MAKSWELL
MAKing Sustainable development and WELL-being frameworks work for policy analysis

www.makswell.eu

Horizon 2020 - Research and Innovation Framework Programme
Call: H2020-SC6-CO-CREATION-2017
Coordination and support actions (Coordinating actions)

Grant Agreement Number 770643

Work Package 3
Regional poverty measurement as a prototype for modern indicator methodology

Deliverable 3.1

Report on methods and data base for regional income and household expenditures

September 2019

Destatis, HCSO, Istat, Statistics Netherlands, Pisa University, Trier University

This project has received funding from the European Union's Horizon 2020 research and innovation programme.

Deliverable D3.1

Report on methods and data base for regional income and household expenditures

Authors

Destatis:
Thomas Zimmermann

HCSO:
Krisztina Kollár, Antónia Schwartz and Balázs Jankó

Istat:
Isabella Siciliani and Fabio Bacchini

Statistics Netherlands:
Jan van den Brakel

Pisa University:
Caterina Giusti, Monica Pratesi

Trier University:
Christopher Caratiola, Hanna Dieckmann, Florian Ertz, Laura Güdemann, Anne Konrad, Ralf Münnich, Anna-Lena Wölwer

Summary

One of the key topics of the Europe 2020 strategy is 'Poverty and social exclusion'. The target is to reduce the number of people at risk of poverty or social exclusion by 20 million until 2020 compared to 2008 (Eurostat, 2018). Furthermore, the Millennium Development Goals focus on the reduction of poverty as well. Changes in poverty and inequality happen at low regional levels. However, poverty and inequality indicators have primarily been estimated at the national level (EU-SILC data). As a large share of the EU's budget is directed to its cohesion policy, a greater focus on European regions is needed. In order to examine regional changes in poverty, well-being, and well-being, accurate regional estimates for indicators are needed and the respective statistical methodology has to be developed.

Therefore, this Deliverable first reviews different indicators used around the world to measure poverty and well-being. Thereafter, an overview of design-based, model-assisted and model-based estimation methods for such indicators is given. It is followed by a description of data sets in use for the estimation of indicators of poverty and well-being. The Deliverable furthermore gives an overview of data and methods currently in use by CBS, DESTATIS, ISTAT, and HCSO for the measurement of poverty and well-being.

1. Introduction

One of the key topics of the Europe 2020 strategy is 'Poverty and social exclusion'. The target is to reduce the number of people at risk of poverty or social exclusion by 20 million until 2020 compared to 2008 (Eurostat, 2018). Furthermore, the Millennium Development Goals focus on the reduction of poverty as well. Changes in poverty and inequality happen at low regional levels. However, poverty and inequality indicators have primarily been estimated at the national level (EU-SILC data). As a large share of the EU's budget is directed to its cohesion policy, a greater focus on European regions is needed. In order to examine regional changes in poverty and well-being, accurate regional estimates for indicators are needed and the respective statistical methodology has to be developed.

This Deliverable is structured as follows: Chapter 2 gives an overview of indicators used to measure poverty and well-being around the world. For poverty measurement it presents the indicators used by the European Union, the OECD, SAIPE, the Worldbank and the United Nations. For well-being measurement it presents multidimensional approaches and composite indicators. Chapter 3 focuses on data and methodologies. First, design-based methods for poverty and well-being indicators are presented. Thereafter, model-assisted and model-based estimation methods for SAE estimation of poverty indicators are presented. Furthermore, data sets in use for the estimation of indicators of poverty and well-being are presented. Chapter 4 shows how data and methods are used in practice at European NSIs. The methods for poverty and welfare measurement used at CBS are presented first, followed by those of DESTATIS, ISTAT, and HSCO. The report concludes with a summary in Chapter 5.

2. Indicators for the measurement of poverty and well-being

2.1. Indicators of poverty

Several international institutions aim to measure different indicators of poverty. Some of the most well-known indicators of poverty are described in the following.

2.1.1. European Union poverty indicators

The information about the European Union poverty indicators is mainly based on Eurostat (2018), European Commission (2010) and Eurostat (2017b).

The Europe 2020 strategy

On 17 June 2010 the European Commission adopted the Europe 2020 strategy as the successor to the Lisbon strategy (Eurostat, 2017b). The Europe 2020 strategy aims at enabling a smart, sustainable and inclusive future. The strategy was put forward in the light of the past economic crises in order to prepare the EU for future challenges. The European Commission (2010) defines the priorities of smart, sustainable and inclusive growth which are aimed to be achieved by 2020. The main objectives are to deliver high levels of employment, productivity and social cohesion in the Member States, while reducing the impact on the natural environment. In order to meet these objectives, eight targets are defined. The targets comprise the areas of employment, research and development, climate change and energy, education and poverty reduction. The European Commission is convinced that the different target areas are strongly interlinked and that e.g. higher educational levels lead to improved employability which leads so poverty reduction. In order to reflect specific member state situations, national targets are elaborated. Therefore, the different targets should be addressed at the same time.

The EU put forward seven flagship initiatives to support the fulfilment of the different targets. In Table 2.1 the three priorities, eight targets and seven flagship initiatives are displayed together. More detailed information about the Europe 2020 strategy and its flagship initiatives is given in European Commission (2010).

To monitor the different targets, nine headline indicators are used (Eurostat, 2018). The headline indicators and their sub indicators are displayed in Table 2.2. A more detailed description of the headline indicators is given in Eurostat (2018). In the yearly reports, the latest are Eurostat (2018, 2017b, 2016), the European Union updates the development of the headline indicators.

Beyond the cope of Europe, the Europe 2020 strategy supports the internationally adopted 2030 Agenda for Sustainable Development (SDGs) as both strategies have several aims in common (for details see Eurostat, 2018).

	Targets	Flagship initiatives
Smart growth	Increasing combined public and private investment in Research & Development to 3 % of GDP Reducing school drop-out rates to less than 10 % Increasing the share of the population aged 30-34 having completed tertiary education to at least 40 %	Innovation: Innovation Union Education: Youth on the move (ended in December 2014) Digital society: A digital agenda for Europe
Sustainable growth	Reducing greenhouse gas emissions by at least 20 % compared to 1990 levels Increasing the share of renewable energy in final energy consumption to 20 % Moving towards a 20 % increase in energy efficiency	Climate, energy and mobility: Resource efficient Europe Competitiveness: An industrial policy for the globalisation era
Inclusive growth	Increasing the employment rate of the population aged 20-64 to at least 75 % Lifting at least 20 million people out of the risk of poverty and social exclusion	Employment and skills: An agenda for new skills and jobs Fighting poverty: European platform against poverty and social exclusion

Table 2.1: The Europe 2020 strategy's key priorities, headline targets and flagship initiatives (Table 0.1, Eurostat, 2017b, slightly modified)

Topic	Headline indicator
Employment	Employment rate age group 20-64, total (% of population) • Employment rate age group 20-64, females (% of population) • Employment rate age group 20-64, males (% of population)
Research & Development	Gross domestic expenditure on R&D (% of GDP)
Climate change and energy	Greenhouse gas emissions (Index 1990 = 100) Share of renewable energy in gross final energy consumption (%) Primary energy consumption (Million tonnes of oil equivalent) Final energy consumption (Million tonnes of oil equivalent)
Education	Early leavers from education and training, total (% of population aged 18-24) • Early leavers from education and training, females (% of population aged 18-24) • Early leavers from education and training, males (% of population aged 18-24) Tertiary educational attainment, total (% of population aged 30-34) • Tertiary educational attainment, females (% of population aged 30-34) • Tertiary educational attainment, males (% of population aged 30-34)
Poverty and social exclusion	People at risk of poverty or social exclusion, EU-27 (Million people) People at risk of poverty or social exclusion, EU-28 (Million people) People at risk of poverty or social exclusion, EU-28 (% of population) • People living in households with very low work intensity, EU-28 (% of population aged 0-59) • People at risk of poverty after social transfers, EU-28 (% of population) • Severely materially deprived people, EU-28 (% of population)

Table 2.2: The Europe 2020 headline indicators (Table 0.1, Eurostat, 2018, slightly modified)

Reduction of poverty

One of the key topics of the Europe 2020 strategy is 'Poverty and social exclusion'. The target is to reduce the number of people in risk of poverty or social exclusion by 20 million in 2020 compared to 2008 (Eurostat, 2018).

Eurostat (2018) describe opposing the trends of risk of poverty or social exclusion in the EU over the past. The economic crisis led to a rise in people at risk in 2009 followed by a downward trend in 2012. 118.0 million people were affected by poverty or social exclusion in 2016 in the EU-28. Compared to 2010, there were hundred thousand more people at risk. Compared to 2015, on the other hand, there were one million people less at risk. Recently, the number of people at risk nearly approached the level observed before the economic crisis.

In Europe, poverty can be caused by different factors. Eurostat (2018) show that the most widespread form of poverty or social exclusion in the EU is monetary poverty (in 2016 about 17.3 % of EU population are at risk of poverty after social transfers) followed by low work intensity (in 10.5 % of EU population) and severe material deprivation (about 7.5 % of EU population aged 0 to 59). The changes in people at risk of poverty and social exclusion are mainly driven by the changes in the number of severely materially deprived people. The groups most likely to be affected by poverty and social exclusion are young people, unemployed and inactive persons, single parents, households consisting of only one person, people with low educational attainment, foreign citizens born outside the EU, and those residing in rural areas as described by Eurostat (2018).

European Union measurement of poverty and social exclusion

The target of 'lifting at least 20 million people out of the risk of poverty or social exclusion' by 2020 compared to 2010 is monitored by the headline indicator 'People at risk of poverty or social exclusion', see Table 2.2. It is furthermore published for several subcategories according to characteristics of sex, activity status, education, country of birth, age group, level of activity limitation, household type, children by education attainment level of their parents and degree of urbanisation (Eurostat, 2018). The indicators of the Europe 2020 strategy mainly stem from official European Social Surveys. For the estimation of poverty indicators, the EU Statistics on Income and Living Conditions (EU-SILC) is used. EU-SILC is described in European Commission (2017). The estimation of people at risk of poverty is described in Eurostat (a).

Eurostat (a) describe that the at-risk-of-poverty or social exclusion rate (AROPE) is calculated as the percentage of people over the total population to which at least one of the following characteristics applies: at-risk-of-poverty, severely deprived or living in a household with very low work intensity. It is hence necessary that at least one of the three input measures is fulfilled. The three input measures to AROPE, namely the percentage of people who are at-risk-of-poverty (ARPR), severely deprived or living in a household with very low work intensity over the total population, are described in the following.

At-risk-of-poverty
The as-risk-of-poverty rate (ARPR) is an indicator related to monetary poverty via the disposable equivalised income. Detailed information about the ARPR is given in Eurostat (d). Eurostat (d) show that the at-risk-of-poverty rate is calculated in two steps. First, the at-risk-of-poverty thresholds are calculated as a percentage of the median and mean equivalised disposable income after social transfers. The specific percentage depends on the threshold to be calculated and can be set to 40-70 %. In a second step, the ARPR is calculated as the percentage of people who are at-risk-of-poverty (calculated for the different thresholds) over the total population. For the Europe 2020 strategy, the ARPR is calculated with a 60 % threshold of the median equivalised disposable income level (Eurostat, a). The overall list of indicators related to monetary poverty is (Eurostat, d):

- At-risk-of-poverty thresholds
- At-risk-of-poverty rate
- At-risk-of poverty rate before social transfers (pensions included in social transfers)
- At-risk-of poverty rate before social transfers (pensions excluded from social transfers)
- Relative at-risk-of poverty gap
- Persistent at-risk-of poverty rate
- At-risk-of poverty rate after deducing housing costs
- Distribution of population by number of years spent in poverty within a four-year period
- At-risk-of-poverty rate anchored at a fixed moment in time At-risk-of-poverty rate for children by citizenship of their parents (population aged 0 to 17 years)
- At-risk-of-poverty rate for children by country of birth of their parents (population aged 0 to 17 years)

Material deprivation
Detailed information about material deprivation is given in Eurostat (c). Eurostat (c) show that material deprivation is measured in two steps. First, in EU-SILC it is asked whether people can afford the following material deprivation items (Eurostat, e):

- to pay their rent, mortgage or utility bills
- to keep their home adequately warm
- to face unexpected expenses
- to eat meat or proteins regularly
- to go on holiday
- a television set
- a washing machine
- a car
- a telephone

The material deprivation rate is then calculated as the percentage of people who are materially deprived, based on their inability to afford to pay no up to nine items from the list of the material deprivation items (Eurostat, c). For the Europe 2020 strategy, the severe material deprivation is used which sets the deprivation threshold to four material deprivation items (Eurostat, a). The overall list of indicators related to material deprivation is (Eurostat, c):

- Material deprivation rate - Economic strain and durables dimension
- Material deprivation rate for the 'Economic strain' dimension
- Mean number of deprivation items among the deprived - Economic strain and durables dimension
- Material deprivation rate for the 'Durables' dimension
- Material deprivation rate for the 'Economic strain' and 'Durables' dimensions
- Material deprivation rate for the 'Housing' dimension
- Material deprivation rate for the 'Environment' dimension
- Severe material deprivation rate

Work intensity

Detailed information about the work intensity threshold is given in Eurostat (b). Eurostat (b) show that the distribution of population living in household with very low work intensity is calculated as the percentage of people living in households with a certain low work intensity threshold, e.g. 20 %. For the Europe 2020 strategy, the very low work intensity threshold is 20 % (Eurostat, a). The overall list of indicators related to the subject area of health and labour conditions are (Eurostat, b):

- Distribution of population aged 18 and over along with several combination of dimensions
- People living in households with very low work intensity
- Labour transitions by labour status
- Labour transitions by type of contract
- Labour transitions by type of contract - changes in employment security
- Labour transitions by pay level
- Labour transitions by employment status and pay level - changes in qualifications

2.1.2. OECD poverty indicators

The OECD applies both an absolute and a relative poverty line. The absolute poverty line is set at 50% of the median income in 2005. The relative poverty line is a given percentage of the median disposable income expressed in nominal terms. Two relative poverty lines are applied: one at 50% of median equivalised disposable income and the other at 60% of median equivalised disposable income. Equivalised disposable income, DI_{ij}, is calculated as $DI_{ij} = \frac{Y_i}{S_i^{\epsilon}}$, with Y_i denoting total disposable income of household i, S_i the number of members in household i and ϵ the equivalence elasticity. When calculating the poverty rate for various population sub-groups, the entire population serves as reference population. The poverty threshold is determined based on the population as a whole (OECD, 2017b). The basis for benchmarking and analysing poverty across countries is the OECD Income Distribution Database (OECD, 2019).

2.1.3. SAIPE poverty indicators

The U.S. Census Bureau's Small Area Income and Poverty Estimates (SAIPE) program provides annual estimates with respect to income and poverty for all US states, counties and school districts. SAIPE produces five county and state estimates and three estimates for school districts. At county and state level they provide the following estimates:

- Absolute number of people living in poverty
- Number of children under 5 years living in poverty (published for states only)
- Number of families with children aged between 5 and 17 living in poverty
- Number of children under 18 years living in poverty
- Median household income

According to Title I of the Elementary and Secondary Education Act, SAIPE produce the following estimates school district (U.S. Census Bureau, 2018b):

- Absolute number of people living in poverty
- Number of children between 5 and 17 years living in poverty
- Number of families with children aged between 5 and 17 living in poverty

In order to obtain county and state estimates, SAIPE sets up a regression model that relates survey data from the American Community Survey to census data and administrative records. The regression predictions are combined with direct estimates using Bayesian estimation techniques (U.S. Census Bureau, 2018a). In order to obtain school district estimates, the SAIPE county-level estimates of poverty are combined with inputs from federal tax information and multi-year American Community Survey estimates (U.S. Census Bureau, 2019).

2.1.4. Worldbank poverty indicators

The World Bank applies an absolute international poverty line of US \$1.90 per person per day in 2011 purchasing power parity to monitor extreme poverty. The level of the international poverty line is determined by the average of the national poverty lines from the poorest 15 countries in 2005. The underlying idea is that if basic needs can met with US \$1.90 in the poorest countries in the world, it can be applied as minimum threshold to all countries. The value of the poverty line is regularly adjusted to price fluctuations. Furthermore, the World Bank introduces two additional sets of poverty lines, the middle income class poverty lines and the societal poverty line. The first set of complementary poverty lines accounts for the fact that the international poverty line might be too low for many countries or that the basic needs have increased. The lower middle income class poverty line and the upper middle income class poverty line are determined by the median values of the respective national poverty lines in 2011. The lower middle income class poverty line is set at US \$3.20 and the upper middle income class poverty line at US \$5.50 person per day in 2011 purchasing power parity. The second set of complementary poverty lines is the societal poverty line. It accounts for the fact that participation in society becomes costlier as countries become richer and is defined as US \$1 per day plus half of the daily median income in that country or the international poverty line, whichever is greater. The societal poverty line is calculated in 2011 purchasing power parity as follows: $SPL = \max \text{US\$1.90}, \text{US\$1.00} \times$ median consumption. Moreover, the Word Bank has taken the inital step towards a multidimensional global poverty measure that goes beyond the monetary dimension. In order to account for potential differences within a household the World Bank promotes individual-data instead of household data. If the whole household were seen as poor because one individual in the household is poor, these differences would be neglected (World Bank, 2018a).

The small-area method of the World Bank was developed by Elbers et al. (2003). They build on small-area literature and combine measures consumption from household surveys and Census data in order to estimate unit level consumption and map poverty.

2.1.5. United Nations poverty indicators

The United Nation Development Program (UNDP) understands poverty in a multidimensional approach which is closely connected with the concept of human development. Within the corresponding nations the concept of poverty is advanced by improving the quality of life of citizens. This multidimensional approach acknowledges that poverty can not only be defined as the lack of material well-being but should also include the lack of opportunities and choices which are part of human development. (cf. UNDP, 1997, p. 5 and Hill and Adrangi, 1999, p. 137)

Two important multidimensional measures of poverty and human development established by the United Nations are the Human Development Index (HDI) and the Multidimensional Poverty Index (MPI). These approaches are based on three dimensions measured with several indicators on different aspects of poverty. Furthermore, in 2016 the Agenda 2030 with 17 goals and 169 associated indicators was implemented. With the Sustainable Development Goal (SDG) 1 the aim to end all form of poverty is expressed. Each of the five targets of SDG 1 are based on measures related to poverty. The HDI, MPI and the SDG 1 are explained in the following.

The Human Development Index

The human development index (HDI) was already conceptualized in 1990 with the aim to measure the national development not only by income per capital but also with indicators informing about health and educational achievements. (cf. UNDP, 2018a, p. 1) This, three dimensional composite indicator was used to rank countries based on their human development and foster the discussion surrounding policy incentives for human development. Based on the HDI other composite indicators have been developed which aim to reflect the development in certain groups such as the Gender Inequality Index (GII). The estimated indicators for each observed country as well as analysis of the development of single components are published each year in the Human Development Report by the UNDP.

In figure 2.1, the three dimensions and corresponding indicators are represented. The ability to lead a long and healthy life is measured with the life expectancy at birth, the ability to acquire knowledge is measured with the mean years of schooling and the expected years of schooling. The third aspect, the ability to achieve a decent standard of living is measured by the gross national income per capita. (cf. UNDP, 2018a, p. 1)

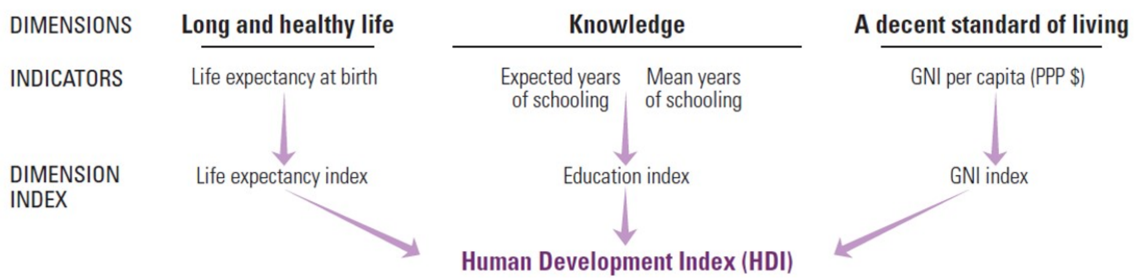

Figure 2.1: The Human Development Index (UNDP, 2018a, p. 1)

The four individual indicators as components of the HDI are estimated by the Human Development Report Office (HDRO) and the data origins from international data agencies which have the mandate, resources and the expertise to collect the national data needed to estimate the indicators. Examples of some of the well-respected international data providers are the Centre for Research on the Epidemiology of Disasters, Economic Commission for Latin America and the Caribbean, Eurostat, Food and Agriculture Organization, Gallup and many more. Furthermore, the UNDP provides information about the national trends of the HDI, for example, by publishing interpolated consistent datasets on the HDI indicators. This is necessary because national and international agencies continually improve the data series and hence, published HDIs from different years with different data sources may not be comparable without adaptation. (cf. UNDP, 2018a, p. 17f.) Discrepancies between the national and the international data is also possible because international agencies often harmonize the country data for country comparisons or produce estimates for missing data on the indicators; for example; with nowcasting or cross country regression methods. Another problem can be that international agencies do not have access to the most up to date data from the national agencies. (cf. UNDP, 2018a, p. 17 and UNDP, 2018b, p. 3)

For the categorisation of the evaluated country performances on the HDI, a classification framework was developed based on fixed points of possible HDI values. HDI value less than 0.550 mark low human development, HDI values between 0.550 and 0.669 mark medium human development. HDI values between 0.700 and 0.799 mark high human development and HDI values greater than 0.800 show very high human development. (cf. UNDP, 2018a, p. 17 and UNDP, 2018b, p. 3)

Individual indicators of the HDI

In the following the individual indicators of the HDI and the respective data sources based on UNDP (2018a) are listed in more detail. (cf. UNDP, 2018a, p. 25 and UNDP, 2018b, p. 2)

- **Life expectancy at birth**
  Calculated as the number of years a new-born infant is expected to live if prevailing patterns of age-specific mortality rates at the time of birth stay the same throughout the infant's life
  Source: UN Department of Economic and Social Affairs (2017).

- **Expected years of schooling**
  Calculated as the number of years of schooling that a child of school entrance age is expected to receive if prevailing patterns of age-specific enrolment rates persist throughout the child's life.
  Source: UNESCO Institute for Statistics (2018), OECD (2017a), ICF Macro Demographic and Health Surveys, United Nations Children's Fund (UNICEF) Multiple Indicator Cluster Surveys.

- **Mean years of schooling**
  Calculated as the average number of years of education received by people aged 25 and older, converted from education attainment levels using official durations of each level.
  Source: UNESCO Institute for Statistics (2018), Barro and Lee (2016), OECD (2017a), ICF Macro Demographic and Health Surveys, UNICEF Multiple Indicator Cluster Surveys.

- **Gross national income (GNI) per capita**
Calculated as aggregate income of an economy generated by its production and its ownership of factors of production, less the incomes paid for the use of factors of production from the rest of the world. The measure is converted to international dollars using PPP rates, divided by the number of midyear population.
Source: World Bank (2018b), International Monetary Fund (2018), UN Statistics Division (2018).

Calculation of the HDI
With the data on the four individual indicators the HDI can be calculated in two steps. Minimum and maximum values are set as goalposts such that the indicator values can be expressed in a range between 0 and 1. Using the minimum and maximum values given in table 2.3 and formula 2.1 the individual indicator values are standardised. (cf. UNDP, 2018b, p. 2)

Table 2.3: Values for the standardisation of the individual indicators of the HDI

Dimension	Indicator	Minimum	Maximum
Health	Life expectancy (years)	20	85
Education	Expected years of schooling (years)	0	18
	Mean years of schooling (years)	0	15
Standard of living	Gross national income per capita (2011 PPP $)	100	75,000

For further information on the choice of the minimum and maximum values see UNDP (2018b). The standardised individual indicator values $x^*_{c,i}$ are calculated as

$$x^*_{c,i} = \frac{x_{c,i} - x_{\min,i}}{x_{\max,i} - x_{\min,i}} \tag{2.1}$$

using the corresponding values $x_{\min,i}$ and $x_{\max,i}$ for each i in table 2.3. After the standardisation of the individual indicator values the three dimension indices can be calculated. The dimension indices I_{Health} and I_{Income} for the health and standard of living dimension are equal to the standardised values x^* since they are measured with one variable each. For the dimension indices I_{Health} the standardised values of each individual indicator are averaged with the arithmetic mean for each corresponding country. Following this, the second step of calculating the HDI entails the aggregation of the dimensional indices as geometric mean (UNDP, 2018a, p. 2)

$$HDI = (I_{\text{Health}} \cdot I_{\text{Education}} \cdot I_{\text{Income}})^{1/3}. \tag{2.2}$$

The multidimensional poverty index
Another indicator which is reported in Human Development report in the light of measuring multidimensional poverty as a deprivation index, is the Multidimensional Poverty Index. It was developed

by UNDP, HDRO and the Oxford Poverty and Human Development Initiative (OPHI). The MPI is an internationally comparable measure of acute poverty and is intended to capture multiple deprivations poor people can experience. Similar to the HDI, the dimensions of this measure are health, education and living standard, but the MPI makes use of different individual indicators to measure these dimensions. The MPI was first introduced in 2010 and relaunched with new features in 2014. In 2018 a revised MPI was developed as a joint version of the 2010 and 2014 version to better align with the Sustainable Development Goals. (cf. Alkire and Jahan, 2018, p. 1ff.) The data which is used to calculate the MPI is collected mainly from two surveys, the Demographic and Health Survey (DHS) and the Multiple Indicators Cluster Surveys (MICS). For some countries national surveys with similar content and questionnaires are used such as the Pan Arab Population and Family Health Survey (PAPFAM) or national surveys for Brazil, China or Mexico. (cf. Alkire and Jahan, 2018, p. 7)

For the calculation of the MPI ten individual indicators are used to measure three dimensions. The ten individual indicators identify each observed individual as deprived based on the joint achievements in the indicators of the household members. Since the MPI is calculated as a deprivation index, it has to be decided for each observed individual and each indicator if the individual is deprived in that corresponding indicator or not. The rules for deciding whether an individual is deprived or not is explained in more detail below. In order to analyse multidimensional poverty a cross-dimensional poverty cut-off of $z = 1/3$ is applied. A person is identified as poor if the weighted deprivation sum is equal or exceeds this poverty cut-off. The weights for each indicator are determined due to the construction of the MPI and are also mentioned below for each individual indicator. Two other cut-offs are frequently applied for the calculation of the MPI, the severe poverty cut-off of 1/2 and the vulnerability cut-off which identifies the individuals deprived in 20 % to 33 % of the individual indicators. Finally, after classifying each individual as multidimensionally poor or not poor, depending on the cut-off value, the MPI for a country is calculated as the adjusted head count ratio. It is expressed as the percentage of individuals which are identified as multidimensionally poor and adjusted by the average share of deprivations among the multidimensionally poor individuals as a measure for the intensity. (cf. Alkire and Jahan, 2018, p. 8 and Alkire and Santos, 2014, p. 9ff.)

