G. Betti

THE EFFECT OF EQUIVALENCE SCALES ON POVERTY AT OBLAST LEVEL IN UKRAINE[1]

This paper aims at properly measuring and evaluating the impact of equivalence scales on poverty and inequality at both national and regional (Oblast) level in Ukraine. A new equivalence scale set is proposed and estimated on the basis of the UHLSC data; for some regions the precision of the estimate results as not being sufficient due to small sub-sample sizes. A variant of EBLUP small area estimation technique is proposed and implemented to estimate poverty measures properly and to reduce standard errors of such estimates; the variant concerned is based on a ratio approach: in this way the effect of the difficult-to-qualify institutional and historical factors, common to the country and its regions, is abstracted.

Keywords: equivalence scales; poverty and inequality; small area estimation

1. Introduction

Before joining the European Union (EU), some Eastern European countries have faced the need for adopting adequate equivalence scales during the transition period. During the period from mid '90s to mid 2000s, this has been the subject of debate in several of them. In particular, in Poland and Romania, the debate has brought to the proposal of new equivalence scales used in this transition period and officially adopted by GUS and INSTAT (see Betti, 1999a,b; Szulc, 2003; Molnar *et al.*, 2003). Under the new project funded by the World Bank[2], Ukraine is now following a similar path, although it seems to be full of hurdles as shown by the daily tension in the Country.

The main goal of this paper is properly measure and evaluate the impact of equivalence scales on poverty and inequality, at both national and regional (Oblast) level. This latter evaluation is particular relevant since measures of poverty and inequality are most useful to policy-makers and researchers when they are finely disaggregated, i.e. when they are estimated for small geographic units, such as regions, provinces, districts or other "local" administrative partitions of the country.

[2] World Bank Project "Development of State Statistical System for Monitoring the Social and Economic Transformations," component 4729/18 "Improvement of methodology and statistics organization."

However, it is essential to properly identify the poverty and inequality measures to be targeted in the evaluation analysis. The present dramatic international economic and financial crisis emphasizes the role of social equity in policymaking, in order to effectively combat the increasing impoverishment of relevant segments of the population due to unemployment, bankruptcy, and difficulties related to bank credit. In the light of this, the recent contribution of Betti and Lemmi (2013) provides a deep description of research themes on poverty and inequality which constitute a consistent and relevant part of the recent economic, econometric and statistical methods and empirical analyses in the international scientific literature. They also play a seminal role in policymaking, given the importance of inequality and poverty reduction, and of improving living conditions within the framework of sustainable economic development.

Such themes are categorised as: i) conventional and multidimensional poverty; ii) longitudinal and chronic poverty; iii) small area estimation methods.

The focus of the present paper is the evaluation of the impact of equivalence scales on conventional monetary measures at national and regional level; in the latter case analysis is performed by means of the Empirical Best Linear Unbiased Estimator (EBLUP); European Commission (2005) and Betti *et al.* (2012) provide a theoretical and empirical base for such an approach at the EU level.

The paper is made up of six sections; after the present introduction, Section 2 defines the concept of equivalence scale and describes the most widely adopted methods for estimating them. In Section 3, we estimate equivalence scales based on the Engel method and the Almost Ideal Demand System (AIDS) on the basis of the 2008 UHLSC survey in Ukraine. On the basis of such empirical evidence, a new set of equivalence is proposed for official national purposes. Section 4 is devoted to the evaluation of the impact of such scales on traditional poverty and inequality measures at national level: here a set of seven alternative equivalence scales are also compared. In Section 5 the regional dimension is analysed further; poverty estimates based on the new equivalence scale proposed in Section 3 are improved by means of the EBLUP technique; as usual, the last section provides concluding remarks and describes possible future research. We aim at evaluating the effect of equivalence scales on multidimensional poverty in future research following the approach of Betti and Verma (2008), and the effect on longitudinal and chronic poverty following the approaches proposed by Cheli and Betti (1999), Betti *et al.* (2002) and Betti *et al.* (2004).

2. Equivalence scales

Equivalence scales are economic index numbers, which discount household total consumption expenditure (or income) in accordance with some household characteristics. Although equivalence scales are generally considered a necessary tool, there is no unanimity in the way they have to be calculated. Buhmann *et al.* (1988) and Hagenaars *et al.* (1994) present a broad classification:

a. normative and social security equivalence scales,

b. equivalence scales based on consumption or expenditure,

c. equivalence scales based on direct welfare measurement.

Equivalence scales evaluated in this paper belong to both class a and b.

2.1. Normative and social security equivalence scales.

Normative equivalence scales are based on some norms set by experts in defining a minimum level of consumption or basket of goods for households of different composition and size. Sometimes these norms directly define the scale values. The OECD-scale, also known as Oxford scale, uses weights equal to one for the first adult, 0.7 for each of the following adults and 0.5 for each child under 14. Hagenaars *et al.* (1994) introduce a modified OECD-scale, which presents lower elasticity of family size: this scale gives a value of 0.5 for each adult except the first and 0.3 for each child. This scale is currently adopted by Eurostat for poverty and inequality in the EU-SILC.

In Ukraine the Academy of Science defines a scale, which is equal to one for the first adult, 0.7 for each of the following adults and 0.7 for each child under 14. The World Bank defines the scale to be equal to one for every person, i.e. consider the per-capita income or consumption expenditure (no economies of scale). This method is, however, used for less developed or developing countries.

2.2 Equivalence scales based on consumption or expenditure.

This is the most widely used methodology in economic literature; equivalence scales are derived using data sets on household expenditure.

Engel (1895) presents the first important work on equivalence scales, based on the assumption that the household welfare, or standard of living of adults, is strongly related to the share of the budget devoted to food. For a fixed characteristic household set the food share is inversely related to total expenditure (Engel's law) and, for a fixed level of total expenditure, the food-ratio is a direct function of the number of children. To restore the food share after the birth of a child the reference household (couple) would need to reach a higher level of total expenditure or income.

The development in constructing models suitable for equivalence scales calculation has been mostly focused on the introduction of demographic variables into demand systems.

