7. INCOME INEQUALITY AND ITS DECOMPOSITION: THE CASE OF ESTONIA

Alari Paulus

Introduction

Various forms of inequality can be observed in every society. Because of the differences between individuals, the existence of inequality is to some extent natural. However, deciding whether it is low or high is not an easy task. This requires coherent definition and measurement of inequality, complemented with an analysis of its causes and consequences.

Economists, who are always looking for optimality, try to assess the potential impact of inequality on long-term economic development. Accordingly, the interrelationship between inequality and economic growth has been extensively researched. This topic already attracted classical economists, but the beginning of more serious theoretical and empirical studies is marked with an influential paper by Kuznets (1955). Yet, there is still no common understanding about whether inequality influences economic growth or vice versa and whether their relationship is positive or negative.

It has been perceived that inequality has increased in the societies characterized by transitional processes. This is also the case for Estonia where deepening gaps between different population groups are considerable. An evaluation of these dynamics requires determination of the extent and sources of inequality. Increasing inequality might be a case for taking certain economic and socio-political measures in order to avoid social tensions.

The aim of this article is to find out how different social and demographic factors are associated with the extent of income inequality in Estonia. As income inequality can be understood and measured in several ways, a considerable attention is paid to these issues as well.

The article consists of four sections. The first is a theoretical discussion demonstrating different definitions of economic inequality and systematizing the factors of inequality. The second section analyzes the measures of income inequality in order to choose appropriate ones for empirical assessment. The third part is an analysis of income inequality dynamics in Estonia and a comparison to other European countries. In the last section, inequality measures are decomposed to explain the main reasons for inequality in Estonia and to parallel them with theoretical insights.

The present empirical analysis is based on the household budget surveys conducted by the Statistical Office of Estonia in the period 1996–2002. The international comparison is based on the data compiled by *Luxembourg Income Study*.

6.1. Economic inequality and its sources

Economic inequality implies that economic resources are not divided equally between individuals. Inequality analysis can be based either on wealth, income, expenditure, or utility distribution. Most often wealth and income inequality are considered. Wealth denotes accumulated assets; income, on the other hand, reflects more the potential of an individual to accumulate resources in the future. Expenditures are mostly considered if adequate data about incomes are missing, and utility is rather difficult if not impossible to measure empirically. Herein the emphasis is on income inequality as the data of wealth distribution are also quite limited.

The research on inequality presumes defining equal distribution. Then the deviation between this and the actual distribution of income can be measured. Equality is usually considered as resources being divided absolutely equally. However, the question remains whether such an approach is correct, considering the large variety of individuals' abilities and needs.

A theoretical discussion adds two alternatives which are relatively intuitive and logical. Firstly, the aim of a society could be to equalize individual utility through satisfied needs. This does not necessarily imply that economic resources must be divided equally between individuals. Differences in individual welfare functions imply that an individual with greater needs (e.g. poor health) should be given a greater share of resources (Sen 1997, p. 78). Although this approach might have a better theoretical background, it could not be applied because of difficulties with defining and measuring the needs.

The second alternative focuses on the principle of distribution according to desert, indicating that an individual's share of total resources should correspond to the value of work done (Sen 1997, p. 87). However, the work contribution is closely related to individual abilities and there are people who are not able to manage themselves. Therefore, for the sake of humanity, compensating for the work contribution cannot be the only criterion for distributing resources.

Considering shortcomings of these alternatives, an equal distribution of resources is usually accepted for the comparison basis. This means that different individual preferences, needs and abilities are not taken into account. Therefore the approach is simple and allows easily to build up mathematical models.

But what are the causes of income inequality? In general, incomes are generated in a process whose major stages are **factor endowments, factor productivity and redistribution of income**. The

distribution of market income results from the first two stages, while the third determines the distribution of disposable income.

The first stage, factor endowments, indicates that income inequality is based on differences in the ownerships of physical and human capital. Human capital is related to individuals' education and experiences; physical capital indicates distribution of wealth accumulated over generations. Formation of human capital is based on individual abilities; their development depends on societal opportunities (e.g. educational and health system). For physical capital, the concern is related to imperfect capital markets (e.g., because of information asymmetry), meaning limited access to credit and investment opportunities for some population groups (Barro 1999, p. 2). Physical capital can also be inherited, which itself is influenced by social norms and traditions (Tanzi 1998, p. 9). In general, the distribution of physical capital is perceived as more persistent and unequal than that of human capital.

The next stage, factor productivity, depends on the labour and capital market. The yield of human capital (mostly wages) is mainly determined by employment and labour market flexibility, the latter depending on labour market institutions, e.g. labour unions.[1] Similarly, the capital market is relevant as its imperfectness could result in different marginal productivity in different businesses (Barro 1999, p. 1).

The last stage in income generation is redistribution of incomes. This can occur both in the public and private sector. In the public sector, redistribution takes place through the tax system (including public goods) and in the private sector through income transfers inside the households, underlining the economies of scale resulting from the structure of a household. The total extent of in-household transfers depends on general demographic characteristics, for

[1] There is empirical evidence that labour unions have an equalizing effect on income distribution and an influence on allocation processes (Wagschal 1997, p. 14).

instance, the average size of families, percentage of single parent households and population's age structure.