Individual indicators of the MPI

The individual indicators of the MPI are listed with their corresponding weights in table 2.4. In the following, it is explained in more detail when an individual is considered as deprived in each of the individual indicators. These explanations are based on Alkire and Jahan (2018).

Table 2.4: Values for the standardisation of the individual indicators of the HDI

Dimension	Indicator	Weight
Health	Nutrition	1/6
	Child mortality	1/6
Education	Years of schooling	1/6
	School attendance	1/6
Living standards	Cooking fuel	1/18
	Sanitation	1/18
	Drinking water	1/18
	Electricity	1/18
	Housing	1/18
	Assets	1/18

- **Nutrition**
  An individual is considered to be deprived, if for any individual under 70 years of age in the household the nutritional information shows undernourishment. Adults between 20 and 70 years are considered to be undernourished if their Body Mass Index (BMI) is below 18.5. Individuals between 5 and 20 years of age are considered to be undernourished if their age-specific BMI is below the age specific BMI cut-off minus two standard deviations. The age specific BMI cut-off values are based on the recommendations of the World Health Organisation (WHO) Expert Committee on Physical Status in World Health Organization Expert Committee on Physical Status (1996). Individuals under 5 years of age are considered to be undernourished if their z-score for height-for-age or weight-for-age is below the value of the median minus two standard deviations of the reference population.

- **Child mortality**
  An individual is deprived if any child in the household has died within the family in the last five years of the corresponding survey from which the information is taken. If the data of child's death is not known the number of all reported death is used instead.

- **Years of schooling**
  The individual is considered to be deprived in this individual indicator if no household member with ten years of age or older has completed six years of school education.

- **School attendance**
  If any school-age child of the household is not attending school up to the age at which the child should complete class two, the corresponding individual is considered to be deprived in this individual indicator.

- **Cooking fuel**
  In case the household is using dung, wood, charcoal or coal to cook the individual of this household is considered to be deprived in this individual indicator.

- **Sanitation**
  The individual is considered to be deprived, if the household has no access or has to share an improved sanitation which is defined as some type of flushed toilet or latrine, ventilated improved pit or composting toilet.

- **Drinking water**
  If the individual lives in a household which has no access improved drinking water or save drinking water within 30-minutes walking distance from home. Here, piped water, public tap, borehole or pump, protected well, protected spring water or protected rain water is considered as improved drinking water.

- **Electricity**
  An individual is considered to be deprived if it lives in an household without electricity.

- **Housing**
  If at least one of the materials for roof, walls and floor are of rudimentary or natural materials the individuals living in this household are considered to be deprived. This applies to floors made of mud, clay, earth, sand or dung or dwellings without any roof or walls. For the roofs and walls this applies if their are made of cane, palm, trunks, sod, mud, dirt, grass, reeds, thatch, bamboo, sticks, carton, plastic, polythene sheeting, loosely packed stones, uncovered adobe, wood, plywood, cardboard, unburnt brick, canvas or tent.

- **Assets**
  The individual is considered to be deprived if the corresponding household does not own more than one of the following assets: TV, telephone, computer, animal cart, bicycle, motorbike or refrigerator, and does not own a car or truck.

The sustainable development goal 1: No poverty

The measurement of poverty is, besides others concepts, also embedded in the Agenda 2030 for Sustainable Development, which was implemented in 2016 as a new agenda and global priorities towards well-being of current and future generations. 17 integrated SDGs and 169 associated targets have been developed to support a joint effort on public, private, domestic and international level to achieve sustainable development. (cf. OECD, 2016, p. 3)

SDG 1: *End poverty in all its forms everywhere* summarises five targets which are each measured with a specified indicator on the progress of human development. A List with the targets and the corresponding indicators from UNDP (2018a) is given below in table 2.5. Some of the HDI and MPI indicators are also suitable for the SDG 1. (cf. UNDP, 2018a, p. 18ff.)

Table 2.5: Targets and indicators of SDG 1: End poverty in all its forms everywhere

Targets	Indicator
1.1 By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day	Population living below income poverty line, PPP $1.90 a day (%)
1.2 By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions	Population living below income poverty line, national poverty line (%) Population in multidimensional poverty, headcount (%) Population in multidimensional poverty, headcount (thousands)
1.3 Implement nationally appropriate social protection systems and measures for all, including floors, and by 2030 achieve substantial coverage of the poor and the vulnerable	Old-age pension recipients (% of statutory pension age population) Old-age pension recipients (female to male ratio) Mandatory paid maternity leave (days)
1.5 By 2030, build the resilience of the poor and those in vulnerable situations and reduce their exposure and vulnerability to climate-related extreme events and other economic, social and environmental shocks and disasters	Internally displaced persons (thousands) Refugees by country of origin (thousands) Homeless people due to natural disaster (average annual per million people)
1.a Ensure significant mobilization of resources from a variety of sources, including through enhanced development cooperation, in order to provide adequate and predictable means for developing countries, in particular least developed countries, to implement programmes and policies to end poverty in all its dimensions	Government expenditure on education (% of GDP)

2.2. Indicators of well-being

As the main focus of this deliverable is on the regional measurement of poverty, we only briefly touch on the related concept of well-being. For a more thorough treatment of well-being and the sustainable development goals as well as the implementation of these concepts in the context of European states, please refer to Deliverable 1.1 (Tinto et al., 2018) and Deliverable 1.2 (Tinto and Baldazzi, 2018) of the MAKSWELL project, dealing with the international and national experiences and main insights and the related existing related databases, respectively.

The measurement of well-being has been limited for decades to the use of the Gross Domestic Product (GDP), under the assumption that the wealthier a country, the higher the well-being of its citizens. Since 2009, the 'Beyond GDP' initiative underlined that aspects essential to good quality of life such as health, social relations and personal security are not measured in GDP. The Commission on the 'Measurement of Economic Performance and Social Progress' published a report with 12 recommendations on how to better measure economic performance, societal well-being and sustainability (Stiglitz et al., 2009). As concerns the measurement of EU citizens' well-being, relevant recommendations are:

- Recommendation 6: Quality of life depends on people's objective conditions and capabilities. Steps should be taken to improve measures of people's health, education, personal activities and environmental conditions;
- Recommendation 7: Quality of life indicators in all the dimensions covered should assess inequalities in a comprehensive way. Inequalities in quality of life should be assessed across people, socio-economic groups, gender and generations, with special attention to inequalities that have arisen more recently, such as those linked to immigration;
- Recommendation 8: Surveys should be designed to assess the links between various quality of life domains for each person, and this information should be used when designing policies in various fields;
- Recommendation 10: Measures of both objective and subjective well-being provide key information about people's quality of life. Statistical offices should incorporate questions to capture people's life evaluations, hedonic experiences and priorities in their own survey;
- Recommendation 12: The environmental aspects of sustainability deserve a separate follow-up based on a well-chosen set of physical indicators.

As a consequence, the measurement of individual well-being in EU countries witnessed a significant revolution. Individual well-being has been conceived as a complex and multidimensional concept that need to be measured to illustrate people's quality of life, giving a picture of people's satisfaction in relation to their jobs, family life, health conditions, and standards of living. From a methodological point of view, traditional measurements of well-being based on single indicators have been often replaced by multidimensional approaches. These new methods aim to embrace the phenomena under study in a wider way, mainly by taking into consideration a large number of dimensions that can be interpreted either as the result of the latent concepts that need to be analysed, or as their cause, depending on the theoretical model chosen (Bollen and Bauldry, 2011).

In particular, multidimensional measures of quality of life has been defined, covering different aspects such as:

- Objective measures of well-being, complementing economic aspects of quality of life (i.e. poverty indicators) with indicators concerning leisure time, educational level, social connections;
- Subjective measures of well-being, based on self-reporting by individuals, complementing objective indicators with direct measures of complex components of individual well-being such as life satisfaction;
- Measurement of the indicators among relevant subgroups of the population, to understand inequality in well-being levels within countries;
- Sustainability well-being measures, trying to capture the sustainability of socio-economic and environmental systems to ensure citizens' well-being to last over time.

Important examples are the OECD multidimensional framework for well-being indicators and the quality of life indicators currently used by Eurostat. We present these two approaches in the next subsection.

Under the multidimensional conceptual framework for the measurement of individual well-being, there has been a growing reference to the capability approach (CA) as an alternative normative framework for the evaluation of human development, well-being and freedom by thinking in terms of human functionings and capabilities. Functionings are "features of the state of an existence of a person" (Hawthorn 1987) while capabilities represent what people are able to do or to be (Sen 1999, p. 18). According to Dean (2009, p. 262), capabilities represent the essential fulcrum between material resources (commodities) and human achievements. Thus, the CA focuses on measuring the well-being of adults whose freedom to choose a life they have reason to value is central to the notion of capabilities (Klasen 2001). In this way the CA engages with both objective and subjective perceptions of well-being. In the next subsections we refer to some research work that has recently engaged with the CA for defining living condition indicators across EU countries.

A last important issue in the measurement of well-being is data availability. To cover all the important aspects of the multidimensional definition of individuals' well-being, indicators computed using sample surveys data are of basic importance. In particular, the European Union statistics on Income and Living Conditions (EU-SILC) survey, is the core survey that allows the computation of well-being indicators across the EU countries in a comparative way. It covers important aspects at household and individual level such as income, housing, material deprivation, education, health, labour. Moreover, in 2013 the 'ad-hoc' module of the survey was designed to measure specific aspects of subjective well-being. Other 'ad-hoc' modules have also being designed to deepen the measurement of specific aspects already included in the main questionnaire of the survey. Another relevant survey data sources used to complement the coverage of the different dimensions of quality of life include the Labour Force Survey (LFS). More details on the data sources currently in use to measure well-being indicators will be presented in chapter 3.

In the following subsections we present some of the main important indicators that are currently used to measure the dimensions of well-being in EU countries, with specific reference to the corresponding data source. We also present some relevant research works that have dealt with important methodological issues related to the definition of multidimensional and composite well-being measures.

2.2.1. Multidimensional approaches to measurement of well-being

Building on the Stiglitz-Sen-Fitoussi Commission recommendations, the European Statistical System's Sponsorship Group on Measuring Progress, Well-being and Sustainable Development endorsed a framework encompassing nine dimensions to measure Quality of Life Eurostat (2017a). Eight dimensions concern the functional capabilities citizens should have available to effectively pursue their self-defined well-being, while the last dimension refers to the personal achievement of life satisfaction and well-being. The first eight dimensions are: Material living conditions; Productive or other main activity; Health; Education; Leisure and social interactions; Economic security and personal safety; Governance and basic rights; Natural and living environment. The last dimension, the subjective one, is Overall experience of life.

The following Tables report the objective indicators that are currently in use to measure the first eight dimensions. The last dimension, Overall experience of life, is measured only using subjective indicators: the full list is reported in a separate Table.

Another important example is the OECD framework for well-being indicators OECD (2011). This framework distinguishes between current material living conditions and quality of life, the 'Current Well-being' domain, and the conditions required to ensure their sustainability over time, the 'Resources for Future Well-being' domain. These two main domains are further divided into sub-domains and dimensions.

For the 'Current Well-being' domain OECD defines two sub-domain. The first one is 'Quality of Life', with the following corresponding dimensions: Housing; Income and wealth; Jobs and earnings; Social connections; Education and skills; Environmental quality; Civic engagement and governance; Health status; Subjective well-being; Personal security; Work-life balance. The second sub-domain, 'Material Living Conditions', includes the following dimensions: Income and wealth; Jobs and earnings; Housing.

The dimensions defining the 'Resources for Future Well-being' domain include four different "capital stocks": Natural capital; Economic capital; Human capital; Social capital.

There is still some lacking in the availability of indicators to measure the 'Resources for Future Well-being' domain. In the volume 'How's life 2017' OECD refers to the following indicators that can currently be monitored over time. For the natural capital: Forest area (square kilometres/thousand people); greenhouse gas emission (tonnes per capita CO_2); CO_2 emissions (tonnes per capita). For the human capital, besides the indicators already in use for the Current Well-being domain, the use of indicators monitoring smoking and obesity are suggested as human capital risk factors. The Economic capital is instead measured through: Produced fixed assets; Intellectual property assets; Investment in R&D; Financial worth of the economy; Household net wealth; Gross fixed capital formation; Financial worth of the general government; Banking sector leverage; Household debt. Finally, the chosen Social capital indicators are: Voter turnout; Trust in government.

As we can see, the Quality of Life dimensions defined by Eurostat and the OECD Well-being framework have many overlaps. Both approaches conceive Quality of Life and Well-being as multidimensional concepts covering the Stiglitz-Sen-Fitoussi Commission recommendations. In both cases objective indicators are complemented by subjective indicators as Life Satisfaction is an important topic under both approaches.

A third important source for multidimensional well-being measurement in the EU, using again both objective and subjective aspects, is the European Quality of Life Survey (EQLS). Over the years, Eurofond EQLS has developed into a valuable set of indicators which complements traditional indicators of economic growth and living standard such as GDP or income. The EQLS, carried out every four years, examines both the objective circumstances of European citizens' lives and how they feel about those circumstances and their lives in general. It considers several domains, such as employment, income, education, housing, family, health and work-life balance. It also consider at subjective topics, such as people's levels of happiness, how satisfied they are with their lives, and how they perceive the quality of their societies. Being repeated the survey every four years, the EQLS makes it also possible to track key trends in the quality of people's lives over time in the EU countries. The 2016 EQLS provided detailed information in three main areas:

- Quality of life: subjective well-being, optimism, health, standard of living and aspects of deprivation, work-life balance
- Quality of society: social insecurity, perception of social exclusion and societal tensions, trust in people and institutions, participation and community engagement, and involvement in training/life-long learning
- Quality of public services: health-care, long-term care, childcare and other public services

As we can see, subjective well-being represent an important topic also under the EQLS (European Commission, 2013). In particular, the EQLS and the EU-SILC 2013 ad-hoc module allow to investigate the aspects of overall well-being measures. This allows to investigate the contribution of the single aspects on the overall well-being using person-level data, also controlling for important personal characteristics, by using suitable methodologies. For example, D'Agostino et al. (2019) used a Structural Equation Modelling approach to investigate the determinants of subjective well-being of young adults in Europe using EU-SILC 2013 data.

2.2.2. Composite indicators of well-being

The two approaches highlighted above rely on the use of a dashboard of indicators that are used to measure each relevant dimension. In some cases, studies that are theoretically grounded on a multidimensional background rely on the use of a composite index (CI), that acts as operational tool aiming to reduce the dimensionality of the data (OECD, 2008). As multidimensional approaches aim to extend old frameworks that based their measurement on a single dimension, it might seem contradictory that the result of such a significant theoretical shift is still an unidimensional value. One of the main reason behind this apparent paradox is the fact that unidimensional values are by far more appealing when it comes to make a comparison between different units or across time, for example for monitoring purposes.

An example is OECD's Better Life Index, which allows for a comparison of well-being across countries. It is based on the 11 Quality of Life key dimensions listed above. Each dimension is evaluated based on one to four statistical indicators, which are assigned equal weight. Scores for each aspect can then be integrated into an overall value: the default is to treat all aspects equally, but users of the interactive website www.oecdbetterlifeindex.org can give greater importance to certain dimensions so as to come up with their own overall country values and rankings OECD (2011).

It is important to underline that definition of CIs, and thus the shift to new multidimensional paradigms for the measurement of well-being, with their growing complexity, introduced many methodological issues that unidimensional approaches did not have to deal with, like the selection of the dimensions that define the underlying concepts, the choice of the theoretical framework or the way the various dimensions interact with each other (Burgass et al., 2017, Cheli and Lemmi, 1995, Mauro et al., 2018, Mazziotta and Pareto, 2016).

In particular, Mazziotta and Pareto (2016) suggested a methodology to define composite indicators as an alternative to the use of the simple arithmetic or geometric mean to combine the value of the indicators referring to the single dimensions. Simple aggregating functions like a simple arithmetic mean assume a full substitutability among the components of the index: a deficit in one dimension can be compensated by a surplus in another. However, this can be considered a strong assumption: Mazziotta and Pareto (2016) introduced their non-compensatory composite index methodology. The authors presented an application to some of the OECD well-being dimensions introduced in the previous subsection. Bacchini et al., 2019 illustrates some drawbacks in the use of this approach both looking at the desirable properties of the composite index and at the relationship amid the unbalance adjustment and the time evolution. The results presented suggest that the current state of the art in composite indices claims for an agenda where the interplay between normalization, aggregation and time dimension is correctly addressed.

Mauro et al. (2018) proposed the Multidimensional Synthesis Indicator (MSI), a new approach for the synthesis and analysis of multidimensional poverty and well-being indicators. The author's perspective was inspired by the theoretical foundations of the Capability Approach and sustainable human development paradigm. The main contribution of the approach is that the degree of substitutability between the dimensions of the multidimensional indicator can be directly linked to the general level of well-being of a person.

Under a different approach, Betti (2017) suggested the use of the fuzzy set methodology firstly introduce by Cheli and Lemmi (1995) to define an overall quality of life indicator using EQLS data. The fuzzy set methodology was born on the assumption that poverty is a multidimensional phenomenon and a vague predicate that manifests itself in different shades and degrees (fuzzy concept) rather than an attribute that is simply present or absent for individuals in the population, as the traditional poverty approach using only the At-Risk-Of-Poverty rate assumes. Betti (2017) describes the statistical methodology applied to use EQLS data for the construction of fuzzy multidimensional indicators of Quality of Life, and the use of fuzzy intersection and union operators to aggregate several composite indicators of Quality of Life simultaneously. This allows the author to perform an overall comparison of Quality of life among different European countries.

The use of the fuzzy set methodology under the Capability Approach framework has been suggested by Potsi et al. (2016) for measuring the deprivation (defined as lack of well-being) of children. In this case the authors used EU-SILC data to define a fuzzy monetary and a fuzzy non-monetary measures for children capability deprivation in Italy. D'Agostino et al. (2018) extended the framework to the measurement and comparison of children well-being in four Mediterranean Countries: Portugal, Italy, Greece and Spain.

Table 2.6: Eurostat objective Quality of Life indicators by dimension, topic and sub-topic. Dimensions: Material living conditions, productive or other main activity, health.

	Material living conditions	
Topic	Sub-topic	Indicator
Income		Median disposable equivalised income
		Income inequality (S80/S20 income quintile ratio)
	At-risk-of-poverty rate	
		At-risk-of-poverty rate anchored at fixed moment in time
Consumption	Consumption	Actual individual consumption (per capita)
	Constrained consumption	Basic expenses in the total household expenditure
Material conditions	Material deprivation	Severe material deprivation rate
	(In)ability to make ends meet	
	Housing conditions	Structural problems of the dwelling
		Space in the dwelling (overcrowding/under-occupation)
	Productive or other main activity	
Topic	Sub-topic	Indicator
Quantity of employment	Employment and unemployment	Employment rate
		Unemployment rate
		Long-term unemployment rate
	Underemployment	People living in households with very low work intensity
		Underemployed part-time workers
Quality of employment	Income and benefits from employment	Low-wage earners
	Temporary work	Temporary contracts
		Involuntary temporary contracts
	Over-qualification	Over-qualification rate
	Health and safety at work	Incidence rate of fatal accidents at work
	Work/life balance	Average number of usual weekly hours of work
		Long working hours (more than 48 per week)
		Atypical working hours (usual work during evenings, nights, Saturdays or Sundays)
		Flexibility of the work schedule
Main reason for economic inactivity	Inactive population	Inactive population by reason of inactivity
	Health	
Topic	Sub-topic	Indicator
Outcomes	Life expectancy	Life expectancy at birth
	Health status	Healthy Life Years
Determinants		Body Mass Index
		Daily smokers
		Hazardous alcohol consumption
		Practice of physical activity
		Consumption of fruits and vegetables

Table 2.7: Eurostat objective Quality of Life indicators by dimension, topic and sub-topic. Dimensions: Education, leisure and social interactions, economic security and physical safety, governance and basic rights, natural and living environment.

	Education	
Topic	Sub-topic	Indicator
Competences and skills	Educational attainment	Educational attainment
		Early leavers from education and training
	Assessed skills	Mean literacy proficiency score
Lifelong learning	Participation in adult education and training	
Opportunities for education		Participation in education of children four-year-olds
	Leisure and social interactions	
Topic	Sub-topic	Indicator
Leisure	Quantity of leisure	Non-participation in culture or sport activities
	Access to leisure	Financial obstacles to leisure participation
Social interactions	Relations with people	Frequency of getting together with friends
	Activities for people	Participation in formal voluntary activities
		Participation in informal voluntary activities
	Social cohesion	Trust in others
		Perception of social inclusion
	Economic security and physical safety	
Topic	Sub-topic	Indicator
Economic security	Wealth (assets)	Population unable to face unexpected financial expenses
	Debt	Population in arrears
	Income insecurity	Perc. of employed in the previous year transitioning to unempl.
Physical safety	Crime	Homicide rate
	Governance and basic rights	
Topic	Sub-topic	Indicator
Discrimination and equal opportunities	Equal opportunities	Gender employment rate gap
		Gender pay gap
		Gap in employment rates between nationals and non-EU citizens
Active citizenship		Active citizenship
	Natural and living environment	
Topic	Sub-topic	Indicator
Pollution (including noise)		Urban pop. exposure to air pollution (PM10)

Table 2.8: Eurostat subjective Quality of Life indicators by dimension, topic and sub-topic.

	Material living conditions	
Topic	Sub-topic	Indicator
Material conditions	Housing conditions	Satisfaction with accommodation
	Productive or other main activity	
Topic	Sub-topic	Indicator
Quality of employment	Over-qualification	Self-reported over-qualification
Quality of employment	Work/life balance	Satisfaction with commuting time
	Assessment of the job quality	Job satisfaction
	Health	
Topic	Sub-topic	Indicator
Outcomes	Health status	Self-perceived health
		Self-reported mental health
Access to healthcare		Unmet needs for medical care
	Education	
Topic	Sub-topic	Indicator
Competences and skills	Self-reported skills	Individuals' level of internet (digital) skills
		Population reporting not to know any foreign language
		Level of best known foreign language
	Leisure and social interactions	
Topic	Sub-topic	Indicator
Leisure	Quantity of leisure	Satisfaction with time use
Social interactions	Relations with people	Satisfaction with personal relationships
	Social support	Help from others (having someone to rely on in case of need)
		Having someone to discuss personal matters with
	Economic security and physical safety	
Topic	Sub-topic	Indicator
Physical safety	Crime	Perception of crime, violence or vandalism in the living area
	Perception of physical safety	Safety feeling (pop. feeling safe when walking alone after dark)
	Governance and basic rights	
Topic	Sub-topic	Indicator
Trust in institutions		Trust in the legal system, the political system and the police
	Natural and living environment	
Topic	Sub-topic	Indicator
Pollution (including noise)		Perception of pollution, grime or other environmental problems
		Noise from neighbors or from the street
Access to green and recreational spaces		Satisfaction with recreational and green areas
Landscape and built environment		Satisfaction with living environment
	Overall experience of life	
Topic	Sub-topic	Indicator
Life satisfaction		Overall life satisfaction
Affects		Negative affects (being very nervous; feeling down in the dumps; etc.)
		Positive affects (being happy)
Meaning and purpose of life		Assessing whether life is worthwhile

Table 2.9: OECD well-being indicators: Current Well-being.

Material living conditions	
Dimension	Indicator
Income and wealth	Household net adjusted disposable income per person
	Household financial net wealth per person
Jobs and earnings	Employment rate
	Long-term unemployment rate
Housing	Number of rooms per person
	Dwelling with basic facilities
	Quality of life
Dimension	Indicator
Health status	Life-expectancy at birth
	Self-reported health status
Work and Life	Employees working very long hours
	Time devoted to leisure and personal care
	Employment rate of women with children
Education and Skills	Educational attainment
	Students' cognitive skills
Social connections	Contacts with others
	Social network support
Civic engagement and Governance	Voter Turn-out
	Consultation on rule-making
Environmental Quality	Air pollution
Personal Security	Intentional homicides
	Self-reported victimisation
Subjective Well-being	Life-satisfaction

3. Overview of methodologies and data

3.1. Design-based methods for indicators

For regional measurements of poverty and well-being, we first present the HT estimator as an estimator for classical parameters of interest such as a mean values. Furthermore, we introduce the concept of domain estimation. Thereafter, we describe some of the most well-known Leaken indicators capable of measuring poverty and well-being, namely the at-risk-of-poverty rate, the quintile share ratio and the Gini coefficient. For further indicators of poverty and social exclusion see Eurostat (2009), Graf et al. (2011).

A comprehensive overview of design-based estimation methods can be found in Särndal et al. (1992) which serves as the main source for the following estimator descriptions in Section 3.1.1 and 3.1.3. A overview about the at-risk-of-poverty rate, the quintile share ratio and the Gini coefficient is given in Eurostat (2009). For estimation of these indicators using EU-SILC data see Alfons and Templ (2013).

Let there be a population $U = \{1, \ldots, j, \ldots, N\}$ consisting of N elements labelled $j = 1, \ldots, N$. We draw a sample $s \subseteq U$ of size n from this population following a specified sampling design. In accordance with the sampling design, we can allocate a first order inclusion probability $\pi_j = P(i \in S)$ to each element of the population indicating the probability of being drawn in a sample s under the specified sampling design. The inverse of the inclusion probability, $w_j = \pi_j^{-1}$ is refereed to as the design weight. The second order inclusion probabilities, i.e. the probabilities that to elements j and l are jointly included in a sample with a specified sampling design, are given by $\pi_{jl} = P(i, l \in S)$.

The variable of interest is called y with $y_1, \ldots, y_j, \ldots, y_N$ as the characteristics of variable y in the population. Hence, the population total of the variable of interest is defined by $\tau_y = \sum_{j=1}^{N} y_j$ whereas the population mean is defined by $\bar{y} = \tau_y / N = 1/N \sum_{j=1}^{N} y_j$.

3.1.1. Horvitz-Thompson estimator

The following description is mainly based on Särndal et al. (1992, Chapter 2.8). After drawing a sample $s \subseteq U$ of size n from population U and calculating the design weights $w_j \ \forall \ i \in S$, a classic estimator of the population total or mean is given by the Horvitz-Thompson (HT) or π-estimator developed by Horvitz and Thompson (1952). A HT estimator of the population total is given by

$$\hat{\tau_y}^{\mathrm{HT}} = \sum_{j=1}^{n} y_j w_j$$

with estimated variance

$$\hat{V}(\hat{\tau_y}^{\mathrm{HT}}) = \sum_{j=1}^{n} \sum_{l=1}^{n} y_j y_l (1 - ((\pi_j \pi_l)/\pi_{jl})).$$

The HT estimator is design-unbiased.

