

28 November 2019

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

Item C of the provisional agenda

Improving response rate and sampling precision in surveys

Draft Report of the Task Force on Disaggregated Poverty Measures

Summary

This working paper contains the consolidated Task Force report in its current state.

Table of Contents

| | |
|---|----------|
| List of Figures | V |
| List of Tables | VII |
| List of Boxes | VIII |
| Acronyms / Abbreviations | 1 |
| 1 Introduction | 1 |
| 1.1 Disaggregation as choice - why this report can be useful | 1 |
| 1.2 Background | 2 |
| 1.3 Outline of the Guide | 3 |
| 2 Standard core variables for disaggregation | 5 |
| 2.1 Policy relevant target groups for poverty disaggregation | 5 |
| 2.2 Defining variables for poverty disaggregation | 12 |
| 2.2.1 Sex | 13 |
| 2.2.2 Age | 14 |
| 2.2.3 Disability status | 15 |
| 2.2.4 Migratory status | 19 |
| 2.2.5 Ethnicity | 21 |
| 2.2.6 Household type | 24 |
| 2.2.7 Current employment status | 26 |
| 2.2.8 Tenure status of the household | 28 |
| 2.2.9 Receipt of cash (or near cash) social transfers | 30 |
| 2.2.10 Degree of urbanisation | 31 |
| 2.2.11 Educational attainment level | 33 |
| 2.3 Examples of applications to the poverty disaggregation | 35 |
| 2.3.1 Example 1. Gender and poverty measurement: approaches and challenges | 35 |
| 2.3.2 Example 2: Disability and more appropriate poverty measures (United Kingdom) | 37 |
| 2.3.3 Example 3: Disaggregation and the use of alternative poverty measures (United States) | 38 |
| 2.3.4 Example 4: Joint disaggregation of child poverty | 40 |
| 2.3.5 Example 5: Presentation of poverty across population groups | 46 |
| 2.3.6 Example 6: Poverty status and receipt of social transfers | 48 |
| 2.3.7 Example 7: Interactive presentation of poverty indicators | 48 |
| 2.3.8 Example 8: Poverty and housing tenure (United Kingdom) | 50 |
| 2.4 Recommendations | 50 |

| | | |
|-------|---|----|
| 3 | Including Hard to Reach Groups in Poverty Measurement | 52 |
| 3.1 | Introduction | 52 |
| 3.2 | Improving inclusiveness of poverty surveys | 53 |
| 3.2.1 | Fieldwork approaches for hard to reach populations | 53 |
| 3.2.2 | Principles to facilitate surveying vulnerable groups | 53 |
| 3.2.3 | Interviewing hard to reach groups..... | 54 |
| 3.2.4 | Training and field work materials | 55 |
| 3.3 | Dedicated survey programmes for hard to reach groups..... | 55 |
| 3.3.1 | Design modifications to include hard to reach groups in standard surveys..... | 56 |
| 3.3.2 | Special survey programmes | 58 |
| 3.3.3 | Sampling target groups | 58 |
| 3.4 | Poverty among children | 69 |
| 3.4.1 | Street children..... | 69 |
| 3.4.2 | Children with disabilities..... | 70 |
| 3.4.3 | Foster children | 70 |
| 3.5 | Homelessness and housing difficulties in the European Union..... | 71 |
| 3.5.1 | Policy justification | 71 |
| 3.5.2 | How to define homelessness | 72 |
| 3.5.3 | Sample Issues | 73 |
| 3.5.4 | Implementation in EU-SILC | 74 |
| 3.5.5 | Plans for the revised EU-SILC | 76 |
| 3.5.6 | Conclusion..... | 76 |
| 3.6 | Surveying Institutional Households | 76 |
| 3.6.1 | Case Study: UK - including non-private household populations (NPHP) | 79 |
| 3.7 | Recommendations | 83 |
| 4 | Assessing and Improving Survey Methods | 84 |
| 4.1 | Survey errors and quality | 84 |
| 4.1.1 | A typology of survey errors..... | 84 |
| 4.1.2 | Errors in estimation..... | 86 |
| 4.1.3 | Non-response errors | 87 |
| 4.1.4 | Multidimensional quality frameworks..... | 88 |
| 4.1.5 | Questions which should be addressed to evaluate the utility of a survey | 88 |
| 4.1.6 | Relationship between different aspects of quality | 92 |
| 4.2 | Improving Quality in each Survey Step | 93 |

| | | |
|-------|---|-----|
| 4.2.1 | Survey design | 93 |
| 4.2.2 | Data collection | 100 |
| 4.2.3 | Data processing and imputation | 104 |
| 4.2.4 | Weighting | 106 |
| 4.2.5 | Variance estimation | 113 |
| 4.2.6 | Dissemination..... | 115 |
| 4.3 | Other methodological issues related to measuring poverty | 118 |
| 4.3.1 | Pooling | 118 |
| 4.3.2 | Small Area Estimation | 121 |
| 4.3.3 | Rapid estimates..... | 122 |
| 5 | Improving accuracy of measurement: the use of supplemental or experimental poverty measures..... | 123 |
| 5.1 | Adjustments to poverty thresholds | 123 |
| 5.1.1 | Spatial differences within countries with regard to consumption and income poverty .. | 124 |
| 5.1.2 | Household size and age of household members | 130 |
| 5.1.3 | Disability status/High medical expenditures | 133 |
| 5.2 | Adjustments to resources | 134 |
| 5.2.1 | Social transfers in kind (STIK)..... | 134 |
| 5.2.2 | Accounting for housing wealth | 142 |
| 5.2.3 | Asset poverty: Why assets should be considered in the measurement of poverty and how it could be done | 147 |
| 5.2.4 | What is the role of unequal sharing of resources within households? | 153 |
| 5.2.5 | Going beyond income: The role of multidimensional poverty measures..... | 161 |
| 5.3 | Recommendations | 166 |
| 5.3.1 | Adjustments to thresholds..... | 166 |
| 5.3.2 | Adjustments to resources | 166 |
| | References | 169 |

List of Figures

| | |
|--|-----|
| Figure 2.1 United States example of measuring race and ethnicity..... | 23 |
| 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 | 36 |
| Figure 2.3 Proportion of families by material deprivation and disability status | 38 |
| Figure 2.4 Percentage of People in Poverty by Different Poverty Measures, 2017 | 39 |
| Figure 2.5 Change in Number of People in Poverty After Including Each Element, 2017 (in millions) | 40 |
| Figure 2.6 At risk-of-poverty rates for the total population, households with dependent children and households with three or more children | 40 |
| 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..... | 41 |
| Figure 2.8 Children in the EU Member States aged 0-17 years at risk of poverty, citizenship of their parents, 2016 (in %) | 42 |
| Figure 2.9 Venn Diagram of Deprivation Overlaps between Nutrition, Water, Housing, by urban and rural areas, children aged 0-4 years..... | 43 |
| Figure 2.10 Venn Diagram of Deprivation Overlaps between Education, Water, Housing, by urban and rural areas, children aged 5-10 years..... | 44 |
| Figure 2.11 Venn Diagram of Deprivation Overlaps between Water, Sanitation, Housing, by urban and rural areas, children aged 5-10 years..... | 44 |
| Figure 2.12 Venn Diagram of Deprivation Overlaps between Education, Water, Housing, by urban and rural areas, children aged 11-17 years..... | 45 |
| Figure 2.13 Venn Diagram of Deprivation Overlaps between Housing, Sanitation, Water, by urban and rural areas, children aged 11-17 years..... | 45 |
| Figure 2.14 Poverty among the basic sectors in the Philippines | 46 |
| Figure 2.15 Program Participation Status of Household – Poverty Status of People in 2017 (% of population) | 48 |
| Figure 2.16 At risk of poverty or social exclusion, Austria, 2014..... | 49 |
| Figure 2.17 Percentage of people (all ages) in relative low income, by housing tenure, United Kingdom, 2016-17..... | 50 |
| Figure 2.18 Percentage of children in relative low income, by housing tenure, United Kingdom, 2016-17..... | 50 |
| Figure 3.1 Sample Question from Survey | 67 |
| Figure 3.2 Self-coded reasons to go abroad in Western Balkans (n=1136 micro-narratives).68 | |
| Figure 3.3 Self-coded reasons to go abroad in Bosnia and Herzegovina (n=203 micro-narratives)..... | 68 |
| Figure 3.4 Pattern of responses for unfavorable treatment based on gender, ethnicity and age, n=510 | 69 |
| Figure 4.1 Increasing Sampling Error/Bias Ratios for disaggregations | 87 |
| Figure 4.2 EU-SILC standardised rotating scheme | 99 |
| Figure 4.3 Distribution of ACS Interviews and Non-interviews | 102 |
| Figure 4.5 Replication Variance Estimation Method for three consecutive years | 120 |

| | |
|--|-----|
| Figure 5.1 Impact of geographic adjustments on poverty rates, 2018 | 124 |
| 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 | 126 |
| Figure 5.3 Relationship of deprivation index and income among two-person households .. | 129 |
| Figure 5.4 Poverty Rate with and without Care Allowances for Austria, 2008 | 134 |
| Figure 5.5 Social transfers in kind by income quintile group, United Kingdom, 2016/17 | 137 |
| Figure 5.6 Original, disposable, post-tax and final income by quintile groups, all households, financial year ending 2016/17 | 138 |
| Figure 5.7 Relative at-risk-of-poverty rates, United Kingdom, 2011/12 | 139 |
| Figure 5.8 Low-income rates by age of household head, Canada 2016 | 146 |
| Figure 5.9 Impact of Differential Thresholds on the U.S. Supplemental Poverty Measure .. | 147 |
| Figure 5.10 Illustration of the two-dimensional poverty index | 148 |
| Figure 5.11 Share of income poor that do not have sufficient liquid assets, by poverty line and reference period, in per cent of the population (individuals), with 95%- confidence interval | 149 |
| Figure 5.12 Share of income poor that do not have sufficient liquid assets for three months, by subgroups, in per cent of the population (individuals), with 95%-confidence interval | 150 |
| Figure 5.13 Share of population indicating material or subjective difficulties, by poverty status, in per cent of the population (individuals), with 95%-confidence interval | 151 |
| Figure 5.14 The standard household income model – exemplary household | 154 |
| Figure 5.15 Dimensions of deprivation for households | 163 |

List of Tables

| | |
|--|-----|
| Table 2.1 Educational attainment levels used in Mexico's National Household Income and Expenditure Survey..... | 33 |
| Table 3.1 Summary of sampling approaches in previous Roma surveys..... | 62 |
| Table 3.2 EU SILC variables on past episodes of homelessness and housing difficulties..... | 75 |
| Table 3.3 Estimated size of each NPHP sub-population in the UK | 80 |
| Table 3.4 Proposed approaches to capture data for the different categories of NPHP | 81 |
| Table 4.1 Concepts and categories used by different organizations to identify dimensions of quality | 88 |
| Table 4.2 Use of administrative data and registers for each domain covered by EU-SILC (2012)..... | 94 |
| Table 4.3 ACS Data Collection Consists of Three Overlapping Phases | 101 |
| Table 4.4 Demographic Tools used by the Canadian Income Survey | 111 |
| Table 5.1 Proposal of new equivalence scales for Tajikistan..... | 132 |
| Table 5.2 At-risk-of-poverty by sex I | 157 |
| Table 5.3 At-risk-of-poverty by sex II | 158 |

List of Boxes

| | |
|--|-----|
| Box 2.1 Approaches to disaggregation of SDG indicators in a global framework | 8 |
| Box 2.2 National responses to improving disaggregation of SDG indicators – example from the United Kingdom..... | 9 |
| Box 2.3 Regional approaches for harmonizing variables for disaggregation across statistical domains – Eurostat’s project on standardization of social variables..... | 9 |
| Box 2.4 OECD survey of current practices to measure ethnicity, race and indigenous identity | 10 |
| Box 2.5 Detailed age disaggregation of poverty measures for children | 14 |
| Box 2.6 United States’ experience with measuring race and ethnicity..... | 23 |
| Box 3.1 Surveying marginalized Roma communities in Slovakia..... | 56 |
| Box 3.2 Case Study: UNDP – Sampling approaches in previous Roma Surveys | 62 |
| Box 3.3 Case Study – UNICEF: Measuring the Roma population..... | 63 |
| Box 3.4 Case Study – UNDP: Micronarratives Roma in the Western Balkans | 66 |
| Box 3.5 Case Study: USA - Measuring Poverty for the Institutionalized Population..... | 77 |
| Box 3.6 Case Study: European Union Agency for Fundamental Rights | 78 |
| Box 4.1 Framework to assess errors in poverty measurements (Verma et al, 2010) | 87 |
| Box 4.2 European Union: Content of Quality Reports as Required by the Law | 89 |
| Box 4.3 Country Case: Use of Administrative Tax Data at Statistics Canada..... | 93 |
| Box 4.4 Country Case: Use of Registers in the European Union..... | 94 |
| Box 4.5 Country Case: Record Linkage through Statistical Identifiers in Austria | 95 |
| Box 4.6 Country Case: Design of the American Community Survey in the United States..... | 97 |
| Box 4.7 Country Case: Canadian Income Survey Design | 98 |
| Box 4.8 Country Case: EU-SILC Sample Design | 99 |
| Box 4.9 Country Case: Collection Methodology of the American Community Survey..... | 100 |
| Box 4.10 Country Case: Calibration in the Canadian Income Survey | 110 |
| Box 4.11 Country Case: Calibration in Household Living Condition Survey 2009 in the Ukraine..... | 113 |
| Box 4.12 Country Case: Calculating Variance for EU-SILC in Austria | 115 |
| Box 4.13 Country Case: United States Suppression Rules..... | 117 |
| Box 5.1 Assessment of the 60 per cent of median threshold for the elderly in Austria | 128 |
| Box 5.2 Two country examples for the estimation of an equivalence scale based on a demand model – Tajikistan and Ukraine..... | 132 |
| Box 5.3 Approaches to the estimation of imputed rent..... | 143 |
| Box 5.4 The National Multidimensional Poverty Index (MPI) of Armenia | 162 |

Acronyms / Abbreviations

| | |
|-----------|---|
| ABS | Australian Bureau of Statistics |
| ACS | American Community Survey |
| ASEC | Annual Social and Economic |
| BEA | Bureau of Economic Analysis |
| BPA | Beijing Platform for Action |
| BRR | Balanced Repeated Replication |
| CPS ASEC | Current Population Survey Annual Social and Economic Supplement |
| 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 | European Union Statistical Office |
| EU-SILC | European Statistics on Income and Living Conditions |
| FEANTSA | European Federation of National Organisations Working with the Homeless |
| FRA | EU Agency for Fundamental Rights |
| GPSDD | Global Partnership for Sustainable Development Data |
| GQ | Group Quarters |
| HBAI | Households Below Average Income |
| 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 |
| 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 |
| 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 |

| | |
|--------|---|
| STIK | Social Transfers in Kind |
| SPM | Supplemental Poverty Measure |
| UK GSS | United Kingdom Government Statistical Service |
| UN | United Nations |
| UNECE | United Nations Economic Commission for Europe |
| UNICEF | United Nations Children's Fund |
| UNSC | United Nations Statistical Commission |

1 Introduction

1.1 Disaggregation as choice - why this report can be useful

The Guide on Poverty Measurement (UNECE, 2017) 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”. It is the intention of this report to provide further guidance to consolidate the measurement of poverty as well as inspiration for the practice of statistical offices.

Official poverty measures are only a subset of the poverty measures which could potentially be produced. Measurement of poverty is a description of the social world. Amartya Sen (Sen, 1980) proposed to understand description as choice. For a good description it is not sufficient to reflect the reality. At the same time, a good description should also not be lost in accurate but useless detail.

For example, it is still a true description that some countries are poorer than others. However, it would be misleading to assume that all people in a poor country are poor. The distributions of income and wellbeing increasingly overlap 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. The discrepancies between regions and population groups become even more significant when absolute poverty tends to disappear in developing countries while relative poverty may be on the rise because large parts of the population benefit from rapidly increasing levels of prosperity. With ongoing progress, it becomes more important to identify fault lines within countries.

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).

Disaggregation and measurement of poverty cannot be reduced to a mere relative or subjective value judgement which rests solely in “the eye of the beholder” (Orshansky, 1969). For example, the European Union has established a legal framework for the measurement of poverty in the Community Statistics on Income and Living Conditions (EU-SILC). This reflects (rather than imposes) ideas about economic growth and employment policy as much as about redistribution or social and human rights. Indicators on poverty 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.

Disaggregation has practical consequences. The potential spread of disease among victims of natural disaster or future development of property prices and labour migration are all related to poverty. Its measurement and disaggregation can aid interventions out of humanitarian motives as much as of mere economic reasoning. Measurement of poverty may hence be called good, if it supports the effective allocation of various resources such as food banks, medicines, teachers or law enforcement forces.

International standards may provide an important point of reference for government and non-government stakeholders when no specific policy objectives are yet 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 to the greatest benefit. 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 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.

Sometimes international standards may be in conflict with capacities as well as the moral or instrumental motivations and priorities present within a single community. To be able to produce good descriptions, statisticians therefore do not only have to select between currently known alternatives in poverty measurement. Their choices also involve the development for new methodologies for future descriptions and better disaggregation which will truly leave no one behind. Especially if national poverty measurement departs from international conventions, it is essential to acknowledge and make users aware of any impacts on quality.

1.2 Background

The UNECE Statistical Division has been involved in poverty statistics starting with an in-depth review of this topic prepared by the Bureau of the Conference of European Statisticians (CES Bureau) in 2012. Since then, expert meetings and capacity-building workshops on poverty measurement were held regularly on an annual basis. In addition, task forces mandated by the CES and composed of UNECE member countries 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, 2017). The Guide provided an overview of 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 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 measure at 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.

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 to establish 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 worked through 2018-2019 to develop **the present Guide on disaggregated poverty measures**.

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 with the purpose of improving poverty-related measures at 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

The ethical and practical purpose of disaggregating poverty measures have at least four implications which also define the structure of this report. Firstly, there should be clarity about the definition of the groups to be considered. Secondly, every effort must be taken to ensure coverage and include groups which 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 which would be obtained from supplemental measures to assess their robustness. Where this appears necessary, complementary measures need to be published together with the conventional measures.

2 provides methodological and practical guidance on standard core variables for poverty disaggregation with the purpose to enhance further international harmonisation. The chapter offers recommendations to improve quality, relevance and use of data, consistent with international human rights norms and principles for 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 a possible implementation and ways of operationalization of data disaggregation in practice.

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.

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. **4** provides analysis on how response rates and sampling precision may be improved in the measurement of poverty among relevant social groups in the UNECE region. 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.

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. 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 accuracy of measurement. Error! Reference source not found. describes individual country experiences with adjustments for difference in cost of living across regions and household circumstances and measurement issues such as social transfer in kind, disability cost or high medical expenditures, housing wealth, imputed rent, assets poverty and unequal sharing of resources within households.

2 Standard core variables for disaggregation

Co-ordinator: Jakub Hrkal

Other contributors:

Trudi Renwick (United States)

Dominic Webber (United Kingdom)

Haoyi Chen (UNSD)

Lara Fleischer, Carlotta Balestra (OECD)

Joanne Bosworth (UNICEF)

Agata Kaczmarek-Firth (EUROSTAT)

Teresa Munzi (LIS)

This section provides methodological and practical guidance on standard core variables with the purpose of enhancing further international harmonisation of disaggregation of poverty-related measures. 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. 2.2 provides guidance on definitions and operationalization of standard core variables recommended for purposes of disaggregation. 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

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, 2017)

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, the principles of self-identification (freedom to self-identify) and participation (involvement of all relevant stakeholders) are arguably the most important.

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.

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.

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.

The list of selected variables describing target groups covered in this report is the following:

- **Sex** (target group of women);
- **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 defined by ethnicity or race);

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).

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 report, such as in 3.

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 Inter-Agency and Expert Group on SDG Indicators (IAEG-SDGs)¹ (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 (UN, 2017a; UNECE, 2015a);
- **Sample surveys:** The Intersecretariat Working Group on Household Surveys² established by the United Nations Statistical Commission (UNSC), the International Household Survey Network³, or Eurostat project on standardization of social variables (see Box 2.3);
- **International sample survey programmes or data collections with an important income or poverty component:** European Statistics on Income and Living Conditions (EU-SILC)⁴ or Living Standards Measurement Study (LSMS);⁵
- 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;⁶ 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);

¹ <https://unstats.un.org/sdgs/iaeg-sdgs/>

² <https://unstats.un.org/iswghs/>

³ <http://www.ihsn.org/>

⁴ <http://ec.europa.eu/eurostat/web/income-and-living-conditions/overview>

⁵ <http://surveys.worldbank.org/lms>

⁶ <http://www.unece.org/statistics/networks-of-experts/task-force-on-measuring-old-age-population-in-institutions.html>

- **Migration:** On-going activities of the Expert Group on Refugee and Internally Displaced Persons Statistics (UNECE, 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),⁷ and the work of the UNECE on gender statistics;⁸
- 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);
- 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 in recommendations for disaggregation of SDG monetary poverty indicators for ECE countries (UNECE, 2017: 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-

⁷ <https://unstats.un.org/unsd/demographic-social/gender/>

⁸ <http://www.unece.org/statistics/areas-of-work/statsoc/gender-statistics.html>

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 – example from the United Kingdom

As part of the United Kingdom's commitment to the "Leaving No One Behind" agenda, the Office for National Statistics (ONS) are 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 opposite to sharing a dwelling 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 conceptualisation of ethnicity as related to nationality and ancestry (but not current citizenship). Poland and the Slovak Republic 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, the Slovak Republic, 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, the Slovak Republic, 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 of 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 which collect data on race (Canada, the United Kingdom, and the United States) enable individuals to declare more than one race. 12 out of 17 countries allow multiple responses for ethnicity, except Israel, Latvia, the Slovak Republic, and Slovenia.

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 e.g. response options between census waves like in Chile.

2.2 Defining variables for poverty disaggregation

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 accessibility of certain population groups, such as persons with disabilities. Measures implemented to protect personal data, i.e. guaranteeing confidentiality of data and their exclusive use for statistical purposes, can have an important impact on the participation of some vulnerable groups.

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).⁹ 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.

⁹ The project is ongoing, and the draft document used is subject to further changes to be considered for purposes of poverty disaggregation once available.

2.2.1 Sex¹⁰

2.2.1.1 Background

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 gender¹¹ are not necessarily the same in all cases - as 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 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

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.

The reporting units are all individuals.