The tax system, in addition to determining the after-tax income distribution, affects individual behaviour (e.g., savings propensity) and the relative productivity of physical and human capital (Atkinson, Stiglitz 1980, p. 280). However, the role of taxation remains somewhat limited, depending on the tax base and the administration of taxes (Chu *et al.* 2000, p. 3). Some theorists also conclude that taxation cannot efficiently reduce inequality through direct redistribution based on taxes and income transfers, but might do so by establishing a stable tax base for public expenditures and thereby avoiding horizontal inequality (Tanzi 1998, p. 15).[2]

The extent of income redistribution resulting from taxation is determined by social norms stating society's understanding of justice and preferred level of equality. Most people would vote for redistribution of incomes from the rich to the poor when the average income exceeds the median. Of course this can take place only if political power is allocated more equally than economic (Barro 1999, p. 3–4).

In summary, economic inequality can be seen as a result of different individual endowments. Additionally, inflexibility of economic institutions may increase inequality, and social norms usually help to limit it. On the other hand, the effectiveness of the latter also depends on the level of economic development.

6.2. Measurement of income inequality

There are several methods for studying the extent of income inequality. In general, these can be divided into graphic and analytical measures (see Figure 1), while the latter can be further

[2] Horizontal equality implies taxing individuals with same income identically.

subdivided into positive and normative measures. Positive measures have an important subgroup of measures based on information theory.

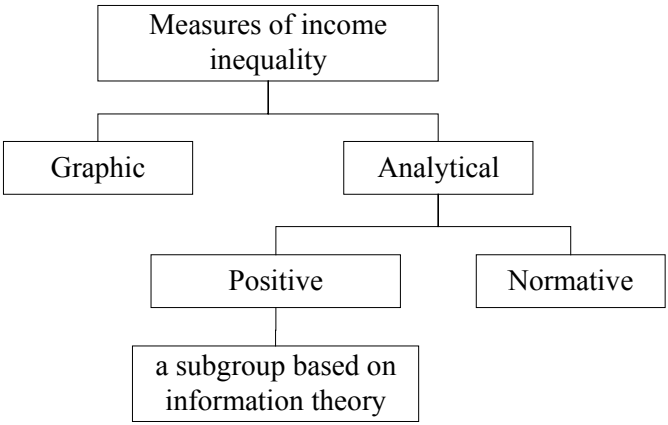

Figure 1. Measures of income inequality.

Graphic methods give a general and illustrative overview of underlying income distributions. Examples of graphic methods include *the parade of dwarfs*, frequency distributions and the Lorenz curve. The parade of dwarfs reflects non-aggregated incomes in ascending order (Cowell 1995, p. 16). Frequency distributions include histogram, density and frequency functions, which differ according to the way the incomes are aggregated. Income distributions are often based on logarithmic incomes as otherwise the incomes lie within a wide range and, for instance, compared to the normal distribution, the distribution of incomes has a higher kurtosis and a long right tail. The distribution of logarithmic incomes is similar to the normal distribution, making the process of modelling more accurate.

The most frequently used graphic method is the Lorenz curve. In the figure then, the horizontal axis shows the percentage of popu-

lation in ascending order of incomes and the vertical axis represents the percentage of income received by the bottom percentage of the population. If incomes are distributed completely equally, then the Lorenz curve will become a diagonal. Otherwise, individuals with bottom incomes have a lower share in total incomes — therefore the curve will lie below the diagonal and its slope will be increasingly rising (Sen 1997, p. 30).

Figure 2 presents an example of the Lorenz curve for Estonia in 2002. It shows that the poorest 40% of the population receives only 20% of the total income, while richest 20% of the population enjoys around 40% of the total income (using equivalence scales), which implies a relatively large inequality.

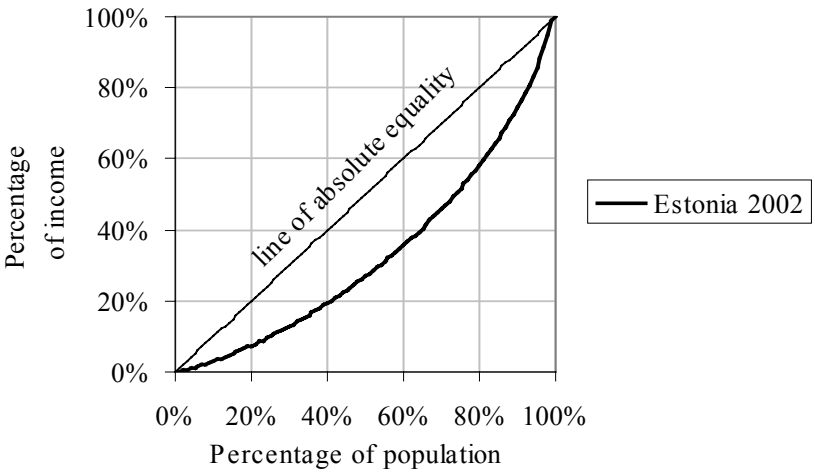

Figure 2. Lorenz curve (data from the Household Budget Survey 2002, author's calculations).

There is also a generalized Lorenz curve differing by its vertical axis, where the percentage of income is multiplied by the average level of income. A generalized Lorenz curve is appropriate for comparing different income distributions if regular Lorenz curves

for those distributions intersect (Shorrocks 1983, p. 15). The Lorenz curve is also well known for being one possible way to define the widely employed analytical measure, the Gini coefficient.

