

INEQUALITY IN SELECTED SEE COUNTRIES:
A DECOMPOSITION ANALYSIS OF TRENDS AND DETERMINANTS

SEBASTIAN LEITNER AND ROBERT STEHRER

The Vienna Institute for International
Economic Studies (wiiw)
Oppolzergasse 6, A-1010 Vienna, Austria.
Phone ++43 1 533 66 10 - 47
Fax ++43 1 533 66 10 - 50
E-mail: Sebastian.Leitner@wiiw.ac.at
E-mail: Robert.Stehrer@wiiw.ac.at

Preliminary and unfinished version: January 21, 2009

Paper prepared as part of the Global Development Network (GDN) project "Inequality and Fiscal Policy in Transition Countries" financed by the Jubiläumsfonds of the Austrian National Bank (OENB), Project Nr. 13168 and Nr. 13171.

Abstract

In this paper we provide a comparative analysis of inequality in household consumption per capita in four Western Balkan countries, i.e. Albania, Bosnia & Herzegovina, Bulgaria, and Serbia. The analysis is based on a comparable dataset based on the World Bank's Living Standards Measurement Study (LSMS) which provide data for at least two years for each of the countries. We apply standard decomposition methods and inequality decomposition based on the Shapley approach. The results suggest that three groups of variables are particularly important for explaining patterns of inequality; these are age and sex, employment status and education.

JEL-Classification: C20; D63.

Keywords: inequality decomposition; Shapley value; Western Balkan countries;

INEQUALITY IN SELECTED SEE COUNTRIES: A DECOMPOSITION ANALYSIS OF TRENDS AND DETERMINANTS

SEBASTIAN LEITNER AND ROBERT STEHRER

1 Introduction

This paper is concerned with the trends and determinants of inequality in a sample of Western Balkan countries. To our knowledge this is the first attempt to provide a comparative analysis for this set of countries based on individual or household data.¹ In doing so we collected data from the World Bank's Living Standards Measurement Study (LSMS) which are described below in more detail and have a close look at the various components with respect to income recipients or consumer units by providing standard measures of inequality and respective decomposition methods. This latter aspect is further expanded by employing a recently researched method based on the Shapley value approach. The inequality measures are well documented and explained in a number of papers and books, notably Jenkins (1995) and Sen (1997). In more global context these methods are applied in Milanovic (2005); for an overview see Anand and Segal (2008). The datasource (LSMS) is well described and explained in detail in Deaton (1997) which also points towards all kind of potential problems and pitfalls. With respect to the analysis of inequality in transition economies the paper by Milanovic (1999) stands out. This however focus more on inequality by income sources rather than recipients.

The paper goes as follows: In Section 2 we present some stylized facts about inequality in the Western Balkan countries. Section 3 then provides a short description of the data we use in this paper. Further we point to several constraints and problems of data availability, quality of data, etc. In Section 4 we calculate some inequality measures using the household dataset and provide some descriptive evidence on income differences by household characteristics. In Section 5 we show the results of various decomposition analysis which allow us to make comparative statements across countries and to analyze the shifts in inequality over time. Section 6 concludes.

¹There are however individual country studies; e.g. INSTAT (2006) and Canova (2006) on Albania.

2 Inequality trends in selected See countries

Let us start with a brief overview of inequality trends in the countries under consideration. We briefly discuss some recent trends in inequality in the Western Balkan countries Albania, Bulgaria, Bosnia & Herzegovina, and Serbia (the countries we study in more detail below) by drawing on already existing databases. Tables 1 and 2 present available information on Gini coefficients from the WIDER database. The first table is based on household data with consumption, expenditures or disposable income per capita (i.e. no equivalence scales were used). The Gini coefficients on

Income/Expenditure definition	Code	Country	Year	Gini (1)	Gini (2)	Equivalence scale
Consumption	ALB	Albania	1996	29.3	28.0	Household per capita
			2002	28.1	28.2	Household per capita
			2004	31.1	31.1	Household per capita
Consumption	BIH	Bosnia & Herzegovina	2001	26.0	26.2	Household per capita
			2005	35.8	35.8	Household per capita
Income, Disposable	SRB	Serbia	2003	40.3	40.3	Household per capita
			2004	39.3	39.3	Household per capita
			2005	39.3	39.3	Household per capita
			2006	38.8	38.8	Household per capita

Notes: Gini (2) reports a Gini coefficient adjusted by WIDER.

Table 1: Gini coefficients from WIDER database: Households

consumption (expenditures) available for Albania, Bosnia & Herzegovina are between 26 and is about 36; further in all countries the Gini index was and was slightly increasing over time. For Serbia the Gini coefficient based on disposable income per capita is higher with about 40 in 2003 and even slightly declining little bit over time. Table 2 presents the Gini coefficient based on personal earnings data which is only available for Serbia. The coefficient was increasing slightly from 31 to a level of about 35 in 2006.

Income/Expenditure definition	Code	Country	Year	Gini (1)	Gini (2)	Equivalence scale
Earnings, Gross	SRB	Serbia	2002	31.1	31.1	No adjustment
			2003	32.0	32.0	No adjustment
			2004	33.0	33.0	No adjustment
			2005	31.3	31.3	No adjustment
			2006	34.7	34.7	No adjustment

Notes: Gini (2) reports a Gini coefficient adjusted by WIDER.

Table 2: Gini coefficients from WIDER database: Persons

3 Data

In this study we use data collected from the The World Bank’s Living Standards Measurement Study (LSMS). From this data source we tried to collect information which is available across countries which then allows to do cross-country comparisons. The countries, years and variables chosen are mainly dictated by the availability of comparable data.² Of course the reliance on LSMS data has some drawbacks as the surveys are not intended for cross country comparisons. Furthermore, the structure of the available data sets implies that the data have to be collected separately for each country and the years covered differ across countries. Table 3 provides an overview of the data collected so far. We have been successful to establish a more or less compa-

Country	Years	
Albania	2002, 2005	Consumption
Bulgaria	1995, (1997), 2001	Consumption, Income
Bosnia & Herzegovina	2001, 2004	Consumption
Serbia	2002, 2007	Consumption, Income

Table 3: Availability of LSMS data

able database for four Western Balkan countries, i.e. Albania, Bulgaria, Bosnia & Herzegovina and Serbia. As one can see the years for which the LSMS was conducted differs across countries. Whereas for Bulgaria the survey data are available for 1995 and 2001, these are available more recent years in the other countries, namely for 2001 or 2002 and 2004, 2005, and 2007 respectively. For some countries - Bulgaria (1995 and 1997) and Albania (2002,2003, and 2004) - a panel structure might be exploited which would allow to control for unobserved heterogeneity in regression analysis. however, as we wanted to compare a broader set of countries we have not yet used this possibility.

With respect to the income variable we decided to use household consumption rather than income data which seems to be more reliable. The drawback on using household consumption data is that we can mainly employ only information about the characteristics of the household head. For this we tried to collect the information by sex ('Male' and 'Female'), by age groups (<25, 25-34, 35-44, 45-54, 55-64, 65-74, >=75), by education (None, low, medium, uppers secondary and tertiary), by ethnical groups for the respective countries, regional classification and

²The usage of the LSMS data was mainly driven by not having access to other household survey data in the respective countries.

employment status of the household head (employed, employer, unemployed, retired and others). Unfortunately, these information is not available for all countries and years. Educational categories, regional groups and employment status is missing for Bosnia & Herzegovina in 2004, and information on ethnicity is missing for Serbia in 2002.

From these data we calculated household consumption per capita per month (i.e. dividing the total consumption with the number of persons; further we also use a second measure by applying an equivalence scale, total consumption divided by the square root of the number of persons, i.e. we assume an equivalence parameter of $\rho = 0.5$ which is commonly used in the literature. Further, for being able to compare data across countries we converted household consumption using purchasing power parity rates (PPPs) for the respective years. One should note that this transformation does not change the inequality measures provided below.

4 Summary statistics and inequality measures

In this section we present some summary statistics for each of the countries under consideration based on the collected data. Before presenting the results we give a summary of the inequality measures used in the paper.

4.1 Inequality measures

There are a number of inequality measures available from the literature (see Sen, 1997, for a detailed overview on the properties of these measures). For a first assessment we use the most common ones, the mean logarithmic deviation (MLD or I_0), the Theil measure (I_1), and the half squared coefficient of variation (I_2); formally,

$$\begin{aligned} I_0 &= \frac{1}{N} \sum_{i=1}^N \ln \frac{\mu}{y_i} \\ I_1 &= \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\mu} \ln \frac{y_i}{\mu} \\ I_2 &= \frac{1}{2} \cdot \frac{1}{N} \sum_{i=1}^N \left[\left(\frac{\mu}{y_i} \right)^2 - 1 \right] = \frac{1}{2} \cdot \frac{\sigma^2}{\mu^2}. \end{aligned}$$

Here, N denotes the number of units (individuals, households), y_i is the income or expenditure category, $\mu = \frac{1}{N} \sum_{i=1}^N y_i$ denotes the mean income or expenditure, and σ denotes the variance.