2.2.1.3 Categories definition

- Male
- Female

2.2.1.4 Implementation guidelines

Due to the importance of sex characteristic, 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, it should be imputed into the data record by attributing male and female interchangeably.

¹⁰ See more in UNECE, 2015b: 93-94 and Eurostat, 2017a: 6.

¹¹ 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.

2.2.1.5 Reference question

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 Age¹²

2.2.2.1 Background

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.

2.2.2.2 Variable definition

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.

The reporting units are all individuals.

2.2.2.3 Categories definition

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

- 0-17 (children)¹³
- 18-24¹⁴
- 25-49
- 50-64
- 65 and over (older persons).

More detailed age disaggregation would 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

¹² See more in UNECE, 2015: 94 and Eurostat, 2017a: 7.

¹³ 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.

¹⁴ 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).

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.

2.2.2.4 Implementation guidelines

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)

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

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

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 status¹⁵

2.2.3.1 Background

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*

¹⁵ See more in UNECE, 2015a: 154-161.

hinder their full and effective participation in society on an equal basis with others" (OHRHC, 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.

2.2.3.2 Variable definition

A conceptual framework for and classification of disability is provided in the International Classification of Functioning, Disability and Health (ICF).¹⁶ 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.

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 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.

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

There is a variety of policy needs and mandates, traditions and practices in data collections on disability. As a result, there are various international and national instruments for measuring disability and for the disaggregation of SDGs. The current situation was recognized by the UNSC, who called upon countries to examine the underlying concepts, purposes and advantages and choose appropriate measurement tools and instruments depending on their national data needs. Furthermore, the UNSC endorsed the creation of a group of experts for the revision of the UN Guidelines and Principles for the Development of Disability Statistics (UN, 2018a) with an expanded mandate to review methodological instruments and provide guidance for data disaggregation by disability status.

2.2.3.5 Reference question

For the purpose of disaggregating poverty measurement, it is recommended to use a two-question instrument 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?"

¹⁶ <http://www.who.int/classifications/icf/en/>

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".

The following text presents approaches of three international entities that could be used for purposes of disaggregation:

- a) The comprehensive scale used by the World Health Organization (Model Disability Survey),
- b) The Washington Group on Disability Statistics - its short set of disability questions is referred to as a preferable tool in the Guide on Poverty Measurement (UNECE, 2017), and
- c) The ESS, which includes the most pragmatic approach reflected in the above reference question.

More methodological and practical guidance for disability statistics in an international context was further developed, for example, by the UN Statistics Division (UN, 2001)¹⁷ and the World Health Organization and ESCAP (WHO/ESCAP, 2008). 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

The Model Disability Survey¹⁸ 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 version¹⁹ 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 Washington Group on Disability Statistics

The Washington Group on Disability Statistics constituted as a city group under the UNSC developed several modules for measuring disability in censuses and surveys.²⁰ The Short Set of Disability Questions²¹ 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). Disability is measured with six questions, each using four graded answer categories: “no difficulty”, “some difficulty”, “a lot of difficulty”, or “cannot do”. A person with a disability is defined as a person who has a lot of difficulty or who cannot do for at least one of the functioning domains.

In order to address measuring children 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

¹⁷ The guidelines were being updated at the time of writing this report.

¹⁸ <http://www.who.int/disabilities/data/mds/en/>

¹⁹ Idem.

²⁰ <http://www.washingtongroup-disability.com/washington-group-question-sets/>

²¹ Idem.

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.

The European Statistical System and the European Health Interview Survey

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 over²² 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.

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 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.

Additional steps were undertaken to measure the health of children within EU Statistics on Income and Living conditions (Eurostat, 2019). 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

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.

The DSQ comprehensively identify ten distinct disability types and quantify the severity level of each type. 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

²² Some activity limitation questions were also adapted for children and were subject to pretesting at the time of drafting this paper.

validity of each question. This was followed by two major quantitative tests to assess the reliability of the DSQ on surveys with different contexts.

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

Within the 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.

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?"

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.

2.2.4 Migratory status²³

2.2.4.1 Background

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 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 country, and reason for migration. Finally, internal migrants and internally displaced persons could also be considered if there is 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.²⁴

²³ See more in UNECE, 2015: 134-147 and Eurostat, 2017a: 36, 40.

²⁴ These variables are usually complemented by variable Country of residence and Country of birth of father/mother.

2.2.4.2 Variable definition

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.

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.

The reporting units for both variables are all individuals.²⁵

2.2.4.3 Categories definitions

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.

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).

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.

2.2.4.4 Implementation guidelines

The categories used for collection of data are the following:

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

Country codes are recorded according to ISO 3166²⁶ three-letter code (alpha-3).

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

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

²⁵ Specific filters (e.g., aged 15+) may apply depending on the survey target population concerning the variable.

²⁶ <https://www.iso.org/iso-3166-country-codes.html>

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.

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

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?"*

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?"*

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 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 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.²⁷

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 Ethnicity²⁸

2.2.5.1 Background

Statistical information on ethnicity (or ethno-cultural characteristics in general) are 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

²⁷ 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.

²⁸ See more in UNECE, 2015a: 148-153.

need for data may even be driven by legal requirements associated with evaluating anti-discrimination policies.

However, the collection of such data may be prohibited by law in some countries. 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.

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.

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 ongoing OECD work related to national practices of measuring selected aspects of diversity (see **Error! Reference source not found.** above).

2.2.5.2 Variable definition

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.

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.

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

2.2.5.3 Implementation guidelines

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.

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*²⁹ 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 2.6.

Box 2.6

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

²⁹ <http://webarchive.nationalarchives.gov.uk/20160106185805/http://www.ons.gov.uk/ons/guide-method/measuring-equality/equality/ethnic-nat-identity-religion/index.html>

What is this person's race?
Mark ☒ one or more boxes **AND** print origins.

☐ White – Print, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.

☐ Black or African Am. – Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.

☐ American Indian or Alaska Native – Print name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.

| | | |
|--|---|--|
| <input type="checkbox"/> Chinese | <input type="checkbox"/> Vietnamese | <input type="checkbox"/> Native Hawaiian |
| <input type="checkbox"/> Filipino | <input type="checkbox"/> Korean | <input type="checkbox"/> Samoan |
| <input type="checkbox"/> Asian Indian | <input type="checkbox"/> Japanese | <input type="checkbox"/> Chamorro |
| <input type="checkbox"/> Other Asian – Print, for example, Pakistani, Cambodian, Hmong, etc. | <input type="checkbox"/> Other Pacific Islander – Print, for example, Tongan, Fijian, Marshallese, etc. | |

☐ Some other race – Print race or origin.

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 the 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>

2.2.5.4 Reference question

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

2.2.6 Household type³⁰

2.2.6.1 Background

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.

³⁰ See more in UNECE, 2015: 176-177 and Eurostat, 2017a: 16-18

2.2.6.2 Variable definition

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

The reporting units are private households.

2.2.6.3 Categories definition

The following categories are recommended for disaggregation:³¹

- One-person household
- Couple without any children
- Couple with one child under 18
- Couple with two or more children under 18
- Lone parent with children under 18
- Other type of household.

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.

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

- 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.

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.

³¹ Guide of poverty measurement 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.

2.2.6.4 Implementation guidelines

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

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”.

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.³²

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

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

- 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 status³³

2.2.7.1 Background

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.

³² The United States regularly reports separate poverty estimates for female householders and male householders (See Fontenot et al, 2018: Table 4).

³³ See more in UNECE, 2015: 103-111.

2.2.7.2 Variable definition

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

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).

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

2.2.7.3 Categories definitions

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

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

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

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 status³⁴. 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 / Other³⁵.

³⁴ See the definition of derived variable Activity Status (ACTSTA) (Eurostat, 2019)

³⁵ The last five categories compose the disaggregation category 'Other outside the labour force'.

2.2.7.5 Reference question

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

The following question could be used:

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

2.2.8 Tenure status of the household³⁶

2.2.8.1 Background

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.

2.2.8.2 Variable definition

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

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

2.2.8.3 Categories definitions

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

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.

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

³⁶ See more in UNECE, 2015: 179-180 and Eurostat, 2017a: 19-20.

mortgage and only the interest remains outstanding, the owner shall be considered as an owner with outstanding mortgage.

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

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.

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.

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.

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

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

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

The recommended question is:

“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

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.

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 are defined as old-age (retirement) and survivors' (widows' and widowers') pensions; unemployment benefits; family-related benefits; sickness and invalidity benefits; education-related benefits; housing allowances; social assistance; and other benefits (for definitions see Eurostat, 2008). This definition is suggested for use. Specific types of analysis should be made with caution, such as analysis of relation to receipt of social transfers and indicators of poverty before/after social transfers³⁷.
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 non-addressed 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

³⁷ Pensions, such as old-age and survivors' (widows' and widowers') benefits, are counted as income (before social transfers) when constructing the European indicator "at-risk-of-poverty rate before social transfers".

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).

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 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.

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

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

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

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.

The reporting units are private households.

2.2.9.3 Categories definitions

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

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

2.2.10 Degree of urbanisation³⁸

2.2.10.1 Background

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

³⁸ For more details, see UNECE, 2015: 88-89 and Eurostat, 2017a: 51-52.

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).

2.2.10.2 Variable definition

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/municipalities³⁹ based on the share of local population living in urban clusters and in urban centres identified by population size and population density.

The reporting units are all individuals or households.

2.2.10.3 Categories definition

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).

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

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”.

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

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

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

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.⁴⁰

³⁹ For EU and some other countries, it corresponds to LAU2 or communes, see <http://ec.europa.eu/eurostat/web/nuts/local-administrative-units>

⁴⁰ 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

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

The educational attainment level is defined according to the ISCED 2011.⁴¹

2.2.11.3 Categories definitions

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).

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

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) |
| 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, |

⁴¹ <http://www.uis.unesco.org/Education/Pages/international-standard-classification-of-education.aspx> – ISCED-A for levels of educational attainment.

| | |
|------------|--|
| | 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 | • 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 | • 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

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

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.

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).

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

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).

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

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.

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.⁴²

2.3.1 Example 1. Gender and poverty measurement: approaches and challenges⁴³

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. Mainly, this refers to the adequate measurement of women's control over resources (including income in kind, assets, debt or individual pension entitlements). This is because female- and male-headed households cover a broad range of situations in terms of demographic and economic composition. "Examples of female-headed households include one-person households, such as older women with a small or no income and young women who are economically independent, households of lone mothers with children who may or may not receive financial support from the father... Because of their different demographic and economic composition, these types of female-headed households tend to have a different risk of poverty. Similarly, different types of male-headed households have a different risk of poverty depending on their demographic and economic composition. Furthermore, as the share of the detailed types of female- and male- headed household in total households varies across countries and within countries, the overall gap between the poverty of female-headed households and that of male-headed households also varies..." (UN, 2016a).

Household-level data on poverty, measured traditionally based on either household income or consumption, is often disaggregated only by the sex of the head of household, which may have limited value in capturing the gender dimensions of poverty. To have a better understanding of this poverty dimension, more detailed disaggregations are needed, including by sex, age, and 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.

The following example shows an analysis of risk at poverty of older women and men benefited from disaggregating poverty data by sex and type of household. 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%

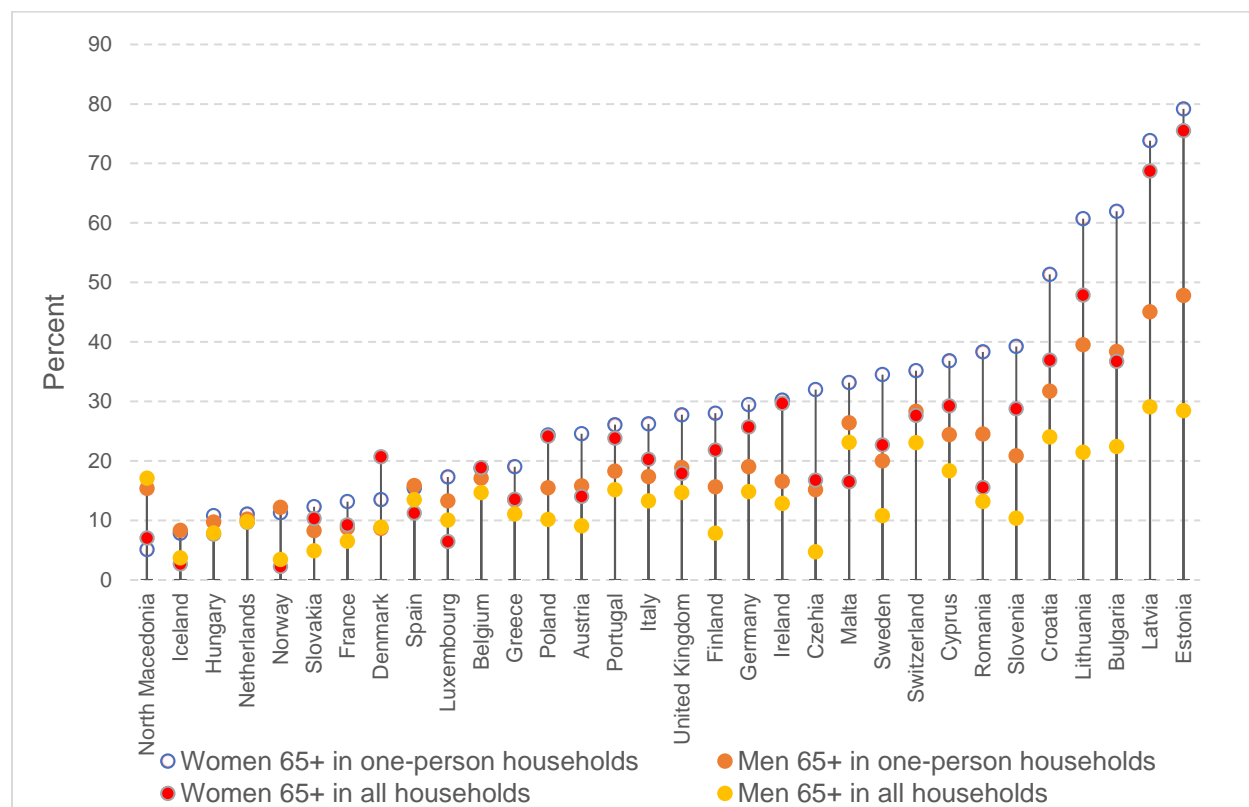
⁴² For additional examples on poverty risk groups see https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Europe_2020_indicators_-_poverty_and_social_exclusion#Which_groups_are_at_greater_risk_of_poverty_or_social_exclusion.3F

⁴³ Adapted from publications: UN 2015 and UN 2016a.

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, 2019.

Note: The data for Iceland is for 2016.

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, 2013), 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.

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. In addition, 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 to the unequal distribution of resources.

Currently, there is no single straightforward measure of poverty from a gender perspective, and no single internationally agreed-upon indicator that can give more meaningful poverty counts for women and men. 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 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.

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.⁴⁴ 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 gender (see also 3).

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

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, 2017: 24).

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 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

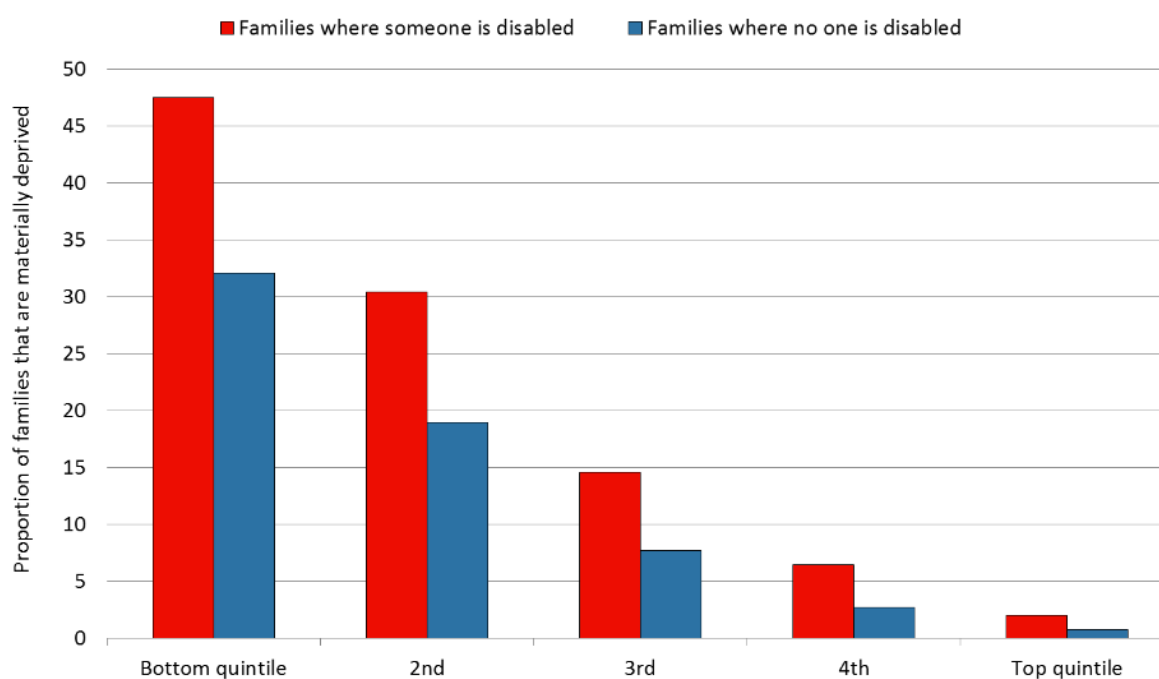
⁴⁴ Face unexpected expenses; one-week annual holiday; avoid arrears; meal with meat; chicken or fish every second day; home adequately warm; a car; replace worn-out furniture; replace worn-out clothes; two pairs of shoes; 'pocket money'; regular leisure activities; get together with friends/family for a drink/meal; have an internet connection.

income quintile, in which someone is disabled, are materially deprived compared with 32% for households where no one is disabled (see Figure 2.3).

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 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.

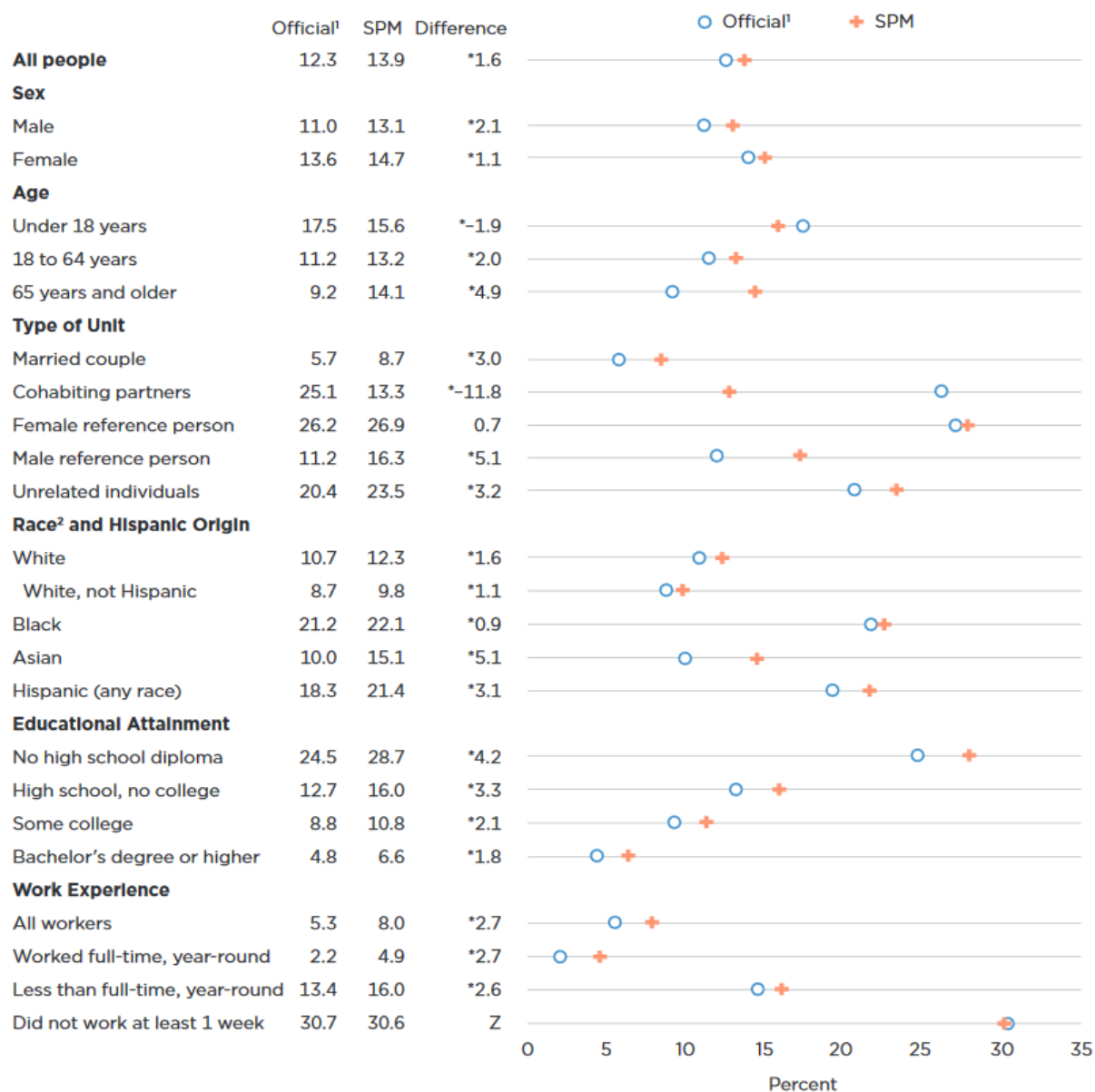
2.3.3 Example 3: Disaggregation and the use of alternative poverty measures (United States)

The impact of the use of an alternative poverty measure on different population groups is shown on 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.