Barten (1964) considers the utility function associated with the household demographic characteristics to be:

$$u = v\left[\frac{q_1}{m_1(\mathbf{z})}, \frac{q_2}{m_2(\mathbf{z})}, \ldots, \frac{q_n}{m_n(\mathbf{z})}\right], \qquad (1)$$

where $m_i(\mathbf{z})$ is the equivalence scale for the particular good i q_i is the corresponding consumption, $\mathbf{z}$ is the vector of household characteristics and n the number of consumer goods; all $m_i(\mathbf{z})$ are equal to unity in the case of the reference household. Although the model is more general than the previous one, there is a drawback due to the evaluation of equivalence scales for goods that are not consumed in the reference household (for example child food). Gorman (1976) presents a modification pointed out in the cost function:

$$x = c[u, p_1 m_1(\mathbf{z}), p_2 m_2(\mathbf{z}), \ldots, p_n m_n(\mathbf{z})] + \sum_i p_i n_i(\mathbf{z}), \qquad (2)$$

where p_i is the price of good i and $n_i(\mathbf{z})$ is the corresponding fixed consumption, so that the added term on the right hand side represents the fixed cost associated with the demographic characteristic vector $\mathbf{z}$. In the last decade, many authors have pointed out the need for setting up new models incorporating demographic variables; for example Pollak and Wales (1981) and Lewbel (1985).

3. Estimation of equivalence scales for Ukraine

In the last fifteen years the State Statistical office of some Eastern European transition countries and the former Soviet Republic have estimated ad hoc equivalence scales for measuring poverty and inequality measures for national purposes (see Betti (1999a,b) for Poland; Molnar *et al.* (2003) for Romania; Betti and Lundgren (2012) for Tajikistan).

The present section aims at estimating equivalence scales based on consumption expenditure from the 2008 Ukrainian Household Living Conditions Survey (UHLSC); it is the main source of information for comprehensive research on the welfare of the Ukrainian population. UHLCS has been working on a permanent basis since 1999, when it was established by the SSCU. Survey methodology in general complies with commonly accepted international standards and requirements for preparation and for carrying out of state sampling surveys on population.

3.1. The Engel method

The first model estimated on 2008 UHLSC data, is based on the so-called food ratio method. This model refers to Van Ginneken (1982):

$$\ln F_i = a + b\ln C_i + c\ln N_i + e_i \qquad (3)$$

where F is the food expenditure, C is the total consumption expenditure and N is the family size. According to Engel's (1895) law — the household's standard of living varies inversely with the food ratio — it is possible to derive the economies of scale ε (for $d(F/C) = 0$):

$$\varepsilon = \frac{\partial \ln C}{\partial \ln N} = \frac{c}{1-b} \qquad (4)$$

The model estimated on 2008 UHLSC data points out quite a high value of the elasticity, which reaches the value of 0.8, as follows:

$$\varepsilon = \frac{\partial \ln C}{\partial \ln N} = \frac{c}{1-b} = \frac{0.80736}{1-(-0.00631)} = 0.802298 \qquad (4')$$

3.2. The complete demand system AIDS

The model considered here is the AI system of Deaton and Muellbauer (1980), whose indirect utility function is defined as follows: $v(x,\mathbf{p}) = \frac{\ln x}{b(\mathbf{p})}$, where $x = \frac{\mu}{a(\mathbf{p})}$ is the expenditure in real terms; $\ln a(\mathbf{p}) = \alpha_0 + \sum \alpha_k \ln p_k + \frac{1}{2}\sum\sum \gamma_{kj} \ln p_k \ln p_j$ and $\ln b(\mathbf{p}) = \sum \beta_k \ln p_k$ are price indices. The associated cost function which results is:

$$\ln C(u,\mathbf{p}) = \alpha_0 + \sum \alpha_k \ln p_k + \frac{1}{2}\sum\sum \gamma_{kj} \ln p_k \ln p_j + \prod p_k^{\beta_k} u, \qquad (5)$$

and from Roy's identity one can obtain the budget shares:

$$w_i = \frac{\partial \ln a(\mathbf{p})}{\partial \ln p_i} + \frac{\partial \ln b(\mathbf{p})}{\partial \ln p_i} \ln x = \alpha_i + \sum \gamma_{ij} \ln p_j + \beta_i(\ln\mu - \ln a(\mathbf{p})). \qquad (6)$$

This simple Engel curve can be extended in several directions, introducing:

— nonlinearities in the Engel curves;

— socio-demographic variables.

Non linearities in the model could be introduced by means of a linear and homogenous specification of $\ln \varphi(\mathbf{p}) = \sum \varphi_k \ln p_k$, which leads to the Banks, Blundell and Lewbel (1997) demand sys-

Table 1

AIDS_PS1 estimates (*p*-values are reported in parentheses)

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
α_i	1.4681 (<.0001)***	–0.1502 (<.0001)***	0.1846 (<.0001)***	0.0730 (<.0001)***	–0.2421 (<.0001)***	–0.2889 (<.0001)***
β_i	–0.0797 (<.0001)***	0.0198 (<.0001)***	–0.0042 (<.0001)***	–0.0024 (<.0001)***	0.0278 (<.0001)***	0.0322 (<.0001)***
$\gamma_{1,i}$	–0.3918 (0.0018)***	0.1076 (0.3791)	–0.2375 (0.0505)*	0.2778 (0.0250)**	0.2211 (0.0697)*	–0.0512 (0.6777)
$\gamma_{2,i}$	0.0807 (0.1155)	–0.0048 (0.9272)	0.0065 (0.9012)	0.0352 (0.5090)	–0.0924 (0.0778)*	–0.0726 (0.1713)
$\gamma_{3,i}$	0.0610 (0.4686)	0.0260 (0.7536)	0.0950 (0.2504)	–0.1479 (0.0789)*	–0.1821 (0.0276)**	0.1480 (0.0771)*
$\gamma_{4,i}$	–0.1171 (0.0470)**	–0.0695 (0.2311)	0.0049 (0.9321)	–0.0029 (0.9610)	0.1540 (0.0078)***	0.0912 (0.1195)
$\gamma_{5,i}$	0.1154 (0.0313)**	0.0045 (0.9313)	0.0561 (0.2843)	–0.1325 (0.0131)**	–0.0966 (0.0662)*	0.0787 (0.1386)
$\gamma_{6,i}$	0.1585 (0.0053)***	–0.0596 (0.2841)	0.0658 (0.2364)	–0.0048 (0.9318)	–0.0112 (0.8409)	–0.1353 (0.0161)**
	$\hat{\tau}_1$ = 0.0884 (<.0001)***	$\hat{\tau}_2$ = 0.1136 (<.0001)***	$\hat{\tau}_3$ = 0.1460 (<.0001)***	$\hat{\tau}_4$ = 0.1921 (<.0001)***	$\hat{\tau}_5$ = 0.1249 (<.0001)***	