These are special cases of the generalized class of entropy measures, defined as

$$I_\alpha = \frac{1}{\alpha(1-\alpha)} \frac{1}{N} \sum_{i=1}^N \left[1 - \left(\frac{y_i}{\mu} \right)^\alpha \right] \quad \text{for } \alpha \neq 0, 1.$$

For $\alpha = 0$ this becomes Theil's 'second' measure (the mean logarithmic deviation or I_0) and for $\alpha = 1$ the Theil measure I_1 . These measures are mainly used as they possess favorable decomposition probabilities (see e.g. Sen, 1997, for a summary and details).

Further we present the Gini index (G) given by

$$G = \frac{N+1}{N-1} - \frac{2}{N(N-1)\mu} \sum_{i=1}^N \rho_i y_i.$$

In this formula ρ_i denotes the rank of unit i in the distribution; the counting is from the top such that the richest unit has rank 1. The formula is equivalent to the more commonly known expression

$$G = \frac{1}{\mu N(N-1)} \sum_{i>j} \sum_j |y_i - y_j|,$$

which however is computationally less convenient.

4.2 Results

Table 4 presents the mean of household consumption per capita and the respective inequality measures for the data not applying and applying the equivalence scale. The mean income is increasing by in between 50 and 100 % when applying the equivalence scale of 0.5. The variation of mean income across countries tends to increase in this case mainly reflecting cross-country differences in household size. In line with the results presented above, most inequality indicators show an increase in inequality over time; the exception to this is again Serbia where inequality decreased slightly. Comparing the inequality measures reveals a clear ranking with the measure I_0 being lowest to the Gini coefficient being higher than the other measures; this reflects the properties of the inequality measures (see e.g. Sen, 1997). Finally, when adjusting consumption for household size inequality tends to be lower for all countries and years (the only exception is Bulgaria where inequality measures show a higher value in 1995 when doing the adjustment).

In the Appendix we provide further detailed results with respect to the characteristics of households.

Income/Expenditure definition	Country	Year	Mean	I_0	I_1	I_2	Gini
No equivalence scale ($\rho = 1$)							
Household consumption per capita	Albania	2002	158.7	0.137	0.140	0.166	0.290
		2005	177.5	0.158	0.166	0.213	0.311
Household consumption per capita	Bulgaria	1995	203.6	0.126	0.128	0.150	0.272
		2001	245.7	0.152	0.156	0.215	0.295
Household consumption per capita	Bosnia & H.	2001	457.9	0.140	0.145	0.181	0.291
		2004		0.131	0.134	0.159	
Household consumption per capita	Serbia	2002	340.1	0.171	0.186	0.265	0.321
		2007	440.6	0.161	0.178	0.242	0.312
Equivalence scale ($\rho = 0.5$)							
Household consumption per capita	Albania	2002	302.6	0.100	0.103	0.118	0.248
		2005	349.2	0.130	0.140	0.184	0.282
Household consumption per capita	Bulgaria	1995	321.9	0.138	0.136	0.156	0.285
		2001	389.7	0.149	0.148	0.178	0.295
Household consumption per capita	Bosnia & H.	2001	730.1	0.102	0.102	0.115	0.248
		2004		0.095	0.096	0.108	
Household consumption per capita	Serbia	2002	554.3	0.171	0.176	0.223	0.320
		2007	717.5	0.158	0.163	0.204	0.309

Table 4: Inequality measures

5 Determinants of inequality

In this section we present the results of various decomposition methods to explain the sources of income inequality by the characteristics of households (income recipients approach). We first use decomposition methods for the inequality measures in the class of generalized entropy measures which allow a decomposition in within and between components. This is followed by a regression based decomposition method based on the Shapley value. We start with a short methodological outline and then present the results.

5.1 Decomposition analysis

5.1.1 Methodological issues

DECOMPOSITION OF MLD

$$I_0 = \frac{1}{N} \sum_{i=1}^N \ln \frac{\mu}{y_i}$$

Sub-group decomposition:

$$I_0 = \sum_k v_k I_{0,k} + \sum_k v_k \ln(1/\lambda_k)$$

where v_k denotes the population share, and $\lambda_k = \mu_k/\mu$.

Change between two years (see Mookherjee and Shorrocks, 1982; Jenkins, 1995):

$$\begin{aligned} \Delta I_0 &= \sum_k \bar{v}_k \Delta I_{0,k} + \sum_k \bar{I}_{0,k} \Delta v_k - \sum_k [\overline{\ln(\lambda_k)}] \Delta v_k - \sum_k \bar{v}_k \Delta \ln(\lambda_k) \\ &\approx \underbrace{\sum_k \bar{v}_k \Delta I_{0,k}}_{\text{Term A}} + \underbrace{\sum_k \bar{I}_{0,k} \Delta v_k}_{\text{Term B}} - \underbrace{\sum_k [\bar{\lambda}_k - \overline{\ln(\lambda_k)}] \Delta v_k}_{\text{Term C}} + \underbrace{\sum_k (\bar{\theta}_K - \bar{v}_k) \Delta \ln(\mu_k)}_{\text{Term D}} \end{aligned}$$

where $\theta_k = v_k \lambda_k$ denotes group k 's share of total population income.

DECOMPOSITION OF I_1 (based on Giammatteo (2007))

$$\begin{aligned} I_1 &= \frac{1}{N} \sum_{i=1}^N \frac{y_i}{\mu} \ln \frac{y_i}{\mu} \\ I_1 &= \sum_{k=1}^K \frac{N_k}{N} \frac{\mu_k}{\mu} \ln \frac{\mu_k}{\mu} + \sum_{k=1}^K \frac{N_k}{N} \frac{\mu_k}{\mu} \left(\frac{1}{N_k} \sum_{i=1}^{N_k} \frac{y_{i,k}}{\mu_k} \ln \frac{y_{i,k}}{\mu_k} \right) \\ &= \underbrace{\sum_{k=1}^K \pi_k s_k \ln s_k}_{\text{Between}} + \underbrace{\sum_{k=1}^K \pi_k s_k I_{1,k}}_{\text{Within}} \end{aligned}$$

5.1.2 Results

// ... TO BE INCLUDED ... //

5.2 The Shapley value approach to decomposition

Recently the literature has provided some additional decomposition methods which is based on regression results and the Shapley value approach as introduced by Shorrocks (1999) in the inequality literature; see Fields and Yoo (2000), Morduch and Sicular (2002), Fields (2003), Wan (2004), and Gunatilaka and Chotikapanich (2006) for such applications. Finally, Israeli (2007) shows how the Shapley-approach is related to the method proposed by Fields (2003) and also points to some advantages of the former which is applied in this paper as well. The most important advantage of the Shapley value approach is that this take the potential correlation amongst regressors into account. One should however note that the latter two contributions aim at decomposing the R^2 of the regressions whereas in this paper we decompose the resulting inequality measure. This is more similar to the contributions by Wan (2004) Gunatilaka and Chotikapanich (2006).

5.2.1 Methodology

We explain the Shapley value approach using a simple example with three explanatory variable. We first regress consumption levels y_i on these explanatory variables x_i ($i = 1, 2, 3$),

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

where ε denotes the error term. The predicted consumption level is then given by

$$\hat{y}_{123} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \hat{\beta}_3 x_3.$$

This predicted value is then used to calculate the Gini coefficient $\hat{G}_{\{123\}}^{(0)}$, where subscripts denote the variables included. In the first round we then eliminate one variable and calculate the predicted consumption levels $\hat{y}_{\{23\}}$, $\hat{y}_{\{13\}}$ and $\hat{y}_{\{12\}}$. The corresponding Gini coefficients are then given by $\hat{G}_{\{23\}}^{(1)}$, $\hat{G}_{\{13\}}^{(1)}$ and $\hat{G}_{\{12\}}^{(1)}$ respectively. Analogously, in the second round we eliminate two variables, thus calculating $\hat{y}_{\{1\}}$, $\hat{y}_{\{2\}}$ and $\hat{y}_{\{3\}}$. The resulting Gini coefficients are $\hat{G}_{\{1\}}^{(2)}$, $\hat{G}_{\{2\}}^{(2)}$ and $\hat{G}_{\{3\}}^{(2)}$. The final round would then be to only include the constant; the resulting Gini coefficient would thus be $\hat{G}_{\{\}}^{(3)} = 0$.

