MODELLING OF THE PROBABILITY DISTRIBUTION OF THE DEPRIVATION INDICES IN EUROPEAN COUNTRIES

Ivana Malá

Abstract

Two deprivation indices were introduced in the Survey of Health, Ageing and Retirement database in Europe (SHARE) in 2013. The comprehensive survey covers the process of ageing in the European population; it forms a panel database of data concerning inhabitants aged above 50. The probability distributions of both - material and social - deprivation indices are modelled for 15 participating countries in wave 5. The countries are clustered according to the characteristics (quartiles and probability of severe deprivation) and shapes of empirical distributions. There are differences between countries in probability distributions reflecting the situation of deprivation in the elderly population, and the participant countries are inhomogeneous concerning this distribution. The distributions are also different for quantified phenomena (material and social deprivation), as the social deprivation index seems to obtain higher values than the material index. There is a significant proportion of respondents without any material problems defined by the index (from 16% in Estonia to 77% in Denmark); for social deprivation, we obtain only 2.5% (from 0.5% in Spain to 5.4% in Sweden). The normal mixture distributions are used to fit distributions.

Key words: SHARE data, distribution fitting, deprivation, ageing, cluster analysis

JEL Code: I32, C43, C38

Introduction

Quantifying subjective phenomena such as deprivation and exclusion is a task for a wide spectrum of researchers. We have similar problems to quantify the quality of life or similar phenomena based on subjective perception. There are two main approaches; we use a question (or questions), or the composite indicators are constructed to obtain a more objective description. The difference between the described state (or its perception) and the quantification (indicator or index) is called an adequacy gap. We have to take it into account in all interpretations. There are different indicators frequently used to describe deprivation, loneliness or, on the other hand, quality of life, especially at a later age. For material and

social deprivation, more composite indicators are used to reflect situation in households reliably; these indicators try to include more characteristics than just the household income (Bellani, D´Ambrosio, 2011; Bellani, 2013). At present, no construct is generally accepted through the scientific community as the best and universally applicable. It is important to remember that such an effort is always just an attempt to quantify a very general situation involving a strong subjective element involving personal perception and feelings. With the ageing of a European population, the indices oriented or adapted to inhabitants at later ages are required for authorities to quantify the situation of this growing part of the population (van Leeuwen et al., 2019 for quality of life at later ages).

In the text, we analyse two indices proposed for the description of material and social deprivation included in the 5th wave of SHARE survey. The included states emphasise the well-being of their inhabitants, and they take place at the top of the world rankings characterising well-being, quality of life or just income per capita (Malá, 2019b). We aim to cluster participating countries with respect to the probability distribution of analysed indices or probability of severe deprivation, combining both problems.

1 Deprivation indices

The European population above 50 is the target population of the extensive survey Survey of Health, Ageing and Retirement in Europe (SHARE Release Guide, 2019; Börsch-Supan, 2016); the analysed indices of material deprivation *depmat* and social deprivation *depsoc* are adapted to the situation of elderly respondents (Adena et al., 2015). The included questions in the questionnaire indicate material inconvenience or problems connected with the availability of services, medical care and social relations. In the 5th wave, 14 European countries and Israel (will be referred to as IL) took part, especially Austria (AT), Belgium (BE), the Czech Republic (CZ), Denmark (DE), Estonia (EE), France (FR), Germany (DE), Italy (IT), Luxembourg (LU), the Netherlands (NL), Slovenia (SI), Spain (ES), Sweden (SE) and Switzerland (CH). When constructing the composite index, hedonic weights are used to build the index of deprivation on the 0-1 scale (Saisana et al., 2005). Higher values of the indices mean a higher level of deprivation; value zero indicates no problems included in the questioned items. The material deprivation index is constructed on the household level (its value is equal for all members of a household, usually for partners). The social deprivation index is based on both – household and individual level.