*** Parameter significant at 99 % level.
** Parameter significant at 95 % level.
* Parameter significant at 90 % level.

Table 2

AIDS_PS2 estimates (*p*–values are reported in parentheses)

	Group 1	Group 2	Group 3	Group 4	Group 5	Group 6
α_i	1.4676 (<.0001)***	–0.1499 (<.0001)***	0.1846 (<.0001)***	0.0730 (<.0001)***	–0.2419 (<.0001)***	–0.2887 (<.0001)***
β_i	–0.0709 (<.0001)***	0.0198 (<.0001)***	–0.0042 (<.0001)***	–0.0024 (0.0005)***	0.0278 (<.0001)***	0.0322 (<.0001)***
$\gamma_{1,i}$	–0.3917 (0.0018)***	0.1076 (0.3793)	–0.2376 (0.0505)*	0.2778 (0.0250)**	0.2211 (0.0698)*	–0.0513 (0.6773)
$\gamma_{2,i}$	0.0840 (0.1157)	–0.0048 (0.9275)	0.0065 (0.9012)	0.0352 (0.5091)	–0.0923 (0.0778)*	–0.0726 (0.1714)
$\gamma_{3,i}$	0.0610 (0.4685)	0.0260 (0.7536)	0.0950 (0.2504)	–0.1479 (0.0789)*	–0.1821 (0.0276)**	0.1480 (0.0771)*
$\gamma_{4,i}$	–0.1171 (0.0470)**	–0.0695 (0.2311)	0.0049 (0.9321)	–0.0029 (0.9610)	0.1540 (0.0078)***	0.0912 (0.1195)
$\gamma_{5,i}$	0.1154 (0.0314)**	0.0045 (0.9314)	0.0562 (0.2843)	–0.1325 (0.0131)**	–0.0966 (0.0663)*	0.0787 (0.1384)
$\gamma_{6,i}$	0.1584 (0.0053)***	–0.0596 (0.2843)	0.0658 (0.2364)	–0.0049 (0.9316)	–0.0111 (0.8411)	–0.1353 (0.0161)**
	$\hat{\tau}_1$ = 0.4062 (<.0001)***	$\hat{\tau}_2$ = 0.6551 (<.0001)***	$\hat{\tau}_3$ = 0.7364 (<.0001)***	$\hat{\tau}_4$ = 0.8005 (<.0001)***	$\hat{\tau}_5$ = 0.6595 (<.0001)***	

*** Parameter significant at 99% level.
** Parameter significant at 95% level.
* Parameter significant at 90% level.

tem Quadratic Almost Ideal (QUAIDS). From this model the corresponding budget shares are:

$$w_i = \alpha_i + \sum \gamma_{ij} \ln p_j + \beta_i \ln x + \frac{\varphi_i}{b(\mathbf{p})} (\ln x)^2. \quad (7)$$

Evidence from UHLSC has shown that there is no need to estimate such a complex model. It has then been decided to keep the AIDS model, and to introduce the socio-demographic variables using the Ray (1983) Price Scaling method: $\ln\left(\frac{x}{m(\mathbf{p},\mathbf{z})}\right)$, where $m(\mathbf{p}, \mathbf{z})$ is defined in two alternative ways:

$$m(\mathbf{p},\mathbf{z}) = 1 + \sum_{i=1}^{5} \tau_i z_i,$$

where z_4 is equal to the number of adults minus one; or: $m(\mathbf{p},\mathbf{z})=\sum_{i=1}^{5}\tau_i z_i$.

In order to estimate the AIDS model, the 12 consumption expenditure components, which follow the COICOP classification, have been collapsed into seven groups. This is because the expenditure pattern in Ukraine is concentrated on food expenditure (about 55 %), and most of the other components rarely reach 5 % each. The aggregation used is the following:

	FROM	TO
Food and non-alcoholic beverages	1	1
Alcoholic beverages, tobacco	2	6
Clothing and footwear	3	2
Housing, water, electricity, gas and other fuels	4	3
Furnishings, household equipment and routine maintenance of the house	5	3
Health	6	4
Transport	7	5
Communication	8	5
Recreation and culture	9	6
Education	10	4
Restaurants and hotels	11	6
Miscellaneous goods and services	12	7

Tables 1 and 2 report the estimated parameters of the two AIDS models.

3.3. A new proposal for equivalence scales in Ukraine

In this proposal, we aim at estimating new equivalence scales for Ukraine; the empirical analysis is carried out on the basis of the Ukrainian Household Budget Survey data for the year 2008 and follows towards two complementary directions.

First of all, a simple regression model representing an Engel food ratio curve has been estimated a là Van Ginneken (1982). This produced an overall elasticity, of the consumption expenditure with respect to the size of the family, very high ($\varepsilon = 0.802298$): this result can allow us to define the equivalence scale of a subsequent adult aged 18–64 as being at least equal to the value 0.8.

In order to estimate the cost of any elderly person or child, an AIDS complete demand system has been estimated, with the introduction of socio-demographic variables conducted with two version of the Price Scaling method. In both models (Tables 1 and 2), all parameters referring to the socio-demographic variables are significantly different from zero.

According to the parameters, the scale for the eldest person in the family should be about 80 % of the value of the first adult aged 19–64, while the scale of any subsequent elder should be about 65 % of that value. The parameter for children aged 15–18 suggests that the cost of those children is similar, but slightly less, than any subsequent adult, while the scale for children aged 6–14 should be about 60 % of the first adult. Finally, the scale of any children aged 0–5 should be equal to 0.45. Table 3 summarizes the results described above.