The marginal contributions are then calculated using the Gini coefficients. The first round marginal contribution of each variable are $C_1^{(1)} = \hat{G}_{\{123\}}^{(0)} - \hat{G}_{\{23\}}^{(1)}$, $C_2^{(1)} = \hat{G}_{\{123\}}^{(0)} - \hat{G}_{\{13\}}^{(1)}$ and $C_3^{(1)} = \hat{G}_{\{123\}}^{(0)} - \hat{G}_{\{12\}}^{(1)}$. The marginal contributions in the second round of the first variable are given by

$$C_1^{(2,1)} = \hat{G}_{\{12\}}^{(1)} - \hat{G}_{\{2\}}^{(2)} \quad \text{and} \quad C_1^{(2,2)} = \hat{G}_{\{13\}}^{(1)} - \hat{G}_{\{3\}}^{(2)}.$$

The average of these contributions is then marginal contribution of the first variable in the second round, i.e. $C_1^{(2)} = \frac{1}{2}(C_1^{(2,1)} + C_1^{(2,2)})$. Similarly, we calculate $C_2^{(2)}$ and $C_3^{(2)}$. The third round contribution then is given by $C_1^{(3)} = \hat{G}_{\{1\}}^{(2)} - \hat{G}_{\{\}}^{(3)} = \hat{G}_{\{1\}}^{(2)}$ as $\hat{G}_{\{\}}^{(3)} = 0$ and analogously for $C_2^{(3)} = \hat{G}_{\{2\}}^{(2)}$ and $C_3^{(3)} = \hat{G}_{\{3\}}^{(2)}$.

Finally, by averaging the marginal contributions of each variables over all rounds results in the total marginal effect of each variable $j = 1, 2, 3$, i.e.

$$C_j = \frac{1}{3} \cdot (C_j^{(1)} + C_j^{(2)} + C_j^{(3)}).$$

The proportion of inequality not explained is then given by

$$C_R = G - \hat{G}_{\{123\}}^{(0)}.$$

The approach can be easily extended to any number of explanatory factors and to other inequality measures as we will show below.

5.2.2 Results of Shapley value decomposition

In the next four tables we present the results of the Shapley value decomposition. The corresponding regression results are shown in the Appendix. For each country we present the results for three measures of inequality - Gini, I_0 and I_2 - and for household consumption per capita and equivalent household income. The results of the individual regressions are shown in the Appendix. Tables 5 to 8 present the results of the decomposition analysis.

These tables present the contributions to the inequality measures under consideration - Gini, I_0 and I_2 - in absolute term and in percent of inequality. Before describing the results in more detail let us note some important properties: First, the contributions in absolute values sum up to the overall inequality measure reported in Table 4 due to the decomposability property of the Shapley value. Second, regarding results for the three inequality measures one can see that the unexplained part of inequality is much larger for the generalized entropy measures I_0 and I_2

Variable	Gini				I0				I2			
	2002		2005		2002		2005		2002		2005	
	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %
No equivalence scale												
Age and sex	0.022	7.5	0.056	18.0	0.003	2.2	0.017	11.1	0.003	1.7	0.019	8.8
Employment status	0.022	7.6	0.013	4.0	0.003	2.4	0.000	0.0	0.003	1.7	0.000	0.0
Education	0.035	12.0	0.048	15.3	0.007	4.9	0.012	7.7	0.007	4.4	0.013	6.2
Ethnicity	0.045	15.4	0.004	1.4	0.015	11.1	0.002	1.1	0.017	10.3	0.002	0.9
Region	0.012	4.2	0.049	15.6	0.001	0.7	0.014	9.1	0.000	0.2	0.015	6.9
Unexplained	0.154	53.3	0.142	45.7	0.108	78.8	0.112	70.9	0.135	81.6	0.164	77.0
Total	0.290		0.311		0.137		0.158		0.166		0.213	
Equivalence scale ($\rho = 0.5$)												
Age and sex	-0.016	-6.3	-0.013	-4.6	-0.008	-8.0	-0.007	-5.7	-0.009	-7.7	-0.008	-4.6
Employment status	0.034	13.6	0.038	13.4	0.006	6.2	0.009	7.1	0.007	5.6	0.012	6.5
Education	0.056	22.5	0.068	24.2	0.013	12.8	0.018	13.9	0.014	12.0	0.020	10.8
Ethnicity	0.024	9.5	0.002	0.7	0.007	7.4	0.001	0.5	0.008	7.1	0.001	0.5
Region	0.014	5.6	0.047	16.8	0.001	1.3	0.012	8.9	0.001	0.9	0.012	6.7
Unexplained	0.136	55.1	0.140	49.6	0.080	80.3	0.098	75.3	0.097	82.1	0.147	80.1
Total	0.248		0.282		0.100		0.130		0.118		0.184	

Table 5: Shapley value decomposition: Albania

Variable	Gini				I0				I2			
	2001		2004		2001		2004		2001		2004	
	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %
No equivalence scale												
Age and sex	0.066	22.5	n.a.	n.a.	0.017	12.1	n.a.	n.a.	0.018	9.7	n.a.	n.a.
Employment status	0.039	13.6	n.a.	n.a.	0.009	6.7	n.a.	n.a.	0.009	5.1	n.a.	n.a.
Education	0.028	9.8	n.a.	n.a.	0.005	3.4	n.a.	n.a.	0.005	2.7	n.a.	n.a.
Ethnicity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Region	0.011	3.9	n.a.	n.a.	0.002	1.8	n.a.	n.a.	0.002	1.3	n.a.	n.a.
Unexplained	0.146	50.2	n.a.	n.a.	0.106	76.0	n.a.	n.a.	0.147	81.3	n.a.	n.a.
Total	0.291		n.a.	n.a.	0.140		n.a.	n.a.	0.181		n.a.	n.a.
Equivalence scale ($\rho = 0.5$)												
Age and sex	0.004	1.5	n.a.	n.a.	-0.001	-1.0	n.a.	n.a.	-0.001	-1.2	n.a.	n.a.
Employment status	0.055	22.0	n.a.	n.a.	0.011	11.2	n.a.	n.a.	0.012	10.6	n.a.	n.a.
Education	0.061	24.5	n.a.	n.a.	0.014	13.8	n.a.	n.a.	0.015	13.3	n.a.	n.a.
Ethnicity	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Region	0.009	3.6	n.a.	n.a.	0.001	1.3	n.a.	n.a.	0.001	1.1	n.a.	n.a.
Unexplained	0.120	48.3	n.a.	n.a.	0.076	74.7	n.a.	n.a.	0.088	76.3	n.a.	n.a.
Total	0.248		n.a.	n.a.	0.102		n.a.	n.a.	0.115		n.a.	n.a.

Table 6: Shapley value decomposition: Bosnia & Herzegovina

Variable	Gini				I0				I2			
	1995		2001		1995		2001		1995		2001	
	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %
No equivalence scale												
Age and sex	0.005	1.9	0.081	27.6	-0.002	-1.4	0.024	15.8	-0.002	-1.2	0.025	11.4
Employment status	0.076	27.8	0.055	18.5	0.019	15.1	0.016	10.3	0.019	12.7	0.017	7.8
Education	0.038	13.9	0.027	9.0	0.008	6.1	0.006	3.7	0.007	4.8	0.005	2.2
Ethnicity	0.000	0.0	-0.004	-1.4	-0.001	-0.8	-0.003	-1.7	-0.002	-1.2	-0.005	-2.4
Region	0.003	1.0	0.001	0.2	0.000	0.1	0.000	0.1	0.000	0.1	0.000	0.0
Unexplained	0.151	55.3	0.135	46.0	0.102	81.0	0.109	71.8	0.127	84.8	0.174	81.0
Total	0.272		0.295		0.126		0.152		0.150		0.215	
Equivalence scale ($\rho = 0.5$)												
Age and sex	-0.010	-3.6	0.085	28.8	-0.015	-11.0	0.030	19.8	-0.017	-10.7	0.026	14.7
Employment status	0.111	38.8	0.066	22.3	0.049	35.4	0.020	13.3	0.046	29.3	0.022	12.3
Education	0.049	17.2	0.028	9.4	0.015	11.2	0.006	4.3	0.013	8.5	0.005	2.9
Ethnicity	-0.011	-3.8	-0.008	-2.6	-0.012	-8.6	-0.008	-5.6	-0.010	-6.5	-0.007	-4.2
Region	0.012	4.3	0.000	0.1	0.004	2.8	0.000	0.0	0.003	1.8	0.000	0.0
Unexplained	0.134	47.1	0.124	42.0	0.097	70.2	0.102	68.2	0.121	77.6	0.132	74.2
Total	0.285		0.295		0.138		0.149		0.156		0.178	