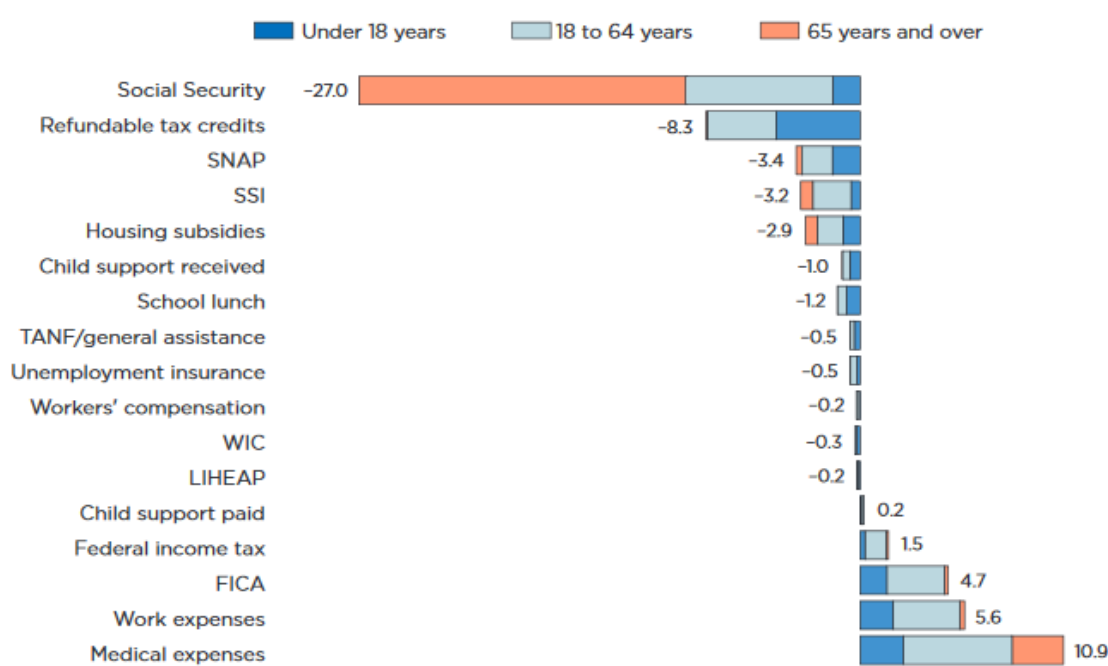
² Federal surveys give respondents the option of reporting more than one race. Therefore, two basic ways of defining a race group are possible. A group such as Asian may be defined as those who reported Asian and no other race (the race-alone or single-race concept) or as those who reported Asian regardless 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

Another advantage of SPM is that it enables showing the impact of not only cash benefits from the government but also the impact of non-cash transfers, taxes and other nondiscretionary expenses on poverty estimates. 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.

Note: For information on confidentiality protection, sampling error, nonsampling error and definitions, see www2.census.gov/programs-surveys/cps/techdocs/cpsmar18.pdf.

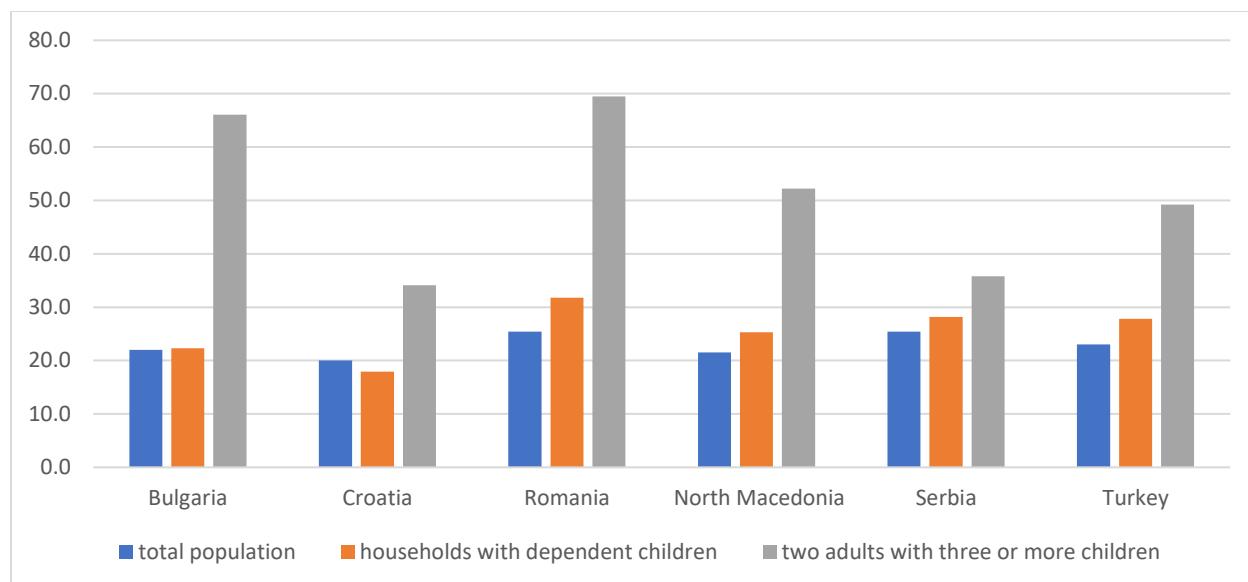
2.3.4 Example 4: Joint disaggregation of child poverty

2.3.4.1 Examples of relevant disaggregation of children poverty

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.

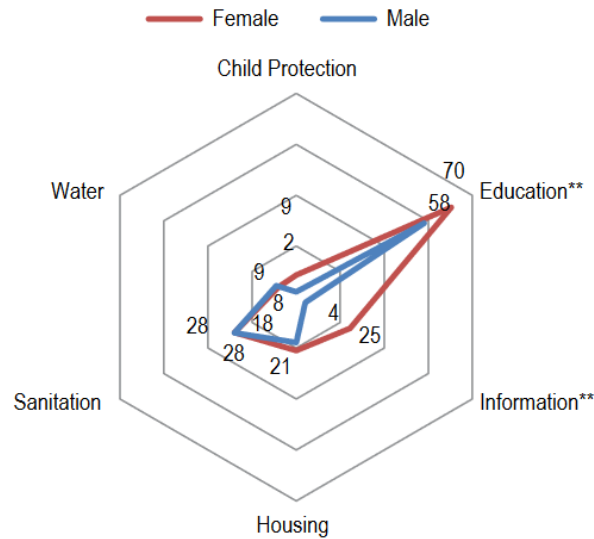
Note: The data for Bulgaria, Croatia, North Macedonia, Romania and Serbia is for 2015, and for Turkey for 2014.

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.

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.

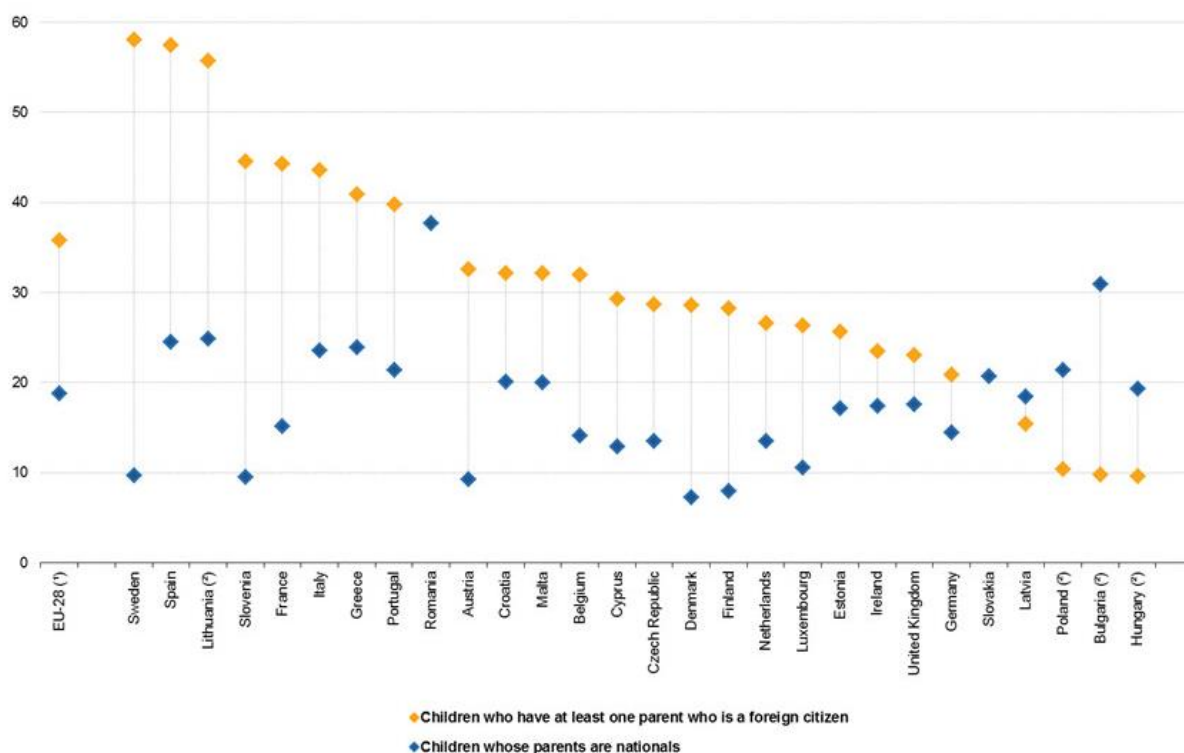
2.3.4.2 Children at risk of poverty according to citizenship of their parents

Eurostat provides an example of combining disaggregation of poverty measure for population groups defined by age and migratory status in international comparison. Figure 2.8 below shows differences in 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.⁴⁵

Figure 2.8

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

⁴⁵ <http://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/DDN-20180319-1>



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 with 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

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. Figure 2.9, Figure 2.10, Figure 2.11, Figure 2.12 and Figure 2.13 below 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

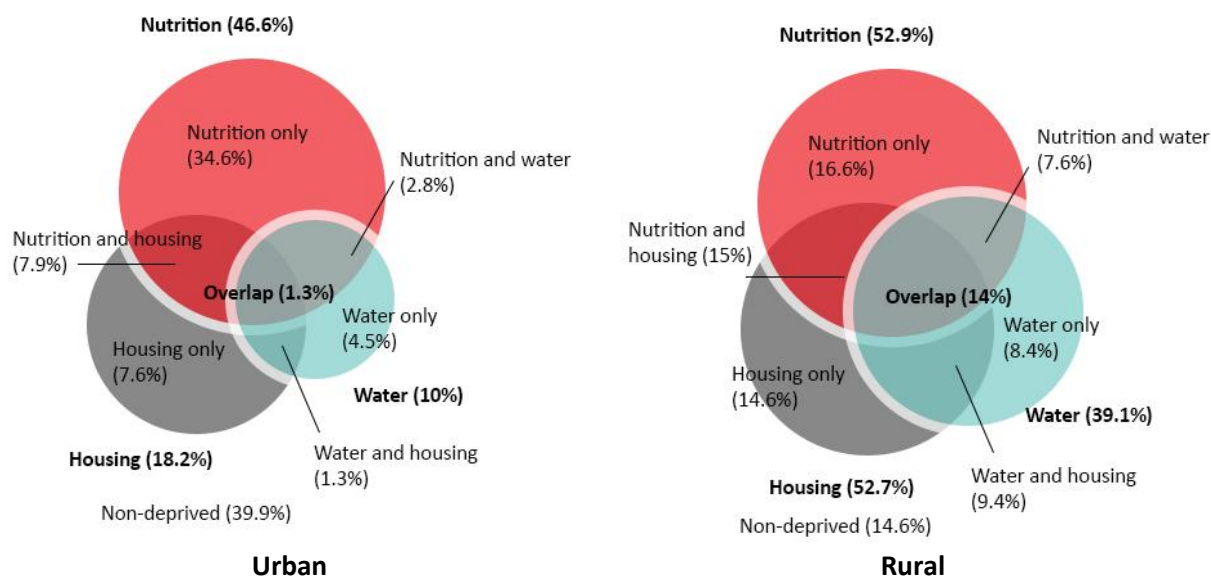


Figure 2.10

Venn Diagram of Deprivation Overlaps between Education, Water, Housing, by urban and rural areas, children aged 5-10 years

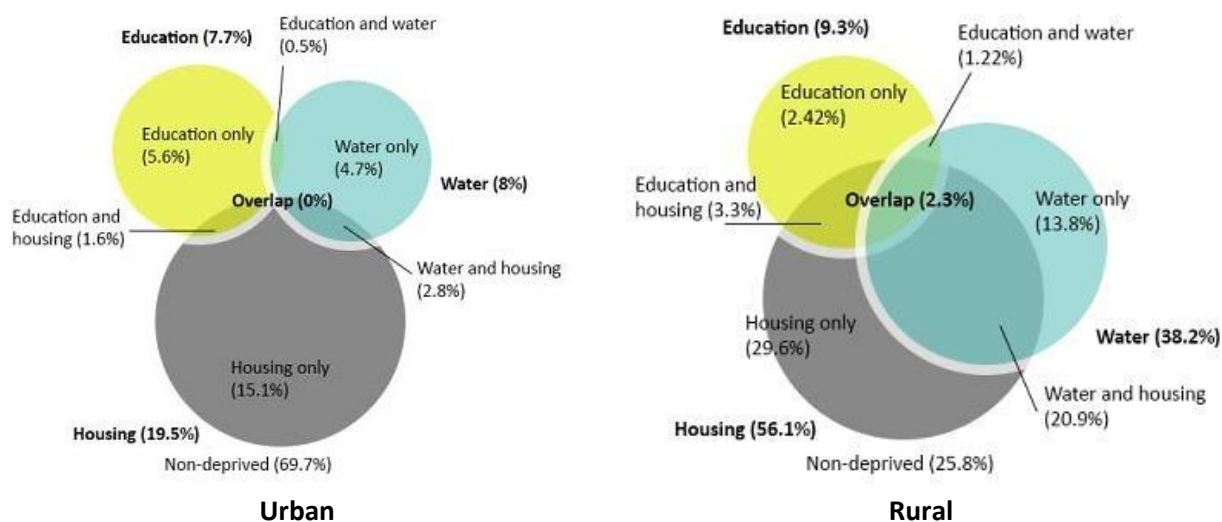


Figure 2.11

Venn Diagram of Deprivation Overlaps between Water, Sanitation, Housing, by urban and rural areas, children aged 5-10 years

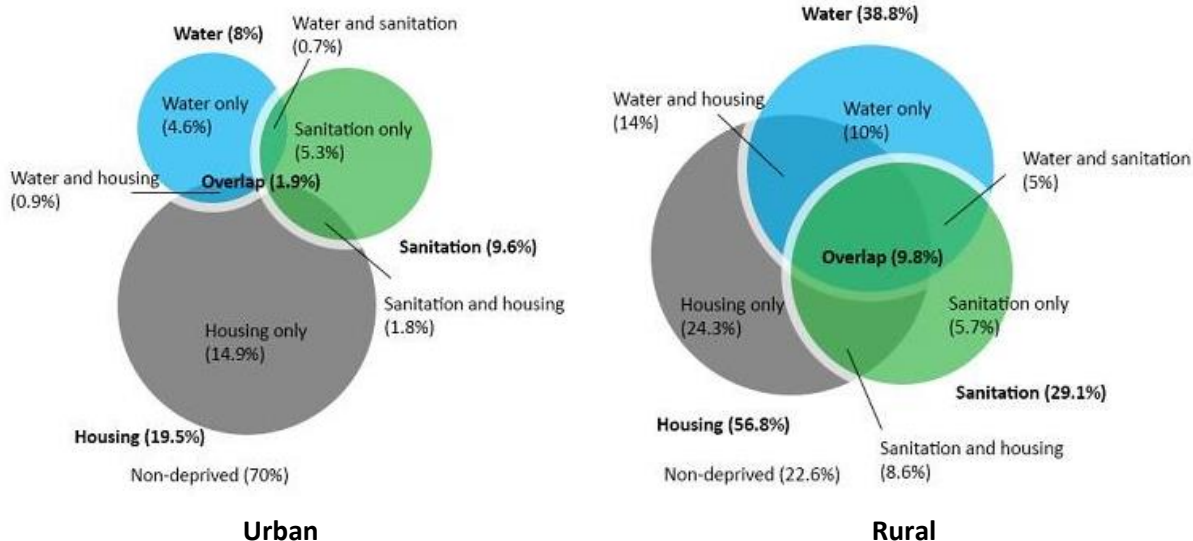


Figure 2.12

Venn Diagram of Deprivation Overlaps between Education, Water, Housing, by urban and rural areas, children aged 11-17 years

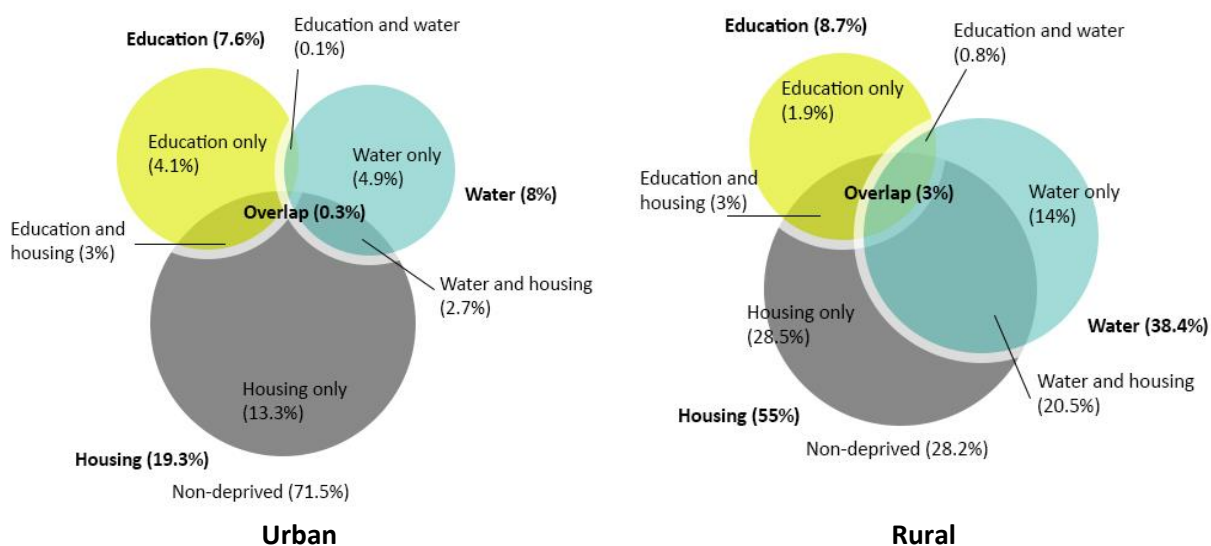
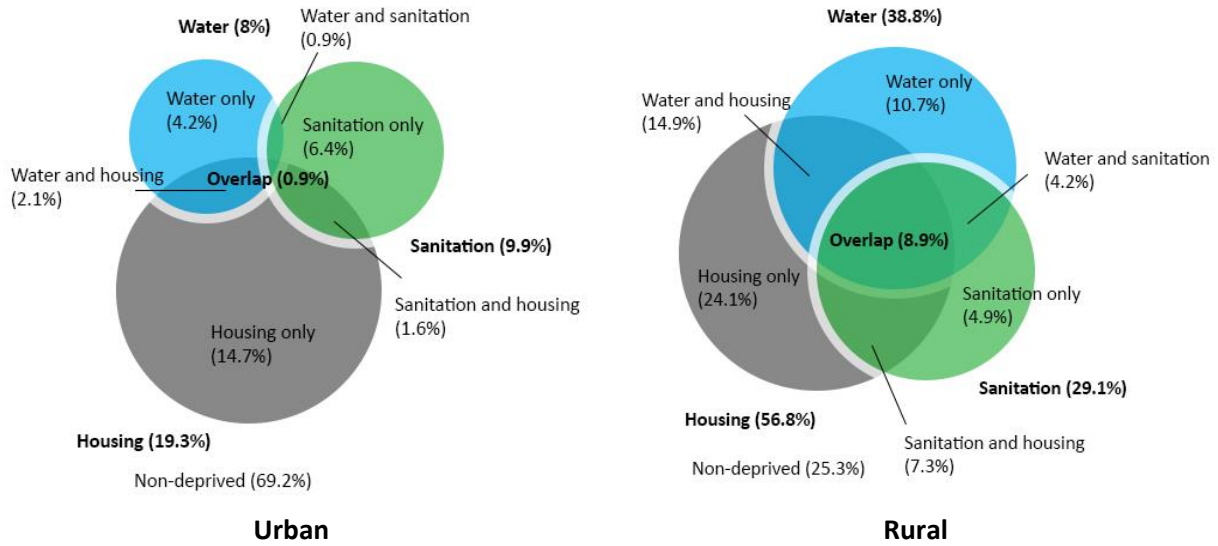


Figure 2.13

Venn Diagram of Deprivation Overlaps between Housing, Sanitation, Water, by urban and rural areas, children aged 11-17 years



Source: UNICEF, MODA, Tajikistan.

2.3.5 Example 5: Presentation of poverty across population groups

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.⁴⁶

Figure 2.14

Poverty among the basic sectors in the Philippines

⁴⁶ <http://psa.gov.ph/poverty-press-releases>

POVERTY AMONG THE BASIC SECTORS IN THE PHILIPPINES¹

Farmers, Fishermen, Children, Self-employed and Unpaid Family Workers, and Women, belonging to poor families, had higher poverty incidence¹ than the general population estimated at 21.6% in 2015.



PSA report provides the estimates of poverty incidence for 9 of the 14 basic sectors identified in Republic Act 8425 or the Social Reform and Poverty Alleviation Act based on the income and sectoral data from the merged Family Income and Expenditure Survey (FIES) and Labor Force Survey (LFS).