Table 3

New equivalence for Ukraine

Category	Scale
First adult aged 18–64	1.00
Any subsequent adult aged 19–64	0.80
First adult aged 65 or more	0.80
Any subsequent adult aged 65 or more	0.65
Children aged 14–17	0.75
Children aged 7–13	0.60
Children aged 0–6	0.45

This is the first original contribution of the paper.

4. Effects of scales on poverty and inequality

This section describes the effect of a set of seven equivalence scales on poverty and inequality at national level. It also shows how this effect could be more evident when calculating measures at regional (Oblast) level, even if for some regions the sub-sample sizes are not sufficiently large: this issue will later be addressed in Section 5.

4.1. Evaluating the effects on poverty

For evaluating the effect of the choice of equivalence scales, five poverty measures have been taken into account, namely the three measures of the FGT class (Foster, Greer and Thorbecke, 1984) with parameters 0 (Head Count Ratio), 1 (Poverty Gap Ratio) and 2 (Severity of Poverty Index); the Sen index and the Gini index of the poor.

Table 4 reports the results obtained on the five poverty measures estimated on the basis of the total household expenditure, equivalised by seven set of equivalence scales:

1) Engel method, with $\varepsilon = 0.802298$;
2) OECD 70–50;
3) OECD-modified 50–30;
4) Academy of Science 70–70;
5) New proposed scale in Section 3;
6) Per-capita household expenditure ($\varepsilon = 1$);
7) No scale, i.e. total household expenditure ($\varepsilon = 0$).

In fact, when $\varepsilon = 0$ economies of scale are perfect (independent of the household size); on the other hand, when $\varepsilon = 1$ there are not economies of scale (per-capita household expenditure).

The analysis has been conducted on the basis of two different poverty lines: 60 % and 75 %

Table 4

Welfare variable: Total expenditure

Poverty line		HCR	Poverty Gap	FGT(2)	SEN	Gini of the poor
60 % median	Engel	0.13413	0.02786	0.00885	0.03915	0.10663
	OECD	0.13406	0.02656	0.00832	0.03784	0.10523
	OECD-mod	0.13015	0.02568	0.00794	0.03641	0.10298
	Academy of Science	0.13637	0.02842	0.00909	0.04002	0.10782
	New proposal	0.14120	0.02922	0.00933	0.04121	0.10750
	Per-capita	0.14816	0.03160	0.01036	0.04456	0.11159
	None $\varepsilon = 0$	0.18600	0.04986	0.01947	0.06892	0.14016
75 % median	Engel	0.26777	0.06254	0.02172	0.08699	0.11932
	OECD	0.26884	0.06149	0.02096	0.08557	0.11632
	OECD-mod	0.27036	0.06069	0.02037	0.08451	0.11378
	Academy of Science	0.26969	0.06316	0.02209	0.08806	0.12074
	New proposal	0.27107	0.06464	0.02272	0.08966	0.12143
	Per-capita	0.27583	0.06730	0.02427	0.09351	0.12591
	None $\varepsilon = 0$	0.30638	0.08943	0.03751	0.12302	0.15495

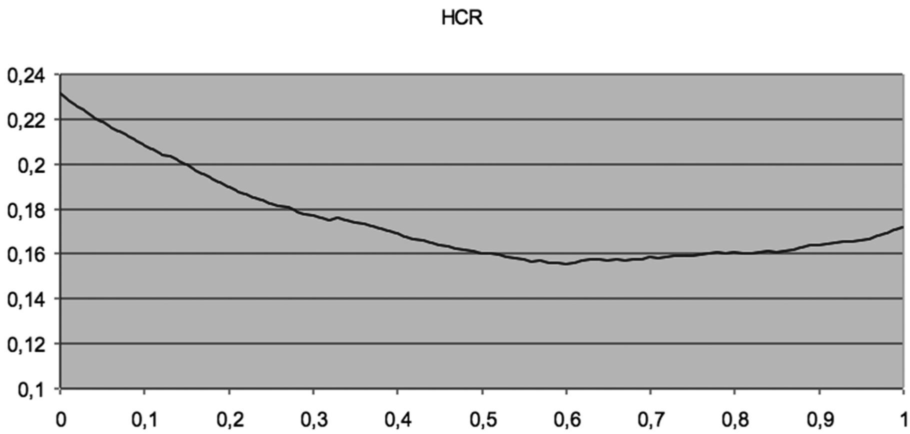

Fig. 1. *Sensitivity analysis: Head Count Ratio and elasticity*

of median equivalised household expenditure. All five measures present a similar pattern when observing the two equivalence scales corresponding to the extremes of the elasticity ε: when $\varepsilon = 0$ total household expenditure is taken into account, the poverty line is higher, and small households are much worse-off compared to the rest of the population. On the other hand, when $\varepsilon = 1$ per-capita household income is taken into account, the poverty line is much lower, and large households are slightly worse-off compared to the rest of the population. The intermediate patterns are shown in Figure 1 where parameter ε varies from 0 to 1 and the poverty line is calculated as 60 % of median equivalised household expenditure. For instance, the minimum is at about $\varepsilon = 0.6$.

The effect of choosing one of the other five equivalence scales is less evident when performing poverty analysis at national level. The effect is much more evident when analysing the poverty profiles, in terms of household composition. In particular, the new proposal is able to take into account also the differences between adults and the elderly, and the difference between children in different age-groups.

4.2. Evaluating the effects on inequality

For evaluating the effect of the choice of equivalence scales, six inequality measures have been taken into account; the two Laeken indicators Gini index and S80/S20, the Theil index of generalised Entropy GE(1), and S90/S10, P80/P20 and P90/P10.

Table 5 reports results of the six inequality measures estimated on the basis of three total household monetary variables, equivalised by the seven equivalence scales described in section 4.1; the three variables are total cash income, total income and total expenditure.

As expected, the two extreme equivalence scales with elasticities $\varepsilon = 0$ and $\varepsilon = 1$ lead to monetary distributions which are more unequal, since the way in which households of different size are treated is unequal. The effect of the other equivalence scales is not so evident observing inequality measures at national level as already shown for poverty measures in section 4.1.

