , is given by $1 + cv^2$, where cv is the coefficient of variation of weights. In this case the Kish weight effect is equal to 1.267, which means an increase of the variance of about 27% compared to a sample with equal weights. Verma and Betti (2011) show that the Kish effect of weights may depend on the statistic under observation, and that such an effect is multiplicative with the overall design effect; they recommend an optimal ratio of about 10 between the highest and the lowest weights in the sample (here we can observe a ratio of 6363).

4.2.5 Variance estimation

448. Among the various types of survey errors, sampling error is unique in that it does not need an external source serving as 'gold standard' in order to be measured; it can be estimated based on the sample design. Moreover, as noted in the first part of this chapter, the sampling variance is often the largest component of error for disaggregated statistics for a domain having a small sample size. Being able to estimate the sampling variance is the basis on which inference can be made in design-based approaches to sample surveys. It is therefore very important to have practical procedures for estimating sampling variance. This section will outline options that are often used for social surveys, including those on income and poverty, while the next section discusses how to provide these estimates to the data users.

449. Practical procedures for estimating sampling errors for social surveys:

- a. Must take into account the actual, complex structure of the design;
- b. Should be flexible enough to be applicable to diverse designs;
- c. Should be suitable and convenient for large-scale application, producing results routinely for diverse statistics and subclasses;
- d. Should be robust against departure of the actual sample design from the ideal model assumed in the computation method;

- e. Should have desirable statistical properties such as small mean squared error of the variance estimator;
- f. Should be economical in terms of effort and cost; and
- g. Suitable computer software should be available for application of the method (Verma, 1991).

450. Linearization methods and replication methods are two broad practical approaches to the computation of sampling errors. These methods are needed when the explicit variance cannot be derived, for instance because the design is too complex. A major advantage of replication methods is that they do not require an explicit expression for the variance of each particular statistic, and hence can more easily handle complex statistics and designs, including multi-wave and longitudinal situations. As a result, replication methods are more commonly used on social surveys that generally use complex designs. Under these methods, the variance is estimated by

- Taking repeated subsamples, or replicates, from the data, each of which reflect the structure of the full sample;
- Re-computing the weighted survey estimates for each replicate and for the full sample; and
- Estimating the variance as a function of the resulting estimates.

451. Examples of replication methods include the bootstrap, Jackknife, and Balanced Repeated Replication (BRR).

452. The variance estimates should also take into account the effect on the variance of aspects of the estimation process by repeating these steps on each of the replicates. In principle, this can include complex effects such as those of imputation and various steps of weighting, though often full repetition of these procedures for each replication is not feasible and practical.

453. Step by step, this means first creating the replicates, taking into account the sample design. The way in which this is done depends on the replication method used. Subsequently, each step of weighting (and imputation if feasible) is redone on each replicate. At the non-response adjustment stage, ideally the entire modelling of non-response is redone for each replicate but, when this is not practical, the same non-response adjustment factor that was applied to the original sample can be applied to each replicate instead. Similarly, for influential values, while it is ideal to re-identify which values are excessively influential on each replicate individually, it is generally not feasible to do so. In this case, the records whose weights were modified because of influential values on the original sample can have the same modification applied to them in each of the replicates. For calibration on the other hand, it is very important to repeat the calibration on each replicate individually since it has such a large impact on the final variability of the survey estimates. In other words, each individual replicate should be calibrated individually to the same control totals as are used for the whole sample.

454. Once the replicates are created and have each gone through the weight adjustment process, estimating the variance can be done in most statistical software packages, such as SAS or Stata. Though the bootstrap is not always supported explicitly, the bootstrap variance can be calculated using the BRR functionality when it exists. More information about replication methods can be found in Wolter (Wolter, 2007), Rust and Rao (Rust and Rao, 1996) or Lohr (Lohr, 1999).

Box 4.13

Calculating variance for EU-SILC in Austria

Statistics Austria has developed its own tool for variance calculation. A package called `surveysd`⁷⁶ is freely available for the open source software R. It was developed specifically to take into account the EU-SILC overlapping sample structure which affects the variance properties of estimators when data is pooled over several years (see 4.3). The package has three basic elements which are described in a hands-on manner in the accompanying documentation⁷⁷.

The first function creates any desired number of bootstrap replicates which consider the sampling design. Each sampling unit receives a slightly altered selection weight. If sampling units are followed over time it can be specifically requested that the longitudinal structure is preserved, having individual sampling units carry their weights as long as they are part of the sample.

The second function specifies controls for calibration and adjusts the original replicate weights accordingly. These controls should be the same as used in the actual survey. They ensure that each sample replicate does indeed represent the same population. This ensures also that the variance estimation will consider the impact of calibration on variance. Depending on the characteristics and the control totals which are used this effect will usually imply a gain in precision over the uncalibrated estimates. This point was considered particularly important for Statistics Austria to be able to demonstrate how regional precision requirements can be met (as stipulated in Annex II of regulation (EU) 2019/1700).

The third element does estimations over the previously defined and calibrated replicates. At Statistics Austria this function is used in combination with a tool which automatically suppresses tables according to customizable filtering and flagging rules to ensure that only reliable estimates will be published.

4.2.6 Dissemination

455. Disseminating disaggregated poverty statistics from survey data can be a challenge. Domains of interest from the analytical and policy perspective may be small, especially when disaggregation variables are considered together. Particular attention must therefore be paid to the accuracy of estimates before they are disseminated. Estimates will have lower precision when the sample contains only a small number of units in the domain of interest. Though the relevance of poverty statistics clearly is increased by producing disaggregated indicators of economic well-being, these statistics are only useful if they are sufficiently accurate for their intended use.

456. As outlined in the first part of this chapter, the guiding principle in evaluating the quality of statistical estimates is fitness-for-use. It is, of course, impossible to anticipate all eventual uses of a survey's data before its publication. Moreover, the accuracy of survey estimates that is required varies by use. It is particularly important to inform users of the quality of estimates and the elements that affect the quality of the survey or surveys from which disaggregated poverty statistics are produced, so that they may determine whether the data is fit for their intended use. Earlier chapters give recommendations on

⁷⁶ <https://github.com/statistikat/surveysd>

⁷⁷ <https://statistikat.github.io/surveysd/articles/surveysd.html>

which statistics are most useful for making international comparisons and by which variables the statistics should be disaggregated. In this section, we outline best practices for what complementary information should be provided to the data users. The principal recommendation is that users should be informed of the quality of the estimates and of the quality of the survey more generally.

457. The published tables should contain not only the point estimates but also an indication of the accuracy of the estimates. Most often, a measure of the sampling variability is used as a measure of the exactitude of the estimates. There are many ways in which to present sampling variability. The standard error, the coefficient of variation (CV) or relative standard error (RSE), and confidence intervals are all possibilities. No matter which indicator is used, it is a best practice to make the indicator available in the same table as the point estimate to which it corresponds, in order to make it easily accessible to users.

458. The most commonly used measure in dissemination tables is generally the CV which, by definition, is the ratio between the standard error and the mean of the variable of interest. The CV is particularly useful for comparing the precision of two estimates that are on different scales, such as different types of income or income from different countries. One common communication strategy for dissemination tables is to classify the quality of the estimates based on CV scale. For example, estimates from the Australian Bureau of Statistics' Survey of Income and Housing are annotated by an asterisk (*) when the CV is between 25% and 50% to indicate that the estimate should be used with caution. Estimates with a CV greater than 50% are annotated with a double asterisk (**) to indicate that the estimates are considered too unreliable for general use and should only be used to aggregate with other estimates to provide derived estimates with RSEs of 50% or less (Australian Bureau of Statistics, 2019).

459. Unfortunately, the CV is not as useful for estimates of proportions, of change or differences, and of statistics that can take on negative values and these are all common type of statistics when analysing poverty. Measures of poverty (such as the At-risk-of-poverty (AROP) for example) often take the form of the proportion of a group whose income is below a threshold. In general CVs tend to understate the quality of estimates of small proportions and overstate the quality of large proportions.

460. Confidence intervals, on the other hand, are appropriate for all types of estimates and have the advantage of being easier to interpret than the CV. Including confidence intervals in the same table as the point estimates is an excellent practice. When the sample is sufficiently large for a central limit theorem to apply, a symmetric confidence interval around the mean of a variable whose width is based on the sampling error, that is $[\bar{y} - z_{\alpha}SE(\bar{y}), \bar{y} + z_{\alpha}SE(\bar{y})]$, may be appropriate. However, this may not be the case for small proportions especially if they are based on a small sample, which is often the type of variables that is desired when disaggregating poverty measures. In this case, alternate methods such as a bootstrap confidence interval or Wilson's method could be used.

461. As mentioned above, some estimates may not be reliable enough to be published. Criteria that can be used to determine which estimates to suppress are always subjective. In addition to release criteria based on the CV as mentioned above, it can be a good idea to suppress cells of a table that are based on too few records to be reliable. For estimates based on a sample that is too small, not only may the estimate be less precise than desired but the estimate of the variance may also be imprecise making it difficult to reliably inform users of the quality of the estimate. As an example, on the Canadian Income Survey, estimates that are derived from fewer than 25 records are suppressed even though the value of the CV

for those estimates might seem acceptable. Here the number of records is either the number of individuals or the number of families depending on the statistic. For some estimates special rules may be applied.

462. EU-regulations on EU-SILC require that the European Commission shall not publish an estimate if it is based on fewer than 20 sample observations, or if non-response for the item concerned exceeds 50%. The data shall be published by the Commission with a flag if the estimate is based on 20 to 49 sample observations, or if non-response for the item concerned exceeds 20% and is lower than or equal to 50%. The data shall be published by the Commission in the normal way when based on 50 or more sample observations and the item non-response does not exceed 20% (European Commission, 2003). Following these guidelines, Statistics Austria puts numbers which are based on less than 50 observations in brackets in its publications from the EU-SILC survey and uses a hyphen (“-”) for cells in tables which contain fewer than 20 observations.

Box 4.14

Suppression rules in the United States

In the United States, one-year estimates from the American Community Survey (ACS) are published for an extensive set of tables for any geography or group with a population of 65,000 or more. Until the series was discontinued due to budgetary reasons, three-year estimates had also been produced for any group/geography of 20,000 or more units. Instead a streamlined set of "supplementary" tables using one-year data for these smaller geographies (>20,000 but < 65,000) is produced regularly. Every other geography/group gets 5-year ACS estimates. The smallest geography published in tables is Census Block Groups which typically have a population of 600 to 3,000 people.

In addition, tables with one-year estimates are limited to those that pass two data quality filtering rules. CVs are calculated for each table's estimates, and the median CV value is determined. If the estimate is 0, a CV of 100% is assigned. First, if more than half of the estimates in the table are not statistically different from 0 (at a 90% confidence level), then the table is not published. Second, if the median CV value for the table is greater than 61 percent, the table is not published. Whenever a table fails these rules, a simpler table that collapses some of the detailed lines together can be substituted for the original. If the simpler table passes, it is released. If it fails, none of the estimates for that table and geographic area are released. These release rules are applied to single year estimates, but are not applied to the 5-year estimates. Tables with 5-year estimates are subject to some disclosure avoidance suppressions when cell sizes are small.

Public use micro data show only public use microdata areas (PUMAs) - which tend to have populations of approximately 100,000 or more.

Public use micro data show only public use microdata areas (PUMAs) - which tend to have populations of approximately 100,000 or more.

463. Depending on the sample design, it can be useful to consider the design effect to determine what is too small of a sample size for dissemination. Though suppression for confidentiality reasons is not addressed here, suppressing cells based on less than a predetermined number of records can be part of a strategy of disclosure control for confidentiality reasons as well. Since the sampling variability is not the only factor influencing the accuracy of survey estimates, additional information of the survey should also

be made readily available to survey users. This additional information should include information on the survey methodology with a focus on aspects that affect the accuracy of estimates. For example, the Australian Bureau of Statistics presents this material in the User Guide for its Survey of Income and Housing (Australian Bureau of Statistics, 2019).

464. Even if poverty statistics that are sufficiently disaggregated are disseminated, these statistics may not address all questions. Making microdata sources available to users can be a complementary way of increasing the usefulness of surveys that can be used to measure poverty. The confidentiality of survey respondents is the most important concern to be addressed for microdata sources. Beyond the necessary first step of removing all personal identifiers, such as name, address and other contact information, there are two main ways of protecting the confidentiality of respondents. The first is to restrict access to the microdata set to certain individuals in a controlled setting. The second is to prepare a confidentialised microdata file that has been treated to protect privacy and confidentiality using a variety of techniques while preserving the variables of interest to the degree possible.

465. Statistics Canada uses both of these strategies. The Canadian Income Survey data is available to researchers at Research Data Centres across the country. At these centres, researchers must have their research projects reviewed and take an oath not to disclose individual information. These researchers are asked to follow the same suppression rule for small sample sizes as used on Statistics Canada tables and their final tables are reviewed to ensure that there is no breach of confidentiality. Additionally, a Public-Use Microdata File (PUMF) is prepared based on the CIS. This file can be shared more broadly and used outside of the Research Data Centres (Statistics Canada, 2019).

466. Outreach activities may support the relevance and quality of the data. For example, Statistics Austria has successfully held several user conferences to stimulate research and obtain feedback on EU-SILC micro data. On a European level, GESIS has established regular user meetings for EU-SILC microdata and other official data sets which bring together expertise from leading academic researchers and data producers.⁷⁸

4.3 Other methodological issues related to measuring poverty

467. This section treats some further aspects that are not always part of the survey process, such as pooling, small area estimation, and rapid estimates.

4.3.1 Pooling

468. For repeated surveys, where there is not enough data to produce reliable enough estimates from a single iteration of a survey, pooling of data may be used. Pooling consists of using data from multiple iterations of the survey together to produce estimates based on a larger sample. The objective of doing

⁷⁸ <https://www.gesis.org/en/services/events/gesis-conferences/european-user-conference-6>

so is to increase the sample size in order to decrease the sampling variance, thereby producing estimates that are more precise. In the context of a survey on economic well-being, pooling is generally done over multiple iterations of a national survey in order to increase the precision of regional estimates. These disaggregated statistics can be of particular interest since the dispersion of regional estimates is a measure of disparity within a country.

469. Pooled estimates are generally obtained by averaging estimates derived independently from each survey iteration or wave. In other words, pooled estimates are relatively simple to calculate by working at a purely macro level. This is generally the preferred approach over concatenating the micro-data sets for the waves being pooled together (Verma et al, 2013). If micro data is used, weights need to be adjusted accordingly (dividing the final weight for every unit by the number of waves being pooled together). When multiple waves of a longitudinal survey are pooled together, this corresponds to keeping multiple lines in the final data set for units that are found in more than one wave, one line for each time a unit is in a wave. Despite the fact that point estimates from pooled data can be calculated at the macro level, it is important to remember that measuring the sampling variance requires access to the micro data.

470. Pooling can be done for both longitudinal and repeated cross-sectional surveys. Regardless of the type of survey, the concerns that must be addressed when pooling data fall into two main categories. The first has to do with the concepts being measured and the second with how to estimate the reduction in sampling variance achieved by pooling.

471. Conceptually, pooling muddies the waters when it comes to specifying what is being measured. For example, income and consumption are central concepts when measuring poverty. However, both can be expected to evolve from one year to the next due to inflation. As a result, income and consumption from one year may not be directly comparable to income and consumption from the next year and, while it may be natural to want to apply an inflation factor to make them more comparable, it is not always obvious what factor to use or even whether to use one factor globally or different factors for different components. Taking this reasoning one step further, variables derived from income, such as poverty status (below a poverty threshold or not), are also affected. As a result, when pooling data, a respondent's poverty status should be done separately for each iteration of the survey.

472. Another way in which conceptual differences between what is measured at different iterations of a survey can occur is through changes in how it is being measured. In a questionnaire, if the way in which a question is worded or if the categorization of responses changes, it may be necessary to harmonize the concepts from different years before they can be pooled together. Similarly, if administrative data is used and the programs on which the data are based change, then there can again be conceptual differences that need to be reconciled.

473. Finally, the population that is being measured also evolves over time. This is another source of conceptual differences. For this aspect, it is important to specify what population the pooled data represents, in particular what year or years the population corresponds to. If desired, recalibration to population control totals for the most recent year can be used to make the years being pooled more comparable.

474. The second major concern is how to estimate the sampling variance of estimates coming from pooled data. The design of a pooled sample is more complex than that of the original samples, which may already use a complex survey design themselves. Moreover, the variance of the pooled sample depends not only on the design of the original samples but also on the dependencies between these samples. In particular, when multiple waves of a longitudinal survey are pooled together, it is essential to take into account the overlap between the samples that is inherent in the longitudinal survey design. Apart from correlations at the individual level, we also have to deal with additional correlation that arises because of the structure (stratification and clustering) of the samples from different iterations of the survey. Such correlation can exist even in cross-sectional surveys that have sample coming from the same clusters even if there is no overlap in terms of individual households.

475. Using a replication method for the estimation of variance, as introduced in Section 4.2.5, can make the process of variance estimation more straightforward for pooled samples. For this purpose, the replicates should be chosen in a coordinated fashion that takes into account the dependencies for the whole pooled sample. In other words, for a longitudinal survey, if a unit is part of multiple waves, either the unit is part of a replicate for all of its waves, with the same multiplicity when applicable, or it is in the replicate for none of its waves. Similarly, for a cross-sectional survey in which the same cluster is selected for multiple survey years being pooled together, the inclusion of that cluster in the replicates should be determined for all survey years together.

476. It should be expected that the reduction of sampling variance from pooling multiple years together will be more substantial for cross-sectional surveys that are based on independent samples for each year than for longitudinal surveys and surveys with a rotating panel design in which the panels are highly correlated.

477. Verma et al (Verma et al, 2013), have implemented pooling for Austria and Spain for 2011 EU-SILC data. In this case, single year estimates for 2011 are compared with the average over three years (2009-2010-2011) and averaging the poverty rate over three waves leads to a variance of this averaged estimator that is 30% less than the variance of the AROP rate estimated from just a single wave for Austria and 35% less for Spain. Similar reduction is also found at regional level (NUTS2).

4.3.2 Small Area Estimation

478. Small area estimation (SAE) refers to a family of techniques that combine a direct survey estimator with a model-based estimator. It is used where direct survey estimates for small domains are not precise enough. The model part is based on the whole sample, not just the part of the sample in the domain of interest, and, since the model borrows strength from the whole sample, the variance of the resulting combined estimator can be smaller than that of the direct estimator alone. Since the goal of SAE is to produce estimates for small domains, it is particularly relevant to the production of disaggregated statistics. The standard reference on small area estimation methodology has recently seen a considerable updated (Rao and Molina 2015). See also, among others, Gosh and Rao (1994) and Henderson (1950).

479. Today, there is a wide variety of small area estimation (SAE) techniques available, and the field is rapidly expanding. In general, the suitability and efficiency of a particular technique depends on the specific situation and on the nature of the statistical data available for the purpose. It appears useful to distinguish unit level and area level approaches.

480. Area level models, such as Fay and Herriot (Fay and Herriot, 1979) and Torabi and Rao's (Torabi and Rao, 2014) two-fold subarea level models can be applied so that poverty estimates (and other indicators) may be obtained as long as there is a regular data collection to measure poverty. These methods do not require Census data and can be executed with administrative and satellite data which may be available more frequently.

481. The advantage presented by unit level models such as Molina and Rao's (Molina and Rao, 2010) EBP, or the World Bank's methods is that these make use of census data which lends more information to the estimates and thus yields more precise estimates. A good starting point on this work can be seen in the report of the World Bank on pinpointing poverty in Europe, which was a part of a large trust fund which aimed to produce poverty estimates at the NUTS3 level for all EU countries for 2011.

482. The applicability of SAE methods is generally very country-specific since it depends greatly on the data available in each country (see Box 4.15 for an example how SAE are implemented in the United States). The most important factor influencing how successfully SAE can be used is the availability of auxiliary information on which the models can be based. Censuses and administrative data are common sources of auxiliary information since they can provide sufficiently detailed information for the small domains that are of interest. The knowledge and experience of national statisticians and other researchers about the specific possibilities and limitations in their own country can therefore be very valuable when determining whether SAE is likely to work.

483. An application of SAE based on the EU-SILC (see Betti et al, 2012) highlights the merits of the procedure but also some serious limitations to the application of SAE methodology for the production of regional estimation in EU-SILC. As merits, it is worth noting that its SAE methods make use of external data aggregated to NUTS2 (area level) only and that these estimates could be performed every year, given that such external sources are available. Although NUTS2 level indicators are often published, their precision is not always satisfactory due to vastly varying sample sizes and poverty mapping made it possible to obtain estimates at NUTS2 level with better precision.

484. On the other hand, four types of limitations had to be faced:

- a) The first concern is the lack of external data for the purpose of using SAE. The methodology used information from census data, which are only available every ten years in many countries. The poverty mapping model, for instance, is mainly used for consumption data, although some applications with income data have been successful. In any case, often such external sources are not sufficiently correlated to the poverty measures under investigation. Also, most of the models assume the external data to be error-free, which is certainly not the case when the data come from other large-scale field studies and surveys.
- b) The methodology tends to be complex and to require specialized knowledge and software.

- c) Another major concern in application to a multi-country undertaking such as EU-SILC is that the results may lack comparability. Generally, the procedures and application would have to be country-specific, and ensuring the application of common standards required for EU-SILC may be very difficult.
- d) An important merit of EU-SILC is the availability of public-use microdata files. A major limitation of using SAE methodology in this context is that the results cannot be replicated by researchers since the microdata files do not include the auxiliary information (nor the software tools) used in constructing the original small area estimates.

Box 4.15

Small Area Income and Poverty Estimates in the United States

The U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program provides annual estimates of income and poverty statistics for all school districts, counties, and states. The main objective of this program is to provide estimates of income and poverty for the administration of federal programs and the allocation of federal funds to local jurisdictions. In addition to these federal programs, state and local programs use the income and poverty estimates for distributing funds and managing programs.

The SAIPE program produces the following county and state estimates:

- total number of people in poverty
- number of children under age 5 in poverty (for states only)
- number of related children ages 5 to 17 in families in poverty
- number of children under age 18 in poverty
- median household income

In addition, in order to implement provisions under Title I of the Elementary and Secondary Education Act as amended, they produce the following estimates for 13,000+ school districts:

- total population
- number of children ages 5 to 17
- number of related children ages 5 to 17 in families in poverty

The estimates are not direct counts from enumerations or administrative records, nor direct estimates from sample surveys. Instead, for counties and states, they model income and poverty estimates by combining survey data with population estimates and administrative records. For school districts, they use the model-based county estimates and inputs from federal tax information and multi-year survey data to produce estimates of poverty. For more information see: <https://www.census.gov/programs-surveys/saipe/about.html>.

4.3.3 Rapid estimates

485. Surveys used to estimate income and poverty are often not as timely as users would like. This is especially the case when they are based on administrative data that is generally not available to the statistical office until many months after the end of the reference year. With this in mind, many countries are looking to develop rapid estimates or nowcasts of key income statistics.

486. In 2017, Eurostat, the statistical office of the European Union, and the United Nations Statistics Division jointly published the “Handbook on Rapid Estimates” (Eurostat, 2017c). It presents four options for producing rapid estimates:

- *Extrapolation*: This method consists of using a historical data series to produce future estimates. It is characterized by good performance under normal conditions but is unable to predict turning points or the effect of changes to government programs.
- *Nowcasts*: This is an increasingly popular method that uses available data to provide early estimates shortly after the end of the reference period. Data that is available from the reference period is used along with modelling to provide an estimate of the relevant statistics.
- *Flash estimates*: Unlike nowcasts, these estimates use the usual statistical process, but with incomplete survey or administrative data. Not waiting for the complete versions of datasets saves time in production, but the quality of the estimates may be lower when incomplete data is used.
- *Leading indicators*: These are indicators linked to the variable of interest that are characterized by better timeliness. Indicators can be based on a variable that is highly correlated with the variable of interest or created through modelling.

487. A number of national statistical offices have started to evaluate rapid estimates in order to produce new statistical products that are timelier than traditionally published statistics. Eurostat and the national statistical offices of the United Kingdom (ONS) and France (INSEE), are three organizations that have developed and started to publish nowcast estimates of income measures over the past few years (Eurostat, 2017c; Fontaine and Fourcot, 2015; INSEE, 2015; INSEE, 2018; ONS, 2015; ONS, 2018c; Stoyanova and Tonkin, 2016).

488. With respect to disaggregation objectives, it is important to note that rapid estimation cannot replace the publication of traditional statistics. Instead, rapid estimates aim to complement traditional statistics and are published while awaiting the official figures. They are generally only produced for a few important but high-level statistics. The traditional survey remains essential for addressing questions that require disaggregate statistics.

4.4 Conclusion and recommendations

489. In conclusion, it is useful to remember that in order to establish trust in poverty measurement and contribute to solid evidence-based policies, Statistical Offices have to regularly assess and

continuously improve the quality of their processes and accuracy of their data. Quality reports which describe the quality criteria and explain any instances in which these criteria could not meet or in which statistical concepts could not be correctly applied will not only facilitate the correct interpretation of survey results but can also provide the basis for future improvements.

490. Overall, in order to understand the impact of and control for both non-sampling and sampling errors, a series of useful recommendations are as follows:

Recommendation 10: Document coverage, precision and any departure from international standards in quality reports to be issued with each data release.

Recommendation 11: Strictly adhere to probability sampling and refrain from substitution. Alternative practices should be reserved exclusively for the purpose of representing specific population groups which otherwise could not be reached with capacities normally available for a developed statistical system.

Recommendation 12: Adjust the sample design and allocation for the required disaggregation. As a minimum, the sample size for each group should never be below 50 units. Ideally, estimation domains will be used as strata so that their size may be controlled.

Recommendation 13: Where available, consider the use of administrative or register data since it can limit response burden and increase accuracy.

Recommendation 14: Establish a clear collection protocol and monitor field work to ensure that it is respected. This should include sufficient interviewer training, ensuring a sufficient number of contact attempts and procedures to address barriers to survey response for at-risk groups (for instance, translations to address language barriers).

Recommendation 15: Ensure that vulnerable populations participate in the design of the study, are protected from any harm and have access to its results.

Recommendation 16: Minimize the impact of item non-response through imputation. Because poverty status is usually derived from several variables that have to be observed simultaneously; it is essential that group representation be preserved by adequate and transparent imputation rules.

Recommendation 17: Compensate for unit non-response and other sources of variation by applying a non-response adjustment and calibration. Ensure that these adjustments remain reasonable by imposing a maximum and minimum factor by which weights are multiplied (for instance between 1/3

Recommendation 19: Establish filtering rules for publications that eliminate unreliable estimates. Do not publish results based on fewer than 20 observations. Similarly, do not publish results with a coefficient of variation that is too high or a confidence interval that is too wide.

5 Supplemental or Experimental Poverty Measures

491. A number of questions become especially pertinent when poverty measures are compared for different groups even within a single country: What is the role of differences in the cost of living? What adjustments do poverty thresholds and resource measures need to adequately capture differences in well-being? What assumptions should be made about resource sharing within the household? What role do assets play in determining well-being? What additional dimensions need to be considered when analysing poverty from a multidimensional perspective? Harmonized standards do not necessarily give satisfactory answers. Nonetheless disaggregation of poverty measures requires careful assessment how it is affected by the methodological problems underlying these questions.

492. Any methodology for measuring poverty relies on a number of assumptions and a multitude of decisions about how to set the poverty threshold, how to define resources and how to implement any particular methodology. Decisions must be made to establish the appropriate threshold or poverty line. Other decisions must be made as to how to vary that threshold, by family size, age, geographic location, or other circumstances. Should the threshold be the same for employed adults as unemployed adults? Should the threshold vary by disability status? What resources should be compared to this threshold? Should assets be taken into account? What assumptions should be made about economies of scale and how household members share or do not share resources?

493. While these decisions might seem “arbitrary”, they mainly rely on normative decisions, deliberations and statistical tests to justify them. How can these be evaluated? One method to assess the impact of any particular assumption or decision is the development of supplemental or experimental poverty measures that change or relax one or more of these elements or parameters. This chapter describes some of the individual country experiences with these supplemental or experimental measures.

5.1 Adjustments to poverty thresholds

494. Even if we were able to agree on a bundle of basic needs that must be met before an individual can be considered not “in poverty”, there are questions about how these basic needs and the cost of fulfilling these basic needs vary across households - by place of residence, family size, age, disability status, receipt of STiK, etc. Adjusting poverty thresholds (or resources) for differences in the cost of living across geographies and household circumstances is one of the thorniest issues in poverty measurement. For example, we may all agree that adequate shelter is a basic need but, depending on the climate, adequate shelter may or may not include central heating or air conditioning. Even within areas with similar climates, the cost of purchasing this adequate shelter may vary significantly due to differences in housing prices, rental rates and energy costs. In addition, the number of bedrooms required for “adequacy” would depend on the size of the household as well as the ages and relationships of the household members. If a member of the household has a disability, there may be some extra shelter requirements (e.g. wheelchair accessibility) that would change shelter costs. Finally, if the household receives a government housing subsidy their out of pocket costs for shelter might be much lower. Ideally poverty thresholds (or the

resource measure compared to these thresholds) should be adjusted to take into account all of these factors for all elements of the basic consumption basket.

5.1.1 Spatial differences within countries with regard to consumption and income poverty

495. Surveys are often collected across many administrative areas within countries. Regional price differences can cause the same bundle of goods to be more expensive in one region than in another. In addition, the bundle of goods necessary to meet basic needs may differ across regions. However, differences in expenditure caused by these regional price and expenditure differences are often not reflected in measured well-being or welfare.

496. Ideally a price index should consider differences in the consumption bundle, differences in the weights of specific items in the consumption bundle, as well as differences in prices across regions. In addition, if the purpose of the price index is to adjust poverty thresholds, it should be constructed using prices paid by households at or close to the poverty threshold. A recent study using data from the Spanish Household Budget Survey found that the cost of attaining a given level of utility in food consumption is greater in the largest cities (Navamuel et al, 2018). A 2018 working paper from the U.S. Bureau of Labor statistics stratified the sample of prices of goods and services into three price level groups finding that over the 36-month study period, changes in prices for the low-price items, on average, were higher compared to high-price items (Cage et al, 2018).

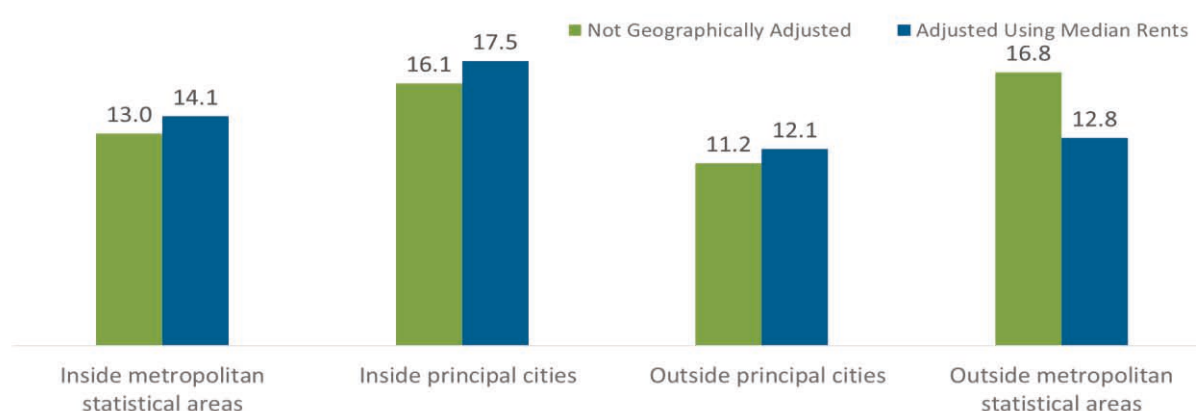
5.1.1.1 United States: Supplemental Poverty Measure

497. The Supplemental Poverty Measure (SPM) published for the United States tackles this problem by adjusting thresholds for geographic differences in shelter costs (rent and utilities). The assumption behind this adjustment mechanism is that differences in cost for the other elements of the thresholds (food, clothing and miscellaneous) do not vary as much as shelter costs. The geographic adjustments are based on 5-year American Community Survey (ACS) estimates of median gross rents for two-bedroom housing units with complete kitchen and plumbing facilities. Separate medians are estimated for the 260 largest metropolitan statistical areas. For each state, a median is estimated for all non-metropolitan areas (47) and for a combination of all smaller metropolitan areas within a state (42). This results in 349 adjustment factors (Renwick et al, 2014; Fox, 2017). Only the housing portion of the SPM threshold is adjusted using this index. The housing shares of the thresholds are derived by the U.S. Bureau of Labor Statistics using the Consumer Expenditure survey and vary across tenure types (renters, owners with a mortgage and owners without a mortgage) but are assumed constant across all geographic areas and family sizes. Prices are measured at the median and therefore may not adequately reflect the rental prices faced by the lower end of the income distribution. Differences in rental costs are used to adjust the thresholds for owners, as well as renters.

