Measuring Small-Scale At-Risk-of-Poverty in Germany – a Methodical Overview

Stefan Kaup

(Stefan Kaup, ILS – Research Institute for Urban and Regional Development, Bruederweg 22-24 44135 Dortmund, stefan.kaup@ils-forschung.de)

1 ABSTRACT

Regarding the EU 2020 initiative of the European Commision, one of the main targets for the next years is to reduce the number of people in or at risk of poverty and social exclusion in Europe by 20 million. (European Commission 2014a). This target aims different aspects of the financial setting and social participation of individuals and groups and has to be operationalizes in mesurable indicators, that could cover the major domains and dimensions of the complex theme (Copus 2014).

The latest publication of the German Federal Statistical Office speaks of 20.3 % of the German population affected by poverty or social exclusion. This term is a multi-variate definition based on indicators related to people at risk of poverty (16.1 %), people affected by massive material depriviation (5.4 %) and people living in households with very low income (9.9 %) (Destatis 2014a). It focuses mainly on the financial aspects. The first of the three indicators is measured by the so called At-Risk-of-Poverty rate, which is defined by Eurostat as the "share of people with an equivalised disposable income (after social transfer) below the At-Risk-of-Poverty threshold, which is set at 60 % of the national median equivalised disposable income after social transfers" (Eurostat 2014a). The underlying datasets for the German indicators come from the EU-SILC, an EU wide annual survey of income and living conditions (Eurostat 2014c). This survey provides the possibility to calculate statistics down to the NUTS 2 regions. A regional level that in Germany is called Government regions. For decision makers on the regional or local level, this computation is not good enough in terms of spatial resolution. So how to come to more useful numbers?

The author here discusses possibilities to create At-Risk-of-Poverty rates in Germany on a higher spatial resolution. An own elaboration based on a linear regression model is cross compared with an approach based on the German Microcensus.

2 POVERTY – THE THEORETICAL AND POLITICAL FRAMEWORK

Poverty as an expression has an absoulte and a relative meaning. Mainly in the context of less developed countries, the absolute term is in focus. It "measures poverty in relation to the amount of money necessary to meet basic needs such as food, clothing, and shelter" (UNESCO 2015). The UN Millenium Project indicates the eradiction of extreme hunger and poverty as its first Millennium Development Goal. For the operationalization of the goal extreme poverty is defined by an income less than 1 US-$ per day (United Nations Development Programme 2001).

In the European context there are no bigger parts of society under such depriviation of basic needs. In this environment poverty is defined as a relative measure for the integration of social groups into the overall economic situation of the whole society. It does not reflect a direct threat for the individuals lifes or well-being but it is a measure for the unequality of societies which could lead to social disharmony. "Relative poverty defines poverty in relation to the economic status of other members of the society: people are poor if they fall below prevailing standards of living in a given societal context." (UNESCO 2015)

"Prevailing standards" as given in the definition of the UNESCO is a pretty vague concept. It has to be underlaid with concrete statistical definitions to be well-measured. Therefore there has to be an agreement about wich information under which statistical procedure calculates the right or best fitting indicator for the concept. In 2001 the European Union has developed a set of 18 indicators in addition to the Lisbon strategy to measure poverty and social exclusion. These indicators are regulary produced for every European (EU) country. In reference to the place where they where agreed on, they are called the Laeken Indicators. (European Commission 2015)

One of the most prominent is the so called At-Risk-of-Poverty rate. It is calculated as the share of people with an equilized disposable income below 60 % of the Median income of the national state. (European Commission 2014b)

ISBN: 978-3-9503110-8-2 (CD-ROM); ISBN: 978-3-9503110-9-9 (Print)
Editors: M. SCHRENK, V. V. POPOVICH, P. ZEILE, P. ELISEI, C. BEYER

3 MEASURING AT-RISK-OF-POVERTY ON A SMALL SCALE

The Laeken Indicators in general and the At-Risk-of-Poverty rate in specific are produced for the EU member states at the national state level. In many fields of decision making, this country wide view does not reflect the level of spatial detail that is needed to make good and suitable decisions. Regional divisions have to be made and therefore statistical procedures are needed to make valid and solid calculations for these (spatial) subgroups, these more or less “small areas”.

