The use of a three level M-quantile model to map poverty at LAU 1 in Poland

Stefano Marchetti
University of Pisa, Italy
E-mail: stefano.marchetti@unipi.it
Maciej Beręsewicz
Poznań University of Economics and Business, Poland
Statistical Office in Poznań, Poland
E-mail: maciej.beresewicz@ue.poznan.pl
Nicola Salvati
University of Pisa, Italy
E-mail: nicola.salvati@unipi.it
Marcin Szymkowiak
Poznań University of Economics and Business, Poland
Statistical Office in Poznań, Poland
E-mail: m.szymkowiak@ue.poznan.pl
Łukasz Wawrowski
Poznań University of Economics and Business, Poland
Statistical Office in Poznań, Poland
E-mail: lukasz.wawrowski@ue.poznan.pl

Abstract. A three level M-quantile model for small area estimation is proposed. The methodology represents an efficient alternative to prediction by using a three level linear mixed model in the presence of outliers and it is based on an extension of M-quantile regression. A modified method of the traditional M-quantile (two level) approach for poverty estimation is also proposed. In addition, an estimator of the mean-squared prediction error is described, which is based on a bootstrap procedure. The proposed methodology, as well as the three level empirical best predictor, are applied to Polish EU-SILC and census data to estimate poverty at LAU 1 level in Poland, i.e. the level for which the Central Statistical Office of Poland has not published any official estimates to date.

Keywords: small area estimation, Polish Census of Population and Housing 2011, European Survey on Income and Living Conditions, bootstrap method, poverty indicators

1. Introduction

The most common indicators used for measuring poverty, such as poverty rate (head count ratio – HCR) or poverty gap are taken from the Foster-Greer-Thorbecke family of indices (FGT; Foster et al. 1984). Other social cohesion indicators were proposed by the European Commission in the Europe 2020 Strategy (European Commission, 2010). Moreover, Eurostat calculates other inequality indicators, for instance quintile share ratio. All measures in this group are based on income, which can be defined in different ways. In this paper we use the household equivalised income, which is computed by using the modified OECD scale (Hagenaars et al., 1994). The modified OECD equivalence scale is the only one officially adopted by Eurostat for the definition of equivalised income used as the basis for poverty indicators such as HCR.

Poverty mapping is useful in economic planning and policy formulation to support policy makers. In European countries the basic source of information about the poverty indicators is the European Survey on Income and Living Conditions (EU-SILC), which provides accurate information on the relevant variables (household disposable income, household composition) and enables reliable estimation of average equivalised household income for large territorial units, such as the NUTS 1 or NUTS 2 regions. Unfortunately, in order to follow an appropriate social strategy, which is consistent with the guidelines of the cohesion policy, it is necessary to measure poverty and provide information about this phenomenon at lower levels of spatial aggregation. In this context, poverty maps are used to support decisions concerning important political issues, such as the allocation of development funds by governments, national ministries of infrastructure and development or international organizations, such as the World Bank.

Since the EU-SILC survey does not cover adequately all the specific areas or population subgroups, the required information could be obtained using small area estimation (SAE) techniques based on the idea of "borrowing strength". In Poland, for instance, EU-SILC data are only sufficient to publish estimates of poverty indicators at the level of the whole country and at the regional level (NUTS 1). However, given the growing demand for information about poverty indicators at lower levels than NUTS 1 in Poland, there is a pressing need to take advantage of appropriate SAE techniques and data from different statistical sources (EU-SILC, census or administrative registers).

The SAE methodology has been developed to produce reliable estimates of different characteristics of interest, such as means, counts, quantiles or ratios for domains for which only small samples are available (Rao and Molina, 2015). The SAE methodology is used by different National Statistical Institutes in different areas, in particular to estimate quantities related to the labour market, agriculture or business statistics. It is also useful for mapping poverty. For instance, the World Bank has used the SAE methodology to prepare poverty maps for more than 60 countries all over the world.

In 2013 the Center for Small Area Estimation, a specialized unit at the Statistical Office in Poznan, in cooperation with the Central Statistical Office of Poland and the World Bank prepared a poverty map of Poland at the level of 66 subregions (NUTS 3), using the Fay-Herriot approach (Statistical Office in Poznan, 2013). For the Central Statistical Office the next step is the application of the SAE methodology to estimate the poverty rate for *powiats* (LAU 1 districts) in Poland using data from the EU-SILC, the census and the Local Data Bank, which is the largest collection of information about the social, economic and demographic situation in Poland.

In the field of poverty mapping there are different approaches to choose from. A comprehensive guide to implementing SAE methods for poverty studies and poverty mapping, the mathematical description of methods and their applications, the pros and cons of specific SAE techniques in the context of poverty can be found in Pratesi (2016). Briefly, common SAE-based poverty mapping methods include (i) direct estimators, which, in general, are inefficient; (ii) Fay-Herriot models (Fay and Herriot, 1979), which enable aggregation, specific modelling, specification of sampling variances; (iii) Elber, Lanjouw, Lanjouw (ELL) method (Elbers et al., 2003), which is used by the World Bank and can perform poorly when auxiliary variables do not explain the entire between-area variation; (iv) the EB approach (Molina and Rao, 2010; Marhuenda et al., 2017) based on a nested-error model, which is very efficient under normality, (v) the Hierarchical Bayes (HB) approach (Molina et al., 2013) based on a nested-error model, which is similar to the EB approach but is less computationally demanding, (vi) M-quantile methods (Tzavidis et al., 2008), which are less sensitive to outliers. It is easy to see that an outlying value which destabilizes a population estimate based on a large survey sample will almost certainly destroy the validity of the corresponding direct estimate for the small area from which the outlier is sourced, since this estimate will be based on a much smaller sample size. This problem does not disappear when the small area estimator is an indirect estimator: large deviations from the expected response (outliers) are known to have a large influence on classical maximum likelihood inference based on generalized linear mixed models (GLMMs).

Indirect methods for poverty mapping are based on models that take into account between area variation beyond what is explained by auxiliary variables: the ELL method is based on a two-stage sampling design with clusters as first stage units, and small areas are aggregations of clusters. The EB and the HB approaches do not take into account within-cluster variability and assume that variability associated with the conditional distribution of y given x can be at least partially explained by a pre-specified hierarchical structure, such as the small areas of interest. The EB approach based on a three-level mixed model (Marhuenda et al., 2017) includes random effects explaining the heterogeneity at two levels of aggregation (cluster and small area variability). 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.

The main aim of the article is to present selected results of a study to map poverty at LAU 1 in Poland, i.e. at the level for which official estimates have not been published to date, by applying different small area estimators. In particular, we propose, first, an extension of the traditional M-quantile model, where we account for three levels of hierarchy in the data, we refer to it as a three-level M-quantile model. Then we define a Monte-Carlo technique to obtain small area estimates of parameters of interest. Secondly, we propose a modification of the Monte-Carlo technique used in Marchetti et al. (2012) to estimate parameters of interest, like the incidence of poverty, using the traditional M-quantile model, which we refer to as the two-level M-quantile approach. Both methods are resistant to outlying values, but the three-level M-quantile model accounts for between-small area and cluster variation beyond what is explained by covariates. Using model-based and design-based simulations – in the supplementary material – we compare estimators obtained under two- and three-level M-quantile models and the Empirical Best (EB) of Molina and Rao (2010) and its extension, which is based on a three-level mixed model proposed by Marhuenda et al. (2017).

The structure of the article is as follows. Section 2 is devoted to the theoretical background of parameters considered and estimated in our article at the level of LAU 1 in Poland. In Section 3 we describe data from the 2011 Polish Census of Population and Housing and EU-SILC 2011, which are used to estimate appropriate parameters of poverty using estimators based on three level M-quantile models and EB estimators based on three level mixed models. In Section 4 we summarize the M-quantile linear model and its application to small area estimation. Then, we introduce an extension of M-quantile model able to mimic a three level mixed model, the three level M-quantile model, and consider the small area estimation process of target parameters. In this section we also introduce a modification of the Monte-Carlo-based estimator proposed by Marchetti et al. (2012), which is based on a two-level M-quantile model. We also explain the notation used for the two-level M-quantile model. In Section 5 we present results of poverty mapping in Poland at LAU 1 level. Section 6 concludes the paper with some final remarks.

2. Notation and poverty indicators

Let U denote a finite population of size N divided into D domains or small areas of sizes N_d, $d = 1, \dots, D$, which are assumed to be known. In what follows we assume the availability of survey data on the outcome variable and explanatory variables which can be used to model it. In addition, the methods assume the availability of micro-level census/administrative data on the same set of explanatory variables. So we assume that a p-vector of auxiliary variables $\boldsymbol{x}_{jd}$ is known for each population unit j in small area d and that values of the outcome variable of interest y_{jd} are available from a random sample s, which includes units from all target domains. Usually y_{jd} is the equivalised income for household j from area d. We denote the sample size, the sampled part of the population and the non-sampled part of the population in each domain by n_d, s_d and r_d respectively, with $U_d = s_d \cup r_d$. We further assume that w_{jd} is the sampling weight of household j from area d and that, conditional on the covariates available, the sampling design is ignorable.

The first parameter we are interested in is the average equivalised income for each small area:

$$\bar{y}_d = N_d^{-1} \sum_{j \in U_d} y_{jd}. \tag{1}$$

A direct estimate of this parameter is given by the Horvitz-Thompson estimator:

$$\hat{\bar{y}}_d = \frac{\sum_{j \in s_d} w_{jd} y_{jd}}{\sum_{j \in s_d} w_{jd}}. \tag{2}$$

The FGT family of measures is based on income and requires fixed poverty threshold (line). The Eurostat recommends setting this threshold at 60% of the national median of equivalised disposable income (Central Statistical Office, 2012). This reference level is used to determine whether a given household is poor – if its equivalised income is lower than the poverty line, the household is considered to be poor. The general formula for the FGT indices is given by:

$$\mathrm{P}_{\alpha d} = \frac{1}{N_d} \sum_{j \in U_d} \Big(\frac{z - y_{jd}}{z}\Big)^{\alpha} I(y_{jd} < z), \tag{3}$$

where z denotes the poverty line, $I(y_{jd} < z)$ is an indicator function, which equals 1 if $y_{jd} < z$ and 0 otherwise. For $\alpha = 0$ equation (3) is the HCR. When $\alpha = 1$, then we obtain the poverty gap, which measures the area mean of the relative distance to the poverty threshold.

The FGT poverty indicator expressed in equation (3) can be estimated by the direct estimator in the following way:

$$\hat{\mathrm{P}}_{\alpha d} = \frac{1}{\hat{N}_d} \sum_{j \in s_d} w_{jd} \Big(\frac{z - y_{jd}}{z}\Big)^{\alpha} I(y_{jd} < z), \tag{4}$$

where $\hat{N}_d = \sum_{j \in s_d} w_{jd}$.

Usually small domains are not planned in the survey design, so we can assume simple random sampling (SRS) within domains. Hence, the variance of estimator (4) is given by:

$$\hat{V}(\hat{\mathrm{P}}_{\alpha d}) = \frac{1}{N_d^2} \sum_{j \in s_d} w_{jd}(w_{jd} - 1) P^2_{\alpha jd}, \tag{5}$$

where $P_{\alpha jd} = 1$ if household j is under the poverty line and $P_{\alpha jd} = 0$ otherwise. In the paper we focus on two measures of poverty: the average equivalised income and the HCR for which the direct estimator and variance are calculated using `sae` package available in `R` (Molina and Marhuenda, 2015). Direct estimation was conducted using SRS. It can be justified by the fact that these domains are unplanned in the sampling scheme of the Polish EU-SILC.

3. Data sources, model specification and diagnostics

In this section we describe the sources of data, i.e. the Polish Census of Population and Housing 2011 (hereinafter referred to as the Census) and EU-SILC 2011, used to estimate the average equivalised income and the HCR at LAU 1 level. We also present diagnostics from fitting a three level linear mixed model to these data. These diagnostics will then allow us to motivate the use of the alternative semi-parametric methodology based on M-quantile models.