498. The geographic adjustments make a difference in SPM rates for particular areas and for particular demographic groups. Nationwide, the geographic adjustments change the poverty status for about 2.7%

of the population but the net change in poverty is relatively small, less than 0.5 percentage points. The largest impact is for those living outside metropolitan statistical areas. The geographic adjustment decreases their poverty rate from 16.8% to 12.8%.

Figure 5.1
Impact of geographic adjustments on poverty rates, 2018



Source: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement.

Note: This figure meets all of the U.S. Census Bureau's Disclosure Review Board (DRB) standards and has been assigned DRB approval number CBDRB-FY19-ROSS-B0032. Renwick (2018).

5.1.1.2 Canada: Market Basket Measure

499. Canada's Market Basket Measure (MBM) of low income adjusts thresholds based on estimates of the price of a basket of goods and services lower-income persons would consume on a regular basis, for various regions of the country.⁷⁹ The MBM threshold represents a cost of specified quantities of food, clothing, shelter, transportation and other items for a reference family of two adults and two children. It is priced for 50 different geographic areas covering about 97% of Canada's population. The MBM threshold is one of three lines used in Canada to produce low-income statistics and was designated as the "official measure of poverty in Canada" by the Government of Canada in 2018.⁸⁰

5.1.1.3 Russian Federation

500. When setting the absolute poverty line in the Russian Federation, the challenge is to address the two causes of interregional disparities – differentiated needs and different levels of consumer prices. For example, the ratio of the maximum to the minimum average regional price for apples was 5.1 while estimates of dietary caloric requirements increase by as much as 20% in colder climates.

501. The consumer basket for the Russian Federation is defined by legislation as a minimum set of food products and proportional expenditure on non-food products and services. The food basket contains in-

⁷⁹ <https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop165-eng.cfm>

⁸⁰ <https://www.canada.ca/en/employment-social-development/news/2018/08/government-of-canada-launches-first-poverty-reduction-strategy.html>

kind units of annual consumption. The nutritional and caloric value of the minimum food basket is further differentiated by three socio-demographic groups: working age, retired persons and children. In general, the cost of the food basket is multiplied by two to get the poverty line. For the working age population, an additional amount is added to account for net income taxes paid. This methodology provides the general framework for determining the level of regional subsistence levels.

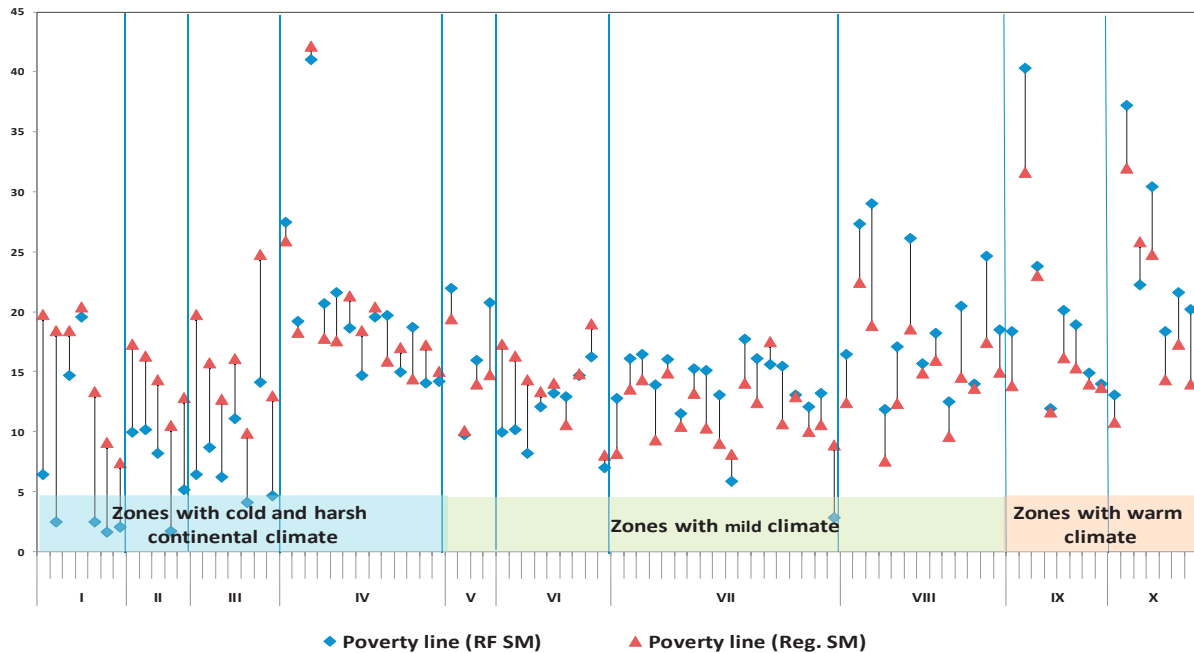
502. Assuming identical needs for goods and services for all regions would not do justice to the extremely varied natural environment and cultural consumption patterns across the territories of the Russian Federation. Regional authorities therefore adjust the basket of goods according to common methodological guidelines to obtain comparable absolute poverty thresholds at regional level. Based on scientific research on regional consumption patterns, 10 different zones were identified which primarily reflect different climate conditions. Regional food baskets are determined using specific nutritional and caloric requirements for each of these zones. Consumption patterns for other goods and services are differentiated by three main climatic zones: cold and harsh, mild and warm. The assessment of needs is further differentiated by socio-demographic group. Prices for regionally representative baskets of goods are obtained by the standardized data collection used to derive the consumer price index.

503. For each climatic zone, the central government recommends a range between 40% and 60% of the total poverty line both for non-food products and for the consumption of services. The resulting poverty line therefore amounts to approximately two times the amount of the food expenditure. Only a few regions exceed those limits. For example, in Moscow, the value of services amounts to 127% of the minimum food basket value for the working-age population and to 75% for retired persons and children.

504. Figure 5.2 compares regional poverty rates when the Russian Federation's subsistence minimum (RF SM) or the regional subsistence minimum (Reg. SM) is used. It is notable that in some regions with cold and harsh climates the regional subsistence minimum implies poverty rates, which are several times higher than if the Russian Federation's subsistence minimum is used. On the other hand, poverty rates tend to decrease in warm climate zones when regional thresholds are applied. The example of the Russian Federation illustrates the sensitivity of disaggregated poverty measures against empirically established differences in the cost of living.

Figure 5.2

Regional poverty rates using subsistence minimum (RF SM) and regional subsistence minimum (Reg. SM) of the Russian Federation for three climatic regions



5.1.1.4 Using Purchasing Power Price Parities

505. Purchasing Power Parities (PPPs) are estimated for many countries in the world by the World Bank's International Comparison Program (ICP) and used by the World Bank to calibrate its international poverty lines and produce estimates of the incidence of global poverty and extreme poverty. The ICP is a worldwide statistical initiative led by the World Bank under the auspices of the United Nations Statistical Commission, with the main objective of providing comparable price and volume measures of gross domestic product (GDP) and its expenditure aggregates among countries within and across regions. Through a partnership with international, regional, sub-regional and national agencies, the ICP collects and compares price data and GDP expenditures to estimate and publish purchasing power parities (PPPs) of the world's economies.

506. A primary concern with using PPPs to adjust poverty thresholds across or within nations is that they are estimated using a broad array of goods and services in their consumption bundle. This bundle may be very different than the bundle implicit in the poverty thresholds. To the extent that the consumption bundle does not reflect the consumption patterns of low-income households and the prices used to develop the index are not the prices typically faced by these lower-income households, they may not be appropriate for use in poverty measurement. Indeed, the Global Poverty Commission, chaired by Tony Atkinson, recommended that "the global poverty estimates should be updated up to 2030 based on of the International Poverty Line for each country set in local currency, and updated in line with the change

in the national CPI or, where available, national index of prices for the poor; the estimates would not be revised in the light of new rounds of the ICP” (World Bank, 2017).

507. For the United States, the Bureau of Economic Analysis (BEA) publishes regional price parities (RPPs) that could be used to adjust the entire SPM thresholds rather than just the housing portion of the threshold. Several papers compare these to the median rent index currently used for the SPM (Renwick et al, 2014; Renwick et al, 2017). Generally, the adjustments using the RPPs were smaller than the adjustments using either median rent. These differences were driven by the differences in the weight assigned to housing in each method. In addition, the BEA has developed a special version of the RPPs that include only food, apparel and shelter, something akin to the poverty-specific PPPs referred to in the Atkinson study. While the expenditure weights for these more limited RPPs were more in line with the expenditure shares for the poverty thresholds, they were based on prices for the general population, not the prices faced by the low-income population.

5.1.1.5 European Union

508. For the European Union, there is a common methodology, and yet no common EU-poverty threshold. Inspired by studies on poverty and deprivation in the 1970s (e.g. Townsend, 1979), the European Union defines the poor as “individuals whose resources are so small as to exclude them from the minimal acceptable way of life of the Member State in which they live”. In practice, the poverty line is usually set to 60% of the national median equivalised income, which implies different values for each Member State. As price levels tend to be more similar across countries than income levels, these differences do not correspond perfectly to similar differences in the cost of living. If for example, thresholds are expressed in PPPs, the threshold in Luxembourg for 2017 was more than 5 times than the one in Romania. This means that Luxembourg is not only more expensive but also that the poor in Luxembourg will be able to purchase a considerably larger basket of goods and services than those who are defined as poor in Romania.

509. The national median based poverty thresholds, which are used in the European Union, are therefore not merely adjusted for the cost of identical goods but represent an adjustment of its own sort. They intend to capture differences of the function of income to guarantee an “acceptable way of living”. This reflects the basic assumption of a relative poverty definition that identical capabilities (Stiglitz et al, 2009) may require quite different means in different societies. The mechanical construction of the relative median threshold can provide only an extremely simplified approximation of that relationship. The 60% of national median income threshold works relatively well for comparisons across EU Member States. Occasionally, however, it may be useful to assess this conventional threshold against some more concrete measures for the acceptable way of living. This holds especially for disaggregation purposes.

Box 5.1

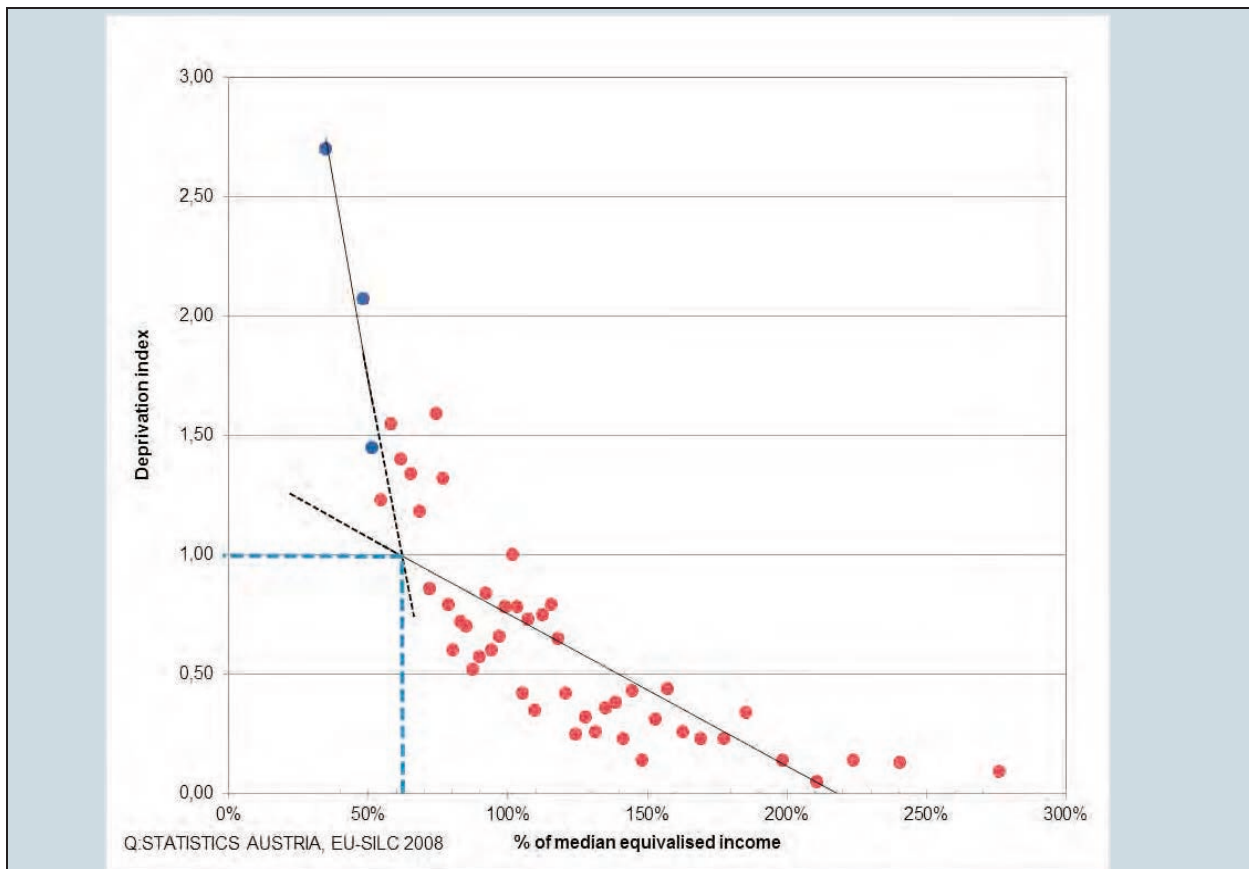
Assessment of the 60% of median threshold for the elderly in Austria

In a study on poverty among the elderly, Statistics Austria (Eiffe et al, 2011) assessed income levels against an index of deprivation.

The deprivation index, which was used in the study represented seven items, including adequate food, heating or payment of regular bills. These items had been declared as essentials for the minimum acceptable way of life in Austria by the majority of respondents in a small nationwide survey. The items were selected out of a broader list, based on qualitative methods, which involved poverty experts, including professionals but also homeless and social assistance recipients. The study also assessed whether these items could be considered equally applicable for all relevant groups. On these grounds, for example, the affordability of a car was excluded from the list of essentials as this appears less important for urban populations and the elderly than for families with children or rural places. Figure 5.3 shows average deprivation scores among 50 income groups of the same size for two person households in Austria. The image suggests that below a level of about 60% of the median average, deprivation scores were disproportionately higher than above 60% of the median. With econometric analysis it this threshold was even found more discriminatory than any higher or lower threshold. This was taken to confirm conventional poverty lines and equivalence scales as roughly appropriate for Austria. Moreover, in more detailed analysis the conventional threshold was also found suitable for disaggregation by age.

Nonetheless, the association between deprivation and income may vary across groups and also over time. While such analysis is clearly helpful for the critical assessment of the conventional poverty line, it may be unrealistic to expect a perfectly robust, scientific poverty line (see Piachaud, 1981).

Figure 5.3
Relationship of deprivation index and income among two-person households



5.1.1.6 Challenges and issues for further research

510. There are a number of areas on this topic that require more research and analysis. It would be helpful to identify the best practices and methods to include non-food prices given the heterogeneity of items within and across nations. In the United States, the SPM is exploring methods to take into account differences in amenities that may improve the quality of life but not be reflected in spatial price differences.⁸¹

Recommendation 20: As a general rule, measurement of monetary poverty has to ensure equivalence of standards of living for all regions and groups within a country, notably with regard to the needs of children/ persons with health impairments or disabilities. When disaggregating poverty estimates for smaller geographies, methodologies that take into account spatial differences in the cost of living are important, particularly in large countries with heterogeneous cultural norms and price levels. Relative thresholds such as those used in the European Union are a convenient practical alternative.

5.1.2 Household size and age of household members

511. Generally, equivalence scales are used to adjust poverty thresholds to reflect differences in household size and composition. These equivalence scales make assumptions about economies of scale in consumption and whether or not needs vary by the ages of household members. Although equivalence scales are an essential tool for poverty measurement, methods to derive such a tool – whether subjective, normative, or behavioural (Hagenaars et al, 1994) – remain a constant concern. The impact of different scales is explored in the Canberra Handbook (2011) and has been further discussed in the *Guide on Poverty Measurement* (UNECE, 2017a). The Guide concluded in the recommendation that the square root of household size should be applied as an equivalence scale for CES countries.

512. In the context of disaggregation for different groups in the population, it is particularly important to understand that simple per capita measures assume no economies of scale and no differences in needs by age. Consequently, per capita measures imply higher thresholds for bigger families with many children than if – for example - household measures would be used. By contrast, per capita measures imply comparatively low thresholds for elderly persons who tend to live in smaller households. Many countries and all recent OECD publications use the square root equivalence scale. This assumes economies of scale (a four-person household needs only two times the resources of a single person household) but no age-based differences in needs. Several alternative equivalence scales take into account age differences by assigning differential weights to adults and children. What is sometimes referred to as “Oxford scale” or “OECD equivalence scale” (but in fact has never been used by the OECD) assigns a value of 1.0 to the first household member, 0.7 to each additional adult and 0.5 to each child. An alternative scale proposed by

⁸¹ While price differences across regions may be apparent, the problem of comparing well-being across regions is more complex. A city with higher out-of-pocket expenditures for rent and utilities may have amenities that compensate for those higher out-of-pocket expenditures. In a world of perfect mobility consumers would move from higher cost to lower cost areas until their utility was maximized. In the real world there are often restrictions on mobility, particularly for individuals living below or near the poverty threshold (Ziliak, 2011).

Haagenars et al (Hagenaars et al, 1994), and often referred to as the OECD-modified scale, assigns a value of 1.0 to the household head, 0.5 to each adult and 0.3 to each child and is used by Eurostat. The U.S. SPM uses a three-parameter scale, which allows for a different adjustment for single parents. Units with one or two adults are assigned $(adults)^{0.5}$, single parents $(adults + 0.8 * first\ child + 0.5 * other\ children)^{0.7}$, and all other families $(adults + 0.5\ children)^{0.7}$.

513. These assumptions regarding economies of scale in consumption may vary significantly depending on the bundle of goods included in the threshold. If public education, health care, and dependent care (whether for children or elders) are included in the consumption bundle, the economies of scale and appropriate equivalence scales should be adjusted appropriately (see Section 5.2.1). For example, using the OECD equivalence scale a family with one adult and one child will require less than a family with two adults (1.7 versus 1.5). If the cost of public education is included in the threshold then the equivalent threshold for the one adult one child family should be considerably higher than the threshold for the family with two adults. If childcare is included in the threshold there would be considerably higher needs for families with infants than families with older children.

514. In related work, Renwick and Garner (Renwick and Garner, 2016) used different equivalence scales for different elements of the SPM poverty thresholds in order to reflect the different economies of scale enjoyed in the consumption of each element. For example, there are probably greater economies of scale in housing than in food consumption. Their research found that customizing the equivalence scales in this manner had very small impact on the subsequent poverty rates.

515. UNICEF notes that including child-specific needs in the threshold calculation would improve the measurement of child poverty. For example, the Russian Federation defines the subsistence minimum for working age adults, pensioners, and children.⁸² The official poverty thresholds for the United States vary by the size of the family and the number of children. For example, the poverty threshold for a family of three adults for 2017 was US\$ 19,173 while the threshold for a family of one adult with 2 children was US\$ 19,749.⁸³

516. However, there is lack of evidence of use of child-specific consumption baskets and consequently poverty lines. Obtaining individual data for consumption is time consuming and expensive, and therefore seldom done. One solution could be to include child-specific items and expenditures in consumption modules of ordinary living conditions household survey. Using information on expenditures on education (uniforms, books, tutoring and private lessons) and health related expenditures plus, imputing prices to items included in material deprivation modules, such as clothing, shoes, etc., it could be possible to derive a more child-specific consumption basket to compare with household expenditure power or income, and derive a measure of child poverty. (Of course, the caveat would be that if these expenditures are already counted in consumption aggregates, then the measure would be endogenous and not valid). Including

⁸² Consumer Basket in the whole Russian Federation, valid till 31 December 2020. Federal Law of the Russian Federation No. 227-Φ3 dd. 03.12.2012 with amendments done by the Federal Law of the Russian Federation No. 421-Φ3 dd. 28.12.2017

⁸³ <https://www.census.gov/data/tables/time-series/demo/income-poverty/historical-poverty-thresholds.html>

child-specific items in consumption/expenditures modules could provide at least an estimation of the percentage of family resources that are allocated to children.

517. Among countries outside the UNECE region that experimented with individual consumption (or age-categorical consumption) there are:

- China Health and Nutrition Survey: includes individual food consumption (2011),
- Bangladesh Integrated Household Survey: includes individual food consumption (2013-2015),
- Brazil HBS 2008-09 includes individual food consumption,
- Senegal – Survey on poverty and structure of families (Agence National de la Statistique et de la Démographie, 2015).

518. Most of these surveys focus on food consumption, which is the biggest source of intra-household differences in consumption. However, other differences in resources allocation regarding clothing, use of items (for example access to media), and child-specific items (such as school supplies and toys) cannot be ignored.

519. As said, a broad classification presented in Hagenaars et al (Hagenaars et al, 1994) could be summarized as follows: normative or social security scales; subjective scales; scales based on demand models. While in much of the existing literature, equivalence scales have been estimated based on consumption/expenditure data, there is a growing body of literature using data on subjective perceptions of economic well-being to derive equivalence scales. Examples include the use of income evaluation and/or minimum income questions (Bishop et al, 2014; Carbonnier, 2019; De Vos and Garner, 1991; Flik and Van Praag, 1991; Garner and De Vos, 1995; Hagenaars et al, 1994; Kapteyn et al, 1988; Martin, 2017), minimum spending question (Garner and Short, 2003; Garner and Short, 2004), income satisfaction (Bütikofer and Gerfin, 2009), or personal evaluations of material well-being (Dang et al, 2019).

520. Mysíková et al (Mysíková et al, 2019) use data from the EU-SILC to derive subjective equivalence scales for various European countries. They find that the subjective equivalence scales generally provide higher economies of scale than the OECD-modified scale and that economies of scale, in fact, differ across countries and for households with differing numbers of adults or children. They find that Eastern European countries generally exhibit lower economies of scale than Western European countries when subjectively assessed minimum income needed by households is considered.

521. The observable variation in economies of scale across countries and assessment methods implies that when identical parameters are used, these cannot be equally representative of the household economies of scale in each country. This fact fundamentally limits the international comparability of all monetary poverty measurement - whether based on consumption or income - even if identical concepts and methodology are applied. Comparable estimates may require country specific equivalence scales and it is advisable to regularly assess the empirical economies of scale in a country and sensitivity of poverty estimates using different equivalence scales (Atkinson et al, 2002). It ought to be emphasized that equivalence scales directly affect disaggregation between population groups who are typically found within households of different sizes, such as the elderly or families with children.

Box 5.2

Equivalence scale based on a demand model in Tajikistan

With higher standards of living, per capita calculations are increasingly problematic. However, the OECD scale which is widely used for international comparisons may also not be fully adequate for the level of development of middle-income countries. Several countries have therefore implemented ad-hoc scales for their specific national context. In Tajikistan and Ukraine scales based on the estimation of demand systems which formulate consumption needs as a function of number of household members and their basic demographic characteristics have been developed. The approach taken in both countries went beyond conventional Engel curves which consider only the demand for food to assess equivalent income. Ideally, such scales need to be reassessed periodically to adjust the scale to the demand implied by current standards of living.

In 2009 the European Union program EUROPEAID funded a project in Tajikistan with the aim of evaluating particularly the impact of remittances on poverty in Tajikistan. In that context, Betti and Lundgren (Betti and Lundgren, 2012), which were involved in planning the sampling design and in training the staff of the Tajik State Statistical Office (TAJSTAT), also estimated official ad-hoc equivalence scales. A complete demand system was estimated, following the Almost Ideal (AIDS) model of Deaton and Muellbauer (Deaton and Muellbauer, 1980), by introducing the socio-demographic variable using the Ray Price Scaling method (Ray, 1983). To estimate the AIDS model, 12 COICOP consumption expenditure components, were collapsed into seven groups. Following the estimated demand parameters, the TAJSTAT scale defines the cost of living of household members as displayed in the table below. The scale is currently not used but can be referred to assess the sensitivity of per capita base poverty estimates which are still common in Tajikistan today. We should note, however, that Tajikistan continues to use per capita income in its official poverty measures.

Table 5.1

Proposal of new equivalence scales for Tajikistan

Category	Scale
First adult aged 15-58 (women) or 15-62 (men)	1.00
Any subsequent adult	0.80
First elderly adult aged 59 + (women) or 63 + (men)	0.80
Any subsequent elderly adult	0.65
Children aged 12-14	0.70
Children aged 7-11	0.60
Children aged 3-6	0.50
Children aged 0-2	0.40

Source: Betti and Lundgren (2012)

A similar exercise was conducted also in Ukraine. Like in Tajikistan, the empirically verified parameters have never been used to replace the established equivalence scale. Small differences with regard to age groups appeared to not have any significant impact on the overall estimates. While the exercise seems to have confirmed the established scale, it is useful to repeat similar assessments from time to time to identify if increased prosperity may have shifted demand parameters.

Recommendation 21:

- a) Equivalence of measurement should be assessed empirically. Sensitivity analysis can compare poverty profiles of official poverty measures with supplemental poverty measures which may be either monetary (based on income, expenditure or assets) or nonmonetary.
- b) The sensitivity of monetary poverty measures should be regularly assessed using alternative equivalence scales. While simple equivalence scales based on the square root of family size are useful for international comparability, countries should explore the empirical development of equivalence scales that go beyond family size to include differences in needs based on age, disability status and health.
- c) If Social Transfers in Kind (STiK) are included in resources, special consideration needs to be given to appropriate equivalence scales.

5.1.3 Disability status/High medical expenditures

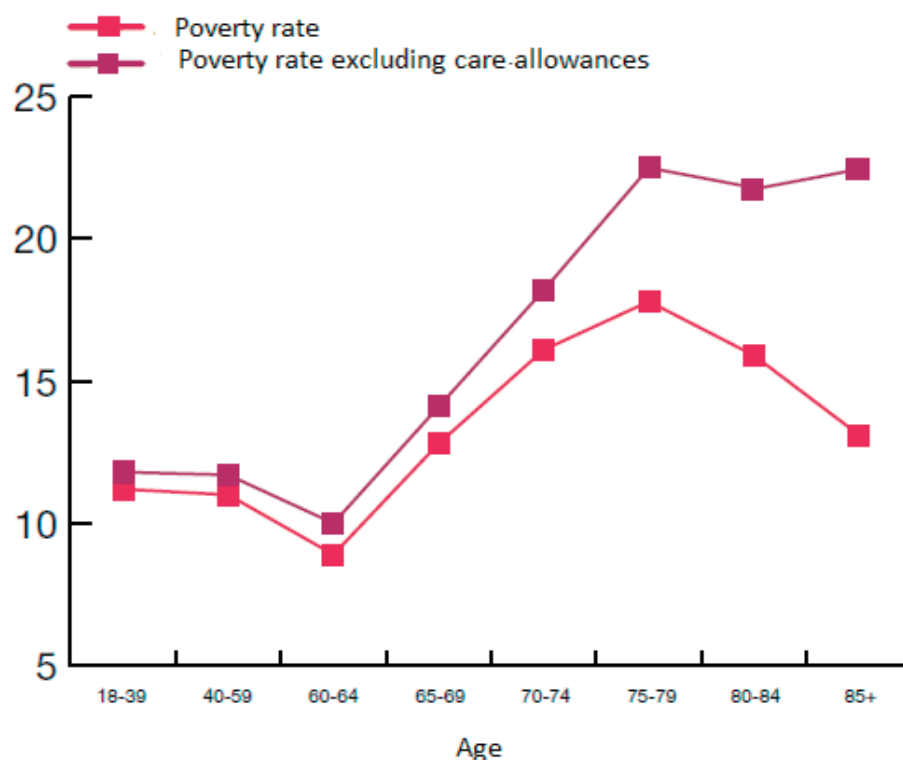
522. Differences in costs and needs for households with disabilities or high medical expenditures is another dimension of this issue. Should individuals with disabilities be given a higher weight in an equivalence scale? Another approach would be to adjust resources to reflect the higher expenditures required by those with a disability. For the U.S. supplemental poverty measure out-of-pocket medical expenses, including insurance premiums, are subtracted from resources before assigning poverty status. Poverty status is based only on resources available for food, clothing, shelter and utilities. Resources required for other necessities, health care but also child care, income taxes, payroll taxes, and commuting expenses are subtracted from the resource measure. Other countries would have to explore adding questions on these types of expenditures to their household surveys. For the United Kingdom, the New Policy Institute published a report that adjusted income to account for disability-related benefits arguing that, for instance, £100 of disability benefit merely makes up for the extra £100 of needs required, and should not be considered as extra income (Marciano et al 2012, MacInnes et al, 2014).

523. In Austria, care allowances (Pflegegeld in German) are intended to compensate for care needs (even that may not always be accomplished) and their inclusion may introduce serious bias for disaggregation of poverty measure. Nonetheless, according to the EU-SILC framework such benefits are to be included in the resource measure. Figure 5.4 shows how poverty rates among elderly citizens may be underestimated if care allowances are considered as income. According to Austrian EU-SILC data from 2008, persons above the age of 75 would have about twice the poverty rate of the average adult poverty

rate if care allowances are excluded from income while the conventional estimate (which includes care benefits as income) is only 30% above the average (Eiffe et al, 2011).

Figure 5.4

Poverty Rate with and without Care Allowances for Austria, 2008



Source: Statistik Austria, EU-SILC 2008.

Recommendation 22: Further research on the special needs of persons with disabilities is necessary to ensure these are taken into account in monetary poverty measures. These adjustments may be done on the thresholds, perhaps through equivalence scales, or on the resource-side of the measure.

5.2 Adjustments to resources

5.2.1 Social transfers in kind (STiK)

524. Social transfers in kind consist of individual goods and services provided as transfers in kind to individual households by government units (including social security funds) and non-profit institutions serving households (OECD). The total value of STiK is recorded in national accounts or other aggregate sources. Based on this information it is possible to make imputations to obtain estimates on the individual level. The Canberra Handbook recommends accounting for the effects of STiK on the distribution of

income statistics for a number of reasons. First, it enables a more robust means of measuring income inequality, and income poverty across countries. For instance, households living in a country with state-provided health and education services will have a higher standard of living than households with similar disposable income (before STiK) but living in a country where these services need to be purchased from the market. Secondly, it improves comparisons within countries where STiK is distributed unevenly across the distribution. Finally, STiK allows improved measurement over longer time series where, for instance, government policy has shifted between contractionary and expansionary fiscal policies.

525. However, because of measurement challenges STiK are often excluded from the welfare measures used for poverty statistics. OECD research (Balestra and Sustova, 2017) shows that the majority of countries that contribute to their OECD Income Distribution Database do not produce any estimates of these transfers, with even smaller numbers including them within their income statistics.

526. Tonkin et al (Tonkin et al, 2014) examine the measurement of two major social transfers - education and health - in the United Kingdom and Finland. They find that accounting for STiK reduces the Gini coefficient - and therefore measured income inequality - with state provided health care having the largest overall impact in both countries. This outcome is mirrored in measured relative poverty rates where the authors finding that the inclusion of STiK within household income reduces the at-risk-of-poverty rates in both countries.

527. There are several groups of measurement challenges to confront when considering STiK in income and poverty measures, whether or not these are used for international comparisons. These challenges include: differences in the types of transfers included by each country; differences in methodologies used to assign a value to these transfers; for poverty measurement, the challenges of ensuring consistency between resource measures and poverty thresholds; concerns about the “fungibility” of transfers; concerns about underreporting of benefits in survey data, and the need to consider an adjustment to equivalences scales when STiK are added to resources.

528. The types of benefit included within social transfers in kind vary from country to country, affecting comparability. For example, Tonkin et al (Tonkin et al, 2014) showed that while both the United Kingdom and Finland include the value of health and education in-kind transfers in their statistics on income (re)distribution, social housing and public transport subsidies were included in the United Kingdom but not in Finland, and elderly care was included in Finland only.