Table 7: Shapley value decomposition: Bulgaria

Variable	Gini				I0				I2			
	2002		2007		2002		2007		2002		2007	
	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %	Total	in %
No equivalence scale												
Age and sex	0.071	22.0	0.077	24.6	0.021	12.1	0.026	15.9	0.020	7.6	0.025	10.4
Employment status	0.010	3.3	0.005	1.5	0.000	-0.1	-0.003	-2.0	0.000	0.0	-0.003	-1.2
Education	0.074	23.1	0.088	28.2	0.021	12.5	0.028	17.4	0.021	7.9	0.028	11.4
Ethnicity	n.a.	n.a.	-0.002	-0.6	n.a.	n.a.	-0.001	-0.8	n.a.	n.a.	-0.001	-0.5
Region	0.004	1.1	0.006	2.0	-0.001	-0.8	0.000	-0.3	-0.002	-0.6	-0.001	-0.2
Unexplained	0.162	50.5	0.138	44.3	0.130	76.3	0.112	69.8	0.225	85.1	0.194	80.1
Total	0.321		0.312		0.171		0.161		0.265		0.242	
Equivalence scale ($\rho = 0.5$)												
Age and sex	0.062	19.5	0.063	20.3	0.022	12.8	0.024	14.9	0.019	8.4	0.021	10.2
Employment status	0.014	4.4	0.005	1.8	0.001	0.4	-0.003	-2.1	0.001	0.4	-0.003	-1.4
Education	0.089	27.9	0.104	33.6	0.029	17.2	0.035	22.0	0.027	12.0	0.033	16.1
Ethnicity	n.a.	n.a.	-0.004	-1.3	n.a.	n.a.	-0.003	-2.0	n.a.	n.a.	-0.003	-1.3
Region	0.003	1.0	0.009	2.8	-0.003	-2.0	0.000	-0.1	-0.003	-1.2	0.000	0.1
Unexplained	0.151	47.1	0.132	42.9	0.123	71.6	0.106	67.3	0.179	80.4	0.155	76.3
Total	0.320		0.309		0.171		0.158		0.223		0.204	

Table 8: Shapley value decomposition: Serbia

compared to the Gini coefficient. For the latter the unexplained part is around 50 %. This is remarkable, first, as the R^2 reported in the regressions ranges between 20 and 30 % in most cases. Second, the explanation for this is likely to be that the OLS regression (on the untransformed data) does not very well predict the values at the tails of the distribution (which is more relevant for the entropy measures).³ For these reasons we thus focus on the results based on the Gini coefficient. Third, the contributions to inequality can seemingly be negative.⁴

For *Albania* education and ethnicity explains the largest part of inequality in 2002. However, there is a remarkable shift towards inequality explained mainly by age and sex, region and education in 2005.⁵ The results when applying equivalence scaling give much more explanatory power to the employment status in both years; education is however the most important source of inequality. The regional aspect again becomes important in 2005 only. For *Bosnia & Herzegovina* we only provide results for 2001 for reasons outlined above. In the case of no equivalence scale age and sex together with employment status are the most important sources of inequality with education having a less important role. However, for consumption data with equivalence scaling employment status and education make up the largest part in the explanation. In *Bulgaria* again employment status and education are the most important variables in 1995; however, in 2001 it is age and sex together with employment status. The results are similar when using the scaled consumption data. Finally, for *Serbia* in both cases and both years age and sex and education are the most important determinants of inequality.

Summarizing, the results suggest a dominant role for basically three variables: age and sex, employment status and education though in most cases only two of these account for almost half of the explained inequality. This is further not stable over the years for which data are available and also differ when applying an equivalence scale. In a technical sense this could point towards either some data problems (e.g. outliers) or correlation amongst the explanatory variables which has to be checked in more detail. On the other hand, ethnicity and regional aspects only play a minor role in explaining inequality from these results.

³We are still working on a proper transformation of the data by using logarithms for example. Wan (2004) suggested to transform the data using Box-Cox or Box-Tidwell transformations. Another route would be to use quantile estimation which is more robust to outliers.

⁴This point will have to be checked in more detail.

⁵The summary statistics in the Appendix show a large increase in one region which might also point towards a data problem. Further, the number of for other ethnical groups are quite small.

6 Conclusions

In this paper we provided some evidence on the determinants of inequality with respect to consumption by households for SEE countries: Albania, Bulgaria, Bosnia & Herzegovina, and Serbia. For this we used data collected from the LSMS database provided by the World Bank. We calculated various inequality measures and then used decomposition methods to consider inequality patterns by household characteristics. The results from the Shapley value approach suggest that mainly three groups of variables - age and sex, employment status and education - are important determinants. Somewhat surprisingly regional and ethnical aspects play a minor role in most cases.

Future research will have to proceed along two main lines: The first is to improve the underlying regression results applying sensible data transformations and maybe other estimation strategies together with a further improvement of the dataset (mainly with respect to potential outliers). Second, these results shall be compared with other methods focusing on the explanation of the R^2 from the regression as suggested by Fields (2003) and Israeli (2007). Additionally, other decomposition methods like the Owen and the Nested-Shapley will have to be considered.

A Summary tables

Category	Group	2002						2005					
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum		
1	All	3599	158.7	137.3	27.4	1219.1	3840	177.5	149.5	29.2	1878.0		
1	Sex	3140	156.1	134.6	27.4	1219.1	3445	174.2	147.7	29.2	1878.0		
2	Female	459	176.4	155.8	35.7	704.3	395	206.0	170.8	34.1	774.4		
1	Age	37	149.4	134.0	27.7	333.4	16	185.0	143.1	70.7	475.4		
2	25-34	385	141.7	124.4	38.3	704.3	303	173.9	133.5	41.4	1182.1		
3	35-44	944	138.8	119.4	30.5	585.9	927	168.7	140.0	30.1	1878.0		
4	45-54	865	165.8	139.7	30.2	1219.1	1117	181.9	154.3	41.4	772.8		
5	55-64	675	180.0	155.0	27.4	834.1	737	180.8	154.6	29.2	1013.1		
6	65-74	501	166.4	154.2	33.2	661.7	517	183.8	156.4	30.8	686.8		
7	=>75	192	165.5	150.2	42.3	536.8	223	169.7	146.4	31.9	695.8		
1	Education	258	147.2	130.3	33.2	622.1	254	149.5	122.4	30.1	649.2		
2	None	518	149.6	130.0	27.4	149.6	502	144.8	125.8	29.2	567.2		
3	Low education	1373	137.3	118.9	27.7	834.1	1467	148.6	128.9	32.2	747.4		
4	Medium education	1047	167.8	149.0	33.0	1219.1	1231	198.4	170.7	36.5	1878.0		
5	Upper Secondary	403	227.0	200.0	45.6	830.6	386	279.9	239.1	46.5	1324.5		
1	University & higher	3546	157.5	136.2	27.4	1219.1	3711	178.6	150.7	29.2	1878.0		
2	Albanian	39	264.3	249.0	75.7	834.1	31	193.6	160.2	53.9	951.1		
3	Greek	4	109.6	104.8	79.3	149.5	12	111.5	101.4	35.0	346.4		
4	Roma	1	283.5	283.5	283.5	283.5	10	150.8	138.4	77.6	242.1		
5	Macedonian	6	212.0	200.6	132.4	316.6	1	206.1	206.1	206.1	206.1		
6	Montenegrin	3	114.7	134.1	42.5	167.7	8	155.1	117.7	66.0	518.7		
7	Vllahë	776	182.4	153.6	27.4	830.6	999	183.1	155.2	30.8	1878.0		
1	Region1	968	168.1	147.8	30.5	834.1	999	170.4	147.1	30.1	1006.4		
2	Tirana	1143	150.3	129.8	33.2	1219.1	1000	143.4	125.8	29.2	747.4		
3	Coastal	712	133.5	118.2	27.7	478.2	640	233.1	204.5	46.1	1324.5		
4	Central	1959	173.3	151.1	27.4	1219.1	1999	202.3	172.6	30.1	1878.0		
1	Region2	1640	141.3	120.8	30.2	834.1	1639	147.3	125.7	29.2	1013.1		
2	Urban	1075	166.1	144.7	33.0	704.3	1996	190.8	160.1	30.1	1878.0		
1	Rural	17	302.6	283.3	195.0	503.6	52	362.3	290.9	131.5	1324.5		
2	Employment	260	126.4	110.7	33.1	740.3	157	155.2	131.2	35.0	969.5		
3	Employer	807	169.7	150.0	27.4	834.1	918	176.0	150.7	30.8	774.4		
4	Unemployed	1440	151.1	128.8	27.7	1219.1	717	136.1	117.3	29.2	603.0		
5	Retired												
6	Others												

Table 9: Summary statistics I: Albania (consumption per capita)