3.1 The data sources

An important precondition to make good small area estimations for the At-Risk-of-Poverty rate is a suitable data souce. In general there are two types of source that estimations can rely on, census and survey.

Censuses have a long history in human societies, since it was an appropriate tool to estimate the tax income for the authorities. In the last decades, Germany had a lack of full census. The last one was held in 1987 (Western Germany) and 1981 (Eastern Germany) which means that there is practically no recent census data that could be used for an up-to-date calculation of At-risk-of-Poverty rates. With the regulation 763/2008 the European parliament and the Council of the European Union established common rules for a decennial provision of comprehensive data on population and housing. In Germany the first results from the first Census under that regulation in 2011 where published in 2014 but limited to aggregation data at city or regional level. The individual and household data sets are announced to be published during 2015. (Research Data Centres of the Federal Statistical Office and the statistical offices of the Länder 2014)

The statistical institutions of Germany and the European Union carry out different surveys including information that could be used for the measurement of poverty. In Germany the so called Microcensus is the biggest one. It is a nearly 1 % sample of the population including 830 000 persons and 370 000 households. The first Microcensus was carried out in 1957 and since 1991 annual data is available for both parts of Germany (Research Data Centres of the Federal Statistical Office and the statistical offices of the Länder 2015). The data sets have a regional identifier, the so called adaption layer (Anpassungsschicht). This regional level is a disjunct tesselation of Germany. Single regions should not be smaller than 500 000 inhabitants. That means large regional clusters for less populated regions. So the federal state of Mecklenburg-Vorpommern is subdivided into just three of them.

The European Union Statistics on Income and Living Conditions (EU-SILC) was launched in 2004 for the EU-15 members. It is a survey “aiming at collecting timely and comparable cross-sectional and longitudinal multidimensional microdata on income, poverty, social exclusion and living conditions” (Eurostat 2014b). The sample size is 28 000 persons in Germany (14 000 households) which is less than 4 % of the Microcensus sample.

3.2 The methodologies

The methodologies to estimate or calculate At-Risk-of-Poverty rates vary by the availability of the data. Despite the scepticism in official statistics on putting already established methods in use there is a lively advancement in the field of small area estimation (Münnich et al. 2013: 187). In general the approaches can be divided into design based and model based ones. Both of them include available informations from neighbouring or similar fields in their estimations. Design based approaches are the base of modern sample theory. They rely on sampling designs including random sampling, two- or more-stages designs. When small sample sizes occure, these design based approaches have very high variances of the estimation function. Model based approaches overcome this uncertainty by replacing the direct estimation by an indirect model based estimation (Münnich et al. 2013: 151 ff). Examples for design based estimators are the Horvitz-Thompson estimator and the generalized regression estimator. Examples for model based estimators are the synthetic estimator, the EBLUP (empirical best linear unbiased predictor) and EBP estimator (LEHTONEN et al. 2011: 13 ff).

It can be stated, that the advantages and disadvantages of design based and model based approaches stand opposite to each other. Therefore most of the modern small area estimations use compound estimation models. The best known approaches are the Fay-Herriot-Estimator (FH) and the Battese-Harter-Fuller-Estimator (BHE). They are both special cases of general mixed models whereas the FH uses data on an aggregated level (area level model) and the BHE on individual level (unit level model). (Münnich et al. 2013: 161 ff)

4 LATEST APPROACHES IN GERMANY

So far there are two recent approaches carried out trying to estimate At-Risk-of-Poverty rates for small areas in Germany. The first one comes from the Cologne Institute for Economic Research (IW), the second one is part of the ESPON project TiPSE – Territorial Dimensions of Poverty and Social Exclusion in Europe. Both approaches where published in 2014. They are built on different data sources and methodologies and therefore provide different distributions and interpretations.