529. The United States incorporates a narrow concept of social transfers in kind in the calculation of its SPM. The SPM thresholds are based on expenditures for food, clothing, shelter and utilities. The SPM resource measure takes into account the value of STiK intended to support the consumption of these same four categories. Since medical expenditures are NOT in the SPM thresholds, the SPM resource measure does not include a value of medical benefits (insurance or actual consumption) in the resource measure. In a similar vein, since educational expenditures are not included in the threshold for the SPM, the value of educational services are not included in the resource measure.⁸⁴ An advantage of the inclusion of STiK in the supplemental poverty measure in the United States is the ability to disaggregate the impact

⁸⁴ For more detail regarding the specific methodologies used by the United States to estimate the value of these social transfers in kind see Johnson et al (Johnson et al, 2010).

of each specific transfer on the poverty rate. For example, in 2018, holding all else the same and assuming no behavioural changes, the major food assistance program (SNAP) kept 3.2 million people out of poverty while housing subsidies reduced poverty by 3.0 million (Fox, 2019).

530. In contrast to the narrow concept used in the United States, researchers in Norway (Aaberge et al, 2017) extend the conventional income distribution studies by accounting for the value of a wide array of public services received by households. Using EU-SILC data and OECD expenditure data they incorporate primary and second education, healthcare services, early childhood education and care and long-term care for the elderly and disabled for 23 countries from 2006 to 2009. They set the value of these transfers at the cost of producing them, assign average benefits by gender and age group and use equivalence scales that recognize that children and elderly have higher needs for basic public welfare services. By replacing disposable income with this extended income concept, they find significantly smaller estimates of inequality in all countries with an approximate reduction of 20% in the Gini coefficients. The estimated proportion of people who are at risk of poverty (income below 60% of median income) is reduced by at least 40% in most countries. The rankings of countries by the poverty headcount were insensitive to the change in the income measure and poverty rose in most countries between 2006 and 2009 irrespective of the income definition used.

5.2.1.1 Measuring social transfers in kind: United Kingdom

531. Since 1961, the UK Office for National Statistics and its predecessor organisations have published analysis showing how taxes and benefits redistribute income between different types of households, in an annual publication called “Effects of Taxes & Benefits on Household Income”. In order to measure as completely as possible, the impact of the tax and benefits systems on households, this analysis includes estimates for the value of various STiK, in addition to the cash benefits received. From a tax perspective, both direct taxes (such as income tax and employees’ National Insurance contributions) and indirect taxes ultimately borne by households (such as fuel and alcohol duties, and VAT/purchase taxes) are included.

532. The STiK that the ONS currently include in their measures are education, health, subsidies for housing, rail and bus, free school meals. The methodology to distribute values to households varies STiK-by-STiK but all involve allocating the total cost of provision borne by the state to households. This analysis is based on the United Kingdom’s Household Budget Survey, which is known as the Living Costs and Foods survey (LCF) since it contains detailed data on both household income and expenditure, facilitating the production of microdata on both social transfers in kind and indirect taxes.

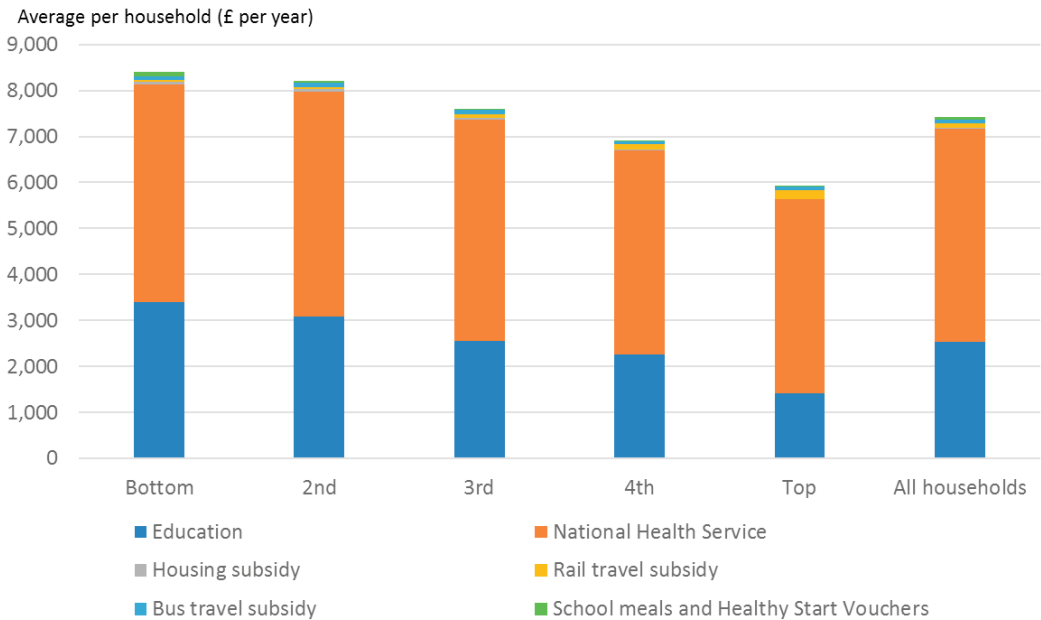
- Education services: ONS use an “actual consumption” approach to allocate education services to households. This involves attempting to measure each household’s likely consumption of these services to distribute total education spending data provided by public bodies, and adjusted to reflect evidence that per capita education expenditure increases by child age.
- Health services: In contrast to education services, an “insurance” approach is taken to allocate health spending to individuals which models individuals’ likelihood of requiring health services based on their characteristics. Each individual in the LCF is allocated a benefit from the National Health Service according to the estimated average use made of these various types of health

service by people of the same age and sex, and according to the total cost of providing those services.⁸⁵

- **Travel subsidies:** Travel subsidies cover the support payments made to bus and train operating companies. Rail subsidy is allocated to households based on their spending on rail travel taken from the LCF. In making these allocations, allowances are also made for the use of rail travel by the business sector, tourists and the institutional part of the personal (household) sector (for example, people who do not live in private households; i.e. prisoners, or people in care homes). Bus travel is calculated in a similar way but additional levels of benefit are allocated to those household containing individuals who indicate in the LCF that they hold a concessionary bus pass. The figures for rail travel subsidy also take into account the Government grant to the infrastructure operator (Network Rail), which enables Network Rail to lower the charges levied on each train operating company, using data supplied and published by the United Kingdom’s Department for Transport.

533. Figure 5.5 highlights the latest figures available for the value of social transfers in kind by income quintile, based on disposable income equivalised using the OECD modified scale. In 2016/17, the poorest fifth of households received the equivalent of around £8,400 per year from social transfers in kind, compared with £5,900 received by the top fifth.

Figure 5.5
Social transfers in kind by income quintile group, United Kingdom, 2016/17



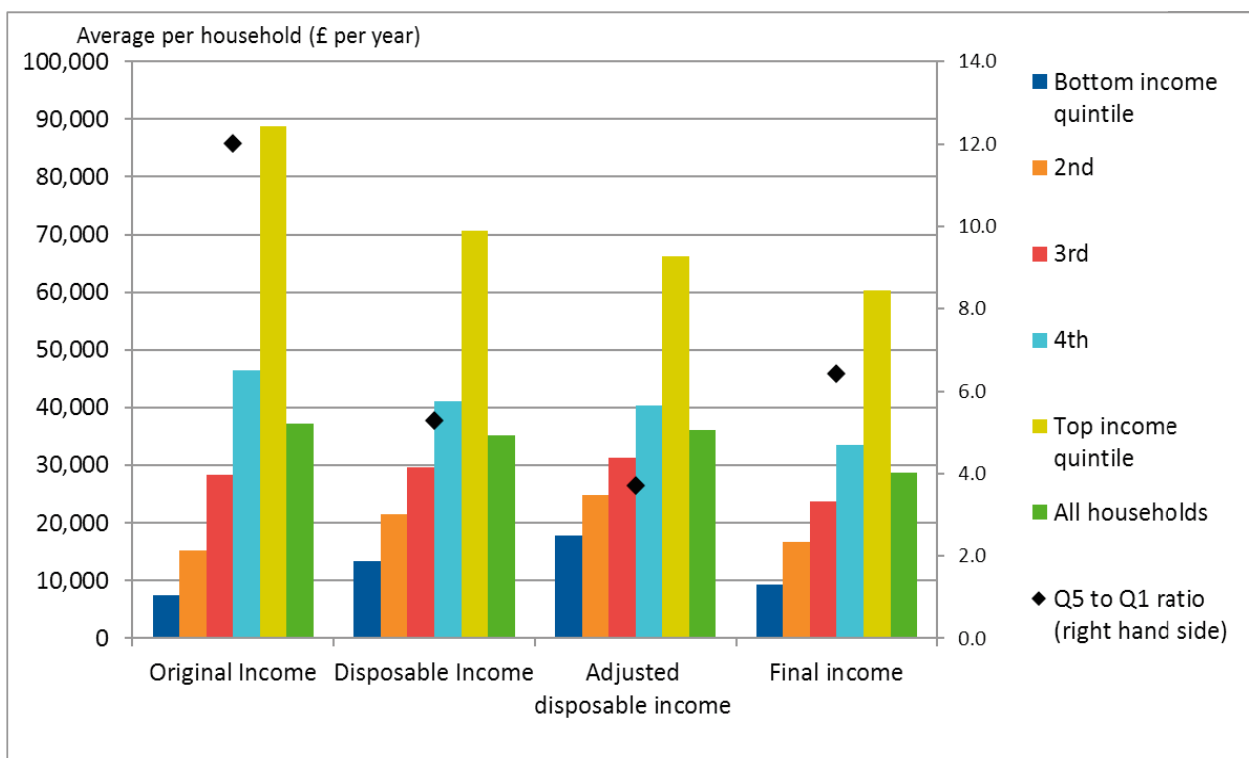
⁸⁵ Ideally, an adjustment should be at least considered for the use of private health care services. However, this is not felt to be feasible from the data available. For example, whilst it is possible to identify individuals paying for private health insurance, it is not possible to ascertain the level of cover provided and therefore the likely impact on that individual/household’s use of the NHS. In addition, it is generally acknowledged that all individuals do benefit from the simple existence of public sector health services, so there is an argument for not taking private health care into account in estimating the value of STIK. The assigned benefit is relatively high for young children, low in later childhood and through the adult years until it begins to rise from late middle age onwards.

534. Figure 5.5 summarises the effect that these STiK has on the United Kingdom income distribution. In this figure, original income is defined as income from employment, private pensions, investments and other non-government sources. In 2016/17, the richest fifth of households had an average original income (i.e. before all taxes and benefits) of £88,800 per year, compared with £7,400 for the poorest fifth – a ratio of 12 to 1. Direct taxes and cash benefits (Disposable Income) lead to income being shared more equally between households, with the richest fifth of households having average disposable incomes that were five times those of the poorest fifth (£70,700 per year and £13,400, respectively).

535. Adjusted disposable income, which includes STiK, was £17,800 per year for the bottom quintile and £66,300 for the top quintile, meaning that the S80/20 ratio was reduced to 3.7 to 1. However, in order to get a complete picture of the effects of the tax and benefits system, it is necessary to also taken into account indirect taxes, which are included in the final income measure. Although richer households pay more in indirect taxes than poorer ones, they pay less as a proportion of their income. This means that indirect taxes act to increase inequality of income, with the S80/20 ratio increasing slightly, to 6 to 1.

Figure 5.6

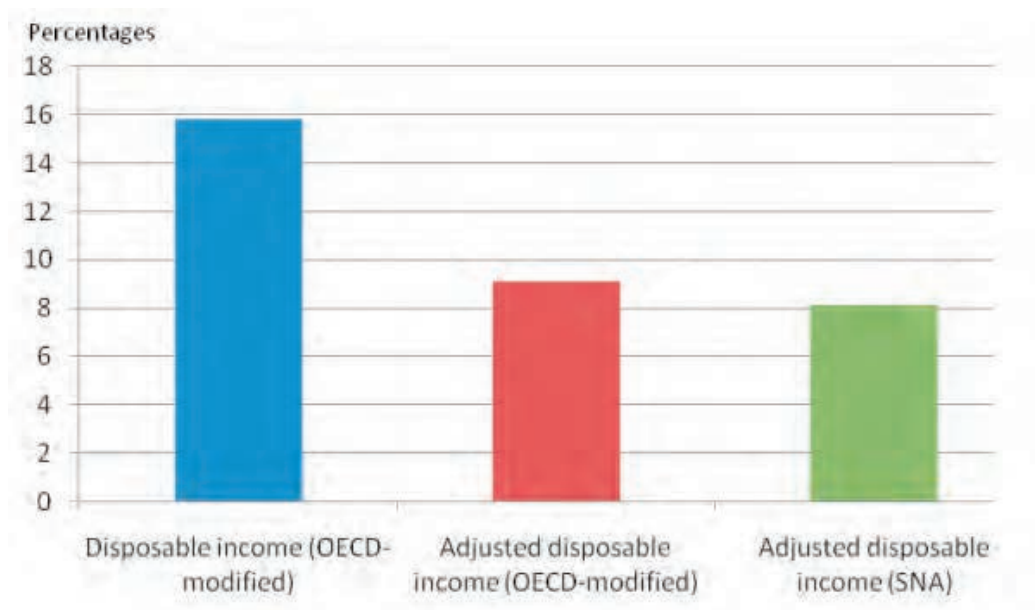
Original, disposable, post-tax and final income by quintile groups, all households, financial year ending 2016/17



536. The United Kingdom's main poverty statistics are produced and published by the United Kingdom's Department for Work and Pensions (DWP) in a publication called Households Below Average Income (HBAI). STiK are not included within the HBAI statistics. Figure 5.7 presents three measures of poverty according to the impact of including STiK within United Kingdom poverty estimates, based on the Effects of Taxes & Benefits data described above presented by Tonkin et al (Tonkin et al, 2013).

537. The three measures presented are disposable income and adjusted disposable income both equivalised using the OECD-modified scale. The third measures equivalises adjusted disposable income with the Simplified Needs Adjusted scale. This scale accounts for the extra non-cash needs when carrying out poverty analysis based on adjusted disposable income. In the case of young children, the OECD-modified scale assigns a smaller value for children than for additional adults in the household, based on assumed needs. However, young children have a relatively higher need for both education services and healthcare (although less than for older people). Therefore, applying a standard equivalisation scale to adjusted disposable income would risk overstating the standard of living of those households with young children.

Figure 5.7

Relative at-risk-of-poverty rates, United Kingdom, 2011/12

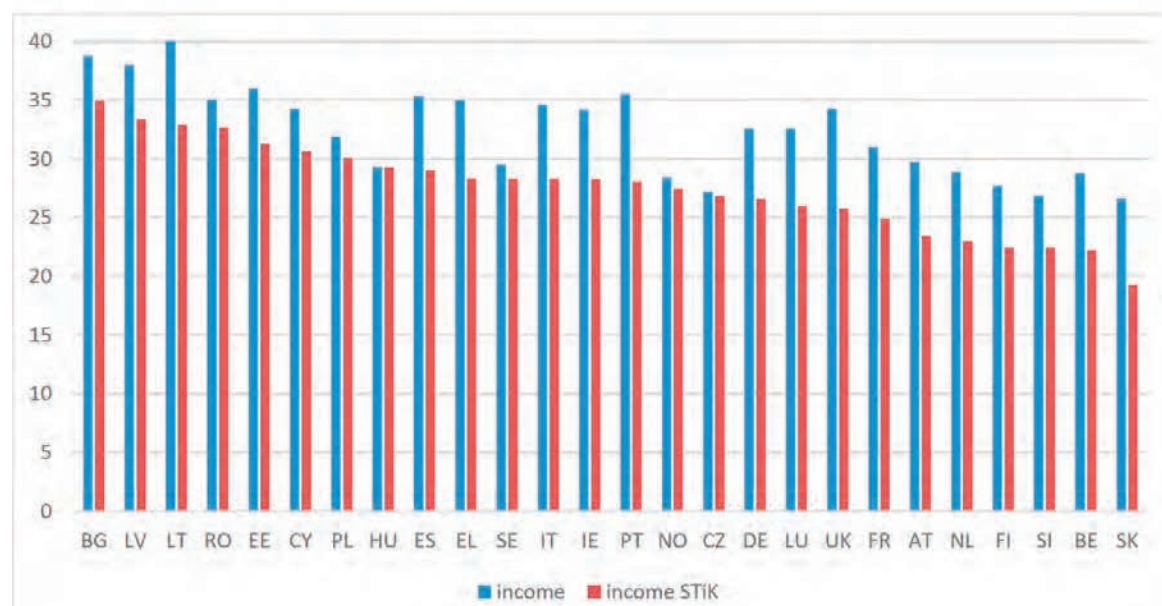
538. In 2011/12, 16% of people were below the at-risk-of-poverty threshold, based on the equivalised disposable income measure. The impact of including STiK but retaining the OECD-modified equivalisation scale reduced the headline rate to 9%. For the overall population, the effect of applying the SNA equivalisation scale was relatively small, reducing the relative at-risk-of-poverty rate for adjusted disposable income to 8%.

539. A note by Eurostat Unit (Grundiza, 2019) examines the impact of STiK on inequality measures using the 2016 EU-SILC survey ad hoc module on access to services. The education STiK values are imputed to the EU-SILC micro data to only those who consume the services, in this case to pupils and to students. For health STiK values are imputed to the EU-SILC micro data to all people by age and sex based on age-gender specific profiles from national data sources provided directly by the Member States through the Ageing Working Group channel (source: European Commission's Directorate General for Economic and Financial Affairs (ECFIN)). The analysis uses a European average equivalence scale that is a population-weighted average of country-specific equivalence scales that combine non-cash equivalence scales with the cash income scales. GINI coefficients are compared between income including and excluding health

and educational STiK. Results show that for almost all countries GINI coefficient for income and educational and health STiK is smaller than those for just income (except for Hungary, and minor changes for the Czech Republic). Larger differences are for the United Kingdom (8.5 pp), Portugal (7.5 pp), Lithuania (7.4 pp), and Slovakia (7.4 pp).

Figure 5.8

GINI coefficients for income (SNA scale) with and without educational and health STiK



5.2.1.2 Mexico – Strategies to measure self-consumption

540. In Mexico, current income is defined as both monetary and non-monetary receipts that are expected to recur regularly in a given period and are available for the acquisition of household consumption of goods and services.

541. In accordance with international statistical standards and practices, labour income is defined as all income received by the members of the household, as a result of their current or previous participation in any activity carried out in an economic unit whose purpose is to produce or provide goods and services for the market, self-consumption or generation of public goods or services (ILO, 2003). This income can be in cash, in kind or in services. The National Survey of Household Income and Expenditures (ENIGH) asks about compensation in kind for each source of income, including income from independent work or self-employment. The value of remuneration in kind uses the respondent's own estimate.

542. The survey also asks respondents to report income from "independent work" or self-employment. Independent work income is all the resources in cash or in kind that members of the household receive regularly for their performance as independent workers in their main job or in their secondary work. Under this concept, a wide variety of economic activities are covered, ranging from those carried out in small businesses in the household sector to produce goods and services for the market, to the production of final goods and services for self-consumption or barter. The factor they have in common is that the

people who receive them work for themselves and are owners of an unincorporated company, and therefore belong to the household sector.

543. In the estimate that is made of independent work, a non-monetary component is included, self-consumption. Its value is established based on the estimates made by the informant of the retail market value of the goods and services produced or marketed by the household and taken from the business, for household consumption, in a given period. This non-monetary component was included in the traditional ENIGH within the monetary income and its treatment is the main difference with this construction of the ENIGH. In this construction, the self-consumption data (including self-supply) is presented in the account of the income from the businesses of households and no longer presented separately, in a specific item of non-monetary income. Effective with the ENIGH 2016, the value of self-consumption is based on the estimate of the value given by the person responsible for the business rather than the value reported by the person who reported on household expenses.

544. Transfers are the monetary entries received by the members of the household and for which the provider or donor does not demand compensation of any kind. The variables included under this concept related to in-kind income include benefits from government programs, in-kind transfers from other households (gifts) and in-kind transfers from institutions. The value of transfers in kind is estimated by the informant based on the retail market value of the goods or services received. In 2002 the collection of data on STiK in the ENIGH was expanded from just two sources (school breakfasts and food pantries) with integration of a section on “supports” into the Basic Questionnaire for each of the expenditure items and in the Daily Expense Booklet. In 2008 the name of this section is changed by transfers in kind. In the 2010 ENIGH the questions of transfers in kind were changed to capture information on the items and services that the members of the household received; the amount and form of presentation in which they received the article or service; whether any of the members of the household made any payment for the items or services received, as well as what they paid if they had to pay something. Likewise, questions of transfers in kind are intended to know the cost that the informant estimates would have been paid in case of having purchased the item or service; the institution that granted the good or service and the frequency with which it receives said goods.

5.2.1.3 Challenges and issues for further research

545. Statistics including a valuation of social transfers in kind are clearly of use to those interested in measuring and understanding poverty, particularly when making cross-country comparisons where levels of in-kind services provided by the state may vary considerably. It is therefore useful for poverty analysts in national and international organisations to be able to consider such measures alongside the official poverty statistics, whether these are based on relative/absolute low (cash) income as in the United Kingdom, or another measure. For within-country comparisons, measuring the impact of STiK, along with that of indirect taxes, is also highly desirable for work aimed at examining, in full, the redistributive effect of the tax and benefits systems. However, analysts looking specifically at income poverty within, or between countries should remain aware that the choices and assumptions made in carrying out analysis of STiK are crucial. The outcomes of any analysis are likely to vary considerably depending on decisions made on factors such as:

- The social transfers in kind included/excluded from the analysis;
- Whether an actual consumption, insurance value, or flat rate approach is taken for valuation⁸⁶;
- Whether values are allocated at the household or individual level;
- Which equivalization scale is used; and
- What at-risk-of-poverty threshold is used.

546. Another issue in the consideration of social transfers in kind is the tendency for these transfers to be underreported in household surveys. Recent work at the U.S. Census Bureau uses administrative records to assess the extent of underreporting. For the most important nutritional assistance program (Supplemental Nutritional Assistance Program - SNAP, formerly known as Food Stamps), 46% of households identified as recipients in the administrative records did not report SNAP receipt on the survey (Stevens et al, 2018). Many researchers “correct” the survey data to bring reciprocity rates and benefit amounts up to administrative targets. The methodologies used to correct for underreporting include microsimulation models and regression-based modelling as well as matching survey data to administrative program data.

547. When estimates of the value of STiK are used in poverty measurement it is important to ensure there is the consistency between the thresholds used for poverty determination and the resource measure. For example, including the value of health benefits in resources is not appropriate if health expenditures are not included in the threshold calculation. For the supplemental poverty measure, the U.S. Bureau of Labor Statistics has done considerable research to impute the value of STiK to the expenditure data in order to make the thresholds consistent with the resources. These efforts have focused only on the benefits designed to meet the consumption items in the threshold (food, clothing, shelter and utilities) (Garner et al, 2016). If the poverty threshold is set as a percent of median income, the value of STiK should be included in the income measure before establishing the median if the value of STiK will be incorporated in the resource measure.

548. A related concern is the fungibility of STiK. A family may benefit from high expenditures for free public education but if their other resources are not sufficient to cover basic needs they should still be considered living in poverty. If the poverty threshold assumes that a family needs \$15,000 per year to cover basic housing costs, the value assigned to housing benefits should not exceed \$15,000. In simple terms, a person with a housing benefit “worth” \$20,000 would still be in poverty if his/her cash resources are not sufficient to purchase the other elements of the thresholds --- food and clothing. In the United States, for the supplemental poverty measure, housing benefits are capped at the housing portion of the threshold.

549. Finally, there are concerns about the appropriate equivalence scale to use when STiK are incorporated in the resource measure. Equivalence scales facilitate the comparison of incomes across heterogeneous households but the equivalence scales appropriate for making comparisons of disposable

⁸⁶ Balestra and Sustova (Balestra and Sustova, 2017) suggest such methodological differences are widespread, with 30% of countries producing estimates of social transfers in kind using the insurance value method, the same proportion using actual consumption, and 40% using a combination of the two.

income may not be appropriate for comparisons of an income concept extended to include STiK. Aaberge et al (Aaberge et al, 2017) use the OECD scale for disposable income but use an alternative scale for public services. The alternative scale is needs-adjusted and country-specific and relies on national mean public spending targeted to different population subgroups, defined by age and gender. They create a simplified need-adjusted scale that can be easily computed for any dataset with household information that includes the age of household members. This scale assigns higher weights to children and the elderly compared to the OECD scale for disposable income because children and the elderly have higher needs for basic public welfare services like education and healthcare.

550. The approach used in the U.S. for the supplemental poverty measure may reduce the need to develop alternative equivalence scales for poverty measurement. Rather than including the cost/value of basic public services (such as health care and education) in the thresholds and the resource measure, the SPM subtracts any out of pocket expenditures for these services from the resources. For example, payments for childcare services are subtracted along with other necessary work expenses before assigning poverty status. Out of pocket expenditures for healthcare are similarly excluded from the resource measure. Since the survey used to measure resources for the SPM asks specifically about these expenditures, the measure can use actual outlays rather than relying on average needs thus mitigating the need for equivalence scales. Public support for these services is recognized implicitly in the level of out of pocket expenditures. An individual with subsidized childcare should have lower out of pocket expenditures for this item than an individual not receiving this benefit.⁸⁷

551. Which STiK should be considered depends on what we are trying to measure and whether or not the transfer is universal. If our goal is to measure the distribution of the entire gross national product, we will want to include a broad array of government services in our resource measure. If our goal is to determine the share of gross national income going towards families with children, it would be very important to include spending on public education. On the other hand, if our goal is to measure the percent of the population with income below a set threshold incorporating the value of a universal benefit might not be useful. To take an extreme example, if we were to add the “value” of national defence spending to each person’s resources we would have to add an equal dollar amount to each person’s poverty threshold. There would be no change in the poverty rate.

⁸⁷ On the other hand, the family that cannot afford childcare may be erroneously categorized as not in poverty since they have no out of pocket expenditures for child care even though their need for child care may be unmet. Similarly, if there are two individuals with equal incomes but one has health insurance and the other does not, the person with health insurance may be considered poorer than the individual without health insurance because his/her resources will reflect any out of pocket payments for health insurance.

Recommendation 23:

- a) While the measurement of Social Transfers in Kind (STiK) continues to pose serious challenges, it is important to develop a mechanism to take them into account when estimating poverty and the impact of these transfers on poverty estimates. Supplemental or alternative poverty measures are important tools for illustrating the impact of these transfers on economic well-being. STiK can be particularly relevant for comparisons between different welfare systems, where STiK are more important than cash transfers in one country (or group) than another.
- b) Figures on total STiK should be presented together with poverty measures wherever possible as a useful indicator in its own right.
- c) STiK should be included in the measurement of poverty if their value can be empirically estimated on household or individual level with sufficient precision. Particularly relevant for poverty measurement are STiKs for food, shelter, clothing, and utilities. Some countries also make provisions for health care and education. If STiK are included in the resource measure this may affect the equivalence scale (see Recommendation 21 above).
- d) If poverty headcounts of relevant groups would change by 10% after STiK some consideration in the poverty measure is highly advisable. If however measurement is very poor or its effect on poverty profiles is within the margin of sampling error, STiK should not be included in poverty measures.
- e) Given the unavoidable and essentially arbitrary methodological choices regarding valuation and distribution of STiK, these need to be made fully transparent in regularly updated quality reports. In any case, users should be given the possibility to assess poverty measures with and without adjustments for STiK.
- f) STiK can be valued at equivalent insurance cost or actual consumption or as a mix. Its total value and estimated number of recipients need to be assessed against administrative data on the total public cost on STiK.
- g) Caution is needed when analysing STiK and their potential distributive consequences. If STiK are included in the resource measure, its value needs to be capped and by all means should not exceed the poverty threshold.
- h) If the value of STiK received is too difficult to obtain, the deduction of out-of-pocket expenses from the resource measure are a viable alternative. In such a situation however some poor individuals who have already curtailed certain expenditure may eventually appear as non-poor.

5.2.2 Accounting for housing wealth

552. Poverty measurement can be improved with a meaningful and standardised measure of resources available to the household to spend on its consumption.⁸⁸ When a household owns its home, it will tend to have lower shelter costs, and therefore have more resources to spend on other necessities. This is the logic for adjusting household income to take account of housing wealth for poverty measurement. Owners without mortgages are not the only households who may pay lower than market rents for their homes. Owners with mortgages may pay less than market rates if they have paid off their mortgage sufficiently.

553. In addition, some households live in subsidised housing – these households also pay less than market rates for their dwellings, and therefore, have more left to pay for other necessities than do comparable households in market rent apartments. Incorporating the value of their housing subsidy to their income could allow for a clearer representation of their poverty relative to other groups and for more transparent accounting of the benefit implied by the in-kind subsidy.

554. Therefore, reflection of housing wealth in income can lead to useful insights on well-being, poverty and inequality of the population. A well-recognised solution to this is the estimation of imputed rents (see Box 5.3). Under this approach, the incomes of homeowners are inflated to account for the rent that they would otherwise have to pay. While this approach may have a good theoretical basis and is analogous to the approach used within national accounts, it is challenging to implement in practice and not uniformly measured internationally. It can be also difficult to communicate to users of income inequality statistics. An OECD study of international practices determined that all but 3 out of 27 countries (Canada, Republic of Korea, and United States) calculated imputed rent regularly as part of their income programs, but often publish them as “secondary or alternative” measures of income or inequality. Countries also followed a variety of different practices in computing imputed rent (Balestra and Sustova, 2017).

Box 5.3

Approaches to the estimation of imputed rent

There are three estimation methods commonly in use in valuing imputed rent. The two main approaches are the rental equivalence (market rent) approach and the user cost (return to capital) approach. The third approach is the self-assessment method.

The objective of the rental equivalence approach is to estimate, using a statistical process, the market rent that would apply to the dwelling, were it to be rented out. The difference between this rent and the actual costs incurred by the household would be the imputed rent or, in the case of subsidized renters, the estimate of the value of the subsidy. The statistical processes used to estimate the market value of rent include the stratification approach, the hedonic regression approach and the Heckman approach. In each approach, the market value is arrived at by predicting rents of owner-occupied dwellings based upon the characteristics of dwellings occupied by renters. Likewise, once the rental

⁸⁸ An alternative approach adjusts the thresholds to reflect the differences by housing tenure. The United States Census Bureau’s Supplemental Poverty Measure uses three distinct thresholds: one for renters, one for owners with a mortgage and one for owners without a mortgage.

equivalence is determined for homeowners, it is necessary to deduct from this the cost of ownership. Costs of ownership could include property taxes, condominium fees, cost of maintenance and upkeep, some insurance costs and potentially other costs incurred by homeowners but not renters. What is included in the cost of ownership can vary from study to study, and data on these costs can vary from survey to survey.

In the user or unit cost approach, net imputed rent is derived through applying a rate of return to home equity. Thus, imputed rent is understood as the rate of return that would have been obtained by owners if the home equity had been invested in a safe investment (Törmälehto and Sauli, 2013; Balcasar et al, 2014). Difficulties in applying the user cost approach include being able to settle upon the current market value of the dwelling, as well as an appropriate rate of return (Balcasar et al, 2014). Garner and Verbrugge (Garner and Verbrugge, 2009) note a divergence between actual rents and user costs for the United States and conclude that the rental equivalence approach should be used in official statistics where possible. Eurostat guidelines indicate that the user cost method should be used only in cases where the rental market comprises a small share (less than 10%) of the total housing market. However, Törmälehto and Sauli, have suggested that the user cost method should be reconsidered due to its “better transparency, relative simplicity and lower production and respondent burden” (Törmälehto and Sauli, 2017: 156), and that “the quality of data about the current market prices of dwellings may be better than that of data about competitive market rents, even if market price values were derived from survey respondents” (Ibid: 156).

The self-assessment method involves asking owner-occupiers to estimate the potential market rent for their dwelling. Strengths and weaknesses of this approach are outlined in Balacasar et al (Balacasar et al, 2014) who conclude with concerns about using owner-assessments of rental value, as owner-occupiers may overestimate the true rental value of their property given their affinity to their property or neighbourhood. This approach is used in the United States based on a question in the Consumer Expenditure Survey, “If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?”

555. An alternative approach to consider is an after-housing costs measure introduced within the United Kingdom’s Department for Work and Pensions (DWP) annual article – HBAI. HBAI reports distributional measures of income that net a range of housing costs including rent and mortgage repayments, and potentially offers an approach to compare the incomes of different groups of the population that is easier for data users to understand. Similar “shelter poverty” measures have been developed for the United States by researchers at the U.S. Department of Housing and Urban Development.⁸⁹

556. A third approach adjusts resources (or the thresholds) to reflect the differences between homeowners with a mortgage and those without a mortgage. The Canadian Market Basket Measure of low income (Employment and Social Development Canada, 2018) makes this type of adjustment. The threshold based upon the costs of food, shelter, clothing transportation and other items for 50 regions of Canada. In the resource measure, a value for “Mortgage Free Advantage” is added in which is conceptually a value for imputed rent for mortgage free homeowners. The U.S. Supplemental Poverty Measure also adjusts the thresholds by tenure status.