Severe deprivation is defined for respondents in upper quartiles (based on the whole sample) for both indices. These values for the material and the social deprivation are, respectively, 0.220 and 0.224.

1.1 Modelling of the probability distribution and logistic regression model

Both analysed indices are discrete variables with many states. It is why so many peaks are visible in kernel estimates in Figures 1 and 2. The lognormal distribution and a normal mixture density model were applied to obtain skewed and multimodal distributions.

The normal mixture density $f(y|\mathbf{x})$ with K components with probabilities $0 \le \pi_j \le 1,\ j = 1, 2, ..., K$, $\sum_{j=1}^{K} \pi_j = 1$ is given by

$$f(y|\mathbf{x}) = \sum_{j=1}^{K} \pi_j f_j(y), \tag{2}$$

where $f_j,\ j = 1, 2, ..., K$ are normal component densities specified by the component parameters (μ_j, σ_j^2). In the text, we apply models with two components (5 unknown parameters), three components (8 parameters), and four components with 11 parameters to be estimated. Estimating parameters of mixtures for the data with local peaks, we meet a problem of the identification of components with extremely small variances. To obtain a smooth estimated density (according to the assumptions), we apply a small number of components strategy. We propose our models such a suitable graphical presentation for clustering without any statistical inference. To select the number of components, only AIC criterion, graphical presentation, and information on numerical problems.

In the logistic regression model fitting probability of being severely deprived, we include explanatory variables for age (5 years groups, 8 categories, 50-54, 55-59, ..., 85+), gender (male, female), size of household (factor with three categories 1, 2, 3+ members) and a country of domicile (15 categories). All variables are coded as factors; the baseline category respondent is a woman from Austria, aged 50-54 and living alone (in a single-member household). We have a logistic regression model with 25 regression parameters. The model and baseline distribution correspond with a logistic model (Malá, 2019a) for the probability that *depmat* = 0 (aggregated probabilities for participating countries, Table 1 in the last column).

All calculations were performed in freeware R (R Core Team, 2020). To fit probability distributions, packages Fitdistrplus (Delignette-Muller, Dutang, 2015) and Mixtools (Benaglia et all., 2009) were used.

1.2 Description of data

The dataset includes 54,930 respondents with both indices observed. The mean age in the data is 69.8 years, with the sample standard deviation being 9.7 years. We obtain Spearman correlation coefficient between indices of 0.39; this value indicates expected but moderate positive dependence between indices reflecting social and material deprivation.

Median values are higher for the social deprivation index for all countries. Moreover, the material deprivation index medians are zero for Austria, Belgium, Denmark, Germany, Luxembourg, Netherlands, Sweden and Switzerland. We obtain the highest value for Estonia (0.289), subsequently for Israel, Slovenia, Spain, and Italy with 0.163. In the whole sample, we obtain 11.1% of respondents in severe deprivation situation based on the definition given in (Adena et al., 2015). The value of this percentage is highly country-specific in the interval 1.7%-27.2% (Table 1). In Table 1, country-specific values of upper quartiles are shown; lower values than the whole sample values are emphasised in red. In the last column, the percentage of respondents with material deprivation index being 0, the value for the whole data exceeds one half of respondents (51.8%).

Tab. 1: Characteristics for severe deprivation according to countries

Country	% of severe	upper quartile depmat	upper quartile depsoc	greater value upper quartiles	material index=0 %
Denmark	1.65%	0.000	0.110	social	76.94%
Sweden	2.95%	0.042	0.144	social	73.87%
Switzerland	3.13%	0.057	0.132	social	65.83%
Netherlands	3.39%	0.023	0.132	social	74.03%
Austria	4.32%	0.163	0.002	material	63.51%
Luxembourg	5.32%	0.058	0.208	social	67.03%
Belgium	6.81%	0.163	0.002	material	60.40%
Germany	8.28%	0.163	0.197	social	56.99%
Slovenia	10.00%	0.333	0.197	material	32.44%
France	12.03%	**0.220**	**0.285**	social	48.69%
Spain	13.03%	0.282	**0.231**	material	39.86%
Czech Republic	17.07%	**0.261**	0.375	social	40.40%