4.1 The IW Cologne approach

The IW approach was based on German Microcensus 2012 data. This dataset consists of individual records including income values and the regional identifier for the adaption layer. The provided sample for the study is a 70 % subsample of the whole Microcensus of that year. The definition of relative poverty is according to the 60 % median income definition given by Eurostat (IW Cologne 2014). Figure 1 shows the calcutaled At-Risk-of-Poverty rates as provided by IW. In the approach, they extend the methodology by including regional prize niveaus as a counterbalance. They argue that poverty rates in regions having a low prize niveau compared to those with a higher one are not comparable. People with an income below that rate are under much higher financial pressure in high cost regions than in low cost regions. But since there is no counterbalancing in the second approach, we here focus on the flat At-Risk-of-Poverty rates as shown in the figure. The regional identifier for the adaption layer as the only one included in the data set causes that there is no intra-regional differentiation for the unterlying NUTS 3 regions in respect to the At-Risk-of-Poverty rate. By showing the NUTS 3 borders, the provided maps of the IW approach pretend to depict a detail of spatial distribution of the indicator that is actually not given. Especially in the more rural regions, higher number of NUTS 3 regions (up to seven) form one unique region for a common value.

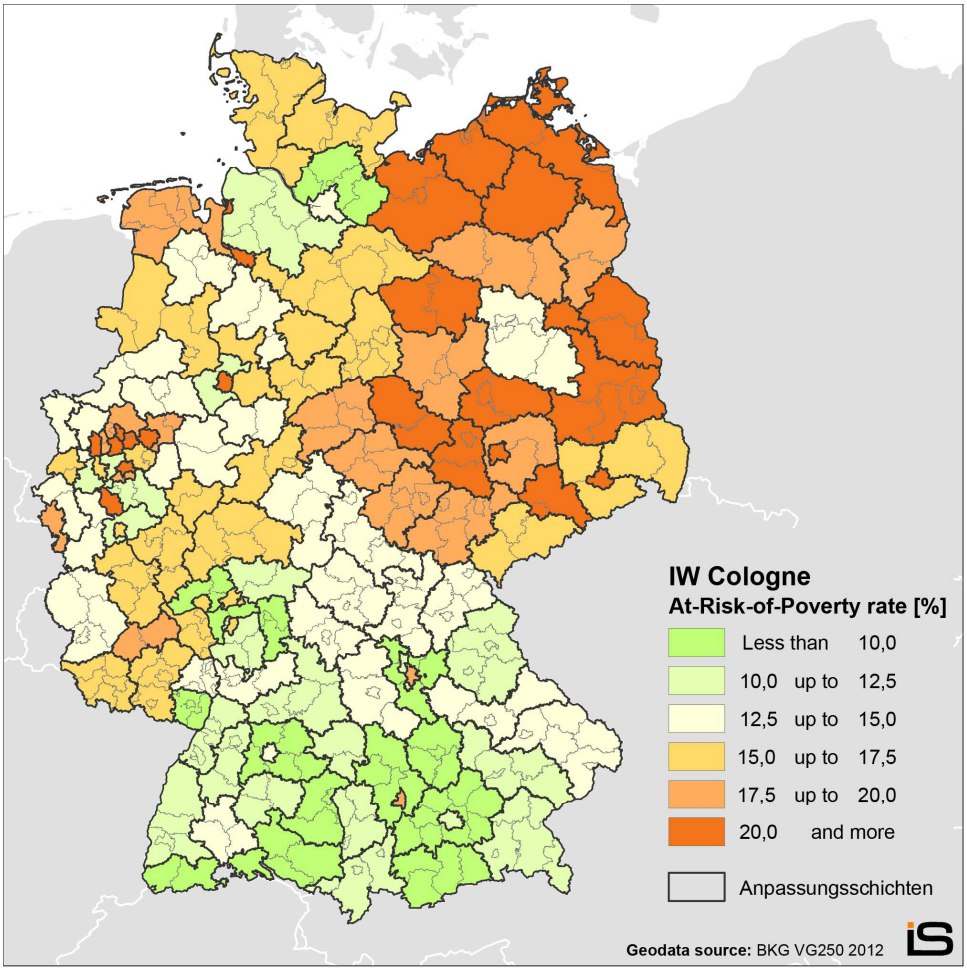

Fig. 1: The IW Cologne estimations of At-Risk-of-Poverty rates 2012 (IW Cologne 2014)

ISBN: 978-3-9503110-8-2 (CD-ROM); ISBN: 978-3-9503110-9-9 (Print)
Editors: M. SCHRENK, V. V. POPOVICH, P. ZEILE, P. ELISEI, C. BEYER

The map of At-Risk-of-Poverty rates clearly shows a difference between the western and eastern parts of Germany. But on top of this west-east gradient an additional south-north one plays an important role. So the highest values can be found in the new federal states and eastern parts of Lower Saxony, lower values in the southern federal states of Bavaria and Baden-Würtemberg. Besides this macro regional differences, higher values seem to be found in core cities rather than in rural areas. Since most of the core cities are integrated into bigger adaption layer regions, this exposition is covered and can mainly be seen for the cities in Northrine-Westfalia (Bielefeld, Cologne, Ruhr area).