Category	Group	2002					2005				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	3599	302.6	271.7	39.0	1493.4	3840	349.2	299.7	67.3	4199.4
1	Sex	3140	305.7	274.4	61.1	1493.4	3445	353.5	303.8	67.3	4199.4
2	Female	459	281.5	254.1	39.0	1194.0	395	311.6	267.6	74.2	1206.6
1	Age	37	250.3	254.9	64.5	472.1	16	316.4	296.2	158.1	672.3
2	25-34	385	273.1	248.0	76.7	932.4	303	328.1	272.6	101.4	1396.5
3	35-44	944	289.7	260.0	39.0	1310.1	927	355.5	302.6	67.3	4199.4
4	45-54	865	322.9	284.1	67.7	1493.4	1117	375.7	327.0	74.2	2363.4
5	55-64	675	322.4	292.9	72.6	1288.1	737	343.3	295.8	75.8	1679.1
6	65-74	501	295.5	275.4	87.8	868.4	517	325.5	292.8	82.9	1189.6
7	=>75	192	292.9	254.5	70.6	1459.0	223	294.6	257.6	73.2	1206.6
1	Education	258	262.6	226.7	70.6	1459.0	254	267.7	226.5	67.3	1206.6
2	None	518	277.3	261.1	72.6	866.9	502	278.8	253.8	75.8	896.1
3	Low education	1373	269.3	248.1	39.0	1288.1	1467	305.0	269.7	73.2	2363.4
4	Medium education	1047	323.5	293.4	84.6	1493.4	1231	391.7	346.8	81.7	4199.4
5	Upper Secondary	403	419.8	389.5	111.6	1438.6	386	524.5	461.4	74.2	2649.1
1	Ethnic	3546	301.5	271.1	39.0	1493.4	3711	351.2	302.6	67.3	4199.4
2	Albanian	39	402.1	417.8	131.1	834.1	31	384.5	310.5	129.9	1647.4
3	Greek	4	244.1	215.2	158.6	387.1	12	215.1	201.0	85.7	489.9
4	Roma	1	567.1	567.1	567.1	567.1	10	284.2	265.7	190.2	454.7
5	Macedonian	6	327.9	345.8	187.2	447.8	11	504.9	504.9	504.9	504.9
6	Montenegrin	3	225.0	290.4	85.0	299.7	8	311.4	231.2	129.2	1037.4
7	Vllahë	776	333.6	292.6	72.6	1438.6	999	360.4	311.2	73.8	4199.4
1	Region1	968	317.6	287.0	39.0	1493.4	999	334.9	292.2	67.3	1821.6
2	Tirana	1143	286.5	259.8	67.7	1459.0	1000	299.7	272.5	83.7	2363.4
3	Coastal	712	274.2	255.6	84.1	888.6	640	431.2	374.5	73.2	2649.1
4	Central	1959	320.4	283.8	70.6	1493.4	1999	383.9	328.7	67.3	4199.4
1	Region2	1640	281.4	258.8	39.0	1459.0	1639	306.8	270.3	73.8	2363.4
2	Rural	1075	325.1	296.5	87.2	1094.1	1996	386.1	333.7	67.3	4199.4
1	Employment	17	590.3	573.5	398.0	872.2	52	749.6	598.6	204.6	2649.1
2	Employed	260	243.8	216.5	84.1	757.1	157	307.7	268.1	83.7	1679.1
3	Unemployed	807	296.3	271.0	70.6	1459.0	918	314.2	280.5	73.2	1206.6
4	Retired	1440	296.5	265.3	39.0	1493.4	717	275.8	250.8	74.2	881.5
5	Others										

Table 10: Summary statistics: Albania (equivalent consumption per capita)

Category	Group	2001					2004				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	5402	457.9	396.9	46.5	4470.9	2945	72.2	62.2	9.5	412.4
1	Sex	4056	430.6	375.5	46.5	4470.9	2035	67.3	57.9	9.5	382.8
2	Female	1346	540.3	468.0	79.5	3220.5	743	86.0	79.8	14.7	412.4
1	Age	34	712.4	665.0	150.3	1793.0	37	115.8	96.6	30.2	310.2
2	25-34	393	434.0	352.8	46.5	2300.5	214	75.7	58.3	16.7	412.4
3	35-44	1036	398.0	350.1	91.2	2825.4	485	63.1	52.8	16.3	336.0
4	45-54	1333	458.5	398.0	48.6	4470.9	693	72.0	63.6	9.5	354.5
5	55-64	1080	492.7	424.0	52.1	3097.4	527	76.3	67.2	11.6	382.8
6	65-74	1136	476.6	424.3	79.5	3389.4	553	73.0	65.1	14.3	215.7
7	>=75	390	466.4	414.8	83.8	1904.0	436	71.0	61.8	17.5	300.8
1	Education	1268	409.4	365.5	52.1	1995.8					
2	Low education	957	397.0	347.7	76.7	2058.7					
3	Medium education	2421	458.4	399.4	46.5	2362.9					
4	Upper Secondary	387	537.2	470.5	98.2	2645.4					
5	University & higher	369	696.7	590.2	178.9	4470.9					
11	Region1	1704	424.1	384.9	46.5	1645.9					
12	Banja Luka	240	461.8	400.2	90.4	1967.0					
14	Doboj	252	425.8	374.1	121.0	1758.0					
15	Vlasenica	360	591.9	513.7	148.3	2074.0					
16	Sarajevo-Romanija	120	294.8	266.6	95.1	886.1					
21	Foca	96	583.9	535.8	142.9	2300.5					
23	Una-Sana	672	469.9	416.7	76.7	2645.4					
24	Tuzla	974	365.9	318.0	48.6	3389.4					
26	Zenica-Doboj	192	480.8	434.9	121.2	1995.8					
29	Central Bosnia	792	585.8	505.5	112.1	4470.9					
1	Employment	2019	454.5	389.3	48.6	4470.9					
2	Employer	268	522.1	451.1	52.1	3389.4					
3	Unemployed	745	363.2	312.2	46.5	2825.4					
4	Retired	1962	479.3	423.3	84.5	2058.7					
5	Others	408	502.9	429.6	79.5	2309.2					

Table 11: Summary statistics: Bosnia & Herzegovina

Category	Group	2001					2004				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	5402	730.1	667.7	84.1	4793.3	2945	111.7	101.5	21.1	541.4
1	Sex										
	Male	4056	739.9	673.2	84.1	4793.3	2035	113.9	102.8	21.1	541.4
2	Female	1346	700.4	641.8	112.4	3761.2	743	106.2	96.6	23.2	412.4
1	Age										
	<25	34	884.2	827.5	232.2	1793.0	37	141.7	125.2	61.1	327.6
2	25-34	393	749.4	676.9	93.1	3761.2	214	118.0	105.5	34.2	412.4
3	35-44	1036	734.4	689.7	159.7	3458.1	485	112.5	103.3	33.7	378.1
4	45-54	1333	803.4	742.6	84.1	4470.9	693	125.7	115.4	21.1	419.9
5	55-64	1080	753.1	691.7	114.1	3097.4	527	114.9	106.0	28.5	541.4
6	65-74	1136	652.5	607.9	112.4	4793.3	553	97.9	91.0	24.8	325.5
7	>=75	390	597.1	537.8	123.2	1904.0	436	88.6	82.7	23.2	300.8
1	Education										
	None	1268	575.2	529.5	112.4	2822.5					
2	Low education	957	642.2	599.1	125.6	2058.7					
3	Medium education	2421	764.6	710.3	84.1	2897.6					
4	Upper Secondary	387	868.9	814.2	195.9	3761.2					
5	University & higher	369	1117.5	1000.1	283.1	4793.3					
11	Region1										
	Banja Luka	1704	700.5	659.5	93.1	2520.4					
12	Doboj	240	695.4	664.7	221.5	2296.7					
14	Vlasenica	252	704.7	646.4	203.5	2376.4					
15	Sarajevo-Romanija	360	882.9	823.1	209.7	2524.3					
16	Foca	120	475.4	434.8	160.3	1062.0					
21	Una-Sana	96	1007.4	1025.5	229.8	2300.5					
23	Tuzla	672	742.5	679.0	123.2	2897.6					
24	Zenica-Doboj	974	587.9	536.1	84.1	4793.3					
26	Central Bosnia	192	777.9	713.9	207.0	2822.5					
29	Sarajevo	792	900.4	830.0	219.1	4470.9					
1	Employment										
	Employed	2019	800.3	740.8	84.1	4470.9					
2	Employer	268	922.3	826.9	137.8	4793.3					
3	Unemployed	745	639.3	573.1	93.1	2897.6					
4	Retired	1962	683.6	637.9	114.1	2822.5					
5	Others	408	645.6	583.6	112.4	2309.2					

Table 12: Summary statistics: Bosnia & Herzegovina (equivalence scale)