⁸⁹ See for example Pelletiere, 2008.

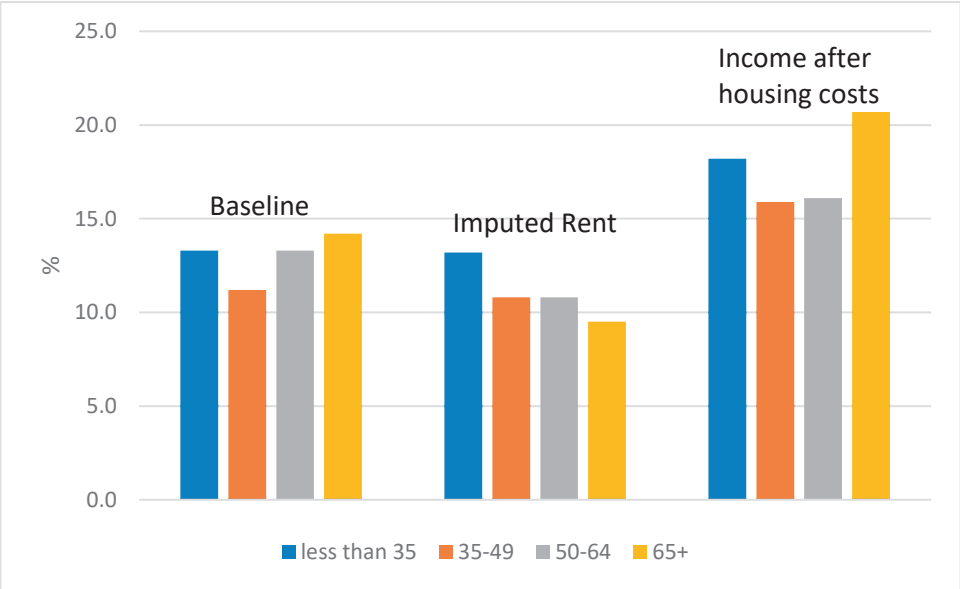
5.2.2.1 Impact of imputed rent on Canadian poverty estimates

557. Using the Canadian Income Survey and the Survey of Financial Security, Heisz and Phipps-Burton (Heisz and Phipps-Burton, 2018) have estimated low-income rates for Canada using three methods to account for the value of housing wealth: rental equivalence, user (or unit) cost and subtracting housing costs from income. Each approach required a number of assumptions, limitations and imputations. For example, for rental equivalence only interest payments on mortgages, property taxes paid and condominium fees were subtracted. The value of interest paid is estimated using a model to predict interest payments as a share of total mortgage payments. For the user costs approach, estimates were done with both a 2% rate of return and a 3% rate of return. In this study, low income (poverty) was defined as having household income below 50% of median equivalised household income.

558. The authors found that the results varied across the different methodologies. Using the rental equivalence approach, the poverty rate of owners without a mortgage fell from 9.4% to 4.0% (4.3% with the Heckman selection correction). With the unit/user cost approach, the poverty rate for this group fell from 7.2% to 3.5% or 4.0% depending on the assumed rate of return. Subtracting housing costs from income increased poverty rates for homeowners without a mortgage from 9.4% to 10.9%. For owners with a mortgage the rental equivalence approach decreased poverty rates from 5.4% to 4.6% while the unit cost approach decreased poverty rates from 5.2% to 4.3% or 3.7% depending on the rate of return assumptions. Subtracting housing costs from income increased poverty rates for this group from 5.4% to 8.4%.

559. Using rental equivalence, poverty rates for older households decreased from a baseline of 14.2% to 9.6%. Using the “income less housing costs” approach the poverty rate for older households increased from 14.2% to 20.7%.

Figure 5.9
Low-income rates by age of household head, Canada 2016



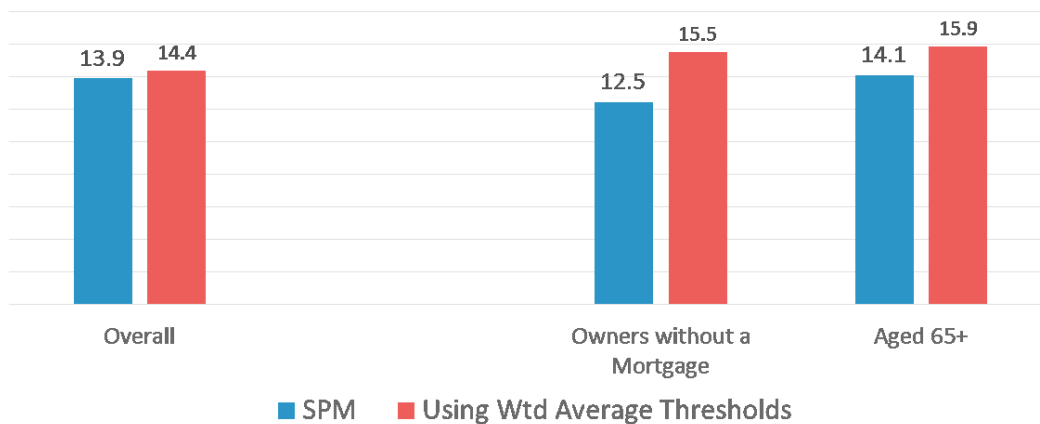
Source: Calculations using the Canadian Income Survey, 2016.

5.2.2.2 Impact of Differential Thresholds on the Supplemental Poverty Measure for the U.S

560. The Supplemental Poverty Measure in the United States tackles this issue by using three different thresholds: one for homeowners with a mortgage, one for homeowners without a mortgage and one for renters. The housing portion of the threshold is estimated using actual expenditures of each type of household. The difference between the threshold for owners without a mortgage and the one for owners with a mortgage is a rough estimate of imputed rent, albeit imperfect in that it treats all homeowners with mortgages the same, regardless of the terms of their mortgages. The differential thresholds decrease 2017 poverty rates for owners without a mortgage 3 percentage points, from 15.5% to 12.5%. Poverty rates for individuals aged 65 or older decrease from 15.9% to 14.1%.

Figure 5.10

Impact of Differential Thresholds on the U.S. Supplemental Poverty Measure, 2017



Source: U.S. Census Bureau, Current Population Survey, 2018 Annual Social and Economic Supplement.

Note: This figure meets all of the U.S. Census Bureau's Disclosure Review Board (DRB) standards and has been assigned DRB approval number CBDRB-FY19-ROSS-B0032. Renwick (2018).

5.2.2.3 Challenges and Issues for Further Research

561. While standardization of methods to deal with housing wealth may be desirable, this is difficult due to different data inputs being available in different countries. Given these limitations, methodologies used in estimation should be clearly described in metadata and made readily available to assist in the interpretation of the results. Because of the sensitivity of results to the methods chosen, and because there is no consensus on whether to include imputed income in poverty measures, poverty estimates including imputed rent or other methods of accounting for housing costs in the income distribution should be presented separately from the standard estimates.

Recommendation 24:

- a) Given that the choice of method for obtaining monetary poverty measures may especially depend on the available data and may contain essentially arbitrary elements, methodological choices need to be made fully transparent in regularly updated quality reports. In any case, users should be given the possibility to assess poverty measures with and without adjustments for imputed rent.
- b) As an alternative to imputing rent, residual income after actual housing cost may be considered as a resource measure which reflects housing equity.
- c) Surveys should consider adding questions to measure housing costs and property values and other characteristics of the dwelling, which facilitate the calculation of imputed rent.

5.2.3 Asset poverty: Why assets should be considered in the measurement of poverty and how it could be done

562. In research, the joint consideration of income and wealth has been recognized as important for measuring poverty (e.g. Weisbrod and Hansen, 1968; Caner and Wolff, 2004; Stiglitz et al, 2009; Brandolini et al, 2010; Azpitarte, 2012). However, up until now, in most national poverty statistics, income from assets and rentals (interest, dividends, rental income etc.) are considered but not the assets themselves. The official measurement of poverty is therefore usually limited to the measurement of income poverty. This is primarily because statistical offices lack the required data but also because suitable international concepts for integrating assets into the measurement of poverty are still under development. Thanks to the European Central Bank's Household Finance and Consumption Survey (HFCS), the availability of data in many European countries has been improved in past years.⁹⁰

563. There are several possibilities to include assets in poverty measurement, and results may vary considerably according to conceptual choices. In the literature, two main variants can be identified in order to incorporate assets into poverty measurement:

- The unidimensional approach: Assets are directly integrated into income so that a single indicator is generated. This can, for example, be achieved through an annuitization of assets (e.g. Weisbrod and Hansen 1968; Haveman and Wolff 2004; Brandolini et al, 2010), or through the application of asset limits above which a person is by definition classified as not poor (Headey et al, 2009).

⁹⁰ However, for most of the countries covered in the HFCS, income is still recorded only as gross income (with the exception of Italy and Finland, where disposable income is used). The Eurostat-OECD Expert Group on measuring the joint distribution of income, consumption and wealth has thus made efforts to perform statistical matching between different statistical sources (e.g. HFCS and EU-SILC).

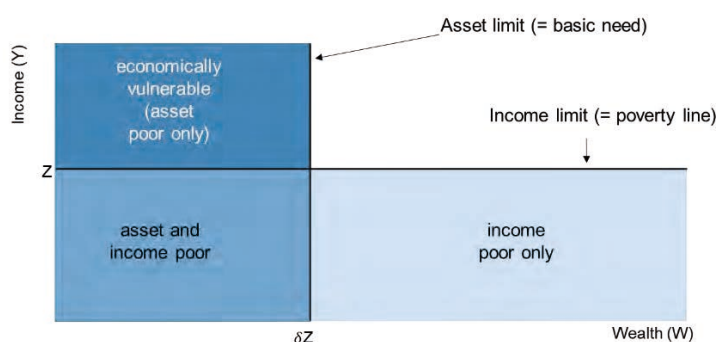
- The two-dimensional approach: A separate asset-based poverty rate is calculated in addition to the income poverty rate (Brandolini et al, 2010). The income poverty rate remains unchanged. This approach is applied, for example, by the OECD (Balestra and Tonkin, 2018) and Eurostat (Eurostat, 2017b) and is also described by the UNECE (UNECE, 2017a).

564. It has already been shown several times that the choice of approach (uni- or two-dimensional approach) has a considerable influence on the results (e.g. Kuypers and Marx, 2016; Azpitarte, 2010; Brandolini et al, 2010; Tonkin et al, 2016; Törmälehto et al, 2013). Ravazzini et al (Ravazzini et al, 2017) have also shown for Switzerland that the unidimensional approach leads to higher percentages of asset poor than the two-dimensional approach.

5.2.3.1 Asset poverty in Switzerland

565. Empirical illustrations for asset poverty are in this Guide referring mainly to Switzerland, but this topic can be assumed to be especially pertinent for all high-income countries. Since asset data are not available in the CH-SILC every year, the two-dimensional approach is better suited to official statistics in Switzerland. An asset-based poverty rate is calculated separately from the income poverty rate but both calculations use the same poverty line. By combining the income and asset-based poverty rates, it is possible to distinguish between income poor people who are (still) able to compensate their low income by using assets (the income poor only) and income poor people who do not have this possibility (the asset and income poor). Furthermore, persons who are not income poor but do not own sufficient assets to prevent them from ending up in a precarious situation in the event of loss of income (the economically vulnerable) can also be identified. Schematically, the combination of both poverty rates is shown as follows (see Figure 5.11):

Figure 5.11
Illustration of the two-dimensional poverty index



Source: Balestra and Tonkin, 2018.

566. For the operationalization of the asset-based poverty rate, various choices still need to be made: which wealth concept will be considered, what type of poverty line is used, for which reference period the assets have to suffice, and how the assets are to be equivalized.

- Wealth concept: For measuring asset poverty, net worth is often used as a starting point, i.e. all assets minus all liabilities. Pursuant to the question in the CH-SILC, liquid assets are defined here

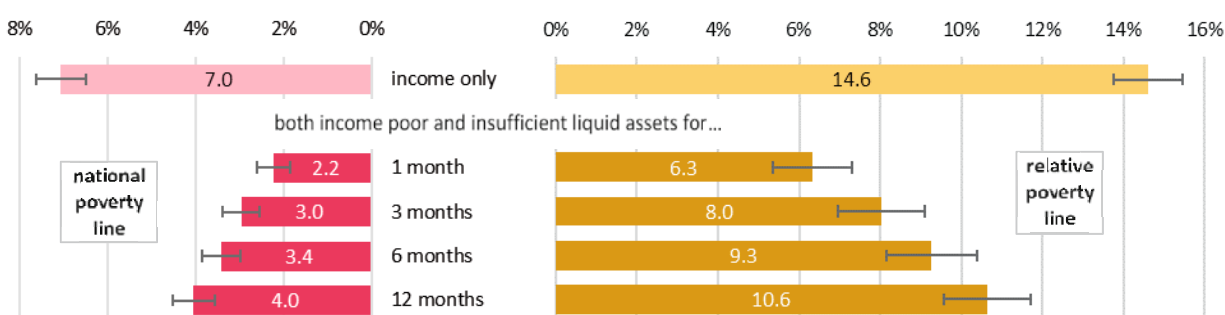
as the household's total assets in bank and postal accounts and the estimated total value of shares, bonds and investment funds.

- **Poverty line:** The Swiss FSO publishes income poverty rates based on various poverty concepts. As in the other SILC countries, a relative poverty line set at 60% of the median equivalized disposable income. A national poverty line is calculated set at the social minimum subsistence level that serves as the basis for measuring public social assistance benefits in Switzerland.
- **Reference period:** The proportion of the asset-based poor also depends on the time span for which the assets should suffice. The results vary considerably with the chosen reference period. This analysis shows asset-based poverty rates for one, three, six and twelve months.
- **Equivalence scale:** While there is no general agreement as to whether the use of an equivalence scale is appropriate for assets, since in this analysis assets are seen as an income substitute the relative poverty line has been adjusted using the modified OECD equivalence scale. For the national poverty line, as amounts are classified by household size, there is no need for an additional equivalence adjustment of assets.

567. The share of persons who have neither sufficient income nor sufficient assets to cover their basic needs for a certain time span varies considerably according to the poverty line and the reference period chosen (see Figure 5.12). With the relative poverty line, the shares are much higher than with the national poverty rate. However, this reflects mainly the differences of the income poverty rates since the relative poverty rate is also around twice as high as the national poverty rate (14.6% vs. 7.0%). The share of the income and asset poor increases when a longer reference period is applied. With a reference period of 12 months, assets need to be 12 times as high to be considered sufficient than when a reference period of 1 month is chosen.

Figure 5.12

Share of income poor that do not have sufficient liquid assets, by poverty line and reference period, in percent of the population (individuals), with 95%-confidence interval



Source: FSO – CH-SILC 2015 (provisional data on wealth, version 07.06.2018).

Notes: Only liquid assets such as deposits, bonds, shares, mutual funds etc. are taken into account. The national poverty line consists of a fixed amount for living expenses, individual housing costs and CHF 100 per month and per person aged 16 or over. The relative poverty line is calculated as 60% of median equivalised income including imputed rent.

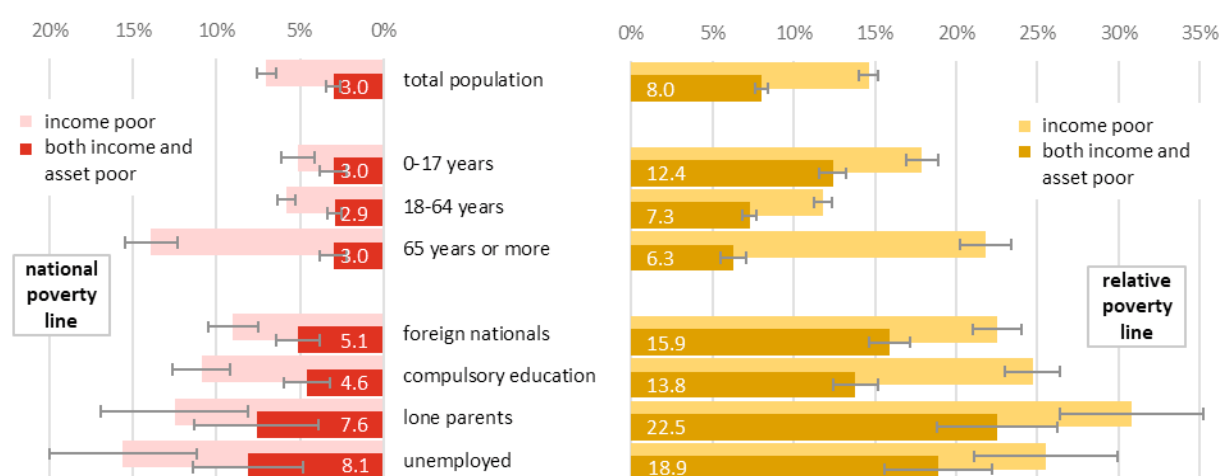
568. For both poverty lines, the inclusion of assets especially reduces the poverty rates of the older population compared with income poverty, meaning that the income poor aged 65 years and over may

fall back on liquid assets particularly often. This finding is in line with many other studies on asset poverty (e.g. Kuypers and Marx, 2016; Azpitarte, 2012; Brandolini et al, 2010; Caner and Wolff, 2004; Balestra and Tonkin, 2018). This also influences analyses by employment status: while old-age pensioners and the unemployed have comparable income poverty rates, the unemployed are more than twice as often income and asset poor than pensioners. Furthermore, foreign nationals, persons with no post-compulsory training, and lone parents with children under the age of 18 are more often at the same time income and asset poor than their respective reference groups.

569. However, differences between the two poverty lines are noticeable for children and young people under the age of 18. When using the relative poverty line, children are considerably more often income and asset poor than the other age groups, while there is no major difference by age when the national poverty line is applied. This reflects again the structure of the respective income poverty rates and is mainly due to the fact that the national poverty line rises less sharply with increasing household size than the relative poverty line.⁹¹ Children typically live in larger households. For the relative poverty line, households with children accordingly have higher income and asset poverty rates than comparable households without children, whereas for the national poverty line, the differences between households with and without children are smaller.

Figure 5.13

Share of income poor that do not have sufficient liquid assets for three months, by subgroups, in percent of the population (individuals), with 95%-confidence interval



Source: FSO – CH-SILC 2015 (provisional data on wealth, version 07.06.2018)

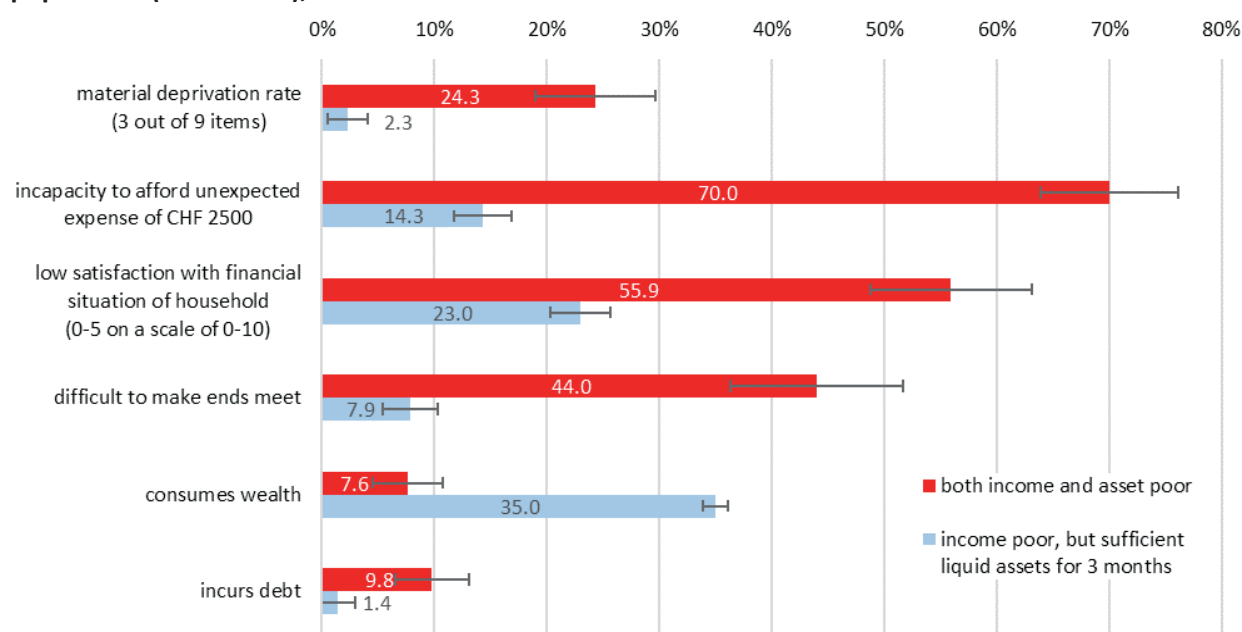
Notes: Only liquid assets such as deposits, bonds, shares, mutual funds etc. are taken into account. The national poverty line consists of a fixed amount for living expenses, individual housing costs and CHF 100 per month and per person aged 16 or over. The relative poverty line is calculated as 60% of median equivalised income including imputed rent.

⁹¹ The difference in the increase of both poverty lines is mainly due to the fact that the national poverty line integrates individual living costs (see Guggisberg et al, 2013). Housing costs do not equally increase with additional persons in the household and are also relatively high even in single person households.

570. The correlation between income poverty rates and other indicators on the standard of living is generally quite low (see e.g. Headey et al, 2009). This changes greatly when assets are taken into consideration: the possession or non-possession of assets divide the income poor population into two groups with clearly different standards of living. The income poor are much less exposed to material or subjective difficulties⁹² if they have assets in reserve compared with those who do not have any such resources to fall back on (see Figure 5.14).

Figure 5.14

Share of population indicating material or subjective difficulties, by poverty status, in percent of the population (individuals), with 95%-confidence interval



Source: FSO – CH-SILC 2015 (provisional data on wealth, version 07.06.2018)

Notes: Results are based on the national poverty line and a reference period of 3 months. Only liquid assets such as deposits, bonds, shares, mutual funds etc. are taken into account. The national poverty line consists of a fixed amount for living expenses, individual housing costs and CHF 100 per month and per person aged 16 or over.

571. The same pattern can be found for the subjective indicators: the income and asset poor are most dissatisfied with their household's financial situation and most often struggle to make ends meet. Together with the economically vulnerable (i.e. people who are asset poor but not income poor), they face the most frequent difficulties in meeting unexpected financial expenses and get into debt most often to cover ongoing household expenses. Despite their low income, income poor with sufficient assets are usually even better off than the population as a whole. As expected, this group also includes the highest

⁹² Material deprivation is defined as “the enforced inability [...] to pay unexpected expenses, afford a one-week annual holiday away from home, a meal involving meat, chicken or fish every second day, the adequate heating of a dwelling, durable goods like a washing machine, colour television, telephone or car, being confronted with payment arrears” (see https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Material_deprivation)

proportion of people who use their assets to finance their ongoing expenses. Thus, the contradiction between the high-income poverty rates of the older population and the positive evaluation of their subjective situation may largely be explained by the fact that this age group very often has financial resources in the form of assets.

572. For all indicators, the second worst off group are the economically vulnerable. Assets therefore appear to have at least as much correlation with the variables used as income: a financial buffer has a greater influence on material living standards than the current income situation and also seems to make a greater contribution to satisfaction. Headey et al (Headey et al, 2009) also came to a similar conclusion: "It may be noted that some recent research has suggested that in several countries, including Australia and Germany, wealth has as much if not more impact on life satisfaction than income". For all indicators considered, those with sufficient income and reserves for at least 12 months are best placed. This can be seen as a strong indication that the combination of income and assets is of great relevance to the standard of living and, consequently, both should be included in the measurement of poverty.

5.2.3.2 Challenges and issues for future research

573. While these findings show that assets should be considered in measuring poverty, the communication of these results poses a major challenge for NSOs, which primarily address the general public. Feedback from users shows that there is already today some uncertainty as to which poverty rate (national poverty rate or relative poverty rate) is most suitable for a certain purpose. The introduction of yet another poverty rate could exacerbate such uncertainties. By introducing the asset-based poverty rate, the income poverty rate may furthermore lose some legitimacy, because in public debate poverty is usually equated with a lack of income and assets, which are also included in the assessment of eligibility for social welfare support. As long as asset data is collected only occasionally, the income poverty rate must be likely to remain the main poverty indicator among high income countries.

574. At the political level, the result that a considerable proportion of the income poor own certain assets could be misinterpreted as though they are actually "wealthy" and further fuel debate on the welfare state. Therefore, it is important to clarify that the focus of asset-based poverty is not on larger asset reserves but on relatively modest amounts that would be used up in a few months without any income. To avoid any misunderstanding, the terminology should therefore be chosen with care and should describe as clearly as possible what is really measured with the new indicator. In this way, the asset-based poverty rate, for example, could be described as the share of people without sufficient reserves for x months whereby x corresponds to the chosen reference period. The term poverty should remain limited to income poverty.

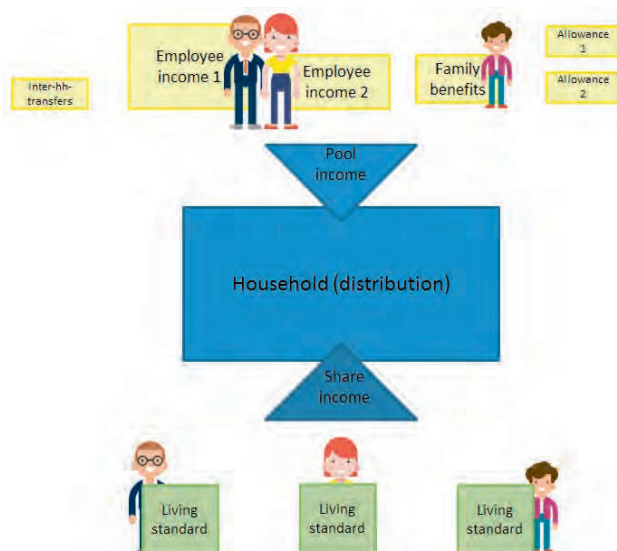
Recommendation 25: Assets are an important resource for determining well-being and countries should continue to experiment with ways to measure both income and asset poverty. At this point, the two-dimensional approach with a separate asset-based poverty rate calculated in addition to income poverty seems to be the easiest to communicate to the public.

5.2.4 What is the role of unequal sharing of resources within households?

575. In measuring poverty, are persons in a household by definition treated equally or should each person be weighted by what he or she actually contributes to the household income or takes from the household income for spending? The question of intra-household distribution is usually ignored by standard poverty measures although the problem has been raised and named an “agenda for action” (Jenkins, 1991). However, so far pooling of income has been implied when measuring a household’s and its members’ living standards by the European concepts for poverty indicators.⁹³ Therefore we speak of the “standard (EU-indicator) approach” when no difference between the household members is made relative to their input of income or their achieved living standard. In the United States, poverty is measured at the family (related by birth, marriage or adoption) rather than the household level but equal sharing of resources within each family is assumed. For unrelated individuals sharing a household, poverty status is based on personal income and a one-person poverty threshold.

Figure 5.15

The standard household income model – exemplary household



576. Given these implicit assumptions, to be precise, we should then not be speaking of persons (or men/women/children) being income poor but of persons (or men/women/children) living in households/families that are income poor (World Bank, 2017: 44). The break-down of the total population in poverty by individual-level characteristics (e.g. by sex) will be distorted in cases where the assumption

⁹³ See also Canberra Handbook (UNECE, 2011): “While income is usually received by individuals, it is normally shared with other household members present e.g. spouse and children.” (p. 9) and “This distribution reflects the assumption that household income is shared equally between all members of the household, and does not reflect the direct receipt of income by individuals. Because many household members receive no money income, e.g. younger children, such an assumption is hard to avoid in practice.” (p. 28).

of equal sharing does not hold and where resources are not pooled for equal benefit and well-being of all household members. In these cases, gender differences, for example, can accurately only be reported for single person households/families.

577. It is therefore interesting to test the assumptions of equal sharing of resources within households and explore how much the assumption of equal sharing hinders the accuracy of poverty measures if disaggregated by social characteristics, e.g. by sex. The aim is to make gender differences visibly that might be covered by the traditional poverty measurement concept. The forthcoming UNECE *Recommendations for Measuring Intra-household Power and Decision-making* includes (UNECE, 2020), among others, an analysis on identifying gaps for which there are currently no adequate indicators of gendered dimensions of intrahousehold power and decision making, including in the area of poverty measurement, and makes proposals to fill such gaps. Core and supporting indicators are proposed, and consideration is given to the specificities of each dimension such as appropriate survey vehicles, and modes, necessary contextual questions, and areas where cultural factors would make items especially relevant or irrelevant in some countries. The “black box” household has to be cracked opened – at least as far as data availability on within household distributions allows for it. This approach poses practical difficulties especially in cases of common social transfers and in the treatment of children and dependent persons in the household.

578. Literature on the assumption of sharing of resources and its implications for economic theory on the household is extensive.⁹⁴ “Pooling of income” addresses the input side of household resources – who brings what amount of money into the household for common or exclusive use -, “sharing of income” addresses the outcome side – what can the household and its members afford. In general, we can either think of the household as an entity with its members acting only as one unit (favoured by economic theory in 1950s until the 1980s) known as the “unitary approach to economic behaviour” where incomes are pooled and shared; or we think of the household members as each trying to maximize his/her own utility function in a “non-unitary model”. This second approach, pioneered by Manser-Brown and McElroy-Horney in the early 1980s, was to model family demands as a cooperative bargaining game. Then there were also other non-unitary approaches, that used either collective models (e.g. Chiappori, 1992) or non-cooperative models.

579. Looking at practically oriented studies and empirical evidence in regard to measuring well-being and poverty, we hope to find some answers to the following two questions:

- 1) When leaving the pooling of resources assumption aside what models can we use instead? How – as social statisticians dealing with household income data – can we arrive at models that adequately implement the degree of (non-)pooling and sharing?
- 2) How big is the difference between both approaches, i.e. how much bias is introduced by acting upon the notion of the pooling of resources assumption vs. a “new” concept of pooling/non-pooling ratio.

580. Survey questions which would provide empirical information for the analysis of unequal consumption patterns within households are rarely found in the literature (Lechene et al 2019, Bargain,

⁹⁴ For a very good overview on potential indicators of intra-household decision-making, related methodological challenges and data availability see Krieger, 2018.

O. and L. Martinoty 2018). Using such information in EU-SILC data Ponthieux (2017) was able to calculate the “at-risk-of-poverty rate” for women and men that live in couple households (with or without children, married or cohabitant). The question which was used asks: “What proportion of your personal income do you keep separate from the common household budget?” Answers are categorical with six choices:

1. All my personal income;
2. More than half of my personal income;
3. About half of my personal income;
4. Less than half of my personal income;
5. None;
6. The respondent has no personal income.

581. This question was included in the ad-hoc module of the EU-SILC survey in 2010 on “Intra-household sharing of resources”. Ponthieux used this information to assess the degree of income pooling and estimate the amount of income which is not pooled. She calculated a modified income in which only the pooled income sources are equivalised using the Eurostat-scale. Personal incomes which are kept separate were added to this equivalised income without equivalisation to obtain an individual level welfare measure. In this modified approach income values can be different for each member of the household when some part of income is kept separately. For example, in the modified approach, children without personal income only get the equivalised pooled income. Husbands and wives which have personal income sources from which they keep a part separate, get the equivalised pooled income plus the unequivalised amount of their income which is not pooled. If no part of income is kept separate the incomes are the same for all household members just as in the conventional approach.

582. Two limitations need to be considered. First, only some income components are available on personal level (employee income, pensions, some social transfers), while others are reported at the household level (inter-household transfers, some social transfers). Second, the conceptual differences between pooling and sharing: “...no pooling does not mean no sharing; conversely, income pooling does not necessarily entail equal or fair sharing.” (p. 187). When we aim to know about the living standards of an individual the knowledge of the part of his/her income that is kept separately is only a proxy for that question of interest. The model proposed by Ponthieux builds on the assumption that income that is kept separate only increases the living standards of the income recipient. This assumption may be violated if for example mothers spend a larger part of their resources on children.

583. The modified approach reveals that men control more than half of the total household income in all Member States of the EU with the exception of Lithuania. The differences between men and women are rather small (biggest in Malta with about 42% for women vs. 58% for men). Better earnings by one (the male) partner appear to be nearly perfectly counterbalanced by the couples’ pooling regimes (p. 184f.) in most countries. Age differences were not considered in the study.