Independent from one another, Sen (1953), Yates and Grundy (1953) developed another variance estimator under fixed sample sizes, the so called Sen-Yates-Grundy (SYG) variance estimator

$$V(\hat{\tau_y}^{\mathrm{HT}})^{\mathrm{SYG}} = -\frac{1}{2}\sum_{j\in U}\sum_{\substack{l\in U\\ j\neq l}}(\pi_j\pi_l - \pi_{jl})(y_j\pi_j^{-1} - y_l\pi_l^{-1})^2$$

and can be estimated by

$$\hat{V}(\hat{\tau_y}^{\mathrm{HT}})^{\mathrm{SYG}} = -\frac{1}{2}\sum_{j\in S}\sum_{\substack{l\in S\\ j\neq l}}(\pi_j\pi_l - \pi_{jl})(y_j\pi_j^{-1} - y_l\pi_l^{-1})^2.$$

Depending on the sampling design, the formulas for the point and variance estimation of the HT estimator can be simplified. For further information on the variance estimation of the HT estimator with an emphasis on complex survey designs see Bruch et al. (2011).

3.1.2. Calibration estimation

The efficiency of the Horvitz-Thompson estimator can be improved by incorporating auxiliary information in the estimation process. A widely used class of estimators incorporating auxiliary information are the calibration estimators. Deville and Särndal (1992) first coined the term calibration. However, the original idea of exploiting auxiliary information within the estimation process goes back to raking estimators introduced by Deming and Stephan (1940). A detailed description of the calibration approach is given by the following definition.

Definition 1. ***Calibration estimators*** *(Särndal, 2007, p. 99)*
The calibration approach to estimation for finite populations consists of

a) a computation of weights that incorporate specified auxiliary information and are restrained by calibration equation(s),
b) the use of these weights to compute linearly weighted estimates of totals and other finite population parameters: weight times variable value, summed over a set of observed units,
c) an objective to obtain nearly design unbiased estimates as long as nonresponse and other non-sampling errors are absent.

The class of calibration estimators has the general form

$$\hat{\tau}_y^{\mathrm{cal}} = \sum_{i\in s} d_i g_i y_i,$$

where the weights g_i are adjusted such as

$$\sum_{i\in s} d_i g_i \boldsymbol{x_i} = \boldsymbol{\tau_x} \tag{3.1}$$

is satisfied with $\boldsymbol{\tau_x}$ as known total vector of dimension Q of the auxiliary variables (for example from a census or other external sources). Equation 3.1 is called *calibration constraints*. These calibration constraints guarantee that the sample sum of the weighted auxiliaries equals their known population totals. The final calibration weight is given by $w_i^{\text{cal}} = d_i g_i$. Note that the weight $g_i = g_i(s)$ depends on the sample. There are two different concepts to compute the calibrated weights: the minimum distance approach and the functional approach.

In the **minimum distance approach**, introduced by Deville and Särndal (1992), the calibrated weights w_i^{cal} are chosen as close as possible to the original design weights d_i. Closeness between both weights is measured via a pre-specified distance function $G(w_i^{\text{cal}}, d_i)$. Requirements for the distance function are (i) $G(w_i^{\text{cal}}, d_i) \geq 0$; (ii) strict convexity; (iii) differentiability with respect to w_i^{cal} with $g(w_i^{\text{cal}}) = \frac{\partial G(w_i^{\text{cal}}, d_i)}{\partial w_i^{\text{cal}}}$, and (iv) $G(1) = g(1) = 0$ (Haziza and Beaumont, 2017, p. 213). The latter property ensures that for $w_i^{\text{cal}} = d_i$ the distance is zero. Then the minimization problem to compute the calibrated weights is formalized as

$$\min_{w_i} \sum_{i \in s} \frac{d_i G(w_i^{\text{cal}}, d_i)}{\alpha_i} \qquad \text{subject to calibration constraints equation (3.1)},$$

where α_i is a positive scale factor indicating the importance of unit i. In most practical situations α is set to 1. The solution of the minimization problem yields the calibration weights

$$w_i^{\text{cal}} = d_i F(\alpha_i \boldsymbol{x_i^T \lambda}), \tag{3.2}$$

where $F(u) = g^{-1}(u)$ is the inverse function of $g(\cdot)$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_Q)^T$ denotes a vector of Lagrange multipliers. Properties (i) and (ii) ensure that the inverse function $F(u)$ exists. The Lagrange multiplier $\boldsymbol{\lambda}$ is determined by solving

$$\sum_{i \in s} d_i \boldsymbol{x_i} F(\alpha_i \boldsymbol{x_i^T \lambda}) = \boldsymbol{\tau_x}. \tag{3.3}$$

As equation (3.3) involves a system of Q equations and Q unknowns, it can be solved via the Newton-Raphson algorithm (Geiger and Kanzow, 2002, p. 235).

Applying the chi-square distance $G(w_i^{\text{cal}}, d_i) = \frac{1}{2}(\frac{w_i^{\text{cal}}}{d_i} - 1)^2$ and assuming $\alpha = 1$, we obtain

$$g(w_i^{\text{cal}}, d_i) = \left(\frac{w_i^{\text{cal}}}{d_i} - 1\right)\frac{1}{d_i} \qquad \text{and} \qquad F(\boldsymbol{x_i^T \lambda}) = g^{-1}(\boldsymbol{x_i^T \lambda}) = 1 + \boldsymbol{x_i^T \lambda}.$$

Inserting $1 + \boldsymbol{x_i^T \lambda}$ into equation (3.2) yields the calibrated weights

$$w_i^{\text{cal}} = d_i(1 + \alpha_i \boldsymbol{x_i^T \lambda})$$

with Lagrange multipliers $\boldsymbol{\lambda} = (\sum_{i \in s} \alpha_i d_i \boldsymbol{x_i x_i^T})^{-1}(\boldsymbol{\tau_x} - \boldsymbol{\hat{\tau}_x}^{\text{HT}})$. Hence, the minimization of the chi-square distance leads to the GREG weights (Särndal, 2007, p. 106).

Deville and Särndal (1992) examined six further distance functions, such as the Hellinger distance or the minimum entropy distance. Deville et al. (1993) introduced generalized raking estimators as a subclass of calibration estimators, which can be used when marginal counts of the auxiliaries are known. In this case, the distance function is multiplicative. The subclass of generalized raking estimators contains the classical raking estimator originated by Deming and Stephan (1940). Raking is equivalent to iterative proportional fitting and the maximum entropy approach by Wittenberg (2010). Further distance functions are discussed in Huang and Fuller (1978), Alexander (1987), Singh and Mohl (1996), and Stukel et al. (1996). Each distance function leads to a specific weighting system despite they meet the same calibration constraints. Deville and Särndal (1992) showed that under mild conditions on $F(\cdot)$ the calibration estimator generated by different distance functions asymptotically equals the GREG estimator. Thus, for large sample sizes, the choice of the distance function has only a minor impact on the properties of the calibration estimator. Singh and Mohl (1996) and Stukel et al. (1996) extended this finding to modest sample sizes.

The **functional form approach** was introduced by Estevao and Särndal (2000). The intention of the functional form approach was that the minimum distance approach does not provide much insight into the properties of the different estimators. Instead of a distance function, a simple functional form is imposed, which depends on the instrumental variables $\boldsymbol{z_i} = (z_{i1}, \ldots, z_{iQ})^T$. The vector $\boldsymbol{z_i}$ has to be of the same dimension as the auxiliary vector $\boldsymbol{x_i}$. Then, the calibrated weights w_i^{calF} are determined by the functional relationship

$$w_i^{\text{calF}} = d_i F(\boldsymbol{\lambda}^T \boldsymbol{z}_i),$$

where $\boldsymbol{\lambda}$ is a vector determined by the calibration constraints $\sum_{i\in s} w_i^{\text{calF}} \boldsymbol{x}_i = \boldsymbol{\tau}_x$. The final weights depend on the design weight, the auxiliary and the instrumental variables. F is a known real-valued function. Examples are $F(u) = 1 + u$ and $F(u) = exp(u)$. For the linear function $F(u) = 1 + u$ the weights are given by

$$w_i^{\text{calF}} = d_i(1 + \boldsymbol{\lambda}^T \boldsymbol{z}_i) \tag{3.4}$$

with $\boldsymbol{\lambda} = (\sum_{i\in s} d_i \boldsymbol{z}_i \boldsymbol{x}_i^T)^{-1}(\boldsymbol{\tau}_x - \hat{\boldsymbol{\tau}}_x^{\text{HT}})$. The resulting calibration estimator is given by

$$\hat{\tau}_y^{\text{calF}} = \sum_{i\in s} w_i^{\text{calF}} y_i.$$

Inserting $\boldsymbol{\lambda}$ in w_i^{calF} gives

$$w_i^{\text{calF}} = d_i + d_i \boldsymbol{z}_i (\sum_{i\in s} d_i \boldsymbol{z}_i \boldsymbol{x}_i^T)^{-1}(\boldsymbol{\tau}_x - \hat{\boldsymbol{\tau}}_x^{\text{HT}})$$

which reminds us of an instrumental variables regression known from econometrics. Therefore, the functional form approach was later termed as **instrument vector approach** by Estevao and Särndal (2006), Kott (2003) and Kott (2006). Irrespective from the choice of z_i , the weights w_i^{calF} satisfy the calibration constraints. Even 'deliberately awkward choices' for $\boldsymbol{z}_i$ give surprisingly good results (Särndal, 2007, p. 106). The following relation is valid: $\hat{\tau}_y^{\text{GREG}} \subseteq \hat{\tau}_y^{\text{calF}} \subseteq \hat{\tau}_y^{\text{cal}}$ (Estevao and Särndal, 2000, p. 382).

The calibration estimator is no longer design-unbiased. However, the bias is typically small (Wu and Lu, 2016). Instead of it is design-consistent under a suitable asymptotic framework (Isaki and Fuller, 1982). Calibration estimators generated by different distance functions share the same large sample design-based variance (Deville et al., 1993, p. 1014). Because Deville and Särndal (1992) proved the asymptotic equivalence of the calibration and the GREG estimator, the design-based variance of the calibration estimator can be approximated by the residual variance of the GREG estimator.

Remarks: Devaud and Tillé (2019) have recently published a review of calibration methods in the past 25 years accompanied with a discussion and rejoinder. A generalized calibration method using soft constraints on different hierarchical levels, which may become important in regional and local estimation of indicators, was provided in Burgard et al. (2019). Though calibration methods are widely used to improve accuracy estimates and compensate for nonresponse while using adequate auxiliary information, this benefit is less efficient when estimation non-linear statistics. The reason is the auxiliary variable must be highly correlated to a linearized variable of the non-linear statistics similar to the procedure presented under variance estimation below. This can often be hardly achieved.

3.1.3. Domain estimation

The following description is mainly based on Särndal et al. (1992, Chapter 10). Using a sample not only estimates for the target population of the sample can be calculated, but also estimates for certain domains. Let population U be partitioned into D sub-populations $U_1, \dots, U_d, \dots, U_d$, called domains, of sizes $N_1, \dots, N_d, \dots, N_d$ such that

$$U = \bigcup_{d=1}^{D} U_d, \quad N = \bigcup_{d=1}^{D} N_d \quad \text{and} \quad n = \bigcup_{d=1}^{D} n_d.$$

Then, the domain totals and means are defined by

$$\tau_{y_d} = \sum_{j=1}^{U_d} y_j \quad \text{and} \quad \bar{y}_d = \tau_{y_d}/N_d, \quad d = 1, \dots, D.$$

The HT estimator of the domain total is defined by

$$\hat{\tau}_{y_d}^{\text{HT}} = \sum_{j=1}^{n_d} w_j y_j$$

with the corresponding variance estimated by

$$\hat{V}(\hat{\tau}_{y_d}^{\text{HT}}) = \sum_{j=1}^{n_d} \sum_{l=1}^{n_d} y_j y_l (1 - ((\pi_j \pi_l)/\pi_{jl})).$$

Since often in domain estimation, the sampling design was not constructed to cover the domains of interest with fixed sample sizes, random sample sizes result. This may yield less accurate domain estimates in design-based approaches. In order to compensate for this, the estimates for N_d can be computed by

$$\hat{N}_d = \sum_{j=1}^{n_d} w_j \cdot 1 \qquad d = 1, \ldots, D.$$

When estimating the population mean, this refers to the Hájek estimator (cf. Särndal et al., 1992, p. 182).

3.1.4. Variance estimation of non-linear statistics

Not only the calculation of estimates, but also their variance estimation are crucial. An overview and detailed description of variance estimation methods is given in Graf et al. (2011), Münnich and Zins (2011). Münnich and Zins (2011) focus on the variance estimation for indicators of poverty and inequality. As the variance estimators of non-linear statistics cannot be given in closed form, resampling and linearisation methods are often used. Münnich and Zins (2011) present the use of influence functions, estimating equations and linearization of poverty and inequality measures including the at-risk-of-poverty rate, the quintile share ratio and the Gini coefficient.

For the introduction of the following non-linear statistics and their corresponding variance estimation methods, some background information on variance estimation on non-linear statistics is necessary. Langel and Tillé (2013) give an overview about the rationale behind linearization techniques used to estimate the variances of non-linear statistics. When θ is a statistics with estimated value $\hat{\theta}$, linearization techniques try to find a linearized variable z_j such that

$$\hat{\theta} - \theta \approx \sum_{j \in S} w_j z_j - \sum_{j \in U} z_j.$$

Then, the variance of $\hat{\theta}$ is approximated by

$$\hat{Z} = \sum_{j \in S} w_j z_j.$$

As the z_js normally depend on population parameters and thus have to be estimated, $\hat{z}_j$ is obtained and plugged in for z_j. Further information for variance estimation of non-linear statistics is given in Glasser (1962), Deville (1999), Deville and Särndal (1992), Isaki and Fuller (1982), Osier (2009), Osier et al. (2013), Kovacevic and Binder (1997), Eurostat (2013).

3.1.5. At-risk-of-poverty or social exclusion rate

Eurostat (a) describe that the at-risk-of-poverty or social exclusion rate (AROPE) is calculated as the percentage of people who are at-risk-of-poverty or severely deprived or living in a household with very low work intensity over the total population. The three input measures to AROPE are the percentage of people who are at-risk-of-poverty, severely deprived or living in a household with very low work intensity over the total population. The definition of the three measures is described in Section 2.1.1.

The formulas for calculating the at-risk-of-poverty rate (ARPR) are described in Graf and Tillé (2014) which is the main source of the following description. First, the at-risk-of-poverty threshold (ARPT) has to be defined. The ARPT is defined by the 60% median income as

$$\text{ARPT} = 0.6F^{-1}(0.5)$$
$$\widehat{\text{ARPT}} = 0.6\hat{Q}_{0.5} = 0.6\hat{m}$$

with $\hat{m} = \hat{Q}_{0.5}$ as the estimated median income of a sample.

The ARPR is then defined as the share of the population with an income below the ARPT

$$\widehat{\text{ARPR}} = \frac{\sum_{y_j < \widehat{\text{ARPT}}} w_j}{\hat{N}},$$

where the weights w_j might have been adjusted, e.g. to nonresponse.

For estimating the variance of the ARPR, Graf and Tillé (2014) use the linearization technique and present the estimated linearized variable of the QSR defined by Osier (2009) as

$$\begin{aligned}\hat{z}_j^{\text{ARPR}} =& \frac{1}{\hat{N}}\left(\mathbf{1}_{y_j} \leq \widehat{\text{ARPR}} - \widehat{\text{ARPR}}\right) - \frac{f(\widehat{\text{ARPR}})}{f(\hat{m})}\frac{0.6}{\hat{N}}\left(\mathbf{1}_{[y_j \leq \hat{m}]} - 0.5\right) \\ =& \frac{1}{\hat{N}}\left(\mathbf{1}_{y_j} \leq \widehat{\text{ARPR}} - \widehat{\text{ARPR}}\right) + f(\widehat{\text{ARPR}})\hat{z}_j^{\text{ARPR}}.\end{aligned}$$

Hence, the income density function has to estimated at the median income and the ARPT.

For variance estimation of AROPE estimates using linearisation see Deville (1999), Osier (2009).

3.1.6. Quintile share ratio

The quintile share ratio (QSR or S_{80}/S_{20}) is described in (Eurostat, 2009, 15):
"The 'S_{80}/S_{20} income quintile share ratio' is the ratio of the sum of equivalised disposable income received by the 20% of the country's population with the highest equivalised disposable income (top inter-quintile interval) to that received by the 20% of the country's population with the lowest equivalised disposable income (lowest inter-quintile interval)."

In order to calculate the QSR, the equivalised disposable income (EQ_INC) of a person is calculated. Then, given all equivalised disposable incomes, the five quintiles persons belong to (QPB) are calculated. The quintile share ratio can then be calculated as (Eurostat, 2009, 15):

$$\text{QSR} = \frac{\sum\limits_{j \in \text{QPB}=5} w_j * EQ\_INC_j}{\sum\limits_{j \in \text{QPB}=1} w_j * EQ\_INC_j},$$

where the weights w_j might have been adjusted, e.g. to nonresponse.

For estimating the variance of the QSR, Graf and Tillé (2014) use the linearization technique and present the estimated linearized variable of the QSR based on Langel and Tillé (2013) as

$$\hat{z}_j^{\mathrm{QSR}} = \frac{y_j - \left\{ y_j H\left(\frac{0.8\hat{N} - \hat{N}_{j-1}}{w_j}\right) + \hat{Q}_{0.8}[0.8 - \mathbf{1}_{[y_j < \hat{Q}_{0.8}]}] \right\}}{\hat{Y}_{0.2}} - \frac{(\hat{Y} - \hat{Y}_{0.8}) \left\{ y_j H\left(\frac{0.2\hat{N} - \hat{N}_{j-1}}{w_j}\right) + \hat{Q}_{0.2}[0.2 - \mathbf{1}_{[y_j < \hat{Q}_{0.2}]}] \right\}}{\hat{Y}_{0.2}^2},$$

where

$$H(x) = \begin{cases} 0 & \text{if} \quad x < 0 \\ x & \text{if} \quad 0 \leq x < 1 \\ 1 & \text{if} \quad x \geq 1 \end{cases}.$$

For variance estimation of QSR estimates using linearisation see also Hulliger and Münnich (2006).

3.1.7. Gini coefficient

The Gini coefficient (GINI) is described in (Eurostat, 2009, 15):
"The Gini coefficient is defined as the relationship of cumulative shares of the population arranged according to the level of equivalised disposable income, to the cumulative share of the equivalised total disposable income received by them."

The GINI ranges from 0 to 1, i.e. from complete equality to complete inequality. In order to calculate the GINI, the equivalised disposable incomes (EQ_INC) are sorted. The GINI is then calculated as (Eurostat, 2009, 16):

$$\mathrm{GINI} = 100 * \left(\frac{2 \sum_{j=1}^{n} \left(w_j EQ\_INC_j \sum_{j=1}^{i} w_j \right) - \sum_{j=1}^{n} w_j^2 EQ\_INC_j}{(\sum_{j=1}^{n} w_j)(\sum_{j=1}^{n} w_j EQ\_INC_j)} - 1 \right),$$

where the weights w_j might have been adjusted, e.g. to nonresponse.

For estimating the variance of the GINI, Graf and Tillé (2014) use the linearization technique and present the estimated linearized variable of the GINI based on Langel and Tillé (2013) as

$$\hat{z}_j^{\mathrm{GINI}} = \frac{1}{\hat{N}\hat{Y}} \left[2\hat{N}_j(y_j - \hat{\bar{Y}}_j) + \hat{\bar{Y}} - \hat{N}y_j - \widehat{\mathrm{GINI}}(\hat{Y} + y_j\hat{N}) \right],$$

where $\hat{\bar{Y}}_j = \sum_{l=1}^{j} w_l y_l / \hat{N}_j$, the values of y_l are sorted and distinct and $\widehat{\mathrm{GINI}}$ denoted the estimated value of GINI from a sample.

For variance estimation of the GINI estimates using Taylor linearisation see also Kovacevic and Binder (1997).

3.2. Model-based estimation methods for indicators

There are many situations were model-based approaches can be of added value for the production of official statistics indicators and specifically of poverty indicators.

The use of models for Small Area Estimation (SAE) of indicators is probably the aspect of major interest when the goal is to estimate indicators at regional (local) level. A general overview of SAE models has been already presented in MAKSWELL Deliverable 2.2 (van den Brakel et al., 2019). In the next subsection we briefly review the major model-assisted and model-based methods for SAE, considering the specific case of estimating poverty indicators based on unit-level models.

Other relevant examples of model-based methods for the production of official statistics indicators include the estimation in the presence of non-sampling errors, such as nonresponse and attrition, discontinuities due to survey transitions and the use of register data (van den Brakel and Bethlehem, 2008).

Model-based approaches are also important in the definition of composite indicators, such as composite well-being indicators presented in section 2. Examples of model-based approaches for the definition of composite indicators include Structural Equation Modelling (D'Agostino et al., 2019, Lauro et al., 2018). Specifically, Lauro et al. (2018) discuss the use of Partial Least Squares Path Modelling, that can be of important value for CI definition. Also approaches based on factor analysis have been used and applied for reducing data dimensionality (Cavicchia and Vichi, 2017, Betti, 2017, Cavicchia and Vichi, 2017, D'Agostino et al., 2018, Potsi et al., 2016).

3.2.1. Model-assisted methods for SAE estimation of poverty indicators

In the last 30 years mixture modes of making inference have become common in survey sampling: in many cases design-based inference is model assisted. Also in the SAE context the model assisted approach has become popular. Pratesi and Salvati (2016) shortly review the most common estimators under this approach.

Among design-based methods assisted by the specification of a model for the study variable there are three families of methods recently applied in poverty mapping. Generalized Regression (GREG) estimators, pseudo-EBLUP estimators and M-quantile weighted estimators.

The Generalized Regression (GREG) approach can be used to estimate several poverty indicators. Anyhow with reference to the estimation of the small area mean, the estimators under this approach share the following structure

$$\hat{\bar{m}}_d^{GREG} = \sum_{j \in U_d} \hat{y}_{jd} + \sum_{j \in s_d} w_{jd}(y_{jd} - \hat{y}_{jd}). \tag{3.5}$$

where w_{jd} is the sampling weight of unit j within area d that is the reciprocal of the respective inclusion probability π_{jd}. Different GREG estimators are obtained in association with different models specified for assisting estimation, i.e. for calculating predicted values $\hat{y}_{jd}$, $j \in U_d$. In the simplest case a fixed effects regression model is assumed: $E(y_{jd}) = \mathbf{x}_{jd}^T\boldsymbol{\beta}$, $\forall j \in U_d$, $\forall d$ where the expectation is taken with respect to the assisting model. Lehtonen and Veijanen (1999) introduce an assisting two-levels model where which is a model with area specific regression coefficients. In practice, not all coefficients need to be random and models with area-specific intercepts mimicking linear mixed models may be used (see Lehtonen et al., 2003). In this case the GREG estimator takes the form (3.5) with $\hat{y}_{jd} = \mathbf{x}_{jd}^T(\hat{\boldsymbol{\beta}} + \hat{\mathbf{u}}_d)$. Estimators $\hat{\boldsymbol{\beta}}$ and $\hat{\mathbf{u}}$ are obtained using generalized least squares and restricted maximum likelihood methods.

Under the pseudo-EBLUP approach the estimators are derived taking into account the sampling design both via the sampling weights and the auxiliary variables in the models. The estimators of the area mean proposed by Prasad and Rao (1999) and You and Rao (2002) are based on the assumption of a population nested error regression model and it is also assumed that the sampling design is ignorable given the auxiliary variables included in the model. As for the error terms it is assumed that $u_d \overset{i.i.d.}{\sim} N(0, \sigma_u^2)$ and $e_{ij} \overset{i.i.d.}{\sim} N(0, \sigma_e^2)$.

By combining a Hájek type direct estimator of $\bar{m}_d$ defined as $\bar{y}_{dw} = \sum_{j \in s_d} \breve{w}_{jd} y_{jd}$ where $\breve{w}_{jd} = w_{jd}\Big(\sum_{j \in s_d} w_{jd}\Big)^{-1}$, and the nested error regression model, Prasad and Rao (1999) obtain the following aggregated area level model:

$$\bar{y}_{dw} = \bar{\mathbf{x}}_{dw}^T \boldsymbol{\beta} + v_d + \bar{e}_{dw}, \tag{3.6}$$

with $\bar{e}_{dw} = \sum_{j \in s_d} \breve{w}_{jd} e_{jd}$ and $\bar{\mathbf{x}}_{dw} = \sum_{j \in s_d} \breve{w}_{jd} \mathbf{x}_{jd}$.

The design consistent pseudo-EBLUP estimator $\hat{\eta}_{dw}$ of the d-th area mean is then given by

$$\hat{\bar{m}}_{dw} = \hat{\gamma}_{dw} \bar{y}_{dw} + \big(\bar{\mathbf{X}}_d - \hat{\gamma}_{dw} \bar{\mathbf{x}}_{dw}\big)^T \hat{\boldsymbol{\beta}}_w, \tag{3.7}$$

where $\hat{\gamma}_{dw} = \hat{\sigma}_u^2(\hat{\sigma}_u^2 + \hat{\sigma}_e^2 \delta_d)^{-1}$, $\delta_d = \sum_{j \in s_d} \breve{w}_{jd}^2$ and

$$\hat{\boldsymbol{\beta}}_w(\hat{\sigma}_u^2, \hat{\sigma}_e^2) = \left(\sum_{d=1}^{D} \sum_{j \in s_d} \breve{w}_{jd} \mathbf{x}_{jd} (\mathbf{x}_{jd} - \hat{\gamma}_{dw} \bar{\mathbf{x}}_{dw}^T) \right)^{-1} \left(\sum_{d=1}^{D} \sum_{j \in s_d} \breve{w}_{jd} (\mathbf{x}_{jd} - \hat{\gamma}_{dw} \bar{\mathbf{x}}_{dw}^T y_{jd}) \right). \tag{3.8}$$

The variance components (σ_u^2, σ_e^2) can be estimated using for example, REML or the fitting-of-constants method. Both Prasad and Rao (1999) and You and Rao (2002) provided formulae for the model-based MSE associated with the pseudo-EBLUP estimators of the area mean. Jiang and Lahiri (2006) noted that these estimators are not second-order correct. Torabi and Rao (2010) derived a second order unbiased predictor for the pseudo-EBLUP estimator.