3.1. Census data

The main objective of the census is to provide the most detailed information on the numbers in the population, its territorial distribution, socio-demographic and professional structures, and the socio-economic specificity of households and families, as well

Table 1. Distribution of the sample size at LAU 1 level for Census data

Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
3,099	5,290	6,233	6,985	7,734	72,990

as their resources and dwelling conditions at all levels of the country's territorial division: national, regional, and local (Central Statistical Office, 2011).

The 2011 Population and Housing Census was conducted by means of the so-called mixed method, i.e. some data were acquired from administrative sources (registers and information systems), while others were collected directly from the population either through full enumeration or a sampling survey. Additionally, two full enumeration surveys were conducted to collect data about persons residing in collective living quarters and the homeless. The mixed mode approach was mostly intended to reduce the census costs and the level of bias with respect to persons covered by the census, without affecting the high quality of the census results.

The full enumeration survey of population and housing was based on administrative registers supplemented by a brief questionnaire to be completed by each respondent. For the first time in Poland 28 administrative sources were used in order to obtain the values of the census variables, both at the stage of creating a specification of census units (population and housing census) and for qualitative comparisons. Thanks to the availability of one system of identifiers (PIN – Personal Identification Number) it was possible to merge data from different registers.

The sample survey conducted as part of the 2011 Census was carried out on a 20% sample of dwellings and an approximately 20% sample of the population. Design weights associated with units drawn to the sample had to be calibrated to known demographic totals from administrative registers (Central Statistical Office, 2011).

For the purpose of this study we treat the 20% census sample as a full census, assuming that the sub-population mimics the characteristics of the total population. The distribution of the sample size (households) at LAU 1 level is presented in Table 1.

Variables for the study were selected by taking into account two factors: the literature on poverty (Haughton and Khandker, 2009), particularly the existing research on poverty in Poland (Central Statistical Office, 2013), and the availability of data in the Polish Census. The following variables were selected: the fraction of males in the household (*males*), child dependency ratio in the household (*children*), the fraction of people aged 30–44 in the household (*people30_44*), the fraction of people aged 65 and over in the household (*people65*), the fraction of unemployed in the household (*unemployed*), the fraction of disabled people in the household (*disabled*), the fraction of people with basic/elementary education in the household (*educ_elementary*), the fraction of people with higher education in the household (*educ_high*), whether the household occupies a flat with only one room (dummy variable, *room1*), whether the household occupies a flat with more than three rooms (dummy variable, *room3*), whether the household lives in a village or a town with a population $<$ 20 k inhabitants (dummy variable, *village_city20*). Table 2 contains summary statistics for Census and EU-SILC auxiliary variables used in the model.

Table 2. Summary statistics for Census, EU-SILC and administrative registers data

Variable	Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
Census data						
children	0.00	0.00	0.00	0.22	0.33	11.00
disabled	0.00	0.00	0.00	0.16	0.20	1.00
educ_elementary	0.00	0.00	0.00	0.17	0.25	1.00
educ_high	0.00	0.00	0.00	0.13	0.14	1.00
males	0.00	0.33	0.50	0.47	0.67	1.00
people30_44	0.00	0.00	0.00	0.19	0.33	1.00
people65	0.00	0.00	0.00	0.20	0.25	1.00
room1	0.00	0.00	0.00	0.08	0.00	1.00
room3	0.00	0.00	1.00	0.63	1.00	1.00
unemployed	0.00	0.00	0.00	0.05	0.00	1.00
village_city20	0.00	0.00	0.00	0.46	1.00	1.00
EU-SILC data						
eq_inc (in k PLN)	-7.81	13.39	19.15	22.48	27.10	535.00
children	0.00	0.00	0.00	0.21	0.33	4.00
disabled	0.00	0.00	0.00	0.05	0.00	1.00
educ_elementary	0.00	0.00	0.00	0.17	0.20	1.00
educ_high	0.00	0.00	0.00	0.13	0.00	1.00
males	0.00	0.33	0.50	0.45	0.60	1.00
people30_44	0.00	0.00	0.00	0.15	0.33	1.00
people65	0.00	0.00	0.00	0.24	0.50	1.00
room1	0.00	0.00	0.00	0.10	0.00	1.00
room3	0.00	0.00	1.00	0.56	1.00	1.00
unemployed	0.00	0.00	0.00	0.05	0.00	1.00
village_city20	0.00	0.00	0.00	0.44	1.00	1.00
Domain level data (based on administrative registers)						
lau1_benefits	0.04	0.10	0.14	0.40	0.58	2.56
lau1_nace_a	0.00	0.03	0.21	0.24	0.41	0.79
lau1_unempl	0.04	0.10	0.13	0.14	0.18	0.37

3.2. EU-SILC data

The survey to collect European Union Statistics on Income and Living Conditions (EU-SILC) was launched in 2003. In Poland it was first administered by the Central Statistical Office in 2005. The main aim of the survey is to deliver comparable data about income, poverty and living conditions of households in EU Member States. EU-SILC data are collected using a questionnaire in face-to-face interviews covering demography, education, health, housing conditions, economic activity, and the level of household incomes and its sources. EU-SILC is a sample-based, representative longitudinal survey, in which the household is the basic statistical unit. In addition, every household member above 16 years old is also surveyed.

The EU-SILC survey has a two-stage sampling scheme with different selection probabilities at the first stage. Census enumeration areas are treated as primary sampling units (PSU), while dwellings are selected at the second stage. Primary sampling units are stratified before selection. In Poland, the strata correspond to the provinces (NUTS 2 units) and within each province primary sampling units are classified by class of locality. In urban areas census areas are categorized by size of town. Big cities constitute a separate stratum, but in the five largest cities districts are treated as strata. In rural areas strata consist of rural municipalities (LAU 2) of a given sub-region (NUTS 3) or a few neighbouring districts (LAU 1). Altogether, 211 strata are distinguished. A detailed description of the sampling scheme can be found in Central Statistical Office (2012).

In 2011 the EU-SILC sample size was equal to 12,871 households and 28,305 persons aged 16 and over participated in individual interviews. Sampled households consisted of 36,720 persons in total.

Like other sample surveys, EU-SILC suffers from non-sampling errors. The quality of about 84% of non-income data was good or very good, according to interviewers (Central Statistical Office, 2012, p. 60). The non-response rate in 2011 for households was equal to 14.9%, for individuals 7.0%, amounting to 20.9% in total. The fraction of missing data for the total gross income was equal to 9.04% and 7.85% for the disposable income. In the case of income data, the EU-SILC methodology requires the imputation of missing information in income data (EU-SILC survey has no replacement sample). Missing data are imputed using stochastic methods (hot-deck method, regression imputation with randomly selected empirical residuals) and deterministic methods (regression deterministic imputation, deduction imputation).

In addition to household variables, three domain level variables are used in the model. These are the registered unemployment rate (*lau1_unempl*), the fraction of people employed in companies classified into Section A (Agriculture, Forestry and Fishing, according to NACE rev. 2) (*lau1_nace_a*) and social benefits paid to natural persons in millions PLN per 1,000 employed persons (*lau1_benefits*). These variables are calculated by the Central Statistical Office based on register data.

Table 2 contains descriptive statistics for EU-SILC and domain auxiliary variables. An average household contained 0.21 children with a maximum of 4, which is lower than the average based on census data, 5% of households contained disable persons, 17% – persons with elementary education, and 13% – persons with higher education. The number of households with disabled persons is lower in comparison to 17% observed in census data. According to EU-SILC data, 10% of households have one room, compared to 8% in the

Table 3. Distribution of the sample size at LAU 1 level for EU-SILC data

Minimum	1st quartile	Median	Mean	3rd quartile	Maximum
1.00	17.00	28.00	34.32	43.00	418.00

census. The distribution of auxiliary variables was compared by Pearson's chi-squared test, which indicated no significant differences in data distributions between the census and EU-SILC for the tested variables. Table 2 indicates that there are (five) households with negative income. This result was due to considerable amounts of regular taxes on wealth, regular inter-household cash transfers and income tax and social insurance contributions declared by those households.

In EU-SILC 2011 only 4 out of 379 districts were not included in the sample. However, owing to the number of observations at LAU 1 level the sample size in districts is not sufficient to provide reliable estimates. Table 3 presents descriptive statistics for domain sample size observed in EU-SILC 2011. The minimum sample size in domains is 1, 25% of all domains contain 17 or fewer observations. The median sample size is 28, with a mean of 34.32, which implies an asymmetry in the distribution. The maximum sample size is 418 in Warsaw, the capital of Poland.

3.3. Model fitting and model diagnostics

We start by fitting a two- and three-level nested error model, each of which takes into account variability within domains with normally distributed random effects to the EU-SILC data. The models are fitted using `R` and `lmer` function from `lme4` package (Bates et al., 2015). In order to account for variability within small areas (LAU 1) we selected LAU 2 level (gminas) for a practical reason: the sampling scheme applied in the EU-SILC does not match the sampling scheme applied in the 20% sample drawn in the Polish Census. It is not possible to match PSUs between these two surveys. We decided to select the lowest level available in both data sources (i.e. LAU 2). For simplicity, we will refer to this level as the PSU level.

The three-level nested error model is characterised by the following three- level hierarchical structure: level 1 - individuals/households (denoted by j, $j = 1, \dots, n_{c_d}$) are nested within PSUs - level 2 (denoted by c, $c = 1, \dots, c_d$), and PSUs are nested within districts - level 3 (denoted by d, $d = 1, \dots, D$). Furthermore, $\sum_{d=1}^{D} c_d = C_s$ and $\sum_{d=1}^{D} \sum_{c=1}^{c_d} \sum_{j=1}^{n_{c_d}} = n$. The model is as follows:

$$y_{jcd} = \boldsymbol{x}_{jcd}^T \boldsymbol{\beta} + u_d + v_{cd} + \epsilon_{jcd} \,, \tag{6}$$

where y is the log shifted transformation of the equivalised income, to avoid the negative income problem, $\boldsymbol{x}_{jcd}$ represents the vector of covariates, $\boldsymbol{\beta}$ is the vector of regression parameters, u_d – level-three errors (districts, powiats), v_{cd} – level-two errors (PSUs), and ϵ_{jcd} – level-one (individual/household) errors. In model (6) u_d, v_{cd} and ϵ_{jcd} are assumed to be Normally distributed with zero mean and variance $\boldsymbol{\Sigma}_u = \sigma_u^2 \boldsymbol{I}_D$, $\boldsymbol{\Sigma}_v = \sigma_v^2 \boldsymbol{I}_{C_s}$, $\boldsymbol{\Sigma}_\epsilon = \sigma_\epsilon^2 \boldsymbol{I}_n$, respectively.

The estimated model parameters and the corresponding test statistics are shown in Table 4 and are in the expected direction. The highest absolute coefficients at unit level are observed for the unemployed (on average a 43% decrease in income), high education

Table 4. Log-linear model fitting results for EU-SILC data

Coefficient	Estimate	SE	Wald t	exp(B)
(Intercept)	10.3242	0.0156	662.4786	30461
males	0.0955	0.0120	7.9701	1.1002
children	-0.1231	0.0085	-14.4836	0.8841
people30_44	0.0604	0.0142	4.2448	1.0622
people65	-0.0431	0.0097	-4.4524	0.9578
unemployed	-0.5524	0.0222	-24.8896	0.5756
disabled	-0.2856	0.0189	-15.1472	0.7515
educ_elementary	-0.1162	0.0113	-10.2648	0.8903
educ_high	0.4695	0.0125	37.4686	1.5992
room1	-0.0983	0.0112	-8.7643	0.9064
room3	0.0857	0.0069	12.4075	1.0895
village_city20	-0.0887	0.0086	-10.2547	0.9151
lau1_unempl	-0.2097	0.0708	-2.9602	0.8108
lau1_nace_a	-0.2676	0.0255	-10.5142	0.7652
lau1_benefits	-0.0504	0.0107	-4.7270	0.9508
σ_u domain level	0.0351			
σ_v psu level	0.0369			
σ_ϵ unit level	0.3457			

(on average a nearly 60% increase in income) and disabled people (on average an almost 25% decrease in income). There is also a difference in the level of income between villages and towns with fewer than 20,000 inhabitants and larger towns. Parameters that represent the overall situation at LAU 1 level are negative as expected. As unemployment at LAU 1 level rises, incomes decrease. A similar pattern can be observed for the share of people employed in companies classified into Section A, which reflects the structure of the labor market in LAU 1 units. Finally, we observe differences in estimates of random effects variances. The variance of the PSU effect is slightly larger than the domain effects. To support the use of the three-level model we calculated conditional AIC (cAIC; Efron 2004) and corrected conditional AIC (ccAIC), as suggested by Greven and Kneib (2010), using `cAIC4` package. Efron's cAIC for the two-level model is equal to 9,315.39, while for the three-level model 9,186.42 and ccAIC equals 9,401.80 for the two-level and 9,394.54 for the three-level model. This means that the inclusion of the PSU effect helps to account for within-area variability and can improve the small area prediction of income and HCR.