Israel	23.21%	0.327	0.375	social	39.04%
Italy	24.45%	0.386	0.375	material	35.45%
Estonia	27.16%	0.482	0.332	material	15.28%

Source: own computations

2 Results

In Table 2, the results of a logistic regression model for the probability of being severely deprived, assuming the same predictors as in (Malá, 2019b) is shown. All regressors are significant (ANOVA analysis); the intervals for odds ratios shifted above one are shown in blue and shifted below 1 in red. Controlling for other explanatory variables, we identify one country indistinguishable to Austria (the Netherlands) and a group of countries with a better situation (Denmark, Sweden, Switzerland). Other countries are worth with respect to severe deprivation than Austria with an odds ratio for Estonia 8.5 followed by Italy (7.5) and Israel (7.2) (Table 2).

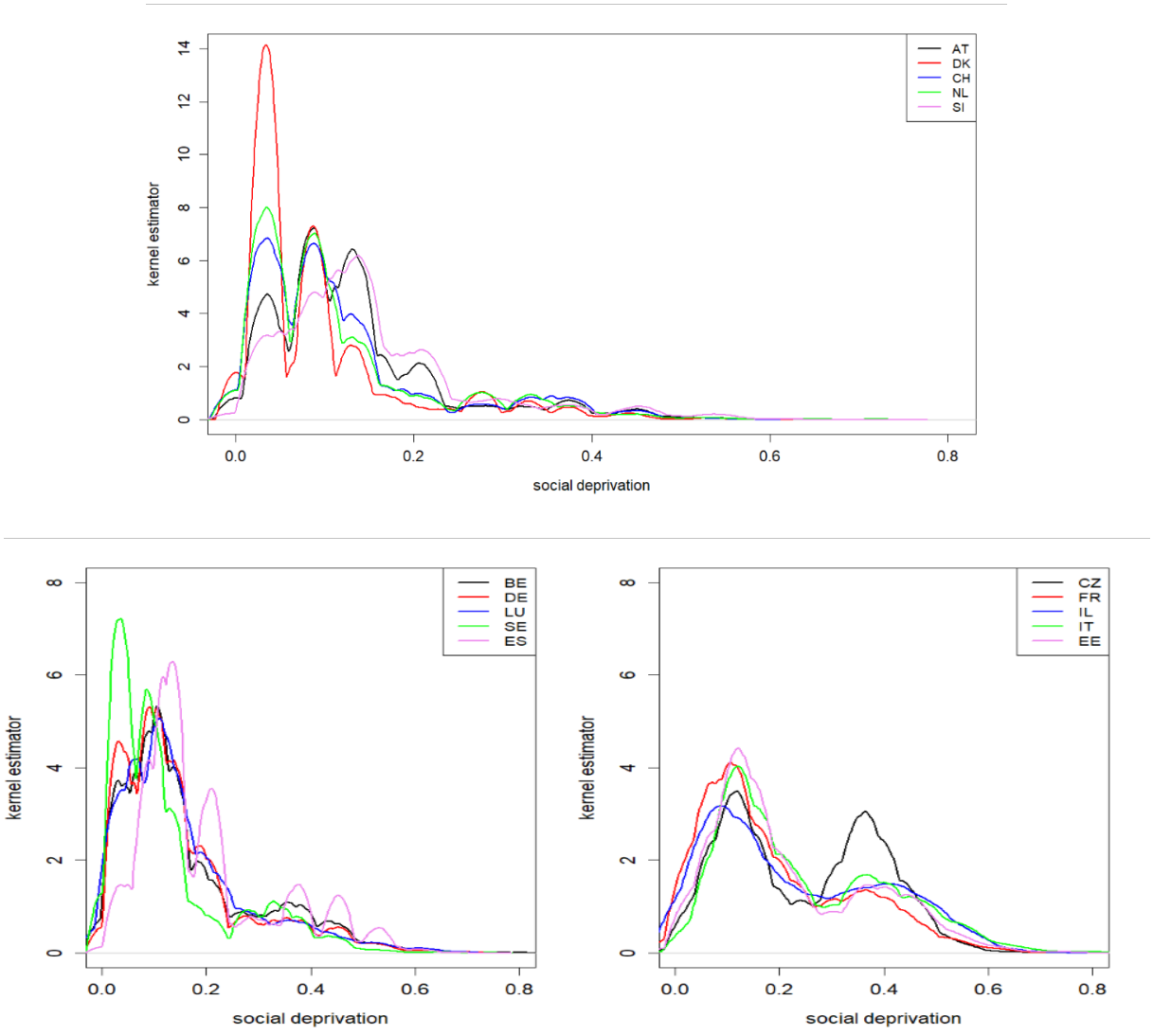
Tab. 2: Results of a logistic regression model for severe deprivation