584. Using the same “modified” income and comparing it to the poverty threshold calculated by the standard methodology (60% of median for the targeted couples) Ponthieux shows a higher “modified” poverty risk for women as compared to men. The study demonstrates that a different model of within

household income distribution can be calculated when treating personal and household level income differently and – if available like here – additional use can be made of a question on the share of the personal income that is used separately for the person him/herself.

585. Guio and Van den Bosch (Guio and Van den Bosch, 2018) put their focus in a recent paper on variables of material deprivation: Do men and women in couples have the same material and financial possibilities, the same risk to be deprived? Their paper is the first to present empirical evidence on this issue for a range of EU countries, using the 2015 wave of EU-SILC. In contrast to the conventional EU-SILC questionnaire which measured deprivation only for the whole household, this wave introduced a number of questions on deprivation which are asked separately for all adults such as whether an individual respondent could afford to have regular leisure activities. The extent to which answers to such questions differed across household members was taken as a measure of intra-couple inequality in deprivation and used to analyse important determinants. Findings point to a higher deprivation rate of women compared to men living together as couples. The only deprivation item which did not exhibit significant inequality within households was access to internet. Therefore, one important finding here is that measuring deprivation through individual rather than only household items (“Can you afford?” versus “Can your household afford?”) unveils some gender differences that otherwise are lost in the black box of the household. So, to our first question we would answer that – when we are able to leave the narrower concept of household income and look deprivation – we can find ways to ask all persons individually on what they have available, can or cannot afford. Since not all goods of consumption are for the sole use of one individual (e.g. heating is of common use whereas clothing is not, some like a family car are in between) we also face the challenge to derive the “sharing rule”. Guio and Van den Bosch summarize the economic literature in this field in their article.

586. On the second question as to what differences this approach of surveying individual rather than household deprivation makes the authors find the following answers: There are also couples where the men are more strongly deprived than women; however, when aggregating the individual items into a deprivation scale they find more couples where the number of enforced lacks is higher for the woman (9.2%) than where the man is disadvantaged (6.5%). The work status of the partners and their relative contribution to the joint income are important determinants of the intra-couple gender deprivation gap.

587. Recent research by Burchardt and Karagiannaki (Burchardt and Karagiannaki, 2018) of the LSE also deals with unequal deprivation within households with the aim to overcome restrictions of conventional indicators on the individual level. Their project is also using micro-data from EU-SILC to examine the sensitivity of poverty, deprivation and inequality estimates across European countries to different assumptions about the intra-household sharing of resources in complex households, and to identify the groups of people for whom intra-household inequality may have the largest impact. Provisional findings say that especially in multi-family households, an individual’s contribution to total household income makes a difference on the living standard. In their cross-cultural study they add the economic, social and cultural dimension to intra-household decisions and their impact.

5.2.4.1 Impact of pooling assumptions on Austrian poverty rates

588. Lamei and Skina-Tabue (Lamei and Skina-Tabue, 2018) have applied the approach proposed by Ponthieux (Ponthieux, 2017) to the Austrian EU-SILC data. As their reference population are couples with or without children, they have recalculated the 60% of the median poverty threshold for this specific subgroup. The first line in Table 5.2 shows the poverty risk of women and men under the standard assumption of total pooling. The second line shows results for the alternative approach. In the standard approach the risk-of-poverty is the same for both partners. In the alternative calculation, the poverty risk of men appears slightly reduced while the risk for women is drastically higher. The gender ratio of the modified poverty risk (women/men) amounts to 1.9.

Table 5.2

At-risk-of-poverty by sex

	At-risk-of poverty in %	
	Men	Women
Standard approach (adjusted threshold for couple households)	9.8	9.8
Personal equivalized income approach	9.2	17.2

Source: Statistics Austria, EU-SILC 2010.

Note: Couple households in the dataset of the EU-SILC module 2010 (with or without children).

589. The approach presented opens new possibilities to assess gender inequalities within the household. However, the analysis is currently restricted to heterosexual couples and does not consider households with more than two adults. It also does not satisfactorily assess the poverty risk for children, as it is not foreseen that they “receive” a share of the parents’ personal income additional to the equivalised part. Details concerning the accumulation of income components on the individual or the household level (e.g. alimonies received by only one partner) are open to discussion as well to get to a more precise estimation of the personal versus the household income. Last but not least this methodology is at the moment restricted to EU-SILC data, more precisely to the dataset of the 2010 module. The knowledge of the pooling system applied in the household is crucial to the calculations, to make this approach universally applicable an equivalent to this has to be developed.

590. The analysis presented by Lamei and Skina-Tabue also includes a comparison with the more radical version proposed by Heuberger and Knittler (Heuberger and Knittler, 2018) which does not assume any pooling at all. The approach was developed for the specific context of measuring in-work poverty. This approach compares (unequalised) personal incomes and shared incomes to the specific poverty threshold in which the respondent lives with the required income per adult to meet the usual 60% of median threshold. This threshold marks the minimum personal income for an adult to avoid poverty in his or her specific household constellation. The measure of individual personal income includes net earnings and social transfers plus the income which cannot be attributed to individual members divided by the number of adults. Resources for children in the household are not considered. When the individual income is lower than the household specific poverty threshold the person is considered to be at-risk-of-poverty. The results reveal a much higher in-work-poverty risk for women than for men, whereas in the

standard approach the risk of in-work-poverty is approximately the same for both sexes (Heuberger and Knittler, 2018: 240). These results point to the hidden financial dependencies within households. In the calculations from Lamei and Skina Tabue (2018) the gender ratio (women/men) of poverty rates among couples amounts to 13.5, compared to 1.9 for the calculations according to Ponthieux.

5.2.4.2 Challenges and issues for further research

591. While measuring poverty at the household level has many advantages, analysis can be enhanced if concerns about poverty and deprivation at the individual level are taken into account in the design phases of questionnaires. Surveys should attempt to collect material living standards at individual level and data on the distribution of income within households.

592. Recent changes to the EU-SILC, are a good example going in this direction of measuring absolute living standards on a personal and household level: individual items to measure material deprivation on the personal level for all adults have gradually been introduced since 2009 to accompany those established indicators on the household level. The EU material and social deprivation indicator based on the new items has been approved in 2017 by the Indicator Sub-Group of the EU Social Protection Committee. It includes five items at personal level retained after a thorough study⁹⁵. The next step in EU-SILC was to consider also the special situation of children (see section 5.2.5). While there has been progress on the design and validation of child specific items (e.g. outdoor leisure equipment, go on school trips), there continue to be measurement issues. In March 2018, a new indicator on child deprivation (the child deprivation rate), based on the child deprivation module of 2014 was agreed at the EU level and will be included in the portfolio of social indicators. However, here we still face some of the measurement problems. If in a given household at least one child lacks an item as surveyed from their parents, it is then assumed that all the children belonging to that household lack that item. Thus, the specific situation of boys and girls or by age of the children is lost due to the question design. We would therefore further recommend to test if at least some of the child specific deprivation items can be measured at the individual level.

593. In order to have an empirical basis to disaggregate household income and assume an individual living standards of household members we have seen that it is useful to have some measure that gives insight into the “black box” of the household and shows how individual income components are treated. In addition, it would be good to know how common household income components (e.g. family transfers) are split. Here much can be done on a modelling basis, but the more empirical evidence there is and the less you have to work with assumptions the better. It is in our view not necessary to have very detailed information on how families treat their household income in every data collection – although it is a very personal decision and by modelling, we are sure to lose some of the variance that exists in reality. But it could also be a possibility to further test and develop several pooling/splitting assumptions with special data collections dedicated to that topic and apply them to general income surveys where necessary.

⁹⁵ Measuring material deprivation in the EU: Indicators for the whole population and child-specific indicators - <http://ec.europa.eu/eurostat/documents/3888793/5853037/KS-RA-12-018-EN.PDF>. The Annex gives an overview of the items and their guidelines as used in EU-SILC.

594. When a data collection is fortunate enough to have both, information on income and its distribution as well as information on individual material living situations of household members, one is able to compare and validate the assumptions of the income distribution model by using material living standards items. However, for individual households a difference in deprivation between the partners might go together with equal distribution of resources or, on the other hand, unequal income resources might not necessarily mean that one partner is more deprived than the other (Guio and Van den Bosch, 2018: 6).

595. Besides income and material items, any other dimension of potential inequality could be explored as regards the sensitivity to gender differences – e.g. wealth including the ownership of a household's dwelling (and the impact of adding/leaving out imputed rent in the household income calculation). Other factors of well-being and living conditions like time use, satisfaction, personal networks and psychological measures are usually gathered for individuals and are therefore not so problematic for profiles of living standards. For the situation of children, however, we often lack individual data and our recommendation that is most in line with practical solutions is to approach their standard of living with material items and affordability questions.

596. Also, for both strands – the individual absolute living standards and the individual income drawn from the household income – there are for sure some measurement issues worth exploring. The issue of interviewing all the relevant individuals in the household – not only one household respondent – is a basic starting point; it may be regarded as costly but collecting information at the individual level is only the way to get valid data on the individual situation. As summed up by Krieger (Krieger, 2018) many studies have pointed to the difference that can be made by the interview situation, e.g. if a partner is present during the interview or not. Also, the issues of proxies, missing information of one partner, mode effects, interviewer effects (e.g. the interviewer's gender) could be explored. Questionnaire design issues are also of uttermost importance to guarantee a common understanding of questions and answer options between respondents.

597. These recommendations are in line with World Bank Recommendation 13 (World Bank, 2017: 114) to disaggregate poverty figures by sex and age, and with the UNECE (2020) Recommendations for measuring intra-household power and decision-making.

598. While this discussion emphasised gender inequalities within households, other aspects, such as the situation of children within the household, would also be worth exploring. It might be interesting to distinguish different household types by age and number of children and/or adults. It would be worthwhile to test the hypotheses some studies have already confirmed that children appear to do better when their mothers control a larger fraction of the family income (Thomas, 1990; Hoddinott and Haddad, 1991).

599. Lastly, the concepts of the household, family and dwelling have practical implications. Another relevant future topic is the measurement of the living standards of children who belong to more than one household because their parents live apart.

Recommendation 26:

- a) In the immediate term, official poverty rates, collected at the household level, should be routinely disaggregated by sex and age of each household member. This shall provide estimates for the poverty rate of prime age women who have children and/or live without a partner (e.g. in lone parent households).
- b) More research is needed to determine how best to address intra-household resource sharing in order to develop sub-household measures of poverty status. This is particularly important for the disaggregation of poverty estimates by sex and age. Surveys should continue to experiment with questions designed to determine individual control of resources and to measure material deprivation at the person level.
- c) Where possible, official poverty rates for men, women, children or other socio-demographic subgroups should be accompanied by results that consider unequal sharing of resources. At a minimum, national statistical offices should carry out sensitivity analysis for poverty profiles contrasting the conventional full pooling assumption with partial pooling and full separation of resources.
- d) To validate assumptions on within-household distribution of economic resources, these should be considered in combination with material living standards, wherever possible. Questions on sharing of personal economic resources and/or personal material living standard have been used in EU–SILC. Before such questions are adapted it should however be ensured by pre-tests that sharing of resources is not considered as sensitive by respondents.

5.2.5 Going beyond income: The role of multidimensional poverty measures

600. Measures of monetary poverty are justified for their correlation with the fulfilment of needs and their direct connection to wages and government spending. While poverty has traditionally been measured in terms of income or consumption, the SDGs define poverty also as having multiple forms, with a target to reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions by 2030.

601. As demonstrated in Alkire et al (Alkire et al, 2015), a mismatch in the identification of the monetary vis-à-vis the non-monetary deprivations have long been identified in the poverty literature (Ruggeri Laderchi, 1997; Stewart et al, 2007), as well as differences in trends in reductions in monetary and non-monetary deprivations. Motivated by Bourguignon et al (Bourguignon et al, 2010), Alkire et al (Alkire et al, 2015) found little association between reductions in monetary poverty and progress in four non-income Millennium Development Goals. This has led to the conclusion that “income poverty

reduction does not ensure reducing deprivations in non-income indicators” and that “income poverty trends do not proxy trends in the reduction of non-income deprivations” (Ibid).

602. To date, most countries use a national income poverty measure for policy decisions, following one or several of the approaches that have been discussed in detail in previous sections. However, an increasing number of countries have adopted official multidimensional measures using non-monetary measures. Multidimensional poverty indices are particularly suitable to combine policy priorities across sectors and tailor measurement to the specific needs in countries or population groups. Multidimensional measures involve technical choices regarding the selection of dimensions, indicators, thresholds and weights. Ideally, these decisions should involve broad participation from stakeholders, and they reflect normative decisions in terms of what constitutes poverty and how best to gauge it. By contrast, technical specifications for monetary measures involve complex operations such as the transformation of national currency in purchasing power parities or establishing scales of equivalence. Non-experts hardly have any say in determining such critical parameters in the measurement of monetary poverty. On the other hand, assumptions in multidimensional measure are more easily challenged and it usually requires time to establish a widely trusted measure. Three different strands of multidimensional poverty measures are highlighted in this Guide:

- A) Multidimensional Poverty Index (MPI);
- B) Scales of Material Deprivation;
- C) Rights Based Assessments (e.g. Multiple Overlapping Deprivations Analysis, MODA).

603. Each of these strands offers different strategies for implementation and no general recommendation can be given which one is more promising for disaggregation. The background and specific problems are briefly summarised below.

5.2.5.1 Multidimensional Poverty Index

604. The global Multidimensional Poverty Index (MPI) has primarily gained acceptance in the middle income countries of Latin America and the Caribbean and is getting increasingly popular among many low income countries. It is based on the Alkire-Foster Method (Alkire and Foster, 2011) to analyse non-monetary deprivations and their interconnections and enables, like all multidimensional poverty measures, the creation of high-impact policies that address multiple interconnected deprivations and accelerate progress towards the SDGs.

605. The global MPI is a poverty measure that is primarily useful for international comparisons. It builds on strictly harmonized micro datasets and it allows tracking progress in poverty reduction over time. It was developed by the Oxford Poverty and Human Development Initiative (OPHI) and released by the United Nations Development Program (UNDP) in the Human Development Report since 2010. It is a powerful measure for policy analysis, dialogue and monitoring. The global MPI, an internationally comparable measure of acute multidimensional poverty, has ten indicators grouped into three dimensions: education (child school attendance and years of schooling), health (nutrition and child mortality), and living standards (electricity, sanitation, drinking water, housing, cooking fuel, and assets). Each dimension is equally weighted, and each indicator in a given dimension is also equally weighted.

People need to be deprived in at least one third of the weighted indicators to be considered multidimensionally poor.

606. Following this trend, the World Bank presented in October 2018 its own multidimensional poverty measure, also using the Alkire-Foster Method (World Bank, 2018). It includes six indicators grouped into three dimensions (monetary poverty, measured as the daily consumption or equivalent income below US\$ 1.90, education, and access to basic infrastructure). The World Bank also presented an exercise expanding its multidimensional measure to also include health and security from crime and natural disasters. However, this expansion is only available for six countries (Ecuador, Indonesia, Iraq, Mexico, Tanzania and Uganda) covering the years 2009-2013 (World Bank, 2018).

607. In turn, several regions have decided to adapt the general framework and create regional MPIs better tailored to their own contexts. In 2014, ECLAC launched a pilot MPI for Latin America and the Caribbean in its flagship publication *Panorama Social*, which included dimensions on housing, living standards, education, employment and social protection. Then, in 2017, the League of Arab States, OPHI and UN-ESCWA launched the Arab MPI to better characterize poverty and guide policies in the Arab States. Similarly, the *Guide on Poverty Measurement* edited by UNECE in 2017 included as a recommendation for countries in the region: “Recommendation 23: Each country should develop a national MPI that suits its national data sources and policy objectives. It is desirable that the national MPI includes the dimensions of living standards, services, health, education, work and the lived environment”. This publication emphasized the interest in a comparable measure for countries in the region, while acknowledging the limitations due to data availability and the priority for countries to first create nationally relevant MPIs.

608. But just like all countries use national income poverty measures for policy, so too many countries are developing National MPIs, which are adapted to reflect their own specific poverty priorities and national contexts. The national MPIs are adopted as official, permanent poverty statistics, and they can be used as an analytical tool to identify the people who suffer simultaneous deprivations, show aspects in which they are deprived and, consequently, to target resources and design policies more effectively. It can also guide national policies against poverty such as budget allocation and monitor policy programs requiring coordination across sectors. Because the MPI measures outcomes directly, it will immediately reflect changes such as school attendance, whereas it can take time for some policy interventions to affect income, thus may show outcomes of effective policy interventions quickly. Importantly, national MPIs explicitly reflect national decisions, strategies and priorities, so they have a clear and meaningful interpretation in a particular national context. As each country is different, such a specifically tailored national measure needs to make use of country-specific data and indicators to provide a better understanding of country-level poverty. The countries that have adopted national MPIs as an official poverty measure and included nationally adapted multidimensional measures of poverty in their long term social development strategies include Colombia, Mexico, Bhutan, Chile, El Salvador, Costa Rica, Honduras, Malaysia, Vietnam, South Africa, Ecuador, Pakistan, Armenia, Mozambique, Panama, Nepal, Dominican Republic, Nigeria, Sierra Leone, the Philippines, and Afghanistan, among others, have implemented while the city of Ho Chi Ming in Vietnam and the states of Minas Gerais in Brazil and Andhra

Pradesh in India have done so at the local level. Many other countries are now working on their national measures of multidimensional poverty.

609. Developing a useful, coherent and robust national MPI requires following a set of political and technical steps that are covered in detail in the OPHI-UNDP guide on ‘How to Build a National Multidimensional Poverty Index (MPI): Using the MPI to inform the SDGs.’⁹⁶

Box 5.4

The national Multidimensional Poverty Index (MPI) of Armenia

In November 2016, Armenia announced the launch of a national Multidimensional Poverty Index (MPI) in Yerevan. The measure was released as part of the annual “Social Snapshot and Poverty in Armenia” Report produced by the National Statistical Service of the Republic of Armenia (NSSRA) in partnership with the World Bank. The report used data from the 2010-2015 Integrated Living Conditions Survey (ILCS), as well as data from the Ministry of Labour and Social Affairs, Ministry of Health, and Ministry of Education and Science. As the report says: “Poverty has been described as a deprivation in wellbeing, a lack of key capabilities, and a type of “economic scarcity” of basic needs. A measure of multidimensional poverty captures the complexity, depth and persistence of poverty and offers important information to complement the analysis of monetary (consumption) poverty.”

The selection of dimensions for the National MPI was based on a series of consultations with stakeholders that sought to identify those deprivations most reflective of the experience of poverty in an Armenian context. The index is composed of the following five dimensions: basic needs, housing, education, labour, and health. The threshold for all members of a household to be considered deprived in a certain dimension is if they are deprived in one quarter or more of the indicators for that dimension. All members of a household are considered multidimensionally poor if they are deprived in more than one quarter of all aggregated indicators.

The main finding of the MPI is that multidimensional poverty has fallen dramatically from 41.2% in 2010 to 29.1% in 2015. Multidimensional poverty was found to be higher in rural areas than urban areas, with a 2015 rate of 32.7% in rural areas, 25.9% in urban areas other than Yerevan, and 28.0% in the capital city. However, this difference was much larger in 2010 than in 2015 and has narrowed in part due to investment in infrastructure improvements in rural areas. The composition of poverty was also found to be different between rural and urban areas: in rural areas the dimension in which most individuals were deprived was housing, while in urban areas it was labour.

Results also showed that, though multidimensional poverty had improved since 2010, the majority of individuals in Armenia still lived in households that experienced deprivations in one or more dimensions in 2015.

⁹⁶ <https://www.undp.org/content/undp/en/home/librarypage/poverty-reduction/how-to-build-a-national-multidimensional-poverty-index.html>

Figure 5.16

Dimensions of deprivation for households

Dimension: Basic needs	A household is deprived, if ...
Extreme poverty	not having access to minimum requirement of food (according to national poverty measurement methodology and FAO recommendations)
Life in dignity	not having funds to buy, when necessary, food and/or cloths
Humanitarian aid	being dependent on humanitarian assistance to ensure basic functioning of living
Remittance dependent	being dependent on remittances to ensure basic functioning of living or being in extreme (food) poverty
Dimension: Housing	A household is deprived, if ...
Satisfaction of housing conditions	not having access to adequate housing: housing conditions are evaluated as bad or very bad
Adequate housing	not having access to adequate housing: available housing requires major repairs, is dump, slum, or old; adequate flooring and adequate walls
Overcrowding	available housing floor space does not exceed 20 sq. meters per person adult equivalent
Healthy heating	household uses wood, carbon or other heating means as primary source for heating
Centralized water system	no access (use) to centralized water system
Centralized sanitation and garbage disposal	no access (use) to centralized sanitation or garbage disposal system
Hot running water	no access (use) of hot running water
Quality of paid public services	not satisfied in one third or more paid services (relative to all answered): water supply, sanitation, garbage collection, telephone, electric supply, post, banking, irrigation, public transportation
Access to transportation	not having access to opportunities: no or poor transportation and road networks (all- year road)
Dimension: Education	A household is deprived, if ...
No secondary education	present: all household member between the age of 15 years and 75 years have less than secondary education (vocational or professional)
Schooling enrollment rate	future: at least one child of compulsory schooling age between 6 and 17 years is not attending school
Access to education services	not having access to kindergarten, complete secondary school, primary (general) school in the neighborhood
Quality of education services	not satisfied with education services
Dimension: Labor	A household is deprived, if ...
Labor market participation	more than half of household members in the working age population do not participate in the labor market
Long term unemployment	at least one household member is not working due to long term unemployment (structural)
Decent jobs	not having access to decent jobs - employment status is own account worker
Underemployment	not having access to a full position in the labor market (underemployment, and seasonal/occasional employment for all members)
Dimension: Health	A household is deprived, if ...
Termination of usual activity	at least one household member did terminate usual activities because of illness, injury, or bad health.
Affordability of health services	not having funds to pay for required health services (excluding dentist) in a health care facility (in case of no or difficult access to free services), tests, examinations and procedures prescribed by a doctor
Access to health services	not having access to health care facility, emergency ambulance services, pharmacies in the neighborhood
Quality of health services	not satisfied with health services

Source: National Statistical Service of the Republic of Armenia (2016).

Recommendation 27: Following the SDGs, countries should implement multidimensional poverty measures to complement existing monetary measures of poverty. These multidimensional measures should be tailored to the national context and policy priorities, and be tracked over time.

Measures of Material Deprivation in the European Union

610. High income countries offer varied consumption choices which influence social opportunities. With ongoing progress, historic poverty definitions therefore tend to become challenged as inadequate. In 1975, the EU Council of Ministers agreed that the poor are *“the persons whose resources are so small as to exclude them from the minimum acceptable way of life in the Member State in which they live”*, with “resources” being defined as *“goods, cash income plus services from public and private sources”* (Council of the European Union, 1975). In 1985, the Council amended this definition and enlarged the concept of “resources” so that the poor were now defined as: *“persons whose resources (material, cultural and social) are so limited as to exclude them from the minimum acceptable way of life in the Member State to which they belong”* (Council of the European Union, 1985).

611. These EU definitions are effectively identical to the definition of deprivation that Peter Townsend (1979) had proposed for the United Kingdom. For Townsend, poverty is a result of a lack of adequate resources and deprivation is an outcome of poverty. This theory has been successfully operationalised in many countries using the ‘Consensual’ or ‘Socially Perceived Necessities’ methodology of Mack and Lansley (1985) which measures deprivation based on an “enforced lack concept”, i.e. the person/household lacks the item for financial reasons, not by choice. Thus, the majority of the deprivations are collected using a three-answer category question:

“Can you tell me if...”:

- You have the item;
- You do not have the item because you cannot afford it; or
- You do not have the item for any other reason

612. In April 2017, the European Union adopted a new thirteen item *Material and Social Deprivation Index* which replaced the 2009 nine item *Material Deprivation Index* (Guio et al, 2012; 2016). To be used as an annual measure, the same set of 13 items and the same threshold (5+) is used for both children and adults. However, when computing deprivation for a child, a lower weight is given to adult items → Among the 5+ deprivations required to be considered a deprived, there needs to be at least three household deprivations (out of the seven household deprivations included in the list).

Box 5.5

Definition of the material and social deprivation measure in the European Union

The inability of the household to:

1. face unexpected expenses;
2. afford a one-week annual holiday away from home;
3. avoid arrears (in mortgage, rent, utility bills and/or hire purchase instalments);
4. afford a meal with meat, chicken or fish or vegetarian equivalent every second day;
5. afford to keep their home adequately warm;
6. have access to a car/van for personal use; and
7. replace worn-out furniture.

The six personal deprivations are the inability of adults (aged 16 and older) to:

1. replace worn-out clothes with some new ones;
2. have two pairs of properly fitting shoes;
3. spend a small amount of money each week on him/herself;
4. have regular leisure activities;
5. get together with friends/family for a drink/meal at least once a month; and
6. have an internet connection.

The indicator is based on the unweighted sum of the 13 items for each person. The scale ranges from 0 (no deprivation) to 13 (enforced lack of all items).

613. The same reasoning that requires adjustment of poverty definitions for countries over time can also be applied to the measurement of poverty among children which are considered as a particularly vulnerable population group. Their situation may be inadequately represented if only adult or household-based measures were used. In March 2018, the European Union adopted its first ever measures of *Child Deprivation* (Guio et al, 2017a; 2018). The two indicators that were agreed were a child deprivation rate and an indicator of child deprivation intensity, based upon a seventeen-item index to be measured every three years. This new child deprivation index includes age appropriate measures of material and social deprivation which reflect the specific situation and changing needs of children as they grow older. Child specific items are complemented by household level items.

614. These new deprivation measures are collected in the EU-SILC and allow living standards comparisons to be made both between and within households, including disaggregated results for men and women in the same household and the comparison of deprivations suffered by children and their parents.

Box 5.6

Definition of the child deprivation measure in the European Union

The child deprivation rate is the percentage of children aged between 1 and 15 years who suffer from the enforced lack of at least three items out of the list of 17 (unweighted) deprivations:

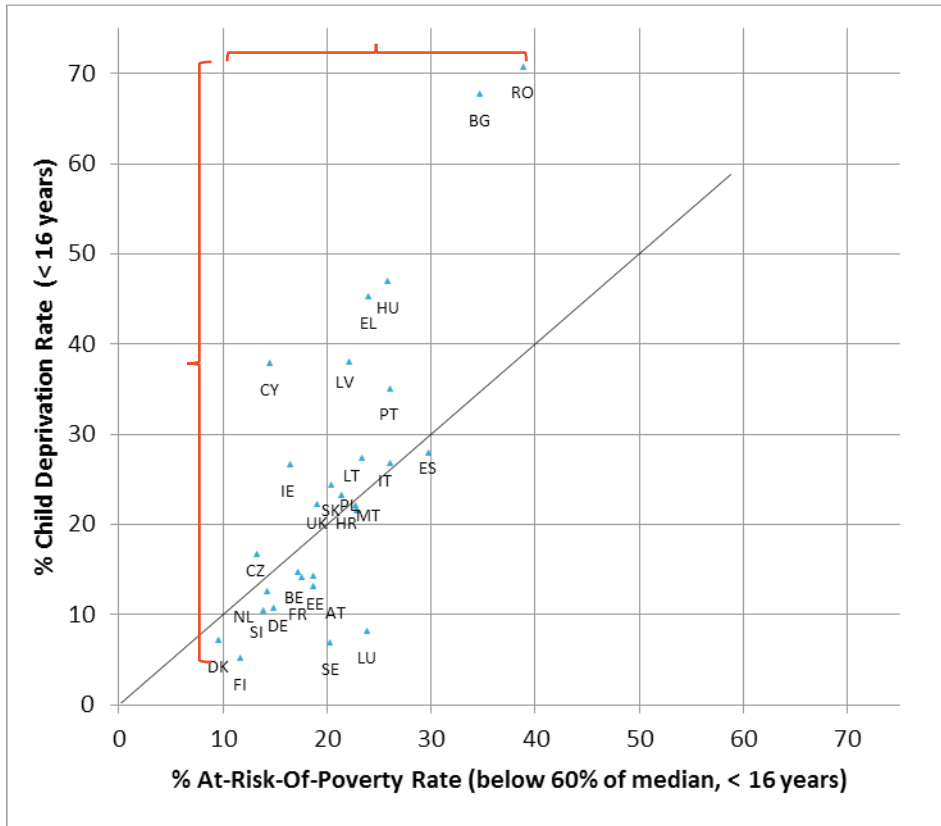
1. Child: Some new (not second hand) clothes;
2. Child: Two pairs of properly fitting shoes;
3. Child: Fresh fruits & vegetables daily;
4. Child: Meat, chicken, fish or the vegetarian equivalent daily;
5. Child: Books at home suitable for the children's age;
6. Child: Outdoor leisure equipment suitable for children' age;
7. Child: Indoor games suitable for their ages;
8. Child: Regular leisure activities that cost money;
9. Child: Celebrations on special occasions;
10. Child: Invite friends around to play and eat from time to time;
11. Child: Participate in school trips/events that cost money;
12. Child: A one-week annual holiday;
13. Household: Replace worn-out furniture;
14. Household: Arrears in last 12 months;
15. Adult: Internet connection;
16. Household: Home adequately warm; and
17. Household: Car.

The child deprivation intensity is the average number of enforced lacks among deprived children, i.e. among children lacking at least three items out of the list of 17 (unweighted) retained items.

615. Figure 5.17 shows how rates of material deprivation among children relate to the monetary poverty risk of children measured at 60% of the median in each Member State of the European Union. The positions of countries appear largely consistent, but it is notable that the differences between countries is much larger.

Figure 5.17

Relationship of Monetary Poverty (below 60% of Median) and Material Deprivation among children below 16 years



Source: Eurostat, EU-SILC 2014, ilc_peps01,

https://ec.europa.eu/eurostat/documents/1012329/8088305/Child_deprivation_rate_2014.xlsx

Box 5.7

Methodology for determining deprivation measures for the European Union

An optimum sub-set of European Union deprivation indicators were selected by drawing on an analytical framework developed for the 1999 and 2012 Poverty and Social Exclusion Survey projects (Gordon et al, 2000; Pantazis et al, 2006; Dermott and Main, 2017). Each deprivation item was tested to ensure that it was a suitable, valid and reliable measure in all 28 EU Member States (Guio et al, 2012; 2017b). Only deprivation items which passed the following tests were included in the final index:

1. The *suitability* of each item, in order to check that respondents in the different Member States (as well as the different population sub-groups within each Member State) consider them necessary to have an “acceptable” standard of living in the country where they live. Here, ‘suitability’ is understood as a measure of ‘face validity’ amongst the EU population. In each Member State, the suitability of each indicator was tested by age group, sex, household type, population density, country of birth, educational attainment and income group.
2. The *validity* of individual items, to ensure that each item exhibits statistically significant relative risk ratios with independent variables known to be correlated with deprivation.

3. The *reliability* of the scale, to assess the internal consistency of the scale as a whole - i.e., how closely related the set of items are as a group. This analysis is based on the Cronbach's Alpha statistic as well as on the Beta and Lambda coefficients and is conducted as part of a Classical Test Theory (CTT) framework. This reliability analysis of the deprivation scale as a whole is complemented with additional tests on the reliability of each individual item in the scale using Item Response Theory (IRT) and by a Hierarchical Omega Analysis.
4. The *additivity* of items, to test that the indicator's components add up – i.e. that someone with a deprivation indicator score of '2' is suffering from more severe deprivation than someone with a score of '1'. Additivity was measured for the items that successfully passed the suitability, validity and reliability tests.

The resulting 13 item all person material and social deprivation index and 17 item child deprivation index have been shown to have good measurement invariance properties across EU member states, i.e. a person with the same deprivation index score in each country has effectively the same low standard of living (Guio et al, 2017b).

The development of these new deprivation indicators took over a decade's work by poverty measurement experts in the 28 National Statistics Offices of the European Union member states and many academic experts. A number of clear lessons can be drawn for the development of multidimensional poverty and deprivation measures and for the disaggregation of such measures.

All poverty measures need to be reliable and valid if they are to be useful for policy, monitoring and evaluation purposes. Reliability is the amount of random error affecting a poverty measure and validity is the amount of systematic error or bias. It is self-evident that a poverty rate of $20\% \pm 1\%$ is of greater use than a poverty rate of $20\% \pm 30\%$. Similarly, an invalid poverty measure which incorrectly identifies certain groups of people or geographic areas as 'poor' is of little use for policy making and may result in the misallocation of money and efforts (Najera and Gordon, 2019a; Najera and Gordon, 2019b).