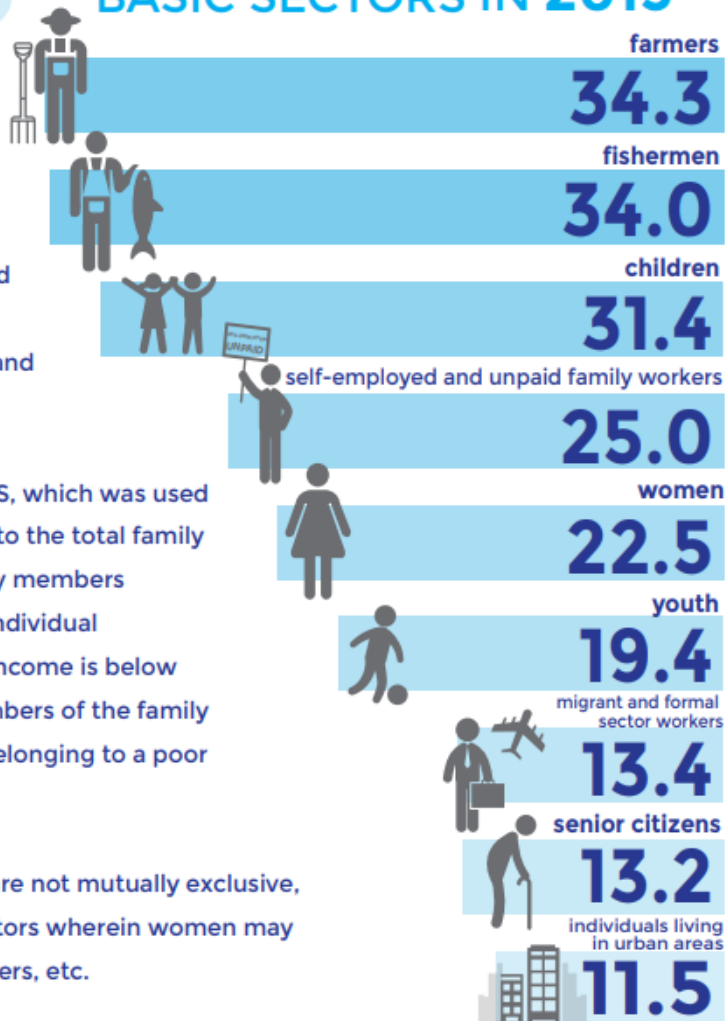
Per capita income derived from the FIES, which was used for classifying poor or non-poor, refers to the total family income divided by the number of family members and not necessarily the income of the individual in a particular sector. If the per capita income is below the poverty threshold, then all the members of the family are considered poor. Hence, a farmer belonging to a poor family is considered poor.

It may be noted that the basic sectors are not mutually exclusive, i.e., there could be overlaps among sectors wherein women may also be counted as senior citizens, farmers, etc.

POVERTY IN THE PHILIPPINES



POOR(%) AMONG THE BASIC SECTORS IN 2015



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 Labor Force Survey

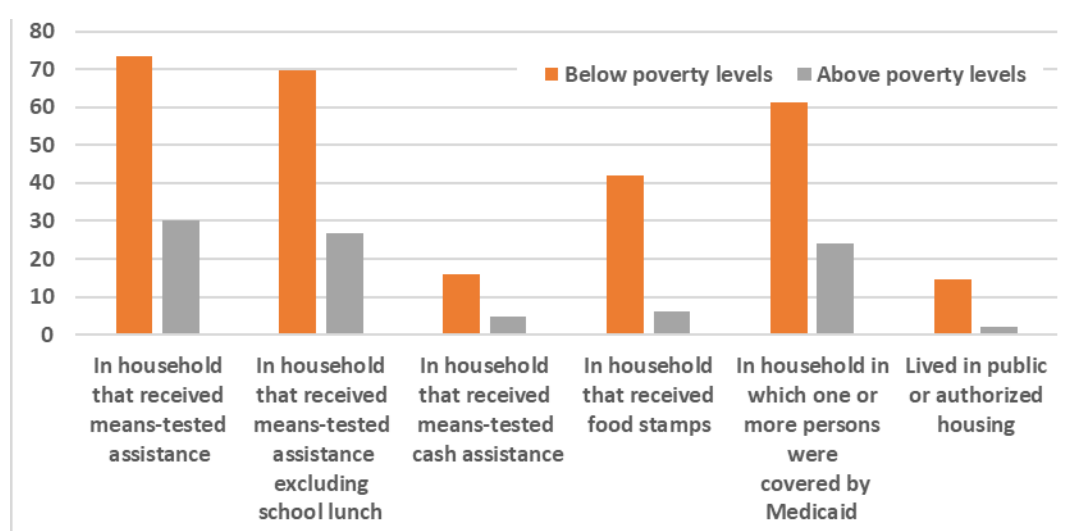
2/ Poverty Incidence is the proportion of families/individuals with per capita income less than the per capita poverty threshold to the total number of families/individuals.

2.3.6 Example 6: Poverty status and receipt of social transfers

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.⁴⁷

Figure 2.15

Program Participation Status of Household – Poverty Status of People in 2017 (% of population)



Source: U.S. Census Bureau, Current Population Survey, Annual Social and Economic Supplement. Available from: <https://www.census.gov/programs-surveys/cps.html>.

2.3.7 Example 7: Interactive presentation of poverty indicators

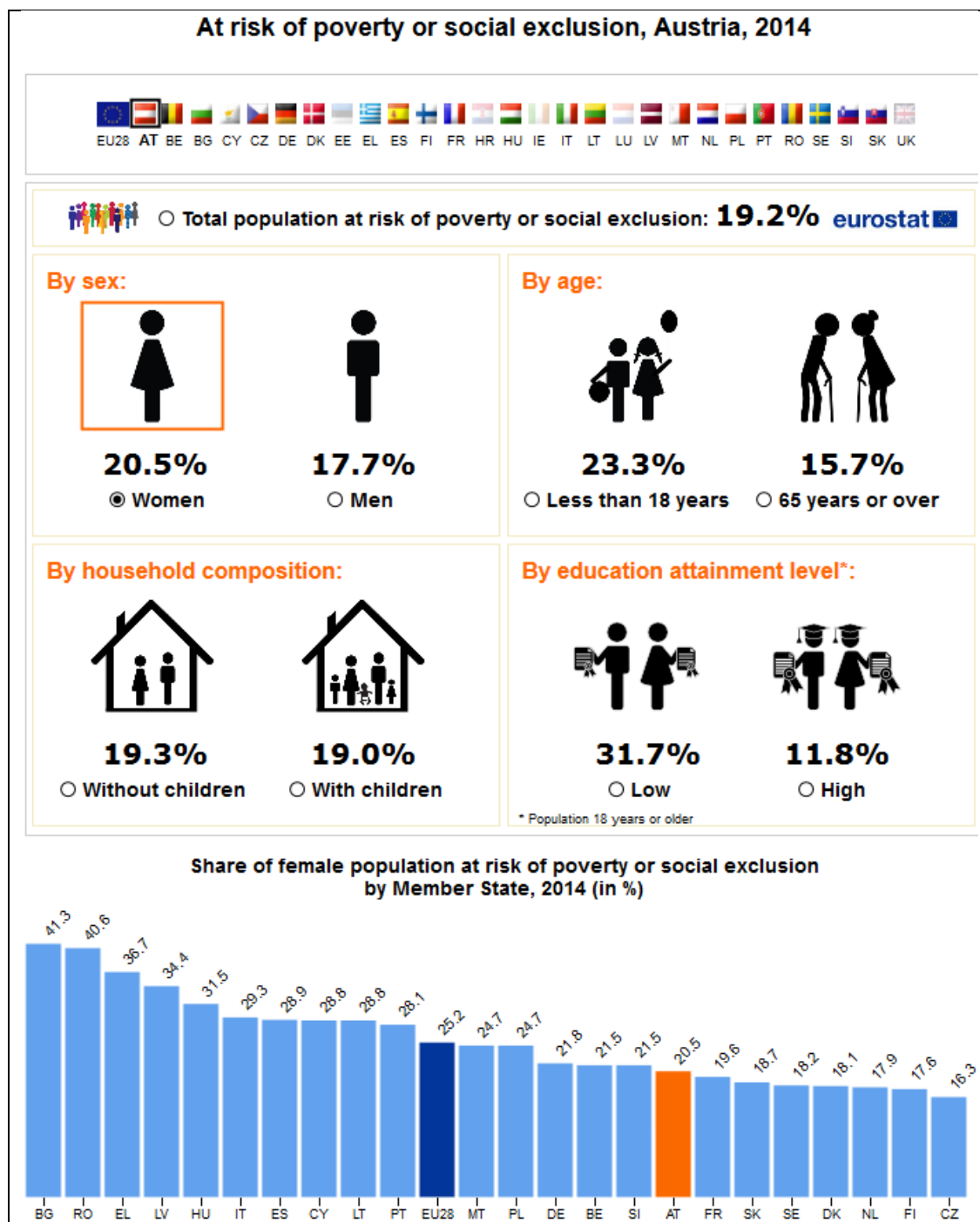
Interactive visualisation tools can facilitate the use of poverty statistics. An example of such tool is a Eurostat interactive tool put in place on the occasion of the International Day for the Eradication of Poverty and enabling visualisation of data on risk of poverty or social exclusion for the EU28 and its member states for the total population as well as for selected breakdowns (sex, age, household composition and education level).⁴⁸

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.

⁴⁷ See more at <https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pov/pov-26.html>.

⁴⁸ <https://ec.europa.eu/eurostat/news/themes-in-the-spotlight/poverty-day>

Figure 2.16
At risk of poverty or social exclusion, Austria, 2014



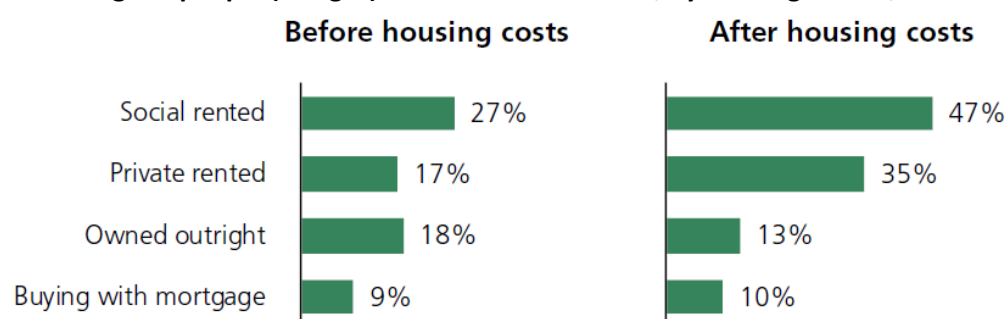
Source: Eurostat. People at Risk of Poverty or Social Exclusion by Age and Sex. Available from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=ilc_peps01&lang=en.

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

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. Figure 2.17 and Figure 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

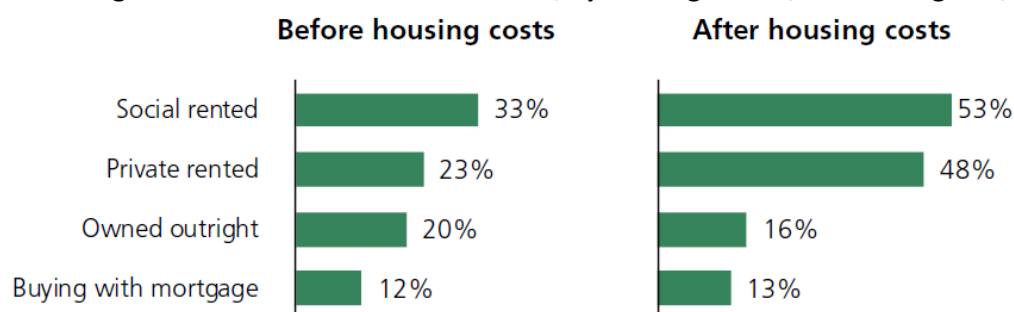
Percentage of people (all ages) in relative low income, by housing tenure, United Kingdom, 2016-17



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

1. 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.
2. In line with the Recommendation 4 in the Guide on Poverty Measurement and considering other emerging needs, especially in the context of SDGs, the following variables are suggested for

poverty disaggregation, wherever relevant and permitted by national legislation: age, sex, disability status, migratory status, ethnicity, household type, employment status, tenure status of the household, receipt of social transfers, educational attainment and degree of urbanisation.

3. The guidance on the definitions and operationalization of the disaggregations presented in this report is supposed 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.

3 Including Hard to Reach Groups in Poverty Measurement

3.1 Introduction

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, 2017) 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.

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.

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, 2017).

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 specific 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 populations⁴⁹

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

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).

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 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 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.

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

⁴⁹ 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, 2017).

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 identifying their Romani ethnic background, the respondents were given opportunity to identify if they feel belonging to another ethnicity in addition to the first identified ethnicity.

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.

The fifth principle is accountability. The UN OHCHR 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

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.

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 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.

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

Fieldwork should provide a safe and respectful environment to ensure the privacy and dignity of respondents. This may include to offer 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.

The *mode of the survey* should be tailored to the needs of the specific population. Sometimes non-probability designs such as opt-in online surveys are the only possibility to reach out to elusive populations (e.g. surveys on LGBTI).

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 Training and field work materials

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.

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.

3.3 Dedicated survey programmes for hard to reach groups

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.

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.

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

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.

Most importantly, the surveys need to include information which allows to identify the target groups. For example, for ethnic groups, the survey needs to ask respondents about their ethnic group, respecting the principle for self-identification. Likewise, information on disability and other characteristics which were presented in 2 needs to be collected.

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 country of birth or ethnicity due to differences in naturalisation policies, or for indigenous people.

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 3.1). 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 3.1

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. 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 Statistics on Income and Living Conditions (SILC) survey”. 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 (SO SR) and Office of the Plenipotentiary of the Government of the Slovak Republic for Roma communities. 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 relating to the MRC (hereinafter referred to as “EU SILC MRC”); the taking place in 2018, and with a second planned for in 2020. The SO SR was responsible for implementing EU SILC MRK and providing data to the Office of the Plenipotentiary of the Government of the Slovak Republic for Roma communities for further processing. The Office of the Plenipotentiary of the Government of the Slovak Republic for Roma communities assisted in developing the data collection methodology, and interviewer guidelines. They were also responsible for analyzing 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 aim group are household members and households of MRC.
- The objective of the survey is to measure income and living conditions of the MRC in the Slovak Republic.
- 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 survey, and then adapted for the MRC conditions.
- The SO SR 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 SO SR created a network of external interviewers for the purpose of collecting data in the field. The SO SR collected management information on interviews including their educational attainment, experience and age. Each interviewer had an identification number, which was recorded in questionnaire.
- The SO SR ensured record and contact information about households (e.g. address) was kept 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 SO SR 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.

3.3.2 Special survey programmes

In case the design of standard surveys cannot be adjusted to capture certain groups, the methods described in the next section based on work of the EU Agency for Fundamental Rights (FRA) can be considered. In 2015, FRA launched EU-MIDIS II which is a good example for 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).

3.3.3 Sampling target groups

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

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.

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.

Several methods to increase the efficiency of screening have been employed in EU-MIDIS:

- **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 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 neighbors (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 neighbors 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 neighboring 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 to locate 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, is it not acceptable to ask people to identify others based on their ethnicity or sexual orientation.⁵⁰ Additionally, if respondents do not know if their neighbors 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 at small geographical areas (e.g. blocks, neighborhoods). The method involves normal screening in PSUs, and then once a member of the target population is found, the interviewer also screens the neighboring household to the left and to the right. If a member of the target population is found in the neighboring household, the interviewer continues to the next neighbor 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 be implemented 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 neighboring households of core households. The weighting needs to adjust the selection probabilities based on the cluster size.

⁵⁰ Self-identification is one of the principles and recommendations for human rights-based data collection. See OHCHR, 2018.

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

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.⁵¹

- **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 centers 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 centers, 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 impacting on the design effect. It is possible that at some locations, interviewers are not

⁵¹ For a more comprehensive overview of methods, see for instance Marpsat and Razafindratsima, 2010; Kalton, 2009 or Verma, 2014.

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).

- **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).

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 from 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 3.2

Case Study: 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 3.1 below, which outlines various sampling approaches and identification techniques in previous Roma surveys.

Table 3.1

Summary of sampling approaches in previous Roma surveys

| Survey | Roma identification | First stage sampling | Second stage sampling |
|--|---|--|---|
| UNDP Regional Roma Survey 2011⁵² | 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 |
| EU 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 ⁵³ . | 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 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, 2017) | 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 |

⁵² <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>

| | | | |
|--|--|--|--|
| | | | make up less than 25 % of the population (in Bulgaria, Hungary and Romania.) |
| MICS Roma settlement Surveys (2010-ongoing in several West Balkan Countries)⁵⁴ | 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. |
| 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. |
| 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. |

Box 3.3

Case Study – UNICEF: Measuring the Roma population

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

⁵⁴ <http://mics.unicef.org/surveys>

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

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

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, 2017).

EU-MIDIS II also used non-response weights based on 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.

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.

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.

3.3.3.5 Proxy indicators

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 standard for comparative research.

For example, FRA uses current monthly household income in EU-MIDIS II to capture current living standard. 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 standard 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.

3.3.3.6 Micronarratives

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.

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.⁵⁵ 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 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.

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.

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.

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 3.4

Case Study – UNDP: Micronarratives 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. 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 micronarrative was based on this open-ended question to prompt the story:

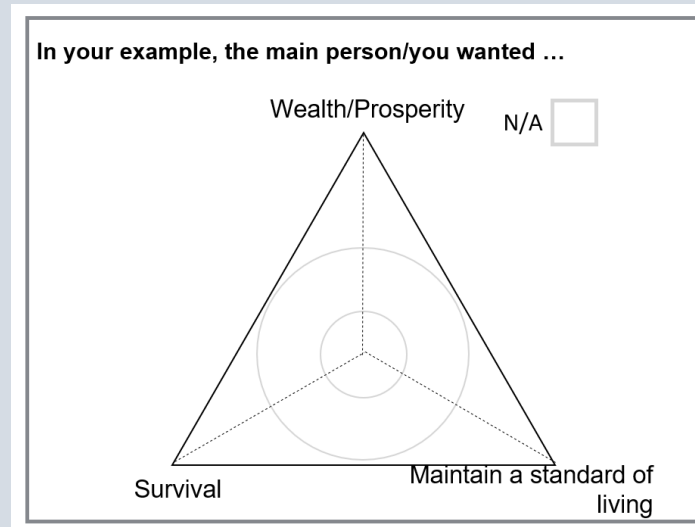
"Give us an example (good or bad) of what life is like for you (or someone you know) after coming back from another country. What happened? Tell us about what happened. Say as much or as little as you wish."

⁵⁵ 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>

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

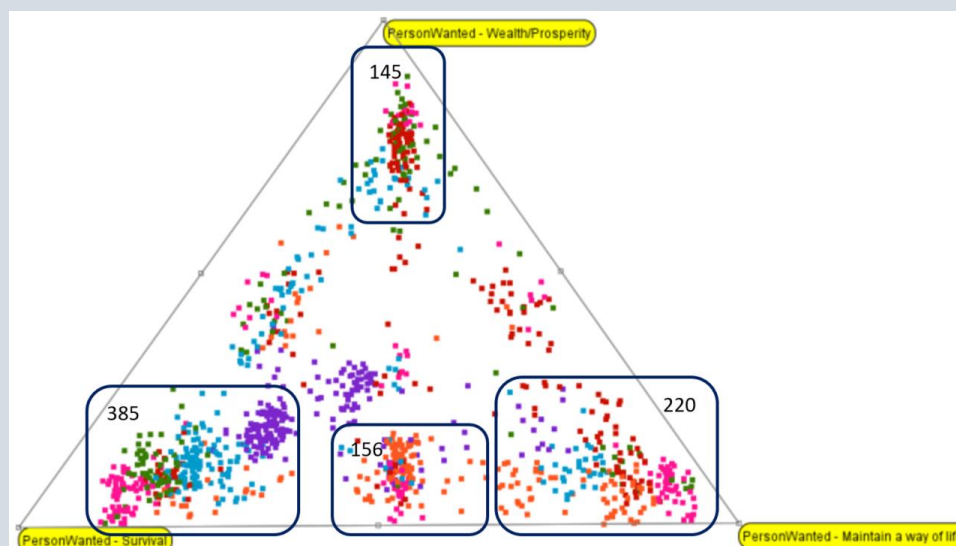
Micronarrative framework



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

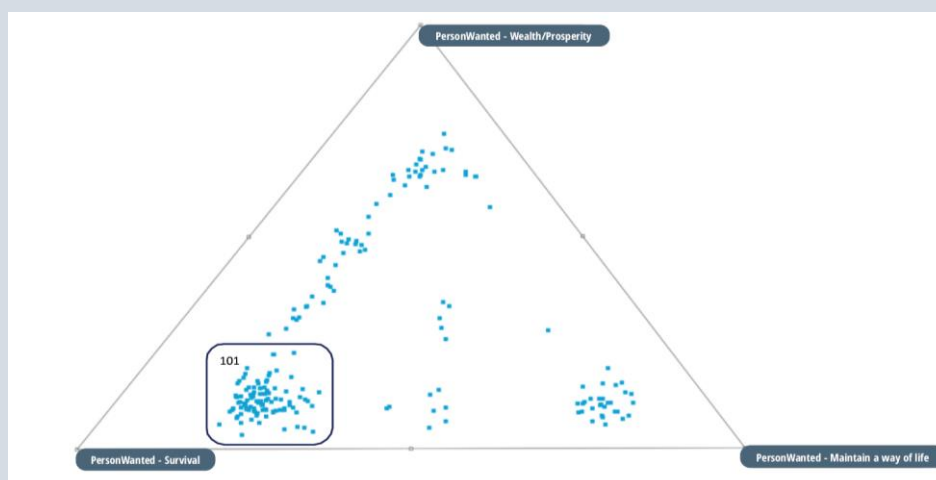
Self-coded reasons to go abroad in Western Balkans (n=1136 micro-narratives)



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.

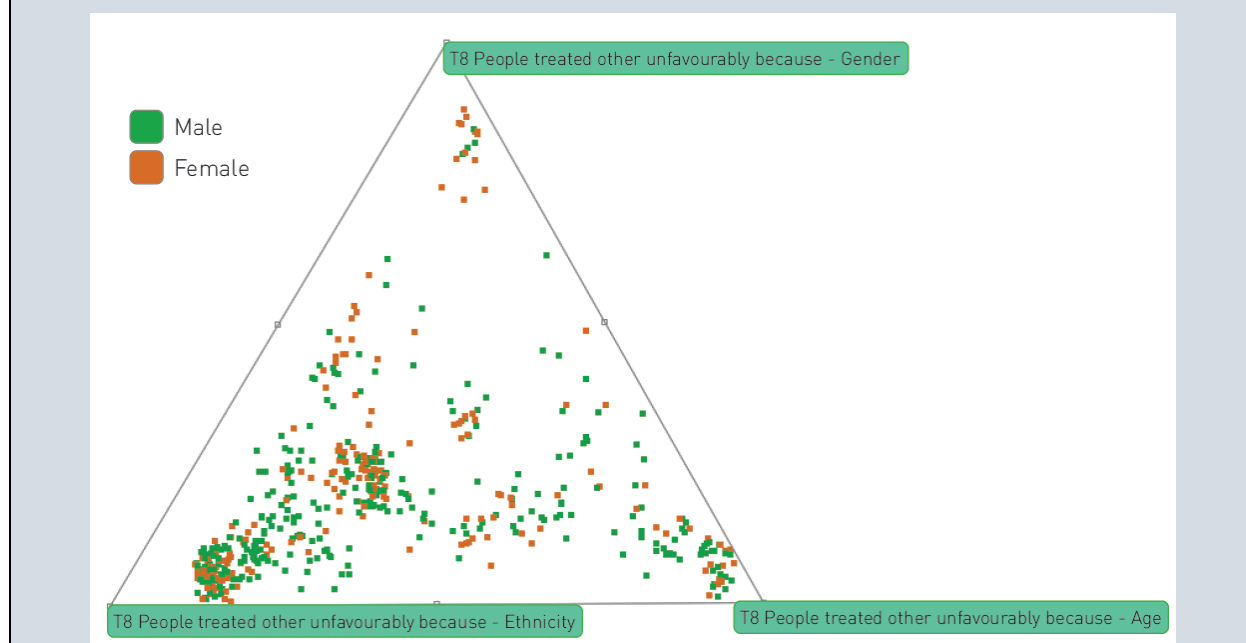
Figure 3.3

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 1000 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 1000) 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 unfavorable treatment based on gender, ethnicity and age, n=510



3.4 Poverty among children

Data on child poverty is essential in fulfilling the United Nations Convention on the Rights of the Child. 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. 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 to include survey question 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, 2018).