Coefficients	estimated value	standard error	95% lower limit	95% upper limit	odds ratio	95% lower limit	95% upper limit
(Intercept)	-2.648	0.093	-2.831	-2.466	0.071	0.059	0.085
gender Male	-0.136	0.029	-0.193	-0.079	0.873	0.825	0.924
age 55-59	-0.065	0.057	-0.177	0.046	0.937	0.838	1.047
age 60-64	-0.127	0.057	-0.239	-0.015	0.881	0.788	0.985
age 65-69	-0.201	0.059	-0.317	-0.085	0.818	0.729	0.918
age 70-74	-0.074	0.060	-0.191	0.044	0.929	0.826	1.045
age 75-79	0.060	0.062	-0.062	0.182	1.062	0.940	1.199
age 80-84	0.155	0.067	0.010	0.300	1.168	1.010	1.350
age 85+	0.270	0.074	0.202	0.339	1.311	1.224	1.404
size hh 2	-0.618	0.035	-0.705	-0.532	0.539	0.494	0.587
size hh 3+	-0.274	0.044	-0.464	-0.084	0.761	0.629	0.920
country Belgium	0.507	0.097	0.333	0.682	1.661	1.395	1.978
country Czech Republic	1.579	0.089	1.278	1.881	4.852	3.588	6.561
country Denmark	-0.929	0.154	-1.094	-0.764	0.395	0.335	0.466
country Estonia	2.148	0.084	1.968	2.329	8.571	7.156	10.264
country France	1.118	0.092	0.934	1.302	3.058	2.544	3.677
country Germany	0.761	0.094	0.561	0.961	2.141	1.753	2.614
country Israel	1.980	0.102	1.809	2.151	7.243	6.107	8.589
country Italy	2.009	0.087	1.731	2.288	7.459	5.647	9.853
country Luxembourg	0.272	0.142	0.029	0.515	1.312	1.029	1.673

country Netherlands	-0.181	0.124	-0.381	0.019	0.834	0.683	1.019
country Slovenia	0.917	0.102	0.744	1.089	2.501	2.104	2.971
country Spain	1.236	0.088	0.995	1.477	3.443	2.705	4.381
country Sweden	-0.321	0.123	-0.586	-0.056	0.725	0.557	0.945
country Switzerland	-0.283	0.135	-0.283	-0.283	0.753	0.753	0.753

Source: own computations

Fig. 1: Kernel estimates of distributions of the social deprivation index (3 clusters)

