The Prince and the Pauper Revisited: A Bootstrap Approach to Poverty and Income Distribution Analysis Using the PACO Data Base [1]

Georges A. HEINRICH (CERT, Dept. of Economics, Heriot-Watt University)

Summary

In this paper, we are using bootstrap methods to perform statistical inference for poverty and inequality indices. Using 1992 cross-sectional household data from the Panel Comparability Project (PACO), we show that the probability that a household headed by a retired person is poor is significantly higher in the United Kingdom than in Hungary or Luxembourg. We also show that the average shortfall from the poverty line is much larger for poor households in Hungary and the United Kingdom than for poor households in Luxembourg. This suggests that poverty spells are also longer in Hungary and the United Kingdom than they are in Luxembourg.

1. Introduction

In the field of poverty analysis, most of the attention focuses on the identification and aggregation problems [Sen (1976), page 219]. Statistical inference for poverty and inequality measures, on the other hand, is widely ignored.

Conclusions about poverty and the distribution of incomes however are typically based on information obtained from sample surveys. These sample surveys are subject to sampling and non-sampling errors. Sampling errors are those errors that are due to the fact that we observe only a sub-set of the total population. Statistical inference deals with sampling

[1] The data used in this study are from the public use version of the PACO datafiles, including data from the British Household Panel Study, the French (Lorraine) Household Panel, the German Socio-Economic Panel, the Hungarian Household Panel, the Luxembourg Household Panel, the Polish Household Panel and the US Panel Study of Income Dynamics. The comparable variables in this datafile were created by the PACO project, co-ordinated by CEPS/INSTEAD in Differdange/Luxembourg.

This paper was made possible by financial support from the Ministry of Education, Luxembourg (BFR 96/064), for which the author is most grateful.

Helpful comments and suggestions were received from Gary Koop, Mark Schaffer and Geoff Wyatt.

errors and allows us to determine whether the estimated poverty and inequality measures represent the true population parameters.

The problem can be formulated as follows. We have obtained a random sample $X = (x_1, x_2,\ldots, x_n)$ from an unknown probability distribution F and we want to estimate a parameter (e.g., a headcount poverty index) $q = t(F)$ on the basis of X. We calculate an estimate $\hat{\theta} = s(X)$ using X. The question we are examining in this paper is: How accurate is $\hat{\theta}$?

Furthermore, we are often interested in analysing poverty and inequality in a dynamic or a cross-country context in order to find answers to questions like: Have poverty and inequality increased in country A over time? or Does country B have a higher incidence of poverty than country C? Thus, we would like to determine the statistical significance of changes in poverty or income distribution indices: $DIFF = P_1 - P_2$. Hypotheses tests conducted on a test statistic *DIFF* involves comparing means from two distributions. In the literature, this is known as the *Fisher-Behrens* problems. Notice that within the classical hypothesis testing framework, there is no straightforward solution to this problem.

In this paper, we will demonstrate the usefulness of bootstrapping techniques for carrying out statistical inference for poverty and inequality measures.

We will analyse poverty and inequality among pensioners in Hungary, Luxembourg and the United Kingdom. We will also compare the results we obtain for households headed by a retired person to the results obtained for households headed by an economically active person.

In Hungary, Luxembourg and the United Kingdom, public pay-as-you-go pension plans remain in place. In addition to providing income insurance in old age, these programs have the unique power to transfer income from the lifetime rich to the lifetime poor. The World Bank [(1994), page 101, emphasis added] argues that "This is their big advantage over other financing arrangements for old age security, and *their success in achieving this poverty alleviation objective may be taken as the litmus test of a well-functioning public plan*".

The problem of poverty in old age is exacerbated by demographic ageing: as the number of elderly and retired people in society grows, there will be an increasing pressure on scarce resources for social security and assistance. Demographic ageing, early retirement and generous pension benefits are deteriorating the financial position of public old age pension systems. In the medium to long term, demographic and system dependency ratios will rise. Pension expenditures relative to GDP will increase and there will be considerable pressure on public finances.

Tsakloglou [(1996), page 272] points out that "A number of policy recommendations aimed at tackling these problems have been suggested in recent policy debates.

Implementation of some of these recommendations is not expected to affect dramatically the living standards of the elderly citizens […] and therefore no serious social unrest is expected. […] However, a policy recommended strongly in some quarters recommends cuts in pensions and other social security benefits targeted towards the elderly. Implementation of such a policy is likely to affect adversely the living standards of large segments of the elderly and to provoke their negative reaction". Thus, we believe that plans to reform public pensions in these countries should be preceded by a careful analysis of the living conditions of the retired.

In this paper, we will use data from the Panel Comparability Project (PACO) to analyse poverty and inequality in Hungary, Luxembourg and the United Kingdom. Section 2 of the paper discusses methodological issues related to poverty analysis. In section 3, we stress the need for performing statistical inference on poverty and inequality indices. In section 4, the bootstrap method is described. Section 5 presents the results from the empirical analysis of poverty and inequality. Finally, section 6 concludes.

2. Poverty analysis and the identification problem

Following Sen [1976], poverty analysis can be broken down into two stages: identifying the poor in the population and summarising this information in a poverty index. The traditional approach to the identification problem involves the use of poverty lines and equivalence scales. Unfortunately, there exists no consensus among researchers as to what constitutes the appropriate poverty line or equivalence scale.