The approach relies on a strong database namely a Micocensus subsample. Since this is the biggest survey carried out by the statistical institutes including income attributes, it is the most reliable. The weakness lies in the regional classification system of the adaption layers. On the one hand, this (Microcensus) specific classification provides comparable regions in term of population number. On the other hand it combines NUTS 3 regions of different structur: cities with rural areas, central with remote areas. Combining these regions under a bigger one causes statistical effects that could cover intra-regional differences regarding the At-Risk-of-Poverty rate.

4.2 The ESPON TiPSE approach

The ESPON TiPSE approach is based on a linear regression model. This model uses precalculated variables from the Eurostat and the German Census 2011 databases. The indicators are available on NUTS 2 as well as on NUTS 3 level. The linear regression model is built on NUTS 2 data. At this level there are 39 regions in Germany representing the degrees of freedom in the model. The independent variables come from socio-demographic domains based on Eurostat, Census and Microcensus data 2010 and 2011. The dependend variable of the model is the precalculated At-Risk-of-Poverty rate from the German System Of Social Reporting (Federal Statistical Office and the statistical offices of the Länder 2013). Figure 2 shows the systematic of the variables used in the model.

Demographic characteristics
Population on 1 January – Less than 15 years [demo_r_pjanaggr3] Population on 1 January – From 15 to 64 years [demo_r_pjanaggr3] Population on 1 January – 65 years or over [demo_r_pjanaggr3]
Socio- economic characteristics
Employment in NACE A – Agriculture, forestry and fishing [nama_r_e3em95r2] Employment in NACE B-E – Industry (except construction) [nama_r_e3em95r2] Employment in NACE C – Manufacturing [nama_r_e3em95r2]
Housing characteristics
2 dwellings in the building Detached house, Semi-detached house, Terraced house, Other type of building
Socio- economic characteristics
Persons in employment, Unemployed persons At risk of poverty rate (NUTS 2)

Fig. 2: The ESPON TiPSE approach data sources

The correlations between the independend and the dependend variables are mainly strong. The weakest correlation is -0.25 (GDP) up to 0.89 (Unemployment rate). The model does not include housing costs and rely on already estimated At-Risk-of-Poverty rates at the higher regional level NUTS 2. These already include an estimation error since they are taken from Microcensus (survey) data.

After the calculation of the values using the covariates of the regression model, an additional factor is being calculated for each NUTS 3 region to synchronize the total number of people at risk of poverty with the total number on the including NUTS 2 region.

The model produces some very strong outliers that should be excluded from the interpretation. Especially in the NUTS 2 region DED5 (Region Leipzig) there are three NUTS 3 regions with very high At-Risk-of-Poverty rates. In addition three regions in Mecklenburg-Vorpommern and the region of Bremerhaven seem to be miscalculated.