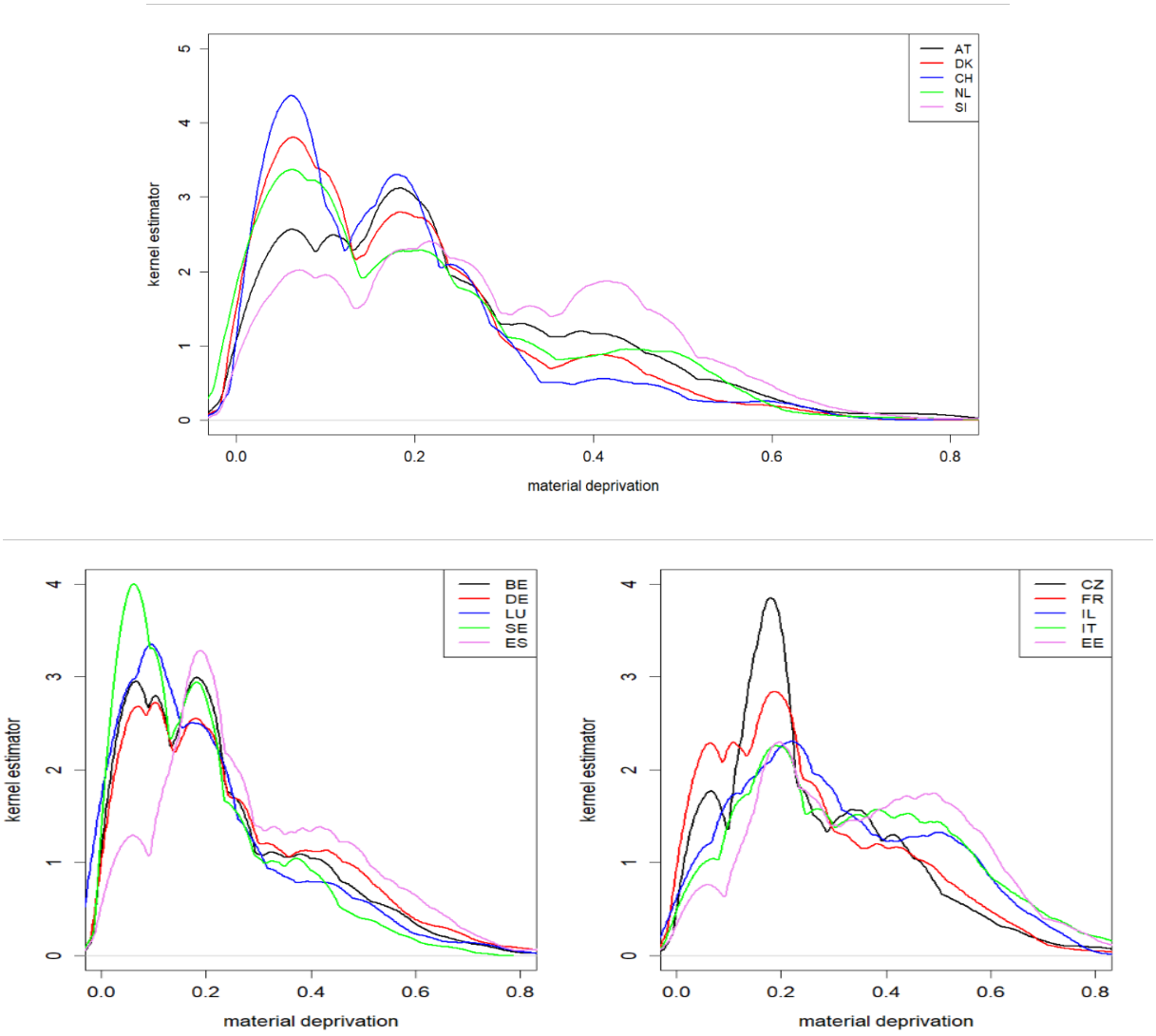


Source: own computations

In Figure 1, the distributions of the social index is plotted for all included countries; in Figure 2, histograms for the material deprivation index (positive values only) are shown. All countries are divided into three groups. The first group includes Austria, Denmark, Switzerland, and the Netherlands; Slovenia is between this and the second group. This group includes Belgium, Germany, Luxembourg, Sweden, and Spain; the last group contains the