Unlike graphics, analytical methods express the extent of inequality by a quantitative estimation. The positive methods define inequality in some objective sense, often using some statistical measure of relative variation of income. The normative methods, on the other hand, attempt to measure inequality in terms of some subjective notion of social welfare, implying that a higher degree of inequality corresponds to a lower level of social welfare (Sen 1997, p. 2)

There are some specific axioms whose fulfilments characterize analytical measures of inequality. Therefore the analytical approach is also known as axiomatic approach. The five key axioms are the following (see Anand 1997, p. 84–89; Champernowne 1974, p. 789–790; Cowell 1995, p. 55–61; Cowell 2000, p. 97–100):

1. **Anonymity** (also symmetry) — inequality measure should not depend on individual characteristics other than income.
2. **Income scale invariance** — inequality measure should be invariant to uniform proportional changes in all incomes (e.g., changing the currency unit).
3. **Population scale invariance** — inequality measure should be invariant to replications of the population.
4. **The Pigou-Dalton transfer principle** — inequality measure should increase (at least not decrease) in the case of mean-preserving increase in dispersion of distribution: an income transfer from a poorer to a richer individual should increase (at least not decrease) the inequality measure, and a transfer from a richer to a poorer individual should decrease (at least not increase) the inequality measure. There are defined a weak and a strong form of this principle. In the first case an income transfer from a richer to a poorer individual always decreases inequality; whereas in the second case the scope of decrease

depends strictly on the distance between the initial incomes of these individuals.

5. **Decomposability** — overall inequality should be consistently related to inequality in population subgroups, i.e. overall inequality should increase if inequality increases in every subgroup. An inequality measure is additively decomposable if overall inequality is a sum of weighted average of inequality within and between subgroups.

There are additional axioms (e.g., inequality measure should be within the range from 0 to 1), which are not fully acknowledged, or are suitable only for a certain problem set.

The main analytical measures are presented in Appendix 1 with definitions for each and fulfilments of the abovementioned axioms, except the anonymity axiom which is satisfied for every noted measure. The only measures satisfying all these axioms are members of the class of generalized entropy measures or its ordinary equivalents (Cowell 1995, p. 60).

Generalized entropy measures $GE(\alpha)$ belong to the group of measures based on information theory. The parameter α represents the weight given to distances between incomes in different parts of the income distribution. $GE(\alpha)$ is more sensitive to changes in the lower tail of the distribution for low parameter values ($\alpha<1$), and more sensitive to changes that affect the upper tail for high parameter values ($\alpha>1$) (Litchfield 1999, p. 3). A minor deficiency is that negative and zero incomes cannot be used with the parameter α smaller than 1. The most often used values of the parameter α are 0, 1 and 2. $GE(0)$ and $GE(1)$ are also Theil's measures of inequality, namely, the mean logarithmic deviation L and the Theil index T. $GE(2)$ can be also defined as half the squared coefficient of variance.

A well-known inequality measure is the Gini coefficient. It is often defined in terms of the Lorenz curve, stated as the ratio of the area between the diagonal and the curve to the triangular area below the

diagonal. In taking differences over all pairs of incomes, the Gini coefficient avoids excessive concentration on differences from the average which could be no one's actual income (Sen 1997, p. 31). This measure only lacks additive decomposability, meaning that in the case of the Gini coefficient the overall inequality cannot be shown strictly as the sum of inequality within and between the subgroups.

The normative measures combine together income and utility. It means that additional assumptions about the individual utility function must be made. A utility function with constant elasticity is often used: $U(y_i) = \frac{y_i^{1-\varepsilon} - 1}{1-\varepsilon}$ (Cowell 1995, p. 37), where y_i indicates the income of *i*-th individual and parameter ε expresses inequality aversion. The most common values for the parameter ε are 0.5, 1 and 2. Examples of normative measures are Dalton's and Atkinson's measure. These measures turn out to be ordinary equivalents to generalized entropy measures with $\alpha = 1 - \varepsilon < 1$ when using the utility function of constant elasticity (Cowell 1995, p. 60).

In summary, the most often used inequality measures are those that comply with the abovementioned axioms — the generalized entropy measures and the Gini coefficient. Although the latter is not additively decomposable, it has other advantages, for example, the intuitive interpretation in terms of the Lorenz curve. Therefore, these measures will also be used in the following empirical application.

6.3. Income inequality in Estonia in 1996–2002

The subsequent empirical analysis concentrates on income inequality in Estonia for the period 1996–2002, using data from the Household Budget Surveys (HBS) conducted by the Statistical

Office of Estonia. It must be noted that the data collection methods were somewhat different in 1996–1999 and 2000–2002. Therefore, some bias may be found in the dynamics of inequality.

The statistical information concerning the observations of household income in HBS is presented in Appendix 2. First, one can notice that the sample size declines roughly twice over the sample period. Second, the percentage of negative and zero incomes in the total observations is less than 1%. Therefore, leaving these observations aside in the following calculations should not cause a considerable bias in the results. The number of households in the sample is around 6 thousand, which is approximately 1% of the country's total.

The statistics of disposable income and equivalent income per household member are pointed out in Appendices 3 and 4, respectively. The *Luxembourg Income Study* techniques are used for calculating equivalent incomes,[3] i.e. the equivalence scale is found as the square root of the household size, with the coding bottom at 1% of mean equivalent income and top at 10 times of median (non-equivalent) disposable income.

Inequality measures are calculated using the equivalent income per household member (over individuals, not households). Proceeding according to the previous section, the dynamics of inequality is described using the Gini coefficient, the mean logarithmic deviation (i.e. generalized entropy measure with α=0), the Theil index (i.e. generalized entropy measure with α=1), and half the squared coefficient of variation (i.e. generalized entropy measure with α=2).