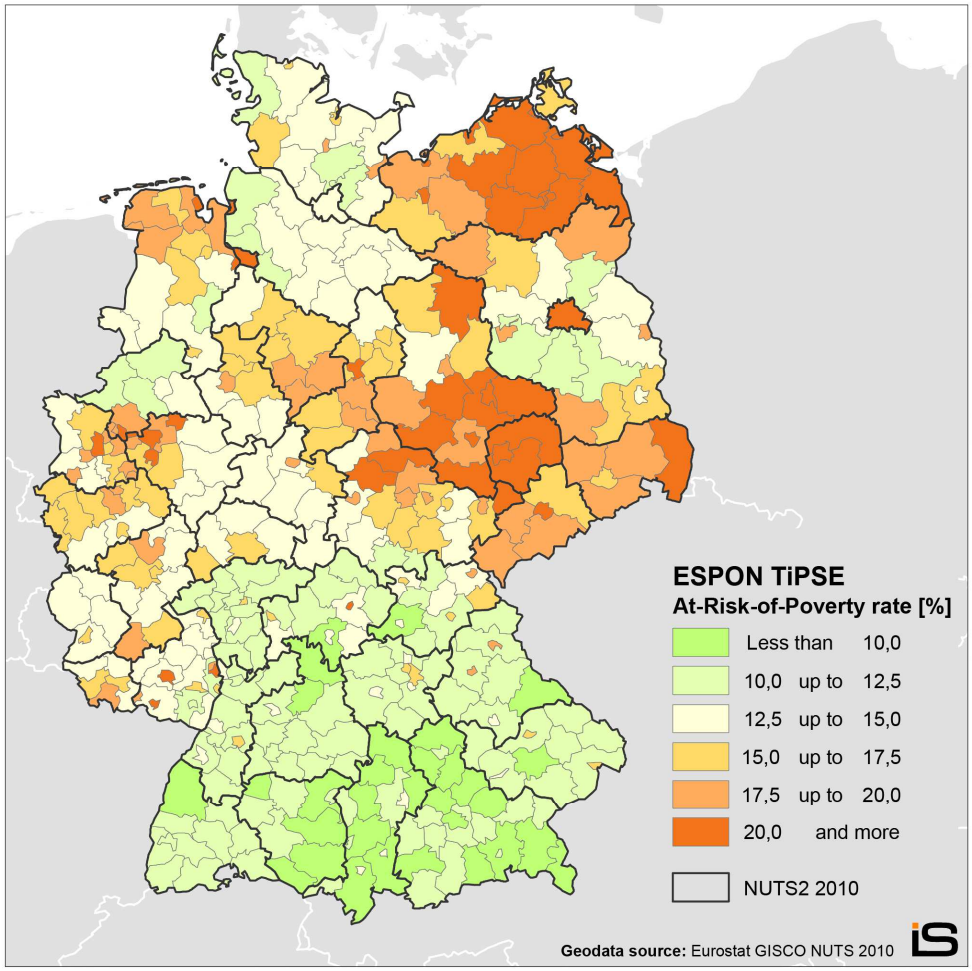

Fig. 3: The At-Risk-of-Poverty rates of the ESPON TiPSE project (ESPON Database 2015)

The map shows a spatial distribution of values, that is quite close to the IW Cologne approach.

Wide areas of the eastern parts of Germany show the highest values including Berlin. As an exception the Berlin surrounding regions of Brandenburg are at a lower level of At-Risk-of-Poverty. The southern regions of Germany in Bavaria, Baden-Würtemberg and Hesse show lower values whereby the northern regions of the western part of Germany are more complex to describe. The distribution does not follow macro regional gradients. Higher values can be found in the core cities of Northrine-Westfalia, the north-west and south-east part of Lower Saxony.

The spatial distribution of the values of course depends on the regression model and the underlying independend variables. High values in one dimension of the factors can cause a high output in the estimation. Although the variables correlate high with the At-Risk-of-Poverty rate, in some cases the dependency could be misleading. For example the old age dependency rate (people over 64 in relation to people between 15 and 64) is high correlated with the dependend variable. That means that where there are more people over 64 in relation to the working group, people are more likely to be at risk of poverty. (Because people aged 64 and older are more likely to be out of work and have lower income). Because some regions with high quality of life at the coast or in rural areas are pulling older people with higher income, the factor is wrong for them but still goes into the estimation for the region. The effect can clearly be seen in the north-western part of Lower Saxony including the islands and the coast.