Table 5

Calculating inequality indices based on 7 scales, for the most relevant monetary variables

Variable		Gini	S80/S20	S90/S10	Theil	P80/P20	P90/P10
Cash income	Engel	0.2761	3.9825	5.9833	0.1384	2.1588	3.2946
	OECD	0.2784	4.0380	6.0670	0.1404	2.1744	3.3242
	OECD-mod	0.2768	4.0214	5.9291	0.1376	2.2059	3.3835
	Academy of Science	0.2801	4.0805	6.2192	0.1426	2.1787	3.3561
	New proposal	0.2729	3.9696	6.0375	0.1361	2.1066	3.3136
	Per-capita	0.2849	4.2024	6.4594	0.1482	2.1842	3.4778
	None $e = 0$	0.3128	5.1876	8.3116	0.1697	2.5907	4.3687
Total income	Engel	0.2554	3.5458	5.1309	0.1195	2.0115	2.9881
	OECD	0.2577	3.5909	5.2064	0.1213	2.0224	2.9982
	OECD-mod	0.2555	3.5584	5.1076	0.1184	2.0349	3.0283
	Academy of Science	0.2599	3.6331	5.3019	0.1236	2.0404	3.0328
	New proposal	0.2538	3.5479	5.1558	0.1184	1.9989	2.9765
	Per-capita	0.2657	3.7639	5.5298	0.1295	2.0827	3.1571
	None $e = 0$	0.2962	4.7087	7.3621	0.1521	2.4214	4.0021
Total expenditure	Engel	0.2886	4.2099	6.2714	0.1560	2.2437	3.4609
	OECD	0.2905	4.2517	6.3641	0.1580	2.2422	3.4830
	OECD-mod	0.2861	4.1538	6.1590	0.1532	2.2281	3.4165
	Academy of Science	0.2936	4.3249	6.4953	0.1612	2.2583	3.5764
	New proposal	0.2898	4.2614	6.3478	0.1567	2.2554	3.5142
	Per-capita	0.3008	4.4950	6.7995	0.1688	2.3110	3.7161
	None $e = 0$	0.3151	5.0826	7.9357	0.1777	2.5188	4.2800

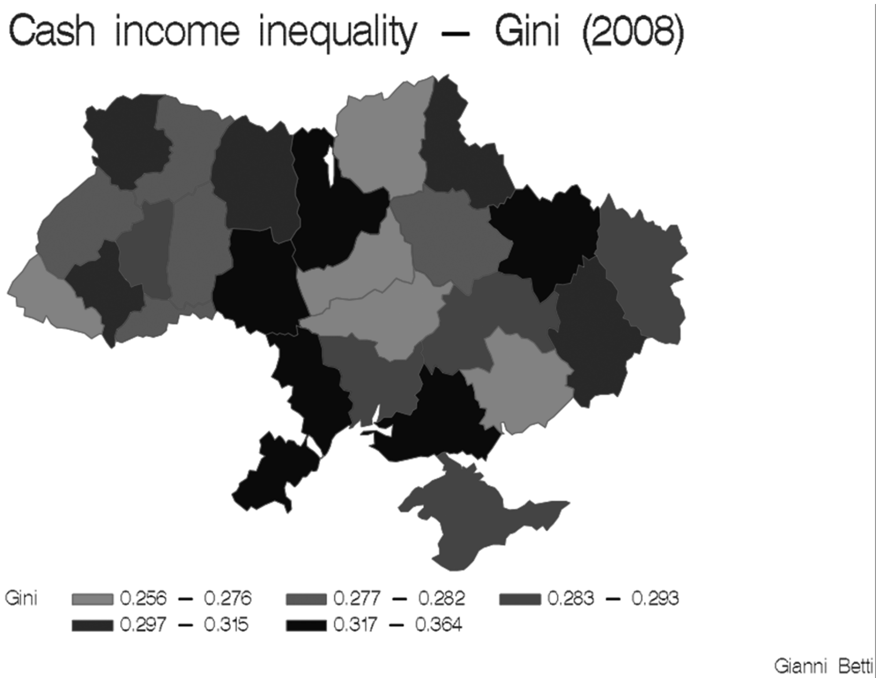

Fig. 2. *Gini inequality index on cash income, scale $\varepsilon = 0$*

It is not worth calculating inequality indices by population groups; on the other hand, it is very useful to perform this analysis at local level, in particular taking into account the Oblast administrative partition of Ukraine.

Figures 2 and 3 show the Gini coefficient calculated on cash income distribution equivalised by the scale with $\varepsilon = 0$ (Figure 2) and by the Academy of Science scale (Figure 3). It is possible to see that the change in the scale naturally changes the inequality measures, but it also changes the ranking among Oblasts. This change in the ranking of Oblasts is due to two main factors: i) the different equivalence scale adopted, and ii) the sampling variability, which is more effective for Oblast with lower subsample sizes.

In order to diminish the effect of sampling variability, the next section presents a method for es-