The results are presented in Figure 3. All the measures reflect a trend of increasing income inequality. However, the inequality has

[3] Equivalent income takes into account the scale of economies in the household. Equivalent income = Disposable income/(Number of household members)^E. Elasticity of equivalence E ranges from 0 to 1 (Gottschalk, Smeeding 1999, p. 13). Herein E=0.5 has been used.

not increased steadily, because of declines in 1999 and 2001. Especially the latter, following the attainment of the highest level of inequality so far, is somewhat confusing and might be the result of the changes in the data collection methods. The dynamics of different measures is similar, though the scale of changes is quite different. Comparing different entropy measures indicates that relatively larger changes took place in the upper tail of income distribution.

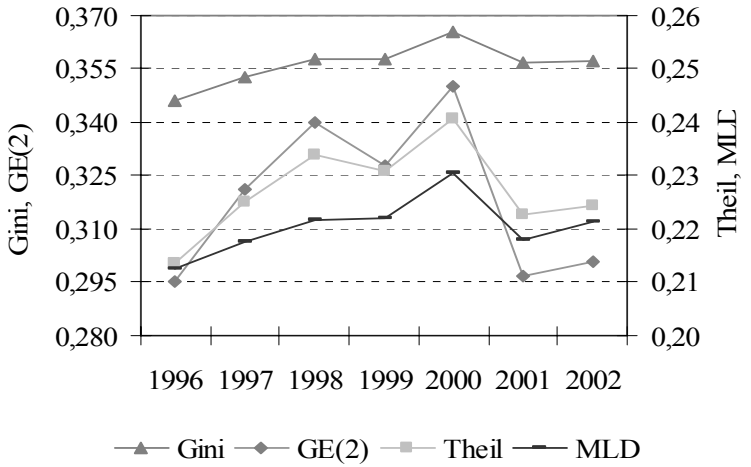

Figure 3. Income inequality in Estonia in 1996–2002: the Gini coefficient, the generalized entropy measure (α=2), the Theil index and the mean logarithmic deviation (data from the Household Budget Surveys 1996–2002, author's calculations).

Figure 4 shows the Gini coefficient together with the mean equivalent income. The latter has steadily increased over time, being strongly correlated with the inequality measure in 1996–2000. Large differences in the dynamics follow again in 2001–2002.

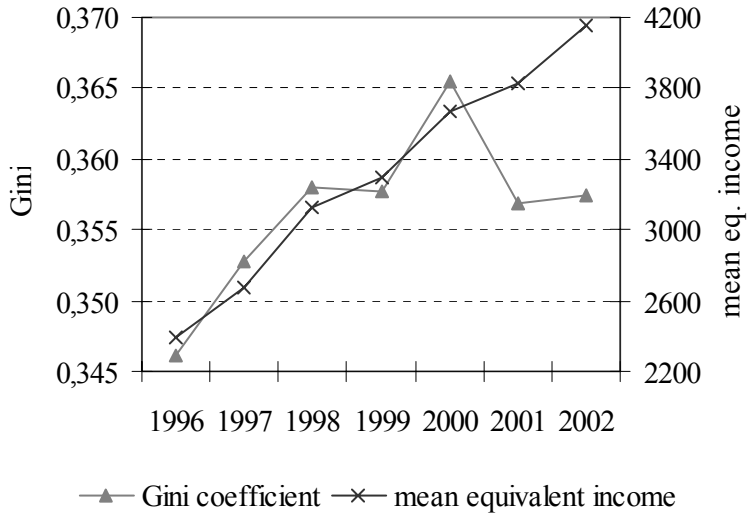

Figure 4. Income inequality in Estonia in 1996–2002: the Gini coefficient and the mean equivalent income (data from the Household Budget Surveys 1996–2002, author's calculations).

The average of the Gini coefficient in 1996–2002 is 0.356. Yet, this single number is not enough to allow us to decide whether the measure indicates high or low inequality in Estonia. Table 1 shows some income inequality measures for 18 European countries. The data stems from *Luxembourg Income Study* and therefore the Gini coefficient is somewhat different for Estonia by comparison with the result attained previously (0.361 and 0.366, respectively).

Table 1 indicates that the level of income inequality in Estonia is one of the highest in Europe, being higher only in Russia. A relatively high level of inequality characterizes also Great Britain and Italy. Additionally, the Gini coefficient for Hungary exceeds 0.3, while in the rest of the countries the measure lies below this value. There are large differences between Estonia and its closest neighbours, Finland and Sweden, who are at the bottom of the table.

Table 1. Income inequality in Europe

Country	**Year**	**Gini**	$q_{90\text{-}100}/\ q_{0\text{-}10}$	$q_{90\text{-}100}/\ q_{0\text{-}50}$	$q_{80\text{-}100}/\ q_{0\text{-}20}$
Russia	1995	0.447	9.39	2.82	3.95
Estonia	2000	0.361	5.09	2.34	2.85
Great Britain	1999	0.345	4.58	2.15	2.85
Italy	1995	0.342	4.77	2.02	2.76
Hungary	1994	0.323	4.19	2.09	2.42
Poland	1999	0.293	3.59	1.88	2.24
France	1994	0.288	3.54	1.91	2.23
Austria	1995	0.277	3.73	1.79	2.33
Germany	1994	0.261	3.18	1.74	2.10
Czech Rep.	1996	0.259	3.01	1.79	2.04
Denmark	1997	0.257	3.15	1.62	2.18
Holland	1994	0.253	3.15	1.73	2.15
Belgium	1997	0.250	3.19	1.70	2.14
Slovenia	1999	0.249	3.15	1.67	2.02
Finland	2000	0.247	2.90	1.64	2.02
Norway	1995	0.238	2.83	1.57	1.95
Luxemburg	1994	0.235	2.92	1.73	1.97
Sweden	1995	0.221	2.61	1.56	1.76

Note: Figures q90-100/q0-10, q90-100/q0-50 and q80-100/q0-20 represent the ratios of corresponding quantiles.
Source: LIS Project Key Figures.