Figure 1 shows Normal probability plots of level 1 (individuals, i.e. household), level 2 (PSU, in Polish gmina), and level 3 (district, in Polish powiat) residuals obtained by fitting three level mixed model to the data. Normal probability plots indicate that the Gaussian assumptions of the mixed model are not met (confirmed by Anderson-Darling Normality test, Anderson and Darling, 1952). The highest residuals are observed for level 1, with a maximum absolute value of 29.32, but only 4.3% of observations have absolute Pearson residuals greater than 2. Analysis of residuals indicates violations of the normality assumption for random effects. In addition to outlying observations, there are also influential values. Figure 2 presents values of Cook's distance, which is used to identify domains and PSUs that influence estimates. A comparison of distances for level

2 and level 3 indicates that more influential observations are observed within domains than in PSUs. Five domains with the highest value of Cook's distance are Warsaw (1465), Łódź (1061), powiat świecki (0414), Katowice (2469), Częstochowa (2464) and and the most influential PSUs are found within these domains. Results indicate that the linear mixed model might not be appropriate for estimating income and HCR at LAU 1 level in Poland. Model diagnostics to evaluate the relative importance of covariates and their collinearity are in the supplementary material.

Hence, even though we used a log transformation, our preliminary analysis indicates that the estimates of model parameters can be potentially driven by influential observations and therefore robust estimation could improve inference under the model. For these reasons we propose a three-level M-quantile model, which is a good compromise between efficiency under normality and robust properties under contamination.

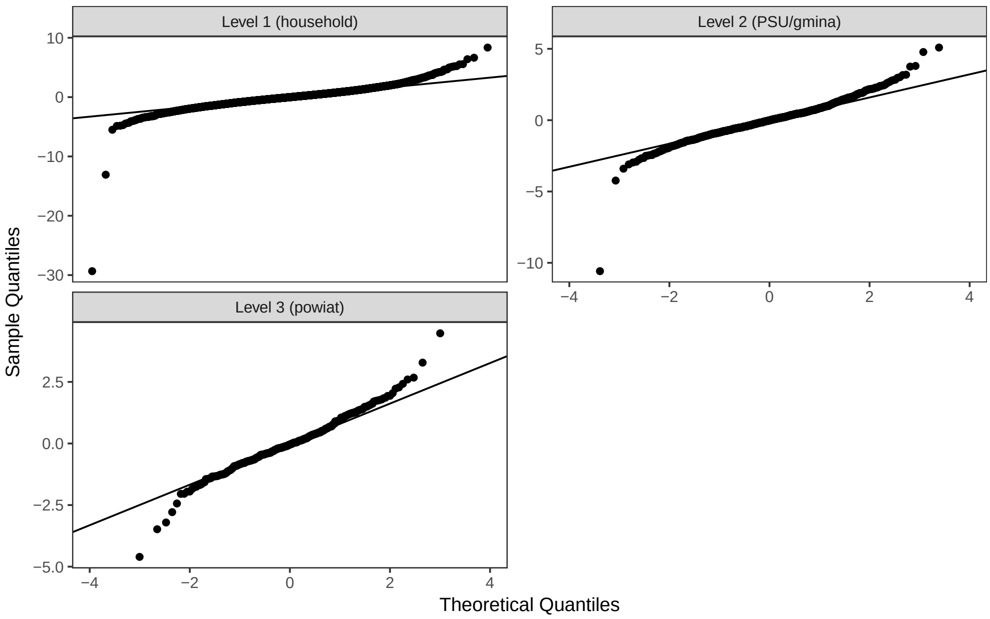

Figure 1. Normal probability plot of (a) estimated district random effects, (b) estimated PSU/gmina random effects and (c) of level 1 Pearson residuals based on a three level log-linear mixed model fitted to the EU-SILC data

4. M-quantile models

In this section we first summarize the M-quantile linear model (Breckling and Chambers, 1988) and its application to small area estimation (Chambers and Tzavidis, 2006). Then we present the two and three level extension of this model for poverty mapping.

The classical regression model summarizes the behavior of the mean of a random variable at each point in a set of covariates. Instead, M-quantile regression summarizes the behavior of different parts of conditional distribution $f(y|\boldsymbol{x})$ at each point in the set of the $\boldsymbol{x}$'s. Let us for the moment and for the sake of simplicity drop the area-specific

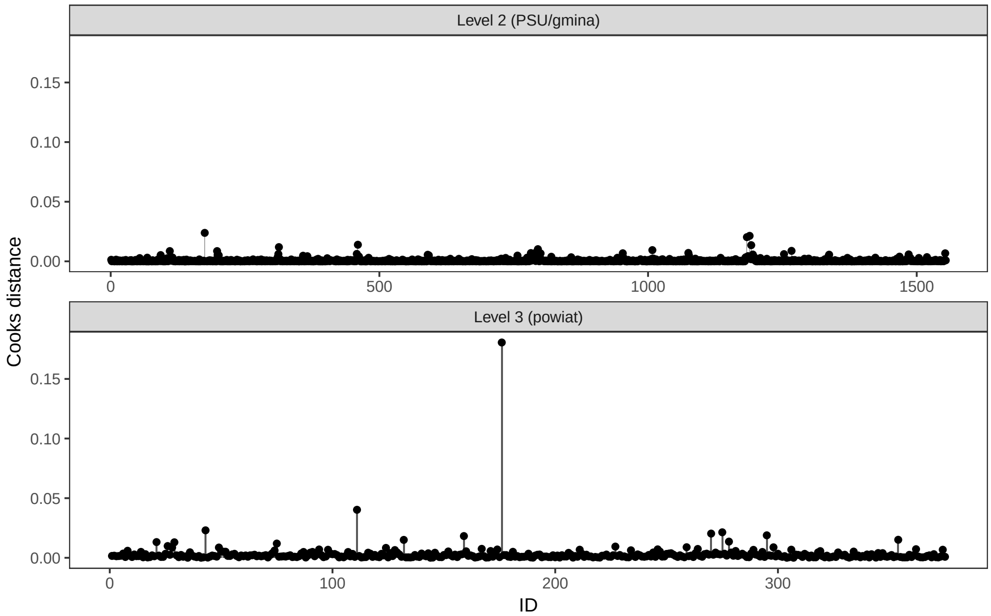

Figure 2. Model fit diagnostics for a three level log-linear mixed model fitted to the EU-SILC data: Cook's distances for PSU/municipality and district random effects

subscript d. Let $(\boldsymbol{x}_j, y_j)$, $j = 1, \ldots, n$ denote the values of a random sample consisting of n units, where $\boldsymbol{x}_j$ is the j-th row of a known design matrix $\boldsymbol{X}$ and y_j is a scalar response variable corresponding to a realization of a continuous random variable with an unknown continuous cumulative distribution function. The M-quantile regression is a "quantile-like" generalization of regression based on influence functions (M-regression). The M-quantile q of the conditional density $f(y|\boldsymbol{x})$, denoted by m, is defined as the solution to the estimating equation:

$$\int \psi_q(y-m) f(y|\boldsymbol{x})\, dy = 0\,,$$

where ψ_q is an asymmetric influence function, which is the derivative of an asymmetric loss function, ρ_q. When a linear relation between M-quantile m and auxiliary variables holds, then the M-quantile regression model is:

$$m_y(q|\boldsymbol{X}) = \boldsymbol{X}\boldsymbol{\beta}_\psi(q)\,.$$

An estimate of $\boldsymbol{\beta}_\psi(q)$ is obtained by solving the set of estimating equations

$$\sum_{j=1}^{n} \psi_q(r_{jq}) = 0\,, \tag{7}$$

with respect to $\boldsymbol{\beta}_\psi(q)$, where $r_{jq} = \{y_j - \boldsymbol{x}_j^T \boldsymbol{\beta}_\psi(q)\}\sigma^{-1}$, σ is a scale parameter, $\psi_q\{r_{jq}\} = 2\psi\{r_{jq}\}[qI(r_{jq} > 0) + (1-q)I(r_{jq} \le 0)]$ and $I(\cdot)$ is the indicator function. The ψ subscript indicates a robust influence function. A different set of regression parameters can be defined for each value of q. In particular, by varying the specifications of the asymmetric influence function ψ, we obtain the expectile, M-quantile and quantile regression models as special cases. In particular, in this paper we use the tilted version of the popular Huber influence function that can be written as:

$$\psi\{r_{jq}\} = \begin{cases} c\ sign(r_{jq}) & |r_{jq}| > c, \\ r_{jq} & |r_{jq}| \leq c\,, \end{cases} \tag{8}$$

where c is a cutoff constant. M-quantile regression models allow us to trade robustness for efficiency by properly tuning the c constant: robustness increases as c decreases, while efficiency increases as c increases. Provided that the tuning constant c is strictly greater than zero, estimates of $\boldsymbol{\beta}_{\psi}(q)$ are obtained using iterative weighted least squares (IWLS).

Chambers and Tzavidis (2006) extended the use of M-quantile regression models to small area estimation. They characterize the conditional variability across the population of interest by the M-quantile coefficients of the population units. For unit j with values y_j and $\boldsymbol{x}_j$, this coefficient is the value θ_j such that $m_y(\theta_j|\boldsymbol{x}_j) = y_j$. The M-quantile coefficients are determined at the population level. Consequently, if a hierarchical structure explains part of the variability in the population data, then we expect units within areas (or domains) defined by this hierarchy to have similar M-quantile coefficients. Using M-quantile coefficients it is possible to define an M-quantile small area model:

$$y_{jd} = \boldsymbol{x}_{jd}^T \boldsymbol{\beta}_{\psi}(\theta_d) + \epsilon_{jd}\,, \tag{9}$$

where $\boldsymbol{\beta}_{\psi}(\theta_d)$ is the unknown vector of M-quantile regression parameters for the unknown area-specific M-quantile coefficient θ_d and ϵ_{jd} is the unit level random error term for which no explicit parametric assumptions are being made. The area-specific M-quantile coefficient θ_d is estimated averaging the M-quantile coefficients of the sample units in area d, so $\hat{\theta}_d = n_d^{-1} \sum_{j=1}^{n_d} \theta_j$. Then, $\boldsymbol{\beta}_{\psi}(\hat{\theta}_d)$ is estimated solving equation (7). Therefore, the predicted values under the M-quantile small area model are $\hat{y}_{jd} = \boldsymbol{x}_{jd}^T \hat{\boldsymbol{\beta}}_{\psi}(\hat{\theta}_d)$. Empirical work indeed indicates that the area-specific M-quantile coefficients are positively and highly correlated with the estimated random area-specific effects obtained with the nested error regression small area model (Chambers and Tzavidis, 2006).

Using this approach to small area estimation it is then possible to obtain estimates for averages, quantiles, poverty indicators, inequality indicators, etc. by taking advantage of the few parametric assumptions needed under this small area method (Chambers and Tzavidis, 2006; Tzavidis et al., 2010; Marchetti et al., 2012; Tzavidis and Marchetti, 2016).