Recommendation 28:

- a) Deprivation measures need to be based upon a clear and explicit theory or normative definition of poverty in order to ensure that each indicator is a valid measure, i.e. that it measures poverty and not some other related (or unrelated) concept such as wellbeing or happiness.
- b) The validity of each indicator should be demonstrated, i.e. the amount of systematic error should be formally assessed and indicators should be dropped if they have a low validity.
- c) The reliability of each indicator should be determined, i.e. the amount of random error should be formally assessed and indicators should be dropped if they have a low reliability.
- d) Normally, a deprivation or poverty index should count each item with an equal weight. Differential weights should only ever be used when this reduces measurement error (i.e. validity and/or reliability of the index are improved).

5.2.5.2 Rights Based Assessment Frameworks- Multiple Overlapping Deprivations Analysis (MODA)

616. MODA is a child-specific tool that uses the child as the unit of analysis, rather than the adult or household, and evaluates child-specific deprivations based on a normative child rights framework (UNICEF 2017b). It was developed in 2012 by UNICEF to address cross-sectoral poverty among children. MODA builds on the MPI and other approaches to multidimensional poverty measurement. It does acknowledge that children suffer specific deprivations which are different from the ones faced by adults. The approach avoids the unwarranted assumption of equal sharing of the available resources in a household and also takes into account that money cannot always purchase what children need.

617. Using indicators such as immunization status (health dimension), exposure to violent discipline (child protection dimension) and overcrowding in the home (housing dimension), these studies suggest that the multidimensional deprivation is higher than monetary poverty for child poverty in Europe and Central Asia region. For example, 63 per cent of children under five in Bosnia and Herzegovina experience deprivation in at least three dimensions; while 64 per cent of children in Armenia are deprived in at least two dimensions. Children who live below monetary poverty lines are more likely to experience multiple deprivations, but the overlap is not complete (UNICEF 2017a)

618. This requires data to be collected for children and not only for adults or on household level. Often MODA is supported by data collected in the Multiple Indicator Cluster Surveys (MICS) which in their 6th round are increasingly supported by UNICEF in countries of the ECE region: Azerbaijan, Kazakhstan, Bosnia and Herzegovina (plus Roma Settlements), Uzbekistan, Kosovo under UNSC res. 1244 (plus Roma settlements), Belarus, Serbia (plus Roma Settlements), Turkmenistan, North Macedonia (plus Roma Settlements), Georgia, Kyrgyzstan, Montenegro (plus Roma Settlements).

619. There is an important conceptual difference from conventional MPI. While in the MPI, disadvantage on one indicator may - within the limits of the weighting scheme - be compensated by other indicators, MODA instead considers each dimension as a fundamental right which cannot be substituted across dimensions. Thus, MODA will always arrive at a headcount of children which are at least in some way deprived of a right as specified in the Charter on the Right of Children (CRC).

620. As MODA is rooted in the CRC it usually produces the same dimensions: Education; Health, Nutrition; Water; Sanitation; Shelter; Information; Protection from violence. To ensure national ownership MODA does however remain flexible for selecting indicators for country-specific analysis (N-MODA) which may be distinct from the standard set used for cross-country analysis (CC-MODA).

621. It is clear that acting on child poverty is the main purpose on monitoring child poverty. UNICEF therefore suggests to begin with the disaggregation of existing measures even if they may inadequately reflect children's needs (UNICEF 2017a). In parallel however, work on more child specific measures should be established so that the impact of policies on children will not be obscured by household level measurement and children needs are acknowledged (UNICEF 2017b). Once such tailored instruments are achieved, they may be integrated and compared with multidimensional adult poverty measures. By considering household characteristics as well as different characteristics of deprivation for adults and children such measures may even be integrated and compared to each other. Instead of having to add or

weight items, this can be achieved by simply assessing across a range of indicators whether or not an individual fulfils any of the criteria for deprivation, including household level criteria which are identical for all members of a household or age specific criteria (Delamonica, 2014).

Recommendation 29:

- a) Countries should use available datasets, such as Multiple Indicator Cluster Surveys (MICS) or household surveys to develop child-specific and life-cycle adapted multidimensional poverty measures that reflect the needs of children at different stages of development and allow for identification of intra-household differences between children. This type of measure can be performed at intervals of 3-5 years to complement more frequent disaggregated national measures, as this will give greater insight into childhood and adolescent poverty.
- b) In order to enhance availability and use of child poverty data, countries should collect data on all key dimensions related to children's rights, including health and nutrition, and introduce lifecycle appropriate indicators to measure the situation of each child in the household. Countries should consider introducing innovative ways to collect, monitor and report on child poverty data, including ways to encourage child participation in the monitoring and discussion of child poverty data and potential policy responses.
- c) In accordance with national definitions of monetary and multidimensional poverty, countries should revise and adopt survey tools to best serve their national needs for poverty measurements. Both household budget surveys (HBS) and MICS are flexible and can be adapted to reflect a national context, but without compromising cross-country comparability. MICS offers the potential to obtain data on a broader list of child-focused indicators that can be used to measure multidimensional child poverty.
- d) Statistical data is an important source for evidence-based decision making by policy makers, not only at national but also at regional and international level. Therefore, it is important to make anonymised statistical data openly available for all users. To meet data protection and confidentiality concerns, efforts must be taken to ensure that the identity of respondents is not disclosed or can be inferred. Hence countries should make all poverty-related data, including micro-data, publicly available and easily accessible for scientific research and production purposes. This would enhance research, policy design and policy innovation in this field, which is of utmost importance for devising policies for poverty reduction.

6 Future Work – measuring what matters

622. The recommendations presented in this Guide focus on the specific issues for which the Task Force on Disaggregation of Poverty Measures has been mandated. It is hoped that they will also contribute to the further improvement and convergence of measuring poverty.

623. National statistical offices operate in quite different legal environments and with often very different capacities. Yet their methodological problems are often rather similar and a model of “mutual learning” is vital to be efficient and effective in achieving the common target of measuring poverty. This Guide is an example of the fruitful collaboration between transatlantic partners which already have powerful statistical systems and reflects the dynamic progress in the countries of Eastern Europe and Central Asia.

624. With the creation of EU-SILC, the European Union has arguably been leading the development of harmonised methodologies for comparative measurement of poverty in the region over the last two decades. Eurostat guidelines proved especially useful for transition countries which previously could not draw on a tradition of statistical instruments for measuring poverty such as EU-SILC. Even more importantly however statistical indicators take a firm place for policy coordination in the European Union. As a rule, it may be said, that what matters for the European Union needs to be measured first.

625. The new European Commission’s framework regulation (EC 2019/1700),⁹⁷ which will become effective in 2021, will have serious impact on EU-Member States. It aims to ensure consistent data collections on income and living conditions as well as labour force, health, education and training, use of information and communication technologies, time use and consumption and breakdowns for disaggregation. The regulation also aims at improving quality, timeliness, regionalisation and responsiveness to emerging policy needs. Priorities like these and the ongoing modernisation of data collections (notably the wider use of registers and online communication) are major challenges for future poverty measurement in the whole region.

626. Strong working relationships between statisticians who work on measuring poverty together needs to be built continuously through workshops, conferences, study visits and collaborative research projects which involve partners from across the whole region. A lot remains to be done. This chapter gives an overview of some of these immediate priorities.

6.1 Pending problems in measurement of resources

627. This Guide presented selected examples how social transfers in kind and housing cost and wealth are accounted for in practice. The adequate inclusion of stocks of wealth and non-monetary resources in the measurement of monetary poverty does however remain an important field for research. This holds especially for comparisons between countries which have quite different volumes of non-monetary

⁹⁷ <http://data.europa.eu/eli/reg/2019/1700/oj>

transfers or different house ownership rates. Those comparisons are critically affected by the way such resources are considered in measurement.

628. It is important to emphasise that even the absence of appropriate adjustments comes down to the unwarranted assumption that such resources are zero. This assumption is clearly violated. For example, in high income countries of the European Union social transfers in kind for education amount to between 5% and 11% of adjusted gross disposable income. Non-monetary transfers in health have been estimated to between 3% and 13% (Eurostat 2019) while housing amounted to between 14% and 25% of adjusted net disposable income in OECD countries (OECD 2014). Together, these components will easily exceed one third of total income. These resources will hence gain increasing importance for countries in transition to higher incomes which are also characterised by increasing aggregate values for social transfers in kind, housing stock or other forms of wealth.

629. Any adjustments regarding resources which at best are known only on an aggregate level will inevitably require critical assumptions on their distribution. There is currently no agreement on methods to impute fictitious rental income or the value of public services. Even among high capacity countries of the European Union such consensus appears far away (Törmälehto and Sauli, 2017). Also, it remains unclear if not for the purpose of measuring poverty it would be more appropriate to instead consider out-of-pocket expenditures for housing or health and education services which can be observed empirically. In either case there may be reasons to demand a capping of the maximum value of resources that will be accounted for in the measurement of poverty. Caution is needed when analysing STiK and their potential distributive consequences. If, for example, all of the resources are STiK, they cannot be assumed as convertible into an acceptable way of living. Otherwise poverty measurement would face the conceptual problem that individuals who happen to be well off according to their imputed resources may in reality be unable to fulfil daily needs which require cash. This is even aggravated for certain approaches of evaluating STiK by the amount of government spending (and not actual consumption). Furthermore, health related STiK (e.g. for cancer treatment) shall never be taken to lift e.g. a homeless person out of evident poverty.

630. In most developed countries income-based poverty measurement has gradually replaced or at least complemented consumption-based poverty measurement. This partly reflects the active role of the government in redistribution through tax and transfers. With further increasing prosperity, inequalities in wealth have emerged which may play an increasingly important role also for the assessment of relative poverty measures. This Guide provided only the first empirical examples. More countries are expected to explore wealth-based resource measures.

631. The implications of different approaches to measuring resources would need to be assessed systematically with empirical data. Instead of only investigating the potential impact within one country however, comparative assessments which compare results from different countries and different approaches are crucially required to inform guidelines that could eventually support countries in adopting a common methodological approach.

6.2 Pending problems in measurement of needs

632. Like the measurement of resources, there remains unfinished business with regard to the assessment of needs. This is most evident with regard to the relationship between the cost of living and characteristics of the household, such as the number of adults, children or persons with special needs due to old age, care or disability. In the short run it will be advisable to establish guidance on a conventional equivalence scale for the region of Eastern Europe and Central Asia. The implicit cost of children reflected in that scale should be compared to empirical estimates for consumer demand, subjective income needs or material deprivation as well as conventional scales used for social security benefits.

633. Within ageing societies the empirical assessment of need for care cannot be limited to children only but has to include the cost for health expenditure, long term care and disability. This research is unlikely to be successful without close cooperation between Central Statistical Offices and agencies which have practical experience with these groups or are responsible for their social protection.

634. Problems in the assessment of needs are also inherent in the definition of multidimensional measures of poverty. The estimates obtained from the global MPI for the region indicate an extremely low incidence and would be hardly useful for practical purposes. Instead the elements of the various existing national MPIs need to be compared and possibly aligned to obtain comparable estimates. The same holds for the need to clarify the definition of a child specific multidimensional poverty index. Increased coordination between national stakeholders in the region as well as OPHI and UNDP on one side and UNICEF on the other appears essential.

6.3 Microsimulation and impact of transfers

635. Over the past decades, microsimulation has been developed as a powerful instrument to study redistribution mechanisms which operate in the modern welfare state. This opens new opportunities both for designing social policy as well as the production of statistics.

636. In the European Union, EUROMOD has become a widely used tool by academics and policy makers to evaluate and/or launch policy reforms. This tool combines policy rules as well as input microdata sets which are commonly used in the measurement of poverty. Through simulation it becomes possible to investigate the anticipated “overnight” effect if for example pensions or child benefits would be increased or entitlements altered. Such simulations also help to anticipate the macroeconomic impact of new social policy instruments or pension reforms. EUROMOD can also be useful to trace to what extent past developments can be attributed to actual policy intervention rather than business cycle or random events.

637. Moreover, the growing accessibility and the use of administrative data in the Statistical Institutes allow new synergies between microsimulation techniques and fiscal data sources. The use of tax returns as benchmark for the micro-simulated outputs together with the sample data (as in the case of the microsimulation model of the University of Siena), in some cases allows to avoid the potential issues in

the exclusive use of a microsimulation model, mainly due to the survey under-reporting, and those ones of the fiscal data that could be affected by incomplete coverage, or tax avoidance and tax evasion.

638. To make such studies beneficial it is advisable to ensure close collaboration between those who set up the simulation machine and the producers of microdata. This involves also many technical aspects such as the level of aggregation as well as the imputation of unobserved income components. On the other hand, microsimulation techniques can help to constantly improve the empirical data base, for example if amounts which are collected only on gross level need to be converted into net values or vice versa (when only net income information are available from the survey).

6.4 Dedicated surveys on hard to reach populations

639. The target of eradicating poverty in all its forms implies that the phenomenon is likely to be more difficult to observe in the future with conventional methods. Remaining pockets of poverty are often particularly difficult to access. Increased attention will therefore be required to ensure that the measurement of poverty does not leave behind those population groups which are at highest risk.

640. Conventional household surveys are usually not designed to represent small or hard to reach populations. Increasing sample size and improving allocation as well as taking measures to better accessibility (e.g. by adequate interviewer training, target group participation, translation) may be helpful. This may not be successful for all population groups, especially if the measures of poverty do not appear appropriate, for instance when the normal subsistence involves irregular or clandestine activity, such as day labour, begging or smuggling, or consumption and income have less importance, such as with persons living in institutions.

641. Dedicated surveys may be used to overcome the constraints of conventional household surveys adopting specific techniques to investigate living conditions of hard to reach populations, as in the case of homeless persons or Roma people. Such special data collections can complement traditional poverty measures. To improve the use of dedicated surveys, increasing coordination and exchange of experiences between the involved organisations appears crucial.

642. It is worth to cite the case of Multiple Indicator Cluster Survey (MICS), as the number of countries which are supported by UNICEF to conduct MICS is increasing. Many of those countries opted to have special samples among the Roma population. Likewise, the EU Agency for Fundamental Rights (FRA) has developed a methodology for sampling Roma and other vulnerable groups. Both operations point beyond conventional monetary measurement of poverty. In order to achieve comparability within the poverty estimates for the whole population it is important to better coordinate activities with such operations. This coordination includes the use of common questionnaire modules as well as sampling.

6.5 Use of registers

643. Most Member States of the European Union have already moved towards an increased use of administrative data for statistical purposes. This move is driven mainly by the need to reduce the cost of data collection, to reduce the burden on respondents, and reuse existing information.

644. The use of administrative data does however require initial investments, including in legal and technical frameworks. The usage of administrative information has important implications on all quality dimensions, including breaks in series, timeliness, comparability and privacy.

645. Administrative sources include population registers, tax registers, social security data, and health and education records. Even if most countries are using one or many of these sources, the degree to which administrative information is considered in the measurement of poverty varies between countries and across statistical domains. Some countries have chosen to design their samples according to the selected respondent model, in which only one person per household is interviewed and information on cohabitants is derived mainly from register information.

646. Moreover, registers in some cases do not provide enough information about poor persons and it is necessary to complement this information with data coming from the surveys. Given the potential impact on results, further exchange of experience in using registers in the measurement of poverty is most urgently required to ensure efficient use of the available information and methodologies but also to potentially improve comparability.

6.6 Link to macro aggregates

647. Monetary poverty is measured on income or consumption. In this Guide, the need to consider aggregates has been discussed especially with regard to social transfers in kind and imputed rent. It turns out that the aggregates and moreover the growth rates are not strictly comparable between microdata and macro data. The latter relate in particular to the UN System of National Accounts. It is noteworthy however, that its definitions are followed to varying degrees by countries and are often subject to revisions.

648. Users who seek robust policy implications would expect that producers have carried out a reconciliation or provide explanations when discrepancies occur to. The necessity of reconciling household survey information with the national accounts has been recognized in the work of international and national statistical agencies (see for example Fesseau, Wolff, and Mattonetti 2013 and Mattonetti 2013 on the work of an OECD-Eurostat Expert Group). A recent Eurostat working paper addressed specifically the reconciliation of EU-SILC and the national accounts (Törmälehto 2019).

649. The discrepancies which are observed empirically are capable of altering poverty profiles quite drastically. For example, ROSSTAT (2017) estimated that in 2012 about 30% of the corresponding macro-economic aggregate was missing in its income survey. As a consequence, ROSSTAT established a special calibration to adjust the income distribution and ensure better coherence. This measure approximately

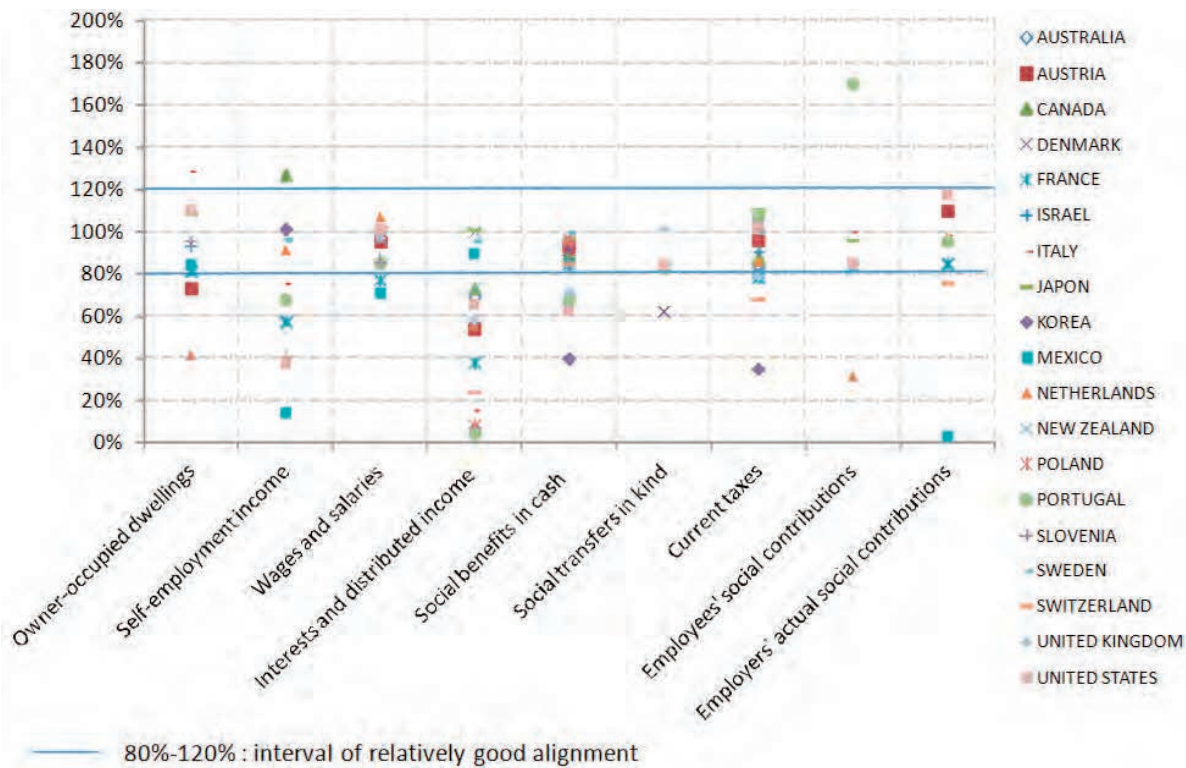
halved the discrepancy with the macroeconomics aggregate but has also reduced the estimated poverty rate from 15.8% to 12.4%. For OECD countries, discrepancies within 80% to 120% of the corresponding national accounts data are considered as a relatively good alignment, given the presence of differences in definition as well as other errors on both sides. For some components of income and in some countries the discrepancy does however exceed by far that reported by the Russian Federation.

650. It is increasingly difficult to obtain quality data from household surveys, due to both falling response rates and concerns with the accuracy of survey responses. One method to evaluate the quality of survey responses is to compare aggregate estimates from surveys to administrative benchmarks, such as aggregates from the national accounts. In addition, some national statistical offices adjust survey estimates to align them more closely with these administrative benchmarks.

651. While these comparisons and adjustments are useful, they should be made with caution given the conceptual differences between survey income concepts and administrative income concepts. For example, for the United States, surveys measure “money income” which is cash regularly available to households while the national accounts measure “personal income” which is the aggregate value of compensation including cash and non-cash benefits. Administrative reports of expenditures on social transfers in kind may include many administrative costs that should not be attributed to individual households. Second, survey sampling universes may not align with the administrative universes. Most surveys include only households in the sampling frame while administrative data has a more inclusive universe.

652. Future work should examine best practices by national statistical offices in the utilization of macro aggregates to evaluate and/or adjust estimates from household surveys. The analysis of inequality would also benefit from integrated data-that allows the joint assessment of distributions of household income, consumption and wealth (Stiglitz et al, 2018b).

Figure 6.1

Difference between national accounts aggregates and income micro data

Source: Fesseau et al, 2013: 21.

6.7 Dynamics of poverty and exclusion (panel data)

653. To understand the ways in and out of poverty which are relevant to policy makers, it is important to track changes over time in the circumstances of individuals. Without longitudinal data it is quite impossible to distinguish between opposite extremes of permanent poverty or perfectly fluid poverty. For example, a poverty rate of 16 per cent would be perfectly consistent with the same one sixth of the same population remaining permanently below the poverty threshold or a population where everyone spends one year in six in poverty.

654. Only panel data can determine how much mobility there is within the poverty population. Available longitudinal data has shown that even when poverty rates do not change there can be substantial turnover (Till and Eiffe, 2010). This has two implications. Firstly, the assumption that all the poor remain persistently in the same disadvantaged condition can be rejected. Secondly, the number of persons who experience poverty is actually larger than can be seen from cross sectional data. The transitions between different statuses reveal that sometimes poverty transitions do not lead very far from the poverty threshold. Hence, the longitudinal study of poverty provides an opportunity to distinguish individuals in persistent, transitory and oscillating poverty from individuals who are very unlikely to ever become poor. The fact that existing longitudinal data remains underused (Eiffe and Till, 2014) indicates a

need to strengthen the capacity within statistical agencies as well as to improve access to longitudinal micro data for researchers.

655. A number of mainly academic panel surveys exist in the region. These operations have cumulated considerable methodological and research experience which should be reflected more critically by official statistics. For example, the Panel Survey of Income Dynamics by the University of Michigan the United States which was started in 1968 is probably the longest running panel survey used to study poverty dynamics (Bane and Ellwood, 1983). In 1983 the US Census Bureau decided to launch its own Survey on Income and Programme Participation (SIPP) to be able to better address the dynamics of poverty in its official measurement of poverty (Mohanty, 2019).

656. Since the 1980s Panel surveys have been widely used for poverty measurement in Europe, including the German Socioeconomic Panel (GSOEP) run by the German economic research institute (DIW); the British Household Panel Study (BHPS) or its successor the Understanding Society Survey run by the University of Essex (Ferragina et al, 2013). The Russia Longitudinal Monitoring Survey (RLMS) started in 1994. When such academic surveys exist in parallel the assessment of their consistency with poverty measures published by statistical offices is essential to maintain credibility towards users.

657. The measurement of poverty with panel data involves special methodological problems, such as panel attrition and the loss of representativity with regard to migration which happened after the panel was started (UNSD 2005). For example, in the case of the RLMS only 29 percent of the original sample were followed up after 20 years (Kozyreva et al, 2016). It is hence extremely important to carefully weigh the purpose of measuring the extent of poverty with the purpose of understanding its dynamics.

658. A genuinely comparative source for poverty statistics in the European Union during the 1990s was the European Community Household Panel Survey (Fourage and Layte, 2005). It was terminated after seven years, partly because the political process of the European Union required more robust cross sectional poverty indicators which should also be anchored firmly in the official statistical system. The rotational EU-SILC design therefore has a much larger cross sectional sample than the ECHP but a relatively short panel duration and a small sample size for the panel component. Research examples include persistent poverty (e.g. Weber, 2019) as well as transitions (e.g. Vaalavuo, 2015, Guggisberg et al, 2019).

6.8 Privacy and access to microdata

659. Steps taken by national statistical offices to prevent any outside entity from identifying individuals or businesses in the statistics they publish are known as disclosure avoidance, disclosure limitation or disclosure control. Disclosure avoidance methods have evolved over time, but historical methods cannot completely defend against threats posed by today's technology. Growth in computing power, advances in mathematics, and easy access to large, public databases pose a significant threat to confidentiality. These forces have made it possible for sophisticated users to ferret out common data points between databases using only our published statistics. If left unchecked, those users might be able to stitch together these common threads to identify the people or businesses behind the statistics. The process is called re-identification, and that threat has become more real with today's technology.

660. For the 2020 Census, the U.S. Census Bureau is moving to a new, advanced, and far more powerful confidentiality protection system, which uses a rigorous mathematical process that protects respondents' information and identity. The new tool is based on the concept known in scientific and academic circles as "differential privacy." It is also called "formal privacy" because it provides provable mathematical guarantees, similar to those found in modern cryptography, about the confidentiality protections that can be independently verified without compromising the underlying protections.

661. "Differential privacy" is based on the cryptographic principle that an attacker should not be able to learn any more about a person from the statistics published using your data than from statistics that did not use your data. After tabulating the data, the Bureau will apply carefully constructed algorithms to modify the statistics in a way that protects individuals while continuing to yield accurate results. This method assumes that everyone's data are vulnerable and provide the same strong, state-of-the-art protection to every record in our database.

662. Future work should examine best practices in disclosure protection techniques, helping national statistical offices balance the competing demands of data accessibility, transparency and privacy protections.

6.9 Comparative quality reporting

663. This Guide has highlighted the need to regularly assess the quality of poverty measurement and inform users about any departures from conventional international standards. The Guide also helps producers to document their work and continuously improve quality of data. If quality reports from different countries are consolidated, this can generate a rich repository of methodological practices. This can also support mutual learning and a long-term convergence of data production and enhanced comparability. To make these possible, international agencies can play an important role. UNECE, Eurostat and CIS-STAT already fulfil important functions to facilitate coordination of methods as well as statistical indicators.

References

- Aaberge, R., Langorgen, A. and Lindgren, P. (2017). The distributional impact of public services in European countries, in Atkinson, Anthony B. Anne-Catherine Guio and Eric Marlier (eds.), *Monitoring social inclusion in Europe: 2017 Edition*, Eurostat, European Union.
- Agence National de la Statistique et de la Demographie (2015). *Enquete pauvreté et structure familiale*. Available from: http://www.ansd.sn/ressources/publications/ANSD_EPSF2_rapportfinal2.pdf
- Alkire, S. and Foster, J. (2011). Counting and multidimensional poverty measurement, *Journal of Public Economics* 95(7-8): 476-487.
- Alkire, S., Foster, J. E., Seth, S., Santos, M. E., Roche, J. M. and Ballon, P. (2015). *Multidimensional Poverty Measurement and Analysis*, Oxford: Oxford University Press.
- Assael, H. and Keon, J. (1982). Nonsampling vs. Sampling Errors in Survey Research. In: *Journal of Marketing*, 46/2: 114-123.
- Atkinson, T., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Social indicators: The EU and social inclusion*. OUP Oxford.
- Atkinson, T., Guio, A.C. and Marlier, E. (eds.) (2017). *Monitoring social inclusion in Europe*. Eurostat Available from: <https://ec.europa.eu/eurostat/documents/3217494/8031566/KS-05-14-075-EN-N.pdf/c3a33007-6cf2-4d86-9b9e-d39fd3e5420c>
- Atkinson, T. and Marlier, E. (eds.) (2011). *Income and living conditions in Europe*. Eurostat. Available from: <https://ec.europa.eu/eurostat/documents/3217494/5722557/KS-31-10-555-EN.PDF/e8c0a679-be01-461c-a08b-7eb08a272767>
- Australian Bureau of Statistics (2019). *Survey of Income and Housing, User Guide, Australia, 2017-18*, <https://www.abs.gov.au/ausstats/abs@.nsf/PrimaryMainFeatures/6553.0?OpenDocument>
- Azpitarre, F. (2010). *Measuring Poverty using both Income and Wealth: Does the Integration Method Matter?* Paper prepared for the 31st General Conference of The International Association for Research in Income and Wealth, St. Gallen, Switzerland, August 22-28, 2010.
- Azpitarre, F. (2012). *Measuring Poverty using both Income and Wealth: A cross-country Comparison between the U.S. and Spain*. In: *Review of Income and Wealth* 58/1: 24-50. Available from: <http://dx.doi.org/10.1111/j.1475-4991.2011.00481.x>
- Baio, G., Blangiardo, C. and Blangiardo, M. (2011). Centre Sampling Technique in Foreign Migration Surveys: A Methodological Note. In: *Journal of Official Statistics*, 27/3: 451-465.
- Balcazar, C. F., Ceriani, L., Olivieri, S., Ranzani, M. (2014). *Rent Imputation for Welfare Measurement: A Review of Methodologies and Empirical Findings*. World Bank Group: Poverty Global Practice Group Policy Research Working Paper.
- Balestra, C. and Tonkin, R. (2018). *Inequalities in household wealth across OECD countries: Evidence from the OECD Wealth Distribution Database*. OECD Statistics Working Paper No. 88. Available from: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC\(2018\)1&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC(2018)1&docLanguage=En)
- Balestra, C., and Fleischer, L. (2018). *Diversity statistics in the OECD: How do OECD countries collect data on ethnic, racial and indigenous identity?* SDD/DOC (2018)9. Organisation for Economic Co-operation and Development Statistics and Data Directorate. Available from: [http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC\(2018\)9&docLanguage=En](http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=SDD/DOC(2018)9&docLanguage=En)
- Balestra, C., and Sustova, S. (2017). *On the measurement of imputed rent and social transfers in kind in the OECD countries*. OECD Statistics Working Papers. Available from: <https://www.oecd.org/els/soc/Session4-1-STiKs-questionnaire.pdf>

-
- Bargain, O. and L. Martinoty (2018). Crisis at home? Effect of the Mancession on intrahousehold distribution. In: *Journal of Population Economics*, 32/1: 277-308 (CNRS: 2, AERES: A). Available from: <http://www.columbia.edu/~pc2167/Publi2017.pdf>
- Barnes, W. (2008). Improving Migrant Participation in the Labour Force Survey: A Review of Existing Practices in European Union Member States. In: *Survey Methodology Bulletin*, 63, Office for National Statistics: 25-38.
- Bethel, J. (1989). Sample allocation in multivariate surveys. In: *Survey Methodology*, 15/1: 47-57.
- Betti, G. and Lundgren, L. (2012). The impact of remittances and equivalence scales on poverty in Tajikistan. In: *Central Asian Survey* 31:4: 395-408
- Betti G., Gagliardi F., Lemmi A., Verma V. (2012). Sub-national indicators of poverty and deprivation in Europe: methodology and applications. In: *Cambridge Journal of Regions, Economy and Society*, 5/1: 149-162.
- Bianchi A., Biffignandi S. and Lynn P. (2016). Web-Face-to-Face Mixed-Mode Design in a Longitudinal Survey: Effects on Participation Rates, Sample Composition, and Costs. In: *Journal of Official Statistics* 33/2 (2017): 385–408.
- Bishop, J. A., Grodner, A., Liu, H., & Ahamdanech-Zarco, I. (2014). Subjective poverty equivalence scales for Euro Zone countries. In: *The Journal of Economic Inequality*, 12/2: 265-278.
- Bizikova, L. (2017). Disaggregated Data is Essential to Leave No One Behind. International Institute for Sustainable Development. Available from: <https://www.iisd.org/blog/disaggregated-data-essential-leave-no-one-behind>
- Bourguignon, F., Bénassy-Quéré, A., Dercon, S., Estache, A., Gunning, J., Kanbur, R., Klasen, S., Maxwell, S., Platteau, J.P. and Spadaro, A. (2010). Millennium Development Goals at Midpoint: Where do we stand and where do we need to go?
- Brandolini, A., Magri, S. and Smeeding, T. (2010). Asset-Based Measurement of Poverty. Paper prepared for the 31st General Conference of The International Association for Research in Income and Wealth, St. Gallen, Switzerland, August 22-28, 2010.
- Burchardt, T. and Karagiannaki, E. (2018). Intra-household inequality and material deprivation and poverty in Europe. Presentation for the Social Situation Monitor Seminar “Multidimensional Poverty in the EU”, March 2018. Additional information on the project from the LSE Website: http://sticerd.lse.ac.uk/case/_new/research/Intra-household/
- Bütikofer, A., Gerfin, M. (2009). The economies of scale of living together and how they are shared: Estimates based on a collective household model. IZA Discussion Paper No. 4327. Bonn: Institute for the Study of Labor
- Cage, R., Klick, J. and Johnson, W (2018). Population Subgroup Price Indexes: Evidence of Heterogeneity or Measurement Error? Bureau of Labor Statistics Working Draft with Preliminary Results.
- Caner, A. and Wolff, E. (2004). Asset Poverty in the United States, 1984-99: Evidence from the Panel Study of Income Dynamics. In: *Review of Income and Wealth*, 50/4: 493-518. Available from: <http://dx.doi.org/10.1111/j.0034-6586.2004.00137.x>
- Carbonnier, C. (2019). Estimating local basic standard of living impact on overall poverty and inequality measures. Paper prepared for and presented at the 8th Meeting of the Society of the Study of Economic Inequality, Paris, France, July 3-5, 2019. Available from: http://www.ecineq.org/ecineq_paris19/papers_EcineqPSE/paper_211.pdf
- Chiappori, P.A. (1992). Collective Labor Supply and Welfare. In: *Journal of Political Economy*, 100/3: 437-467
- CIS-STAT (2019). Disaggregation of poverty indicators for the monitoring of the SDGs in the Commonwealth of Independent States: the practice by countries and the outlook. Economic Commission for Europe. Conference of European Statisticians. Meeting of Experts in Measuring Poverty and Inequality: SDGs 1 and 10. Geneva, Switzerland 5 - 6 December 2019 Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/EN-2._CISSTAT.pdf