Comparison of Estonia and some other transitional countries (Hungary, Poland, the Czech Republic and Slovenia) does not indicate the transition process to be a major factor for the high level of inequality, although it can have some importance.

6.4. Decomposition of income inequality in Estonia

Inequality measures and their dynamics characterize overall income distribution. Studying the causes of inequality implies analyzing income distribution in population subgroups. Herein decomposition is used, meaning that the whole distribution is first divided into subgroups, which is followed by calculations of inequality within and between these groups. Usually additively decomposable inequality measures are preferred, as they are easier to interpret. In this case, overall inequality is the sum of inequality within and between subgroups.

Subsequently three additively decomposable entropy measures are used: the mean logarithmic deviation $GE(0)$, the Theil index $GE(1)$ and half the squared coefficient of variation $GE(2)$. The decomposition formula for the generalized entropy measures is the following (Mookherjee, Shorrocks 1982, p. 889):

$$(1) \qquad GE(\alpha) = \sum_{k=1}^{K} v_k \lambda_k^{\alpha} GE_k(\alpha) + \frac{1}{\alpha(\alpha-1)} \sum_{k=1}^{K} v_k \left[\lambda_k^{\alpha} - 1\right], \quad \alpha \neq 0;1.$$

Here K is the number of subgroups, v_k the share of population in the subgroup k and λ_k the ratio of the average income in the subgroup k to overall average income. The first term on the right hand side represents inequality within groups and the second term inequality between groups.

Specifically,

$$(2) \qquad GE(0) = \sum_{k=1}^{K} v_k GE_k(0) + \sum_{k=1}^{K} v_k \log(\frac{1}{\lambda_k}),$$

$$(3) \qquad GE(1) = \sum_{k=1}^{K} v_k \lambda_k GE_k(1) + \sum_{k=1}^{K} v_k \lambda_k \log \lambda_k ,$$

$$(4) \qquad GE(2) = \sum_{k=1}^{K} v_k \lambda_k^2 GE_k(2) + \frac{1}{2} \sum_{k=1}^{K} v_k \left[\lambda_k^2 - 1\right].$$

The decomposition covers ten variables for different social and demographic household characteristics with a possible influence on income inequality for the year 2002. A complete description of these variables is included in Appendix 5. The importance of a variable in explaining income inequality is denoted as a ratio of the inequality between subgroups to overall inequality. The results are presented in Table 2 (in descending order).

Table 2. The ratio of inequality between subgroups to overall inequality

(%)

Variable	GE(0)	GE(1)	GE(2)
Social group of household	26.6	23.7	16.5
Employment status of the head of household[4]	19.0	16.7	11.3
Educational level of the head of household	12.2	11.6	8.5
County	10.1	9.7	7.1
Age of the head of household	6.3	5.7	3.9
Type of household	5.5	5.1	3.5
Number of household members	4.7	4.2	2.9
Nationality of the head of household	3.0	2.9	2.1
Type of settlement	2.8	2.7	2.0
Gender of the head of household	2.1	2.1	1.5

Source: data from the Household Budget Survey 2002, author's calculations.

[4] The head of a household is considered to be the member of the household who has made the largest contribution to the household's income in the long run.

According to the table, the main causes of income inequality are quite trivial, concerning the social group of household and employment status of the head of household. Both reflect large income differences between employed and unemployed individuals. On this basis, the social group of household is more important as it includes all household members, not only the head.

Likewise the educational level of the head of the household and the location of the household (namely the county) have some importance. The former characterizes the formal side of human capital, the latter the differences between regional markets influencing the productivity of (human) capital. Also, the age of the head of household (indirect human capital), the type and size of household are somewhat significant. The rest of the variables, such as nationality and gender of the head of household and type of settlement have only marginal importance. Next, the first four variables will be analyzed in more detail.

The social group of household shows the significance of working household members (Table 3, in ascending order by the mean income). The average income of subgroups is in logical correlation with the number of employed household members. The groups of "other inactive" and "unemployed" have the lowest incomes followed by the group of "retired". The income level of households with one employed member is close to but still below the average. Only households with two or more employed members earn above the average income.

However, income inequality follows a somewhat different pattern (for better expressiveness the Gini coefficient is also included). The most equal income distribution characterizes households of retired people, implying that the retirement benefits are only slightly differentiated. The average level of inequality is characteristic of the households with employed members and a high inequality is present in the groups of "unemployed" and "other inactive", especially in the case of the measure $GE(2)$ which emphasizes

large differences among top incomes in the group. This result could be an implication of shadow economy.

Table 3. Decomposition of income inequality by the social group of the household

Social group of the household	% of population	Mean income	Relative mean	GE(0)	GE(1)	GE(2)	Gini
other inactive	3.0	1942	0.467	0.264	0.333	0.733	0.379
unemployed	5.2	1950	0.469	0.268	0.324	0.793	0.378
retired	14.5	2341	0.563	0.055	0.055	0.061	0.179
one employed member	33.1	3724	0.896	0.176	0.184	0.259	0.317
two or more employed members	44.3	5481	1.319	0.171	0.171	0.216	0.314
within groups				0.164	0.171	0.251	–
between groups				0.059	0.053	0.050	–
Total	**100.0**	**4156**	**1.000**	**0.222**	**0.224**	**0.301**	**0.357**

Source: data from the Household Budget Survey 2002, author's calculations.