An alternative family of model-assisted small area estimators is based on the M-quantile methodology (Chambers and Tzavidis, 2006). Recently, under this model, Fabrizi et al. (2014) proposed a design consistent estimators of area-specific poverty indicators using the Rao-Kovar-Mantel estimator of the distribution function of income F_i defined as

$$\hat{F}_d^{WMQ/RKM} = N_d^{-1}\Big[\sum_{j \in s_d} w_{jd} I(y_{jd} \le t) + \sum_{j \in U_d} I(\mathbf{x}_{jd}^T \hat{\boldsymbol{\beta}}_{w\bar{\theta}_d} \le t) - \sum_{j \in s_d} w_{jd} I(\mathbf{x}_{jd}^T \hat{\boldsymbol{\beta}}_{w\bar{\theta}_d} \le t)\Big]. \tag{3.9}$$

where $\hat{\boldsymbol{\beta}}_{wq}$ is a design consistent estimator of $\boldsymbol{\beta}_q$. In the application of M-quantile regression to small area estimation Chambers and Tzavidis (2006) characterize the variability across the population, beyond what is accounted for by the model covariates, by using the so-called M-quantile coefficients of the population units. For unit j in area d, this coefficient is the value θ_{jd} such that $Q_{\theta_{jd}}(y_{jd}|\mathbf{x}_{jd}) = y_{jd}$, where $Q_q(y_{jd}|\mathbf{x}_{jd})$ is the conditional M-quantile that is assumed to be a linear function of the auxiliary information. The authors observe that if a hierarchical structure does explain part of the variability in the population data, units within areas defined by this hierarchy are expected to have similar M-quantile coefficients. Average area coefficients $\bar{\theta}_d$ may be calculated and this represents an alternative approach to estimating area random effects without the need for using parametric assumptions.

More specifically, the weighted M-quantile-based small area estimator of the mean from (3.9) is

$$\hat{m}_d^{WMQ} = \int td\hat{F}_d^{WMQ/RKM}(t) = \frac{1}{N_d}\sum_{j\in s_d} w_{jd}y_{jd} + \Big(\frac{1}{N_d}\sum_{j\in U_d}\mathbf{x}_{jd}^T - \frac{1}{N_d}\sum_{j\in s_d} w_{jd}\mathbf{x}_{jd}^T\Big)\hat{\boldsymbol{\beta}}_{w\bar{\theta}_d}. \tag{3.10}$$

The M-quantile method can be also used for estimating the Head Count Ratio and the Poverty Gap. Using t to denote the poverty line, different poverty indicators are defined by the area-specific mean of the variable derived:

$$f_{jd}(\alpha,t) = \Big(\frac{t-y_{jd}}{t}\Big)^{\alpha} I(y_{jd}\leq t), d = 1,\dots,D;\ j = 1,\dots,N_d. \tag{3.11}$$

The population-level small area-specific poverty indicator can be decomposed as:

$$F_d(\alpha,t) = N_d^{-1}\Big[\sum_{j\in s_d} f_{jd}(\alpha,t) + \sum_{j\in r_d} f_{jd}(\alpha,t)\Big]. \tag{3.12}$$

The first component in (3.12) is observed in the sample, whereas the second component has to be predicted by using the M-quantile model. Tzavidis et al.(2014) propose a non-parametric approach by using a smearing-type estimator. More specifically:

$$F_d(\alpha,t) = N_d^{-1}\Big[\sum_{j\in s_d} f_{jd}(\alpha,t) + \sum_{j\in r_d} E(f_{jd}(\alpha,t))\Big]. \tag{3.13}$$

For simplicity let us focus on the simplest case when $\alpha = 0$. An estimator of $F_d(0,t)$ is obtained by substituting an estimator of $E(f_{jd}(\alpha,t))$ in (3.13) leading to

$$\hat{F}_d(0,t) = N_d^{-1}\Big[\sum_{j\in s_d} w_{jd}f_{jd}(0,t) + \frac{1}{\sum_{j\in s_d} w_{jd}}\sum_{k\in r_d}\sum_{j\in s_d} w_{jd}I(\mathbf{x}_{kd}^T\hat{\boldsymbol{\beta}}_{w\bar{\theta}_d} + \hat{e}_{jd}\leq t)\Big], \tag{3.14}$$

where $\hat{e}_{jd}$s are the estimated residuals from the M-quantile fit. The same approach can be followed to estimate $\hat{F}_d(1,t)$ or any other of the FGT poverty measures.

For the estimation of the variance of the MQ predictors see Fabrizi et al. (2014) where two alternative estimators of the variance of the MQ predictors are proposed.

Even if the use of design consistent estimators in SAE is somewhat questionable because of the small sample sizes in some or all of the areas, as Pfeffermann (2013) noted, the families of methods we described above offer generally design consistent estimators.

The three approaches previously described give practical solutions to benchmarking, they face with the presence of outliers, the estimates that they provide are differently affected by the shrinkage effect, and they all offer out-of-sample predictions.

Also to protect against possible model failures, *benchmarking* procedures make the total of small area estimates match a design consistent estimate for a larger area. For what concerns benchmarking all the families of methods offer a solution.

There are two kinds of benchmarked estimators: estimators that are internally benchmarked (or self-benchmarked) and those that are externally benchmarked. Self benchmarked predictors are the GREG estimator and the pseudo-EBLUP introduced by You and Rao (2002) and also discussed in Rao (2003). The externally benchmarked ones are more common under the model-based approach, as we better describe in the next section.

The GREG procedure uses the higher level totals as auxiliary data in calculating survey weights, thereby adjusting the lower level weights so that the total and subtotal estimates are consistent. In addition, the weights that are used for direct estimation using survey data in GREG expression are often constructed using calibration methods. Often benchmarking to auxiliary totals is used together with weight equalization.

Benchmarking (forcing certain estimates to match known totals) has been shown to reduce variances for statistics correlated with the auxiliary characteristics, and weight equalization (forcing the weights within higher-level units to be equal) has been shown to further reduce variances for statistics measured on the higher-level units (Lehtonen and Veijanen, 1999).

The pseudo-EBLUP estimators satisfy the benchmarking property without any adjustment in the sense that they add up to the direct survey regression estimator when aggregated over the areas. A drawback of this type of self benchmarked estimators is that they force the use of the same auxiliary information used for the direct usually GREG-type estimator also for the model-based small area predictors, whereas it could be very profitable to allow for different auxiliary variables at the small area level.

Coming to the M-quantile approach note that expression (3.10) has a GREG-type form. This is the basis to see that the MQ predictors do not satisfy the benchmarking property as it is shown in Fabrizi et al. (2014). Here the authors propose a method of constraining M-quantile regression. It can be applied to obtain benchmarking MQ small area estimates.

The treatment of the *outliers* is not the focus of the estimators of GREG type nor of those under the pseudo-eblup approach, while the weighted M-quantile approaches this issue.

There are studies under the AMELI (2018) project that illustrate the behaviour of the GREG-like estimators in the presence of different models of outlier-contamination of the observed data. The results show that even if a robust method of fitting the logistic mixed model was not available, the poverty rate estimators are fairly robust: this happens both under a simple random sampling design and under a complex sampling design.

To deal with outliers, Beaumont and Alavi (2004) use the weighted generalized M-estimation technique to reduce the influence of units with large weighted population residuals. For what concerns the empirical pseudo best approach recalled before there is no contribution addressing the robustification of the estimates against the presence of outliers. Jiang et al. (2011) relaxed some of the classical EBLUP model to obtain robust model based predictors. These relaxations may work also under the pseudo-eblup approach but until now no evidence of it has been produced. AMELI project provide evidence also on the behaviour of the Empirical Best Predictor type estimator based on a logistic mixed model. This estimator is least affected by contaminations when the data comes from a simple random sample but it is not based on the pseudo-eblup approach.

As it concerns the M-quantile estimator with respect to GREG-S popular in small area literature (see Rao, 2003, Section 2.5), note that: *i*) the use of an area-specific coefficient ($\bar{\theta}_d$) in M-quantile regression accounts for area characteristics not explained by the auxiliary variables; and *ii*) the use of M-estimation offers outlier robust estimation. Specifically, the recourse to M-quantile regression reduces the impact that outlier observations have on the estimated regression coefficients and thereby on the small area means.

The models which are assisting the estimation under the design based approach can have the tendency to *under/over-shrinkage* of small area estimators.

The desirable property of neutral shrinkage is not achieved under the pseudo-eblup approach. In this case it is reasonable that is confirmed the over-shrinking behaviour of the Empirical Best predictors. The understatement of extreme values, referred to as over-shrinkage in this context, is problematic when the goal is the description of the overall distribution among areas. However, this tendency can be adjusted and it is likely that the adjustment can work even under the pseudo-eblup approach, but until now no evidence of it has been produced. The tendency of GREG estimators is similar to that of direct estimators and in contrast with that of the over-shrinking EB predictors, as the results of the EURAREA project have shown.

The behaviour of M-quantile based predictors is then more similar to that of direct estimators and GREG. Anyhow, Fabrizi et al. (2014) propose an adjustment of the benchmarked MQ predictors in order to obtain estimators with approximately neutral shrinkage. They extend the methodology of Fabrizi et al. (2012) to obtain estimates that enjoy ‘ensemble’ properties, that is properties related to the estimation of a functional of an ensemble of parameters (Frey and Cressie, 2003). An ensemble of estimators is said to be neutral with respect to shrinkage if the variance of the ensemble of the parameters can be unbiasedly estimated by the variance of the ensemble of the estimators. This guarantees a correct representation of the geographical variation of the variable in question. Otherwise, this geographical variation may be over or underestimated. Neutral shrinkage is important when small area estimators are used to create ‘maps’.

For the set $E = \{d|n_d = 0\}$ of *the out of sample areas* i.e., areas where $n_d = 0$, the GREG-like estimators cannot be computed. The pseudo-eblup approach provide predictors under the specified models which are likely to underestimate the variability of the estimates among areas. Consistently with Chambers and Tzavidis (2006), the small area estimator $\hat{\bar{m}}_d^{WMQ}$ can be defined as $N_d^{-1}\sum_{j\in U_d}\mathbf{x}_{jd}^T\hat{\boldsymbol{\beta}}_{w0.5}$, that is a synthetic estimator based on the weighted M-regression.

MSE estimation of model-assisted SAE methods

The MSE of the pseudo-ELBUP second order unbiased can be obtained through a linearization method (Torabi and Rao, 2010). The MSE of $\hat{\bar{m}}_{dw}$ can be expressed as

$$MSE(\hat{\bar{m}}_{dw}) = E[(\tilde{\bar{m}}_{dw}^B - \bar{m}_d)^2] + E[(\tilde{\bar{m}}_{dw} - \tilde{\bar{m}}_{dw}^B)^2] + E[(\hat{\bar{m}}_{dw}^{TR} - \tilde{\bar{m}}_{dw})^2] + 2C_{1dw}(\sigma_u^2, \sigma_e^2) + 2C_{2dw}(\sigma_u^2, \sigma_e^2),$$

where $\tilde{\bar{m}}_{dw}^B = \bar{\mathbf{X}}_d^T\beta + \gamma_{dw}(\bar{y}_{dw} - \bar{\mathbf{x}}_{dw}^T\beta)$, $\tilde{\bar{m}}_{dw} = \tilde{\bar{m}}_{dw}(\delta) = \bar{\mathbf{X}}_d^T\tilde{\beta}_w(\sigma_u^2, \sigma_e^2) + \gamma_{dw}(\bar{y}_d - \bar{\mathbf{x}}_{dw}\tilde{\beta}_w(\sigma_u^2, \sigma_e^2))$, $C_{1dw}(\sigma_u^2, \sigma_e^2) = E[(\tilde{\bar{m}}_{dw} - \tilde{\bar{m}}_{dw}^B)(\tilde{\bar{m}}_{dw}^B - \bar{m}_d)]$ and $C_{2dw}(\sigma_u^2, \sigma_e^2) = E[(\hat{\bar{m}}_{dw}^{TR} - \tilde{\bar{m}}_{dw})(\tilde{\bar{m}}_{dw} - \bar{m}_d)]$.

The term $C_{1dw}(\sigma_u^2, \sigma_e^2)$ can be written as follows

$$\begin{aligned} C_{1dw}(\sigma_u^2, \sigma_e^2) = \sigma_e^2 w_{d.}\gamma_{dw}(\bar{\mathbf{X}}_d - \gamma_{dw}\bar{\mathbf{x}})^T\Big\{\sum_d\sum_j w_{jd}x_{jd}(x_{jd} - \gamma_{dw}\bar{\mathbf{x}}_{dw})^T\Big\}^{-1} \\ \times\Big(\sum_j \breve{w}_{jd}^2 x_{jd} - \bar{\mathbf{x}}_{dw}\delta_d\Big), \end{aligned}$$

where $w_{d.} = \sum_j w_{jd}$. When $\breve{w}_{jd} = 1/n_d$ then $C_{1dw}(\sigma_u^2, \sigma_e^2) = 0$, because $\delta_d = 1/n_d$ and $\bar{\mathbf{x}}_{dw} = \bar{\mathbf{x}}_d$.

Torabi and Rao (2010) proposed an analytical approximation based on Taylor expansion to $C_{2dw}(\sigma_u^2, \sigma_e^2)$, based on the REML estimator of (σ_u^2, σ_e^2):

$$\begin{aligned} C_{2dw}(\sigma_u^2, \sigma_e^2) \approx E\Big[\Big\{ &-2\Big(\frac{\partial\tilde{\bar{m}}_{dw}}{\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\mathbf{A}^{-1}\mathbf{M}(\sigma_u^2,\sigma_e^2) \\ &+\Big(\frac{\partial\tilde{\bar{m}}_{dw}}{\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\mathbf{A}^{-1}\Big(\frac{\partial\mathbf{M}}{\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\mathbf{A}^{-1}\mathbf{M}(\sigma_u^2,\sigma_e^2) \\ &+\frac{1}{2}\Big(\frac{\partial\tilde{\bar{m}}_{dw}}{\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\mathbf{A}^{-1}\mathrm{col}_{1\le j\le 2}\big[\big\{\mathbf{M}^T(\sigma_u^2,\sigma_e^2)(\mathbf{A}^{-1})^T E\Big(\frac{\partial^2 M_j}{\partial(\sigma_u^2,\sigma_e^2)\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\big\}^T\big] \\ &\times\mathbf{A}^{-1}\mathbf{M}(\sigma_u^2,\sigma_e^2) + \frac{1}{2}\mathbf{M}^T(\sigma_u^2,\sigma_e^2)(\mathbf{A}^{-1})^T\Big(\frac{\partial^2\tilde{\bar{m}}_{dw}}{\partial(\sigma_u^2,\sigma_e^2)\partial(\sigma_u^2,\sigma_e^2)^T}\Big)\mathbf{A}^{-1}\mathbf{M}(\sigma_u^2,\sigma_e^2)\Big\} \\ &\times\Big\{(\bar{\mathbf{X}}_d - \gamma_{dw}\bar{\mathbf{x}}_{dw})^T(\tilde{\beta}_w - \beta) + \tilde{u}_{dw} + u_d\Big\}\Big] \\ &\equiv \tilde{C}_{2dw}(\sigma_u^2,\sigma_e^2), \end{aligned}$$

where $\tilde{u}_{dw} = \gamma_{dw}(\bar{y}_{dw} - \mathbf{x}_{dw}^T\tilde{\beta}_w(\sigma_u^2,\sigma_e^2))$, $\mathbf{A} = E[\partial\mathbf{M}(\sigma_u^2,\sigma_e^2)/\partial(\sigma_u^2,\sigma_e^2)^T]$, $\mathrm{col}_{1\le j\le 2}\{c_j\}$ represent a 2×2 matrix with jth column as c_j, $\mathbf{M}(\sigma_u^2,\sigma_e^2) = (\mathbf{M}_1(\sigma_u^2,\sigma_e^2), \mathbf{M}_2(\sigma_u^2,\sigma_e^2))^T$ with $\mathbf{M}_j$ the jth component of $\mathbf{M}$:

$$\mathbf{M}_1(\sigma_u^2,\sigma_e^2) = -\frac{1}{2}tr(\mathbf{PJ}) + \frac{1}{2}\mathbf{y}^T\mathbf{PJPy} \quad \mathbf{M}_2(\sigma_u^2,\sigma_e^2) = -\frac{1}{2}tr(\mathbf{P}) + \frac{1}{2}\mathbf{y}^T\mathbf{PPy}.$$

Here, $\mathbf{P} = \mathbf{V^{-1}} - \mathbf{V^{-1}X(X^TV^{-1}X)^{-1}X^TV^{-1}}$, with $\mathbf{V}^{-1} = $ block diag $(\mathbf{V}_1^{-1}, \ldots, \mathbf{V}_D^{-1})$, $\mathbf{X} = (\mathbf{X}_1^T, \ldots, \mathbf{X}_D^T)^T$ and $\mathbf{J} = $ diag $(\mathbf{J}_{n_1}, \ldots, \mathbf{J}_{n_D})$, $\mathbf{J}_{n_d}$ is an $n_d \times n_d$ matrix of one's.

Torabi and Rao (2010) argue that a second order approximation of $MSE(\hat{\tilde{m}}_{dw})$ is given by

$$MSE(\hat{\tilde{m}}_{dw}) = g_{1dw}(\sigma_u^2, \sigma_e^2) + g_{2dw}(\sigma_u^2, \sigma_e^2) + g_{3dw}(\sigma_u^2, \sigma_e^2) + 2C_{1dw}(\sigma_u^2, \sigma_e^2) + 2\tilde{C}_{2dw}(\sigma_u^2, \sigma_e^2),$$

where $g_{1dw}(\sigma_u^2, \sigma_e^2) = (1-\gamma_{dw})\sigma_u^2$, $g_{2dw}(\sigma_u^2, \sigma_e^2) = (\bar{\mathbf{X}}_d - \gamma_{dw}\bar{\mathbf{x}}_{dw})^T \mathbf{\Phi}_w (\bar{\mathbf{X}}_d - \gamma_{dw}\bar{\mathbf{x}}_{dw})$, $g_{3dw}(\sigma_u^2, \sigma_e^2) = \gamma_{dw}(1-\gamma_{dw})^2 \sigma_e^{-4}\sigma_u^{-2} h(\sigma_u^2, \sigma_e^2)$, with $h(\sigma_u^2, \sigma_e^2) = \sigma_e^4 var(\hat{\sigma_u^2}) + \sigma_u^4 var(\hat{\sigma_e^2}) - 2\sigma_u^2\sigma_e^2 cov(\hat{\sigma}_u^2, \hat{\sigma}_e^2)$ and

$$\begin{aligned}\mathbf{\Phi}_w = {} & \sigma_e^2 \Big(sum_{d=1}^D \sum_{j=1}^{n_d} x_{jd}\mathbf{t}_{jd}^T \Big)^{-1} \Big(sum_{d=1}^D \sum_{j=1}^{n_d} \mathbf{t}_{jd}\mathbf{t}_{jd}^T \Big) \Big\{ \Big(sum_{d=1}^D \sum_{j=1}^{n_d} x_{jd}\mathbf{t}_{jd}^T \Big)^{-1} \Big\} \\ & + \sigma_u^2 \Big(sum_{d=1}^D \sum_{j=1}^{n_d} x_{jd}\mathbf{t}_{jd}^T \Big)^{-1} \Big\{ \sum_{d=1}^{D} \Big(\sum_{j=1}^{n_d} \mathbf{t}_{jd} \Big) \Big(\sum_{j=1}^{n_d} \mathbf{t}_{jd} \Big)^T \Big\} \\ & \times \Big\{ \Big(sum_{d=1}^D \sum_{j=1}^{n_d} x_{jd}\mathbf{t}_{jd}^T \Big)^{-1} \Big\},\end{aligned}$$

with $\mathbf{t}_{jd} = w_{jd}(\mathbf{x}_{jd} - \gamma_{dw}\bar{\mathbf{x}}_{dw})$. Replacing (σ_u^2, σ_e^2) with their REML estimates $(\hat{\sigma}_u^2, \hat{\sigma}_e^2)$ a nearly unbiased MSE estiamtor can be as follows:

$$mse(\hat{\tilde{m}}_{dw}) = g_{1dw}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + g_{2dw}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + 2g_{3dw}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + 2C_{1dw}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + 2\tilde{C}_{2dw}(\hat{\sigma}_u^2, \hat{\sigma}_e^2).$$

As an alternative a double bootstrap procedure can be used to estimate the MSE (Torabi and Rao, 2010).

The MSE of the weighted M-quantile-based small area estimator of the mean, $\hat{\tilde{m}}_d^{WMQ}$, can be approximated analytically following Särndal (1982) and noting that $\hat{\tilde{m}}_d^{WMQ}$ is nearly unbiased:

$$mse(\hat{\tilde{m}}_d^{WMQ}) = \frac{1}{N_d^2} \sum_{j=1}^{n_d} \sum_{k=1}^{n_d} \frac{\pi_{jkd} - \pi_{jd}\pi_{kd}}{\pi_{jkd}} w_{jd}^* \hat{e}_{jd} w_{kd}^* \hat{e}_{kd},$$

where π_{jd} is the inclusion probability of unit j of area d, π_{jkd} is the joint inclusion probability,

$$w_{jd}^* = \mathbf{W}\mathbf{1}_d + \mathbf{W}\mathbf{C}(\bar{\theta}_d)\mathbf{X}(\mathbf{X}^T\mathbf{W}\mathbf{C}(\bar{\theta}_d)\mathbf{X})^{-1} \Big(\sum_{j=1}^{N_d} \mathbf{x}_{jd}^T - \sum_{j=1}^{n_d} w_{jd}\mathbf{x}_{jd}^T \Big),$$

with $\mathbf{1}_d$ the n-vector with jth component equal to one whenever the corresponding sample unit is in area i and is zero otherwise, $\mathbf{C}(\bar{\theta}_d)$ is the diagonal matrix of order n defined by the weights obtained from the iterative reweighted least square algorithm used to fit the design-weighted M-quantile regression coefficient of area d, $\bar{\theta}_d$. The proposed estimator of the MSE is a first-order approximation because it does not take into account both of the variability due to the estimation of $\bar{\theta}_d$ and that associated with the estimation of the weighted regression coefficients. Fabrizi et al. (2014) noted that $mse(\hat{\tilde{m}}_d^{WMQ})$ underestimates the actual $MSE(\hat{\tilde{m}}_d^{WMQ})$, but if the overall sample size (on which estimation of the M-quantile model is based) is at least moderate and the sampling variance of the y_{jd} and $\mathbf{x}_{jd}$ dominates that associated to the uncertainty in estimating θ_d, the underestimation is likely to be small.

Alternative estimators of the design-based MSE of $\hat{\tilde{m}}_d^{WMQ}$ may be obtained using bootstrap methods

(Fabrizi et al., 2014).

3.2.2. Model-based methods for SAE estimation of poverty indicators

The most popular method used for model-based SAE employs linear mixed models. In the general case such a model has the form

$$y_{jd} = \mathbf{x}_{jd}^T \boldsymbol{\beta} + u_d + e_{jd}, \tag{3.15}$$

where u_d is the area-specific random effect and e_{jd} is an individual random effect. The empirical best linear unbiased predictor (EBLUP) of m_d (see Rao, 2003, Chapter 7) is then

$$\hat{\hat{m}}_d^{LM} = N_d^{-1} \Big[\sum_{j \in s_d} y_{jd} + \sum_{j \in r_d} \{ \mathbf{x}_{jd}^T \hat{\boldsymbol{\beta}} + \hat{u}_d \} \Big], \tag{3.16}$$

where $\hat{\beta}$, $\hat{u}_d$ are defined by substituting an optimal estimator for the covariance matrix of the random effects in (3.15) in the best linear unbiased estimator of $\boldsymbol{\beta}$ and the best linear unbiased predictor (BLUP) of u_d respectively. A widely used estimator of the mean squared error (MSE) of the EBLUP is based on the approach of Prasad and Rao (1990). This estimator accounts for the variability due to the estimation of the random effects, regression parameters and variance components.

Assuming model (3.15) on the logarithmically transformed values of income y_{jd}, the most widely used method for small area poverty mapping is the so-called World Bank (WB) or ELL method (Elbers et al., 2003). The model is fitted to clustered survey data from the population of interest, with the random effects in the model corresponding to the cluster used in the survey design. Once the model has been estimated using the survey data, the ELL method uses the following bootstrap population model to generate L synthetic Censuses:

$$y_{jd}^* = \mathbf{x}_{jd}^T \hat{\boldsymbol{\beta}} + u_d^* + e_{jd}^*, \;\; u_d^* \sim N(0, \hat{\sigma}_u^2), \;\; e_{jd}^* \sim N(0, \hat{\sigma}_e^2) \tag{3.17}$$

For each draw, using the synthetic values of the welfare variable y_{jd}^*, values of the poverty indicators of interest for the different small areas are calculated. These are averaged over the L Monte Carlo simulations to produce the final estimates of the poverty quantities, with the simulation variability of these estimates used as an estimate of their uncertainty.

Molina and Rao (2010) point out that when small areas and clustered coincide, in the simplest case of estimating a small area mean, the ELL method leads to a synthetic regression estimator that, in many cases, could be less efficient than the alternative model-based estimators. Molina and Rao (2010) propose a modification of ELL method (the EBP method) introducing random area effects (rather than random cluster effects) into the linear regression model for welfare variable, and also simulated out of sample data by making independent draw from the conditional distribution of the out of sample data, given the sample data.

An alternative approach to EBLUP has been discussed in Chandra and Chambers (2005) and it is based on the use of model-based direct estimation (MBDE) within the small areas. In this case an estimate for a small area of interest corresponds to a weighted linear combination of the sample data for that area, with weights based on a population level version of the linear mixed model. These weights 'borrow strength' via this model, which includes random area effects. Provided the assumed small area model is true, the EBLUP is asymptotically the most efficient estimator for a particular small area. In practice however the 'true' model for the data is unknown and the EBLUP can be inefficient under misspecification. In such circumstances, Chandra and Chambers (2005) note that MBDE offers an alternative to potentially unstable EBLUP. In particular, MBDE is easy to implement, produces sensible estimates when the sample data exhibit patterns of variability that are inconsistent with the assumed model (e.g. contain too many zeros) and generates robust MSE estimates.

A different approach has been proposed in the literature for further robustification of the inference by relaxing some of the model assumptions. This approach is based on M-quantile regression (Breckling and Chambers, 1988). It provides a 'quantile-like' generalization of regression based on influence functions (Breckling and Chambers, 1988). A linear M-quantile regression model is one where the qth M-quantile $Q_q(y_{jd}|\mathbf{x}_{jd})$ of the conditional distribution of y given x satisfies

$$Q_q(y_{jd}|\mathbf{x}_{jd}) = \mathbf{x}_{jd}^T\boldsymbol{\beta}_q \tag{3.18}$$

That is, it allows a different set of regression parameters for each value of q. For specified q and continuous influence function ψ, an estimate $\hat{\boldsymbol{\beta}}_q$ of $\boldsymbol{\beta}_q$ can be obtained via an iterative weighted least squares algorithm.