Category	Group	1995					2001				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	2471	203.6	179.5	0.0	1157.9	2500	245.7	214.6	16.5	3881.8
1	Sex	1951	204.5	179.9	0.0	1157.9	1880	245.0	213.6	16.5	3881.8
2	Female	518	200.1	177.0	0.0	870.3	620	247.6	218.3	29.3	1203.2
1	Age	47	200.3	175.1	44.4	465.9	55	366.6	296.3	44.8	1147.3
2	25-34	196	214.5	178.1	38.5	1157.9	218	286.1	208.2	43.5	3881.8
3	35-44	435	220.4	194.7	25.0	860.7	451	251.0	226.0	16.5	1118.0
4	45-54	529	224.8	198.0	27.4	1016.9	534	256.6	225.8	25.5	1024.2
5	55-64	500	209.1	189.3	14.5	1007.9	498	250.7	223.7	26.8	1203.2
6	65-74	514	179.9	166.1	14.7	573.8	453	221.3	202.8	29.3	845.7
7	>=75	250	158.6	145.3	0.0	661.9	291	193.6	179.8	29.1	946.5
1	Education	51	134.5	120.3	14.5	559.2	40	132.6	113.2	25.5	549.4
2	Low education	328	162.9	152.6	0.0	562.9	188	173.9	162.1	16.5	693.4
3	Medium education	737	177.2	163.0	0.0	508.2	712	197.9	181.1	23.4	974.5
4	Upper Secondary	910	223.5	194.2	45.1	1016.9	1095	262.6	230.0	41.0	3881.8
5	University & higher	432	244.9	208.5	60.7	1157.9	465	317.7	273.1	80.9	1147.3
1	Ethnic	2201	210.5	186.1	0.0	1157.9	2182	258.9	224.1	28.4	3881.8
2	Bulgarian	158	164.9	148.7	0.0	559.2	145	166.1	153.6	40.7	647.3
3	Turk	82	98.2	84.3	14.5	321.7	133	117.9	85.7	16.5	550.7
4	Roma/Gypsy	28	186.7	162.4	51.1	485.3	40	236.5	169.3	81.6	1085.2
1	Region1	385	212.9	185.5	31.1	702.6	400	283.2	234.1	49.6	3881.8
2	Sofia city	242	213.2	185.4	15.0	1157.9	250	247.2	217.1	24.6	1024.2
3	Bourgass	261	200.1	180.6	0.0	661.6	275	241.5	211.5	39.4	1074.0
4	Varna	296	180.8	157.9	0.0	1007.9	300	248.5	229.4	45.3	1053.1
5	Lovetch	182	183.8	166.6	33.9	779.7	185	245.3	224.7	39.8	880.5
6	Montana	340	205.8	181.2	25.5	1016.9	344	237.9	201.0	16.5	1178.7
7	Plovdiv	209	202.4	178.7	30.7	562.9	210	237.8	214.7	28.4	852.4
8	Rousse	299	209.0	188.8	25.0	870.3	296	228.8	201.5	24.9	1130.4
9	Sofia region	257	216.0	196.2	14.5	763.4	240	221.7	197.6	29.3	1203.2
1	Haskovo	1647	209.6	182.8	0.0	1157.9	1701	262.9	226.3	23.4	1203.2
2	Urban	824	191.6	172.5	0.0	901.4	799	208.9	185.2	16.5	3881.8
1	Region2	15	234.7	218.5	78.3	419.3	790	277.6	240.2	25.5	1178.7
2	Employment	25	176.6	168.5	25.5	334.1	122	389.0	297.6	26.8	3881.8
3	Employer	1140	183.2	167.6	0.0	1007.9	1198	223.4	202.1	28.4	1203.2
4	Unemployed	1291	221.6	193.1	14.5	1157.9	390	204.4	176.3	16.5	1147.3
5	Retired										
5	Others										

Table 13: Summary statistics: Bulgaria

Category	Group	1995					2001				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	2471	321.9	285.0	0.0	1761.4	2500	389.7	340.7	29.1	3881.8
1	Sex										
2	Male	1951	341.7	303.6	0.0	1761.4	1880	410.2	365.6	29.1	3881.8
2	Female	518	246.4	212.9	0.0	1507.4	620	327.4	282.2	29.3	2041.5
1	Age										
1	<25	47	259.5	221.6	99.2	798.6	55	472.3	406.1	81.4	1220.8
2	25-34	196	373.9	321.6	76.9	1637.5	218	461.3	384.8	87.9	3881.8
3	35-44	435	392.1	358.9	62.5	1158.5	451	450.6	415.9	36.9	1824.1
4	45-54	529	384.2	343.5	42.7	1761.4	534	439.9	402.6	52.9	1609.4
5	55-64	500	315.9	286.4	32.4	1031.9	498	387.2	340.1	80.3	1687.9
6	65-74	514	247.9	222.3	14.7	928.2	453	316.1	281.5	29.3	1288.9
7	>=75	250	200.6	179.6	0.0	693.5	291	252.5	227.8	29.1	946.5
1	Education										
1	None	51	180.3	174.3	14.7	559.2	40	192.4	164.2	61.0	549.4
2	Low education	328	227.8	204.7	0.0	739.3	188	251.3	212.0	36.9	1101.3
3	Medium education	737	280.7	261.3	0.0	1130.1	712	311.2	285.5	29.1	1687.9
4	Upper Secondary	910	368.4	330.3	47.4	1761.4	1095	426.8	382.1	59.1	3881.8
5	University & higher	432	380.2	321.1	81.7	1637.5	465	495.2	442.2	133.4	1824.1
1	Ethnic										
1	Bulgarian	2201	329.3	290.7	0.0	1761.4	2182	405.4	355.3	29.1	3881.8
2	Turk	158	290.1	275.6	0.0	965.6	145	296.5	277.2	70.5	989.3
3	Roma/Gypsy	82	186.0	174.2	32.4	643.5	133	236.1	173.7	36.9	1101.3
4	Other	28	313.6	263.5	114.2	970.6	40	381.9	307.0	131.6	1095.8
1	Region1										
1	Sofia city	385	321.1	271.3	31.1	1226.8	400	430.5	372.7	66.0	3881.8
2	Bourgass	242	345.2	293.3	21.3	1637.5	250	407.1	356.2	67.6	1687.9
3	Varna	261	327.7	289.7	0.0	997.4	275	387.7	329.4	70.5	1385.9
4	Lovetch	296	279.6	242.9	0.0	1125.7	300	383.3	354.4	61.0	1824.1
5	Montana	182	271.6	255.7	33.9	782.4	185	380.7	332.1	81.1	1525.0
6	Plovdiv	340	338.5	296.6	47.4	1761.4	344	390.0	328.2	36.9	2041.5
7	Rousse	209	304.2	283.3	43.8	741.1	210	369.0	338.2	68.6	1476.4
8	Sofia region	299	333.0	304.3	66.1	1507.4	296	376.2	336.0	29.1	1598.6
9	Haskovo	257	358.2	321.7	32.4	1012.7	240	354.8	305.2	29.3	1273.2
1	Region2										
1	Urban	1647	334.5	292.2	0.0	1761.4	1701	419.6	372.9	29.3	1824.1
2	Rural	824	296.4	270.4	0.0	1130.1	799	326.0	284.8	29.1	3881.8
1	Employment										
1	Employed	15	360.4	326.4	156.6	726.2	790	470.9	425.7	67.6	2041.5
2	Employer										
3	Unemployed	25	326.3	318.1	62.5	551.6	122	654.8	573.1	80.3	3881.8
4	Retired	1140	258.2	229.8	0.0	1031.9	1198	324.9	288.4	29.1	1687.9
5	Others	1291	377.2	341.1	14.7	1761.4	390	341.1	303.5	36.9	1277.8

Table 14: Summary statistics: Bulgaria (equivalence scale)

Category	Group	2002						2007					
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum		
1	All	6386	340.1	282.2	23.3	5768.9	5557	440.6	365.7	41.2	5528.7		
1	Sex												
	Male	4853	331.3	276.4	37.8	5768.9	4033	428.2	356.3	41.2	5528.7		
2	Female	1533	367.8	300.8	23.3	3384.4	1524	473.4	387.0	46.5	3520.2		
1	Age												
	<25	44	548.5	436.5	136.8	1720.7	44	924.8	757.7	209.5	3213.7		
2	25-34	412	424.9	327.3	46.0	3384.4	282	552.5	436.7	56.2	5528.7		
3	35-44	917	355.1	297.2	37.8	2702.2	723	455.4	373.5	53.0	3924.3		
4	45-54	1481	364.5	308.2	49.0	5768.9	1316	465.6	391.4	56.1	3693.5		
5	55-64	1373	334.0	281.3	23.3	3203.7	1274	451.4	378.1	76.1	3520.2		
6	65-74	1493	309.1	264.6	31.7	2470.3	1202	398.5	341.1	46.5	2718.6		
7	>=75	666	280.6	231.3	43.6	2123.5	716	357.5	307.8	41.2	2043.1		
1	Education												
	Low education	1430	261.1	224.8	23.3	2586.9	1087	319.7	283.4	41.2	1707.3		
2	Medium education	1286	281.5	241.2	37.8	2409.5	1164	347.8	304.0	56.1	1999.9		
3	Upper Secondary	1223	330.3	278.7	55.7	3384.4	964	431.5	369.0	80.8	3693.5		
4	University & higher	1573	394.0	337.2	75.7	3203.7	1516	489.0	406.2	92.7	3543.0		
5	None	874	471.9	396.7	92.5	5768.9	826	652.4	564.2	90.5	5528.7		
1	Ethnic												
	Serbian					4779	444.2	370.9	41.2	5528.7			
2	Montenegrin					50	601.0	446.9	235.5	3213.7			
3	Bosnian					80	364.1	303.5	122.0	1891.6			
4	Albanian					32	389.0	338.0	135.1	1017.7			
5	Hungarian					228	416.9	343.1	115.5	2005.1			
6	Bulgarian					23	337.5	307.9	163.0	759.6			
7	Croatian					78	460.2	371.0	133.0	2176.3			
8	Romanian					29	433.6	353.9	229.6	896.3			
9	Roma					96	217.8	182.9	53.0	677.5			
10	Slovakian					51	389.2	306.2	128.8	2770.1			
11	Others					111	559.7	460.4	84.9	3924.3			
1	Region1												
	Belgrade	1214	418.7	352.6	49.0	5768.9	921	569.4	465.0	56.1	3924.3		
2	Vojvodina	1673	350.6	291.2	43.6	2280.7	1388	449.2	372.9	53.0	3693.5		
3	West Serbia	707	300.2	255.7	37.8	2470.3	734	372.8	324.9	76.1	5528.7		
4	Sumadija	1156	320.4	273.9	53.0	2702.2	1032	436.1	371.6	101.7	3543.0		
5	East Serbia	671	341.7	278.7	23.3	3384.4	655	416.9	348.6	46.5	1999.2		
6	South-east Serbia	965	274.4	228.6	31.7	2971.3	827	367.4	317.7	41.2	3520.2		
1	Region2												
	Urban	3525	379.6	322.0	46.0	5768.9	2960	496.8	410.6	53.0	3924.3		
2	Rural	2861	291.4	239.6	23.3	3384.4	2597	376.6	316.4	41.2	5528.7		
1	Employment												
	Employed	2249	373.1	312.4	45.5	5768.9	1811	487.1	407.0	53.0	3924.3		
2	Employer	678	356.4	279.7	45.1	2702.2	885	440.2	354.1	90.0	5528.7		
3	Unemployed	381	277.8	239.6	58.2	1689.1	328	359.2	331.4	56.1	1676.5		
4	Retired	2739	321.6	271.2	23.3	3203.7	2227	420.6	351.7	41.2	3520.2		
5	Others	339	306.9	227.4	37.8	2971.3	306	399.3	324.1	46.5	2565.8		