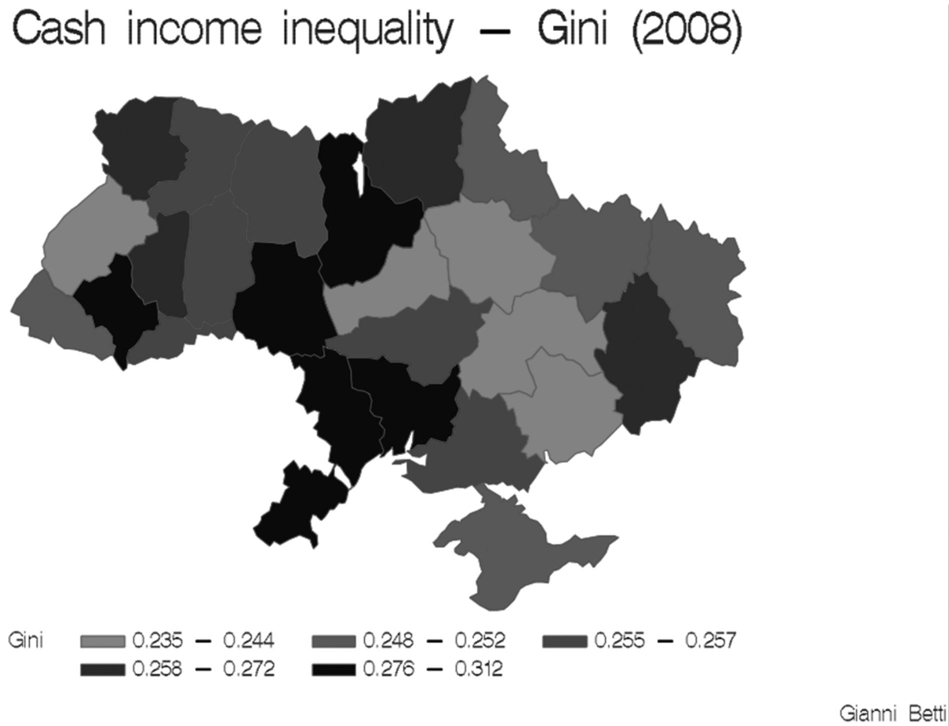

Fig. 3. *Gini inequality index on cash income, scale Academy of Science*

timating poverty (and inequality) measures at local level (Oblast in Ukraine).

5. Poverty estimated at Oblast level: the EBLUP technique

For the estimation of measures at regional level via small area estimation techniques, we believe that a good procedure to use is the Empirical Best Linear Unbiased Predictor (EBLUP) estimator, with appropriate procedures to evaluate the robustness of such measures.[1]

In this methodology, an intensive and small-scale survey such as UHLCS provides direct inequality and poverty-related information at the micro (unit) level; this information can be aggregated to areas such as NUTS regions where the survey contains some sample units and the areas identified are available in the microdata. On the other hand, correlates of poverty-related characteristics of the areas can be derived from aggregated statistics (such as Census or other administrative sources). The two sources can be combined to produce composite estimates, provided that (i) the survey data contain information for the identification of the area to which each unit belongs (which, unfortunately, is not always the case in UHLCS data files), and (ii) the aggregate data on the correlates are available for all the areas in the population of interest (which fortunately is the case for many correlates available in external data sources).

The approach can be to apply area level random-effect models relating small area direct estimates (from UHLCS) to domain specific covariates, considering the random area effects as independent. The basic area-level model includes random area specific effects and the area specific covariates, $x_i = (x_{i,1}, x_{i,2}, \ldots, x_{i,p})$, $i = 1, \ldots, m$, are related to the target parameters θ_i (totals, means, proportion, etc.) as: $i = x_i\beta + z_i\nu_i$, where z_i are known positive constants, β is the regression parameter vector px_1, ν_i are independent and identically distributed random variables with 0 mean and variance σ_v^2. The model assumes that the direct estimators $\hat{\theta}_i$ are available and design unbiased, in the form:

$$\hat{\theta}_i = \theta_i + e_i,$$

where ei are independent sampling errors with zero mean and known variance ψ_i. The BLUP estimator is a weighted average of the design-based

[1] In the literature small area models are classified as: (i) area-level random effect models, which are used when auxiliary information is available only at area level (such as the prevailing unemployment rate); (ii) nested-error unit level regression models, used if unit specific covariates (such as the individual's or the household's employment situation) are available at unit level.
On the basis of empirical work, it appears that area-level synthetic estimates tend to produce better results than their unit-level counterparts. This is because regression coefficients calculated at unit-level do not always correctly reflect the relationship between the area-level averages involved in the synthetic estimator. In any case, the type of data available for poverty analysis at regional level generally precludes the use of unit (household or person) level models.

estimator and the regression synthetic estimator $\tilde{\theta}_i(\sigma_v^2) = \gamma_i \hat{\theta}_i + (1+\gamma_i) x_i \hat{\beta}$, where:

$$\gamma_i = \frac{\sigma_v^2}{\sigma_v^2 + \psi_i}$$

is a weight (or 'shrinkage factor') which assumes values in the range [0–1]. This parameter measures the uncertainty σ_v^2 in modelling θ_i in relation to the total uncertainty including the variance of the direct estimator φ_i. The mean square error of the BLUP estimator depends on the variance parameter σ_v^2, which in practice is replaced by its estimator; hence the estimator obtained is called Empirical BLUP (EBLUP).

5.1. The proposed ratio approach in constructing SAEs

We can expect the predictive power of the model at the regional level to be substantially improved when the target variables, as well as the covariates, are expressed in terms of their values at the preceding higher level. Thus for NUTS1 region i (OBLAST), all target variables and all covariates in the model could be expressed in the form of the ratio:

$$R_j = Y_i / Y_0,$$

where (Y_i, Y_0) refer to the actual values of the variables for NUTS1 *i* and its country (Ukraine), respectively. We propose to define this as the 'ratio approach', which constitutes the second original contribution of the paper. In this way, the effect of the difficult-to-qualify institutional and historical factors, common to the country and its regions, is abstracted. Similarly, in going from NUTS1 region *i* to its NUTS2 region *j*, we could express the model variables in the form $R_{ij} = Y_{ij} / Y_i$, and similarly from NUTS2 to NUTS3 (Rajons) in the form $R_{ijk} = Y_{ijk} / Y_{ij}$.

The same ideas are extended to the modelling of subpopulations, such as children, the elderly, single person households, etc. We may simply model the ratio of the subpopulation measure to the total population measure.

5.2. The calculation of the standard errors in the ratio approach

As noted above, it could be more efficient to model the small area estimates in a hierarchical manner. Instead of estimating the absolute value of any poverty statistic (say p_1), we estimate the ratio ($r = p_1 / p$) of the statistic at one level such as Oblast, to its estimate at the preceding (higher) level such as Ukraine. The objective is to obtain var(r), given var(p_1). We have:

$$\operatorname{var}(r) = \operatorname{var}\left(\frac{p_1}{p}\right) =$$

$$= \frac{1}{p^2}\left(\operatorname{var}(p_1) + r^2 \operatorname{var}(p_1) - 2r\operatorname{cov}(p, p_1)\right).$$

The covariance is easily evaluated by noting that sample "1" is just a subsample of the HBS, with the same measurements so that correlation between them is 1.0. It can be shown that with n1 as the size of the subsample of sample *n*:

$$\operatorname{cov}(p, p_1) = \left[\operatorname{var}(p)\operatorname{var}(p_1)(n_1/n)\right]^{1/2}.$$

Since var(p) and var(p_1) are variances of complex measures (poverty and inequality) from a complex survey (multistage sampling design), they cannot be calculated with simple techniques of variance estimation; we have applied the Jackknife Repeated replication (JRR) technique as described in Verma and Betti (2011).