Czech Republic, France, Italy, Israel, and Estonia. To model the distribution in the first group, it is necessary to smooth observed peaks in the data for values 0.03, 0.11, 1.13, and 0.21 (especially for Slovenia). One lognormal distribution is possible, but we obtain a better model fit with a normal mixture model with 2-4 components (for Austria, see Figure 3). Similar distributions can be used for the second group, but the distribution should cover the longer interval (we need a positive probability up to 0.8), and the decline is slower. One additional normal component is necessary to fit side modes at the interval 0.35 - 0.39 depending on the country (see Figure 3 for Spain). The third group needs at least two components to fit the distribution, as the distribution is bimodal (Estonia in Figure 3).

Fig. 2: Kernel estimates of distributions of the material deprivation index (3 clusters defined in Figure 1)

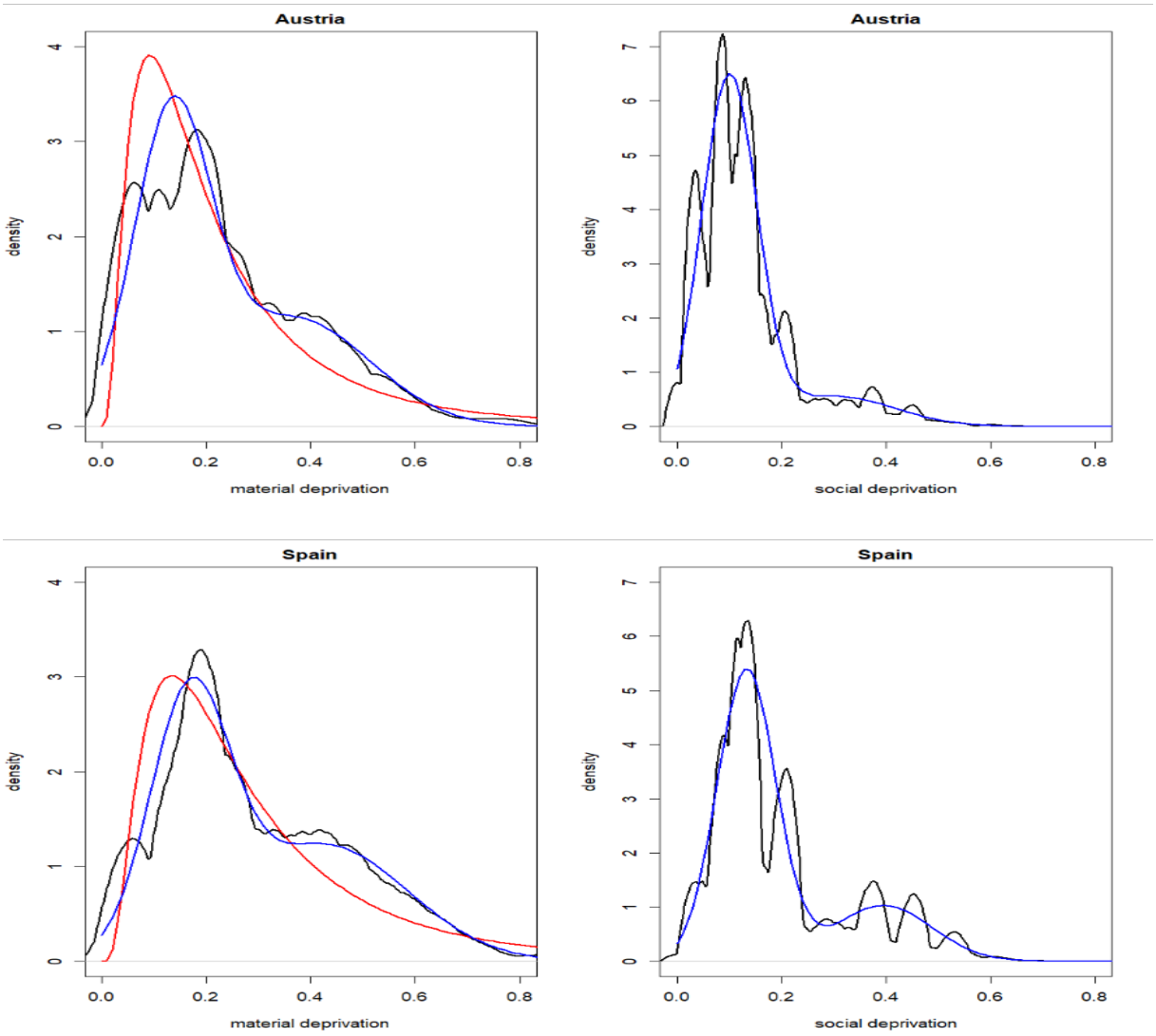


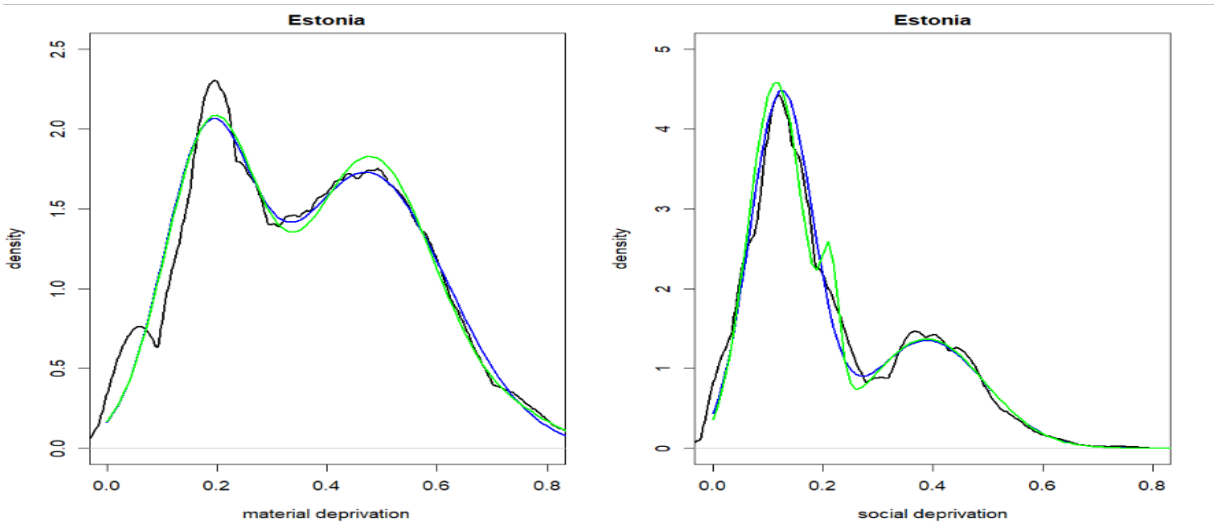
Source: own computations

Conclusion

Quantifying different subjective phenomena such as quality of life, health, deprivation, loneliness, and many more is an important and challenging issue. The results can help compare the situation of different subgroups of population, regions or states, or the quantitative information can help eliminate the problem, reduce deprivation and improve the

Fig. 3: Fitted models for Austria, Spain, and Estonia. Lognormal distribution in red, normal mixtures: two components in blue, three components green.





Source: own computations

quality of life. Well-functioning indices should take into account the specifics of an ageing population.

In the text, we analyse empirical distributions of two indices – index of material and index of social deprivation modified for inhabitants aged above 50 and provided by the 5th wave of SHARE survey. The developed welfare European countries consider not only material but also social exclusion and deprivation as a severe problem, for this reason we consider important to analyse separately material and social deprivation. The paired values of our sample also provide opportunity to deal with the relationship between indices and to construct a characteristic for deprivation based on both components – severe deprivation in this text.