As said in previous section, extending this line of thinking to SAE, Chambers and Tzavidis (2006) observed that if variability between the small areas is a significant part of the overall variability of the population data, then units from the same small area are expected to have similar M-quantile coefficients. In particular, when (3.18) holds, and $\boldsymbol{\beta}_q$ is a sufficiently smooth function of q, these authors suggest a predictor of m_j of the form

$$\hat{\hat{m}}_d^{MQ} = N_d^{-1}\Big[\sum_{j\in s_d} y_{jd} + \sum_{j\in r_d} \hat{Q}_{\bar{\theta}_d}(y_{jd}|\mathbf{x}_{jd})\Big], \tag{3.19}$$

where $\hat{Q}_{\bar{\theta}_d}(y_{jd}|\mathbf{x}_{jd}) = \mathbf{x}_{jd}^T\hat{\boldsymbol{\beta}}_{\bar{\theta}_d}$ and $\bar{\theta}_d$ is an estimate of the average value of the M-quantile coefficients of the units in area d. Typically this is the average of estimates of these coefficients for sample units in the area. When there is no sample in area, we can form a 'synthetic' M-quantile predictor by setting $\bar{\theta}_d = 0.5$. Tzavidis et al. (2010) refer to (3.19) as the 'naïve' M-quantile predictor and note that it can be biased, therefore, they proposed a bias adjusted M-quantile predictor of m_d.

The M-quantile small are model are used also for estimating the poverty indicators as HCR and PG (Marchetti et al., 2012) by using a smearing-type estimator (Duan, 1983). A small area estimator of the HCR is obtained as:

$$\hat{F}_d(0,t) = N_d^{-1}\Big[\sum_{j\in s_d} f_{jd}(0,t) + \hat{E}[f_{jd}(0,t)]\Big] \tag{3.20}$$

where

$$\hat{E}[f_{jd}(0,t)] = \int I(\mathbf{x}_{jd}^T\hat{\boldsymbol{\beta}}_{\bar{\theta}_d} + \hat{e}_{jd} \le t)d\hat{F}(\hat{e}) = n^{-1}\sum_{k\in r_d}\sum_{j\in s_d} I(\mathbf{x}_{kd}^T\hat{\boldsymbol{\beta}}_{\bar{\theta}_d} + \hat{e}_{jd} \le t)$$

with the distribution function estimated as $\hat{F}(\hat{e}) = n^{-1}\sum_{j=1}^{n} I(\hat{e}_j \leq e)$. The same approach can be used to estimate the PG indicator or any other of the FGT poverty measures.

Both the EB and M-quantile approaches can be extended to include two nested sources of variability using a three-level modelling approach. The superpopulation model underpinning these two approaches is as follows:

$$y_{jid} = \mathbf{x}_{jid}^T\beta + u_d + v_{id} + w_{jid}^{-1/2}e_{jid}, \quad j = 1, \ldots, N_{id}, i = 1, \ldots, M_d, d = 1, \ldots, D.$$

Here, M_d clusters are nested in the small area d, for example municipalities nested in a province. When poverty figures are sought at two different aggregation levels, domains and subdomains, it is reasonable to assume a twofold nested error model including random effects explaining the heterogeneity at the two levels of aggregation.

Marhuenda et al. (2017) assume y is a transformation of a welfare variable W such that $h(W) = y \sim N$, then they assume also $u_d \sim N(0, \sigma_u^2)$, $v_{id} \sim N(0, \sigma_v^2)$ and $e_{jid} \sim N(0, \sigma_e^2)$, with u, v, e mutually independent. The unit level error can be heteroscedastic when $w_{jid} \neq 0$ for some units. Under these assumptions Marhuenda et al. (2017) derive an EB estimator for poverty indicators and its predictor (EBP). When the assumptions of Marhuenda et al. (2017) are violated the EBP can be suboptimal. Therefore, Marchetti et al. (2018) proposed a robust alternative to prediction based on an extension of M-quantile regression model, which can account for two nested source of variability.

Among the other issues we focus here on benchmarking, the excess of zero values in the data and the treatment of geographic information.

Model based estimators usually do not have the *benchmarking* property under a complex sampling design. Given a small area estimator Y_i, that does not show the benchmarking property, a first simple way of achieving benchmarking is by a ratio or a difference type adjustment. These approaches have been developed mainly using area level models for small area estimation.

Rao and Molina (2015) present a review of the approaches that can be used to modify the EBLUPs of the area means to obtain the benchmarking property. Externally benchmarked predictors are obtained through an a-posteriori adjustment of model-based predictors. A first possibility are the simple adjustment of the area EBLUPs leading to ratio or difference benchmarking, which however have several limitations. “Optimal” externally restricted benchmarked estimators of small area mean can be instead be obtained following the approach by Pfeffermann and Barnard (1991), Wang et al. (2008) or by Datta et al. (2011), who also show an application to the estimation of poverty rates.

Two-stage benchmarking is a relevant approach when the population have a hierarchical structure consisting of areas and subareas within areas, and it is of interest to estimate both area means and subarea means. This can be a relevant issue in poverty studies. Ghosh and Steorts (2013) extended the method of Datta et al. (2011) to two-stage benchmarking.

Self-benchmarking, originally proposed in the context of the basic unit level model You and Rao

(2002) has been extended to area-level models by Wang et al. (2008). However, it has been noted that self-benchmarking does not necessarily protect against misspecification of the linking model. It is often claimed that adjusted benchmarked estimators provide robustness to model failure. Pfeffermann (2013) provide some support for this claim in the time series context.

Pfeffermann et al. (2014) provide a good review and study the properties of single-stage cross-sectional and time series benchmarking procedures that have been proposed in the literature in the context of small area estimation, and also review cross-sectional methods proposed for benchmarking hierarchical small areas and develop a new two-stage benchmarking procedure for hierarchical time series models. The author present an application for the estimation of unemployment rates.

Schmid et al. (2017) use a benchmark approach to achieve the internal consistency with the direct estimator at national level in the estimation of the share of illiterates disaggregated by gender in Senegal. The authors consider a benchmarked transformed FH estimator following the approach by Datta et al. (2011) and they apply an inverse sine transformation to restrict the share of literates in each area within the interval $[0, 1]$. Therefore, the work of Schmid et al. (2017) can also be of interest for the estimation of poverty indicators.

Considering unit level models, as for EB predictors, also M-quantile regression based estimators do not fulfil the benchmarking property. Fabrizi et al. (2012) propose a modification of the M-quantile predictors estimation algorithm to obtain MQ benchmarked estimates of the small area mean.

Many variables of interest in economics surveys on poverty and living conditions are semicontinuous in nature, i.e. they either take a single fixed value (typically 0, zero) or they have a continuous, often skewed, distribution on the positive real line. They present an *excess of zero-values.* A semicontinuous variable is quite different from one that has been left- censored or truncated, because the zeros are valid self-representing data values, not proxies for negative or missing responses.

A two part random effects model (Olsen and Schafer, 2001) is widely used for small area estimation with zero-inflated variables, see for example, Pfeffermann et al. (2008) and Chandra and Sud (2012). Chambers et al. (2014) propose a small area estimation method for semicontinuous variables under a two part random effects model.

In poverty studies observations that are spatially close may be more alike than observations that are further apart. One approach for incorporating spatial information in *spatial modelling* and in a small area regression model is to assume that the model coefficients themselves vary spatially across the geography of interest and/or the random effects of the model be correlated. Both EBLUP predictors and MQ predictors can be extended to include the effect of the spatial characteristics of the data.

When geography is included as an auxiliary information in modelling, the spatial correlation and the consequent correlation between the random effect in the EBLUP model require the extension of EBLUP estimator to the SEBLUP estimator (Petrucci and Salvati, 2006, Pratesi and Salvati, 2009) when the spatial process is stationary, and to the geographical weighted empirical best linear unbiased predictor

(GWEBLUP) when the spatial process is non-stationary (Chandra et al., 2012).

Under the MQ approach the reference to the Geographically Weighted Regression (GWR) (Brunsdon et al., 1996) helps in modelling spatial variation. This uses local rather than global parameters in the regression model. That is, a GWR model assumes spatial non-stationarity of the conditional mean of the variable of interest. Salvati et al. (2012) propose an M-quantile GWR model, i.e. a local model for the M-quantiles of the conditional distribution of the outcome variable given the covariates. This approach is semi-parametric in that it attempts to capture spatial variability by allowing model parameters to change with the location of the units, in effect by using a distance metric to introduce spatial non-stationarity into the mean structure of the model. The model is then used to define a predictor of the small area characteristic of interest. As a consequence, it integrates the concepts of bias-robust small area estimation and borrowing strength over space within a unified modelling framework. By construction, the model is a local model and so can provide more flexibility in SAE, particularly for out of sample small area estimation, i.e. areas where there are no sampled units.

When studying the spatial distribution of local poverty indicators obtained by SAE methods, it can be relevant to consider the possible effect of *the modifiable areal unit problem* (MAUP). This last is a source of statistical bias that can radically affect the results of statistical analysis. It affects results when point-based measures of spatial phenomena (e.g. population density) are aggregated into larger areas. The resulting summary values (e.g. totals, rates, proportions) are influenced by the choice of the areas's boundaries. For example, point-based census or survey data may be aggregated into census enumeration districts, or post-code areas, or any other spatial partition (thus, the 'areal units' are 'modifiable').

The topic has not yet been treated explicitly in the current literature on SAE. The only empirical study is due to Pratesi and Petrucci (2014) who study the scale effect on SAE predictors by a simulation experiment. They provide evidence to assess the robustness of SAE methods to different scale of aggregation of the point-based measures inside the pre-defined small areas (domains) of interest. The rationale of this simulation study is to verify to what extent we can aggregate the individual values inside the small areas and still have an acceptable accuracy of the estimate of the small area parameter. Under this simulation experiment, methods that are naturally robust to outliers and not linked to distributional assumption on the study variable as M-quantile methods perform better than the alternative methods for SAE and to be resilient to changing scale of analysis. This is likely due to the fact that the changes in geography do not affect the M-quantile coefficients at area level.

MSE estimation of model-based SAE methods

The traditional analytical estimator of the MSE of $\hat{\bar{m}}_d^{LM}$ has been proposed by Prasad and Rao (1990):

$$mse(\hat{\bar{m}}_d^{LM}) = g_{1d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + g_{2d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) + 2g_{3d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2),$$

where the leading term $g_{1d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) = \hat{\gamma}_d(\hat{\sigma}_e^2/n_d)$, $g_{2d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) = (\bar{\mathbf{X}}_d - \hat{\gamma}_d\bar{\mathbf{x}}_d)^T\left(\sum_{d=1}^D \mathbf{X}_d^T\mathbf{V}^{-1}\mathbf{X}_d\right)^{-1}(\bar{\mathbf{X}}_d - \hat{\gamma}_d\bar{\mathbf{x}}_d)$ and $g_{3d}(\hat{\sigma}_u^2, \hat{\sigma}_e^2) = n_d^{-2}(\hat{\sigma}_u^2 + \hat{\sigma}_e^2/n_d)^{-4}(\hat{\sigma}_e^2\bar{\mathbf{V}})_{uu} + \hat{\sigma}_u^2\bar{\mathbf{V}})_{ee} - 2\hat{\sigma}_u^2\hat{\sigma}_e^2\bar{\mathbf{V}}_{eu})(\bar{y}_d - \bar{\mathbf{x}}_d^T\hat{\beta})^2$, with $\bar{\mathbf{V}}_{ee}$ an estimator of the asymptotic variances of $\hat{\sigma}_e$, $\bar{\mathbf{V}}_{uu}$ an estimator of the asymptotic variances of $\hat{\sigma}_u$ and

$\bar{\mathbf{V}}_{eu}$ an estimator of the asymptotic covariance between $\hat{\sigma}_e$ and $\hat{\sigma}_u$.

The MSE of the unit level EBLUP has been discussed widely in the literature. Among others, Datta and Lahiri (2000) extended the estimation of Prasad and Rao (1990) to the more general mixed linear model, Das et al. (2004) extend the model of Datta and Lahiri (2000) by relaxing the assumption of independence of the error terms between the areas. See Datta (2009) for an extensive review of methods of estimating the MSE of the EBLUP and EB under linear mixed models. Moreover, Chambers et al. (2011) develop conditional bias-robust MSE estimators for the case where the small area estimators can be expressed as weighted sums of sample values. One of these case is the estimator $\hat{\bar{m}}_d^{MQ}$. Militino et al. (2007) developed an unbiased MSE estimator for the EBLUP in the case where the sampling fraction is *non* negligible.

Also resampling methods have been proposed in literature. Among others Hall and Maiti (2006a) used a matched-moment bootstrap method to get around the difficulty with the linearization method (see also Hall and Maiti (2006b)). Erciulescu and Fuller (2014) give some advice on how to manage the replication in the double bootstrap of Hall and Maiti (2006a).

An analytic estimator for $\hat{\bar{m}}_d^{MQ}$ according to Chambers et al. (2011) is as follows:

$$mse(\hat{\bar{m}}_d^{MQ}) = N_d^{-2} \sum_{j=1}^{n_d} \{a_{jd}^2 + (N_d - n_d)n^{-1}\} \hat{\lambda}_{jd}^{-1} (y_{jd} - \mathbf{x}_{jd}^T \hat{\beta}_{\bar{\theta}_d}),$$

where $a_{jd} = N_d w_{jd} - I(j \in d)$, $\hat{\lambda}_{jd} = (1 - \phi_{jjd})^2 + \sum_{k=1,-j}^{n_d} \phi_{kjd}^2$ with ϕ_{kjd} such that $\sum_{k=1}^{n_d} \phi_{kjd} y_{kd} = \mathbf{x}_{jd}^T \hat{\beta}_{\bar{\theta}_d}$.
The weights w_{jd} here is the jth element of the vector
$\mathbf{w}_d = n_d^{-1} \boldsymbol{\Delta}_d + (1 - N_d^{-1} n_d) \mathbf{C}(\bar{\theta}_d) \mathbf{X} (\mathbf{X}^T \mathbf{C}(\bar{\theta}_d) \mathbf{X})^{-1} (\bar{\mathbf{x}}_{r_d} - \bar{\mathbf{x}}_{s_d})$, with $\boldsymbol{\Delta}_d$ is the vector of size n that "picks out" the sampled units in area d. An alternative estimator of the MSE is based on first-order approximations to the variances of solutions of robust estimating equations, as proposed in **?**.

We now focus on specific method for poverty mapping (i.e. estimating poverty incidence and intensity), namely the ELL, the EBP (one- and two-folds), the MQ (one- and two-folds) and the MBDE.

The ELL method decompose the total prediction error into three parts: idiosyncratic error, model error and computation error. However, if one is not interested in each of the component, the total prediction error can be estimated drawing from the sampling distribution of the parameters; more details in Elbers et al. (2003).

To estimate the MSE of target $F_d(\alpha, t), \alpha = \{0, 1, 2\}$ under the EBP approach, Molina and Rao (2010) proposed a parametric bootstrap that is an alternative to the double bootstrap method of Hall and Maiti (2006), which could provide a better MSE estimator in terms of relative bias, but for large populations this method might not be computationally feasible. Also for the EBP under the two-folds mixed model the MSE of target estimator is obtained using parametric bootstrap. The choice between the two-folds or one-fold (traditional) model can be based on the results of testing for the significance of the variances at each level or by comparing the models on the basis of goodness-of-fit measures. However, from the simulation study in Marhuenda et al. (2017), which included the cases where only the domain or the subdomain effects are significant, which is equivalent to a case where covariates explain all the between-domain or between-subdomain variation, they have learnt that there is virtually no loss of efficiency by preserving a random factor even if it is not significant.

The estimate of the MSE of M-quantile based poverty measures as in equation (3.20) can be obtained according to the non-parametric bootstrap technique proposed in Marchetti et al. (2012). Moreover, the authors noted that using the proposed bootstrap technique the second-order propertied of the MSE estimator for the M-quantile mean estimator are better than the analytic one proposed in Chambers et al. (2011). The bootstrap technique in Marchetti et al. (2012) has been extended in Marchetti et al. (2018) to estimate the MSE of the M-quantile poverty estimator under the two-fold M-quantile approach.

To estimate the MSE of the MBDE Chandra and Chambers (2009) used the work of Royal (1976) that shows that among linear prediction unbiased estimators the variance of the prediction error is minimized by weights of a specified form.

The estimators presented here that use spatial information are the SEBLUP and the M-quantile GWR. Spatial small area estimators can be classified according to spatial stationary or not. When the spatial process is stationary, then the MSE of the spatial EBLUP (under unit-level approach) can be obtained analytically or by bootstrap. Chandra, Salvati, and Chambers (2007) studied the unit level model with spatially correlated random effects following a SAR model. They developed EBLUP estimators and associated MSE estimators using the Taylor linearization approach. When the spatial process is not stationary, (Chandra et al., 2012) propose the geographical weighted empirical best linear unbiased predictor (GWEBLUP). Its MSE estimator is based on the extension of the approach of Chambers et al. (2011) to estimating the conditional MSE of GWEBLUP. This approach is motivated by re-expressing the GWEBLUP in a pseudo-linear form, i.e. as a weighted sum of the sample values of y, and then applying heteroscedasticity-robust prediction variance estimation methods that treat these weights, which typically depend on estimated variance components, as known.

Finally, the MSE of the M-quantile GWR based estimators can be estimated following the analytic method based on pseudo-linear form of the estimator proposed in Salvati et al. (2012).

3.2.3. Quality issues for auxiliary information in SAE models

The use of SAE modelling for the estimation of local indicators is a growing area of interest both for researchers and NSIs. Examples include the estimation of unemployment and poverty rates, that for policy reasons may require a finer geographical level with respect to that guarantee by direct estimates. However, it is important to ensure that the adopted or developed methods are appropriate for the actual users' needs. Tzavidis. et al. (2018) present a framework for the production of small area estimates in official statistics. They cover and discuss all the relevant step that should be followed when developing a SAE model: specification, analysis/adaptation, evaluation.

As Tzavidis. et al. (2018) discuss, in its essence small area estimation is about the efficient combination of sample survey and auxiliary information, whose quality is essential. Several issues related to the quality of the auxiliary information can impact on the final SAE estimates. First of all, as the authors underline, is the complexity of the targets of estimation that determines the data requirements for SAE. If the interest is in using area level models to estimate domain level totals, there is no need to have access to covariate microdata. However, when the target is in the estimation of non-linear indicators, as for many poverty indicators, access to census or administrative microdata may be needed. This can be a relevant issue when access to microdata is not possible for privacy and confidentiality constrains. Even when the auxiliary covariates are available, several quality issues arise, as underlined by Fabrizi in the discussion of the paper by Tzavidis. et al. (2018):

- auxiliary variables should be measured consistently in the survey and auxiliary dataset (e.g. in the census);
- linking survey and administrative archives can be non-trivial and prone to linkage errors, unless a unique, error-free, identifier is available;
- the frequency of updates of the auxiliary information is crucial, as the use of not regularly updated covariates can lead to biased estimation in non-census years;
- the predictive power of auxiliary information plays an important role in improvements in efficiency.

When instead census or administrative data are not available, one possibility is to assume a model for the observed covariates and to impute the missing values from that model (Sverchkov and Pfefferman, 2018). Pfefferman and Sikov (2011) developed a non-parametric alternative. Another possibility is to use alterative sources of data in SAE modelling. Indeed, in the last year there has been a growing interest in the use of alternative sources of data as covariates in small area estimation methods. The use of new data sources in SAE modelling can be relevant also when the interest is in estimating at the local level poverty indicators, has it has been already discussed in MAKSWELL Deliverable 2.2 (van den Brakel et al., 2019).

From a methodological point of view, the use of auxiliary information not coming from census has many implications since the standard SAE methodology, for example the Fay-Herriot model, requires that the covariates are measured without error. This can be an important limitation in many situations where updated census information is not available (see the discussion in MAKSWELL Deliverable 2.2). Therefore, in the last years several alternative SAE estimators have been proposed to be able to deal with the use of covariates measured with error in SAE models. These methods can be used to deal with many of the covariates quality issues listed above.

When auxiliary information is measured with error, an estimator accounting for the measurement error in the covariates has been proposed in Ybarra and Lohr (2008). In their seminal paper the authors suggest a suitable modification to the Fay-Herriot estimator that accounts for sampling variability in the auxiliary information, and derive its properties, in particular showing that it is approximately unbiased. Marchetti et al. (2015) show an application of the Ybarra-Lohr estimator to measure the ARPR of small areas in Italy using covariates coming from the EU-SILC survey and from a GPS tracking system (see also MAKSWELL Deliverable 2.2). From a methodological point of view, following the work by Ybarra and Lohr (2008), Arima et al. (2015) rewrite the measurement error model as a hierarchical Bayesian model, developing a Bayesian treatment of the univariate Fay-Herriot model when some of the covariates are measured with error.

Arima and Polettini (2019) discuss the same issue of measurement error in the covariates but considering unit-level small area models and the situation when covariates subject to measurement error are of categorical nature. Adopting a Bayesian approach, the authors extend the unit-level model in order to account for measurement error in both continuous and categorical covariates. For discrete covariates, measurement error means misclassification. It is interesting to note that misclassification error may be artificially induced for disclosure limitation purposes by NSIs (Polettini and Arima, 2015).

Arima et al. (2017) present a Bayesian analysis of a multivariate Fay-Herriot model with functional measurement error, allowing for both joint modelling of related characteristics and accounting for random observation error in some of the covariates. The paper provides a multivariate generalization of the approach of Arima et al. (2015). Arima et al. (2017) apply the proposed estimator to modelling 2010 and 2011 poverty rates of school-aged children for US counties, for predicting 2011 poverty rates and the 2010–2011 changes.

In line with the previous works, Burgard et al. (2019) introduce a three-stage FH model by assuming that the vector of true domain means of auxiliary variables differs from the corresponding vector of direct estimators in a zero-mean multivariate normally distributed random error. The authors consider a functional measurement error model that can be considered as the adaptation of the Ybarra-Lohr model to a parametric inference setup with normally distributed measurement errors. They present an application to estimate poverty proportions in the Spanish Living Condition Survey with auxiliary information from the Spanish Labour Force Survey.

Further contributions to the literature on measurement error in SAE models were provided by Ghosh et al. (2006), Ghosh and Sinha (2007), Torabi et al. (2009), Datta et al. (2010). Borssoi et al. (2017) present a new approach to incorporate measurement errors in mixed models with elliptical errors. Pratesi (2015) present a review on missing spatial information in SAE modelling.

3.3. Data sets in use for indicators of poverty and well-being

The indicators of well-being presented in section 2 are mainly measured using data from major sample surveys. The main source for the Eurostat is the EU-SILC survey, covering the dimensions of Material living conditions, Leisure and social interactions and Economic security and physical safety. The EU-SILC is also the main source for all the subjective indicators of Quality of Life: in this case the ad-hoc modules of the survey are of fundamental importance to get data on subjective aspects concerning Leisure and Social Interaction, Economic security and physical safety, Natural and living environment. Another important survey is the Labour Force Survey (LFS), that is the main data source for the Productive or other main activity and Education dimensions. The dimension Governance and basic rights is measured using indicators both from the EU-SILC and the LFS.

The tables below show the specific data source for each of the Eurostat Quality of Life indicators. As we can see, the EU-SILC and LFS are complemented in some dimensions by other surveys such as the Household Budget Survey (HBS), the European Health Interview Survey (EHIS), the Adult Education Survey (AES). Some of these surveys are not carried out every year, so that the corresponding indicators are updated less frequently than those measured using the main EU-SILC and LFS core surveys.

Also for the OECD well-being indicators, sample surveys are the main data source. In this case, due to the broader geographical context, harmonized international surveys are used. For subjective indicators the Gallup World Poll is the main source.

Table 3.1: Eurostat objective Quality of Life indicators by dimension, topic and sub-topic. Dimensions: Material living conditions, productive or other main activity, health

	Material living conditions	
Topic	Indicator	Source
Income	Median disposable equivalised income	EU-SILC
	Income inequality (S80/S20 income quintile ratio)	EU-SILC
	At-risk-of-poverty rate	EU-SILC
	At-risk-of-poverty rate anchored at fixed moment in time	EU-SILC
Consumption	Actual individual consumption (per capita)	National Accounts
	Basic expenses in the total household expenditure	HBS
Material conditions	Severe material deprivation rate	EU-SILC
	(In)ability to make ends meet	EU-SILC
	Structural problems of the dwelling	EU-SILC
	Space in the dwelling (overcrowding/under-occupation)	EU-SILC
	Productive or other main activity	
Topic	Indicator	Source
Quantity of employment	Employment rate	EU-LFS
	Unemployment rate	EU-LFS
	Long-term unemployment rate	EU-LFS
	People living in households with very low work intensity	EU-SILC
	Underemployed part-time workers	EU-LFS
Quality of employment	Low-wage earners	The Structure of Earnings Survey (SES)
	Temporary contracts	EU-LFS
	Involuntary temporary contracts	EU-LFS
	Over-qualification rate	EU-LFS
	Incidence rate of fatal accidents at work	European statistics on accidents at work (ESAW)
	Average number of usual weekly hours of work	EU-LFS
	Long working hours (more than 48 per week)	EU-LFS
	Atypical working hours (usual work during evenings, nights, Saturdays or Sundays)	EU-LFS
	Flexibility of the work schedule	EU-LFS ad hoc modules
Main reason for economic inactivity	Inactive population by reason of inactivity	EU-LFS
	Health	
Topic	Indicator	Source
Outcomes	Life expectancy at birth	Demographic data,
	Healthy Life Years	Demographic data and EU-SILC
Determinants	Body Mass Index	EHIS (not every year)
	Daily smokers	EHIS (not every year)
	Hazardous alcohol consumption	EHIS (not every year)
	Practice of physical activity	EHIS (not every year)
	Consumption of fruits and vegetables	EHIS (not every year)

Table 3.2: Eurostat objective Quality of Life indicators by dimension, topic and sub-topic. Dimensions: Education, leisure and social interactions, economic security and physical safety, governance and basic rights, natural and living environment

	Education	
Topic	Indicator	Source
Competences and skills	Educational attainment	EU-LFS
	Early leavers from education and training	EU-LFS
	Mean literacy proficiency score	PIAAC (not every year)
Lifelong learning	Participation in adult education and training	EU-LFS
Opportunities for education	Participation in education of children four-year-olds	Administrative data collection on education
	Leisure and social interactions	
Topic	Indicator	Source
Leisure	Non-participation in culture or sport activities	EU-SILC ad hoc modules
	Financial obstacles to leisure participation	EU-SILC ad hoc modules
Social interactions	Frequency of getting together with friends	EU-SILC ad hoc modules
	Participation in formal voluntary activities	EU-SILC ad hoc modules
	Participation in informal voluntary activities	EU-SILC ad hoc modules
	Trust in others	EU-SILC ad hoc modules
	Perception of social inclusion	EU-SILC ad hoc modules
	Economic security and physical safety	
Topic	Indicator	Source
Economic security	Population unable to face unexpected financial expenses	EU-SILC
	Population in arrears	EU-SILC
	Perc. of employed in the previous year transitioning to unempl.	EU-SILC
Physical safety	Homicide rate	Data collected by Eurostat
	Governance and basic rights	
Topic	Indicator	Source
Discrimination and equal opportunities	Gender employment rate gap	EU-LFS
	Gender pay gap	Structure of Earnings Survey (not every year)
	Gap in employment rates between nationals and non-EU citizens	EU-LFS
Active citizenship	Active citizenship	EU-SILC ad hoc modules
	Natural and living environment	
Topic	Indicator	Source
Pollution (including noise)	Urban pop. exposure to air pollution (PM10)	European Environment Agency (EEA)

Table 3.3: Eurostat subjective Quality of Life indicators by dimension, topic and sub-topic.