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.

3.4.1 Street children

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

3.4.2 Children with disabilities⁵⁶

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.

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.

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 to understand the differences and extent of multidimensional poverty between children with functioning difficulties and the rest.

3.4.3 Foster children

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 **Error! Reference source not found.**) potential within household inequalities will however remain hidden.

The United States Census Bureau used the Supplemental Poverty Measure (SPM) which includes children in foster case. Estimates show that generally foster children have a lower risk of poverty (Pac et al, 2017).

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 absence of children who live in institutions. It

⁵⁶ For more information see: UNICEF, 2018.

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).

3.5 Homelessness and housing difficulties in the European Union

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.

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 services⁵⁷.

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.

3.5.1 Policy justification

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.

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;

⁵⁷ <http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/ad-hoc-modules>

- 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.

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 harmonization in the definition and methodology, the data available is not comparable across countries.

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.

3.5.2 How to define homelessness

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.

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.

The major challenge for EU SILC was to agree 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 "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.

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 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.

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.

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.

3.5.3 Sample Issues

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.

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.

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.

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.

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.

3.5.4 Implementation in EU-SILC

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.

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.

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.

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.

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.

The 5 agreed variables on past housing difficulties are reported in Table 3.2⁵⁸. 13 Member States⁵⁹ declared that they will implement the optional housing difficulties variables in the 2018 EU-SILC ad hoc module.

Table 3.2

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 |
| 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 |

⁵⁸ 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)

⁵⁹ Belgium, Bulgaria, Germany, Ireland, Greece, Spain, Hungary, Malta, Portugal, Romania, Slovakia, the United Kingdom and Denmark (Denmark partially, only the first 3 variables)

3.5.5 Plans for the revised EU-SILC

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.

3.5.6 Conclusion

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.

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.

3.6 Surveying Institutional Households

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.

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

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.

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.

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 3.5

Case Study: USA - Measuring Poverty for the Institutionalized Population

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. In contrast, for the noninstitutionalized population, poverty status is based on family income and the threshold appropriate for each family configuration. 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.

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, old people's homes, residential homes, prisons, military barracks, shelters for refugees or homeless persons, boarding houses and workers' hostels. The proportion of the individuals living in these types of accommodation varies between countries significantly and is estimated to be increasing in many countries due to aging populations.⁶⁰ 20 out of 30 European countries reported a share of collective households higher than 1% of the entire population. Only in seven countries (Bulgaria, 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 the persons 65 years and older were living in institutions in 2010, which is expected to increase to 28% by 2035 (Rodrigues et al, 2012).

The SERISS network⁶¹ (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).⁶²

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 tend to be covered in headline statistics. The UN Convention on the Rights of Persons with Disabilities (CRPD) Article 31 obliges State Parties to collect appropriate information, including statistical and research data, to enable them to formulate and implement policies to give effect to the present Convention.

Box 3.6

Case Study: European Union Agency for Fundamental Rights

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 of this project 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 process of deinstitutionalisation and to identify the drivers and barriers for that persons with disabilities have the opportunity to choose their place of residence and where and with whom they live on an equal basis with others and are not obliged to live in a particular living arrangement. The study asks for an improvement of availability, comparability and methodological adequacy of data on persons with disability, especially with regard to persons living in institutions.

⁶⁰ For definition see the OECD glossary: <https://stats.oecd.org/glossary/detail.asp?!ID=1372>

⁶¹ <https://seriss.eu/about-seriss/project-overview/>

⁶² https://seriss.eu/wp-content/uploads/2017/02/SERISS_WP2_D2.16_Survey-inventory-institutionalized-population.xlsx

3.6.1 Case Study: UK - including non-private household populations (NPHP)

Following a number of developments which raised concerns about destitution and severe poverty in the UK, 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-being⁶³. 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.

Currently, UK estimates of living standards and quality of life derive mainly from household surveys. Unlike many other countries, the UK 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.

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.

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.

The address-based sampling used in UK 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.

3.6.1.1 Defining the populations of interest

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.

⁶³ 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 full report of this scoping work is available under <https://researchportal.hw.ac.uk/en/publications/scoping-project-to-investigate-the-alternatives-for-including-non->

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 3.3

Estimated size of each NPHP sub-population in the UK

| 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 |

Table 3.3 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.

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

Although, collectively referred to as the NPHP, it is important to recognise 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.

3.6.1.2 Potential ways forward

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 UK-wide coverage.

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.

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 3.4.

Table 3.4

Proposed approaches to capture data for the different categories of NPHP

| 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 |

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.

3.6.1.3 Methodological considerations

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

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 year.

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.

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.

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.

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.

3.6.1.4 Next steps

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.

3.7 Recommendations

1. Countries are encouraged to first identify, and then quantify, parts of the population not sufficiently covered in their poverty statistics, including institutional households. They should further explore the causes and challenges why these populations are not covered or underrepresented.
2. Countries should develop suitable methods to facilitate equal 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.
3. 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.
4. 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.
5. Interviewer training for surveying minority groups should address cultural and groups specific aspects. Persons of the target populations should be involved in the set-up, and development of fieldwork materials.
6. Further research should be directed at creating supplementary measures of poverty to reflect that the living circumstances of target groups don't always neatly fit with standard household definitions.
7. Further research is needed to explore the possibilities for alternative sampling approaches, such as non-probability designs including online surveys, to measure hard-to-reach populations.

4 Assessing and Improving Survey Methods

4.1 Survey errors and quality

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 met, or statistical concepts could not be correctly applied will not only assist 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 to investigate 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

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.

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 for 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).

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

4.1.1.1 Errors in measurement

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 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 in so far as the components differ in nature and in methods of assessment and control.

4.1.1.2 Measurement bias

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.

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 equal 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 **Error! Reference source not found.**).

4.1.1.3 Measurement variance

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.

4.1.2 Errors in estimation

4.1.2.1 Coverage and related errors

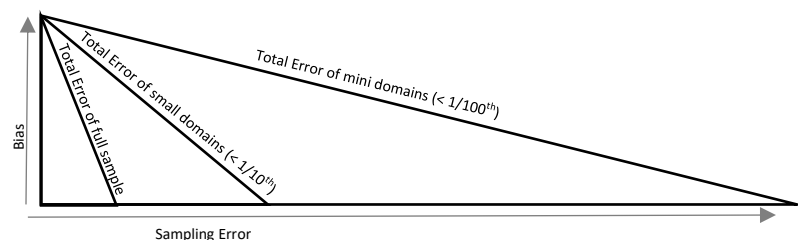
Coverage errors arise from discrepancies between the target and the frame populations, and also from errors in the way the sample is selected from the frame. Valid inference is only ensured by probability samples which meet all of these criteria: (a) the survey population is fully and correctly represented 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).

4.1.2.2 Sampling error

Sampling error is a measure of the 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. For estimates based on small samples, this component 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.

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 much increased by increases of 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



4.1.3 Non-response errors

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.

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.

Proposed solutions for item non-response will be presented in 2.3 on imputing missing values, while coverage and related errors and unit non-response will be treated in 4.2.4 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.

4.1.4 Multidimensional quality frameworks

Quality reports are essential tools to assess, improve and communicate the quality of poverty measurement. As 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).

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

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 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)

4.1.5.1 How relevant is the data?

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.

4.1.5.2 How timely and punctual are results available?

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*."

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 (Fellegi, 2001) identifies credibility as a "survival" issue for a statistical organisation.

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.

4.1.5.3 How precise are the results?

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.

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).”

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. As an absolute minimum this should include an indication of the design effect, which is calculated as the variance of the main poverty estimate divided by the variance of the same indicator if simple random sampling had been used. For disaggregation of poverty measures, it is especially important to identify clearly those results which provide only limited accuracy. Indicators with unacceptable inaccuracy should never be published. More information on best practices for dissemination are discussed in 4.2.6.

4.1.5.4 To what degree are findings comparable?

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, 2017). This 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”.

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

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.”

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 incoherences 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.

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.

4.1.5.6 Are data accessible and clear?

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.

4.1.5.7 What else would users need to know?

A variety of other aspects are also covered in the 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.

4.1.6 Relationship between different aspects of quality

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.

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

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

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.

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 3.

4.2.1.1 Choice of data source

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.

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.

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

Country Case: Use of Administrative Tax Data at Statistics Canada

At Statistics Canada, administrative tax data has been the primary source of income data for about 20 years. Using this source has numerous advantages. The Statistics Canada's surveys used to measure 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 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 included on tax form such as full amounts of spousal support (alimony) and other transfers between households. Collection is also required to obtain variables of interest for disaggregation of poverty statistics such as family composition.

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 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.

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 transfer 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

Country Case: 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: FR, IT, LV, CH, IE, AT and ES. Overall, it was found as 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 | BG BE DK EE ES FI IT LT LV NL AT SE SI IS NO | 15 |
| Education data | DK FI SI IS NO | 5 |
| Labour data | BG DK NL SI IS NO | 6 |
| Housing/dwelling data | DK AT UK | 3 |
| Income data | BG BE CY DK ES FI FR IE IT LT LV MT NL AT SE SI IS NO CH | 19 |

| | | |
|-------------------------------|--|---|
| Other | MT (electricity and water consumption) | 1 |
| Not using administrative data | CZ DE EL HU LU PL PT SK | 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 than 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 required to link survey information with registers is missing for up to 5% of the sample persons. The information was more often missing for certain groups: Persons under age 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 inevitable to put specific questions on alimonies.

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

Country Case: 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 the sampling unit, but household membership is always verified by face to face interviewers. Individuals who are not living at the selected address are added and registered individuals not living at this address are not further considered. This sometimes involves cumbersome enquiries to obtain linkable person identification numbers (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

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 which is responsible for the allocation of regional funds had specified regional precision requirements such that EU-SILC. These requirements are partly reflected in Annex II of a newly established framework regulation which applies also to EU-SILC (EU, 2019/1700).

Where disaggregation variables are available on the survey frame, the sample design may use them as stratification variables. This can help improving 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 is often best to use different sampling rates by regions. For example, within the European Union, differences in sample size are much smaller than the differences in population size countries. This helps make comparisons between Member States.

It ought to be understood that different sampling rates 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 many factors 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 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).

While the theory is well-established, in practice there will not be one ideal 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 may not be for another. 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 (Bethel, 1989) discusses allocation in the context of multivariate surveys. Additionally, when the goal is to provide disaggregate statistics, 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 difference in the allocation since the final number of respondents in each stratum will be different from the number selected due to non-response.

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 micro data 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 has been substantially complicated because this

sample design information is not accessible for all countries for Eurostat (Verma et al, 2010; Goedeme 2013; Trinidad and Goedeme, 2016).⁶⁴

Box 4.6

Country Case: Design of the American Community Survey in the United States

The American Community Survey (ACS) (United States Census Bureau, 2014) is a relatively new survey conducted by the U.S. Census Bureau. 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. Initially, five years of samples were required to produce these small-area data. Once the Census Bureau released its first 5-year estimates in December 2010, new small-area statistics now are produced annually. 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 which has an annual Social and Economic Supplement which is commonly used for poverty statistics.

4.2.1.3 Repeated surveys

Surveys on poverty and income 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.

For the measurement of income and 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).

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 immigration. This is a particular challenge for disaggregation. Without adequate sample refreshments the population of new migrants – which are often vulnerable to poverty – can not be represented which can lead to serious bias (Glaser et al, 2015). Non-response due to respondent fatigue is a particular challenge

⁶⁴ See also the important resources accessible here <https://timgoedeme.com/eu-silc-standard-errors/>

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.

There exist intermediate solutions between a cross-sectional and longitudinal survey and allow for objectives of both to be balanced. In rotating panels designs part of the sample is replaced at each iteration of the survey. This allows for 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 rotation panel design. The EU-SILC uses this type of design.

4.2.1.4 Mandatory and voluntary surveys

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 only few countries participation is mandatory. From the perspective of measurement errors, it is not always advisable to make survey participation mandatory (see Glaser et al, 2015).

Box 4.7

Country Case: Canadian Income Survey Design

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 cross-sectional 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 LFS is composed of six independent samples as rotation groups. A sixth of the sample, corresponding to one rotation group, is replaced every month. 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 is a mandatory survey, CIS is not.

All 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, 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

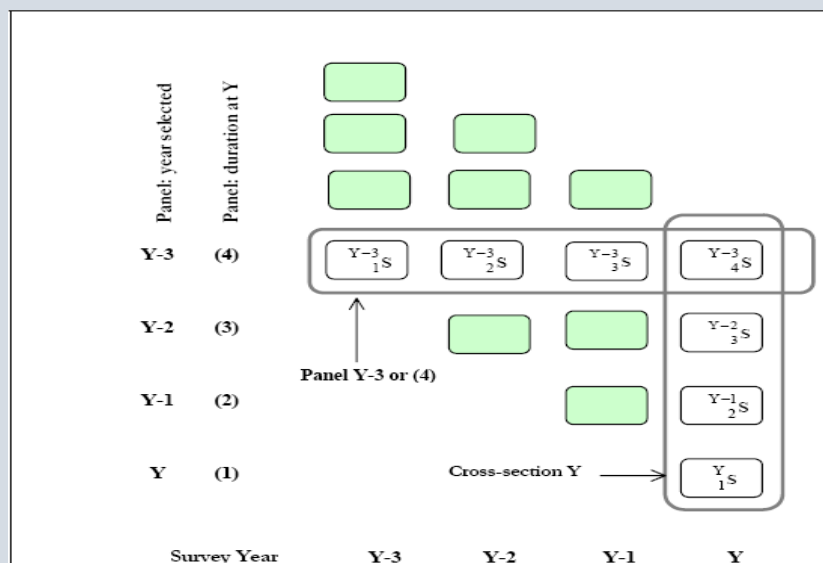
Country Case: EU-SILC Sample Design

The EU-SILC uses a rotating panel design that allows for the production of both cross-sectional and longitudinal statistics. In most EU-SILC, surveys a period of four years is taken as the duration for longitudinal follow-up at the micro level. A standard integrated design has been adopted by nearly all the participating countries.

This integrated design involves a rotational panel in which a new sample of households and persons is introduced each year to replace a part (normally one quarter) of the existing sample (see Figure 4.2 below). Persons enumerated in each new sample are followed-up in the survey normally for four years. A common rotational sample of this type yields each year a cross-sectional sample as well as longitudinal samples of various durations (Verma, 2001; Verma and Betti, 2006).

At any one time, the sample is made up of 4 subsamples or panels. Each year one new panel is added to stay in the survey for 4 years, and then dropped to be replaced by another new panel. Movers from the original sample are followed-up to their new location for up to the time their panel remains in the survey. This scheme provides both cross-sectional and longitudinal data from the same common set of units. 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.

4.2.2 Data collection

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.

Lower response rates 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 can be very useful for understanding longer-term trends in poverty but come with the extra challenge of minimizing attrition due to respondent fatigue.

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.

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 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.

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

Country Case: Collection Methodology of 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 5 weeks after the first mail, the Census Bureau follows up with computer-assisted telephone interviewing (CATI) when a telephone number is available. If the Census Bureau is unable to reach an occupant using CATI, or if the household refuses to participate, 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 three-phase process operates in continuously overlapping cycles so that, during any given month, three samples are in the mail/Internet phase, one is in the CATI phase, and one is in the CAPI phase.

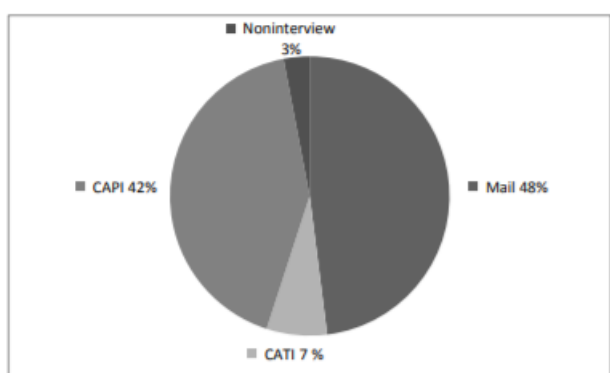
Table 4.3

ACS Data Collection Consists of Three Overlapping Phases

| ACS sample panel | Month of data collection | | | | | |
|------------------|--------------------------|----------------|----------------|----------------|----------------|----------------|
| | 2013 | | | | | |
| | January | February | March | April | May | June |
| November 2012 | Personal Visit | | | | | |
| December 2012 | Phone | Personal Visit | | | | |
| January 2013 | Mail/Internet | Phone | Personal Visit | | | |
| February 2013 | | Mail/Internet | Phone | Personal Visit | | |
| March 2013 | | | Mail/Internet | Phone | Personal Visit | |
| April 2013 | | | | Mail/Internet | Phone | Personal Visit |
| May 2013 | | | | | Mail/Internet | Phone |
| June 2013 | | | | | | Mail/Internet |

Table 4.3 summarizes the distribution of interviews and non-interviews for the 2012 ACS. Among the ACS sample addresses eligible for interviewing in the United States, approximately 48% were interviewed by mail, 7% by CATI, and 42% were represented by CAPI interviews. 3% were non-interviews.

Figure 4.3
Distribution of ACS Interviews and Non-interviews



Source: 2012 ACS Sample

The Current Population Survey (CPS) uses CAPI/CATI interviewing with interviewers trained to probe respondents. It takes several months to get the ACS responses from a given month as nonresponse is followed up and there can be delays with the mail back. Increasingly internet responses are received which might alleviate some of these operational issues in the future. The ACS has a 3-month data collection window and not the one month or one-week collection period that CPS uses. This poses its own estimation challenges because each month's worth of interviews is a somewhat random collection of interviews from three different sample panels depending on the self-response rates and whether the housing units are interviewed in months 1, 2, or 3.

To reduce the 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 should be calculated. A tool that can be used to evaluate non-response is the R-indicator.⁶⁵ 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 groups 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).

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

⁶⁵ <https://www.cmi.manchester.ac.uk/research/projects/representative-indicators-for-survey-quality/>

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.

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 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 also, 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)

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.

When no statistical registers are available, software such as Italian RELAIS (Record Linkage at ISTAT) and Statistics Canada’s G-Link (Statistics Canada, 2017) 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.

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 state 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.

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).

4.2.3 Data processing and imputation

Raw data that is received from collection is invariantly 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 for many reasons, for instance, because of refusals to answer certain questions, because the respondent does 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 first steps 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 a 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.

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.

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. The survey practitioner should not aim to produce dataset that is free of relationships that seem inconsistent at first view. Not only would he find the task near impossible, in trying to do so he would be likely to impose preconceived relationship on the data that may turn out to be incorrect.

Having cleaned the data, the next step is imputation, that is, assigning values to replace missing data. There are many imputation methods available. Two examples of imputation for income variables are given

below, the first is uses a regression-based approach and the second 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. For example, on income surveys, the relationship between labour forces status and the types of income received should be respected. The process should be automated and objective, avoiding the application of preconceived models on the records to be imputed as manual intervention can. Additionally, imputation rates should be reported as part of the survey's quality report as an important complement to the response rates.

The most flexible imputation procedure for imputing missing data (unit non-response) in income variables in EU-SILC is based on the "sequential regression multivariate imputation" (SRMI). This approach was implemented by the University of Michigan in the imputation software (IVE-ware) which can be used with SAS, STATA, SPSS and R packages or as a standalone in Windows, Linux or Mac OS (except SAS) 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 Missing at Random (MAR) mechanism, infinite sample size and simple random sampling.

The procedure is a variant of the estimation-maximisation (EM) algorithm and follows a Bayesian paradigm. The sequential multivariate model used 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 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 variable 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.

The model is as follows. With U as the matrix containing variables with no missing data (including as a result previous imputation), and $Y_1, Y_2 \dots Y_k$ are variables with increasing rates of missing data, the sequence of imputations is determined by the following factorisation:

$$[Y_1 | U] [Y_2 | U, Y_1] \dots [Y_k | U, Y_1, \dots, Y_{k-1}]$$

where $[Y | X]$ is the conditional joint distribution of Y where x is known.

The form of regression depends on the nature of Y , such as a generalised linear regression for continuous variables (as in the case of income amounts), a logistical regression for binary variables, etc.

- 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.

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.

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.

At each phase, the income variables are imputed by nearest neighbour imputation using the Canadian Census Edit and Imputation System (CANCEIS), which is available freely to users outside of Statistics Canada. Nearest neighbour imputation 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 or may be selected randomly from among a set of potential donor each of which is close to the receiver.

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.

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.

4.2.4 Weighting

The calculation of weights is usual 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.

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

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.

For example, when a sample of households (or of addresses or other units containing households) is selected, the household design weight is computed as

$$\omega^{(HD)} = \frac{1}{\text{probability of selection of the household}}$$

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

The next step of the weighting process in 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.

The problem of (unit) non-response can be particular problematic in some household survey 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 wave 1 being excluded from subsequent waves and so on. Good, efficient procedures to re-weight the responding cases is therefore a critical requirement at wave 1. 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.

The non-response adjustment procedure involves estimating response rates or propensities to response 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.

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.

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.

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.

In surveys where substitution has been allowed, 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.

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 tree, 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:

$$\omega^{(HN)} = \frac{\omega^{(HD)}}{R_K}$$

where R_K is the response rate in weighting class k.

In this expression, the response rates should be computed with data weighted by the design weights:

$$R_K = \frac{\text{sum of design weights of responding units in cell } k}{\text{sum of design weights of selected units in cell } k}.$$

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 overall response rate. No absolute rule exists but it may be to useful 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.