4.1. Extension of M-quantile models for three level nested errors

In this section we present an extension of the M-quantile small are model that mimics a three level nested error model. By this extension it is possible to take into account different source of variability related to different hierarchies in the data using the M-quantile approach.

As stated in Section 3.3, let a population be divided in D areas and in C PSUs/clusters (hereafter clusters), where clusters are partitions of an area - i.e. no intersections between clusters exist. Consider C_d clusters in area d, $d = 1, \dots, D$ and in each cluster N_{c_d} units. A possible specification of a three level M-quantile model is as follows:

$$y_{jcd} = \boldsymbol{x}_{jcd}^T \boldsymbol{\beta}_\psi(\theta_d) + \boldsymbol{z}_{jcd}^T \boldsymbol{\gamma}_\psi(\phi_{cd}) + \epsilon_{jcd}, \tag{10}$$

where y_{jcd} is the target (continuous) variable for unit j in cluster c in area d, $\boldsymbol{x}_{jcd}$ and $\boldsymbol{z}_{jcd}$ are respectively a p-vector and a q-vector of auxiliary variables known for all the units in the population. The $\boldsymbol{z}$ vector can contain auxiliary variables at cluster level. The θ_d parameter is the unknown area M-quantile coefficient, $\boldsymbol{\beta}_\psi(\theta_d)$ is the unknown p-vector of area-level regression parameters, ϕ_{cd} is the unknown cluster M-quantile coefficient, $\boldsymbol{\gamma}_\psi(\phi_{cd})$ is the unknown q-vector of cluster-level regression parameters and ϵ_{jcd} is a unit error term.

Bearing in mind the M-quantile approach to small area estimation summarized before, the quantity θ_d, $\boldsymbol{\beta}_\psi(\theta_d)$, ϕ_{cd} and $\boldsymbol{\gamma}_\psi(\phi_{cd})$ in equation (10) are unknown. They can be estimated as follows:

1. Starting from the following M-quantile linear model:
$$y_{jcd} = \boldsymbol{x}_{jcd}^T \boldsymbol{\beta}_\psi(\theta_d) + u_{jcd}\,, \tag{11}$$
estimate θ_d and then $\boldsymbol{\beta}_\psi(\theta_d)$ according to the M-quantile approach to small area estimation presented earlier. Let estimates be $\hat{\theta}_d$ and $\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d)$.
2. Then, compute residuals $\hat{u}_{jcd} = y_{jcd} - \boldsymbol{x}_{jcd}^T \hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d)$.
3. Next, using residuals $\hat{u}_{jcd}$ as the target variable in the following M-quantile linear model:
$$\hat{u}_{jcd} = \boldsymbol{z}_{jcd}^T \boldsymbol{\gamma}_\psi(\phi_{cd}) + \epsilon_{jcd}\,, \tag{12}$$
estimate ϕ_{cd} and then $\boldsymbol{\gamma}_\psi(\phi_{cd})$ using again the M-quantile approach to small area estimation. Indeed, $\hat{u}_{jcd}$ is the target variable, ϕ_{cd} is the domain (e.g. cluster) M-quantile coefficient and $\boldsymbol{\gamma}_\psi(\phi_{cd})$ is the vector of regression parameters. Let these estimates be $\hat{\phi}_{cd}$ and $\hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd})$. The residual of the M-quantile three level model is $\hat{\epsilon}_{jcd} = y_{jcd} - \boldsymbol{x}_{jcd}^T \hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) - \boldsymbol{z}_{jcd}^T \hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd})$.

If there are no auxiliary variables at cluster level, it is possible to use the same set of auxiliary variables in step 1 and 3. Therefore, when $\boldsymbol{z}_{jcd} = \boldsymbol{x}_{jcd}$ for all units, then model (10) mimics a three level nested error linear mixed model as expressed in (6).

It is possible to mimic a three level nested-error model using a common set of auxiliary variables since the traditional (two level) M-quantile small area model is built in two separate steps, as summarized in Section 4. In the first step unit level M-quantile coefficients are computed for each unit in the sample, in the second step these coefficients are "aggregated" (averaged) at a given hierarchical level (e.g. area) to get an M-quantile "area" coefficient. Then, model parameters are estimated and residuals are computed. The M-quantile model in its part $\boldsymbol{x}_{jcd}^T \hat{\boldsymbol{\beta}}(\hat{\theta}_d)$ captures the variability explained by the auxiliary variable together with between-area variability, that is the variability explained by the hierarchical structure in the data (e.g. area level hierarchy). Therefore, the residuals of this model, $\hat{u}_{jcd}$s, include residual variability due to unobserved variables, unobservable factors and other sources of variability related to different hierarchies in the data.

This last source of variability can be captured in step 3 of the proposed procedure. Using the same set of auxiliary variables, $\boldsymbol{x}_{jcd}$s, on the residuals $\hat{u}_{jcd}$s, obtained in step 2, it is possible to capture residual variability explained by auxiliary variables together with between-cluster variability (cluster level hierarchy) by the "aggregation" mechanism in the M-quantile small area model, i.e. $\boldsymbol{x}_{jcd}^T\hat{\gamma}(\hat{\phi})_{cd}$.

The described behavior is heuristically confirmed by empirical evaluation (not reported here) that showed a highly linear correlation between the estimated areas' M-quantile coefficients $\hat{\theta}_d$ and the estimated level-three errors of (6) $\hat{u}_d$, and between the estimated clusters' M-quantile coefficients $\hat{\phi}_{cd}$ and the estimated level-two errors of (6) $\hat{v}_{cd}$.

4.2. Estimation of parameters of a finite population using the three level M-quantile model

In the last years Monte-Carlo based estimators have been used widely to estimate income distribution and poverty indicators for small areas, see for example Elbers et al. (2003); Molina and Rao (2010); Marchetti et al. (2012); Giusti et al. (2012) and Marhuenda et al. (2017).

Using the Monte-Carlo approach it is possible to estimate the parameters that are a function of the (continuous) target variable y, say $h(y)$. Examples of such parameters can be small area means, quantiles, poverty and inequality indicators. To get the desired estimates it is assumed we have data from a random sample drawn from the target population, where target variable y and a set of auxiliary variables $\boldsymbol{x}$ are observed. Moreover, suppose we know the auxiliary variables for all the units of the population, for example from the Census or population registers. In many developed countries surveys on consumption and income are carried out yearly, while censuses are carried out every 10 years (or more). However, population registers are getting more and more rich of information and can represent a valid source of micro-data (person- or household-level data) available for all units of the population.

In our case study we treat LAU2 units (in Polish gmina) as primary sampling units and households (residing in a given gmina) as secondary sampling units. The sampling design ensures reliable estimates at the national and regional level. However, an optimal allocation of resources may require reliable estimates at a higher level of disaggregation, so recourse to small area estimation techniques can be a valid alternative. In our case the goal is to get estimates at district level (LAU 1) (in Polish powiat), where districts are clusters of gminas. A three-level M-quantile model can take into account between-powiat and between-gmina variability. Given that many national statistical agencies usually adopt a two-stage sampling design or a similar scheme, our method can prove useful in many real-life applications.

In the Polish 2011 EU-SILC the sample included units from $D = 375$ districts (Powiats, small areas) out of the total of 379 (4 out-of-sample areas). Let c_d be the number of sampled gminas (clusters) out of C_d clusters in area d, $d = 1, \ldots, D$. Cluster c_d has sample size n_{c_d}. Estimates of the model parameters are available only for sampled areas and clusters. In order to micro-simulate the population using the auxiliary variables and estimates of the model parameters we chose a non-parametric approach.

Let the parameter of interest in area d be a function h of observations and be equal to

$$h_d(y) = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{jcd}) + \sum_{j\in r_d} h(y_{jcd})\Big\}, \tag{13}$$

where s_d is a set of sampled units and clusters in area d (i.e. $c = 1,\dots,c_d$ and $j = 1,\dots,n_{c_d}$) and r_d is a set of non-sampled units in area d (i.e. $c \in C_d - c_d$ and $j \in N_{c_d} - n_{c_d}$), N_{c_d} is the number of units in cluster c of area d and $N_d = \sum_{c=1}^{C_d} N_{c_d}$. Moreover, let s be a set of sampled units and r a set of non-sampled units. The quantity $\sum_{j\in r_d} h(y_{jcd})$ is unknown and has to be estimated.

Taking the conditional expectation of (13), the predictor takes the form:

$$\hat{h}_d(y) = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{jcd}) + \sum_{j\in r_d} \hat{h}(\hat{y}_{jcd})\Big\},$$

where $\hat{h}(\hat{y}_{jcd})$ is the predictor of $h(y_{jcd})$ given by $E_{\boldsymbol{y}_r}[h(y_{jcd})|\boldsymbol{y}_s]$, the expected values under the conditional (predictive) distribution of y_{jcd} given the data vector $\boldsymbol{y}_s$.

To obtain $\hat{h}_d(y)$ we propose a Monte Carlo approximation as follows:

1. Estimate the model (10) using sample data.
2. Generate a synthetic population of target values $\boldsymbol{y}_d$, predicting non-sampled units according to model (10):

   $$\hat{y}^{syn}_{jcd} = \boldsymbol{x}^T_{jcd}\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) + \boldsymbol{z}^T_{jcd}\hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd}),$$

   $c = 1,\dots,C_d - c_d$ and $j = 1,\dots,N_{c_d} - n_{c_d}$. To predict values of $\hat{y}^{syn}_{jcd}$ for the $\sum_{d=1}^{D}(C_d - c_d)$ out-of-sample clusters we use synthetic estimation, therefore $\hat{y}^{syn}_{kc_{\text{out}}d} = \boldsymbol{x}^T_{kc_{\text{out}}d}\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) + \boldsymbol{z}^T_{kc_{\text{out}}d}\hat{\boldsymbol{\gamma}}_\psi(0.5)$.
3. Generate the k Monte-Carlo values for non-sampled units by adding a disturbance to the synthetic values for each source of variability, $\hat{y}^k_{jcd} = y^{syn}_{jcd} + u^*_d + v^*_{cd} + \epsilon^*_{jcd}$. The value ϵ^*_{jcd} can be obtained by sampling with replication from residuals of model (10), $\hat{\epsilon}_{jcd}$, conditional or unconditional on areas; u^*_d can be obtained by sampling with replication from the pseudo-area effects $\boldsymbol{x}^T_{jcd}[\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) - \hat{\boldsymbol{\beta}}_\psi(0.5)]$; v^*_{cd} can be obtained by sampling with replication from the pseudo-cluster effects $\boldsymbol{z}^T_{jcd}[\hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd}) - \hat{\boldsymbol{\gamma}}_\psi(0.5)]$.
4. Compute $h(\boldsymbol{y}_d)$ on the kth Monte Carlo population:

   $$\hat{h}^k_d(y) = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{jcd}) + \sum_{j\in r_d} h(\hat{y}^k_{jcd})\Big\}.$$

5. Repeat steps 3 and 4 K times and then estimate $h_d(y)$ by averaging over the K Monte Carlo populations

   $$\hat{h}_d(y) = K^{-1}\sum_{k=1}^{K}\hat{h}^k_d(y).$$

The disturbances added in step 3 are justified since $\hat{y}^{syn}_{jcd}$ is the expected value under the model (10) of unknown quantity y_{jcd} ($j \in r_d$) that has an unknown distribution. By adding a pseudo-area and a pseudo-cluster error as well as a unit-level error we mimic non-parametrically the unknown distribution of y_{jcd} $(j \in r_d)$; it is an approach similar in spirit to that used by Molina and Rao (2010) and Marhuenda et al. (2017).

Sometimes, it may be better to model a transformed target variable. For example, if W is the income, a log transformation like $y = \log W$ may produce a better model fit. Although the M-quantile approach is resistant to outliers and violations of the normality assumption, such a transformation can improve the model fit and small area estimates. When a transformation $y = g(W)$ is suitable for the target variable W, in step 4 compute $h(g^{-1}(y))$ instead of $h(y)$.

4.3. A Bootstrap scheme proposal to estimate the MSE

In this section we adapt the bootstrap technique of Marchetti et al. (2012) to the proposed three-level M-quantile small area model to estimate the MSE of $\hat{h}_d(y)$.