Table 15: Summary statistics: Serbia

Category	Group	2002					2007				
		Nr.	Mean	Median	Minimum	Maximum	Nr.	Mean	Median	Minimum	Maximum
1	All	6386	554.3	467.0	23.3	5768.9	5557	717.5	614.7	58.2	6397.3
1	Male	4853	577.0	492.2	61.6	5768.9	4033	741.6	634.5	58.2	6397.3
2	Female	1533	482.7	391.4	23.3	4819.0	1524	653.7	547.9	80.5	4471.9
1	Age	44	706.0	575.0	193.5	1886.2	44	1203.9	1027.7	337.6	4137.1
2	25-34	412	689.5	593.1	114.3	3384.4	282	865.8	767.6	112.4	5528.7
3	35-44	917	641.5	555.2	75.5	4819.0	723	806.7	694.5	142.3	4450.0
4	45-54	1481	643.3	558.7	69.2	5768.9	1316	816.6	703.6	103.8	6397.3
5	55-64	1373	541.2	461.1	23.3	3658.4	1274	742.4	641.3	108.1	4386.1
6	65-74	1493	463.6	389.7	54.8	3249.3	1202	602.4	502.5	80.5	3844.7
7	>=75	666	373.3	312.8	46.4	2210.6	716	506.0	426.0	58.2	4471.9
1	Education	1430	391.7	329.4	23.3	3658.4	1087	471.7	406.0	58.2	2728.6
2	Low education	1286	459.7	392.3	75.5	4819.0	1164	578.4	509.4	129.2	4471.9
3	Medium education	1223	562.8	480.8	73.4	3963.6	964	731.5	631.1	103.8	6397.3
4	Upper Secondary	1573	651.4	572.6	141.9	3283.1	1516	805.0	708.1	140.2	4450.0
5	University & higher	874	773.2	675.0	156.2	5768.9	826	1060.1	936.7	90.5	5540.3
1	Ethnic	4779	726.8	626.3	58.2	626.3	4779	726.8	626.3	58.2	6397.3
2	Serbian	50	962.2	800.1	287.7	800.1	50	962.2	800.1	287.7	3479.3
3	Montenegrin	80	608.6	532.3	211.3	2675.1	80	608.6	532.3	211.3	2675.1
4	Bosnian	32	801.2	706.4	310.3	706.4	32	801.2	706.4	310.3	1762.7
5	Albanian	228	615.3	555.1	156.5	555.1	228	615.3	555.1	156.5	3472.9
6	Hungarian	23	537.1	402.5	182.9	402.5	23	537.1	402.5	182.9	1519.1
7	Bulgarian	78	685.7	597.9	172.4	597.9	78	685.7	597.9	172.4	2176.3
8	Croatian	29	671.2	549.5	229.6	549.5	29	671.2	549.5	229.6	1792.6
9	Romanian	96	416.4	379.5	103.8	379.5	96	416.4	379.5	103.8	1052.3
10	Roma	51	656.4	486.2	223.1	486.2	51	656.4	486.2	223.1	5540.3
11	Slovakian	111	831.7	669.9	84.9	669.9	111	831.7	669.9	84.9	3924.3
1	Others	921	912.1	765.0	140.2	765.0	921	912.1	765.0	140.2	4955.4
1	Region1	1214	669.0	582.2	69.2	5768.9	921	912.1	765.0	140.2	4955.4
2	Belgrade	1673	556.3	469.4	46.4	3836.0	1388	717.5	613.1	84.9	6397.3
3	Vojvodina	707	487.2	419.6	75.5	2470.3	734	613.6	540.8	108.1	5528.7
4	West Serbia	1156	553.3	478.8	78.7	3821.5	1032	716.1	618.8	119.2	3720.7
5	Sumadija	671	561.4	472.1	23.3	4819.0	655	676.3	581.6	80.5	4450.0
6	East Serbia	965	452.2	385.1	54.8	2971.3	827	627.6	563.1	58.2	3520.2
1	Region2	3525	611.3	528.7	86.7	5768.9	2960	801.3	686.7	84.9	5540.3
2	Urban	2861	484.1	401.9	23.3	4819.0	2597	622.0	535.9	58.2	6397.3
1	Employment	2249	655.0	570.7	65.0	5768.9	1811	839.4	741.7	112.4	6397.3
2	Employed	678	626.9	506.2	81.1	4819.0	885	760.5	629.3	103.8	5528.7
3	Employer	381	490.1	423.4	89.9	2648.9	328	621.6	559.4	140.2	2763.6
4	Unemployed	2739	477.7	401.9	23.3	3658.4	2227	631.7	532.9	58.2	4471.9
5	Retired	339	433.5	327.9	46.4	2971.3	306	599.4	498.7	80.5	4137.1

Table 16: Summary statistics: Serbia (equivalence scale)

B Regression results

		No equivalence scale		Equivalence scale ($\rho = 0.5$)		
		2002	2005	2002	2005	
Age and sex	Age	0.682 (0.334)	-1.711* (0.076)	2.497** (0.029)	1.553 (0.387)	
	Age ²	0.001 (0.913)	0.022** (0.018)	-0.016 (0.143)	-0.008 (0.641)	
	Male=1	-29.258*** (0.000)	-51.979*** (0.000)	2.046 (0.775)	-1.973 (0.855)	
Employment status	EmpSh	64.211*** (0.000)	75.942*** (0.000)	99.809*** (0.000)	91.906*** (0.000)	
	Employed	-3.798 (0.294)	26.778*** (0.000)	-7.327 (0.210)	58.075*** (0.000)	
	Employer	121.040*** (0.000)	171.344*** (0.000)	240.722*** (0.000)	377.459*** (0.000)	
	Unemployed	-4.287 (0.475)	27.726*** (0.003)	-23.660** (0.015)	33.903* (0.054)	
	Retired	14.447*** (0.005)	27.569*** (0.000)	10.785 (0.196)	31.036** (0.017)	
Education	Low	15.358** (0.022)	-5.071 (0.561)	15.132 (0.161)	-8.498 (0.601)	
	Medium	24.536*** (0.000)	11.714 (0.170)	28.188*** (0.009)	16.534 (0.299)	
	Upper	55.707*** (0.000)	51.889*** (0.000)	81.728*** (0.000)	79.615*** (0.000)	
	Tertiary	103.417*** (0.000)	114.535*** (0.000)	162.239*** (0.000)	186.303*** (0.000)	
Ethnicity	Albanian=1	-72.377*** (0.000)	-8.276 (0.413)	-70.061*** (0.000)	-6.756 (0.719)	
Region	Tirana	28.841*** (0.000)	-19.504*** (0.000)	31.141*** (0.000)	-21.668** (0.029)	
	Coastal	26.126*** (0.000)	-32.596*** (0.000)	33.33*** (0.000)	-46.798*** (0.000)	
	Central	10.102** (0.011)	-51.933*** (0.000)	3.117 (0.629)	-69.887*** (0.000)	
Constant	Constant	138.091*** (0.000)	204.898*** (0.000)	181.579*** (0.000)	213.593*** (0.000)	
		<i>F</i>	50.552	67.438	48.823	56.159
		<i>R</i> ²	0.184	0.230	0.179	0.199
		Obs.	3599	3638	3599	3638