Table 4 reflects employment more narrowly, focusing on the head of the household. In this case, the group of "unemployed" has the lowest income and the group of "other inactive" is in a somewhat better position. In between these two groups is the group of individuals with the retired head of household. Individuals belonging to the group of households whose head is employed or otherwise active have above average incomes. The lowest inequality describes the group of "retired", followed by "employed". The groups of "other inactive", "other active" and "unemployed" have the highest inequality levels. These groups are small and with rather various compositions.

Table 4. Decomposition of income inequality by the employment status of the head of household

Employment status of the head of household	**% of population**	**Mean income**	**Relative mean**	**GE(0)**	**GE(1)**	**GE(2)**	**Gini**
unemployed	5.6	2168	0.521	0.266	0.238	0.272	0.380
retired	21.2	2574	0.619	0.109	0.130	0.222	0.249
other inactive	2.7	2738	0.659	0.404	0.402	0.602	0.476
employed	64.6	4780	1.150	0.177	0.177	0.218	0.321
other active	5.9	5557	1.337	0.292	0.311	0.484	0.402
within groups				0.180	0.187	0.267	–
between groups				0.042	0.037	0.034	–
Total	**100.0**	**4156**	**1.000**	**0.222**	**0.224**	**0.301**	**0.357**

Source: data from the Household Budget Survey 2002, author's calculations.

The educational level of the head of household reflects the distribution of human capital and resulting inequality. The figures in Table 5 indicate that higher incomes and inequality are associated with a higher educational level. It is possible to distinguish between three clusters on the basis of differences in the level of inequality — individuals belonging to the households whose head has a) no education, b) primary or secondary education, or c) higher education.

Table 5. Decomposition of income inequality by the educational level of the head of household

Educational level of the head of household	**% of population**	**Mean income**	**Relative mean**	**GE(0)**	**GE(1)**	**GE(2)**	**Gini**
no education	4.9	2393	0.576	0.122	0.101	0.113	0.242
primary	18.9	2999	0.722	0.165	0.170	0.222	0.306
general secondary	17.2	3922	0.944	0.184	0.171	0.201	0.319
special secondary	29.6	4077	0.981	0.193	0.184	0.220	0.330
higher	29.4	5416	1.303	0.236	0.238	0.317	0.369
within groups				0.195	0.198	0.275	–
between groups				0.027	0.026	0.026	–
Total	**100.0**	**4156**	**1.000**	**0.222**	**0.224**	**0.301**	**0.357**

Source: data from the Household Budget Survey 2002, author's calculations.

Finally, decomposition by counties is presented in Table 6 (Tallinn as the capital separately). The highest incomes are characteristic of Tallinn, the Harju County (in the neighbourhood of the former), and the Tartu County. The lowest incomes are observed in South-East and North-East Estonia. However, inequality is not distributed in the same way. Beside the three richest areas, the Gini coefficient is the highest in the counties of Viljandi and Järva, and the lowest in the counties of Saare, Rapla, and Ida-Viru.

Additionally, the ranking of counties by different inequality measures differs. For instance, *GE*(2) implies that inequality is highest in the counties of Viljandi and Võru, and lowest in the Pärnu County. This means that the gap between the richest and the average households is not the largest in the wealthiest areas. However, a conclusion can be drawn that incomes and inequality are somewhat lower in the peripheries than in the centres and neigh-

bouring areas of the latter. This also emphasizes the importance of human capital as this is concentrated in the centres and the people's educational level in the rest of the regions is probably lower and less dispersed.

Table 6. Decomposition of income inequality by county

County	% of population	Mean income	Relative mean	GE(0)	GE(1)	GE(2)	Gini
Jõgeva	2.8	2,678	0.644	0.209	0.166	0.187	0.314
Ida-Viru	13.1	2,953	0.710	0.132	0.136	0.169	0.279
Valga	2.6	3,081	0.741	0.169	0.163	0.184	0.315
Põlva	2.4	3,155	0.759	0.202	0.190	0.226	0.332
Võru	2.9	3,203	0.771	0.179	0.206	0.325	0.323
Lääne	2.1	3,581	0.862	0.209	0.199	0.270	0.325
Hiiu	0.8	3,709	0.893	0.164	0.161	0.196	0.300
Viljandi	4.2	3,723	0.896	0.199	0.224	0.326	0.346
Lääne-Viru	5.0	3,730	0.897	0.161	0.165	0.232	0.297
Järva	2.8	3,758	0.904	0.200	0.212	0.293	0.345
Pärnu	6.7	3,842	0.924	0.166	0.145	0.155	0.296
Saare	2.6	3,873	0.932	0.144	0.141	0.157	0.295
Rapla	2.7	4,062	0.977	0.177	0.153	0.175	0.295
Tartu	11.0	4,139	0.996	0.212	0.198	0.229	0.346
Harju	9.0	5,123	1.233	0.255	0.231	0.270	0.372
Tallinn	29.4	5,140	1.237	0.235	0.237	0.319	0.368
within groups				0.200	0.203	0.280	–
between groups				0.022	0.022	0.021	–
Total	**100.0**	**4,156**	**1.000**	**0.222**	**0.224**	**0.301**	**0.357**

Source: data from the Household Budget Survey 2002, author's calculations.

With each variable, the income inequality within groups is much higher than the inequality between groups (as was also suggested by Table 2). These variables are mostly related to human capital and its productivity in the previously considered theoretical framework. Therefore, differences in employment and educational level are more important for explaining inequality than the factors characterizing the labour market (discrimination by gender and nationality) and the structure of households (type and size).