Material living conditions		
Topic	Indicator	Source
Material conditions	Satisfaction with accommodation	EU-SILC
Productive or other main activity		
Topic	Indicator	Source
Quality of employment	Self-reported over-qualification	EU-LFS ad hoc modules
Quality of employment	Satisfaction with commuting time	EU-SILC ad hoc modules
	Job satisfaction	EU-SILC ad hoc modules
Health		
Topic	Indicator	Source
Outcomes	Self-perceived health	EU-SILC
	Self-reported mental health	EHIS (not every year)
Access to healthcare	Unmet needs for medical care	EU-SILC
Education		
Topic	Indicator	Source
Competences and skills	Individuals' level of internet (digital) skills	Survey on ICT (Information and Communication Technologies)
	Population reporting not to know any foreign language	Adult Education Survey (not every year)
	Level of best known foreign language	Adult Education Survey (not every year)
Leisure and social interactions		
Topic	Indicator	Source
Leisure	Satisfaction with time use	EU-SILC ad hoc modules
Social interactions	Satisfaction with personal relationships	EU-SILC ad hoc modules
	Help from others (having someone to rely on in case of need)	EU-SILC ad hoc modules
	Having someone to discuss personal matters with	EU-SILC ad hoc modules
Economic security and physical safety		
Topic	Indicator	Source
Physical safety	Perception of crime, violence or vandalism in the living area	EU-SILC ad hoc modules
	Safety feeling (pop. feeling safe when walking alone after dark)	EU-SILC ad hoc modules
Governance and basic rights		
Topic	Indicator	Source
Trust in institutions	Trust in the legal system, the political system and the police	EU-SILC ad hoc modules
Natural and living environment		
Topic	Indicator	Source
Pollution (including noise)	Perception of pollution, grime or other environmental problems	EU-SILC ad hoc modules
	Noise from neighbours or from the street	EU-SILC ad hoc modules
Access to green and recreational spaces	Satisfaction with recreational and green areas	EU-SILC ad hoc modules
Landscape and built environment	Satisfaction with living environment	EU-SILC ad hoc modules
Overall experience of life		
Topic	Indicator	Source
Life satisfaction	Overall life satisfaction	EU-SILC ad hoc modules, then yearly after 2018
Affects	Negative affects (being very nervous; feeling down in the dumps; etc.)	EU-SILC ad hoc modules, then yearly after 2018
	Positive affects (being happy)	EU-SILC ad hoc modules, then yearly after 2018
Meaning and purpose of life	Assessing whether life is worthwhile	EU-SILC ad hoc modules, then yearly after 2018

Table 3.4: OECD well-being indicators: data sources for the current well-being domain.

	Material living conditions	
Dimension	Indicator	Suorce
Income and wealth	Household net adjusted disposable income per person	National Accounts data
	Household financial net wealth per person	National Accounts data
Jobs and earnings	Employment rate	Labour Force Statistics database;
	Long-term unemployment rate	Labour Force Statistics database;
Housing	Number of rooms per person	EU-SILC
	Dwelling with basic facilities	EU-SILC
Quality of life Health status	Life-expectancy at birth	OECD Health Data
	Self-reported health status	EU-SILC
Work and Life	Employees working very long hours	OECD Babies and Bosses.Reconciling Work and Family life
	Time devoted to leisure and personal care	OECD Time-use survey database
	Employment rate of women with children	OECD Family database;
Education and Skills	Educational attainment	OECD Education at a glance
	Students' cognitive skills	OECD PISA
Social connections	Contacts with others	EU-SILC
	Social network support	Gallup World Poll
Civic engagement and Governance	Voter Turn-out	International Institute for Democracy and Electoral Assistance (IDEA)
	Consultation on rule-making	OECD Government at a Glance
Environmental Quality	Air pollution	World Bank
Personal Security	Intentional homicides	UNODC
	Self-reported victimization	Gallup World Poll
Subjective Well-being	Life-satisfaction	Gallup World Poll

4. Data and methods in practice

4.1. Data and methods for poverty and welfare at CBS

Poverty and welfare indicators in the Netherlands are derived from household income, which in its turn is completely derived from registrations. Since all the required information is available from several registers, a complete enumeration of the population is possible. In the past, however, the IT infrastructure was insufficient to produce timely income statistics based on a complete enumeration of the Dutch population. Therefore the income distributions were derived from a panel, which was based on a sample of about 100,000 persons. Since 2011 Statistics Netherlands has access to the complete register on income components. From that year on, income statistics are based on a complete enumeration. In Subsection 4.1.1 it is explained how household income is defined and from which data sources this income is derived. In Subsection 4.1.2 the different poverty indicators used in the Netherlands are described. In Subsection 4.1.3 the methodology used to compute poverty indicators during the period that a sample was in place, is briefly summarised.

4.1.1. Data sources for income

Disposable income of a household is defined as the total income of all household members minus paid premiums and taxes. More precise, disposable household income contains:

- gross income (income from work of all household members plus income out of wealth) minus:
  - current transfers paid (alimony payment to ex-partner),
  - income insurance premiums,
  - health insurance premiums,
  - tax on income and wealth
- gross transfers received, i.e. received unemployment benefits, social benefits, pension benefits, child benefits, rent allowance and care allowance

Income from work includes wages and salaries of employees plus the attributed pay for self-employed people and family members working in the family business. Income from wealth includes the sum of the income from financial assets, income from real estate, and income from other property, minus interest paid. The aforementioned income components are derived from registrations. Received or paid child alimony and parental contributions for children living away from home are not observed and excluded from the disposable income definition.

The household composition is derived from the Dutch population register in combination with the information from tax administrations and the Labour Force Survey. The Dutch population register is an accurate list of all residents in the country, since Dutch citizens are required by law to report changes in their demographics to their municipalities.

Finally the standardised household income is derived, which is defined as the disposable household income corrected for differences in household size and composition. In the literature, this is also known as the equivalised spendable income. In this way household income is corrected for the advantages of scale due to running a multi-person household.

4.1.2. Poverty indicators in the Netherlands

Statistics Netherlands uses three definitions for poverty indicators:

1. Poverty based on low-income threshold
2. Poverty based on social minimum threshold
3. Poverty based on the European poverty threshold

The most important poverty indicator published by Statistics Netherlands is based on the low-income threshold. This is the social security level for a single person household in 1979 corrected for price inflation. Statistics are published on an annual frequency. The poverty rate for year t and region j is now defined as:

$$P_{j,t}^{LI} = \frac{\sum_{i=1}^{N_{j,t}} \delta_{i,j,t}^{LI}}{\sum_{i=1}^{N_{j,t}} \delta_{i,j,t}^{I}} \tag{4.1}$$

with

- $N_{j,t}$number of households residing in region j at 1 January year t
- $\delta_{i,j,t}^{LI}$ an indicator variable equal to 1 if household i in region j and year t is no student household and if its equivalised disposable household income is below the low-income threshold of year t and zero otherwise
- $\delta_{i,j,t}^{I}$ an indicator variable equal to 1 if household i in region j and year t is no student household and receives an income during the entire year t and zero otherwise

The main advantage of this poverty indicator is that households with different compositions are classified as poor if their purchasing power drops below a level that is comparable for all groups and that poverty rates based on this definition are comparable over time sine the purchasing power of the low-income thresholds is comparable over time. In order to measure long term poverty, there is a related definition that calculates the rate of households having an income that is below the low-income threshold for at least four subsequent years. This poverty rate is defined as in (4.1) and $\delta_{i,j,t}^{LI}$ defined as an indicator variable equal to 1 if equivalised disposable household income of household i in region j and year $t-q$ is below the low-income threshold of year $t-q$ for q= 0, 1, 2, and 3 or longer and zero otherwise.

Statistics Netherlands also produces poverty rates that are based on the social minimum threshold, as determined in political decision making. The social minimum threshold deviates between households with different compositions. The poverty rate for the social minimum threshold for year t and region j is now defined as:

$$P_{j,t}^{PT} = \frac{\sum_{i=1}^{N_{j,t}} \delta_{i,j,t}^{PT}}{\sum_{i=1}^{N_{j,t}} \delta_{i,j,t}^{I}} \tag{4.2}$$

with $\delta^{PT}_{i,j,t}$ an indicator variable equal to 1 if disposable household income of household i in region j and year t is below the social minimum threshold that applies to that particular household of year t and zero otherwise. In this definition *disposable household income* instead of *equivalised disposable household income* is used. To correct for scale benefits due to running a multi-person household different thresholds apply to different types of households. Nevertheless the level of purchasing power that corresponds to the different social minimum thresholds are larger compared to the low income threshold. Also the purchasing power of the social minimum thresholds for a particular group deviates over time, which hampers temporal comparisons of poverty rates. These figures are important for the implementation of the municipal poverty policy. Note that Statistics Netherlands increases the social minimum threshold with 1% in order to ensure that households receiving social benefits with little extra income are still classified as poor.

Also for this poverty definition there is a poverty rate for households with an income that is below the social minimum threshold for at least the last four years to measure long term poverty. Similarly to the low-income threshold, this poverty rate is defined as in (4.2) with $\delta^{PT}_{i,j,t}$ defined as an indicator variable equal to 1 if disposable household income of household i in region j in year $t-q$ is below the social minimum threshold that applies to that household type of year $t-q$ for $q=$ 0, 1, 2, and 3 or longer and zero otherwise.

Finally poverty rates are calculated based on the European poverty threshold which is defined as 60% of the median equivalised disposable household income for that particular year. This threshold is recalculated each year and therefore follows price as well as welfare developments. The poverty rate for the European poverty threshold for year t and region j is now defined as:

$$P^E_{j,t} = \frac{\sum_{i=1}^{N_{j,t}} \delta^E_{i,j,t}}{\sum_{i=1}^{N_{j,t}} \delta^I_{i,j,t}} \tag{4.3}$$

with $\delta^E_{i,j,t}$ an indicator variable equal to 1 if *equivalised disposable household income* of household i in region j and year t is below the 60% of the median equivalised disposable household income for that particular year t and zero otherwise. To measure long term poverty, there is also the four-year European poverty rate which is defined as in (4.3) with $\delta^E_{i,j,t}$ an indicator variable equal to 1 if equivalised disposable household income of household i in region j and year $t-q$ is below the 60% of the median equivalised disposable household income for that particular year $t-q$ for $q=$ 0, 1, 2, and 3 or longer and zero otherwise.

4.1.3. Methodology for poverty

This section briefly describes the sample design and weighting procedures that were used for the household panel until 2011 to produce income and poverty statistics. As mentioned before, since 2011, statistical information on poverty and welfare is based on a complete enumeration of the Dutch population.

Households are often considered as the sampling units in panels conducted to collect information at the level of households and persons. Such panels are used for longitudinal analysis as well as the production of cross-sectional estimates. Using households as the sampling units in a panel design has, however, some major disadvantages due to their instability over time. As time proceeds, households might disintegrate, join or split, new members might enter the households and other members might leave the households for different reasons. These changes can affect the selection probabilities of the households in the sample. Reconstruction of the correct inclusion probabilities of the sampling units is essential to derive correct weights for analysis purposes, in particular if the panel is used for producing cross-sectional estimates.

To avoid the problems with panels using households as sampling units, an alternative design was applied. Instead of households, so-called core persons are drawn with an equal probability design, who are followed over time. All household members belonging to the household of a core person at each particular period are included in the sample. This results in a sample design where households are drawn proportionally to the household size and households could be selected more than once, but with a maximum that is equal to the household size. This design is an application of indirect sampling (Lavallée, 1995, 2007).

The target population of the panel are all natural persons residing in the Netherlands. The sample frame was a register containing all natural persons aged 15 years and over residing in the Netherlands as far as they are known to the Tax Office. From this register a stratified simple random sample of so-called core persons was drawn using proportional stratification. Neighbourhoods are used as the stratification variable.

This panel was drawn in 1994. To keep the panel representative of the target population, it was determined on a yearly basis which part of the population has entered the target population through birth and immigration. From this subpopulation, a stratified simple random sample of core persons was selected and added to the panel, with the purpose to maintain a representative sample.

Point and variance estimates are based on the general regression estimator. Therefore first and second order inclusion expectations for the sampling units under the above described sampling design are derived (van den Brakel, 2016). Weights are obtained by means of the GREG estimator to use auxiliary variables which are observed in the sample and for which the population totals are known from other sources (Särndal et al., 1992). Consequently, the weights reflect the (unequal) inclusion expectations of the sampling units and an adjustment such that for auxiliary variables the weighted observations sum to the known population totals. Details of the GREG estimator, expressions for variances, minimum required sample size and expected number of unique persons and households under this sample design are given in van den Brakel (2016).

4.2. Data and methods for poverty and well-being at DESTATIS

4.2.1. Data and methods for poverty measurement

In Germany, the system of social reporting in official statistics provides comparable data on the national and the state levels on issues such as minimum social security as well as poverty and social exclusion (Statistische Ämter des Bundes und der Länder, 2019). The data are jointly provided by DESTATIS (the German Federal Statistical Office) and the statistical offices of the Länder (states). The system comprises two major components, where the first component relates to the publication of reports on

Reference	Indicators
A1.	At-risk-of-poverty rates
A2.	At-risk-of-poverty thresholds
A3.	Gini coefficients
A4.	High-income rates
B1.	Minimum social security rates
C1.	Early school leavers
C2.	Persons with low educational attainment
D1.	Persons in households without persons in employment
D2.	Unemployment rates
D3.	Long-term unemployment rates
D4.	Labour force participation rates
D5.	Employment/population ratios

Table 4.1: Indicators covered in the German system of social reporting in official statistics (reprint from `http://www.amtliche-sozialberichterstattung.de/`)

minimal social security in Germany, while the second component focuses on the provision of indicators measuring poverty and social exclusion at national as well as regional levels. The data are published on the website `http://www.amtliche-sozialberichterstattung.de/` and cover selected indicators from four major areas (Statistische Ämter des Bundes und der Länder, 2019):

A. Income poverty and income distribution
B. Dependence on benefits of minimum social security
C. Qualification level
D. Labour force participation.

An overview of the indicators provided by the German system is given in Table 4.1. Note that various at-risk-of-poverty rates are available on subnational levels. They differ with respect to the level at which the at-risk-of-poverty threshold (60 per cent of the median of the equivalised disposable income) is computed, i.e. the national-, the state- or the NUTS2-level.

Most of the indicators presented in Table 4.1 are obtained from the German Microcensus, which is the largest annual household survey in Germany with a sampling fraction of approximately 1 per cent of all households (Bihler and Zimmermann, 2017). A notable exception are the indicators related to minimum social security rates, which are obtained from administrative records and, therefore, not subject to sampling errors. The German Microcensus uses a complex sampling design with equal selection probabilities, which can be described as a stratified cluster sampling. The sampling frame is stratified along two dimensions to facilitate precise estimates. The frame is stratified into regions such that each regional stratum contains at least 200,000 inhabitants in general. The aim of this procedure is to avoid highly heterogeneous strata in terms of their population size and, most importantly, very small strata. In a second step, the addresses are further stratified with respect to the number of dwellings within an address. In total there are 972 strata, resulting as the combination of 243 regions with 4 address-size-classes. As the German Microcensus samples clusters within strata, the cluster should be of similar size in order to foster precise estimates. Hence, selection districts were constructed from the addresses, such that the number of dwellings and persons is not too heterogeneous between selection districts. Details regarding the construction of selection districts can be found in Bihler and Zimmermann (2017).

In each stratum, 1 per cent of the selection districts is selected by means of a systematic probability sampling. Before the sample is drawn, the selection districts within a stratum are sorted according to more detailed regional information. The aim of this sorting procedure is to achieve approximately balanced samples even for analysis at a very granular level. Furthermore, the sample design includes a rotation pattern leading to a sample overlap of 75 per cent in two consecutive years, i.e. one quarter of the sampled units is replaced each year and a sampled unit remains in the survey for four consecutive years (Destatis, 2018c). This sample overlap has operational advantages, since it leads to a reduction of costs and increases the efficiency of estimates of changes between two consecutive time points at the same time.

The estimation methodology in the German Microcensus follows a two-step procedure, where in the first step a non-response adjustment is carried out and in the second stage the weights are calibrated to reproduce known population totals for auxiliary variables. As pointed out by Haziza and Lesage (2016), this two-step approach is more robust than performing the non-response adjustment and calibration to known totals in a single step. The calibration to known auxiliary totals is carried out at the level of 147 so-called "adjustment strata", which are aggregates of the 243 regional strata with an average population size of about 500,000. As known marginals, counts of 3 age-classes (0 - 14 years, 15 - 44 years, 45 years and older) and 4 nationality-groups (German, Turkish, EU-25, All others), both cross-classified with gender, are used. These auxiliary margins are generally obtained from the intercensal population updates, while the structures for the Non-German residents are taken from the central register of Foreigners. The weights which are used for the production of survey estimates are obtained from a generalised regression (GREG) estimator. This estimator belongs to the class of model-assisted estimation procedures, which allow for efficiency gains in the presence of a correlation between the target variables and the auxiliary information. Nevertheless, the GREG estimator is approximately design-unbiased and therefore robust against a potential misspecification of the working model (cf. Särndal et al., 1992 for further details). The weights which are obtained from the GREG estimator are then used to compute the indicators which are reported in the German system of social reporting in official statistics.

A special characteristic of the German Microcensus is that the net household income, which is the basis for many indicators on poverty and social exclusion, is not available as a metric variable. Instead, the income variable consists of 24 income classes, which have to be taken into account when computing the indicators. As an example, we shall consider the at-risk-of-poverty rate defined as the share of people with an equivalised disposable income (after social transfers) below 60 per cent of the national median equivalised disposable income after social transfers (Eurostat, 2019). To compute the median equivalised disposable income an estimate of the empirical distribution function is required. To arrive at this estimate, for all household members the upper and lower bound of the disposable household income in the income class are divided by the sum of the equivalence weights of all household members.

This process yields equivalence classes for each person in the data set, such that upper and lower bounds for the equivalised per-capita income is known. Under the assumption that the equivalised per-capita income is uniformly distributed within equivalence classes, estimates of the distribution function are obtained. Further details and an example are given in ITNRW (2018).

The second major survey that is used in German official statistics to produce poverty estimates and indicators is called "Leben in Europa", which is the German component of the EU-SILC survey. Whereas the Microcensus is used as a data source to provide information for the German system of social reporting in official statistics, the German component of EU-SILC contributes to EU-wide provision of harmonized microdata and indicators measuring living conditions, poverty and social exclusion as a basis for decision-making (Destatis, 2018b).

The German EU-SILC component is collected by means of a stratified random sampling from an access panel, the so-called Dauerstichprobe befragungsbereiter Haushalte (DSP). The DSP consists of households which declared their consent to take part in surveys of German official statistics and serves as a sampling frame for voluntary household surveys (Czajka and Rebeggiani, 2014). It is constructed as follows. After rotating out of the Microcensus, the households leaving the Microcensus are asked whether they are willing to participate in further surveys. Those who agree are then included in the DSP.

To comply with the EU regulation 1177/2003, the German SILC component aims at effective sample sizes of 8250 households for cross-sectional analyses and 6000 households for longitudinal analyses, respectively (Destatis, 2018b). Owing to the cluster effects present in the Microcensus and, therefore, also in the sampling frame DSP, a design effect of 1.3 is used when calculating the effective sample sizes. Furthermore, a panel mortality of 10 per cent of the households is assumed. Altogether, this implies a net sample size of approximately 14,000 households to fulfil the requirements regarding the effective sample sizes. The sampling frame is stratified according to various socio-demographic characteristics and the states as well (Destatis, 2018b). The rotation mechanism applied in the German SILC component is identical to the one in the Microcensus described above.

The estimation procedures applied to the German SILC component follow the guidelines of the European commission presented in European Commission (2016). The estimation process follows a two-step approach with a similar structure to the approach applied in the Microcensus, i.e. in the first step the weights are adjusted for non-response and then in the second step, they are calibrated to achieve coherence with known totals. As participation in the German SILC component is voluntary, potential biases owing to the self-selection in the sampling frame and the sample have to be accounted for (Czajka and Rebeggiani, 2014). Therefore, the non-response adjustment in the first step of the estimation process is crucially important for the ability to produce reliable estimates using the German SILC sample. In the second step, various different calibrated weights are computed which reproduce known totals of auxiliary variables. These totals of auxiliary variables, which can also be considered as calibration constraints, are obtained from the Microcensus. Further details on the estimation procedure and the calibration constraints can be found in Horneffer and Kuchler (2008).

The calibrated weights derived in the two-step procedure are also used when calculating indicators regarding living conditions, poverty and social exclusion. For this purpose, SAS programs provided by Eurostat are used. A detailed description can be found in Santourian and Ntakou (2014), which also includes a list of the indicators covered by EU-SILC. Further information on variance estimation methods for linear and non-linear indicators using SILC are given in Osier et al. (2013).

At the time of writing this report (summer 2019), Destatis and the statistical offices of the Länder are in the process of redesigning the German system of household statistics. Amongst the anticipated changes is a new sampling design for the German Microcensus, which will require new estimation methods as well. Furthermore, the SILC survey will be integrated into the Microcensus, which will also change its survey mode such that the participation in the SILC survey will be mandatory.

4.2.2. Data for well-being

In the following, we will elaborate on data sources used to produce indicators covered in the report "Nachhaltige Entwicklung in Deutschland" conducted by Destatis on behalf of the Federal Government of Germany (Destatis, 2018a). The report describes the progress of the indicators which are part of the German sustainable development strategy. The first version of this strategy was enacted in 2002 and a major revision approved by the Federal Cabinet in 2017 (Blumers and Kaumanns, 2017). Altogether the report provides information on 63 indicators with targets, which cover the 17 sustainable development goals laid out in the Agenda 2030 of the United Nations. The supplementary material to the report on the indicators is also available in English (Destatis, 2019b). Furthermore, Destatis launched a preliminary version of national reporting platform in July 2019 (Destatis, 2019a).

It should be noted that a certain amount of overlaps exists between the German SDG indicators and the indicators mentioned in the previous section aimed at measuring social exclusion, poverty and living conditions. For those overlapping indicators, the data sources mentioned in the previous section are used, i.e. mainly the Microcensus and the German component of EU-SILC. To compile the other non-overlapping indicators, Destatis integrates further data sources, which are obtained from two major sources. The first source are other data sets within the realms of Destatis and the statistical offices of the Länder. Examples include the use of the quarterly survey of earnings to estimate the indicator on the gender pay gap, the use of school statistics (a complete enumeration) to estimate the share of foreign school leavers who leave with a school degree or the indicator on energy consumption and $CO2$ emissions due to private households which are reported using data from the system of environmental economic accounts. The second data source are data from other national authorities (e.g. data from the Federal Environmental Agency to produce the indicator on greenhouse gas emissions) and also from non-governmental organisations (e.g. using the corruption perception index from Transparency International).

The few examples mentioned in the previous paragraph clearly indicate that in order to provide information on the different aspects of well-being captured by the indicators in the German sustainable development strategy, a vast number of different data sources are needed. Owing to the very heterogeneous characteristics of the data sources, the methodologies used in deriving indicator values are very heterogeneous as well.

4.3. Data and methods for poverty and well-being at ISTAT

Poverty and welfare indicators in Italy are mainly derived from household income, which in its turn is jointly derived from an annual survey and tax data. Starting from the last revision (Istat 2008) the annual survey is based on a sample design related to a rotating panel selected to return poverty measures at regional level.

It is important to stress that, according to a common strategy developed at Istat aiming to improve the Register system, a specific project to built up a register on households income is currently on going.

In Subsection 4.3.1 it is explained the main characteristics of the survey and sample design while Subsection 4.3.2 illustrates the indicators on poverty that are included in the annual report on Italian *well-being*[1] and finally Subsection 4.3.3 reports the list of the 12 indicators used by the Government to track the implication of the policy measures on *well-being*.

4.3.1. The Italian Statistics on Income and Living Conditions

Regulation of the European Parliament no. 1177/2003 is one of the main sources of data for periodic reports on the social situation of the European Union and the spread of the risk of poverty in UE member countries. EU-SILC is a multi-purpose instrument which focuses mainly on income and social exclusion, with a particular attention on aspects of material deprivation. In Italy the EU-SILC data are collected yearly since 2004. Although the EU-SILC Regulation requires national level estimates, the Italian survey allows for reliable estimates at regional level as well. The survey is conducted through household and personal interviews. Since 2011, interviews have been carried out by a private company according to a CAPI (Computer Assisted Personal Interview) technique instead of the PAPI (Paper and Pencil Interview) previously used. Since 2015 a share of the interviews is carried out by CATI (Computer Assisted Telephone Interview) technique, which in 2017 is about 54% of the households.

The sample design is based on a *two-stages scheme* (municipalities and households), where the primary sample units, municipalities, are stratified by population size within each region. Rotational design is used for households; the whole sample is composed of four rotational groups, each group is included in the sample for four waves of the survey. Each year one fourth of the sample is renewed. replacing the group entered in the sample four years before, while the remaining three fourths are made of households and individuals selected one, two or three years before, interviewed respectively for the second, third or fourth time. The overall sample is statistical representative of the population residing in Italy and, in 2017, it amounts to 22,226 households (48,819 individuals), residing in about 680 municipalities.

Data collection is carried out through an electronic questionnaire, structured in three parts:

- General form to collect demographic information related to each household member (sex, date and place of birth, citizenship etc.) and some information for each household member aged less than 16 years (type of school attended, formal and informal childcare etc.);
- Household questionnaire to collect information about housing conditions, housing expenses, economic situation, material deprivation, household income components;

[1] For a more general presentation of the Italian annual report on well-being we refer to Bacchini et al. 2019 and Istat 2019

- Personal questionnaire for each household member aged at least 16 years to collect information on education, health, current or previous labour income by detailed components (employee, self-employment, pensions and other social transfers, financial and real capital, private transfers). Income data collected by interviews are integrated with administrative register data. A micro-simulation model allows to obtain further gross income values.