Starting from model (10) we estimate parameters $\boldsymbol{\beta}_\psi$, θ_d, $\boldsymbol{\gamma}_\psi$, ϕ_{cd} using sample data. We then compute model residuals $\hat{\epsilon}_{jcd} = y_{jcd} - \boldsymbol{x}_{jcd}^T \hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) - \boldsymbol{z}_{jcd}^T \hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd})$, pseudo-area effects $\hat{u}_d = \boldsymbol{x}_{jcd}^T[\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) - \hat{\boldsymbol{\beta}}_\psi(0.5)]$ and pseudo-cluster effects $\hat{v}_{cd} = \boldsymbol{z}_{jcd}^T[\hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd}) - \hat{\boldsymbol{\gamma}}_\psi(0.5)]$.

Given an estimator of the distribution of model residuals $\hat{G}_\epsilon$, pseudo-area effects $\hat{G}_u$ and pseudo-cluster effects $\hat{G}_v$, a bootstrap population, $\Omega^* = \{y^*_{jcd}, \boldsymbol{x}_{jcd}, \boldsymbol{z}_{jcd}\}$, can be generated by sampling from $\hat{G}_\epsilon$, $\hat{G}_u$ and $\hat{G}_v$ to obtain ϵ^*_{jcd}, u^*_{jcd} and v^*_{jcd} respectively,

$$\begin{aligned} y^*_{jcd} &= \boldsymbol{x}_{jcd}^T \hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) + \boldsymbol{z}_{jcd}^T \hat{\boldsymbol{\gamma}}_\psi(\hat{\phi}_{cd}) + u^*_d + v^*_{cd} + \epsilon^*_{jcd} \\ & d = 1, \ldots, D, c = 1, \ldots, C_d, j = 1 \ldots, N_{c_d}. \end{aligned} \tag{14}$$

For details and discussion of the estimation of the distribution of model residuals, see Marchetti et al. (2012); the extension to obtain estimates of pseudo-area and cluster effects distributions is straightforward.

Let $h^*_d(y)$ be the bootstrap population parameter of interest. This parameter can be estimated using the Monte Carlo approach presented in Section 4.2, by selecting a sample without replacement from the bootstrap population Ω^*. Therefore, we obtain the bootstrap estimate $\hat{h}^*_d(y)$ of the bootstrap population small area parameter $h^*_d(y)$.

The bootstrap procedure is conducted in the following steps: starting from sample s, we fit model (10) and obtain estimates of $\boldsymbol{\beta}_\psi$, θ_d, $\boldsymbol{\gamma}_\psi$, ϕ_{cd} which are used to generate B bootstrap populations, Ω^{*b}, using (14). From each bootstrap population Ω^{*b} we select L samples using simple random sampling in such a way that the number of clusters and units sampled corresponds to that of the original sample. Using the bootstrap samples we obtain estimates of the target parameter $h_d(y)$. Bootstrap estimators of the bias and variance of the estimated target parameter, $\hat{h}_d(y)$ are defined as follows:

$$\widehat{Bias}(\hat{h}_d(y)) = B^{-1}L^{-1} \sum_{b=1}^{B} \sum_{l=1}^{L} \big(\hat{h}^{*bl}_d(y) - h^{*b}_d(y)\big),$$

$$\widehat{Var}(\hat{h}_d(y)) = B^{-1}L^{-1} \sum_{b=1}^{B} \sum_{l=1}^{L} \big(\hat{h}^{*bl}_d(y) - \bar{\hat{h}}^{*bl}_d(y)\big)^2,$$

where $h_d^{*b}(y)$ is the small area parameter of interest of the bth bootstrap population, $\hat{h}_d^{*bl}(y)$ is the small area parameter estimated by using the Monte Carlo approach described in Section 4.2 with the lth sample of the bth bootstrap population and $\bar{\hat{h}}_d^{*bl}(y) = L^{-1}\sum_{l=1}^{L}\hat{h}_d^{*bl}(y)$. The bootstrap MSE estimator of $\hat{h}_d(y)$ is defined as:

$$mse(\hat{h}_d(y)) = \widehat{Var}(\hat{h}_d(y)) + \widehat{Bias}(\hat{h}_d(y))^2.$$

Empirical evaluation of the proposed bootstrap estimator is presented in the supplementary material.

A function that fits the three level M-quantile regression model and a function that produces small area estimates and their MSE have been written in `R` and they are available from the authors upon request.

4.4. Estimation of parameters of a finite population using the two level M-quantile model

The method to estimate poverty indicators at small area level using the two level M-quantile model – i.e. the traditional M-quantile approach to small area estimation – is discussed in Marchetti et al. (2012). The authors propose, among others, a Monte-Carlo based estimator for FGT poverty indexes and other statistics for which we propose a slight modification. The notation for the two-level M-quantile model follows directly from the three-level model notation.

Let the parameter of interest in area d be equal to:

$$h_d(y) = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{jd}) + \sum_{j\in r_d} h(y_{jd})\Big\}. \tag{15}$$

The quantity $h_d(y)$ can be estimated using the Monte-Carlo approximation:

1. Estimate the model (9) using sample data.
2. Generate a synthetic population of target values $\boldsymbol{y}_d$, predicting non-sampled units
$$\hat{y}_{jd}^{syn} = \boldsymbol{x}_{jd}^T\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d), \quad j \in N_d - n_d.$$
3. Generate the k Monte-Carlo values for non-sampled units by adding a disturbance to the synthetic values for each source of variability, $\hat{y}_{jd}^k = y_{jd}^{syn} + u_d^* + \epsilon_{jd}^*$. The value ϵ_{jd}^* can be obtained by sampling with replication from residuals of model (9), $\hat{e}_{jd}$, conditional or unconditional on areas; u_d^* can be obtained by sampling with replication from pseudo-area effects $\boldsymbol{x}_{jd}^T[\hat{\boldsymbol{\beta}}_\psi(\hat{\theta}_d) - \hat{\boldsymbol{\beta}}_\psi(0.5)]$.
4. Compute $h(\boldsymbol{y}_d)$ on the kth Monte Carlo population:
$$\hat{h}_d^k(y) = \frac{1}{N_d}\Big\{\sum_{j\in s_d} h(y_{jd}) + \sum_{j\in r_d} h(\hat{y}_{jd}^k)\Big\}.$$
5. Repeat steps 3 and 4 K times and then estimate $h_d(y)$ by averaging over K Monte Carlo populations:
$$\hat{h}_d(y) = K^{-1}\sum_{k=1}^{K}\hat{h}_d^k(y).$$

The modification consists in adding pseudo-area effects to the synthetic population in step 3. The motivation is the same as that discussed in Section 4.2. The bootstrap technique used in Marchetti et al. (2012) to estimate the MSE is adapted accordingly to the bootstrap estimator in Section 4.3.

5. Results

The sample size in the EU-SILC survey is only sufficient for the publication of social cohesion indicators at NUTS 1 level in Poland. Estimation at lower levels of (spatial) aggregation may result in unreliable estimates due to small sample size. In this section we focus on the estimation of the average equivalised income and the HCR at LAU 1 level by fitting the proposed three level M-quantile model.

Estimates at LAU 1 are crucial from a point of view of local authorities. In 2011 there were 379 LAU 1 units in Poland and information at this level of aggregation will be a big advance in comparison to data published at present. In 2011 EU-SILC only 4 districts were not included in the sample and 12 districts showed no variability in the sense that all individuals were below the poverty line. It is impossible to obtain the Horvitz-Thompson estimates for out-of-sample units and estimates for domains without variability because direct estimator is not reliable.

Figure 3 presents the coefficients of variation (CV) of the direct estimator of the HCR and the average equivalised income by sample size at area level. Variance of the direct estimator was calculated assuming simple random sampling within areas according to equation (5).

The sample size in small areas varies from 1 to 418 households. For the smallest areas the coefficient of variation is close to 100%. A fourth of the CVs of direct estimates of the HCR are lower than or equal to 43%. Half of the CVs are lower than or equal to 52% and another fourth are greater than or equal to 67%. Direct estimates of the average equivalised income are characterized by a better precision. The lower quartile of CV values is equal to 23%, the median is equal to 28% and the upper quartile is 35%. However, the precision of direct estimates at district (powiat) level is not acceptable for publication by the Central Statistical Office, hence we resort to model-based unit-level small area methods to improve the precision of the estimates (Central Statistical Office, 2011). On the basis of the analysis carried out in Section 3, instead of direct estimation, one can use a predictor based on the three-level M-quantile approach proposed in this work. Although previous analysis revealed the presence of outliers and departures from the normality assumption of linear mixed models, even after using a log transformation, we have also decided to use the empirical best estimator based on three-level mixed models proposed by Marhuenda et al. (2017).

5.1. Small area estimates of HCR and average equivalised income at LAU 1 level

Estimates of the HCR and the average equivalised income at LAU 1 level are presented in two maps. Figure 4 presents the spatial distribution of three-level M-quantile estimates of HCR and Figure 5 presents the equivalised income. The four districts marked in white are domains that are not present in the EU-SILC 2011 sample.

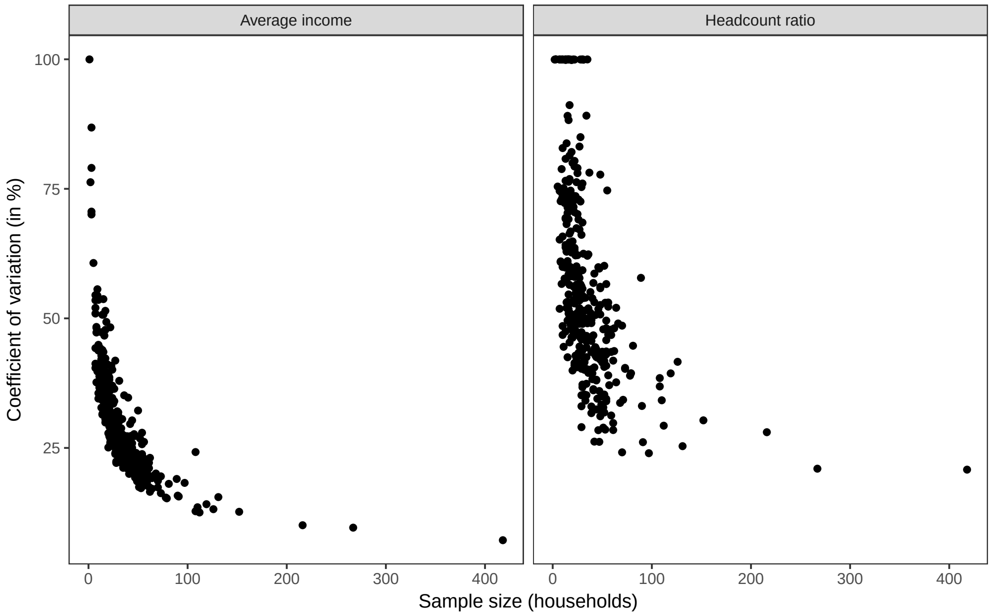

Figure 3. The relationship between sample size (the number of households) and the coefficient of variation of the direct estimates of average equivalised income (left) and HCR (right) within domains

Figure 4 presents a map of districts classified into six groups based on the k-means method, which is implemented in function `classIntervals` in the R package `classInt`, Bivand (2015). The first group consists of 48 districts with the lowest poverty rate, from 7.4% to 13.1%. They include the Polish capital (Warsaw) and provincial capital cities (i.e. Poznan, Wroclaw, Krakow, Gdansk and Szczecin), which are the most economically prosperous districts in Poland. The economy of these cities is dominated by commerce, financial services, educational services and the real estate market. These districts also enjoy a relatively small unemployment rate (Poznan has the lowest unemployment rate in the country). For these reasons HCR values are lower in these districts. A similar situation can be observed in districts surrounding the biggest cities. This pattern can be explained by the fact that large cities function as poles of growth and socio-economic development for the surrounding areas. For peripherally located districts, as the distance from capital cities increases, so does the poverty rate. Other districts with low poverty are characterized by well-developed industry in their area or in neighboring areas (copper, coal and brown coal basin, aviation and chemical industry). This is most evident in the Upper Silesian Agglomeration, which is the most important industrial region of Poland, characterized by the highest level of urban development in the country.