5.3. Empirical analysis for SAE for regional indicators in Ukraine

In this Section, we present some empirical results of EBLUP estimates based on UHLCS 2008 data for NUTS1 regions (OBLAST) in Ukraine. The variable of interest is the Head Count Ratio, defined as the percentage of individuals belonging to poor families. Poor families are defined as those families with Total "equivalent" consumption expenditure (adopting the new proposed equivalence scale) below the poverty line defined as 75 % of the median. The regressors used in the model have been supplied by SSCU and are reported in the following Table 6.

In order to choose the best regressors for the analysis, some simple regressions (OLS) have been performed. The best estimated model is provided upon request.

Table 6

Covariates available at NUTS1 (OBLAST) level

1	Disposable income	Average monthly wage
2	GDP	GDP per capita 2007
3	Activity rate	Activity rate for 2008; Males, Females and Total
4	Unemployment rate	Unemployment rate 2008; Males, Females and Total
5	Urbanisation	Percentage of urban population
6	Population density	Population closeness, persons in 1 km^2
7	IMR	Infant mortality rate 2008; death rate of children under 1 year old
8	HH Size	Mean size of household, 2008
9	Turnover	Turnover for one person 2008, hrn
10	Youths	Percentage of children under 14
11	Elderly people	Percentage of people 65 years old and over

Table 7

Small area (EBLUP) estimates of at-risk-of-poverty rates for Oblasts

	Oblast	n	est	se	gamma	est	stat_se	ratio_est	ratio_MSE
		(1)	(2)	(3)	(4)	(5)	(6)	(7)=(5)/(2)	(8)=(6)/(3)
1	AR Crimea	462	24.89 %	3.50 %	0.74	23.75 %	3.36 %	0.95	0.96
5	Vinnytska	426	24.30 %	3.38 %	0.75	26.79 %	3.22 %	1.10	0.95
7	Volynska	287	47.30 %	5.92 %	0.50	43.10 %	5.08 %	0.91	0.86
12	Dnipropetrovska	760	27.22 %	3.59 %	0.73	29.38 %	3.43 %	1.08	0.96
14	Donetska	734	25.58 %	2.85 %	0.81	25.02 %	2.79 %	0.98	0.98
18	Zhytomyrska	326	34.22 %	6.04 %	0.49	34.59 %	4.82 %	1.01	0.80
21	Zakarpatska	310	20.30 %	4.99 %	0.58	23.17 %	4.80 %	1.14	0.96
23	Zaporizka	441	24.39 %	4.04 %	0.68	22.91 %	3.70 %	0.94	0.92
26	Ivano-Frankivska	300	20.09 %	3.41 %	0.75	22.41 %	3.34 %	1.12	0.98
32	Kyivska	350	21.17 %	4.86 %	0.60	18.39 %	4.52 %	0.87	0.93
35	Kirovogradska	291	47.09 %	5.05 %	0.58	42.42 %	4.37 %	0.90	0.86
44	Luganska	566	29.20 %	3.12 %	0.78	29.45 %	3.08 %	1.01	0.99
46	Lvivska	563	29.25 %	2.81 %	0.82	28.96 %	2.73 %	0.99	0.97
48	Mykolaivska	312	19.81 %	3.53 %	0.74	20.34 %	3.30 %	1.03	0.93
51	Odeska	393	38.48 %	4.53 %	0.63	36.91 %	4.18 %	0.96	0.92
53	Poltavska	432	27.47 %	3.23 %	0.77	28.10 %	3.16 %	1.02	0.98
56	Rivnenska	287	39.28 %	6.02 %	0.49	39.23 %	5.22 %	1.00	0.87
59	Sumska	315	29.07 %	5.92 %	0.50	27.85 %	4.78 %	0.96	0.81
61	Ternopilska	250	42.77 %	7.16 %	0.41	38.32 %	5.86 %	0.90	0.82
63	Kharkivska	585	20.59 %	2.28 %	0.87	20.43 %	2.26 %	0.99	0.99
65	Khersonska	319	32.96 %	4.66 %	0.62	30.44 %	4.03 %	0.92	0.86
68	Khmelnytska	328	29.79 %	4.02 %	0.69	28.49 %	3.59 %	0.96	0.89
71	Cherkaska	394	18.37 %	3.72 %	0.72	20.23 %	3.44 %	1.10	0.93
73	Chernivetska	237	29.60 %	4.27 %	0.66	28.74 %	3.80 %	0.97	0.89
74	Chernigivska	366	28.62 %	4.34 %	0.65	29.22 %	3.94 %	1.02	0.91
80	Kyiv	494	9.63 %	1.63 %	0.93	9.56 %	1.66 %	0.99	1.02
85	Sevastopil	94	2.70 %	3.23 %	0.77	4.81 %	3.20 %	1.78	0.99

5.4. Performance measures

Table 7 shows some performance measures of the SAE Model, where three interesting measures are shown:

— the model parameter gamma (γ). It is the ratio between the model variance and the total variance, and is the share of the weight given to the direct survey estimate in the final composite estimate;

— the ratio between the EBLUP estimated value and the corresponding direct estimate. This is to check the extent to which the modelling changes the input direct estimates;

— the ratio between mean square error (MSE) of the EBLUP estimate of the Oblast, and the MSE of the direct survey estimate (which in this case is simply the variance, since the estimates are unbiased). This is to check the extent to which the modelling has improved the precision of the estimates.

As far as the weights given to direct estimate (gamma) are concerned, they are lower for those Oblasts with lower sub-sample sizes.