From the data, we can derive that countries with excellent material deprivation situation might exhibit relatively high social deprivation index values and vice versa. The probability distributions of indices are not homogenous through participating countries. It follows that minimising deprivation and improving the situation of still growing population of people at later ages should be addressed by the European institutions emphasising the individual countries specifics. Furthermore, it is necessary to pay attention not only to material but also to social conditions.

The distributions of indices in society are very heterogenous and country-specific. However, it is possible to cluster these distributions into subgroups; in the text, three more homogenous subgroups are detected. The normal mixtures with two to four components are introduced as a suitable model for the distributions.

Acknowledgment

The support of the grant no. F4/45/2020 provided by Internal grant agency of the University of Economics, Prague is gladly acknowledged.

References

Adena, M., Myck, M. & Oczkowska, M. (2015). Innovation for better understanding deprivation index. In: Ageing in Europe - Supporting Policies for an Inclusive Society, De Gruyter.

Bellani, L. & D´Ambrosio, C. (2011). Deprivation, social exclusion and subjective well-being. *Social Indicators Research*, pp. 67-86.

Bellani, L. (2013). Multidimensional indices of deprivation: the introduction of reference groups weights, *Econ Inequal*, pp. 495–515.

Benaglia, T., Chauveau, D., Hunter, D. & Young, D. (2009). Mixtools: An R Package for Analysing Finite Mixture Models. *Journal of Statistical Software*, 32. Retrieved from: http://www.jstatsoft.org/v32/i06/.

Blane, D., G. Netuveli & M. Bartley (2007). Does quality of life at older ages vary with socio-economic position? *Sociology-the Journal of the British Sociological Association*, 41**,** pp. 717-726.

Börsch-Supan, A. (2019) Survey of Health, Ageing and Retirement in Europe (SHARE) Wave 5. Release version: 5.0.0. SHARE-ERIC. Data set.

Delignette-Muller M.L. & Dutang, C. (2015). Fitdistrplus: An R Package for Fitting Distributions. *Journal of Statistical Software*, 64, pp. 1-34.

Malá, I. (2019a) Modelling of the distribution of the material deprivation index based on SHARE data. In: International Days of Statistics and Economics (MSED2019) [online]. Praha, 05.09.2019 – 07.09.2019. Slaný: Libuše Macáková, Melandrium, 2019, s. 1044–1053. Retrieved from: https://msed.vse.cz/msed_2019/article/145-Mala-Ivana-paper.pdf.

Malá, I. (2019b) Modelling Deprivation of the 50-plus Population of the Czech Republic Based on the Share Survey. *Statistika*. 2019, roč. 99, č. 2, pp. 117–128. Retrieved from: https://www.czso.cz/documents/10180/88506448/32019719q2_117_mala_analyses.pdf.

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: https://www.R-project.org/.

Saisana, M., Saltelli, A. & Tarantola, S. (2005). Uncertainty and Sensitivity Analysis Techniques as Tools for the Quality Assessment of Composite Indicators. *Journal of the Royal Statistical Society. Series A (Statistics in Society),* vol. 168, pp. 307-323.

SHARE Release Guide 7.0.0. [online]. [26.4.2020] http://www.share-project.org/data-documentation/waves-overview/wave-7. html.

van Leeuwen, K. M., van Loon, M. S., van Nes, F. A., Bosmans, J. E., de Vet, H. C. W., Ket, J. C. F., Widdershoven, G. A. M. & Ostelo, R. W. J. G. (2019). What does quality of life mean to older adults? A thematic synthesis, PLoS ONE, vol. 14, no. 3, pp. 117-128.

Contact

Ivana Malá
Prague University of Economics and Business
W. Churchill pl. 4, Prague 3, Czech Republic
malai@vse.cz