Conclusions

Economic inequality is an important aspect of economic development. The study of the sources and consequences of inequality, and its dynamics over time can give valuable insights for understanding and evaluating the overall economic and social developments. This article dealt with several issues of income inequality in Estonia. In order to find appropriate methods for assessing the extent and sources of inequality, some theoretical aspects were examined.

Firstly, the approach was limited by defining economic inequality as a difference between the actual and absolutely equal distribution. Additionally, because of the data constraints, economic inequality was understood as income inequality.

The theoretical discussion led to identification of the major factors affecting income inequality. On the one hand, the imperfectness of capital and labour market can increase the inequality, on the other hand, the tax system and general social norms usually limit or reduce it. These results were attained from the analysis of three stages in the process of income generation — those of factor endowments, factor productivity and income redistribution.

Secondly, different inequality measures and their properties were analyzed. The current literature proposes several standard axioms which look for consistent estimates of inequality. As only the

general entropy measures comply with these axioms, three measures from this group (the mean logarithmic deviation, the Theil index and half the squared coefficient of variation), and the Gini coefficient were chosen for an empirical analysis. Although the Gini coefficient does not satisfy all of the axioms, it has an intriguing interpretation in relation with the Lorenz curve.

These measures were applied on the data of the Estonian household budget surveys for the period from 1996 to 2002. The results are somewhat limited. Although in general it can be concluded that the income inequality has grown over the period, the increase has not been uniform. There was a considerable decline in 2001 and a smaller one also in 1999. The changes in the data collection methods seem to have influenced the results. Different measures suggest a similar pattern, except for the decline in 2001, when the decrease was more significant for the changes in the upper part of income distribution according to half the squared coefficient of variation.

A comparison of Estonia with some other European countries revealed that its income inequality is relatively high. Among the sample countries, only Russia exceeded Estonia. Additionally, the international comparison implied that the transition process itself is not necessarily involved in the increasing inequality, for in several other transitional countries the level of inequality is relatively low. However, this conclusion is not strong because of the lack of corresponding time series.

The further empirical discussion focused on several social and demographic factors possibly related to income inequality. Three entropy measures were decomposed on the basis of ten different characteristics of households. The decomposition showed that the most important sources of income inequality are associated with one's employment status and educational level. More generally, income inequality in Estonia results mainly from educational and regional disparities. In general, this accords with the theoretical insights that the factor endowments and factor productivity play a

significant role in the genesis of income inequality. Unfortunately, the data were too limited to assess the impact of redistribution. However, an important implication for policies is to look for a more equal provision of education-related public goods in order to achieve a more equal distribution of human capital, and to focus more on regional differences in order to enable a more productive use of the factor endowments.

References

Anand, S. The Measurement of Income Inequality. — Measurement of Inequality and Poverty, ed. S. Subramanian. Oxford: Oxford University Press, 1997, pp. 81–105.

Atkinson, A. B., Stiglitz, J. E. Lectures on Public Economics. Singapore: McGraw-Hill Book Co, 1980, 619 pp.

Barro, R. J. Inequality, Growth, and Investment. NBER Working Paper, 1999, No. 7038, 52 pp.

Champernowne, D. G. A Comparison of Measures of Inequality of Income Distribution. — The Economic Journal, 1974, Vol. 84, Is. 336, pp. 787–816.

Chu, K.; Davoodi, H.; Gupta, S. Income Distribution and Tax and Government Social Spending Policies in Developing Countries. IMF Working Paper, 2000, No. 62, 48 pp.

Cowell, F. A. Measurement of Income Inequality. — Handbook of Income Distribution. Volume 1, ed. A. B. Atkinson, F. Bourguignon. Amsterdam: North-Holland, 2000, pp. 87–116.

Cowell, F. A. Measuring Inequality. London: Prentice Hall & Harvester Wheatsheaf, 1995, 194 pp.

Gottschalk, P.; Smeeding T. M. Empirical Evidence on Income Inequality in Industrialized Countries. LIS Working Paper, 1999, No. 154, 76 pp.

Kuznets, S. Economic Growth and Income Inequality. — The American Economic Review, 1955, Vol. 45, Is. 1, pp. 1–28.

Household Budget Surveys, 1996-2002. Statistical Office of Estonia.

Litchfield, J. Inequality: Methods and Tools. World Bank, 1999, 16 pp. [http://www.worldbank.org/poverty/inequal/methods/litchfie.pdf]. 30/01/2004.

Luxembourg Income Study (LIS) Key Figures. [http://www.lisproject.org/keyfigures.htm]. 30/01/2004.

Mookherjee, D.; Shorrocks, A. A Decomposition Analysis of the Trend in UK Income Inequality. — The Economic Journal, 1982, Vol. 92, Is. 368, pp. 886–902.

Sen, A. K. On Economic Inequality. Oxford: Clarendon, 1997, 260 p.

Shorrocks, A. F. Ranking Income Distribution. — Economica, 1983, Vol. 50, Is. 197, pp. 3–17.

Tanzi, V. Fundamental Determinants of Inequality and the Role of Government. IMF Working Paper, 1998, No. 178, 24 pp.

Wagschal, U. Income Distribution, Inequality and Unemployment. LIS Working Paper, 1997, No. 152, 26 pp.