-
- Cochran, W.G. (1977). *Sampling Techniques*. New York: John Wiley & Sons, Inc.
- Dang, H.-A., Abanokova, K., Lokshin, M. (2019). The important role of equivalence scales: Household size, composition, and poverty dynamics in Russia. Paper being prepared for the IARIW conference in Moscow, Russia, September 17-18, 2019. Available from: http://iariw.org/moscow/ADL_final.pdf
- Deaton, A. and Muellbauer, J. (1980). An Almost Ideal Demand System. In: *American Economic Review*, 70/3: 312-26
- Deaton, A. and Zaidi, S. (2002). Guidelines for Constructing Consumption Aggregates for Welfare Analysis. LSMS Working Paper 135. Washington, D.C.: The World Bank. Available from: <https://openknowledge.worldbank.org/handle/10986/14101>
- Delamonica, E. (2014). Separating and Combining Child and Adult Poverty: Why? How? Crop Policy Brief. Available from: <https://www.crop.org/viewfile.aspx?id=636>
- De La O Campos, A.P., Villani, C., Davis, B. and Takagi, M. (2018). Ending extreme poverty in rural areas – Sustaining livelihoods to leave no one behind. Rome, FAO. (pp. 84). Available from: <http://www.fao.org/3/ca1908en/CA1908EN.pdf>
- De Vos, K., Garner, T. I. (1991). An evaluation of subjective poverty definitions: Comparing results from the U.S. and the Netherlands. *Review of Income and Wealth*, 37(3): 267–285.
- Di Giglio, S., Kore, Z., Karaj Bolduc, E., Bocaj, J., Majko, A., Hima, F. and Jonuzaj, M. (2014). National Study on children in street situation in Albania. Available from: https://resourcecentre.savethechildren.net/node/8217/pdf/final_research_report_english.pdf
- Di Meglio E., and Montaigne F. (2013). Registers, timeliness and comparability: Experiences from EU-SILC, in Markus Jäntti, Veli-Matti Törmälehto and Eric Marlier (2013, eds.), *The use of registers in the context of EU-SILC: challenges and opportunities*, Luxembourg: Publications Office of the European Union, 2013.
- Dijkstra, L. and Poelman, H. (2014). A harmonised definition of cities and rural areas: the new degree of urbanisation. European Commission. Available from: https://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf
- Dillman, D. A., Smyth, J. D. and Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. Hoboken, NJ: John Wiley & Sons.
- Duflo, E. (2012). Women empowerment and economic development. In: *Journal of Economic Literature*, 50/4: 1051–1079. Available from: <https://economics.mit.edu/files/7417>
- ECLAC (2014). *Panorama Social de America Latina*. Available from: https://repositorio.cepal.org/bitstream/handle/11362/37626/6/S1420729_es.pdf
- Eiffe, F. and Till, M. (2014). The Longitudinal Component of EU-SILC: still Underused? December 2012 Conference: Conference: EU-SILC Conference 2012. Net-Silc2 Working Paper 3/2014. Statistics Austria, Vienna. Available from: http://www.statistik.at/web_de/frageboegen/private_haushalte/eu_silc/071274.html
- Eiffe, F., Till, M., Datler, G., Heuberger, R., Glaser, T., Kafka, E., Lamei, N., Skina, M. and Till-Tentschert, U. (2011). *Soziale Lage älterer Menschen in Österreich. Sozialpolitische Studienreihe Band 11*, Ministry of Social Affairs, Vienna. Available from: https://www.sozialministerium.at/cms/site/attachments/3/8/1/CH3434/CMS1459846615201/14_sozialpolitische_studienreihe_band_11.pdf
- Eltinge, J. L. and Yansaneh, I. S. (1997). Diagnostics for formation of nonresponse adjustment cells, with an application to income nonresponse in the U.S. consumer expenditure survey. In: *Survey Methodology*, 23: 33-40.
- European Commission (2003). Commission Regulation (EC) No 1982/2003 of 21 October 2003 implementing Regulation (EC) No 1177/2003 of the European Parliament and of the Council concerning Community statistics on income and living conditions (EU-SILC) as regards the sampling and tracing rules. Available from: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32003R1982>

-
- European Commission (2007). Measurement of Homelessness at European Union Level.
- (2011). On EU indicators in the field of youth. Commission staff working document. Available from http://ec.europa.eu/assets/eac/youth/library/publications/indicator-dashboard_en.pdf
 - (2014). Report on the implementation of the UN Convention on the Rights of Persons with Disabilities (CRPD) by the European Union
- European Union (2019). Regulation (EU) 2019/1700 of the European Parliament and of the Council of 10 October 2019. Official Journal of the European Union.
- Eurostat (2008). ESSPROS Manual. The European System of Integrated Social Protection Statistics. Methodologies and working papers. Available from: <https://ec.europa.eu/eurostat/documents/3859598/5903173/KS-RA-07-027-EN.PDF/11794cb0-3f17-4058-b785-bc2c4f2c91a5>
- (2011). Implementing core variables in EU social surveys. Methodological guidelines. European Commission. Available from: <https://circabc.europa.eu/sd/a/69497b9d-ec9e-4b59-ad7b-1ab709c57d99/CORE%20VARIABLES%20UPDATED%20GUIDELINES%20May%202011.pdf>
 - (2013a). Income pooling and equal sharing within the household – What can we learn from the 2010 EU-SILC module? Methodologies and Working papers. Luxembourg: European Commission. Available from: <https://ec.europa.eu/eurostat/documents/3888793/5857781/KS-RA-13-013-EN.PDF>
 - (2013b). Handbook on precision requirements and variance estimation for ESS households surveys. Methodologies and Working papers. Luxembourg: European Commission. Available from: <https://ec.europa.eu/eurostat/documents/3859598/5927001/KS-RA-13-029-EN.PDF/>
 - (2014). Methodological Guidelines and Description Of EU-SILC Target Variables, Available from: <https://circabc.europa.eu/sd/a/2aa6257f-0e3c-4f1c-947f-76ae7b275cfe/DOCSILC065%20operation%202014%20VERSION%20reconciliated%20and%20early%20transition%20October%202014.pdf>
 - (2017a). Working group on income and living conditions – household budget survey. Luxembourg: European Commission. Available from: https://circabc.europa.eu/webdav/CircaBC/ESTAT/hbs/Library/working_groups/Working%20group%202017-09-28/LC-ILC-193-17%20HBS%20WG%20Standardisation%20of%20social%20variables.docx.pdf
 - (2017b). Interaction of household income, consumption and wealth – statistics on main results, Statistics Explained Experimental, 15.12.2017. Available from: <http://ec.europa.eu/eurostat/statistics-explained/pdfscache/59548.pdf>
 - (2017c). Handbook on Rapid Estimates. Luxembourg: Publications Office of the European Union. Available from: <https://ec.europa.eu/eurostat/documents/3859598/8555708/KS-GQ-17-008-EN-N.pdf>
 - (2018a). European Health Interview Survey (EHIS wave 3). Methodological manual. Available from: <http://ec.europa.eu/eurostat/documents/3859598/8762193/KS-02-18-240-EN-N.pdf/5fa53ed4-4367-41c4-b3f5-260ced9ff2f6>
 - (2018b). EU Labour Force Survey. Explanatory notes. Luxembourg: European Commission. Available from: <https://ec.europa.eu/eurostat/documents/1978984/6037342/EU-LFS-explanatory-notes-from-2017-onwards.pdf>
 - (2018c). Item 4.5. Standardisation of social variables – Progress Report. Luxembourg: European Commission. Available from: <https://circabc.europa.eu/sd/a/e1116065-18c6-4f5b-b2a4-dc30713ccbd6/DSS-2018-Mar-4.5%20Standardisation%20of%20social%20variables.pdf>
 - (2018d). Methodological manual on territorial typologies. Manuals and guidelines. Available from: <https://ec.europa.eu/eurostat/documents/3859598/9507230/KS-GQ-18-008-EN-N.pdf/a275fd66-b56b-4ace-8666-f39754ede66b>

-
- (2018e). Flash estimates of income inequalities and poverty indicators for 2017 (FE 2017) - Experimental results. <https://ec.europa.eu/eurostat/documents/7894008/8256843/Flash-estimates-of-income-inequalities-and-poverty-indicators-experimental-results-2017.pdf>
 - (2019a). EU-SILC methodology – concepts and contents. Statistics Explained. Available from: <https://ec.europa.eu/eurostat/statistics-explained/pdfscache/35201.pdf>
 - (2019b). Social transfers in kind for education and health - imputation into EU - SILC data. Conference of European Statisticians Expert meeting on measuring poverty and inequality: SDGs 1 and 10 Geneva, Switzerland 5–6 December 2019. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/6._Eurostat.pdf
 - (2020). To be added once published.
- Eurostat/OECD (2014a). Guidelines on the measurement of educational attainment according to ISCED 2011. Joint Eurostat – OECD guidelines on the measurement of educational attainment in household surveys. Available from: [https://circabc.europa.eu/sd/a/89d0582c-742d-48c4-859d-8c3930ec9df8/Guidelines%20on%20EA%20version%20September%202014%20final%20\(2014-09-09\).pdf](https://circabc.europa.eu/sd/a/89d0582c-742d-48c4-859d-8c3930ec9df8/Guidelines%20on%20EA%20version%20September%202014%20final%20(2014-09-09).pdf)
- Eurostat/OECD (2014b). Eurostat statistics – classification of ISCED level 3 programmes of partial level completion in EU surveys and indicators. Joint Eurostat – OECD guidelines on the measurement of educational attainment in household surveys. Annex for EU countries. Available from: [https://circabc.europa.eu/sd/a/3ffb91ba-725f-4da1-a7b6-9ce3f6ecc23e/Guidelines%20on%20EA%20-%20Annex%20for%20EU%20final%20\(2014-07-17\).pdf](https://circabc.europa.eu/sd/a/3ffb91ba-725f-4da1-a7b6-9ce3f6ecc23e/Guidelines%20on%20EA%20-%20Annex%20for%20EU%20final%20(2014-07-17).pdf)
- FAP and FEANTSA (2015). An Overview of Housing Exclusion in Europe. Available from: <https://www.feantsa.org/en/report/2015/10/19/an-overview-of-housing-exclusion-in-europe-2015?bcParent=27>
- Fay, R.E. and Herriot, R.A. (1979). Estimation of Income from Small Places: An Application of James-Stein Procedures to Census Data. In: *Journal of the American Statistical Association*, 74: 269-277.
- Fellegi, I. P. (2001). Comment. In: *Journal of Official Statistics*, 17: 43-50.
- Ferragina, E., Tomlinson, M., & Walker, R. (2013). Poverty, participation and choice. JRF, Joseph Rowntree Foundation. Available from: <http://www.jrf.org.uk/publications/poverty--participation--and--choice>.
- Fesseau, M., Wolff, F. and Mattonetti, M. L. (2013). A Cross-Country Comparison of Household Income, Consumption and Wealth between Micro Sources and National Accounts Aggregates. OECD Statistics Working Paper 2013/03. Organisation for Economic Co-operation and Development, Paris.
- Flik, R. J. and van Praag, B. (1991). Subjective poverty line definitions. In: *De Economist*, 139/3: 311– 331.
- Fontaine M. and Fourcot J. (2015). Nowcasting du taux de pauvreté par la micro-simulation. Institut National de la Statistique et des Études Économiques, Série des documents de travail de la Direction des Statistiques Démographiques et Sociales. Available from: <https://www.insee.fr/fr/statistiques/1304142>
- Fontenot, K., Semega, J. and Kollar, M. (2018). Income and Poverty in the United States: 2017. Current population reports. United States Census Bureau. United States Department of Commerce. Economics and Statistics Administration. Available from: <https://www.census.gov/content/dam/Census/library/publications/2018/demo/p60-263.pdf>
- Foster, J., Seth, S., Lokshin, M. and Sajaia, Z. (2013). A unified approach to measuring poverty and inequality: Theory and practice. The World Bank. Available from: <https://openknowledge.worldbank.org/bitstream/handle/10986/13731/9780821384619.pdf?sequence=1&isAllowed=y>
- Fouarge, D. and Layte, R. (2005). Welfare Regimes and Poverty Dynamics: The Duration and Recurrence of Poverty Spells in Europe. In: *Journal of Social Policy*, 34/03: 407-426. Available from: <https://doi.org/10.1017/S0047279405008846>.

-
- Fox, L. (2017). The Supplemental Poverty Measure: 2016. Current Population Reports. United States Census Bureau. Available from:
<https://www.census.gov/content/dam/Census/library/publications/2017/demo/p60-261.pdf>
- (2018). The supplemental poverty measure: 2017. Current population reports. United States Census Bureau. United States Department of Commerce. Economics and Statistics Administration. Available from:
<https://www.census.gov/content/dam/Census/library/publications/2018/demo/p60-265.pdf>
- (2019). The supplemental poverty measure: 2018. Current population reports. United States Census Bureau. United States Department of Commerce. Economics and Statistics Administration. Available from:
<https://www.census.gov/content/dam/Census/library/publications/2019/demo/p60-268.pdf>
- FRA (2007). European Union minorities and discrimination survey. Available from:
<https://fra.europa.eu/en/project/2011/european-union-minorities-and-discrimination-survey>.
- FRA (2013). Technical report: FRA survey - Discrimination and hate crime against Jews in EU Member States: experiences and perceptions of antisemitism - Survey methodology, sample and questionnaire. Available from: <http://fra.europa.eu/en/publication/2013/technical-report-fra-survey-discrimination-and-hate-crime-against-jews-eu-member>.
- (2013b). EU LGBT Survey. Technical report. Available from: <http://fra.europa.eu/en/publication/2013/eu-lgbt-survey-technical-report>.
- (2014). Roma Pilot Survey – Technical report: methodology, sampling and fieldwork. Available from:
<https://fra.europa.eu/en/publication/2014/roma-pilot-survey-technical-report-methodology-sampling-and-fieldwork>
- (2017a). Second European Union Minorities and Discrimination Survey (EU-MIDIS II): Technical Report. Available from: <http://fra.europa.eu/en/publication/2017/eumidis-ii-technical-report>.
- (2017b). Second European Union Minorities and Discrimination Survey (EU-MIDIS II): Main results. Available from: <http://fra.europa.eu/en/publication/2017/eumidis-ii-main-results>.
- (2017c). Roma and Travellers Survey 2018-2019. Available from: <https://fra.europa.eu/en/project/2018/roma-and-travellers-survey-2018-2019>
- (2018a). Combating Child Poverty: An Issue of Fundamental Rights. Available from:
https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-combating-child-poverty_en.pdf
- (2018b). Shifting perceptions: towards a rights-based approach to ageing. In FRA Fundamental Rights Report 2018. Available from: https://fra.europa.eu/sites/default/files/fra_uploads/fra-2018-fundamental-rights-report-2018_en.pdf
- Frazer, H. and Marlier, E. (2017). Progress across Europe in the implementation of the 2013 EU Recommendation on “Investing in children: Breaking the cycle of disadvantage” - A study of national policies, European Social Policy Network (ESPN), Brussels: European Commission.
- Frere-Smith, T., Luthra, R. and Platt, L. (2014). Sampling Recently Arrived Immigrants in the UK: Exploring the effectiveness of Respondent Driven Sampling. Institute for Social and Economic Research, No. 2014-25.
- Garner, T. I. and Verbrugge, R. (2009). Reconciling User Costs and Rental Equivalence: Evidence from the U.S. Consumer Expenditure Survey. U.S. Department of Labor, U.S. Bureau of Labor Statistics Office of Price and Index Number Research. BLS Working Paper.
- Garner, T. I. and De Vos, K. (1995). Income sufficiency vs. poverty: Results from the United States and the Netherlands. *Journal of Population Economics*, 8(2): 117–134
- Garner, T. I. and Short, K. (2003). Personal assessments of minimum income and expenses: What do they tell us about “minimum living” thresholds and equivalence scales? In J. A. Bishop and Y. Amiel (eds.): *Inequality, Welfare and Poverty: Theory and Measurement*, Vol. 9 of the Series Research on Economic Inequality, Oxford, UK: Elsevier Science, pp. 191–243.

-
- Garner, T. I. and Short, K. (2004). Economic well-being based on income, consumer expenditures and personal assessments of minimum needs. In J. A. Bishop and Y. Amiel (eds.): *Studies on Economic Well-being: Essays in the Honor of John P. Formby*, Vol. 12 of the Series Research on Economic Inequality, Oxford, UK: Elsevier Science, pp. 319–361.
- Garner, T. I., Gudrais, M. and Short, K. (2016). Supplemental poverty measure thresholds and noncash benefits. Available from: <https://www.bls.gov/pir/spm/smp-thresholds-and-noncash-benefits-brookings-paper-4-16.pdf>
- Ghosh, M., and Rao, J.N.K. (1994). Small area estimation: an appraisal (with discussion). In: *Statistical Science*, 9: 65-93.
- Glaser, T., Kafka, E., Lamei, N., Lyberg, L. and Till, M. (2015). European Comparability and National Best Practices of EU-SILC: A Review of Data Collection and Coherence of the Longitudinal Component. Available from: http://statistik.at/wcm/idc/idcplg?IdcService=GET_PDF_FILE&dDocName=081094
- Goedemé, T. (2013). How much Confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators. In: *Social Indicators Research*, 110/1: 89-110.
- Gordon, D. and Najera, H. (2019). Reply to Santos and colleagues “The importance of reliability in the Multidimensional Poverty Index for Latin America (MPI-LA)”. In: *Journal of Development Studies*. <https://doi.org/10.1080/00220388.2019.1663176>.
- Grunditza, S. (2019) Social transfers in kind for education and health - imputation into EU - SILC data Note by European Commission Eurostat Unit F/4. Economic Commission for Europe Conference of European Statisticians Expert meeting on measuring poverty and inequality: SDGs 1 and 10 Geneva, Switzerland 5–6 December 2019. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/6._Eurostat.pdf
- Guio, A-C., Gordon, D. and Marlier, E. (2012). Measuring material deprivation in the EU indicators for the whole population and child-specific indicators. Luxembourg: Publications Office of the European Union.
- Guio, A-C., Gordon, D., Marlier, E., Fahmy, E., Nandy, S. and Pomati, M. (2016). Improving the measurement of material deprivation at EU level. In: *Journal of European Social Policy*, 26/3: 219-333.
- Guio, A-C., Gordon, D. and Marlier, E. (2017a). Measuring child material deprivation in the EU. In Atkinson, A., Guio, A-C. and Marlier, E. (eds). *Monitoring social exclusion in Europe*. Luxembourg: Publications Office of the European Union: 209-224.
- Guio, A-C., Gordon, D. Catalan, H.N. and Pomati, M. (2017b). Revising the EU material deprivation variables. Luxembourg: Publications Office of the European Union.
- Guio, AC., Gordon, D, Marlier, E, Catalan, H.N. and Pomati, M. (2018). Towards an EU measure of child deprivation. In: *Child Indicators Research*, 11/3: 835–860.
- Groves R. M., Fowler Jr. F. J., Couper M. P., Lepkowski, J. M., Singer, E. and Tourangeau, R. (2004). *Survey Methodology*. John Wiley & Sons.
- Guggisberg, M. and Häni, S. (2018). Does the inclusion of assets resolve apparent contradictions between income poverty and other indicators on the standard of living? Paper prepared for the UNECE Expert meeting on measuring poverty and inequality, Vienna, Austria, 29-30 November 2018.
- Guggisberg, M., Häni, S. and Fleury, S. (2013). Poverty Measurement in Switzerland. Paper prepared for the UNECE Seminar “The way forward in poverty measurement”, Geneva, Switzerland, December 2-4, 2013.
- Guggisberg, M, Häni, S. and Fleury, S (2019). Dynamics of poverty in Switzerland: current and future analyses by the Federal Statistical Office. Economic Commission for Europe Conference of European Statisticians Expert meeting on measuring poverty and inequality: SDGs 1 and 10 Geneva, Switzerland, 5–6 December 2019. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/13._Switzerland.pdf

-
- Guio, A.-C. and Van den Bosch, K. (2018). Deprivation among couples: sharing or unequal division? Draft paper for the Net-SILC3 International Conference Athens, April 2018.
- Hagenaars, A., de Vos, K. and Zaidi, M.A. (1994). Poverty Statistics in the Late 1980s: Research Based on Micro-data, Office for Official Publications of the European Communities. Luxembourg.
- Harand, J. and Kerstin, A. (2019). Inklusives und barrierefreies Design der “Teilhabebefragung” in Privathaushalten. In: Berufliche Rehabilitation, 33/2.
- Haughton, J. and Khandker, S. R. (2009). Handbook on poverty+ inequality. World Bank Publications. Available from: <http://documents.worldbank.org/curated/en/488081468157174849/pdf/483380PUB0Pove101OFFICIAL0USEONLY1.pdf>
- Haveman, R. and Wolff, E. N. (2004). The Concept and Measurement of Asset Poverty: Levels, Trends and Composition for the U. S., 1983-2001. In: The Journal of Economic Inequality, 2/2: 145-169.
- Headey, B., Krause, P. and Wagner, G. G. (2009): Poverty redefined as low consumption and low wealth, not just low income: psychological consequences in Australia and Germany. Paper prepared for the Joint OECD-University of Maryland Conference, “Measuring Poverty, Inequality and Social Exclusion: Lessons from Europe” Paris, France, March 16-17, 2009.
- Heckathorn, D. D. (1997). Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations. In: Social Problems, 44/2: 174-199.
- Heisz, A. and Phipps-Burton, S. (2018). Accounting for housing wealth in low income statistics for Canada. Paper prepared on UNECE Expert meeting on measuring poverty and inequality, Vienna, Austria, 29-30 November 2018.
- Henderson, C. R. (1950) Estimation of Genetic Parameters. In: Annals of Mathematical Statistics, 21: 309-310.
- Heuberger R., Glaser T. and Kafka E. (2013). The use of register data in the Austrian SILC survey, in Markus Jäntti, Veli-Matti Törmälehto and Eric Marlier (2013, eds.), The use of registers in the context of EU-SILC: challenges and opportunities, Luxembourg: Publications Office of the European Union, 2013.
- Heuberger, R. and Knittler, K. (2018). Armut und Erwerbsarbeit – ein neuer Indikator. In: Statistische Nachrichten, 03/2018: 226-241.
- Hoddinott, J. and Haddad, L. (1991). Household Expenditures, Child Anthropometric Status and the Intrahousehold Division of Income: Evidence from the Cote d'Ivoire.
- Husmanns, R., Mehran, F. and Verma, V. (1990). Surveys of the Economically Active Population, Employment, Unemployment and Underemployment, International Labour Organisation, Geneva.
- Iacovou, M. and Lynn, P. (2013), Implications of the EU-SILC following rules, and their implementation, for longitudinal analysis. ISER Working Paper Series 2013-17, Institute for Social and Economic Research. Available from: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2013-17.pdf>
- ILO (2003). Report II: Household income and expenditure statistics. Seventeenth International Conference of Labour Statisticians Geneva, 24 November-3 December 2003. Available from: <https://www.ilo.org/public/english/bureau/stat/download/17thicls/r2hies.pdf>
- INSEE - Institut National de la Statistique et des Études Économiques (2015). Des indicateurs précoces de pauvreté et d'inégalités - Résultats expérimentaux pour 2014. Available from: <https://www.insee.fr/fr/statistiques/1304063>
- (2018). Estimation avancée du taux de pauvreté et des indicateurs d'inégalités. Available from: <https://www.insee.fr/fr/statistiques/3623841>
- Jäntti, M. and Törmälehto, V. M. (2013). Survey- and register-based estimates of income distribution and poverty, in Jäntti M., Törmälehto V.M. and Marlier E. (2013, eds.). The use of registers in the context of EU-SILC: challenges and opportunities, Luxembourg: Publications Office of the European Union, 2013.

-
- Jäntti, M., Törmälehto, V. M. and Marlier, E. (2013, eds.). The use of registers in the context of EU–SILC: challenges and opportunities, Luxembourg: Publications Office of the European Union, 2013. Available from: <https://ec.europa.eu/eurostat/documents/3888793/5856365/KS-TC-13-004-EN.PDF>
- Jenkins, S. P. (1991). Poverty measurement and the within-household distribution: Agenda for action. In: *Journal for Social Policy*, 20/4: 457-483.
- Johnson, P. D., Renwick, T. and Short, K. (2010). Estimating the Value of Federal Housing Assistance for the Supplemental Poverty Measure. SEHSD Working Paper #2010-13. United States Census Bureau. Available from: <https://www.census.gov/library/working-papers/2010/demo/SEHSD-WP2010-13.html>
- Juran J. M. and Gryna F. M. (1970). *Quality Planning and Analysis*. McGraw Hill.
- Kaczmarczyk, P. (2013). Money for Nothing? Ukrainian Immigrants in Poland and their Remitting Behaviours. IZA Discussion Paper No. 7666, October 2013.
- Kalton, G. (1983). *Introduction to survey sampling*, 35. Sage.
- Kalton, G., McMillen, D. and Kazprzyk, D. (1986). Non-sampling error issues in SIPP. In: *Proceedings of the Bureau of the Census Second Annual Research Conference Washington, D.C.*: 147-164.
- Kalton, G. (2009), Methods for oversampling rare subpopulations in social surveys. *Survey methodology*, 35, 125–141.
- Kapteyn, A., Kooreman, P., Willemse, R. (1988). Some methodological issues in the implementation of subjective poverty definitions. In: *The Journal of Human Resources*, 23/2: 222– 242
- Kish, L. (1965), *Survey Sampling*. Wiley.
- (1987a). *Statistical Research Design*, New York: Wiley.
- (1988). Multipurpose sample designs. In: *Survey Methodology*, 14: 19-32.
- (1992). Weighting for unequal Pi. In: *Journal of Official Statistics*, 8: 183-200.
- Kozyreva, P., Kosolapov, M. and Popkin, B.M. (2016). Data Resource Profile: The Russia Longitudinal Monitoring Survey—Higher School of Economics (RLMS-HSE) Phase II: Monitoring the Economic and Health Situation in Russia, 1994–2013. In: *International Journal of Epidemiology*: 395–401.
- Krieger, M. (2018). Literature Review: Measuring Intra-Household Decision-Making. Unpublished working document of the UNECE Task Force on measuring household power and decision-making.
- Kuypers, S. and Marx, I. (2016). Estimation of Joint Income-Wealth Poverty: A Sensitivity Analysis. Paper prepared for the 34th General Conference of The International Association for Research in Income and Wealth, Dresden, Germany, 21-27 August 2016.
- Lamei, N. and Skina-Tabue, M. (2018). What is the role of unequal sharing of resources within households? Presentation for the UNECE Expert Meeting on Measuring Poverty and Inequality, Geneva, Switzerland, 30 November 2018.
- League of Arab States, OPHI, UNICEF and UN-ESCWA (2017). Arab Multidimensional Poverty Report. Available from: https://ophi.org.uk/wp-content/uploads/LAS_et_al_2017_Arab_MP_Report_ENG.pdf
- Lee D. and Shon A. (2001). Korea’s experiences in statistical quality assessment. *Proceedings, Statistics Canada Symposium “Achieving data quality in a statistical agency: a methodological perspective”*.
- Lechene, V., Krishna, P. and Wolf, A. (2019). OLS Estimation of the Intra-Household Distribution of Consumption. Unpublished Manuscript. Available from: http://www.sfu.ca/~pendakur/papers/LPW_2019_1.pdf
- Lemaître, G. and Dufour, J. (1987). An integrated method for weighting persons and families. In: *Survey Methodology*, 13: 199-207
- Little, R. J. A. (1986). Survey Nonresponse Adjustments for Estimates of Means. *International Statistical Review*, 54/2: 139-157.
- Lohr, S. L. (1999). *Sampling: Design and Analysis*. Duxbury Press, Pacific Grove, CA.

-
- Luiten, A. and Schouten, B. (2019). Impact of panel attrition on representativeness of EU-SILC. Paper presented at Net-SILC3 International Best Practice Workshop “Unit non response and weighting” and “Item, non-response and imputation” University of Essex, United Kingdom, 20-22 February, 2019.
- Lundberg, S.J., Pollak, R.A. and Wales, T.J. (1997). Do Husbands and Wives Pool Their Resources? Evidence from the United Kingdom Child Benefit. In: *The Journal of Human Resources*, XXXII/3: 463-480.
- MacInnes, T., Tinson, A., Gaffney, D., Goretti, H. and Baumberg, B. (2014). Disability, long term condition and poverty. A report for the Joseph Rowntree Foundation. New policy Institute. Available from: https://www.npi.org.uk/files/7814/0490/1005/Disability_long_term_conditions_and_poverty.pdf
- Mack, J. and Lansley, S. (1985). *Poor Britain*. London: George Allen & Unwin.
- Mader K., Schneebaum A., Skina-Tabue M., Till-Tentschert U. (2012). Intrahaushaltsverteilung von Ressourcen. Geschlechtsspezifische Verteilung von Einkommen und Entscheidungsmacht. In: *Statistische Nachrichten*, 12/2012: 983-994.
- Mader, K. and Schneebaum, A. (2013). Zur geschlechtsspezifischen Intrahaushaltsverteilung von Entscheidungsmacht in Europa. In: *Wirtschaft und Gesellschaft*, 39/3: 361-403. Available from: <http://epub.wu.ac.at/4867/>
- Meraner, A., Gumprecht, D. and Kowarik, A. (2016). Weighting Procedure of the Austrian Microcensus using Administrative Data. In: *Austrian Journal of Statistics*, 45/3: 3-14.
- Marpsat, M. and Razafindratsima, N. (2010), Survey methods for hard-to-reach populations: Introduction to the special issue. In: *Methodological Innovations Online*, 5/2: 3–16.
- Martin, H. (2017). Calculating the standard of living of a household: One or several equivalence scales? *Economics and Statistics*, 2017 (491-492): 93–108.
- Mattonetti, M. L. (2013). *European Household Income by Groups of Households*. Eurostat Methodologies and Working Papers, Publications Office of the European Union, Luxembourg.
- McGuinness, F. (2018). *Poverty in the UK: statistics*. Briefing paper. House of Commons Library. Available from: <https://researchbriefings.files.parliament.uk/documents/SN07096/SN07096.pdf>
- McKenzie, D. and Mistiaen, J. (2009). Surveying migrant households: A comparison of census- based, snowball, and intercept point surveys. *Journal of the Royal Statistical Society*, 172/2: 339–360.
- Morciano, M., R. Hancock, and S. Pudney. 2012. “Disability Costs and Equivalence Scales in the Older Population.” ISER Working Paper Series 2012-09, Institute for Social and Economic Research, University of Essex, U.K. Available from: <https://www.iser.essex.ac.uk/research/publications/working-papers/iser/2012-09.pdf>
- Mohanty, A (2019). *Poverty Dynamics: An Overview of Longitudinal Poverty Estimates Produced by the United States Census Bureau*. U.S. Census Bureau1 | SEHSD Working Paper 2019-38 issued on November 12, 2019 Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/14._US_Census_Bureau_Abinash_Mohanty.pdf
- Mysíková, M., Želinský, T. Garner, T. Fialová K. (2019). Subjective Income Poverty and Equivalence Scales in Eastern vs Western European countries. Economic Commission for Europe Conference of European Statisticians Expert meeting on measuring poverty and inequality: SDGs 1 and 10 Geneva, Switzerland 5–6 December 2019. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/15._US_Bureau_of_Labor_Statistics.pdf
- Najera, H. and Gordon, D. (2019). The importance of reliability and construct validity in multidimensional poverty measurement: An illustration using the Multidimensional Poverty Index for Latin America (MPI-LA). In: *Journal of Development Studies*. <https://www.tandfonline.com/doi/full/10.1080/00220388.2019.1663176>
- National Statistical Service of the Republic of Armenia (2014). *Social Snapshot and Poverty in Armenia: Statistical Analytical Report*. Yerevan.