ISBN: 978-3-9503110-8-2 (CD-ROM); ISBN: 978-3-9503110-9-9 (Print)
Editors: M. SCHRENK, V. V. POPOVICH, P. ZEILE, P. ELISEI, C. BEYER

5 COMPARISON

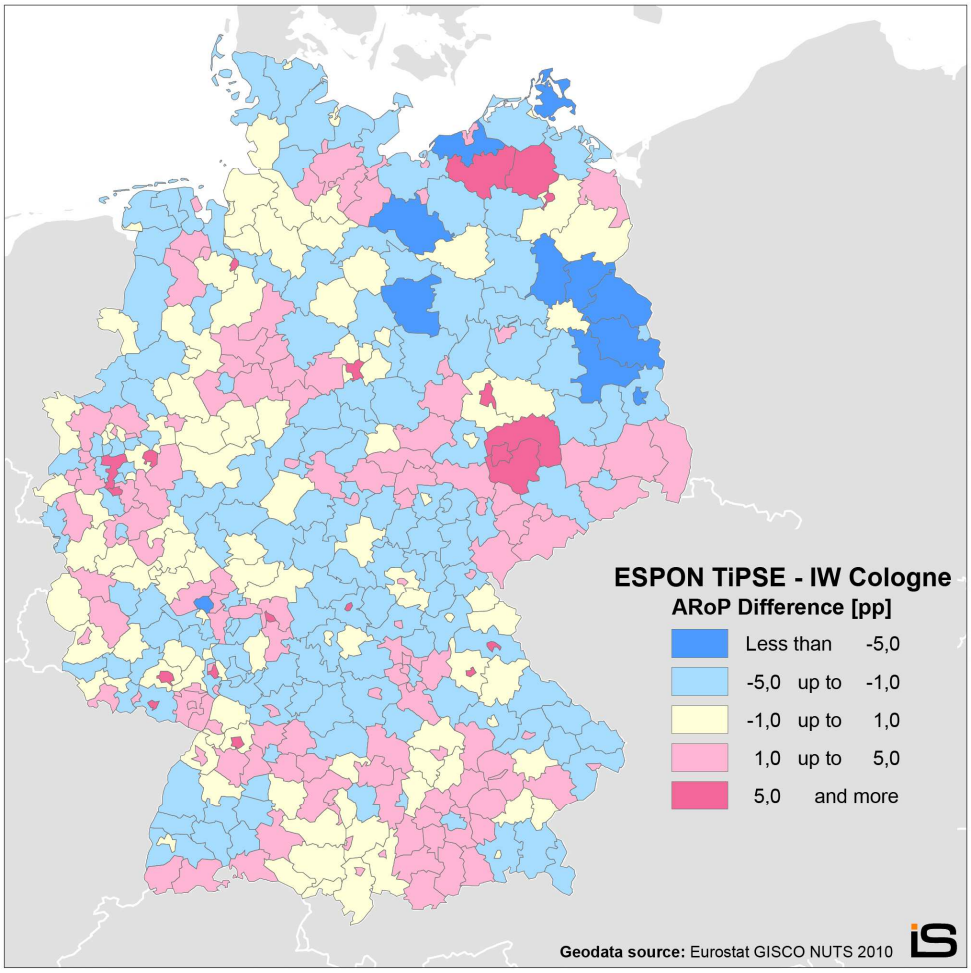

Fig. 4: Differences between the IW Cologne and the ESPON TiPSE approaches (Own calculation)

As already mentioned, there are clearly some outstanding overestimations in the ESPON TiPSE approach due to methodological restrictions and availability of variables. This has to be pointed for the NUTS 2 region DED5 (Region Leipzig) and for two regions in Mecklenburg-Vorpommern which should be excluded from the comparison. Besides that the both approaches do not neccesarily estimate the same values. Figure 4 shows the differences between them in percentage points (pp), where the blue colours mean higher values for the IW Cologne approach, the red colours higher values for ESPON TiPSE. Yellow coloured regions are at the same level of estimation.

The class of higher estimations of the IW approach with more than 5 pp (higher than ESPON TiPSE) is pretty small and include mainly regions in eastern parts of Germany, in particular the eastern, Berlin surrounding regions of Brandenburg. In the western federal states it is only the city of Frankfurt am Main.

The class of higher estimations of ESPON TiPSE with more than 5 pp (higher than IW Cologne) mainly consists of core cities in western parts of Germany that don't build their own addaption layer region. This could be a statistical effect of the regionalisation in the Microcensus data that combines regions with high and regions with low values which average in the bigger adaption layer region.