Appendix 1. Analytical measures of income inequality

Method	**Definition**	**Pigou-Dalton principle**	**Additively decomposable**	**Income and population scale invariance**
Relative mean deviation	$M = \frac{1}{n\mu}\sum_{i=1}^{n}\lvert\mu - y_i\rvert$	No	No	Yes
Variance	$\sigma = \frac{1}{n}\sum_{i=1}^{n}(y_i - \mu)^2$	Strong	Yes	No
Coefficient of variance	$CV = \sqrt{\sigma}/\mu$	Weak	No	Yes
Variance of logarithms	$v_1 = \frac{1}{n}\sum_{i=1}^{n}(\log y_i - \log\tilde{\mu})^2$	No	No	Yes
Logarithmic variance	$v = \frac{1}{n}\sum_{i=1}^{n}(\log y_i - \log\mu)^2$	No	No	Yes
Gini coefficient	$G = \frac{1}{2n^2\mu}\sum_{i=1}^{n}\sum_{j=1}^{n}\lvert y_i - y_j\rvert$	Weak	No	Yes
Generalized entropy measure	$GE(\alpha) = \frac{1}{n}\frac{1}{\alpha(\alpha-1)}\sum_{i=1}^{n}\left[\left(\frac{y_i}{\mu}\right)^{\alpha} - 1\right]$	Strong	Yes	Yes
Mean logarithmic deviation	$GE(0) = L = \frac{1}{n}\sum_{i=1}^{n}\log\frac{\mu}{y_i}$	Strong	Yes	Yes
Theil index	$GE(1) = T = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\mu}\log\frac{y_i}{\mu}$	Strong	Yes	Yes
Half the squared CV	$GE(2)=CV^2/2$	Strong	Yes	Yes

Appendix 1 (continued)

Method	**Definition**	**Pigou-Dalton principle**	**Additively decomposable**	**Income and population scale invariance**
Dalton's measure	$D = 1 - \frac{1}{nU(\mu)} \sum_{i=1}^{n} U(y_i)$	Weak	Yes	No
Atkinson's measure	$A(\varepsilon) = 1 - \left[\frac{1}{n} \sum_{i=1}^{n} \left(\frac{y_i}{\mu} \right)^{1-\varepsilon} \right]^{\frac{1}{1-\varepsilon}}$	Weak	Yes	Yes

Note: n denotes the number of individuals, y_i (i = 1,..,n) the income of i-th individual; the arithmetic mean is $\mu = \frac{1}{n} \sum_{i=1}^{n} y_i$ and the geometric mean $\tilde{\mu} = \exp(\frac{1}{n} \sum_{i=1}^{n} \log y_i)$. *Log* is for natural logarithm.

Source: Cowell 1995, p. 66, 139–140; author's supplements.

Appendix 2. Statistics of Household Budget Surveys in 1996–2002

Year	Household income observations				Weighted observations	
	pos	neg	zero	total	households	individuals
1996	11,228	59	39	11,326	592,271	1,416,321
1997	9,999	43	38	10,080	603,927	1,396,396
1998	9,602	35	34	9,671	614,320	1,382,924
1999	8,280	26	28	8,334	594,133	1,369,039
2000	6,028	36	4	6,068	561,754	1,361,688
2001	5,826	24	4	5,854	562,689	1,356,395
2002	5,481	8	11	5,500	561,460	1,350,538

Source: HBS 1996–2002, author's calculations.

Appendix 3. Disposable income per household member in 1996–2002

Year	Weighted	Average	St.dev.	Min	Max
1996	1,405,489	1,449	1,213	6	28,650
1997	1,385,803	1,659	1,985	3	110,938
1998	1,372,652	1,931	1,644	3	24,121
1999	1,360,553	2,031	1,729	8	50,284
2000	1,352,773	2,208	1,932	17	39,760
2001	1,350,887	2,301	1,811	38	24,102
2002	1,347,076	2,495	2,004	0	31,610

Note: Only incomes with positive values used.
Source: data from Household Budget Surveys 1996–2002, author's calculations.

Appendix 4. Equivalent income per household member in 1996–2002

Year	Weighted	Average	St.dev.	Min	Max
1996	1,405,489	2,389	1,836	24	33,991
1997	1,385,803	2,680	2,148	27	37,000
1998	1,372,652	3,125	2,578	31	36,343
1999	1,360,553	3,298	2,670	33	44,120
2000	1,352,773	3,666	3,067	37	49,760
2001	1,350,887	3,824	2,947	75	31,830
2002	1,347,076	4,156	3,225	42	44,703

Note: Only incomes with positive values used.
Source: data from Household Budget Surveys 1996–2002, author's calculations.

Appendix 5. Decomposition variables

Social group of household 1. One employed household member 2. Two or more employed household members 3. Household of unemployed people 4. Household of retired people 5. Other inactive household
Employment status of the head of household 1. Employed 2. Unemployed 3. Other active 4. Retired 5. Other inactive
Educational level of the head of household 1. Without primary 2. Primary 3. General secondary 4. Special secondary 5. Higher
County Tallinn (the capital), Harju, Hiiu, Ida-Viru, Jõgeva, Järva, Lääne, Lääne-Viru, Põlva, Pärnu, Rapla, Saare, Tartu, Valga, Viljandi, Võru
Age of the head of household (by the 1st of January 2002) 1. Younger than 26 years 2. 26–35 years 3. 36–45 years 4. 46–55 years 5. 56–65 years 6. Over 65 years
Type of household (a child is younger than 16 years by the 1st of January 2002) 1. An adult 2. A single parent 3. A childless couple 4. A couple with a child 5. A couple with 2 children

6. A couple with 3 or more children 7. Other
Number of household members 1. One 2. Two 3. Three 4. Four or more
Nationality of the head of household 1. Estonian 2. Non-Estonian
Type of settlement 1. City 2. Town 3. Settlement 4. A bigger village (having a school, shop or post office) 5. Village, a detached farm
Gender of the head of household 1. Male 2. Female

Source: Household Budget Survey of Estonia in 2002, author's modifications.