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. Each responding unit weight is adjusted by the inverse of the estimated response propensity, in the same way as by cell response rates in the previous method:

$$\omega_i^{(HN)} = \frac{\omega_i^{(HD)}}{R_K}.$$

When many auxiliary variables are available, this approach is often preferable to simply using sampling strata or a geographic partition.

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 is a best practice 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 having similar response propensity. Once these classes are defined, the adjustment can then proceed as in the first method, within the cells. This is known as the score method (Little, 1986; Eltinge and Yanseneh, 1997).

Regardless of 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 increase the variance of the estimates.

4.2.4.3 Calibration

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.

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. 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.

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.

Mathematically, calibration is an optimization problem. The goal is to find weights 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 special cases of calibration. There are numerous software packages that have been built to implement calibration. Examples include CALMAR which is widely used in the EU and Statistics Canada’s G-Est, both of which are available at no cost.

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.

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, often referred to as g-factors) should be neither too small nor too large. It is therefore a good practice to apply bounds to the calibration totals. While there is no specific rule for which range of calibration totals 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 4.10

Country Case: 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 Tools used by the Canadian Income Survey

| | |
|--|---|
| Counts of individuals in 15 sex × age groups | $\left. \begin{array}{l} \text{male} \\ \text{female} \end{array} \right\} \times \left\{ \begin{array}{l} 0 - 6 \text{ (both sexes)} \\ 7 - 17 \\ 18 - 24 \\ 25 - 34 \\ 35 - 44 \\ 45 - 54 \\ 55 - 64 \\ 65 + \end{array} \right.$ |
| 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 |

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 in which of these six classes each survey respondent who has received wages and salaries the respondent belongs. If the respondent did not receive wages and salaries, he is 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 control totals, 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 compensate for higher non-response rates at the top and bottom of the income distribution that cannot be sufficiently corrected for 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

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.

At each step of the weighting procedure, the distribution of the weight adjustments and of 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 to 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.

It can also be useful to reduce the weights of unit 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 the weight.

There is no rigorous or absolute procedure for general use for determining the limits for trimming or windsorising 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.

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;
- $\omega_i^{(HN)}$ the weight determined after non-response adjustment;
- $\bar{\omega}^{(HD)}, \bar{\omega}^{(HN)}$ their respective mean values; and
- any computed non-response weights outside the following limits are recoded to the boundary

$$\text{of these limits: } 1/C \leq \frac{\omega_i^{(HN)} / \bar{\omega}^{(HN)}}{\omega_i^{(HD)} / \bar{\omega}^{(HD)}} \leq C.$$

A reasonable value for the parameter is $C=3$.

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.

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 suffices.

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 4.11

Country Case: Calibration in Household Living Condition Survey 2009 in the 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 not 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 (Betti, 2011) show that the Kish effect of weights may depends to 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

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 sample size. Being able to estimate the sampling variance is the basis on which inference can be made in a 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.

Practical procedures for estimating sampling errors for such a survey: (i) must take into account the actual, complex structure of the design; (ii) should be flexible enough to be applicable to diverse designs; (iii) should be suitable and convenient for large-scale application, producing results routinely for diverse statistics and subclasses; (iv) should be robust against departure of the actual sample design from the ideal model assumed in the computation method; (v) should have desirable statistical properties such as small mean-square error of the variance estimator; (vi) should be economical in terms of effort and cost; and (vii) suitable computer software should be available for application of the method (Verma, 1991).

Linearization methods and replication methods are two broad practical approaches to the computation of sampling errors. 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 survey 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.

Examples of replication methods include the bootstrap, Jackknife, and Balanced Repeated Replication (BRR).

The variance estimates should also take into account the effect on 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.

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, 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.

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 4.12

Country Case: Calculating Variance for EU-SILC in Austria

Statistics Austria has developed its own tool for variance calculation. A package called `surveysd`⁶⁶ 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. A typical workflow with this package consists of three steps which are described in a hands-on manner⁶⁷ in the accompanying documentation:

- Bootstrap samples are drawn with rescaled bootstrapping in the function `draw.bootstrap()`.
- These samples can then be calibrated with an iterative proportional updating algorithm using `recalib()`.
- Finally, estimation functions can be applied over all bootstrap replicates with `calc.stError()`.

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 this can be specifically requested so that the longitudinal structure is preserved. 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 controls 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 implemented in `surveysd` repeats 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

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

⁶⁶ <https://github.com/statistikat/surveysd>

⁶⁷ <https://statistikat.github.io/surveysd/articles/surveysd.html>

disaggregated indicators of economic well-being, these statistics are only useful if they are sufficiently accurate for their intended use.

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 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.

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, 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.

The most commonly used measure used in dissemination tables is generally the CV which, by definition, is the ratio between 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, as can be 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). Statistics Austria has developed a special R-package (surveysd) for obtaining sampling errors using a bootstrap algorithm that is suitable also for longitudinal data collections such as EU-SILC. An extension of this package is currently under development which should automatically flag cells to be suppressed because of their large standard errors.

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 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.

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.

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 Statistics Canada's Canadian Income Survey, estimates that are derived from fewer than 25 records are suppressed. Here the number of records is either the number of individuals or the number of families depending on the statistic.

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 4.13

Country Case: United States Suppression Rules

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 also three-year estimates had 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 are Census Block Groups which typically have a population of 600 to 3,000 people. Public use micro data show only PUMAs - public use microdata areas - which tend to have populations of approximately 100,000 or more. In addition, tables are limited to those that pass a number of data quality filtering rules. Firstly, if more than half of the estimates in the table are not statistically different from 0 (at a 90% confidence level), then the table fails to meet the rule's requirements and is restricted from publication. Secondly, if the median CV value for the table is less than or equal to 61 percent, the table passes for that geographic area and is published; if it is greater than 61 percent, the table fails and is not published. (If the estimate is 0, a CV of 100% is assigned). CVs are calculated for each table's estimates, and the median CV value is determined. 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.

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 fewer 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).

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.

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 Centers (Statistics Canada, 2019).

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.⁶⁸

4.3 Other methodological issues related to measuring poverty

This section treats some further aspects that are not always part of the survey process, such as pooling, small area estimation (SAE), and rapid estimates.

4.3.1 Pooling

According to Verma et al (Verma et al, 2013), two types of measures can be constructed at the regional level by aggregating information on individual elementary units:

⁶⁸ <https://www.gegis.org/en/services/events/gegis-conferences/european-user-conference-6>

- a) Average measures, i.e. ordinary measures such as totals, means, rates and proportions constructed by aggregating or averaging individual values (such as population proportions in the area having certain characteristics relating to welfare).
- b) Distributional measures, such as measures of variation or dispersion among households and persons in the region. Such measures may depend on the distribution of characteristics in each region, or on the overall distribution in the whole national population.

The patterns of variation and relationship for the two types of measures can differ from each other, and hence involve separate statistical considerations. Average measures are often more easily constructed or are available from alternative sources. Distributional measures tend to be more complex and are less readily available from sources other than complex surveys; at the same time, such measures are more pertinent to the analysis of poverty, social exclusion and other aspects of well-being.

Where there is not enough data to produce reliable estimates from a single iteration of a survey, pooling of data may be used. For instance, using three-years of data pooled could produce useful estimates in this context. Both measures of types a) and b) could be performed using pooled data.

An important point to note is that, more than at the national level, many measures of averages can also serve as indicators of disparity and deprivation when seen in the regional context: the dispersion of regional means is of direct relevance in the identification of geographical disparity. In particular, the interest is in pooling over survey waves in a national survey in order to increase the precision of regional estimates.

A difficulty in pooling samples is that, in the presence of complex sampling designs, proper variance estimation may not be possible, at least in an almost exact way as introduced in 4.2.5, because the structure of the resulting pooled sample can become too complex or even be unknown. In any case, different waves of a survey (such as, for instance, EU-SILC) do not necessarily correspond to exactly the same population. The problem is akin to that of combining samples selected from multiple frames.

Pooling of estimates from each wave rather than of micro data sets is generally the appropriate approach to aggregation over time from such surveys. Under this approach, for each wave, a person's poverty status (poor or non-poor) is determined based on the income distribution of that wave separately, and the proportion of poor for each wave is computed. These proportions are then averaged over a number of consecutive waves.⁶⁹

Now, when we consider the variance, the issue is to quantify the gain in sampling precision from such pooling, given that data from different waves of a rotational panel are highly correlated.

A large proportion of the individuals are common in the different cross-sections of the panel. However, a certain proportion of individuals are different from one wave to the other. The cross-sectional samples are not independent, resulting in correlation between measures from different waves. Apart from correlations at the individual level, we have to deal also with additional correlation that arises because of the same structure (stratification and clustering) of the waves of a panel. Such correlation would exist, for

⁶⁹ It may be clarified that when averaging over waves is normally done at the macro level (e.g. poverty rates), and not at micro level (e.g. poverty status of individual persons). Nevertheless, estimation of the measures and their sampling precision requires access to micro data.

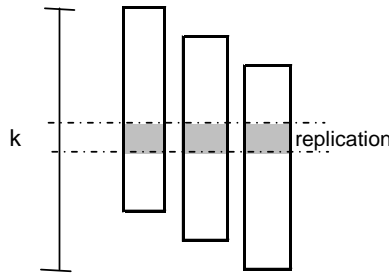
instance, in samples coming from the same clusters even if there is no overlap in terms of individual households.

The estimation of measures such as differences or averages of poverty measures can be straightforward when a replication method is used for the variance estimation. For this purpose, a replication variance estimation method, as introduced in 4.2.5, can be easily extended by building a coordinated set of replicates for the whole pooled sample along the following lines.

The total sample of interest is formed by the union of all the cross-sectional samples being compared or aggregated. Using as basis the common structure of this total sample, a set of replicates is defined in the usual way. Each replicate is formed such that when a unit is to be excluded in its construction, it is excluded simultaneously from every wave where the unit appears. For each replicate, the required measure is constructed for each of the cross-sectional samples involved, and these measures are used to obtain the required averaged measure for the replicate, from which variance is then estimated in the usual way. So, as example, if we have a dataset with three consecutive years and we want to estimate the average of the three years, we proceed as follows.

Figure 4.4

Replication variance estimation method for three consecutive years



It is suggested to construct a common structure of strata and PSUs from the union of the three datasets and assign to this common structure new weights equal to the average of the weights of the three years:

$$w_t^{(Common)} = (w_t)^{Average} = (w_1 + w_2 + w_3) / 3$$

For each year (t) and for each replication (k), we can estimate $y_k^{(t)}$ where $t=1,2,3$ and from this, the required statistic $y_k^{Average} = \sum_t a_t y_k^{(t)}$; that in our case is just $y_k^{Average} = (y_k^1 + y_k^2 + y_k^3) / 3$.

Verma et al (Verma et al, 2013), have implemented such methodology for Austria and Spain for 2011 EU-SILC data. In this case single year estimates for 2011 are compared with the averages of three years (2009-2010-2011). Also, in this case, averaging the poverty rate over three waves leads to a variance of this averaged estimator that is 30% less than the variance of the ARPR estimated from just a single wave for Austria and 35% less for Spain. Similar reduction is also found at regional level (NUTS2).

While the pooling of data can be quite easy to implement and can be used for both average and distributional measures. It is important to remember that access to micro data is required to properly estimate variance of pooled samples and that variance gains are not as large when then the yearly

estimates are highly correlated as is the case with rotating panel design which are common for surveys on income and poverty.

4.3.2 Small Area Estimation

There is a wide variety of small area estimation (SAE) techniques available, and the field is rapidly expanding. SAE has been used successfully on poverty rates and, in fact has its roots in the field, as demonstrated by the now classical example of Fay and Herriot (Fay and Herriot, 1979). 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. A standard reference on small area estimation methodology is Rao (Rao, 2003). See also, among others, Gosh and Rao (Gosh and Rao, 1994) and Henderson (Henderson, 1950).

It is, of course, not possible within the framework of this project to develop and evaluate SAE models for diverse poverty and related indicators in the specific situation of individual countries. Nor would it be appropriate to make such an attempt, given that the applicability of SAE methods is generally very country-specific since it depends greatly on the data available in each country. The knowledge and experience of national statisticians and other researchers about the specific possibilities and limitations in their own country can be expected to be superior to the approach from the single EU-SILC example mentioned below.

There are some serious limitations to the application of SAE methodology in the context of regional estimation in EU-SILC. But let us first note some potential merits of the procedure:

- a) SAE methods such as EBLUP make use of external data aggregated to NUTS2 (area level) only. R-codes are available under projects funded by the EU 7th Framework program (such as SAMPLE, AMELIE, etc.); estimates could be performed every year, given that such external sources are available.
- b) Poverty mapping permit to obtain estimation at NUTS2 level with high precision (very low standard errors).

There are four types of limitation to be faced:

- a) The first concern is the lack of external data for the purpose of making SAEs. The methodology needs information from census data, which are usually 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 correlated sufficiently highly 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 require specialized knowledge and software.
- c) The 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) A most important merit of EU-SILC is the provision 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.

4.3.3 Rapid estimates

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.

In 2017, Eurostat, the statistical office of the European Union, and the United Nations Statistics Division jointly published a document called 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.

A number of national statistical offices have started to evaluate rapid estimates so they can produce new statistical products that are timelier than traditional 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).

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.

5 Improving accuracy of measurement: the use of supplemental or experimental poverty measures

What is the role of differences in the cost of living? What adjustments need to be made to poverty thresholds and resource measures 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?

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 as 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?

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

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

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.

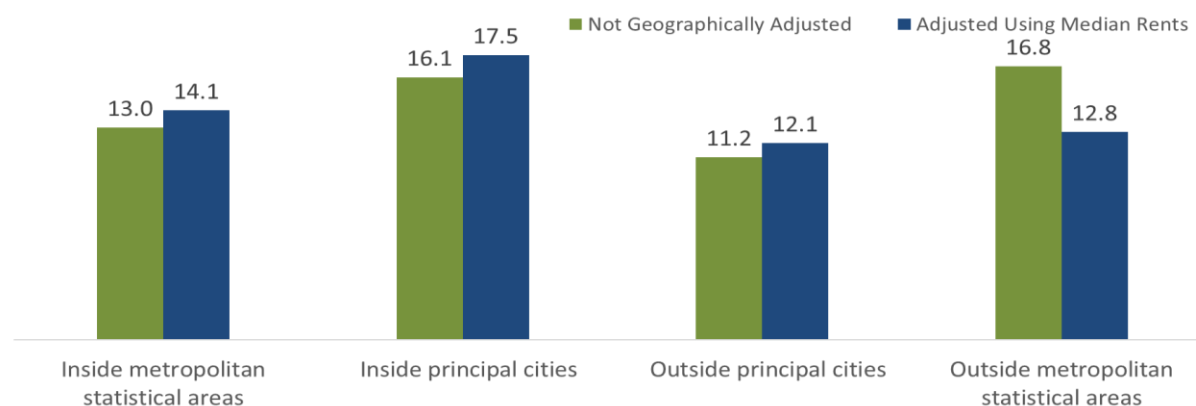
Ideally a price index should consider differences in the consumption bundle, differences in the weights of specific items in the consumption bundle, as well as differences in prices across regions. In addition, if the purpose of the price index is to adjust poverty thresholds, it should be constructed 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

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.

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.

5.1.1.2 Canada: Market Basket Measure

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.⁷⁰ 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. MBM thresholds are used in Canada to generate low-income statistics and that were designated as "Canada's Official Poverty Measure" by the Government of Canada in 2018.⁷¹

5.1.1.3 Russian Federation

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.

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-kind units of annual consumption. The nutritional and caloric values 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.

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

⁷⁰ <https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/pop165-eng.cfm>

⁷¹ <https://www.canada.ca/en/employment-social-development/news/2018/08/government-of-canada-launches-first-poverty-reduction-strategy.html>

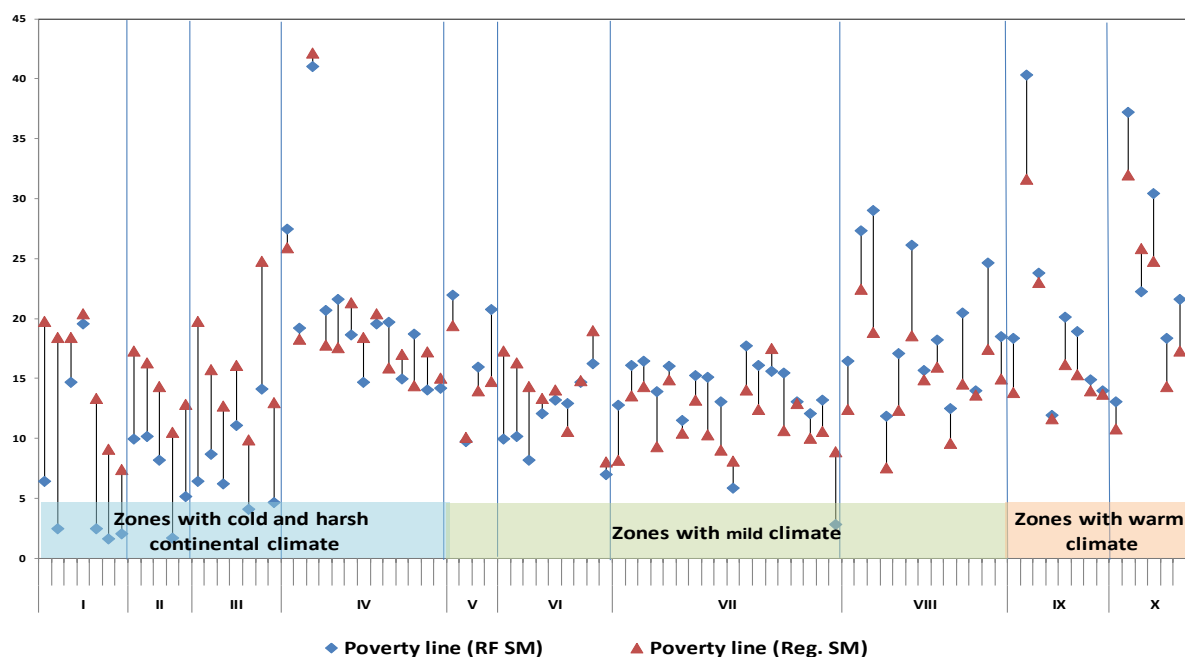
methodological guidelines to obtain comparable absolute poverty thresholds at regional level. Based on scientific research on regional consumption patterns, 10 different zones were identified which primarily reflect different climate conditions. Regional food baskets are determined using specific nutritional 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.

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.

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

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.

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 household 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 on the basis 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" (Atkinson, 2017).

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

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.

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

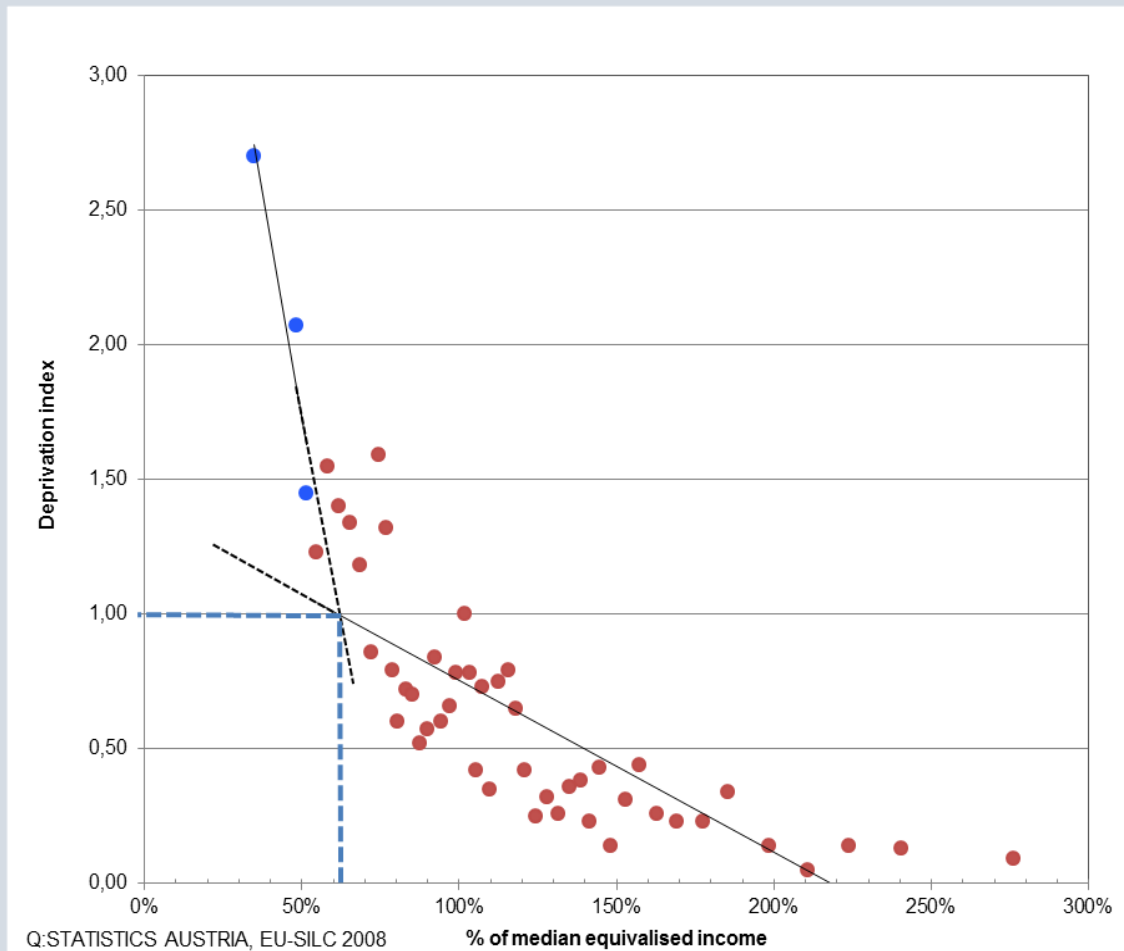
Assessment of the 60% of median threshold for the elderly in Austria

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. **Error! Not a valid bookmark self-reference.** 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 disproportionally higher than above 60% of the median. 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

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.⁷²

⁷² 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).

Recommendations: As a general rule, poverty measurement has to ensure equivalence of standards of living for all regions and groups within a country, notably with regard to 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 when nations are large 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

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, 2017). The Guide concluded in the recommendation that the square root of household size should be applied as an equivalence scale for CES countries.