The last two groups of districts, characterized by the highest HCR (over 25%) consist of 84 districts (22% of all districts in Poland). More than a fourth of households in these districts live under the poverty line. Moreover, a strong spatial clustering is visible in the east of Poland. In particular, the poorest districts are situated in the south-eastern part along the border with Ukraine. This is the consequence of the rural character of these districts: much of their economy used to be based on post-socialist state-owned farms and industrial plants that have been closed down. As a result, most districts in this region are affected by higher unemployment and higher poverty rate. Another group of districts with high HCR are located in the west-central part along the border with Germany. These districts are characterized by poorly developed industry and are mainly rural and agricultural areas.

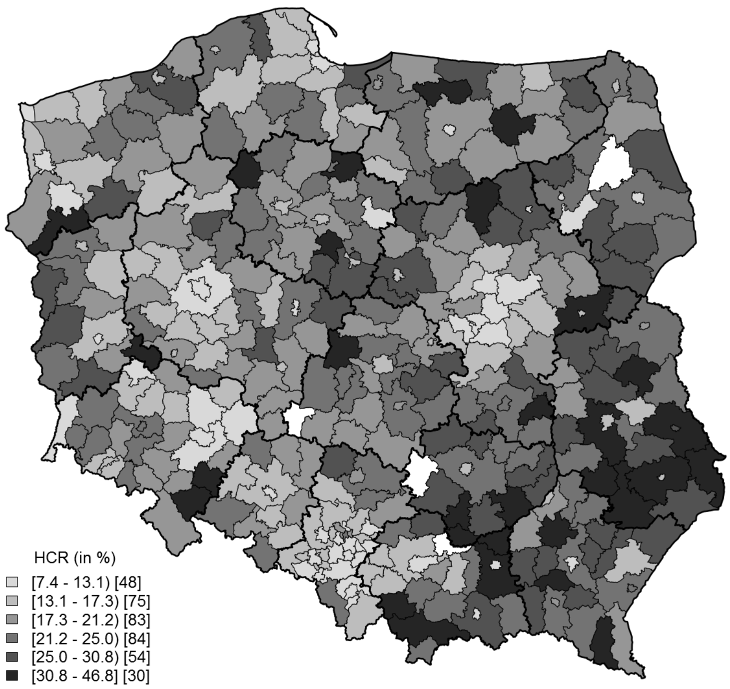

Figure 4. A map of three-level M-quantile based estimates of the HCR in districts of Poland

Based on Figure 5 it is also possible to identify poor and prosperous districts in terms of the average income. The highest average income is observed in big cities (Warsaw, Poznan, Krakow, Gdansk or Wroclaw) and surrounding districts (for instance poznanski powiat). The strong influence of a large city on the economic development of neighboring powiats is also observed. As in the case of HCR, a very good situation in terms of income is visible in powiats that are part of the Upper Silesian Agglomeration. In total, the average equivalised income for 30 powiats, estimated using a three-level M-quantile model was the highest and exceeded the threshold of 25,495 PLN. In terms of spatial distribution, powiats with the lowest level of average equivalised income according to the adopted model, are concentrated in the eastern and south-eastern part of Poland. There are two exceptions: the city of Rzeszow, which is the capital of the province of Podkarpackie, mielecki powiat (home to aviation industry), and the city of Tarnobrzeg along with the entire powiat tarnobrzeski, where a special economic zone is situated. The

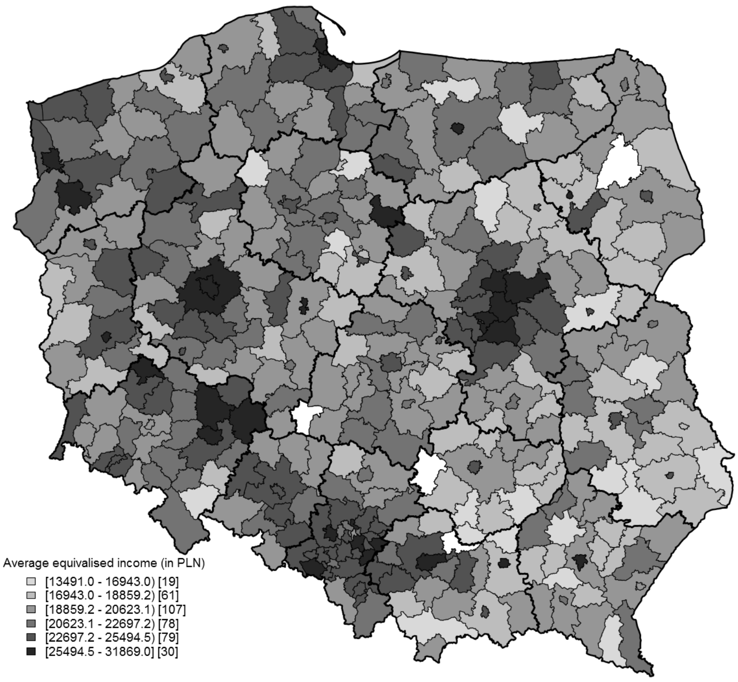

Figure 5. A map of three-level M-quantile based estimates of the average equivalised income in districts of Poland

Pearson coefficient of correlation between the estimated HCR and the average equivalised income was negative and equal -0.95. This means that, in general, powiats with a higher estimated average equivalised income, the poverty rate was lower. One exception is mysliborski powiat situated in the West Pomeranian Province along the German border, where the poverty rate was high – 35% and the average equivalised income was also relatively high.

5.2. *Small area diagnostics*

Although the above maps are an effective tool for summarizing small area estimates, they do not answer one fundamental question: how good are the estimates produced by using the proposed three-level M-quantile and three-level empirical best estimators? To answer this question, we note that the model-based estimates should be: (a) "close" to direct estimates with a reasonable sample size and (b) more precise than direct estimates.

Following Brown et al. (2001), we assess condition (a) by computing a goodness-of-fit diagnostic. This is based on the idea that if the model-based estimates are close to the small area value of interest, then unbiased direct estimates can be considered as random variables, whose expected values are equal to the values of the corresponding model-based estimates. The goodness-of-fit diagnostic is computed by means of the Wald

Table 5. Results of Wald test for HCR and average equivalised income (AEI) by small area sample size

Statistics	all data	$n > 20$	$n > 30$	$n > 40$	$n > 50$
No. of districts	375	249	153	104	63
Wald statistics, M-quantile					
HCR	821.02	501.43	294.72	143.10	82.34*
AEI	85.98*	63.86*	40.57*	30.94*	22.15*
Wald statistics, Empirical best					
HCR	958.70	479.89	229.77	139.42	84.47*
AEI	83.22*	52.43*	28.77*	22.83*	16.47*
$\chi^2_{0.95}$	421.15	286.81	182.86	128.80	82.53
* indicates not significant differences					

statistic for different sample sizes and the results for HCR and average equivalised income (AEI) are presented in Table 5. M-quantile estimates are close to direct ones: results indicate that the distribution of income within domains is not significantly different from that obtained from direct estimation. However, in the case of HCR, the similarity of distributions between M-quantile and direct estimates is observed only for districts with a sample size greater than 50. The same applies to empirical best estimates. The main reason for this is high variability of direct estimates of HCR. These results are confirmed by Figures 8 and 9.

Figure 6 presents the relationship between direct and M-quantile and Empirical best estimates of the number of poor households and the average equivalised income. The black line in the plots is used as reference. Pearson's coefficient of correlation for the number of poor households is equal to 0.63 for M-quantile estimates and 0.57 for empirical best estimates. The level of correlation is due to powiats that had a low poverty rate according to direct estimates (or even 0%). In general, poverty rates obtained using both the M-quantile and the empirical best approaches are higher than those resulting from direct estimates. The mean of direct estimates is 20.2% and 20.6% for M-quantile estimates and 23.1% for empirical best estimates. The median of the HCR was 16.6% for direct estimates and 20.4% for M-quantile estimates and 23.1% for empirical best estimates. These results indicate that on average about one fifth of Polish households in 2012 were poor.

For the average income the correlation is higher than for the number of poor households and is equal to 0.75, both for the M-quantile and Empirical best estimates. Figure 6 shows that small area estimates are more or less equally spread around the identity line for both the M-quantile and Empirical best estimators. Indeed, 51.7% of the direct estimates are lower than the M-quantile estimates, while for the Empirical best this proportion is 49.3%. Finally, the average difference in the measures of central tendency is about 250 PLN for the M-quantile estimates and about 290 PLN for the Empirical best estimates.

To assess condition (b), i.e. the potential gains in precision from using model-based estimates instead of the direct estimates, we examine the distribution of the estimated CVs of direct and model-based estimates for the EU-SILC data. In Table 6 we report the number of districts with CV values below 16.6%, between 16.6% and 33.3% and over

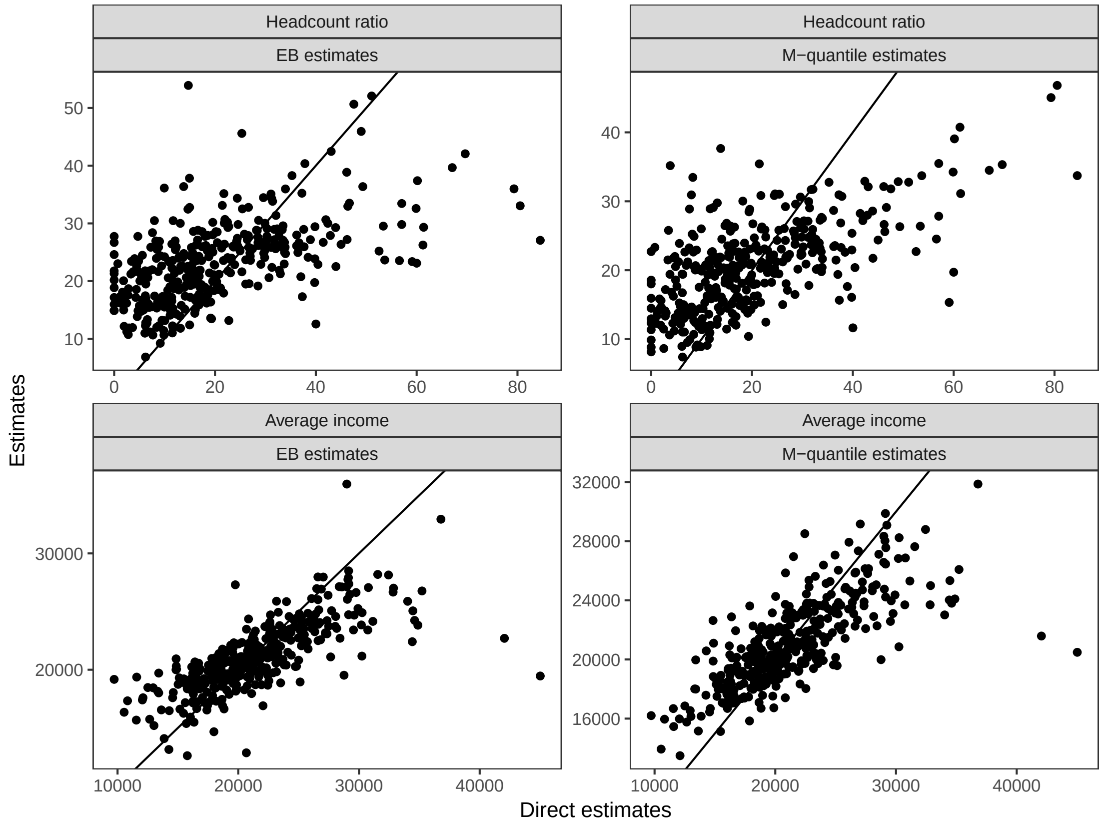

Figure 6. M-quantile and Empirical best estimates vs direct estimates of the number of poor households and the average equivalised income. Note that scales are not equal between panels

33.3% for direct, M-quantile and Empirical best estimators. These values are suggested by Statistics Canada (2007) as quality level guidelines for publishing tables: estimates with CVs below 16.6% are considered reliable for general use. Estimates with CVs between 16.6% and 33.3% should be accompanied by a warning to users. Estimates with CVs larger than 33.3% are deemed to be unreliable.