-
- (2016). Social Snapshot and Poverty in Armenia. Available from:
https://www.armstat.am/file/article/poverty_2016_eng_00.pdf
 - Navamkuel, E. L., Rubiera Morrollon, F. and Fernandez Vazquez, E. (2018). Does the Urban Population Pay More for Food? Implications in Terms of Poverty. *Applied Spatial Analysis and Policy*. Available from:
<https://doi.org/10.1007/s12061-018-9254-x>.
 - Nicaise, I. and Schockaert, I. (2014). The hard-to-reach among the poor in Europe. Lessons from Eurostat's EU-SILC survey in Belgium. In: Tourangeau, R., Edwards, B., Johnson, T. P., Wolter, K. M. and Bates, N. (Eds.) (2014). *Hard-to-Survey Populations*. Cambridge: Cambridge University Press.
 - OECD (2013). *OECD Framework for Statistics on the Distribution of Household Income, Consumption and Wealth*, Paris: OECD Publishing.
 - (2014). *Society at a Glance 2014: OECD Social Indicators*. Paris, France: OECD Publishing.
 - (2017a). *Linking Indigenous Communities with Rural and Regional Development*. Available from:
<http://www.oecd.org/cfe/regional-policy/Indigenous-Communities-project.pdf>
 - (2017b). *Migration Policy Debates. How does having immigrant parents affect the outcomes of children in Europe?* Available from: <http://www.oecd.org/els/mig/How-does-having-immigrant-parents-affect-the-outcomes-of-children-in-Europe.pdf>
 - (2018). *Linking Indigenous Communities with Rural and Regional Development*. Book forthcoming.
 - ONS (2012). *Measuring Benefits in Kind - Methodological Changes in the Measurement of Benefits in Kind, 2005/06-2010/11*. Available from:
https://webarchive.nationalarchives.gov.uk/20160109114622/http://www.ons.gov.uk/ons/dcp171766_273847.pdf
 - (2015). *Nowcasting household income in the UK: initial methodology*. Available from:
<https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/nowcastinghouseholdincomeintheukinitialmethodology/2015-07-01>
 - (2018a). *The Effects of Taxes and Benefits on Household Income, 2016/17*. Available from:
<https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/theeffectsoftaxesandbenefitsonhouseholdincome/financialyearending2017>
 - (2018b). *The Effects of Taxes and Benefits On Household Income, financial year ending 2017: technical report*. Available from:
<https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/articles/theeffectsoftaxesandbenefitsonhouseholdincome/technicalreport>
 - (2018c). *Effects of taxes and benefits on UK household income – flash estimate: financial year ending 2018*. Available from:
<https://www.ons.gov.uk/releases/theeffectsoftaxesandbenefitsonhouseholdincomeflashestimatefinancialyearending2018>
 - Open Society Foundations (2014). *Ethnic Origin and Disability Data Collection in Europe: Measuring Inequality, Combating Discrimination. Policy Report: Equality Data Initiative (November)*, Open Society European Policy Institute, Brussels. Available from: <https://www.opensocietyfoundations.org/sites/default/files/ethnic-origin-and-disability-data-collection-europe-20141126.pdf>.
 - Orshansky, M. (1969). How poverty is measured. In: *Monthly Labor Review*, 92/2: 37-41.
 - Pac, J., Waldfogel, J. and Wimer, C. (2017). *Poverty Among Foster Children: Estimates Using the Supplemental Poverty Measure*. In: *Social Service Review*, 91/1: 8-40.
 - Pantazis, C., Gordon, D. and Levitas, R. (2006). *Poverty and Social Exclusion in Britain*. Bristol: The Policy Press.
 - Pelletiere, D. (2008). *Getting to the Heart of Housing's Fundamental Question: How Much Can A Family Afford? A Primer on Housing Affordability Standards in U.S. Housing Policy*. National Low Income Housing Coalition.
 - Piachaud, D. (1981). Peter Townsend and The Holy Grail. In: *New Society*: 419-21.

-
- Ponthieux, S. (2017). Intra-household pooling and sharing of resources: a tentative “modified” equivalised income. In: Atkinson, A.B., Guio, A.-C., Marlier, E. (2017): Monitoring Social Inclusion in Europe. Eurostat: 175-189.
- Pratt, B. M., Hixson, L. and Jones, N. A. (2015). Measuring race and ethnicity across the decades: 1790 – 2010. United States Census Bureau. Available from: <https://www.census.gov/newsroom/blogs/random-samplings/2015/11/measuring-race-and-ethnicity-across-the-decades-1790-2010.html>
- Psihoda, S. Hora, K. Till, M. Schober, C. Sprajcer, S. Baumgartner, I. Freudensprung, M. , Groß, M. Gruber, E. Hinner, K. Hobel, b. Kopolnai, U. Paškvan, M. Piribauer, J. Ritter, V. Sturmlehner, E., Unger, M. Zechmeister, E. (2018) „Eine Stimme für alle“ – Barrieren in Forschung und Sozialstatistik abbauen, Research Report Statistics Austria. Vienna. Available from: http://www.statistik.at/wcm/idc/idcplg?IdcService=GET_PDF_FILE&RevisionSelectionMethod=LatestReleased&dDocName=119914
- Purcell, N. J. and Kish, L. (1980). Postcensal Estimates for Local Areas (Small Domains). In: International Statistical Review, 48: 3-18.
- Raghunathan, T. E., Lepkowski, J. L., Van Hoewyk, J. H., Solenberger, P. W. (2001). A multivariate technique for imputing the missing values using a sequence regression models. In: Survey Methodology, 27: 85-95.
- Rao J. N. K. and Molina, I. (2015). Small Area Estimation. John Wiley & Sons, Inc., Hoboken, New Jersey.
- Ravazzini, L., Kuhn, U., Brulé, G. and Suter, C. (2017). The influence of wealth on poverty and inequality in Switzerland. Paper prepared for the UNECE Expert Meeting on measuring poverty and inequality, Budva, Montenegro, 26-27 September 2017
- Ray, R. (1983). Measuring the cost of children: An alternative approach. In: Journal of Public Economics, 22: 89–102
- Rio Group. (2006). Compendium of best practices in poverty measurement. Expert Group on Poverty Statistics. Available from: https://www.ibge.gov.br/poverty/pdf/rio_group_compendium.pdf
- Reichel, D. and Morales, L. (2017). Surveying immigrants without sampling frames – evaluating the success of alternative field methods. In: Comparative Migration Studies, 5/1. Available from: <https://doi.org/10.1186/s40878-016-0044-9>.
- Relais 1.0. User's guide, Istat. Available from: http://www.istat.it/strumenti/metodi/software/analisi_dati/relais/
- Renwick, T. (2010). Improving the Measurement of Family Resources in a Modernized Poverty Measurement. Paper presented at the January 2010 Annual Convention of the Allied Social Sciences Association. Available from: <https://www.census.gov/content/dam/Census/library/working-papers/2010/demo/renwicksge2010.pdf>
- Renwick, T., Aten, B. H., Figueroa, E. B. and Martin, T. (2014). Supplemental Poverty Measure: a comparison of geographic adjustments with regional price parities vs. median rents from the American Community Survey. SEHSD Working Paper No. 2014-22. United States Census Bureau. Available from: <https://www.census.gov/library/working-papers/2014/demo/SEHSD-WP2014-22.html>
- Renwick, T. and Mitchell, J. (2015). Estimating the value of federal housing assistance for the supplemental poverty measure. United States Census Bureau. Social, Economic and Housing Statistics Division. SEHSD Working Paper # 2017-38. Available from: <https://www.census.gov/content/dam/Census/library/working-papers/2017/demo/SEHSD-WP2017-38.pdf>
- Renwick, T. and Garner, T. I. (2016). Changing the Housing Share of Poverty Thresholds for the Supplemental Poverty Measure: Does Consumer Unit Size Matter? SEHSD Working Paper #22018-06. Available from: <https://www.census.gov/content/dam/Census/library/working-papers/2018/demo/SEHSD-WP2018-06.pdf>
- Renwick, T., Figueroa, E. B. and Aten, B. H. (2017). Supplemental Poverty Measure: a comparison of geographic adjustments with regional price parities vs. median rents from the American Community Survey: An Update. Available from: <https://www.census.gov/library/working-papers/2017/demo/SEHSD-WP2017-36.html>

-
- Rodrigues, R., Huber, M. and Lamura, G. (2012). Facts and Figures on Healthy Ageing and Long Term Care. Available from: https://ec.europa.eu/eip/ageing/library/facts-and-figures-healthy-ageing-and-long-term-care_en
- Rosstat (2017). Development of methodology in the field of measuring inequality and poverty in transition to new sources of information: Experiences and problems. Conference of European Statisticians Sixty-fifth plenary session Geneva, 19-21 June 2017. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_33-New_methodologies_for_new_sources.pdf.
- Ruggeri Laderchi, C. (1997). Poverty and its many dimensions: The role of income as an indicator. In: Oxford Development Studies 25/3: 345-360.
- Rust, K. F. and Rao, J. N. K. (1996), Variance estimation for complex surveys using replication techniques. In: Statistical Methods in Medical Research, 5: 283-310.
- Särndal, C.-E., B. Swensson and J. Wretman (1992). Model Assisted Survey Sampling. Springer. New-York.
- Schouten, B., Cobben F. and Bethlehem J. (2009). Indicators for the representativeness of survey response. In: Survey Methodology, June 2009, 35/1: 101-113.
- Sen, A. (1980). Equality of What?. In: McMurrin, S. T. (Eds.) (1980). Lectures on Human Values, Volume 1. Cambridge: Cambridge University Press.
- SERISS Report WP2 (2017). Report on sampling practices for the institutionalized population in social surveys. Available from: https://seriss.eu/wp-content/uploads/2017/02/SERISS_WP2_D2.16_Sampling-the-institutionalized-population.pdf
- Shelton, A. (2015). Fact Sheet. Social Security: Who's counting on it? AARP Public Policy Institut. Available from: <https://www.aarp.org/content/dam/aarp/ppi/2015/whos-counting-on-social-security-FS.pdf>
- Statistics Austria (2017). Broken series and coherent communication - Lessons learnt in Austria. Note prepared for the 65. Plenary Session of the Conference of European Statisticians, Geneva, Switzerland, 19-21 June, 2017. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/2017/CES_31-Poverty_measurement_lessons_learnt__Austria_.pdf
- Statistics Belgium (2002). G-CALIB; Generalised Calibration under SPSS®. Statistics Belgium, Release 1.0. April 2002.
- Statistics Canada (2003). Survey Methods and Practice. Ottawa: Ministry of Industry, Statistics Canada, Social Survey Division.
- (2016). A New Survey Measure of Disability: The Disability Screening Questions (DSQ). Available from: <https://www150.statcan.gc.ca/n1/pub/89-654-x/89-654-x2016003-eng.htm>
- (2017a). G-Link A Probabilistic Record Linkage System. Available from: <https://pdfs.semanticscholar.org/abd4/47d16fa837714d582ccd43b18d2285980a8f.pdf>
- (2017b). Ethnic Origin Reference Guide, Census of Population, 2016. Available from: <https://www12.statcan.gc.ca/census-recensement/2016/ref/guides/008/98-500-x2016008-eng.cfm>
- (2019). Canadian Income Survey (CIS). Available from: <https://crdcn.org/datasets/cis-canadian-income-survey>
- SERISS (2017). Work Package 2 Deliverable 2.16: Report on sampling practices for the institutionalised population in social surveys. Available from: https://seriss.eu/wp-content/uploads/2017/02/SERISS_WP2_D2.16_Sampling-the-institutionalized-population.pdf
- Stevens, K., Fox, L. and Heggeness, M. (2018). Does Source Matter? Using State-Level SNAP Administrative Records and the Transfer Income Model (TRIM3) to Evaluate Poverty Measurement. Poster/paper prepared for the Population Association of America annual meeting.
- Stewart, F., Harriss-White, B. and Saith, R. (Eds.) (2007). Defining Poverty in the Developing World. Palgrave Macmillan, UK.

-
- Schepers, W., Juchtmans, G. and Nicaise, I. (2017). Reaching out to hard-to survey groups among the poor. Survey protocols, statistical issues and research design (Milestone 20.5). Leuven: InGrID. Available from: <https://inclusivegrowth.be/downloads/output/ms87-m20-5-hard-to-survey-groups.pdf>.
- Signore, M. (2019). Final methodological report summarizing the results of WP 1-5. Istat. Available from: <https://www.istat.it/it/files//2011/07/MIMOD-project-Final-report-WP1-WP5.pdf>
- Stiglitz, J. E., Sen, A. and Fitoussi, J. P. (2009). Report by the Commission on the Measurement of Economic Performance and Social Progress, Available from: <http://ec.europa.eu/eurostat/documents/118025/118123/Fitoussi+Commission+report>
- Stiglitz, J., Fitoussi, J. and Durand, M. (2018a). Beyond GDP: Measuring What Counts for Economic and Social Performance, OECD Publishing, Paris, <https://doi.org/10.1787/9789264307292-en>.
- Stiglitz, J., J. Fitoussi and M. Durand (eds.) (2018b), For Good Measure: Advancing Research on Well-being Metrics Beyond GDP, OECD Publishing, Paris, <https://doi.org/10.1787/9789264307278-en>.
- Stokes, S., Reiser, C., Compton, E., Bentley, M. and Rothhaas, C. (2012). 2010 Census Alternative Questionnaire Experiment (AQE): Census 2000 Form Replication Panel. Available from: https://www.census.gov/content/dam/Census/library/publications/2012/dec/2010_cpex_169.pdf
- Stoyanova, S. and Tonkin, R. (2016). Nowcasting Household Income in the UK: Financial Year Ending 2015. Presented at the 34th International Association for Research on Income and Wealth General Conference. Available from: <http://www.iariw.org/dresden/stoyanova.pdf>
- Talland, L. and Evans, J. (2019). What is the difference between sex and gender? Exploring the difference between sex and gender, looking at concepts that are important to the Sustainable Development Goals. United Kingdom Office for National Statistics. Available from: <https://www.ons.gov.uk/economy/environmentalaccounts/articles/whatisthedifferencebetweensexandgender/2019-02-21/pdf>
- The Washington Group on Disability Statistics and UNICEF (2018). Module on Child Functioning. A new way to measure child functioning. Available from: <https://data.unicef.org/resources/module-child-functioning/>
- Thomas, D. (1990). Intra-Household Resource Allocation: An Inferential Approach. In: Journal of Human Resources, University of Wisconsin Press 25/4: 635-664.
- Till, M. and Eiffe, F. (2010). Progress of living conditions—a dynamic model of material deprivation for a European society. Income and living conditions in Europe, Eurostat: 241-264. Available from: <https://ec.europa.eu/eurostat/documents/3217494/5722557/KS-31-10-555-EN.PDF>
- Till, M. and Keindl, S. (2015). The Austrian Inclusion Indicators – Putting Multidimensional Poverty Measurement into Practice. Working paper presented at the UNECE Seminar on poverty measurement, Geneva, Switzerland, 5-6 May 2015. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2015/WP_19_EN_Austria_rev.pdf
- Tonkin, R., Lewis, J. and Thomas, N. (2013). Measuring Social Transfers in Kind in the United Kingdom. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2013/WP_11_UK_D_En.pdf
- Tonkin, R., Serafino, P. and Davies, R. E. (2016). Using statistical matching to facilitate the comparison of poverty estimates using income, consumption and wealth. Paper prepared for the UNECE Seminar on poverty measurement, Geneva, Switzerland, 12-13 July 2016.
- Tonkin, R., Tormalehto, V. M., Thomas, N., Lewis, J. and Kajantie, M. (2014). Social Transfers in Kind in the United Kingdom and Finland: Micro-level Measuring and Distributional Impact. Paper presented at the IARIW 33rd General Conference, 24-30 August 2014. Rotterdam, the Netherlands. Available from: <http://www.iariw.org/papers/2014/tonkinPaper.pdf>.
- Torabi, M., & Rao, J. N. K. (2014). On small area estimation under a sub-area level model. Journal of Multivariate Analysis, 127, 36-55.

-
- Törmälehto, V. M., and Sauli, H. (2013). The distributional impact of imputed rent in EU-SILC 2007-2010. Eurostat Methodologies and Working Papers, European Commission.
- Törmälehto, V. M., and Sauli, H. (2017). The distributional impact of imputed rent in EU-SILC 2007-2012. Eurostat. 2017. Monitoring social inclusion in Europe. Eurostat Statistical Books. Ch.7, pp. 141-157
- Törmälehto, V. M., Kannas, O. and Säylä, M. (2013). Integrated measurement of household-level income, wealth and non-monetary well-being in Finland. Working Papers 1/2013, Statistics Finland. Available from: http://www.stat.fi/tup/julkaisut/tiedostot/julkaisuluettelo/ywrrp1_201300_2013_10518_net.pdf (16.03.2018).
- Törmälehto, V. M. (2019). Reconciliation of EU statistics on income and living conditions (EU-SILC) data and national accounts. Statistical Working Paper. Eurostat. Available from: <https://ec.europa.eu/eurostat/documents/3888793/9959642/KS-TC-19-004-EN-N.pdf/cd90cd0f-ebcf-43b6-ab4b-c28bcd44f46>
- Tourangeau, R., Johnson, T., Bates, N. and Wolter, K. (Ed.) (2014). Hard-to-Survey Populations. Cambridge University Press.
- Townsend, P. (1979). Poverty in the United Kingdom, Harmondsworth: Penguin Books.
- Trindade Z. L. and Goedemé, T. (2016). Notes on updating the EU-SILC UDB sample design variables 2012-2014, CSB Working Paper 16/02, Antwerp: Herman Deleeck Centre for Social Policy, University of Antwerp.
- UNDP (2019). Human Development Report 2019 Beyond income, beyond averages, beyond today: Inequalities in human development in the 21st century. Available from: <http://www.hdr.undp.org/en/2019-report>
- UNDP and Ipsos (2019). Regional Roma Survey 2017 technical report. Available from: <https://www.eurasia.undp.org/content/rbec/en/home/library/roma/regional-roma-survey-2017-technical-report.html>
- UNDP and OPHI (2019). How to Build a National Multidimensional Poverty Index (MPI): Using the MPI to inform the SDGs. United Nations Development Programme (UNDP) and Oxford Poverty and Human Development Initiative (OPHI), University of Oxford. Available from: https://ophi.org.uk/wp-content/uploads/How_to_Build_Handbook_2019_PDF.pdf
- UNECE (2010). Developing Gender Statistics: A Practical Tool. Available from: http://www.unece.org/fileadmin/DAM/stats/publications/Developing_Gender_Statistics.pdf
- (2011). Canberra Group Handbook on Household Income Statistics: Second Edition. Geneva: United Nations. Available from: https://www.unece.org/fileadmin/DAM/stats/groups/cgh/Canberra_Handbook_2011_WEB.pdf
- (2015a). Conference of European Statisticians. Recommendations for the 2020 Censuses of Population and Housing. New York and Geneva: United Nations. Available from: https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECECES41_EN.pdf
- (2015b). Indicators of Gender Equality. New York and Geneva: United Nations. Available from: https://www.unece.org/fileadmin/DAM/stats/publications/2015/ECE_CES_37_WEB.pdf
- (2016). Recommendations on Ageing-related Statistics. New York and Geneva: United Nations. Available from: http://www.unece.org/fileadmin/DAM/stats/publications/2016/ECECESSTAT20164_ENG_web.pdf
- (2017a). Guide on Poverty Measurement. New York and Geneva: United Nations. Available from: <https://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20174.pdf>
- (2017c). A world free from child poverty. A guide to the tasks to achieve the vision. Available from: https://www.unicef.org/media/49406/file/Child_Poverty_SDG_Guide-Complete-March_2017.pdf
- (2019). In-depth Review on Measuring Gender Identity. Prepared by Canada and the United Kingdom. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/bur/2019/February/03_In-depth_review_of_Measuring_Gender_Identity.pdf

-
- (2020). Recommendations for Measuring Intra-household Power and Decision-making. Available from <http://www.unece.org/statistics/publications.html>
 - UNFPA (2014). Methodological Guidelines for the Gender Analysis of National Population and Housing Census Data. Available from: <https://www.unfpa.org/sites/default/files/pub-pdf/141006-UNFPA-GenderManual2014-02-SCREEN.pdf>
 - United Kingdom Government Statistical Service (2017). Harmonised Concepts and Questions for Social Data Sources. GSS Harmonised Principle. Ethnic group. United Kingdom Office for National Statistics. Available from <https://gss.civilservice.gov.uk/wp-content/uploads/2017/08/Ethnic-Group-June-17.pdf>
 - UNSD (2014). Fundamental Principles of Official Statistics. Available from https://unstats.un.org/unsd/dnss/hb/E-fundamental%20principles_A4-WEB.pdf
 - UN (2001). Statistics on Special Population Groups. Guidelines and Principles for the Development of Disability Statistics. Department of Economics and Social Affairs. UN Statistics Division. New York and Geneva: United Nations. ST/ESA/STAT/SER.Y/10. Available from: https://unstats.un.org/unsd/publication/SeriesY/SeriesY_10e.pdf
 - (2005). Handbook on Poverty Statistics: Concepts, Methods and Policy Use. New York: United Nations Statistics Division, United Nations. Available from: https://unstats.un.org/unsd/methods/poverty/pdf/un_book%20final%2030%20dec%2005.pdf
 - (2015a). Resolution 70/1 of UN General Assembly. Transforming our world: the 2030 Agenda for Sustainable Development. Available from: <https://undocs.org/A/RES/70/1>
 - (2015b). The world's women 2015, Trends and Statistics. Department of Economic and Social Affairs. New York: United Nations. ST/ESA/STAT/SER.K/20. Available from: https://unstats.un.org/unsd/gender/downloads/worldswomen2015_report.pdf
 - (2016a). Integrating a gender perspective into statistics. Department of Economic and Social Affairs. New York: United Nations. Studies in Methods, Series F No 111. Available from: <https://unstats.un.org/unsd/demographic/standmeth/handbooks/05323%20Integrating%20a%20Gender%20Perspective%20into%20Statistics%20Web%20Final.pdf>
 - (2016b). Inter-Agency and Expert group on the Sustainable Development Goal Indicators. Working group on geospatial information. Available from: <http://ggim.un.org/documents/ToR%20WG%20on%20Geospatial%20Information%20Final.pdf>
 - (2017a). Principles and Recommendations for Population and Housing Censuses. Revision 3. New York: United Nations. Available from: https://unstats.un.org/unsd/publication/seriesM/Series_M67Rev3en.pdf
 - (2017b). Report of the United Kingdom of Great Britain and Northern Ireland on ageing-related statistics and age-disaggregated data. Note by the Secretary-General. Economic and Social Council. E/CN.3/2018/19. Available from: <https://unstats.un.org/unsd/statcom/49th-session/documents/2018-19-Ageing-E.pdf>
 - (2017c). Disability statistics. Joint report of the Secretary-General and the Washington Group on Disability Statistics. Economic and Social Council. E/CN.3/2018/17. Available from: <https://unstats.un.org/unsd/statcom/49th-session/documents/2018-17-Disability-E.pdf>
 - (2017d). Report of the Expert Group on Refugee and Internally Displaced Persons Statistics. Note by the Secretary-General. Economic and Social Council. E/CN.3/2018/16. Available from: <https://unstats.un.org/unsd/statcom/49th-session/documents/2018-16-Refugee-E.pdf>
 - (2017e). Improving Migration Data in the Context of the 2030 Agenda. United Nations Expert Group Meeting. UN Statistics Division. Available from: https://unstats.un.org/unsd/demographic-social/meetings/2017/new-york--egm-migration-data/EGM%20Recommendations_FINAL.pdf
 - (2017f). Resolution adopted by the General Assembly on 6 July 2017. General Assembly. A/RES/71/313. Available from: https://digitallibrary.un.org/record/1291226/files/A_RES_71_313-EN.pdf

-
- (2017g). Guidelines on measuring asset ownership from a gender perspective. Evidence and data for gender equality (EDGE). Regional Workshop on the UN Methodological Guidelines on the Production of Statistics on Asset Ownership from a Gender Perspective. UN Statistics Division and UN Women. Available from: <https://unstats.un.org/edge/meetings/July2018/docs/6%20Data%20Sources.pdf>
 - (2018a). Report on the forty-ninth session. Note by the Secretary-General. Statistical Commission. Economic and Social Council. Available from: <https://unstats.un.org/unsd/statcom/49th-session/documents/Report-on-the-49th-session-E.pdf>
 - (2019a). Data disaggregation and SDG indicators: Policy priorities and current and future disaggregation plans. Prepared by the Inter-Agency and Expert Group on Sustainable Development Goal Indicators (IAEG-SDGs). Statistical Commission. Available from: <https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3a-Data-Disaggregation-E.pdf>
 - (2019b). World Survey on the Role of Women in Development. Report of the Secretary-General Why addressing women's income and time poverty matters for sustainable development. General Assembly. Seventy-fourth session. Available from: <https://undocs.org/en/A/74/111>
 - (2020). Report on the fifty-first session. Note by the Secretary-General. Statistical Commission. Economic and Social Council. Available from: <https://unstats.un.org/unsd/statcom/51st-session/documents/2020-34-DisabilityStats-E.pdf> UN Women (2019). Progress of the World's Women 2019-2020. Families in a changing world. Available from: <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2019/progress-of-the-worlds-women-2019-2020-en.pdf?la=en&vs=3512>
- UN Women and World Bank (2018). Gender Differences in Poverty and Household Composition through the Life Cycle. Available from: <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2018/sdg-report-spotlight-01-gender-differences-in-poverty-and-household-composition-en.pdf?la=en&vs=4813>
- United Nations Children's Fund (2015). Wellbeing of Children in Kosovo (UNSCR 1244), Poverty and Deprivation among Children using the Multiple Overlapping Deprivation Analysis (MODA). Available from: https://www.unicef.org/kosovoprogramme/media/171/file/MODA_ENG.pdf
- (2017b). Child poverty in Europe and Central Asia region: definitions, measurement, trends and recommendations. Available from: <https://www.unicef.org/eca/media/3396/file/Child-poverty-regional-report.pdf>
- OHCHR (2008). Convention on the Rights of Persons with Disabilities. Available from: <https://www.ohchr.org/EN/HRBodies/CRPD/Pages/ConventionRightsPersonsWithDisabilities.aspx>
- OHCHR (2018). A Human Rights-Based Approach to Data. Leaving no one behind in the 2030 Agenda for Sustainable Development. Available from: <https://www.ohchr.org/Documents/Issues/HRIndicators/GuidanceNoteonApproachtoData.pdf>
- United States Census Bureau (2014). American Community Survey Design and Methodology. Available from: https://www2.census.gov/programs-surveys/acs/methodology/design_and_methodology/acs_design_methodology_report_2014.pdf
- United States Census Bureau (2018). Questions Planned for the 2020 Census and American Community Survey. Federal legislative and Program uses. United States Department of Commerce. Economics and Statistics Administration. Available from: <https://www2.census.gov/library/publications/decennial/2020/operations/planned-questions-2020-acs.pdf>
- Vanderhoeft, C. (2002). G-Calib Generalised Calibration Under SPSS. Statistics Belgium.
- Vaalavuo, Maria (2015): Poverty Dynamics in Europe: From What to Why. European Commission Working Paper 03/2015. Available from: <http://ec.europa.eu/social/BlobServlet?docId=14776&langId=en>
- Verma V. and Betti, G. (2006). EU Statistics on Income and Living Conditions (EU-SILC): Choosing the survey structure and sample design. *Statistics in Transition*, 7/5: 935-970.

-
- (2011). Taylor linearization sampling errors and design effects for poverty measures and other complex statistics. In: *Journal of Applied Statistics*, 38/8: 1549-1576.
- Verma V., Betti, G. and Gagliardi F. (2010). An assessment of survey errors in EU-SILC, Eurostat Methodologies and Working papers, Luxembourg: Publications Office of the European Union.
- Verma, V. (1981). Assessment of errors in household surveys. In: *Bulletin of the International Statistical Institute*, 49/2: 905-919.
- (2001). Leslie Kish: Development of Statistics Internationally. Proceedings of Statistics Canada Symposium 2001 Achieving Data Quality in a Statistical Agency: A Methodological Perspective. Available from: <https://www150.statcan.gc.ca/n1/pub/11-522-x/2001001/session4/6242-eng.pdf>
- Verma, V., Gagliardi, F. and Ferretti, C. (2013). Cumulation of poverty measures to meet new policy needs. In: Torelli, N., Pesarin, F. and Bar-Hen, A. (2013, eds.). *Advances in Theoretical and Applied Statistics*. XIX, Springer
- Verma, V. (2014). Sampling elusive populations: Applications to studies of child labour. Geneva, International Labour Office.
- Wargan, K. and Dershem, L. (2009). Don't Call Me a Street Child Estimation and Characteristics of Urban Street Children in Georgia. Available from: <https://resourcecentre.savethechildren.net/node/6347/pdf/6347.pdf>
- Weber, D. (2019) Trends in Persistent Poverty in the United Kingdom. Economic Commission for Europe Conference of European Statisticians Expert meeting on measuring poverty and inequality: SDGs 1 and 10 Geneva, Switzerland 5–6 December 2019. Available from: https://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.15/2019/mtg2/12._UK.pdf
- Weisbrod, B. A. and Lee Hansen, W. (1968). An Income-Net Worth Approach to Measuring Economic Welfare. In: *The American Economic Review*, 58/5: 1315-1329.
- Wolter, K. M. (2007). *An Introduction to Variance Estimation*. Springer.
- World Bank (2017). Monitoring Global Poverty: Report of the Commission on Global Poverty. Available from <https://elibrary.worldbank.org/doi/book/10.1596/978-1-4648-0961-3>
- (2018). *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle*. Washington, DC: World Bank.
- World Health Organization, United Nations Economics and Social Commission for Asia and the Pacific (2008). Training manual on disability statistics. Available from: <https://www.unescap.org/sites/default/files/Front-matter-Disability-Statistics-E.pdf>
- Ziliak, J. P. (2011). Cost of Living and the Supplemental Poverty Measure: a research forum submitted to the Office of the Assistant Secretary for Planning and Evaluation, U.S. Department of Health and Human Services. http://www.ukcpr.org/sites/www.ukcpr.org/files/Supplemental_poverty_measures.pdf
- Van Oyen, H.; Bogaert, R; Yokota, T.C; Berger, V. (2018) Measuring disability: a systematic review of the validity and reliability of the Global Activity Limitations Indicator (GALI). *Archives of Public Health* 76.1: 25. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5985596/>
- Zimmermann, L. (2012). Reconsidering gender bias in intrahousehold allocation in India. In: *The Journal of Development Studies* 48/1: 151–163. Available from: <https://www.tandfonline.com/doi/full/10.1080/00220388.2011.629652>

Poverty Measurement Guide to Data Disaggregation

This publication provides guidance on applying various approaches to data disaggregation for measuring poverty and aims to improve the international comparability of poverty statistics. Regardless of how accurate the national measures of poverty are, their usefulness will be limited if they mask existing disparities within societies.

The introductory **Chapter 1** lays down the motivation behind the choice of disaggregation for poverty measurement.

Chapter 2 suggests a standard set of core variables for poverty disaggregation with the purpose to enhance further international harmonisation. Reference questions are provided to illustrate the implementation of data disaggregation in practice.

Chapter 3 discusses methods to adapt data collection in order to reach beyond the traditional and established survey methodologies and capture the groups that are most exposed to the risk of poverty.

Chapter 4 addresses the need for assessment of data accuracy and the quality of statistical processes as well as their continuous improvement. The chapter discusses how to improve response rates and sampling precision in the measurement of poverty among relevant social groups.

Chapter 5 describes individual country experiences with adjustments for difference in cost of living across regions and household circumstances. It provides analysis on measurement issues such as social transfers in kind, disability cost or high medical expenditures, housing wealth, imputed rent, assets poverty and unequal sharing of resources within households, and multi-dimensional poverty.

Chapter 6 presents an overview of the areas envisaged for further work.

The publication mainly targets national statistical authorities and provides useful information for policymakers, researchers and other users of poverty data.

Information Service
United Nations Economic Commission for Europe

Palais des Nations
CH - 1211 Geneva 10, Switzerland
Telephone: +41(0)22 917 12 34
E-mail: unece_info@un.org
Website: <http://www.unece.org>

ISBN 978-92-1-117260-7



9 789211 172607