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 the “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 Haagenars et al (Hagenaars et al, 1994) 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 $(adults + 0.5\ children)^{0.7}$.

These assumptions regarding economies of scale in consumption may vary significantly depending on the bundle of goods included in the threshold. As discussed in the section 5.2.1 Social transfers in kind (STIK), 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. 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 child care is included in the threshold there would be considerably higher needs for families with infants than families with older children.

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.

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, for pensioners, and for children.⁷³ 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.

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 child-specific items in consumption/expenditures modules could provide at least an estimation of the percentage of family resources that are allocated to children.

Among the countries that experimented with individual consumption (or age-categorical consumption) there are (not in the region):

- 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 Demographie, 2015)

Most of these surveys focus on food consumption, which is the biggest source of intra-household differences in consumption. However, we cannot ignore 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.

As said, a broad classification presented in Hagenaars et al (Hagenaars et al, 1994) could be summarised as follows: normative or social security scales; subjective scales; scales based on demand models.

⁷³ 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

Box 5.2

Two country examples for the estimation of an equivalence scale based on a demand model – Tajikistan and Ukraine

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 which can be used to validate the sensitivity of per capita base poverty estimates which are still common in Tajikistan today.

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)

Recommendations: Equivalence of measurement should be assessed empirically. Sensitivity analysis can compare poverty profiles of official poverty measures with supplemental poverty measures using alternative equivalence scales.

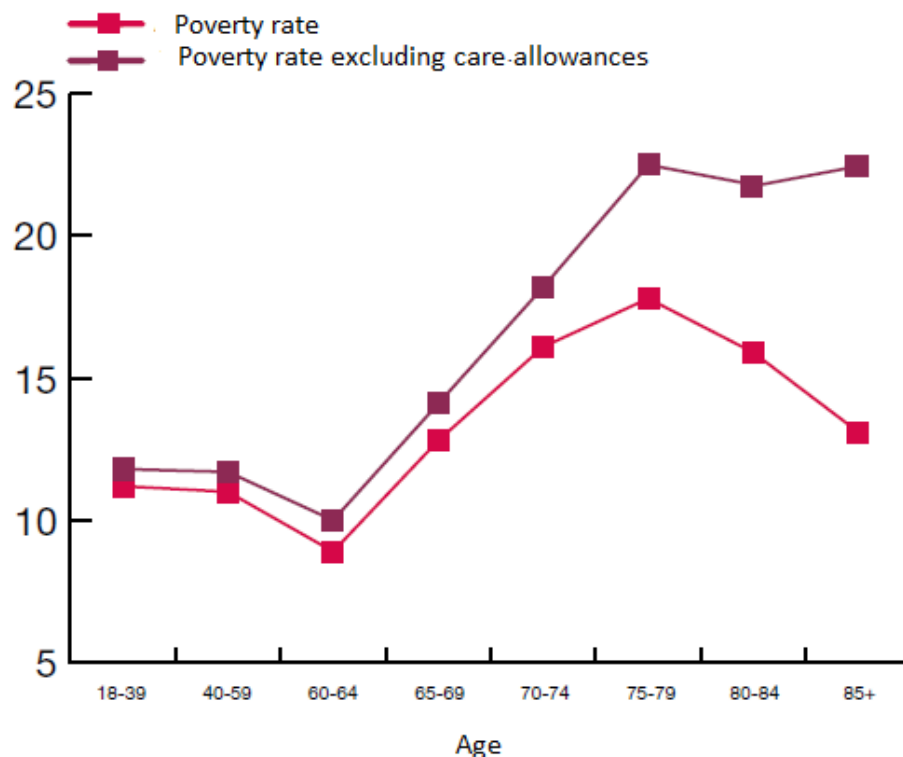
While simple equivalence scales based on the square root of family size are useful for international comparability, countries should explore the empirical development of equivalence scales that go beyond family size to include differences in needs based on age, disability status and health. If Social 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

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, 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 (MacInnes et al, 2014).

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: Further research on the special needs of the disabled is necessary to ensure these are taken into account in 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)

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 (NPISHs) (OECD). 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.

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.

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.

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.

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.

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.⁷⁴ An advantage of the inclusion of STIK in the supplemental poverty measure in the United States is the ability to disaggregate the impact of each specific transfer on the poverty rate. For example, in 2016, holding all else the same and assuming no behavioural changes, the major food assistance program (SNAP) kept 3.6 million people out of poverty while housing subsidies reduced poverty by 3.1 million (Fox, 2016).

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

⁷⁴ 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).

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 was 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

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.

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 UK’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 households’ 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.⁷⁵
- 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).

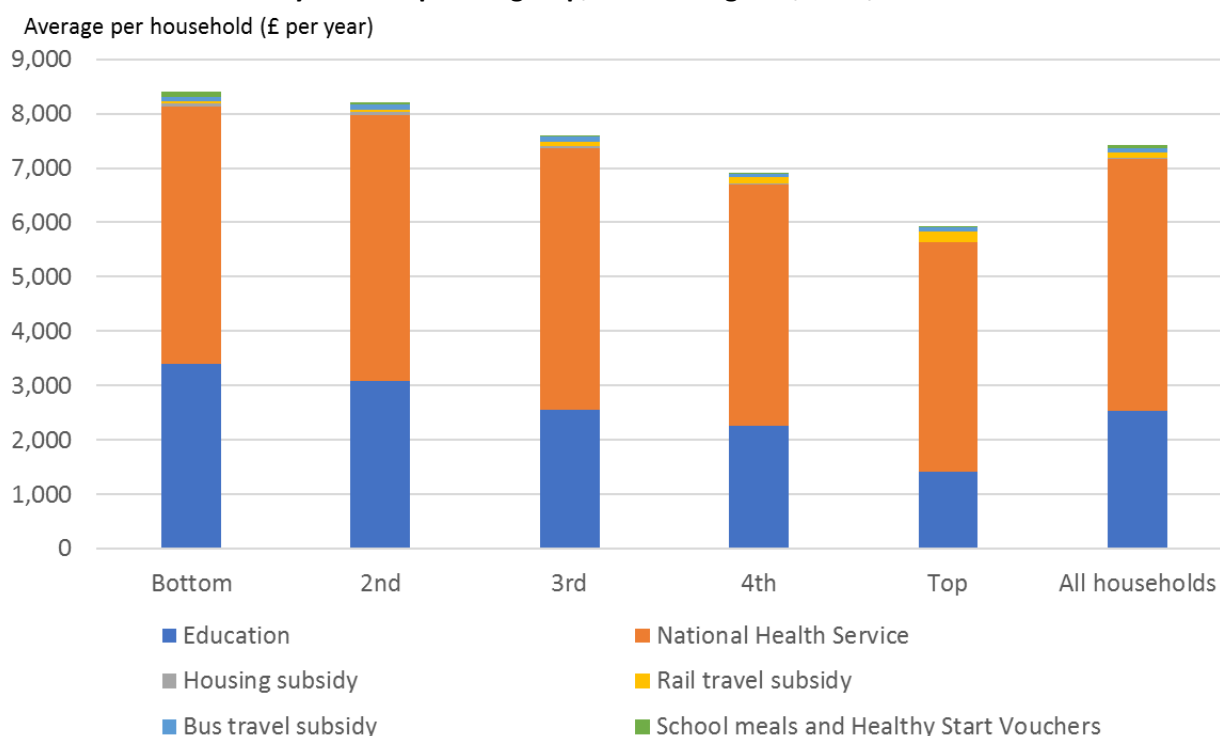
⁷⁵ Ideally, an adjustment should be at least considered for the use of private health care services. However, this is not felt to be feasible from the data available. For example, whilst it is possible to identify individuals paying for private health insurance, it is not possible to ascertain the level 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.

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.

- Impact of STIK on income inequality and poverty: 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



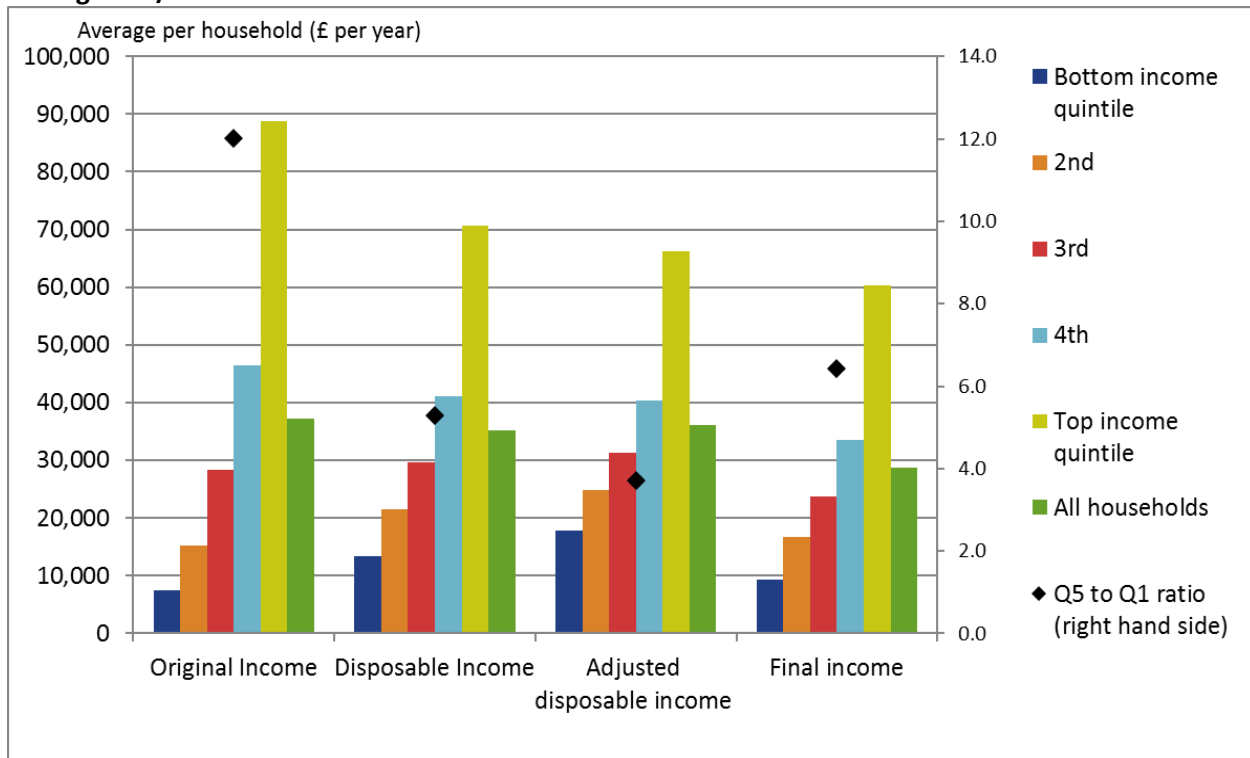
Error! Not a valid bookmark self-reference. summarises the effect that these STIK has on the United Kingdom income distribution. In this figure, original income is defined as income from employment, private pensions, investments and other non-government sources. In 2016/17, the richest fifth of households 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).

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 shown above. 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

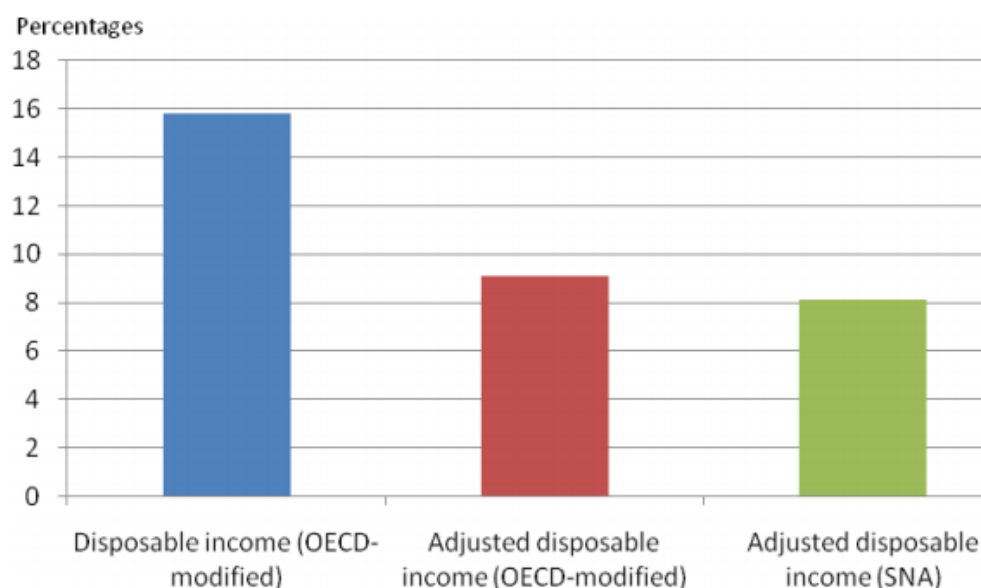


The United Kingdom's main poverty statistics are produced and published by the United Kingdom's Department for Work and Pensions (DWP) in a publication called Households Below Average Income (HBAI). STIK are not included within the HBAI statistics.

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).

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



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%.

5.2.1.2 Mexico – Strategies to measure self-consumption

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. Since 2016, in the National Survey of Household Income and Expenditures (**ENIGH**) current income is no longer disaggregated in its monetary and non-monetary components. The new presentation is due to the new treatment given to the self-consumption variable.

In accordance with international statistical standards and practices, the 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 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 is established in accordance with the estimate made by the informant himself, in relation to the retail market value of the goods and services received in a given period, by virtue of the performance of some activity economic subordinate.

The survey also asks respondents to report income from “independent work” or self-employment. 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 common factor they have is that the people who work for them work for themselves and are owners of an unincorporated company, and therefore belong to the household sector. Thus, independent work income is all those incomes 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.

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 of time. 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.

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 which 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

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 UK, 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 valuation⁷⁶;
- Whether values are allocated at the household or individual level;
- Which equivalization scale is used; and
- What at-risk-of-poverty threshold is used.

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.

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 benefits in kind 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.

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.

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 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

⁷⁶ 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.

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.

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 child care 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 child care should have lower out of pocket expenditures for this item than an individual not receiving this benefit.⁷⁷

Which social transfers in kind 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 defense 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.

5.2.2 Accounting for housing wealth

Low income (poverty) measurement requires a meaningful and standardised measure of resources available to the household to spend on its consumption. 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.

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.

⁷⁷ On the other hand, the family that cannot afford child care 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 will be considered poorer than the individual without health insurance because his/her resources will reflect out of pocket payments for health insurance.

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 (Sustova, 2016).

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 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 is less than 10%. 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 concludes 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?”

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.⁷⁸

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

⁷⁸ See for example Pelletiere, 2008.

imputed rent for mortgage free homeowners. The U.S. Supplemental Poverty Measure also adjusts the thresholds by tenure status.

Recommendations: 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.

Figures on total STIK should be presented together with poverty measures wherever possible as a useful indicator in its own right.

Social transfers in kind (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 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.

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.

The value of STIK needs to be capped to a meaningful maximum. If STIKs are included in the resource measure, its value should never exceed the poverty threshold.

If STIK are included in the resource measure this may affect the equivalence scale.

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.

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.

5.2.2.1 Impact of imputed rent on Canadian poverty estimates

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

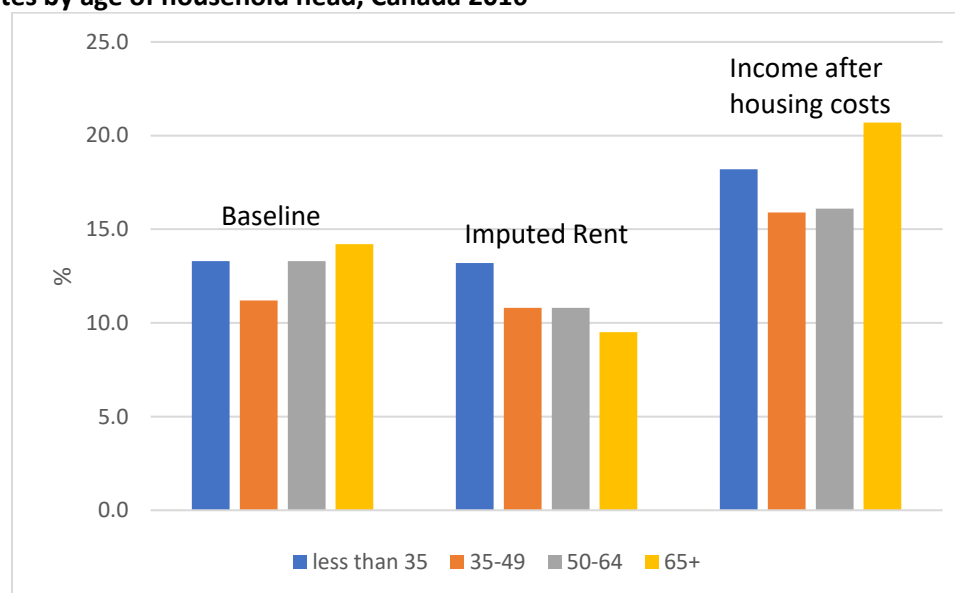
of total mortgage payments. For the user costs approach, estimates were done with both a 2% rate of return and a 3% rate of return. In this study, low income (poverty) was defined as having household income below 50% of median equivalised household income.

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 the rate of return assumptions. Subtracting housing costs from income increased poverty rates for this group from 5.4% to 8.4%.

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 5.8

Low-income rates by age of household head, Canada 2016

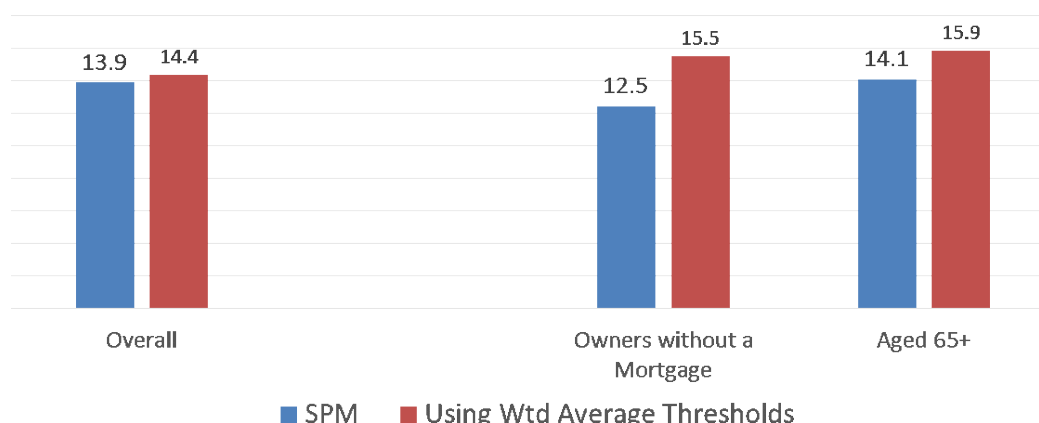


Source: Calculations using the Canadian Income Survey, 2016.

5.2.2.2 Impact of Differential Thresholds on the Supplemental Poverty Measure for the U.S

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 thresholds are 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 5.9

Impact of Differential Thresholds on the U.S. Supplemental Poverty Measure

Source: U.S. Census Bureau, Current Population Survey, 2019 Annual Social and Economic Supplement.

5.2.2.3 Challenges and Issues for Further Research

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.

Recommendations: As an alternative to imputing rent, residual income after actual housing cost may be considered as a resource measure which reflects housing equity.

Surveys should consider adding questions to measure housing costs and property values to facilitate the calculation of imputed rent.

Given that the choice of method may 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.

5.2.3 Asset poverty: Why assets should be considered in the measurement of poverty and how it could be done

In research, the joint consideration of income and wealth has been recognized as important for in 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.

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).
- 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, 2017).

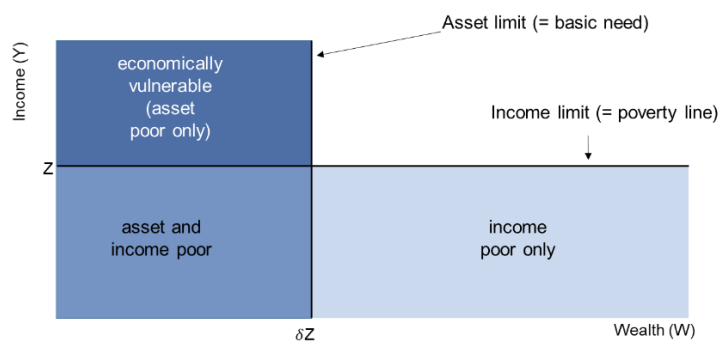
It has already been shown several times that the choice of approach (uni- or two-dimensional approach) has a considerable influence on the results (e.g. Kuypers and Marx, 2016; Azpitarte, 2010; Brandolini et al, 2010; Tonkin et al, 2016; Törmälehto et al, 2013). Ravazzini et al (Ravazzini et al, 2017) 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

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 5.10):

Figure 5.10

Illustration of the two-dimensional poverty index



Source: Balestra and Tonkin, 2018.

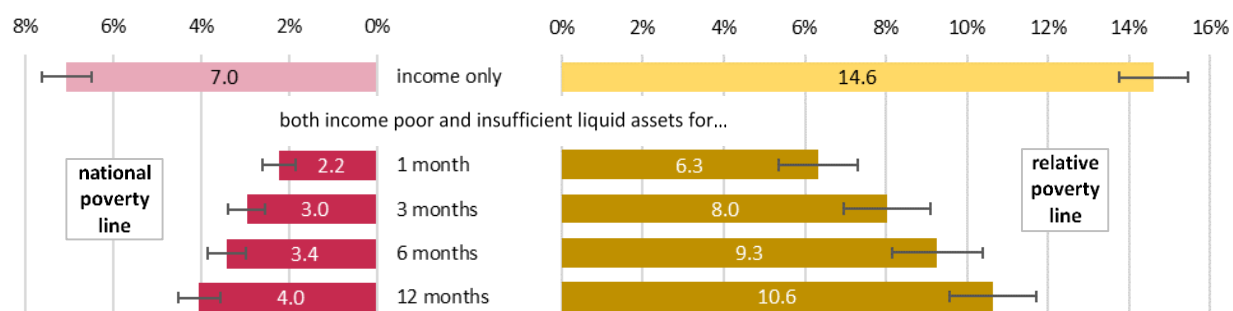
For the operationalization of the asset-based poverty rate, various choices still need to be made: which wealth concept will be considered, what type of poverty line is used, for which reference period the assets have to suffice, and how the assets are to be equivalized.