Precision is also evaluated by ratios of estimated CVs of direct and model-based estimates. A value of this ratio greater than 1 indicates that the estimated CV of the model-based estimate is smaller than that of the direct estimate. Figure 7 shows the relationship between these ratios, for both M-quantile and empirical best estimates, and the sample size in the EU-SILC sample in each district.

Figure 7 shows no districts for which direct estimates of HCR are, on average, more precise than M-quantile estimates. The precision of empirical best estimates is worse than that of direct estimates in 8 out of 375 districts for the HCR and in 7 for the average equivalised income.

With respect to HCR, the gain in precision within the lowest CV interval can be considered good. Direct estimation does not produce any results with CVs below 16.6%, while CVs of M-quantile estimates for 150 districts are below this threshold. The gain

Table 6. The number of districts by classes of CVs for the direct estimator, the M-Quantile and Empirical best predictors for HCR and average income

Estimator	below 16.6%	16.6%–33.3%	33.3% and more	N/A
Head count ratio				
Direct	0	27	337	11
M-quantile	150	209	16	0
Empirical best	85	269	21	0
Average equivalised income				
Direct	16	242	117	0
M-quantile	366	8	1	0
Empirical best	355	18	2	0

in precision between direct and M-quantile estimates can also be observed in the second CV interval – the number of districts for which estimates are at this level of precision is up from 27 to 209, which means that the number of districts with the lowest precision is greatly reduced. CVs of M-quantile estimates for only 16 districts are greater than 33.3%. The gain in precision is evident also for empirical best estimates, although it is worse than for M-quantile estimates – 85 districts with CVs below 16.6%, 269 districts with CVs in the second interval and 21 with CVs greater than 33.3%. From the perspective of quality, it is reasonable to publish HCR estimates for 27 districts using direct estimation, for 359 districts using M-quantile estimation and for 354 districts using empirical best estimation.

M-quantile and empirical best estimation are even more efficient for the average income. Estimates obtained by the M-quantile estimator for 366 out of 375 (97.6%) districts are included in the lowest CV interval (below 16.6%). Similarly, empirical best estimates for 355 districts are included in the first CV interval. In contrast, direct estimates are mostly found in the second and third CV interval.

The MSE of M-quantile estimates has been estimated using the bootstrap procedure described in section 4.3 with $B = 4$ (bootstrap populations) and $L = 100$ (bootstrap samples). The small number of bootstrap population and the limited number of bootstrap samples are due to computational reasons, since the sample and population dimension are very large. However, simulations in supplementary material showed that reasonable approximation of the true MSE can be obtained with even $B = 1$. The MSE of the EB has been estimated according to the bootstrap techniques described in the supplementary material of Marhuenda et al. (2017) setting the number of samples and populations equal to 100, since, also in this case, the use of larger samples is computationally unfeasible.

Figures 8 and 9 present a comparison of direct and M-quantile estimates of the poverty rate and the average equivalised income. The figures show districts arranged into four groups according to sample size: (below 15 – 73 districts, (15,20] – 99 districts, (25,50] – 140, over 50 – 63). The purpose of these plots is to verify differences between point estimates and estimation precision as sample size increases. The M-quantile and empirical best estimates tend to be more stable in comparison with the direct estimates for both variables. This is particularly evident in the groups up to-15 and $(15, 25]$ (fewer than 25 households in the sample). This is to be expected, especially for districts that are really small.

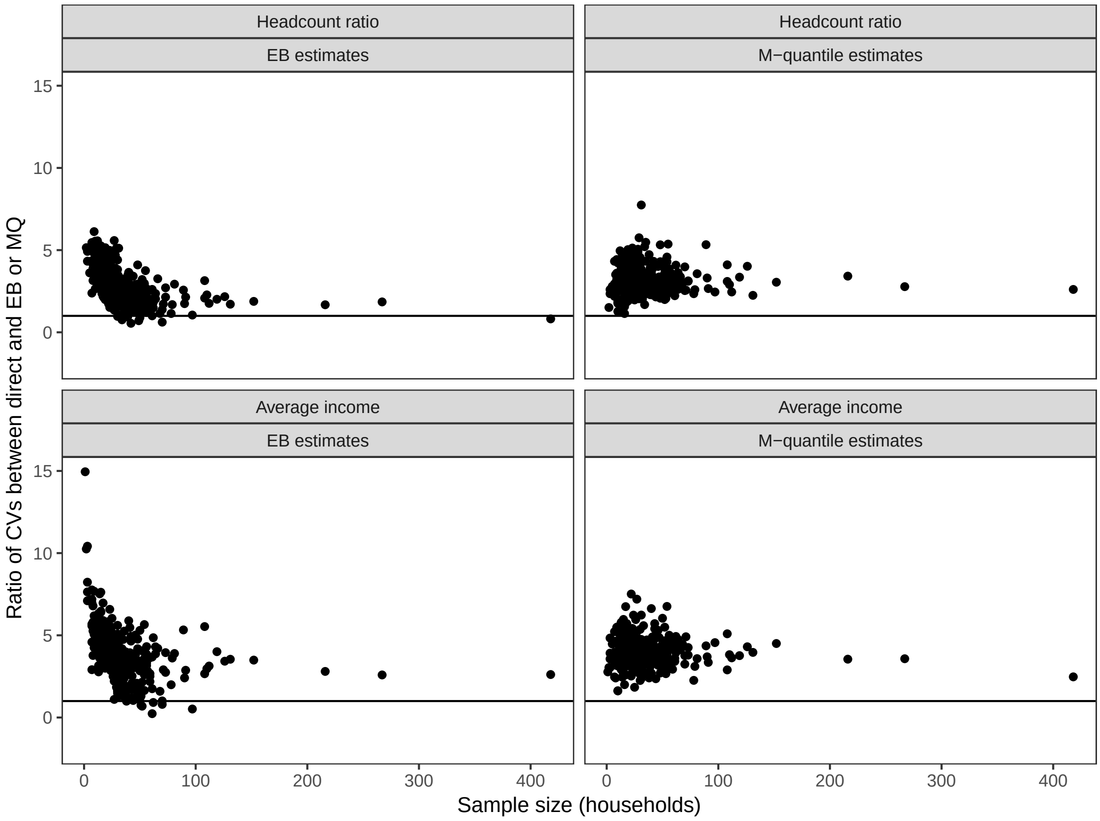

Figure 7. The ratio of estimated CVs of direct estimates and M-quantile and Empirical Best estimates of the HCR and average equivalised income for each district

Looking at the precision of HCR, (the four bottom plots in Figure 8), we can see that both M-quantile and empirical best estimates are better than direct estimates, as expected. Moreover, for districts with sample size below 15, the empirical best estimator performs a little bit better – in terms of CV – than the M-quantile estimator, while in all other sample size classes the M-quantile estimator tends to perform better than the empirical best estimator, which is evident for sample size over 50. Similar conclusions can be drawn for the average equivalised income. The highest CV for the M-quantile estimator of HCR (over 200%) is observed for a domain with only one sampled household.

As can be seen in Table 6, it is not possible to obtain reliable direct estimates for any districts. However, with model-based estimation it is possible to obtain estimates of CVs for all domains.

The validity of model-based inference depends on the validity of the model. The preceding analyses of EU-SILC data are sample specific, which makes generalization difficult. Moreover, in the literature there are alternatives unit-level model-based estimation approaches to those used in this application. Therefore, we have empirically evaluated the properties of widely used small area predictors: the traditional (i.e. two-level) M-quantile small area estimator, the proposed three-level M-quantile estimator

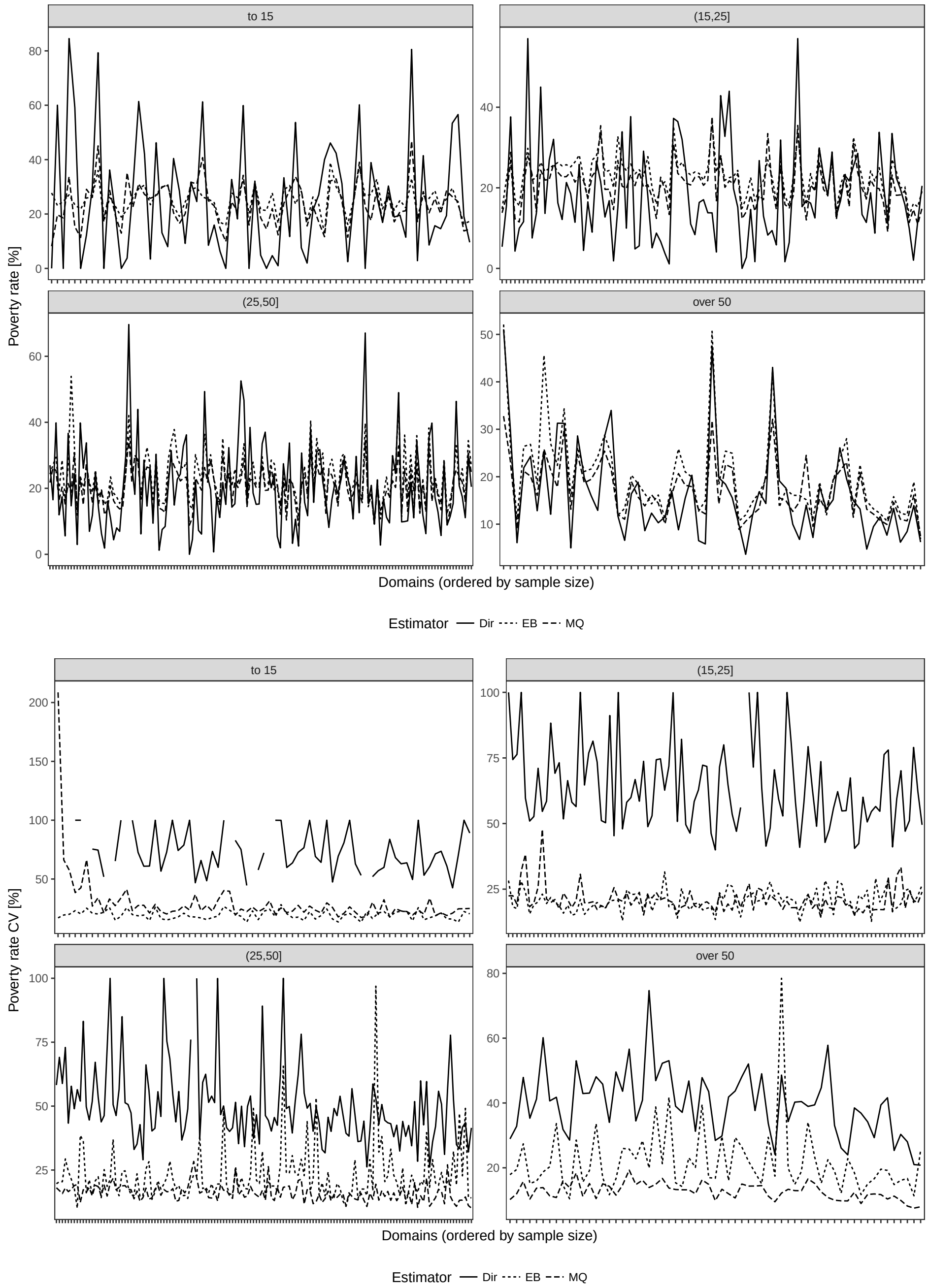

Figure 8. A comparison of direct, M-quantile and empirical best estimates of the poverty rate and their corresponding CV values grouped by sample size. Note that the plots and panels do not have the same Y scale

and the empirical best estimator based on two- and three-level mixed models. In the supplementary material we evaluate the performance of the four small area predictors