4.3.2. Italian indices on Income and living conditions included in the well-being framework

The Italian National Institute of Statistics (Istat), together with the National Council for Economics and Labor (CNEL), launched in December 2010 an *inter-institutional* initiative aimed at developing a *multi-dimensional* approach for the measurement of *equitable and sustainable wellbeing* (*Bes - benessere equo e sostenibile*), in line with the recommendations issued by the OECD and the Stiglitz Commission (see Stiglitz et al., 2009).

In the context of recent international initiatives, the approach adopted with the Bes has been characterized by a participative process, involving civil society and national experts in the definition of the framework and in the selection of indicators[2].

Since the preliminary steps, Bes has had the ambition to measure not only the level of well-being, through the analysis of all relevant aspects of quality of life of the population, but also its equity amid social groups and geographic areas of the Country, and sustainability for future generations. This approach increases the complexity of the measurement but allows a more accurate analysis of the evolution of well-being in Italy.

The results of the consultations, together with the evidences coming from international experiences, supported the Steering Group that identified a total of 12 domains.

The 12 selected domains are divided into 2 typologies, 9 of them are defined as outcome domains and are those related to dimensions which have a direct impact on human and environmental well-being while the remaining 3 domains are defined as drivers of well-being, measuring functional elements to improve the well-being of the community and the surrounding environment. The domains are:

- Outcome: health; education and training; work and life balance; economic well-being; social relationship; security; landscape and cultural heritage; environment; subjective well-being;
- Driver: politics and institutions; innovation, research and creativity; quality of services.

The 12 domains refers to 130 indices Istat, 2018. Particularly the indices related to Eu-silc are the following:

Very low work intensity: number of persons living in a household having a work intensity below a threshold set at 0.20. The work intensity of a household is the ratio of the total number of months that all working-age household members have worked during the income reference year and the total

[2] For further details on the characteristics of Bes see Bacchini et al., 2019

number of months the same household members theoretically could have worked in the same period. A working-age person is a person aged 18-59 years. with the exclusion of students in the age group between 18 and 24 years. Households composed only of children. of students aged less than 25 and/or people aged 60 or more are completely excluded from the indicator calculation.

Severe material deprivation rate: indicator that measures the inability to afford some items considered by most people desirable or even necessary to lead an adequate life. It measures the percentage of the population that cannot afford at least four of the following nine items:

1. to pay their rent. mortgage or utility bills;
2. to keep their home adequately warm;
3. to face unexpected expenses;
4. to eat meat or proteins regularly;
5. to go on a week holiday;
6. a television set;
7. a washing machine;
8. a car;
9. a telephone.

At-risk-of-poverty rate it is the share of people with an equivalised disposable income (after social transfers) below the at-risk-of-poverty threshold, which is set at 60% of the national median equivalised disposable income after social transfers. The disposable income does not include imputed rent, non-cash employee income (other than company car) and income from household production of goods for own consumption. In 2017 the at-risk-of-poverty threshold (computed on 2016 incomes) is 9,925 euros per year (827 euros per month) for a one adult member household.

Income quintile share ratio or S80/S20 ratio: measure of the inequality of income distribution. It is calculated as the ratio of total income received by the 20% of the population with the highest income (the top quintile) to that received by the 20% of the population with the lowest income (the bottom quintile). All incomes are compiled as equivalised disposable income.

Severe housing deprivation rate: is the percentage of population living in the dwelling which is considered as overcrowded, while also exhibiting at least one of the housing deprivation measures. Housing deprivation is a measure of poor amenities and is calculated by referring to those households with a leaking roof, no bath/shower and no indoor toilet, or a dwelling considered too dark.

Inability to make ends meet rate: is the percentage of population living in households who declare to make ends meet with great difficulty

4.3.3. The Italian budget law and well-being (poverty) indicators in the policy cycle

In 2016 the Italian Law reforming the budget law (163/2016), establishing that well-being indicators have to be considered in the economic policy process[3]. In particular, the law indicates that an analysis

[3] The Italian example will be fully explored in WP5

of the recent trend has to be performed, together with ad-hoc simulations of the expected evolution in two scenarios, one just projecting past trends (trend scenario), the other one taking into consideration the impact of new policy measures on well-being (policy scenario) (see G.U., 2016).

The new law's requirements lead to two annual reports. The first one, in April, corresponds to the presentation of the Planning Document on Economic and Financial Policy (DEF–Documento di Economia e Finanza), where the Government outlines the policy actions to be undertaken in the subsequent three-years period. In an annex, also the indicators measuring equitable and sustainable well-being are analysed and projected in the trend and in the policy scenario. In February, following the approval of the Budget law for the current year, a second report is presented to the Parliament, updating findings and forecasts presented in the DEF in light of the specific measures set out in the Budget law in force (usually approved by the end of the previous year).

An high level Commission was set up in order to select a list of indicators to be used in the policy cycle[4]. The final list of indicators was approved unanimously by the parliamentary committees.

The final result of the Commission's work is a selection of 12 indicators out of the 130 included in the Bes framework, namely [5]

1. Mean adjusted income (per capita)
2. *Income inequality (quintile ratio)*
3. *Incidence of Absolute poverty*
4. Life expectancy in good health at birth
5. Overweight and obesity
6. Early school leavers
7. Non-participation in employment
8. Employment rate of women aged 25-49 with preschool children vs women without children
9. Victims of predatory crime
10. Mean length of civil justice trials
11. CO_2 and other greenhouse gas emissions (tons x inhab.)
12. Illegal Building

Comparing this list with the indicators coming from the Eu-silc survey included in the Bes, it emerges that *Income inequality* is a key indicator for policy as well as the *Incidence of absolute poverty*, that is drawn from quarterly survey on household expenditure. However, current policy evaluation is analysing also the dynamics of the *At-risk-of-poverty rate* indicator.

4.4. Data and methods for poverty and well-being at HCSO

4.4.1. Data and methods for poverty

The Hungarian Statistics on Income and Living Conditions

[4] The full report is available (Comitato per gli indicatori di benessere equo e sostenibile 2017) so that the whole process is public and transparent (see also Bacchini et al. 2018).

[5] At this stage it is important to underline that a subset of this indicators is in common with those one selected by the Macroeconomic Imbalance Procedure (MIP) as auxiliary indicators. We will explore this point in WP4

The main objective of EU-SILC - European Survey on Income and Living Conditions is to apply EU comparable data on the living conditions of the population. EU-SILC provides the basic source of information used for the calculation of indicators, among others those related to income, poverty and social exclusion, for the EU member states. Since the lunch of the Europe 2020 strategy for smart, sustainable and inclusive growth the importance of EU-SILC has grown rapidly.

As an EU Member State, Hungary follows the EU regulations on the measurement of poverty and well-being. EU-SILC organisation and methodology is governed by the Regulation (EC) No. 1177/2003 of the European Parliament and of the Council of 16 June 2003 (with amendments included in regulation No. 1553/2005) concerning Community Statistics on Income and Living Conditions (EU-SILC) along with regulations of the European Commission corresponding to that legal act.

EU-SILC was implemented by Hungarian Central Statistical Office in 2005. A gradual merger of HBS and SILC sample started in 2007. 1/4 of the total sample was common in both surveys in 2007-2008. Then 1/2 of the sample became common in 2009; 3/4 of the sample was common in in 2010-2011. By 2012 the merger was completed totally which makes us possible to analyse the joint distribution of income and expenditure. The survey unit is a household and all the household members who had completed 16 years of age by 31 December of the year preceding the survey. The EU-SILC fieldwork is carried out every year from March to May, the fieldwork period is 2.5 months. The reference period for income is the previous calendar year. Data collection is carried out by face to face interviews (CAPI) or Computer Assisted Web Interview (CAWI). The dominant survey mode is CAPI.

The sample of EU-SILC consists of households that successfully participated in HBS in the previous calendar year. It counts about 9,000 households and it represents about 13,000 persons. The participation in EU-SILC is voluntary for the households. The sample covers approximately 300 settlements in the country. Number of interviewers conducting the survey counts about 200 persons. The EU-SILC survey data are representative at regional level. According to the legislation in force, the survey should collect the data allowing for both the cross-sectional and longitudinal analyses.

The survey collects all the data directly from respondents and according to the current lay-out do not incorporate any data from administrative registers. There is an on-going work of test of using employee income from administrative source in the period of 2019-2020.

Methodology of poverty measurement

The source of poverty measures in Hungary is the EU-SILC survey. Since 2014 the sampling frame of the Hungarian EU-SILC consists of the dwelling units of the 2011 census. The SILC has a two-stage sample of households. At the first stage municipalities are selected with probability proportional to size. The population of smaller municipalities is stratified by county and size while the biggest settlements are certainty PSUs. There are approximately 300 municipalities in the sample. In the second stage households are selected with stratified simple random sampling. There are 3 strata in the self-representing (selected by a probability of 1) settlements according to the characteristics of the head of the households: 1.) the head of household is over 60 years, 2.) he/she is younger than 60 years and has a university/college degree, 3.) he/she is younger than 60 years and doesn't have a university/college degree. For the 2 strata of the non self-representing municipalities only the age of the head of the household matters: 1.) he/she is over 60 years or 2.) under 60 years. We apply proportional allocation for the calculation of the strata's sample size. In order to preserve an approximate proportional allocation in the responded sample we have to use nonresponse multipliers. They are calculated from the previous sample by county, type of the settlement and household strata. Actually these are the main characteristics of the sampling plan of the Household Budget Survey which is the prerequisite for qualifying into the EU-SILC. Those households which have an accepted "budget diary" from the previous year are assigned for the first wave of the EU-SILC taking place next spring. The SILC has a rotation panel scheme: households are asked to participate in the survey for four consecutive years. The size of the cross-sectional sample in 2019 was 6911 households.

Many aspects of the survey - especially weighting - are regulated in detail by Eurostat. The main steps of the weighting process also can be found in European Commissions' (2016) methodological guideline. Of course there are some unique features in the countries' practice. At the cross-sectional calibrations we use a relatively simple primary weight which is the ratio of the total number of households and the number of households participating in the survey in strata. These are constructed by region, type and size of the settlements and type of household. In order to reduce sampling error and the bias caused by nonresponse we adjust the primary weights by the means of a raking ratio calibration to demographic and economic activity control totals. These totals stem from the projection of the last census and the Labour Force Survey.

We use these calibration weights for the calculation of different poverty measures. Regarding the properties of the estimator we may think of it as a GREG estimation so approximate design unbiasedness is expected of it. A formula based standard error estimation is carried out taking main sampling features (stratification, multi-stage selection) and calibration effect into account. For the geographical level of estimates also the EU-regulation is decisive, which specifies estimations with given precision requirements for the NUTS2 regions from 2021.

Disposable income of household

The basis of measuring poverty is the size of the household's disposable income.

Disposable income

Disposable income in the survey is defined as a sum of the net (after deduction of income tax prepayment, tax on income from property, social and health insurance contributions) annual monetary incomes (in case of hired employment taking into account also non-monetary profit from the use of the company car) gained by all the household members reduced by: property tax, inter-household cash transfers paid and balance of offsetting settlements with the Tax Office. The disposable income includes:

1. Income from work (including employee income, self-employment income);
2. Social benefits (including family/children-related allowances; housing allowances; unemployment

benefits; old-age benefits; survivors' benefits; sickness benefits; disability benefits; education-related allowances; social exclusion not elsewhere classified);

3. Other income including regular inter-household cash transfers received, income from the financial property, income received by people aged under 16.

Indicators of poverty and social exclusion

Based on EU-SILC the results on poverty and social inclusion are published each year.

People at risk of poverty or social exclusion

The number of people who are at risk of poverty or social exclusion combine three separate measures and covers those persons who are at least in one of these three situations as follows:

1. *At-risk-of-poverty rate*
   Percentage of persons with an equalised annual disposable income (after social transfers) below the at-risk-of-poverty threshold set at 60% of the national median of equalised annual disposable income
2. *Severely materially deprived people*
   Percentage of persons in households declaring inability to meet at least 4 out of 9 following needs due to financial reasons:
   - Go on a week holiday of all households members away from home once a year
   - Eat meat, fish (or vegetarian equivalent) every second day
   - Keep home adequately warm
   - Face unexpected expenses (in the amount of the monthly values 60% of the national median of equalised disposable income)
   - Timely adjust payments related to housing, repayment instalments and credits;
   - A colour television
   - A car
   - A washing machine
   - A telephone

   We also adopted the new material and social deprivation indicator which covers the following six personal deprivations are the inability for the person to:
   - Replace worn-out clothes with some new ones
   - Have two pairs of properly fitting shoes
   - Spend a small amount of money each week on him/herself ("pocket money")
   - Have regular leisure activities
   - Get together with friends/family for a drink/meal at least once a month
   - Have an internet connection
3. *People living in households with very low work intensity*
   Percent of persons aged 0-59 living in households with very low work intensity, where the adults (aged 18-59) work less than 20% of their total work potential during the past year.

Inequality of income distribution S80/S20 (income quintile share ratio)

Ratio of total income received by the 20% of the population with the highest income (top quintile) to that received by the 20% of the population with the lowest income (lowest quintile). In EU-SILC this indicator is calculated for equalised annual disposable income of households.

Gini coefficient

The measure of income distribution inequality; it ranges between 0 and 1 (or if multiplied by 100 - between 0 and 100). This indicator would be 0 (homogenous distribution) if all the persons had the same income, whereas it would be 1 if all the persons except one had 0 income. Thus the higher the indicator, the higher the income concentration and therefore, the greater the income inequalities. In EU-SILC this indicator is calculated for equalised annual disposable income of households.

How to cover the regular expenses of the household

On a scale of 1-6 (very easily - at the expense of great difficulty), we measure how difficult it is for a household to cover its usual expenses.

Overcrowding

The overcrowding rate is defined as the percentage of the population living in an overcrowded household. A person is considered as living in an overcrowded household if the household does not have at its disposal a minimum number of rooms equal to:

- one room for the household;
- one room per couple in the household;
- one room for each single person aged 18 or more;
- one room per pair of single people of the same gender between 12 and 17 years of age;
- one room for each single person between 12 and 17 years of age and not included in the previous category;
- one room per pair of children under 12 years of age.

4.4.2. Data for well-being

Quality of life surveys take a prominent place in official statistics across Europe. In the course of its long history besides objective measurement tools (e.g. GDP) indicators aiming at the measurement of the subjective dimension have become more and more prominent which supplementing objective indicators are more sensitive to measuring people's life-related assessments and experiences. In Hungary, for the first time in 2013, a detailed series of questions measuring individual well-being was surveyed, as part of the SILC survey module coordinated by Eurostat. Hungary considered it important to get a broader picture of the subjective well-being of the Hungarian population, therefore, in the national SILC survey, it repeatedly asked for a series of questions. Thus, after 2013, the subjective opinion of the population in terms of their well-being was also assessed in 2015. Starting in 2015, some key indicator of well-being, such as life satisfaction and household financial health, were also present in the transition years. In 2016, well-being questions previously defined by Eurostat with national specificities were also part of the Hungarian Microcensus (Mikrocensus 2016), which was surveyed by 10% of the addresses assigned to the Census. More than 50,000 people completed the questionnaire. As a result, the HCSO was able to draw conclusions about the well-being of the population from a particularly large sample. In 2018, it was re-queried in the SILC survey as defined by Eurostat, as this year the ad-hoc module repeatedly focused on the topic of subjective well-being.

Official statistics at international as well as national level tend to use a complex system of indicators to measure the quality of life that incorporates biological, psychological and social factors. The Hungarian Central Statistical Office has developed a system of indicators of subjective well-being based on international and national research results and recommendations. The indicator system affects work and leisure; material living conditions; education, knowledge; the health; mental health; the living environment and infrastructure; dimensions of human relationships, social participation and social renewal.

Detailed subjective well-being questionnaire[6]:

- Overall life satisfaction
- Perceived social exclusion
- Material help
- Non-material help
- Satisfaction with financial situation
- Satisfaction with personal relationships
- Satisfaction with time use (amount of leisure time)
- Satisfaction with job
- Trust in others
- Feeling lonely
- Being very nervous
- Feeling down in the dumps
- Feeling calm and peaceful
- Feeling downhearted or depressed
- Being happy
- Satisfaction with the quality of living environment
- Satisfaction with accommodation
- Does he/she feel that the things he/she do in his/her life are worthwhile
- Has anyone to discuss personal matters
- Feeling safe if he/she walk alone in the area around his/her home after dark
- Trust in national institutions (political system, legal system, police)

Permanent questions in the national SILC questionnaire:

[6] Recent status of the national subjective well-being questionnaire block (2018).

- Overall life satisfaction
- Satisfaction with financial situation
- Trust in others
- Feeling safe if he/she walk alone in the area around his/her home after dark

5. Summary

The Deliverable is motivated by one of the EU's top priorities, namely the reduction of poverty and strengthening of well-being. These goals are manifested in the Europe 2020 strategy and the Millennium Development Goals. Although transitions in poverty and inequality happen at low regional levels, poverty and inequality indicators have primarily been estimated at the national level (EU-SILC data). Hence, a greater focus on European regions is needed, also justified by the fact that a large share of the EU's budget is directed to its cohesion policy. For the regional estimation of poverty and well-being, many different indicators are available which are used worldwide. These indicators and their usage had to be gathered, which this deliverable does. In addition to reviewing many different indicators, the document focussed on the discussion of statistical methods to estimate such indicators on a regional level and their specific data needs. Examples from the practice of four European NSIs (CBS, DESTATIS, ISTAT, HSCO) complete the picture.

As illustrated in Table 5.1, the national measures on poverty provided by the NSI belonging to the consortium present an high level of heterogeneity both in the way in which the sample design is developed as well as in the territorial level for which indicators are available. Please note that in the case of Germany, we refer to the indicators covered in the German system of social reporting in official statistics as indicated in Section 4.2.

Concerning the methods present in this deliverable experiments are running for example in Italy to extend the actual measures at a more disaggregated level (large municipalities) by means of a Small Area Estimation. In Germany, model-assisted estimation procedures are used to produce the estimates obtained from sample surveys mentioned in Table 5.1.

Table 5.1: Methods and data on poverty by countries

Country	Poverty measure	Source	Territorial level	Methodology
Italy	At-risk-of-poverty rate	EU-SILC	Region	Rotating panel
	Share ratio S80-S20 ratio	EU-SILC	Region	Rotating panel
	Incidence of absolute poverty	HBS	Region	Sample design
	Very low work intensity	EU-SILC	Region	Rotating panel
	Severe material deprivation rate	EU-SILC	Region	Rotating panel
	Severe housing deprivation rate	EU-SILC	Region	Rotating panel
	Inability to make ends meet rate	EU-SILC	Region	Rotating panel
Germany	At-risk-of-poverty rate	Microcensus	Raumordnungsregion[a]	Rotating panel
	At-risk-of-poverty threshold	Microcensus	Raumordnungsregion	Rotating panel
	Gini coefficient	Microcensus	States	Rotating panel
	High-income rate	Microcensus	States	Rotating panel
	Minimum social security rate	Administrative data	States	Administrative data
	Early school leavers	Microcensus	States	Rotating panel
	Persons with low educational attainment	Microcensus	States	Rotating panel
	Persons in households without persons in employment	Microcensus	States	Rotating panel
	Unemployment rates	Microcensus	States	Rotating panel
	Long-term unemployment rates	Microcensus	States	Rotating panel
	Labour force participation rates	Microcensus	States	Rotating panel
	Employment/population ratios	Microcensus	States	Rotating panel
Netherlands	Poverty based on low-income threshold	Administrative data	Neighbourhood, and aggregates [b]	Administrative data
	Poverty based on social minimum threshold	Administrative data	Neighbourhood, and aggregates	Administrative data
	Poverty based on the European poverty threshold	Administrative data	Neighbourhood, and aggregates	Administrative data
	Longterm poverty based on low-income threshold	Administrative data	Neighbourhood, and aggregates	Administrative data
	Longterm poverty based on social minimum threshold	Administrative data	Neighbourhood, and aggregates	Administrative data
	Longterm poverty based on the European poverty threshold	Administrative data	Neighbourhood, and aggregates	Administrative data
Hungary	Poverty threshold	EU-SILC	Region	Rotating panel
	At-risk-of-poverty rate	EU-SILC	Region	Rotating panel
	Relative at-risk-of-poverty gap	EU-SILC	Region	Rotating panel
	Gini coefficients	EU-SILC	Region	Rotating panel
	Severe material deprivation rate	EU-SILC	Region	Rotating panel
	Share of persons living in overcrowded dwelling	EU-SILC	Region	Rotating panel
	Rate of those living in a household with very low work intensity	EU-SILC	Region	Rotating panel
	People at risk of poverty or social exclusion	EU-SILC	Region	Rotating panel
	At risk of poverty or social exclusion rate	EU-SILC	Region	Rotating panel

[a] There are 96 Raumordnungsregionen in Germany, which are aggregates of NUTS3 regions.

[b] In case of the Netherlands, *aggregates* refers to *municipality, province, national level.*

Bibliography

Alexander, C. H. (1987). A class of methods for using person controls in household weighting. *Survey Methodology 13*(2), 183–198.

Alfons, A. and M. Templ (2013). Estimation of social exclusion indicators from complex surveys: The R package laeken. *Journal of Statistical Software 54*(15), 1–25. doi:10.18637/jss.v054.i15.

Alkire, S. and S. Jahan (2018). The new global MPI 2018: Aligning with the Sustainable Development Goals. Technical report, UNDP Human Development Report Office (HDRO). Occasional Paper.

Alkire, S. and M. E. Santos (2014). Measuring acute poverty in the developing world: Robustness and scope of the multidimensional poverty index. *World Development 59*, 251–274.

AMELI (2018). *AMELI: Advanced Methodology for European Laeken Indicators 2008 Proj.n. SSH-CT-2008-217322. FP7-SSH-2007-1.*

Arima, S., W. Bell, G. Datta, C. Franco, and B. Liseo (2017). Multivariate Fay-Herriot Bayesian estimation of small area means under functional measurement error. *Journal of the Royal Statistical Society - Series A 180*(4), 1191–1209.

Arima, S., G. Datta, and B. Liseo (2015). Bayesian estimators for small area models when auxiliary information is measured with error. *Scandinavian Journal of Statistics 42*(2), 518–529.

Arima, S. and S. Polettini (2019). A unit level small area model with misclassified covariates. *Journal of the Royal Statistical Society - Series A 182*(4), 1439–1462.

Bacchini, F., B. Baldazzi, R. De Carli, L. Di Biagio, M. Savioli, and M. P. Sorvillo (2019). The italian framework to measure well-being: towards the 2.0 version. *GROWINPRO Working Paper n.21*.

Bacchini, F., B. Baldazzi, and L. Di Biagio (2019). The evolution of composite indices of well-being: an application to italy. *GROWINPRO Working Paper n.22*.

Bacchini, F., M. G. Calza, M. Gandolfo, M. P. Sorvillo, and A. Tinto (2018). MAKSWELL: An EU project on MAKing Sustainable development and WELl-being frameworks work for policy analysis. Technical report, Paper presented at ISQOLS 2018.

Barro, R. and J.-W. Lee (2016). Dataset of educational attainment, february 2016 revision.

Beaumont, J. F. and A. S. M. A. Alavi (2004). Robust generalized regression estimation. *Survey Methodology 30*(2), 195–208.

Betti, G. (2017). Fuzzy measures of quality of life in Germany: a multidimensional and comparative approach. *Quality & Quantity 51*, 23–34.

Bihler, W. and D. Zimmermann (2017). Die neue Mikrozensusstichprobe ab 2016. *Wirtschaft und Statistik* (6), 20–29.

Blumers, M. and S. C. Kaumanns (2017). Neuauflage der deutschen Nachhaltigkeitsstrategie. *Wirtschaft und Statistik* (1), 96–109.

Bollen, K. A. and S. Bauldry (2011). Three Cs in measurement models: causal indicators, composite indicators, and covariates. *Psychological methods 16*, 265–284.

Borssoi, J, A., G. Paula, and M. Galea (2017). Elliptical linear mixed models with a covariate subject to measurement error. *Statistical Papers*, 1–39.

Breckling, J. and R. Chambers (1988). M-quantiles. *Biometrika 75*(4), 761–771.

Bruch, C., R. Münnich, and S. Zins (2011). *Variance Estimation for Indicators of Poverty and Social Exclusion.* Research Project Report WP3 (D3.1). FP7-SSH-2007-217322 AMELI. `https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Ameli_Delivrables/AMELI-WP3-D3.1-20110514.pdf` [30.07.2019].

Brunsdon, C., A. S. Fotheringham, and M. E. Charlton (1996). Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical analysis 28*(4), 281–298.

Burgard, J., M. Esteban, D. Morales, and A. Perez (2019). A Fay-Herriot model when auxiliary variables are measured with error. *Test*, 1–30.

Burgard, J. P., R. Münnich, and M. Rupp (2019). A Generalized Calibration Approach Ensuring Coherent Estimates with Small Area Constraints. Research Papers in Economics 2019-10, University of Trier, Department of Economics.

Burgass, M. J., B. S. Halpern, E. Nicolson, and E. J. Milner-Gulland (2017). Navigating uncertainty in environmental composite indicators. *Ecological Indicators 75*, 268–278.

Cavicchia, C. and M. Vichi (2017). Model-based synthesis of indicators. statistical composite indicators to convey consistent policy messages. Presented at the Workshop ‘The Impacts and Methodology of Indicators and Scoreboards’, 22-23 March 2018, Ispra.

Chambers, R., H. Chandra, N. Salvati, and N. Tzavidis (2014). Outlier robust small area estimation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 76*(1), 47–69.

Chambers, R., H. Chandra, and N. Tzavidis (2011). On bias-robust mean squared error estimation for pseudo-linear small area estimators. *Survey Methodology 37*(2), 153–170.

Chambers, R. and N. Tzavidis (2006). M-quantile models for small area estimation. *Biometrika 93*(2), 255–268.

Chandra, H. and R. Chambers (2005). Comparing EBLUP and C-EBLUP for small area estimation. *Statistics in Transition 7*(3), 637–648.

Chandra, H. and R. Chambers (2009). Multipurpose weighting for small area estimation. *Journal of Official Statistics 25*(3), 379–395.

Chandra, H., N. Salvati, R. Chambers, and N. Tzavidis (2012). Small area estimation under spatial nonstationarity. *Computational Statistics and Data Analysis 56*.

Chandra, H. and U. C. Sud (2012). Small area estimation for zero-inflated data. *Communications in Statistics-Simulation and Computation 41*(5), 632–643.