- **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 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.

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 5.11). 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 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 5.11

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.

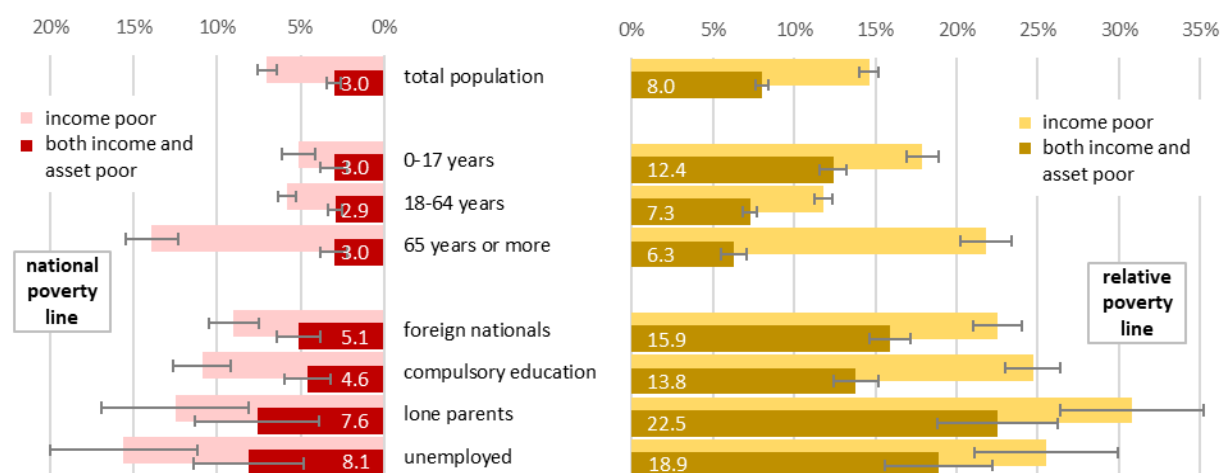
For both poverty lines, the inclusion of assets especially reduces the poverty rates of the older population compared with income poverty, meaning that the income poor aged 65 years and over may fall back on liquid assets particularly often. This finding is in line with many other studies on asset poverty (e.g. Kuypers and Marx, 2016; Azpitarte, 2012; Brandolini et al, 2010; Caner and Wolff, 2004). 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.

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.⁷⁹ 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 5.12

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

⁷⁹ 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.



Source: FSO – CH-SILC 2015 (provisional data on wealth, version 07.06.2018)

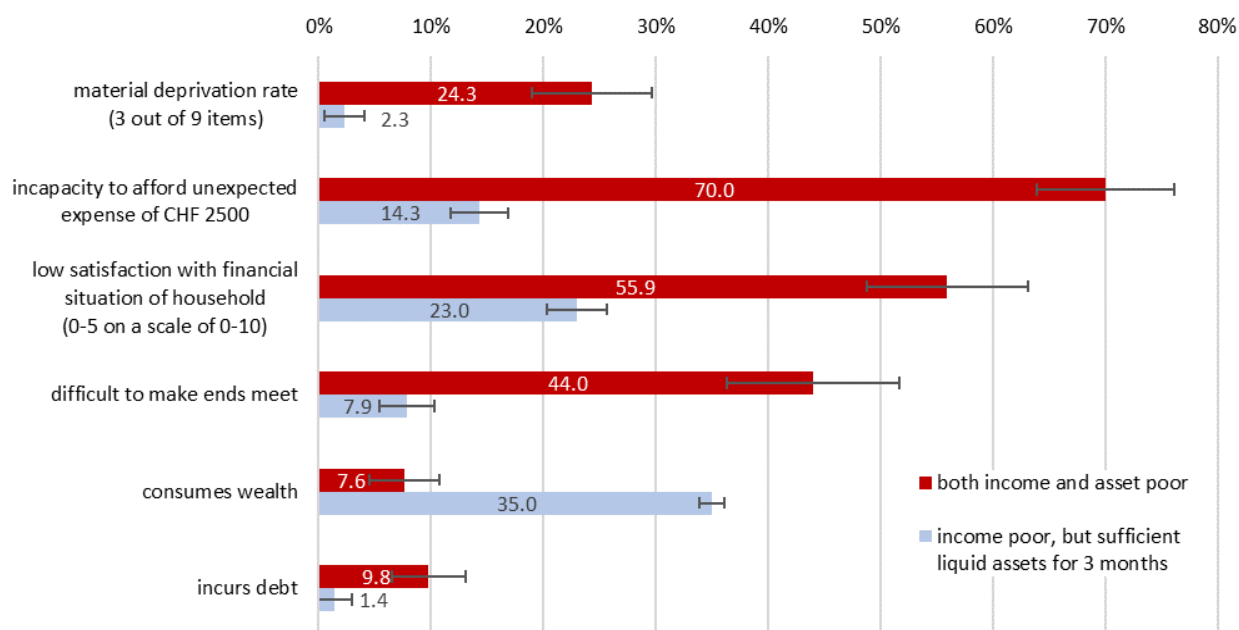
Notes: Only liquid assets such as deposits, bonds, shares, mutual funds etc. are taken into account. The national poverty line consists of a fixed amount for living expenses, individual housing costs and CHF 100 per month and per person aged 16 or over. The relative poverty line is calculated as 60% of median equivalised income including imputed rent.

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 difficulties⁸⁰ if they have assets in reserve compared with those who do not have any such resources to fall back on (see Figure 5.13).

Figure 5.13

Share of population indicating material or subjective difficulties, by poverty status, in percent of the population (individuals), with 95%-confidence interval

⁸⁰ 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)



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.

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 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.

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 standard 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

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 show 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, and in Switzerland assets are also included in the assessment of eligibility for social welfare support. Since there is not access to asset data every year, the income poverty rate must, however, remain the most important poverty statistic indicator for the time being.

At the political level, the new results could also further fuel debate on the welfare state if the fact that a considerable proportion of the income poor in Switzerland own certain assets were misinterpreted as though they were actually "wealthy". 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.

Recommendations: 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 the 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?

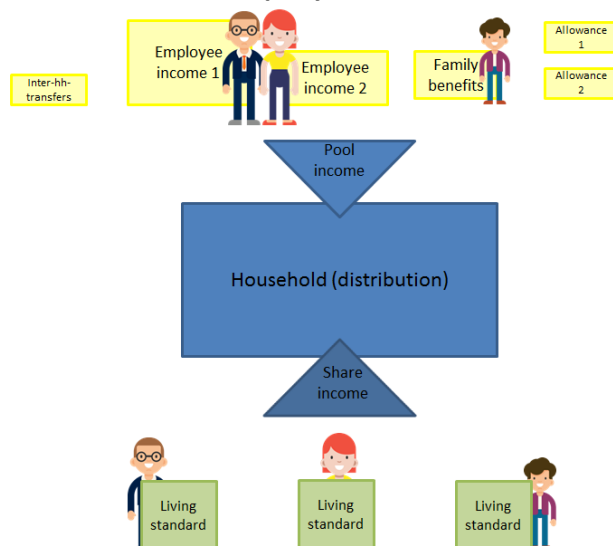
In measuring poverty, are persons in a household by definition **treated equally or should each person be weighted by what he or she actually contributes to the household income or takes from the household income for spending?** The question of intra-household distribution is usually ignored by standard 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 standard by the European concepts for poverty indicators.⁸¹ Therefore we speak of the "standard (EU-indicator) approach" when no difference between the household members is made relative to their

⁸¹ 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).

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 5.14

The standard household income model – exemplary household



Given these implicit assumptions, to be correct, 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 social characteristics on an individual basis is therefore distorted. Gender differences, for example, can accurately only be reported for single person households/families.

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 “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.

Literature on the assumption of sharing of resources and its implications for economic theory on the household is extensive and cannot be summarized here in great detail.⁸² “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

⁸² For a very good overview on potential indicators of intra-household decision-making, related methodological challenges and data availability see Krieger, 2018. This work has been done by the UNECE Task Force on measuring household power and decision-making.

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.

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.

An important contribution to answering each of these questions can be found in the work by **Ponthieux** (Ponthieux, 2017) using EU-SILC data. The author tries to decompose the well-established poverty indicator “at-risk-of-poverty rate” by sex for women and men that live in couple households. She makes use of a question of the ad-hoc module of the EU-SILC survey in 2010 on “Intra-household sharing of resources” that 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

This question proves useful to assess the degree of income pooling. However, two problems arise if one wants to use this information for computations. 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 our aim is to know about the living standard 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 that is proposed with all practical restrictions of the data works on the assumption that income that is kept separate only increases the living standard of the income recipient. The part of the income that is pooled and all income components received at household level are attributed to both partners of a couple equally by equalisation using the Eurostat-scale. What is kept separately is not equalized. Thus, the author arrives at different income values for each partner using this “modified” approach. As a first result she shows the share of income of men and women in the total household income: As expected due to usually higher employment income (and pensions) of men they have 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).

The explanation given is that better earnings by one (the male) partner are nearly perfectly counterbalanced by the couples' pooling regimes (p. 184f.) – at least when looking at country averages.

Using the same “modified” income and comparing it to the poverty threshold calculated by the standard methodology (60% of median for the targeted couples) the author shows a higher “modified” poverty risk for women as compared to men. For example, in Austria women with this approach have a poverty rate 1.4 times larger than men. (p. 186f).

We can learn from this study that 1) 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. 2) For the extent of the bias in the poverty rates we see **large country differences ranging from practically no gender differences for persons living in couple households in Lithuania to a factor of 1.5 in Greece and even 1.8 in Malta.**

Another attempt to disaggregate the common at-risk-of-poverty indicator for a special research issue, namely in-work-poverty, was presented by Heuberger and Knittler (Heuberger and Knittler, 2018). This approach uses **an alternative risk-of-poverty threshold** that takes into account the number of people contributing to the household income. In contrary to the standard approach, the at-risk-of-poverty threshold is applied to a newly computed **individual personal income** for each member of the household. To compute this alternative at-risk-of poverty threshold the authors use the common at-risk-of-poverty threshold of 60% of median income to reset the equivalisation by multiplying it by the OECD's equivalent scale for each household and then dividing by the number of earners in the household. This threshold marks the minimum income each adult member of the household has to obtain on an individual level not to be identified as at-risk-of-poverty. According to the number of adults and dependent children living in a household, results differ to the standard approach; for single or one adult households there is no difference in the adult's results.

To compute the individual income earned income as well as social transfers and income from social security on the individual level are taken into account in a first step. Then to estimate the individual share of the income components on household level their sum is divided by the number of adults in the household. The sum of earned personal income and the share of household income for each adult are compared to the household specific poverty threshold to identify whether a person is at-risk-of-poverty or not.

For their specific group of interest, working adults, Heuberger and Knittler found out that especially for women there is a great difference in the poverty risk when applying their alternative approach. Throughout all socio-demographic groups they are reported to have a higher in-work-poverty risk than men (and as in the standard approach), whereas in the standard approach the risk of in-work-poverty is approximately the same for both sexes (Heuberger and Knittler, 2018: 240). Combining both concepts allows them to understand financial dependencies within a household, for example when a woman is at-risk-of-poverty with her individual income but not with the equivalised income, which assumes, that she is financially supported by her partner or other household members.

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**, which contains a number of items on deprivation at the individual level. They map the extent of intra-couple inequality in deprivation, and analyse its determinants. Their findings point to a higher deprivation rate of women compared to men living together as couples. Only access to internet is not significantly unevenly distributed. 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 at absolute living standards and enforced lack – 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.

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.

Recent research by **Burchardt and colleagues** (Burchardt and Karagiannaki, 2018) of the LSE also deal 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 individual’s contribution to total household income make 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 income pooling assumptions on Austrian poverty rates

The following section presents some **analysis of the impact** of poverty measurement on poverty rates within couple households using two different approaches to adjust for differing assumptions regarding income pooling and resource sharing. The first uses the methodology developed by Ponthieux which relies on the income sharing questions included in the 2010 EU SILC. The second methodology developed by Heuberger and Knittler (Heuberger and Knittler, 2018) looks at in-work-poverty.

Table 5.2 shows the poverty risk of women and men in the given sample with the Eurostat standard approach and with an adjusted at-risk-of poverty threshold as well as for the alternative approach developed by Ponthieux et al (Ponthieux et al, 2017).

Table 5.2

At-risk-of-poverty by sex I

| | At-risk-of poverty in % | |
|--|-------------------------|-------|
| | Men | Women |

| | | |
|--|-----|------|
| Standard approach | 8.0 | 8.0 |
| Standard approach with adjusted at-risk-of poverty 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.

Whereas in the standard approach the risk-of-poverty is the same for both partners the alternative calculations show a gender ratio of the modified poverty risk (women/men) of 1.9.

While Ponthieux's approach opens new possibilities to assess gender inequalities within the household, this analysis is 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.

An alternative approach is to apply the Heuberger and Knittler method to the sample of EU-SILC 2010 in a very basic way, using the personal income and the household income as computed by Ponthieux.⁸³ As the dataset comprised only households with couples (and maybe dependent children) the at-risk-of-poverty threshold was in all cases multiplied by the sum of the household's equivalent weights and divided by two. The personal income was complemented by the household income divided by two, which does not consider the existence of children in the household. Results show a far higher at-risk-of-poverty-rate for women than for men in couple households, regardless of the number of dependent children. Relative to Ponthieux's approach, non-pooling is assumed for all households here, which means that individuals are expected not to profit from their partners income at all. The gender ratio (women/men) of this alternative threshold approach rises to 13.5 versus 1.9 for the calculations according to Ponthieux. Focusing on households where both partners have a personal income the gender ratio is at 12.2.

Table 5.3

At-risk-of-poverty by sex II

| | At-risk-of poverty in % | |
|---|-------------------------|-------|
| | Men | Women |
| Standard approach | 8.0 | 8.0 |
| Individual personal income approach and alternative threshold | 2.5 | 33.8 |

⁸³ Thus, using the same data as in the calculation according to the first model above for the sake of comparability of results.

| | | |
|---|-----|------|
| Individual personal income approach and alternative threshold for couples with personal income >0 | 1.9 | 23.2 |
|---|-----|------|

Source: Statistics Austria, EU-SILC 2010.

Note: Couple households in the dataset of the EU-SILC module 2010

This approach shows, that there is a possibility to combine household and individual perspective even when focusing on individual incomes. Established for the context of working poor it is a supplement to existing indicators. To measure income gaps within households in general, further development could additionally consider the number of children in the household for the computation of the threshold as well as for the income situations of adults. Maybe the construction of an adjustable at-risk-of-poverty threshold for households in different living conditions apart from the one using the Eurostat scale could also contribute to further results.

5.2.4.2 Challenges and issues for further research

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 the persons level and data on the distribution of income within households.

Recent changes to the EU-SILC, are a good example going in this direction of measuring absolute living standard 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 households 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 study⁸⁴. The next step in EU-SILC was to consider also the special situation of children. 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.

In order to have an empirical basis to disaggregate household income and assume an individual living standard 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

⁸⁴ 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.

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.

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 standard 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).

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 standard. 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.

Also, for both strands – the individual absolute living standard 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 common understanding of questions and answer options between respondents.

These recommendations are in line with World Bank Recommendation 13 (World Bank, 2017: 114) to disaggregate poverty figures by gender and age. Although this discussion has put an emphasis on the dimension of gender inequalities within households, there are many more breakdowns by individual characteristics that might be worth exploring, especially as to what concerns the situation of children within their households. 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). And, lastly, in general the concept of household vs. family vs. dwelling has practical implications: A future topic because of its societal relevance has to be how to measure the living standard of children who belong to more than one household (because of their parents living apart).

Recommendations: 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.

Official poverty rates for men, women, children (or other socio-demographic subgroups) should therefore be accompanied with results which consider unequal sharing of resources. At a minimum, NSIS should carry out sensitivity analysis for poverty profiles contrasting the conventional full pooling assumption with partial pooling and full separation of resources.

Wherever possible, income and material living standards should be considered in combination to validate assumptions of within household income distributions. Existing questions on sharing of personal income and/or personal material living standard can be adapted with relatively little additional effort from EU -SILC.

5.2.5 Going beyond income: The role of multidimensional poverty measures

While poverty has traditionally been measured in terms of income, the SDGs define poverty 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.

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).

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 the Alkire-Foster Method (Alkire and Foster, 2011) to analyse non-monetary deprivations and their interconnections and enable the creation of high-impact policies that address multiple interconnected deprivations and accelerate progress towards the SDGs.

The Global Multidimensional Poverty Index (MPI) 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, 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.

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).

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.

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 MPI can be used as an analytical tool to identify the most vulnerable people, show aspects in which they are deprived and, consequently, to target resources and design policies more effectively. Because the MPI measures outcomes directly, it will immediately reflect changes such as school attendance, whereas it can take time for this to affect income, thus may show outcomes of effective policy interventions quickly. Finally, in the understanding that each country is different, the MPI can be tailored to local context using country-specific data and indicators to provide a better understanding of country-level poverty. The governments of 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 and included nationally adapted multidimensional measures of poverty in their long term social development strategies, 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.

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.

Figure 5.15

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).

Recommendations.

Following the SDGs, countries should move to 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.

5.3 Recommendations

5.3.1 Adjustments to thresholds

5.3.1.1 Spatial differences within countries with regard to consumption and income poverty

As a general rule, poverty measurement has to ensure equivalence of standards of living for all regions and groups within a country, notably with regard to 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 when nations are large with heterogeneous cultural norms and price levels. Relative thresholds such as those used in the European Union are a convenient practical alternative.

5.3.1.2 Household size and age of household members

Equivalence of measurement should be assessed empirically. Sensitivity analysis can compare poverty profiles of official poverty measures with supplemental poverty measures 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. If Social Transfers in Kind (STIK) are included in resources, special consideration needs to be given to appropriate equivalence scales.

5.3.1.3 Disability status/High medical expenditures

Further research on the special needs of the disabled is necessary to ensure these are taken into account in poverty measures. These adjustments may be done on the thresholds, perhaps through equivalence scales, or on the resource-side of the measure.

5.3.2 Adjustments to resources

5.3.2.1 Social Transfers in Kind (STIK)

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.

Figures on total STIK should be presented together with poverty measures wherever possible as a useful indicator in its own right.

Social transfers in kind (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 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.

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.

The value of STIK needs to be capped to a meaningful maximum. If STIKs are included in the resource measure, its value should never exceed the poverty threshold.

If STIK are included in the resource measure this may affect the equivalence scale.

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.

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.

5.3.2.2 Accounting for Housing Wealth

As an alternative to imputing rent, residual income after actual housing cost may be considered as a resource measure which reflects housing equity.

Surveys should consider adding questions to measure housing costs and property values to facilitate the calculation of imputed rent.

Given that the choice of method may 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.

5.3.2.3 Asset Poverty

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 the income poverty seems to be the easiest to communicate to the public.

5.3.2.4 What is the role of unequal sharing of resources within households

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.

Official poverty rates for men, women, children (or other socio-demographic subgroups) should therefore be accompanied with results which consider unequal sharing of resources. At a minimum, NSIS should carry out sensitivity analysis for poverty profiles contrasting the conventional full pooling assumption with partial pooling and full separation of resources.

Wherever possible, income and material living standards should be considered in combination to validate assumptions of within household income distributions. Existing questions on sharing of personal income and/or personal material living standard can be adapted with relatively little additional effort from EU - SILC.

5.3.2.5 Going beyond income: the role of multi-dimensional poverty measures

Following the SDGs, countries should move to 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.

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. (2017). *Monitoring Global Poverty: Report of the Commission on Global Poverty*. The World Bank. Available from: <https://openknowledge.worldbank.org/bitstream/handle/10986/25141/9781464809613.pdf>
- 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>
- Azpitarte, 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.
- Azpitarte, 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>
- 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
- 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.
- 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/
- 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>
- Chiappori, P.A. (1992). Collective Labor Supply and Welfare. In: Journal of Political Economy 100/3: 437-467
- Cochran, W.G. (1977). Sampling Techniques. New York: John Wiley & Sons, Inc.

- 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>
- De La O Campos, A.P., Villani, C., Davis, B., 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>
- 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. Journal of Economic Literature, vol. 50, no. 4, pp. 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., 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>
- (2013). 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>
- (2014), Methodological Guidelines and Description Of EU-SILC Target Variables, <https://circabc.europa.eu/sd/a/2aa6257f-0e3c-4f1c-947f-76ae7b275cfe/DOCSILC065%20operation%202014%20VERSION%20reconciliated%20and%20early%20transmission%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. <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>
 - (2019). EU-SILC methodology – concepts and contents. Statistics Explained. Available from: <https://ec.europa.eu/eurostat/statistics-explained/pdfscache/35201.pdf>
- 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.
- 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>
- 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>
- 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>
- (2017). 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>
- (2018). 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
- 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., 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.
- 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.
- 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.
- 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, 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.
- 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 π . In: *Journal of Official Statistics*, 8: 183-200.
- 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, August 21-27, 2016.
- 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"*.
- 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, 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

- 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.
- 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.
- 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.
- (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>
- 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.
- 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>
- 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. (2003). *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. *Journal of Public Economics*, 22: 89–102
- 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
- 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>
- (2017). G-Link A Probabilistic Record Linkage System. Available from: <https://pdfs.semanticscholar.org/abd4/47d16fa837714d582ccd43b18d2285980a8f.pdf>
- (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>
- 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>
- Sustova, S. (2016). The measurement of imputed rents and social transfers in kind in the OECD: results from a questionnaire. OECD Centre for Opportunity and Equality. Powerpoint.
- 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 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>.
- 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).
- Tourangeau, R., Johnson, T., Bates, N. and Wolter, K. (Ed.) (2014). Hard-to-Survey Populations. Cambridge University Press.
- 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 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>
- 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/Canbera_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

- (2017). Guide on Poverty Measurement. New York and Geneva: United Nations. Available from: <https://www.unece.org/fileadmin/DAM/stats/publications/2018/ECECESSTAT20174.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
- 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>
- United Nations (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
- (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>
 - (2019). 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>
- 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
- OHRHC (2008). Convention on the Rights of Persons with Disabilities. Available from: <https://www.ohchr.org/EN/HRBodies/CRPD/Pages/ConventionRightsPersonsWithDisabilities.aspx>
- OHRHC (2018). A Human Rights-Based Approach to Data. Leaving no one behind in the 2030 Agenda for Sustainable Development. Available from: <http://hrbportal.org/wp-content/files/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.
- 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>
- 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

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>