In these cases, the gain in terms of MSE can reach 20 % for Oblasts like Zhytomyrska, Sumska and Ternopilska. Moreover, the direct estimates for the City of Sevastopil (2.70 %), is considered too low value by any expert in poverty analysis. The final estimate (4.81 %) should be a much more unbiased value. Here, the gain in terms of MSE is not large, since the reduction in the original standard error, is compensated by the increase of the real MSE, which is obviously proportional to the magnitude of the estimated measure.

6. Concluding remarks

In the present paper, we have evaluated the impact of equivalence scales on poverty and inequality measures based on monetary variables from the Ukrainian HBS. A preliminary analysis of household expenditure behaviour in the HBS has led to the need for a scale which is not based only on the number of adults and children, but which also takes into account their age. In fact, the first original contribution of the paper can be considered the estimation of scales, where indices are different between adults and the elderly, and are different between children in different age groups.

The analysis of evaluating the effect of scales on poverty and inequality at national level has shown a big impact when choosing “extreme” scales (i.e. corresponding to non-equivalised total expenditure or to per-capita expenditure), but more limited effect for scales traditionally used, such as OECD, Engel-based, Academy of Science, etc… The new proposed scale has a much higher impact when analysing poverty profiles, i.e. when taking into account household composition and — indirectly — household expenditure pattern.

The choice of the scale also has a relevant impact when measuring poverty and inequality at regional (Oblast) level. However, for some regions, such an effect could be mixed-up with high sampling variability of the estimates.

Therefore, another original contribution of the paper consists in proposing a small area estimation technique (EBLUP) based on a ratio approach; in this way the effect of the difficult-to-qualify institutional and historical factors, common to the country and its regions, is abstracted.

The methodology applied to the poverty rate at Oblast level shows a significative reduction in standard errors up to 20 % for some Oblasts. This allows us to better identify the impact of equivalence scales on measures estimated at Oblast level.

Possible further research could consist in evaluating the effect of the choice of the scale on the net change of poverty over time (in the presence of repeated cross-sectional surveys such as HBS), and the effect on longitudinal and/or chronic poverty (in presence of panel surveys such as EU-SILC).

References

1. Banks J., Blundell R., Lewbel A. (1997). Quadratic Engel Curves and Consumer Demand. *The Review of Economics and Statistics,* 79, 527-539.

2. Barten A. P. (1964). Family Composition, Prices and Expenditure Patterns. Hart P. E., Mills F. and Whitaker J. K. (eds.). *Econometric Analysis for National Economic Planning.* London: Butterworths.

3. Betti G. (1999). Nonparametric equivalence scales with application to Poland. *Statistics Research Report* LSERR49, London School of Economics.

4. Betti G. (1999). A Non-parametric approach to equivalence scales. *Statistics in Transition,* 4(2), 181-194.

5. Betti G., Cheli B., Cambini R. (2004). A statistical model for the dynamics between two fuzzy states: theory and application to poverty analysis. *Metron,* 62(3), 391-411.

6. Betti G., D'Agostino A., Neri L. (2002). Panel Regression Models for Measuring Multidimensional Poverty Dynamics. *Statistical Methods and Applications,* 11(3), 359-369.

7. Betti G., Gagliardi F., Lemmi A., Verma V. (2012). Sub-national indicators of poverty and deprivation in Europe: methodology and applications. *Cambridge Journal of Regions, Economy and Society,* 5(1), 149-162.

8. Betti G. Lemmi A. (2013) (Eds.). *Poverty and Social Exclusion: New Methods of Analysis.* London and New York: Routledge.

9. Betti G., Lundgren L. (2012). The impact of remittances and equivalence scales on poverty in Tajikistan. *Central Asian Survey,* 31(4), 395–408.

10. Betti G., Verma V. (2008). Fuzzy measures of the incidence of relative poverty and deprivation: a multi-dimensional perspective. *Statistical Methods and Applications,* 17(2), 225-250.

11. Buhmann B., Rainwater L., Schmaus G., Smeeding T. (1988). Equivalent Scales, Well-Being, Inequality and Poverty: Sensitivity Estimates across Ten Countries Using the Luxembourg Income Study Database. *The Review of Income and Wealth,* 32, 115-142.

12. Cheli B., Betti G. (1999). Fuzzy analysis of poverty dynamics on an italian pseudo panel 1985–1994. *Metron,* 57(1-2), 85-105.

13. Deaton A. S., Muellbauer J. (1980). An Almost Ideal Demand System. *American Economic Review,* 70, 312-326.

14. Engel E. (1895). Die Lebenskosten Belgischer Arbeitfamilien früher und jetzt, *International Statistical Institute Bulletin,* 9, 1-74.

15. Gorman W. (1976). Tricks with Utility Functions. Artis M., Nobay A. (Eds.), *Essays in Economic Analysis,* Cambridge University Press, Cambridge.

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

17. Lewbel A. (1985), A Unifying Approach to Incorporating Demographic or Other Effects into Demand Analysis. *Review of Economic Studies,* 52, 1-18.

18. Molnar M., Betti G., Panduru F. (2003). Consumption pattern and equivalence scales in Romania. *Romanian Economic Review,* 12(2), 207-221.

19. Pollak R. A., Wales T. J. (1981). Demographic Variables in Demand Analysis. *Econometrica,* 49, 1533-1551.

20. Szulc A. (2003), Is it possible to estimate reliable equivalence scales? *Statistics in Transition,* 6(2), 589-611.

21. Van Ginneken W. (1982), Generating Internationally Comparable Income Distribution Data: Evidence from the Federal Republic of Germany (1974), Mexico (1968) and the United Kingdom (1979), *The Review of Income and Wealth,* 28, pp. 365-379.

22. Verma V., Betti G. (2011). Taylor linearization sampling errors and design effects for poverty measures and other complex statistics. *Journal of Applied Statistics,* 38(8) 1549-1576.

Information about athor

Betti Gianni (Siena, Italy) — PhD in Applied Statistics, Associate Professor in Statistics and Economics, University of Siena (P.zza San Francesco 7, 53100, Siena, Italy, e-mail: gianni.betti@unisi.it).