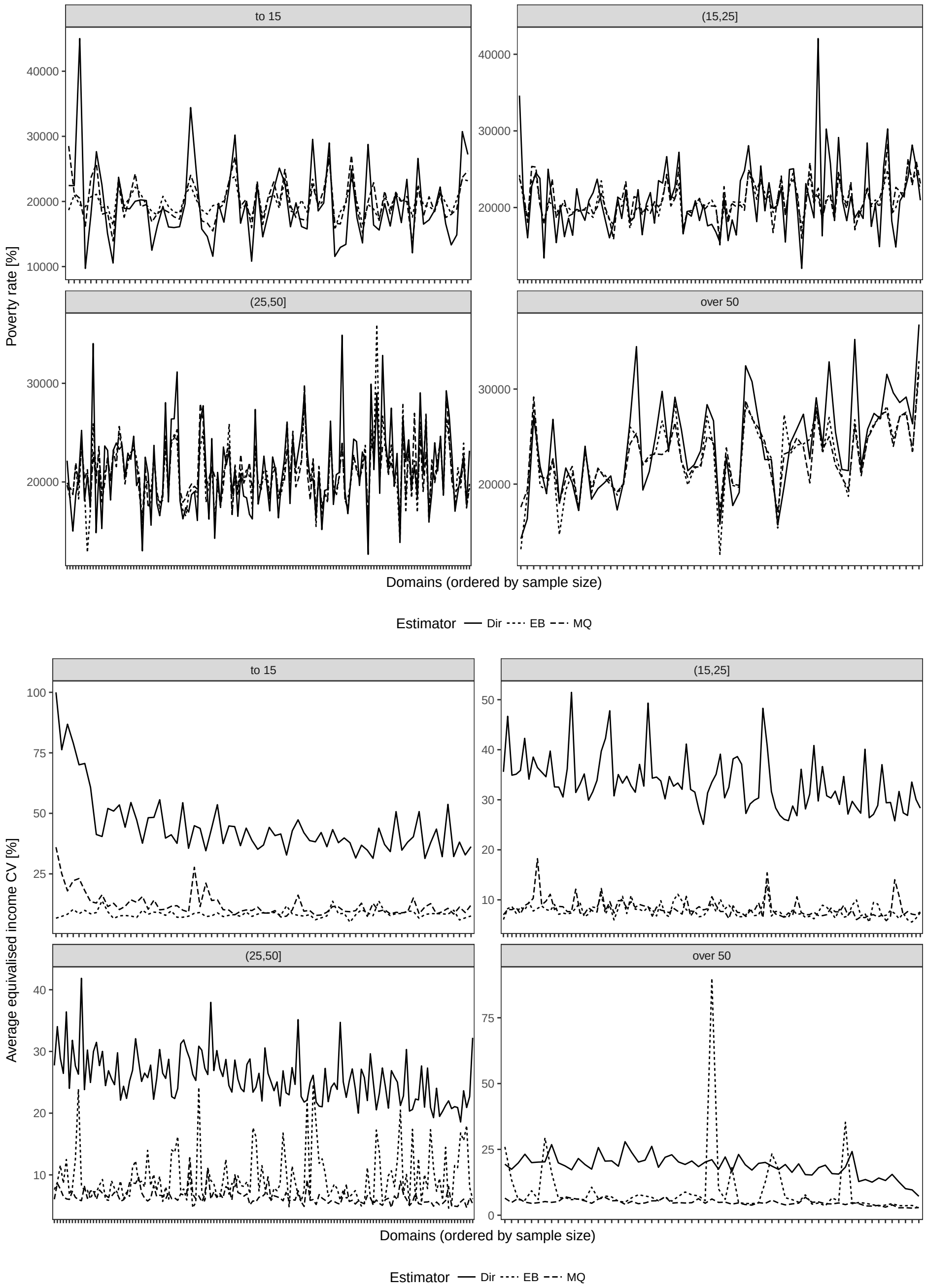

Figure 9. A comparison of direct, M-quantile and Empirical best estimates of the average equivalised income and their corresponding CV values grouped by sample size. Note that scale Y is not equal between plots

by means of model-based simulations, following the procedure described in Marhuenda et al. (2017) and by a design-based simulation experiment, based on EU-SILC data used in the application.

6. Conclusion

In 2013 the Central Statistical Office published poverty indicators for 66 subregions (NUTS 3) using the Fay-Herriot approach. However, the implementation of effective social policy requires information on poverty at lower levels of spatial aggregation (LAU 1). This goal can be achieved using small area estimation methods, which are becoming increasingly relevant for economic analysis and social policy as the availability of reliable estimates of economic aggregates for small geographical regions is crucial for policy planning or evaluation.

Indirect methods used for poverty mapping take into account between-area (EB, HB, M-quantile approaches) or between-cluster (ELL approach) variation beyond what is explained by auxiliary variables, but models do not consider between-area and between-cluster variability simultaneously. Moreover, small area estimation methods are often criticized for their reliance on model assumptions and may be affected by model failures such as those caused by the presence of outliers.

For these reasons in this paper we have proposed a three- level M-quantile model to estimate the HCR and the average equivalised income in Poland at district level (LAU 1). The spatial distribution of poverty in Poland obtained by using the three-level M-quantile model confirms the traditional division of Poland into the affluent, urbanized, and rapidly developing Western part and the less developed, rural and economically retarded Eastern part. The division of Poland in terms of poverty obtained by means of the SAE methodology also confirms the already recognized division of Poland into metropolitan areas and peripheral ones. The largest cities in Poland and surrounding districts are characterized by a lower poverty rate and a higher average equalized income. On the other hand, districts situated along the Eastern border are poorer and less developed.

The proposed methodology can be easily extended to the estimation of population quantiles. In the paper we have proposed a bootstrap MSE estimator; an analytic estimator of the MSE of estimates following the approach described by Chambers et al. (2014) will be a challenge. Finally, it would be good to consider the problem of benchmarking to assure consistency between a collection of small area estimates generated by a-three level M-quantile model with reliable estimates obtained by means of ordinary design-based methods for a collection of districts. It can be, for instance, achieved using the approach described in Fabrizi et al. (2012a) and Fabrizi et al. (2012b).

Acknowledgments

The authors are grateful to the anonymous referee and the associate editor for valuable comments and suggestions. These have led to a considerably improved paper. The study was funded from the European Commission's 7^{th} Framework Programme (FP7/2013-2017) under grant agreement no. 312691, InGRID – Inclusive Growth Research Infrastructure Diffusion. The work of Salvati has been developed under the support of the project PRIN- SURWEY `http://www.sp.unipg.it/surwey/` (Grant 2012F42NS8, Italy). The work of Marchetti and Salvati has been developed under the support of the European Union's Horizon 2020 research and innovation programme under grant agreement No 730998, InGRID-2 – Integrating Research Infrastructures for European expertise on Inclusive Growth from data to policy. The views expressed in this paper are those of the authors and do not necessarily reflect the policies of the Central Statistical Official in Poland.

References

Anderson, T. and Darling, D. A. (1952). Asymptotic theory of certain "goodness-of-fit" criteria based on stochastic processes. *Ann. Math. Statist.*, 23:192–212.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Softw.*, 67(1):1–48.

Bivand, R. (2015). *classInt: Choose Univariate Class Intervals.* R package version 0.1-23.

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

Brown, G., Chambers, R., Heady, P., and Heasman, D. (2001). Evaluation of small area estimation methods – an application to unemployment estimates from the UK LFS. *In Proc. Statistics Canada Symp. Achieving Data Quality in a Statistical Agency: a Methodological Perspective.*

Central Statistical Office (2011). The methodology of the 2011 National Population and Housing Census – selected aspects. Technical report, Central Statistical Office.

Central Statistical Office (2012). Income and Living Conditions of the Population of Poland. Report from the EU-SILC survey of 2011. Technical report, Central Statistical Office in Poland.

Central Statistical Office (2013). Jakość życia, kapitał społeczny, ubóstwo i wykluczenie społeczne w Polsce. Technical report, Central Statistical Office.

Chambers, R., Chandra, H., Salvati, N., and Tzavidis, N. (2014). Outlier robust small area estimation. *J. R. Statist. Soc. B*, 76:47–69.

Chambers, R. and Tzavidis, N. (2006). M-quantile models for small area estimation. *Biometrika*, 93(2):255–68.

Efron, B. (2004). The estimation of prediction error: covariance penalties and cross-validation. *J. Am. Statist. Ass.*, 99(467):619–632.

Elbers, C., Lanjouw, J. O., and Lanjouw, P. (2003). Micro-level estimation of poverty and inequality. *Econometrica*, 71(1):355–364.

European Commission (2010). Europa 2020 — A European strategy for smart, sustainable and inclusive growth. Technical report, European Commission.

Fabrizi, E., Giusti, C., Salvati, N., and Tzavidis, N. (2012a). Mapping average equivalized income using robust small area methods. *Pap. Reg. Sci.*, 93:685–702.

Fabrizi, E., Salvati, N., and Pratesi, M. (2012b). Constrained small area estimators based on M-quantile methods. *J. Off. Stat.*, 28:89–106.

Fay, R. and Herriot, R. (1979). Estimation of income for small places: An application of james-stein procedures to census data. *J. Am. Statist. Ass.*, 74 (366):269–277.

Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica*, 3(52):761–766.

Giusti, C., Marchetti, S., Pratesi, M., and Salvati, N. (2012). Robust small area estimation and oversampling in the estimation of poverty indicators. *Surv. Res. Methods-Ger.*, 6(3):155–163.

Greven, S. and Kneib, T. (2010). On the behaviour of marginal and conditional aic in linear mixed models. *Biometrika*, 97(4):773–789.

Hagenaars, A., de Vos, K., and Zaidi, M. (1994). Poverty statistics in the late 1980s: Research based on micro-data. *Office for Official Publications of the European Communities*, page Luxembourg.

Haughton, J. and Khandker, S., editors (2009). *Handbook on Poverty and Inequality.* World Bank.

Marchetti, S., Tzavidis, N., and Pratesi, M. (2012). Non-parametric bootstrap mean squared error estimation for m-quantile estimators of small area averages, quantiles and poverty indicators. *Comput. Stat. Data. An.*, 56(10):2889–2902.

Marhuenda, Y., Molina, I., Morales, D., and Rao, J. (2017). Poverty mapping in small areas under a twofold nested error regression model. *J. R. Statist. Soc. A.*

Molina, I. and Marhuenda, Y. (2015). sae: An R package for small area estimation. *The R Journal*, 7(1):81–98.

Molina, I., Nandram, B., and Rao, J. (2013). Small area estimation of general parameters with application to poverty indicators: a hierarchical Bayes approach. *Ann. Appl. Stat.*, 8:852–885.

Molina, I. and Rao, J. (2010). Small area estimation of poverty indicators. *Can. J. Stat.*, 38(3):369–385.

Pratesi, M., editor (2016). *Analysis of Poverty Data by Small Area Estimation.* Wiley.

Rao, J. N. K. and Molina, I. (2015). *Small Area Estimation.* Wiley.

Statistical Office in Poznan (2013). Poverty Maps at the Subregional Level in Poland Based on Indirect Estimation. Technical report, Statistical Office in Poznan.

Statistics Canada (2007). 2005 survey of financial security - public use microdata file user guide. *Published by authority of the Minister responsible for Statistics Canada, 2007, http://www.statcan.gc.ca/pub/13f0026m/13f0026m2007001-eng.htm.*

Tzavidis, N. and Marchetti, S. (2016). Robust domain estimation of income-based inequality indicators. In Pratesi, M., editor, *Analysis of Poverty Data by Small Area Estimation*, pages 171–186. Wiley.

Tzavidis, N., Marchetti, S., and Chambers, R. (2010). Robust estimation of small area means and quantiles. *Aust. NZ. J. Stat.*, 52(2):167–186.

Tzavidis, N., Salvati, N., Pratesi, M., and Chambers, R. (2008). M-quantile models with application to poverty mapping. *Stat. Method. Appl.*, 17(3):393–411.