The classes of inequality up to 5 pp show a more disperse picture. The most visible effect is that the ESPON TiPSE approach calculates higher values for the NUTS 3 regions being part of bigger metropolitan regions but not the core cities themselves. This can be states for the regions of Munich, Stuttgart, Frankfurt am Main, Dresden, Hannover, Hamburg (northern parts), and the Rhein-Ruhr area. The IW approach estimates higher values for the more remote areas.

6 REFERENCES

COPUS, Andrew (2014): TiPSE – The Territorial Dimension of Poverty and Social Exclusion in Europe. Draft Final Report. Online: http://www.espon.eu/export/sites/default/Documents/Projects/AppliedResearch/TIPSE/DFR/TIPSE_Draft_Final_Report.pdf

DESTATIS (2014a): 20,3 % der Bevölkerung Deutschlands von Armut oder sozialer Ausgrenzung betroffen. Online: https://www.destatis.de/DE/PresseService/Presse/Pressemitteilungen/2014/12/PD14_454_634.html

DESTATIS (2014b): System of Social reporting. Online: http://www.amtliche-sozialberichterstattung.de/index_en.html

EUROPEAN COMMISSION (2014a): Europe 2020 in an nutshell. http://ec.europa.eu/europe2020/europe-2020-in-a-nutshell/index_en.htm

EUROPEAN COMMISSION (2014b): Glossary:At-risk-of-poverty rate. Online: http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At-risk-of-poverty_rate

EUROPEAN COMMISSION (2015): Income and living conditions (ilc). Online: http://ec.europa.eu/eurostat/cache/metadata/en/ilc_esms.htm

EUROSTAT (2014a): Glossary:At-risk-of-poverty rate. Online: http://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:At-risk-of-poverty_rate

EUROSTAT (2014b): European Union Statistics on Income and Living Conditions (EU-SILC). Online: ec.europa.eu/eurostat/web/microdata/european_union_statistics_on_income_and_living_conditions

EUROSTAT (2014c): Income and living conditions. Online: http://ec.europa.eu/eurostat/web/income-and-living-conditions/overview

FEDERAL STATISTICAL OFFICE AND THE STATISTICAL OFFICES OF THE LÄNDER (2013): System Of Social Reporting. Online: http://www.amtliche-sozialberichterstattung.de/index_en.html

IW COLOGNE (2014): Einkommensarmut in Deutschland aus regionaler Sicht. Online: http://www.iwkoeln.de/_storage/asset/180421/storage/master/file/5047291/download/PK%20Einkommensarmut%20Materialien.pdf

LEHTONEN, Risto; VEIJANEN, Ari; MYRSKYLÄ, Mikko and VALASTE, Maria (2011): Small Area Estimation of Indicators on Poverty and Social Exclusion.

MÜNNICH, Ralf; BURGARD, Jan and VOGT, Martin: Small Area-Statistik: Methoden und Anwendungen. In AStA Wirtschafts- und Sozialstatistisches Archiv, Vol. 6, issue 3-4, pp. 149–191

RESEARCH DATA CENTRES OF THE FEDERAL STATISTICAL OFFICE AND THE STATISTICAL OFFICES OF THE LÄNDER (2014): Mikrodaten des Zensus künftig über die FDZ verfügbar Online: http://www.forschungsdatenzentrum.de/bestand/zensus_2011/index.asp

RESEARCH DATA CENTRES OF THE FEDERAL STATISTICAL OFFICE AND THE STATISTICAL OFFICES OF THE LÄNDER (2015): Datenangebot Mikrozensus Online: http://www.forschungsdatenzentrum.de/bestand/mikrozensus/

THE WORLD BANK (2014): Poverty Research. Online: http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTPROGRAMS/EXTPOVRES/0,,menuPK:477905~pagePK:64168176~piPK:64168140~theSitePK:477894,00.html

UNESCO (2015): Poverty. Online: http://www.unesco.org/new/en/social-and-human-sciences/themes/international-migration/glossary/poverty/

UNITED NATIONS DEVELOPMENT PROGRAMME (2001): UN Millenium Project. Millennium Development Goals. Online: http://www.unmillenniumproject.org/goals/index.htm

ISBN: 978-3-9503110-8-2 (CD-ROM); ISBN: 978-3-9503110-9-9 (Print)
Editors: M. SCHRENK, V. V. POPOVICH, P. ZEILE, P. ELISEI, C. BEYER