Cheli, B. and A. Lemmi (1995). A totally fuzzy and relative approach to the multidimensional analysis of poverty. *Economic Notes 24*, 115–134.

Comitato per gli indicatori di benessere equo e sostenibile (2017). Relazione finale. Technical report, . `http://www.istat.it/it/files//2017/12/relazione_comitato_fin.pdf`.

Czajka, S. and L. Rebeggiani (2014). Die Dauerstichprobe befragungsbereiter Haushalte als Auswahlgrundlage für EU-SILC. *Wirtschaft und Statistik* (10), 621–629.

D'Agostino, A., C. Giusti, and A. Potsi (2018). Gender and children's wellbeing: Four mediterranean countries in perspective. *Child Indicators Research 11*(5), 1649–1676.

D'Agostino, A., G. Grilli, and A. Regoli (2019). The determinants of subjective well-being of young adults in Europe. *Applied Research Quality Life 14*, 85–112.

Das, K., J. Jiang, and J. N. K. Rao (2004). Mean squared error of empirical predictor. *The Annals of Statistics 32*(2), 818–840.

Datta, G. (2009). *Handbook of Statistics: Sample Surveys: Inference and Analysis*, Volume 29B, Chapter Model-based approach to small area estimation, pp. 251–288. North-Holland.

Datta, G. and P. Lahiri (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica 10*, 613–627.

Datta, G., J. Rao, and M. Torabi (2010). Pseudo-empirical bayes estimation of small area means under a nested error linear regression model with functional measurement errors. *Journal Statistical Planning and Inference 140*(11), 2952–2962.

Datta, G. S., R. Ghosh, M. ana Steorts, and J. Maples (2011). Bayesian benchmarking with applications to small area estimation. *Test 20*(3), 574–588.

Deming, W. E. and F. F. Stephan (1940). On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *The Annals of Mathematical Statistics 11*(4), 427–444.

Destatis (2018a). Nachhaltige Entwicklung in Deutschland - Indikatorenbericht 2018. `https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Nachhaltigkeitsindikatoren/Publikationen/Downloads-Nachhaltigkeit/indikatoren-0230001189004.pdf;jsessionid=0A3F2E7222EF2E7362FB0129E3DD5C89.internet721?__blob=publicationFile` [29.07.2019].

Destatis (2018b). Qualitätsbericht Leben in Europa - 2017. `https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Einkommen-Konsum-Lebensbedingungen/leben-in-europa-2017.pdf?__blob=publicationFile` [29.07.2019].

Destatis (2018c). Qualitätsbericht Mikrozensus - 2017. `https://www.destatis.de/DE/Methoden/Qualitaet/Qualitaetsberichte/Bevoelkerung/mikrozensus-2017.pdf?__blob=publicationFile` [29.07.2019].

Destatis (2019a). National Data for UN-SDGs. https://sustainabledevelopment-germany.github.io/ [29.07.2019].

Destatis (2019b). Sustainable development in Germany - Data to the indicator report 2018. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Nachhaltigkeitsindikatoren/Publikationen/Downloads-Nachhaltigkeit/data-relating-indicator-report-2018.pdf?__blob=publicationFile [29.07.2019].

Devaud, D. and Y. Tillé (2019). Deville and Särndal's calibration: revisiting a 25-years-old successful optimization problem. *Test 28*, 1033–1065.

Deville, J. C. (1999). Variance estimation for complex statistics and estimators: linearization and residual techniques. *Survey Methodology 25*(2), 193–204.

Deville, J.-C. and C.-E. Särndal (1992). Calibration estimators in survey sampling. *Journal of the American statistical Association 87*(418), 376–382.

Deville, J.-C., C.-E. Särndal, and O. Sautory (1993). Generalized raking procedures in survey sampling. *Journal of the American statistical Association 88*(423), 1013–1020.

Duan, N. (1983). Smearing estimate: a nonparametric retransformation method. *Journal of the American Statistical Association 78*(383), 605–610.

Elbers, C., J. O. Lanjouw, and P. Landjouw (2003). Micro-Level Estimation of Poverty and Inequality. *Econometrica 71*(1), 355–364.

Erciulescu, A. L. and W. A. Fuller (2014). Parametric bootstrap procedures for small area. In *Proceedings of the Survey Research Methods Section.*

Estevao, V. M. and C.-E. Särndal (2000). A functional form approach to calibration. *Journal of Official Statistics 16*(4), 379–399.

Estevao, V. M. and C.-E. Särndal (2006). Survey estimates by calibration on complex auxiliary information. *International Statistical Review 74*(2), 127–147.

European Commission (2010). Communication from the commission. Europe 2020 - a strategy for smart, sustainable and inclusive growth. COM(2010) 2020 final. https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:2020:FIN:EN:PDF [15.07.2019].

European Commission (2013). *Quality of life in Europe: Subjective well-being.* Luxembourg: Publications Office of the European Union.

European Commission (2016). Methodological guidelines and description of EU-SILC target variables. https://circabc.europa.eu/sd/a/afb4601b-4e5c-4f40-86bb-0c3d0d94aa12/DOCSILC065%20operation%202015%20VERSION%20November%202015.pdf [29.07.2019]. 2015 operation.

European Commission (2017, May). Methodological guidlines and description of EU-SILC target variables. DocSILC065 (2016 operation). https://circabc.europa.eu/sd/a/165c80b9-5631-4f5b-b847-29c638715c0e/DOCSILC065%20operation%202016%20VERSION%2022-05-2017.pdf [17.07.2019].

Eurostat. EU statistics on income and living conditions (EU-SILC) methodology - Europe 2020 target on poverty and social exclusion. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_Europe_2020_target_on_poverty_and_social_exclusion#Description [17.07.2019].

Eurostat. EU statistics on income and living conditions (EU-SILC) methodology - health and labour conditions. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_health_and_labour_conditions [17.07.2019].

Eurostat. EU statistics on income and living conditions (EU-SILC) methodology - material deprivation by dimension. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_material_deprivation_by_dimension [17.07.2019].

Eurostat. EU statistics on income and living conditions (EU-SILC) methodology - monetary poverty. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=EU_statistics_on_income_and_living_conditions_(EU-SILC)_methodology_-_monetary_poverty#Description [17.07.2019].

Eurostat. Glossary: Material deprivation. https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Material_deprivation [17.07.2019].

Eurostat (2009). *Algorithms to compute Overarching indicators based on EU-SILC andadopted under the Open Method of Coordination (OMC)*. Unit F-3: Living conditions and social protection, Directorate F: Social and informationsociety statistics, Eurostat, Luxembourg: Doc. LC-ILC/39/09/EN-rev.1.

Eurostat (2013). *Handbook on precision requirements and variance estimation for ESS households surveys*. https://ec.europa.eu/eurostat/documents/3859598/5927001/KS-RA-13-029-EN.PDF [20.07.2019].

Eurostat (2016). Smarter, greener, more inclusive? indicators to support the Europe 2020 strategy. https://ec.europa.eu/eurostat/documents/3217494/7566774/KS-EZ-16-001-EN-N.pdf/ac04885c-cfff-4f9c-9f30-c9337ba929aa [12.07.2019].

Eurostat (2017a). *Final report of the expert group on quality of life indicators*. Luxembourg: Publications Office of the European Union.

Eurostat (2017b). Smarter, greener, more inclusive? indicators to support the Europe 2020 strategy. https://ec.europa.eu/eurostat/documents/3217494/8113874/KS-EZ-17-001-EN-N.pdf/c810af1c-0980-4a3b-bfdd-f6aa4d8a004e [12.07.2019].

Eurostat (2018). Smarter, greener, more inclusive? indicators to support the Europe 2020 strategy. https://ec.europa.eu/eurostat/documents/3217494/9087772/KS-02-18-728-EN-N.pdf/3f01e3c4-1c01-4036-bd6a-814dec66c58c [12.07.2019].

Eurostat (2019). Glossary: At-risk-of-poverty rate. https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At-risk-of-poverty_rate [29.07.2019].

Fabrizi, E., C. Giusti, N. Salvati, and N. Tzavidis (2014). Mapping average equivalized income using robust small area methods. *Papers in Regional Science 93*(3), 685–701.

Fabrizi, E., N. Salvati, and M. Pratesi (2012). Constrained small area estimators based on m-quantile methods. *Journal of Official Statistics 28*(1), 89–106.

Fabrizi, E., N. Salvati, M. Pratesi, and N. Tzavidis (2014). Outlier robust, model-assisted small area estimation. *Biometrical Journal 56*(1), 157–175.

Frey, J. and N. Cressie (2003). Some results on constrained bayes estimators. *Statistics & probability letters 65*(4), 389–399.

Geiger, C. and C. Kanzow (2002). *Theorie und Numerik restringierter Optimierungsaufgaben*. Springer. Springer Berlin Heidelberg.

Ghosh, M. and K. Sinha (2007). Empirical bayes estimation in finite population sampling under functional measurement error models. *Journal Statistical Planning and Inference 137*(9), 2759–2773.

Ghosh, M., K. Sinha, and D. Kim (2006). Empirical and hierarchical bayesian estimation in finite population sampling under structural measurement error models. *Scandinavian Journal of Statistics 33*(3), 591–608.

Ghosh, M. and R. Steorts (2013). Two-stage bayesian benchmarking as applied to small area estimation. *Test 22*(4), 670–687.

Glasser, G. J. (1962). Variance formulas for the mean difference and coefficient of concentration. *Journal of the American Statistical Association 57*(299), 648–654.

Graf, E. and Y. Tillé (2014). Variance estimation using linearization for poverty and social exclusion indicators. *Survey Methodology 40*(1), 61–79.

Graf, M., A. Alfons, C. Bruch, P. Filzmoser, B. Hulliger, R. Lehtonen, B. Meindl, R. Münnich, T. Schoch, M. Templ, M. Valaste, A. Wenger, and S. Zins (2011). *State-of-the-art of Indicators on Poverty and Social Exclusion - the Laeken Indicators*. Research Project Report WP1 (D1.1). FP7-SSH-2007-217322 AMELI. `https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Ameli_Delivrables/AMELI-WP1-D1.1-20110418.pdf` [30.07.2019].

G.U. (2016). Legge 4 agosto 2016, num. 163. *Gazzetta Ufficiale della Repubblica Italiana 198*. `http://www.gazzettaufficiale.it/eli/id/2016/08/25/16G00174/sg`.

Hall, P. and T. Maiti (2006a). Nonparametric estimation of mean-squared prediction error in nested-error regression models. *The Annals of Statistics 34*(4), 1733–1750.

Hall, P. and T. Maiti (2006b). On parametric bootstrap methods for small area prediction. *Journal of the Royal Statistical Society: Series B (Statistical Methodology) 68*(2), 221–238.

Haziza, D. and J.-F. Beaumont (2017). Construction of weights in surveys: a review. *Statistical Science 32*(2), 206–226.

Haziza, D. and É. Lesage (2016). A discussion of weighting procedures for unit nonresponse. *Journal of Official Statistics* (1), 129–145.

Hill, R. P. and B. Adrangi (1999). Global poverty and the United Nations. *Journal of Public Policy & Marketing 18*(2), 135–146.

Horneffer, B. and B. Kuchler (2008). Drei Jahre Panelerhebung EU-SILC. *Wirtschaft und Statistik* (8), 650–661.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American statistical Association 47*(260), 663–685.

Huang, E. and W. Fuller (1978). Non-negative regression estimation in sample survey data. *Proccedings Social Statitics Section, American Statistical Association*, 300–305.

Hulliger, B. and R. Münnich (2006). Variance estimation for complex surveys in the presence of outliers. In *Proceedings of the Section on Survey Research Methods. Survey Research Methods*, pp. 3153–3156.

International Monetary Fund (2018). World economic outlook database.

Isaki, C. T. and W. A. Fuller (1982). Survey design under the regression superpopulation model. *Journal of the American Statistical Association 77*(377), 89–96.

Istat (2008). L'indagine europea sui redditi e le condizioni di vita delle famiglie (Eu-Silc). Istat.

Istat (2018). *Rapporto Bes: Il benessere equo e sostenibile in Italia*. Istat. `http://www.istat.it/it/archivio/224669`.

Istat (2019). *Rapporto Bes: Il benessere equo e sostenibile in Italia*. Istat. `https://www.istat.it/it/archivio/236714`.

ITNRW (2018). Berechnung von Armutsgefährdungsquoten auf Basis des Mikrozensus. `http://www.amtliche-sozialberichterstattung.de/pdf/Berechnung%20von%20Armutsgefaehrdungsquoten%20auf%20Basis%20des%20Mikrozensus.pdf` [29.07.2019].

Jiang, J. and P. Lahiri (2006). Estimation of finite population domain means: A model-assisted empirical best prediction approach. *Journal of the American Statistical Association 101*(473), 301–311.

Jiang, J., T. Nguyen, and J. S. Rao (2011). Best predictive small area estimation. *Journal of the American Statistical Association 106*(494), 732–745.

Kott, P. S. (2003). A practical use for instrumental-variable calibration. *Journal of Official Statistics 19*(3), 265.

Kott, P. S. (2006). Using calibration weighting to adjust for nonresponse and coverage errors. *Survey Methodology 32*(2), 133.

Kovacevic, M. S. and D. A. Binder (1997). Variance estimation for measures of income inequality and polarization - the estimating equations approach. *Journal of Official Statistics 13*(1), 41–58.

Langel, M. and Y. Tillé (2013). Variance estimation of the gini index: revisiting a result several times published. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 176*(2), 521–540.

Lauro, C. N., M. G. Grassia, and R. Cataldo (2018). Model based composite indicators: New developments in partial least squares-path modeling for the building of different types of composite indicators. *Social Indicators Research 135*, 421–455.

Lavallée, P. (1995). Cross-sectional weighting of longitudinal surveys of individuals and households using the weight share method. *Survey Methodology 21*(1), 25–32.

Lavallée, P. (2007). *Indirect Sampling*. Springer.

Lehtonen, R., C. E. Särndal, and A. Veijanen (2003). The effect of model choice in estimation for domains, including small domains. *Survey methodology: a journal of Statistics Canada 29*(1), 33–44.

Lehtonen, R. and A. Veijanen (1999, August). Domain estimation with logistic generalized regression and related estimators. In *IASS Satellite Conference on Small Area Estimation*, pp. 121–128. Latvian Council of Science Riga.

Marchetti, S., M. Beresewicz, N. Salvati, M. Szymkowiak, and W. Lukasz (2018). The use of a three-level m -quantile model to map poverty at local administrative unit 1 in poland. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Marchetti, S., C. Giusti, M. Pratesi, N. Salvati, F. Giannotti, D. Pedreschi, S. Rinzivillo, L. Pappalardo, and L. Gabrielli (2015). Small area model-based estimators using big data sources. *Journal of Official Statistics 31*(2), 263–281.

Marchetti, S., N. Tzavidis, and M. Pratesi (2012). Non-parametric bootstrap mean squared error estimation for m-quantile estimators of small area averages, quantiles and poverty indicators. *Computational Statistics and Data Analysis*.

Marhuenda, Y., I. Molina, D. Morales, and J. Rao (2017). Poverty mapping in small areas under a twofold nested error regression model poverty mapping in small areas under a twofold nested error regression model. *Journal of the Royal Statistical Society: Series A (Statistics in Society) 180*(4), 1111–1136.

Mauro, V., M. Biggeri, and F. Maggino (2018). Measuring and monitoring poverty and well-being: A new approach for the synthesis of multidimensionality. *Social Indicators Research 135*, 75–89.

Mazziotta, M. and A. Pareto (2016). On a generalized non-compensatory composite index for measuring socio-economic phenomena. *Social Indicators Research 127*, 983–1003.

Militino, A., M. Ugarte, and T. Goicoa (2007). A blup synthetic versus and eblup estimator: an empirical study of a small area estimation problem. *Journal of Applied Statistics 34*, 153–165.

Molina, I. and J. N. K. Rao (2010). Small area estimation of poverty indicators. *Canadian Journal of Statistics 38*(3), 369–385.

Münnich, R. and S. Zins (2011). *Variance Estimation for Complex Surveys*. Research Project Report WP3 (D3.2). FP7-SSH-2007-217322 AMELI. `https://www.uni-trier.de/fileadmin/fb4/projekte/SurveyStatisticsNet/Ameli_Delivrables/AMELI-WP3-D3.2-20110515.pdf` [30.07.2019].

OECD (2011). *Compendium of OECD Well-Being Indicators.* Paris: OECD Publications.

OECD (2016). *Better Policies for Sustainable Development 2016: A New Framework for Policy Coherence.* OECD Publishing, Paris.

OECD (2017a). *Education at a Glance 2017: OECD Indicators.*

OECD (2017b). Terms of Reference: OECD Project on the Distribution of Household Incomes. `https://www.oecd.org/els/soc/IDD-ToR.pdf` [03.06.2019].

OECD (2019). OECD Income Distribution Database (IDD): Gini, poverty, income, Methods and Concepts. `https://www.oecd.org/els/soc/income-distribution-database.htm` [03.06.2019].

Olsen, M. K. and J. L. Schafer (2001). A two-part random-effects model for semicontinuous longitudinal data. *Journal of the American Statistical Association 96*(454), 730–745.

Osier, G. (2009). Variance estimation for complex indicators of poverty and inequality using linearization techniques. *Survey Research Methods 3*(3), 167–195.

Osier, G., Y. Berger, and T. Goedemé (2013). Standard error estimation for the EU-SILC indicators of poverty and social exclusion. Technical report, Eurostat, `https://ec.europa.eu/eurostat/documents/3888793/5855973/KS-RA-13-024-EN.PDF` [20.07.2019].

Petrucci, A. and N. Salvati (2006). Small area estimation for spatial correlation in watershed erosion assessment. *Journal of agricultural, biological, and environmental statistics 11*(2), 169.

Pfefferman, D. and A. Sikov (2011). Imputation and estimation under nonignorable nonresponse in household surveys with missing covariate information. *Journal of Official Statistics 27*(2), 181–209.

Pfeffermann, D. (2013). New important developments in small area estimation. *Statistical Science 28*(1), 40–68.

Pfeffermann, D. and C. H. Barnard (1991). Some new estimators for small-area means with application to the assessment of farmland values. *Journal of Business & Economic Statistics 9*(1), 73–84.

Pfeffermann, D., A. Sikov, and R. Tiller (2014). Single and two-stage cross-sectional and time series benchmarking procedures for small area estimation. *Test 23*(4), 631–666.

Pfeffermann, D., B. Terryn, and F. A. Moura (2008). Small area estimation under a two-part random effects model with application to estimation of literacy in developing countries. *Survey Methodology 34*(2), 235–249.

Polettini, S. and S. Arima (2015). Small area estimation with covariates perturbed for disclosure limitation. *Statistica 25*(1), 57–72.

Potsi, A., A. D'Agostino, C. Giusti, and P. L. (2016). Childhood and capability deprivation in Italy: a multidimensional and fuzzy set approach. *Quality & Quantity 50*, 2571–2590.

Prasad, N. G. N. and J. N. K. Rao (1999). On robust small area estimation using a simple random effects model. *Survey Methodology 25*, 67–72.

Prasad, N. N. and J. N. Rao (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American statistical association 85*(409), 163–171.

Pratesi, M. (2015). Spatial disaggregation and small area estimation methods for agricultural surveys: Solutions and perspectives. Technical report, FAO Technical Report Series GO-07-2015.

Pratesi, M. and A. Petrucci (2014). Methodological and operational solutions to gaps and issues on methods for producing agricultural and rural statistics at small domains level and on methods for aggregation, disaggregation and integration of different kinds of geo-referenced data for increasing the efficiency of agricultural and rural statistics. Technical report, FAO report.

Pratesi, M. and N. Salvati (2009). Small area estimation in the presence of correlated random area effects. *Journal of Official Statistics 25*(1), 37.

Pratesi, M. and N. Salvati (2016). *Analysis of Poverty Data by Small Area Estimation*, Chapter Introduction on Measuring Poverty at Local Level Using Small Area Estimation Methods, pp. 1–18. John Wiley & Sons.

Rao, J. N. K. (2003). *Small Area Estimation.* New York: Wiley.

Rao, J. N. K. and I. Molina (2015). *Small Area Estimation.* Hoboken, New Jersey: John Wiley & Sons.

Royal, R. (1976). The linear least-squares prediction approach to two-stage sampling. *Journal of the American Statistical Association 71*, 657–664.

Salvati, N., N. Tzavidis, M. Pratesi, and R. Chambers (2012). Small area estimation via M-quantile geographically weighted regression. *Test 21*(1), 1–28.

Santourian, A. and E. Ntakou (2014). Working paper with the description of the 'income and living conditions dataset'.

Särndal, C. E. (1982). Implications of survey design for generalized regression estimation of linear functions. *Journal of Statistical Planning and Inference 7*, 155–170.

Särndal, C.-E. (2007). The calibration approach in survey theory and practice. *Survey Methodology 33*(2), 99–119.

Särndal, C.-E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling.* New York: Springer Series in Statistics.

Schmid, T., F. Bruckschen, N. Salvati, and T. Zbiranski (2017). Constructing sociodemographic indicators for national statistical institutes by using mobile phone data: estimating literacy rates in senegal. *Journal of the Royal Statistical Society - Series A 180*(4), 1163–1190.

Sen, A. R. (1953). On the estimate of the variance in sampling with varying probabilities. *Journal of the Indian Society of Agricultural Statistics 5*, 119–127.

Singh, A. and C. Mohl (1996). Understanding calibration estimators in survey sampling. *Survey methodology 22*(2), 107–115.

Statistische Ämter des Bundes und der Länder (2019). Sozialberichterstattung der amtlichen Statistik. http://www.amtliche-sozialberichterstattung.de/ [29.07.2019].

Stiglitz, J. E., A. Sen, and J. P. Fitoussi (2009). *Report by the commission on the measurement of economic performance and social progress.* Paris: Commission on the Measurement of Economic Performance and Social Progress.

Stukel, D., M. Hidiroglou, and C.-E. Särndal (1996). Variance estimation for calibration estimators: a comparison of jackknifing versus taylor linearization. *Survey Methodology 22*(2), 117–126.

Sverchkov, M. and D. Pfefferman (2018). Small area estimation under informative sampling and not missing at random non-response. *Journal of the Royal Statistical Society - Series A 181*(4), 981–1008.

Tinto, A., F. Bacchini, B. Baldazzi, A. Ferruzza, J. A. van den Brakel, R. M. A. Willems, N. Rosenski, T. Zimmermann, Z. Andrási, M. Farkas, and Z. Fábián (2018). MAKSWELL Deliverable 1.1 - Report on international and national experiences and main insight for policy use of well-being and sustainability framework. Technical report.

Tinto, A. and B. Baldazzi (2018). MAKSWELL Deliverable 1.2 - Definition of the existing database on Beyond GDP initiatives within official statistics. Technical report.

Torabi, M., G. Datta, and J. Rao (2009). Empirical bayes estimation of small area means under nested error linear regression model with measurement errors in the covariates. *Scandinavian Journal of Statistics 36*(2), 355–368.

Torabi, M. and J. N. Rao (2010). Mean squared error estimators of small area means using survey weights. *Canadian Journal of Statistics 38*(4), 598–608.

Tzavidis, N., S. Marchetti, and R. Chambers (2010). Robust estimation of small area means and quantiles. *Australian and New Zeland Journal of Statistics 52*(2), 167–186.

Tzavidis., N., L. Zhang, A. Luna, T. Schmid, and N. Rojas-Perilla (2018). From start to finish: a framework for the production of small area official statistics. *Journal of the Royal Statistical Society - Series A 181*(4), 927–979.

UN Department of Economic and Social Affairs (2017). World population prospects: The 2017 revision. https://esa.un.org/unpd/wpp/ [29.07.2019].

UN Statistics Division (2018). National accounts main aggregate database. http://unstats.un.org/unsd/snaama [29.07.2019].

UNDP (1997). *Human Development Report 1997.* New York, Oxford University Press.

UNDP (2018a). *Human Development Indices and Indicators: 2018 Statistical Update.*

UNDP (2018b). *Human Development Indices and Indicators: 2018 Statistical Update. Technical Notes.*

UNESCO Institute for Statistics (2018). Data centre. http://data.uis.unesco.org [29.07.2019].

U.S. Census Bureau (2018a). Small Area Income and Poverty Estimates (SAIPE) program: 2010-2017 state-level estimation detail. `https://www.census.gov/programs-surveys/saipe/technical-documentation/methodology/counties-states/state-level.html` [28.05.2019].

U.S. Census Bureau (2018b). Small Area Income and Poverty Estimates (SAIPE) program: About. `https://www.census.gov/programs-surveys/saipe/about.html` [28.05.2019].

U.S. Census Bureau (2019). Small Area Income and Poverty Estimates (SAIPE) program: 2010-2017 overview of school district estimates. `https://www.census.gov/programs-surveys/saipe/technical-documentation/methodology/school-districts/overview-school-district.html` [03.06.2019].

van den Brakel, J. (2016). Register-based sampling for household panels. *Survey Methodology 42*(1), 137–159.

van den Brakel, J. and J. Bethlehem (2008). *Model-Based Estimation for Official Statistics*. Statistics Netherlands.

van den Brakel, J. A., P. A. Smith, N. Tzavidis, R. Iannaccone, D. Zurlo, F. Bacchini, L. Di Consiglio, T. Tuoto, M. Pratesi, C. Giusti, S. Marchetti, S. Bastianoni, G. Betti, A. Lemmi, F. M. Pulselli, and L. Neri (2019). MAKSWELL Deliverable 2.2 - Methodological aspects of using Big data. Technical report.

Wang, J., W. Fuller, and Y. Qu (2008). Small area estimation under restrictions. *Survey Methodology 34*(1), 29–36.

Wittenberg, M. (2010). An introduction to maximum entropy and minimum cross-entropy estimation using Stata. *Stata Journal 10*(3), 315–330.

World Bank (2018a). *Poverty and Shared Propserity 2018: Piecing Together the Poverty Puzzle*. Washington, DC: World Bank.

World Bank (2018b). World development indicators database. `http://data.worldbank.org` [29.07.2019].

World Health Organization Expert Committee on Physical Status (1996). The use and interpretation of anthropometry. physical status: Report of a WHO expert committee. Technical Report 854, WHO Technical Report Series.

Wu, C. and W. W. Lu (2016). Calibration weighting methods for complex surveys. *International Statistical Review 84*(1), 79–98.

Yates, F. and P. M. Grundy (1953). Selection without replacement from within strata with probability proportional to size. *Journal of the Royal Statistical Society: Series B (Methodological) 15*(2), 253–261.

Ybarra, L. and S. Lohr (2008). Small area estimation when auxiliary information is measured with error. *Biometrika 95*(4), 919–931.

You, Y. and J. Rao (2002). A pseudo-empirical best linear unbiased prediction approach to small area estimation using survey weights. *Canadian Journal of Statistics 30*(3), 431–439.