Table 17: Regression output: Albania

		No equivalence scale		Equivalence scale ($\rho = 0.5$)	
		2001	2004	2001	2004
Age and sex	Age	-3.916**	n.a.	4.927**	n.a.
		'(0.026)	n.a.	'(0.023)	n.a.
	Age ²	0.049***	n.a.	-0.056***	n.a.
		'(0.002)	n.a.	'(0.004)	n.a.
	Male=1	-136.042***	n.a.	-44.759***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
Employment status	EmpSh	268.479***	n.a.	373.519***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
	Employed	-102.729***	n.a.	-90.766***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
	Employer	-9.764	n.a.	70.384***	n.a.
		'(0.660)	n.a.	'(0.010)	n.a.
	Unemployed	-61.132***	n.a.	-68.166***	n.a.
		'(0.001)	n.a.	'(0.002)	n.a.
	Retired	-0.108	n.a.	-8.531	n.a.
		'(0.994)	n.a.	'(0.647)	n.a.
Education	Low	35.954***	n.a.	44.739***	n.a.
		'(0.002)	n.a.	'(0.002)	n.a.
	Medium	108.008***	n.a.	151.009***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
	Upper	175.001***	n.a.	252.859***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
	Tertiary	328.356***	n.a.	474.414***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
Ethnicity	Ethnic		n.a.		n.a.
			n.a.		n.a.
Region	Banhja Lukay = 1	-48.301***	n.a.	-46.837***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
Constant	Constant	530.638***	n.a.	518.643***	n.a.
		'(0.000)	n.a.	'(0.000)	n.a.
	<i>F</i>	95.510	n.a.	128.705	n.a.
	<i>R</i> ²	0.187	n.a.	0.237	n.a.
	Obs.	5402	n.a.	5402	n.a.

Table 18: Regression output: Bosnia & Herzegovina

		No equivalence scale		Equivalence scale ($\rho = 0.5$)	
		1995	2001	1995	2001
Age and sex	Age	2.496*** (0.008)	-4.813*** (0.000)	7.093*** (0.000)	-0.449 (0.788)
	Age ²	-0.028*** (0.001)	0.032*** (0.004)	-0.081*** (0.000)	-0.022 (0.150)
	Male=1	-14.258*** (0.008)	-13.939** (0.049)	47.725*** (0.000)	50.290*** (0.000)
Employment status	EmpSh	91.662*** (0.000)	135.170*** (0.000)	109.459*** (0.000)	158.876*** (0.000)
	Employed		-55.627*** (0.000)		-33.365** (0.031)
	Employer	-66.709 (0.231)	64.335*** (0.000)	-27.309 (0.751)	154.276*** (0.000)
	Unemployed	-45.603 (0.378)		-37.627 (0.637)	
	Retired	-69.623 (0.180)	-36.784*** (0.001)	-37.612 (0.639)	-42.852*** (0.007)
Education	Low	0.381 (0.957)	11.759 (0.313)	1.928 (0.858)	10.863 (0.500)
	Medium	29.445*** (0.000)	40.049*** (0.001)	46.935*** (0.000)	50.742*** (0.003)
	Upper	56.969*** (0.000)	84.222*** (0.000)	75.495*** (0.000)	111.097*** (0.000)
Ethnicity	Bulgarian=1	41.015*** (0.000)	67.695*** (0.000)	50.987*** (0.000)	90.546*** (0.000)
Region	Sofia	-11.099* (0.065)	3.400 (0.684)	-26.057*** (0.005)	1.209 (0.916)
Constant	Constant	142.280** (0.014)	299.193*** (0.000)	87.518 (0.326)	285.554*** (0.000)
	<i>F</i>	36.323	48.670	58.469	71.942
	<i>R</i> ²	0.152	0.190	0.224	0.258
	Obs.	2448	2500	2448	2500

Table 19: Regression output: Bulgaria

		No equivalence scale		Equivalence scale ($\rho = 0.5$)	
		2002	2007	2002	2007
Age and sex	Age	-6.634*** (0.000)	-10.793*** (0.000)	-0.528 (0.799)	-4.143 (0.115)
	Age ²	0.039*** (0.003)	0.067*** (0.000)	-0.038** (0.040)	-0.015 (0.519)
	Male=1	-64.957*** (0.000)	-67.660*** (0.000)	25.379** (0.016)	27.543** (0.032)
	EmpSh	146.174*** (0.000)	201.766*** (0.000)	196.721*** (0.000)	259.684*** (0.000)
Employment status	Employed	-66.269*** (0.000)	-98.201*** (0.000)	-51.626** (0.021)	-85.301*** (0.002)
	Employer	3.272 (0.844)	-40.524** (0.042)	58.643** (0.016)	3.230 (0.911)
	Unemployed	-80.500*** (0.000)	-101.983*** (0.000)	-97.634*** (0.000)	-127.374*** (0.000)
	Retired	20.395 (0.134)	32.334* (0.061)	23.405 (0.238)	39.262 (0.119)
	Low	-197.051*** (0.000)	-305.842*** (0.000)	-286.813*** (0.000)	-470.608*** (0.000)
Education	Medium	-180.369*** (0.000)	-287.947*** (0.000)	-274.336*** (0.000)	-436.134*** (0.000)
	Upper	-126.080*** (0.000)	-197.181*** (0.000)	-190.039*** (0.000)	-305.381*** (0.000)
	Tertiary	-84.365*** (0.000)	-160.239*** (0.000)	-122.149*** (0.000)	-254.623*** (0.000)
	Serbian = 1		14.862 (0.207)		51.664*** (0.003)
Region	Belgrade	89.278*** (0.000)	123.613*** (0.000)	149.958*** (0.000)	185.931*** (0.000)
	Vojvodina	61.765*** (0.000)	70.311*** (0.000)	88.697*** (0.000)	86.689*** (0.000)
	West Serbia	30.023*** (0.008)	1.945 (0.890)	39.776** (0.016)	-21.329 (0.297)
	Sumadija	45.263*** (0.000)	62.098*** (0.000)	101.091*** (0.000)	86.375*** (0.000)
	East Serbia	70.156*** (0.000)	47.485*** (0.001)	127.581*** (0.000)	58.487*** (0.006)
Constant	Constant	678.687*** (0.000)	959.536*** (0.000)	737.464*** (0.000)	1098.405*** (0.000)
	<i>F</i>	65.363	76.469	91.440	95.537
	<i>R</i> ²	0.149	0.199	0.196	0.237
	Obs.	6386	5557	6386	5557

Table 20: Regression output: Serbia

References

- Anand, S. and P. Segal (2008). What do we know about global income inequality? *Journal of Economic Literature XLVI*(1), 57–94.
- Canova, L. (2006). The many dimensions of poverty in Albania: Income, wealth and perceptions. *Unpublished Paper*.
- Deaton, A. (1997). *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy*. Washington, D.C.: The World Bank.
- Fields, G. (2003). Accounting for income inequality and its change: A new method, with application to the distribution of earnings in the United States. *Res. Labor Econ.* 22, 1–38.
- Fields, G. S. and G. Yoo (2000). Falling labor income inequality in Korea’s economic growth: Patterns and underlying causes. *Review of Income and Wealth* 46(2), 139–159.
- Giammatteo, M. (2007). The bidimensional decomposition of inequality: A nested Theil approach. *Working paper, University of Rome 'La Sapienza'*.
- Gunatilaka, R. and D. Chotikapanich (2006). Inequality trends and determinants in sri lanka 1980-2002: A shapely approach to decomposition. *Monash University Working Paper 6/06*.
- INSTAT (2006). Albania: Trends in poverty and inequality, 2002-2005. *Working Paper*.
- Israeli, O. (2007). A Shapely-based decomposition of the R-square of a linear regression. *Journal of Economic Inequality* 212, 199–.
- Jenkins, S. P. (1995). Accounting for inequality trends: Decomposition analysis for the UK, 1971-86. *Economica* 62, 29–63.
- Milanovic, B. (1999). Explaining the increase in inequality during transition. *Economics of Transition* 7(2), 299–341.
- Milanovic, B. (2005). *Worlds Apart. Measuring International and Global Inequality*. Princeton: Princeton University Press.
- Mookherjee, D. and A. F. Shorrocks (1982). A decomposition analysis of the trend in UK income inequality. *Economic Journal* 92, 886–992.

- Morduch, J. and T. Sicular (2002). Rethinking inequality decomposition, with evidence from rural China. *Economic Journal* 112, 93–106.
- Sen, A. (1997). *On Economic Inequality*. Oxford, UK: Oxford University Press.
- Shorrocks, A. F. (1999). Decomposition procedures for distributional analysis: A unified framework based on the shapely value. *Draft paper, University of Essex*.
- Wan, G. (2004). Accounting for income inequality in rural China: A regression-based approach. *Journal of Comparative Economics* 32, 348–363.