A poverty line is a pre-defined cut-off point for income. Households with income equal to or above the poverty line are non-poor while households with income below the poverty line are poor. Callan and Nolan [1991] present a comprehensive survey of operative poverty lines and outline the advantages and disadvantages of each method.

Following Goedhart et al. [1977], there are four general types of poverty lines: *absolute* poverty lines, *official* poverty lines, *subjective* poverty lines and *relative* poverty lines. Using an absolute poverty line, we identify a group of commodities necessary for the subsistence of the individual. The poverty line is then defined as the minimal amount of money that enables the individual to purchase this commodity bundle. The poverty line may also be defined in relation to a government transfer aimed at income maintenance payment (e.g., an unemployment benefit, a minimum pension benefit). This type of poverty line is called an official poverty line. In the case of a subjective poverty line, individuals are asked directly to identify the minimum level of resources necessary to reach a certain standard of living. The poverty line is then constructed using this information on the expressed preferences of the respondents. Finally, the relativist approach to poverty measurement defines the poverty line as a fraction of the median or mean welfare of society.

For the purposes of this paper, we are using the latter concept i.e., a relative poverty line. Relative poverty lines are widely accepted as a tool of analysis for poverty in developed economies. We have chosen 50 percent of the median income per equivalent adult as the cut-off point. This poverty line is widely used in empirical work [see e.g., Blackburn (1994)]. Following Callan and Nolan [(1991), page 253], "The general rationale [for the use of a relative poverty line] is that those falling more than a certain 'distance' below the average or normal income level in the society are unlikely to be able to participate fully in the life of the community".

The distinction between absolute and relative poverty is particularly relevant from a policy point of view: Are we primarily concerned with the standard of living of those who receive low incomes or are we concerned with the unequal distribution of these incomes? Absolute poverty is eliminated by making everybody better off i.e., by shifting the income distribution upwards. Relative poverty, on the other hand, is eliminated by redistributing income from the rich to the poor.

The second choice pertains to the choice of an equivalence scale. Households are differing according to their socio-demographic characteristics e.g., size, composition by age, number of dependants, place of residence. The needs of the household members are likely to vary with these characteristics: children and adults have different needs, women have different needs than men and the cost of living is usually higher in urban areas than in rural areas. If we do not take these differences into account, our poverty estimates are likely to be biased.

A simple headcount measure also ignores economies of scale in consumption. This is obvious for commodities like rent or heating but economies of scale may also exist for other commodities like e.g., food. If economies of scale in consumption exist, the marginal cost of an additional household member is not constant but decreasing. The approach commonly used to take into account differences in needs and economies of scale is to standardise the income data by using an equivalence scale factor.

Following Deaton and Muellbauer [1980], the equivalence scale m_k is defined by:

$$m_k = m\left(P, u, A, A_k\right) = \frac{C\left(P, u, A_k\right)}{C\left(P, u, A_r\right)}$$

where P is a vector of prices, u is the utility level, A_t is a vector of demographic attributes of household t and C is a cost function. Thus, the equivalence scale is the ratio of the costs required to achieve a given utility level for households of differing compositions.

Equivalence scales used in applied work and for policy purposes are extremely varied in how they allow for differences in family composition. There is no consensus as to what scale should be used and equivalisation is often criticised for of its *ad hoc* nature [see Nelson (1993)]. Following Buhman et al. [(1988), page 119], these differences are

summarised by a single parameter: the family -or household- size elasticity of needs. Household incomes are standardised as follows:

$EI = D/S^e$

where *EI* is the equivalent income per household member, *D* is the household disposable income, *S* is the household size and *e* is the elasticity of the scale which varies between 0 and 1. If *e = 0*, there are perfect economies of scale i.e., the cost for a household of two to achieve a given utility level is the same as the cost for a single-person household. If *e = 1*, there are no economies of scale i.e., a household of two needs twice as much income to achieve a given utility level than a single-person household.

The equivalence scale used in this paper is referred to in the literature as the "international experts' scale" [see Burkhauser et al. (1997)] and uses an elasticity scale factor *e = 0.5*.

Our approach to poverty measurement is arguably a simplistic one. Pryke [1995] provides a comprehensive critique of the poverty line and equivalence scale approach to poverty analysis. However, we shall adopt Sen's [(1973), page 78] dictum and avoid "the danger of falling prey to a kind of nihilism [which] takes the form of noting, quite legitimately, a difficulty of some sort, and then constructing from it a picture of total disaster". Notwithstanding, we do believe that it is very important to state these limiting assumptions explicitly and to take them into account when analysing and discussing our results.

3. Statistical inference for poverty and inequality indices

The interest of this paper centres on properties of the distribution of incomes in three countries: Hungary, Luxembourg and the United Kingdom. The respective probability distributions of the populations are unknown but we have obtained three cross-sectional samples with observations drawn from these distributions.

Based on the information contained in the samples, our objective is to draw inferences about the properties of the distribution of incomes at the population level. This is a standard problem of statistical analysis. The sample mean and sample standard deviation are used to construct a point estimate and a confidence interval for the unknown population mean i.e., $\hat{m} = \bar{x}$ and $[\bar{x} - 1.96\, s\,;\, \bar{x} + 1.96\, s]$ for a 95 percent confidence interval.

This is where the problems with classical statistical inference begin. Classical inference assumes a normal distribution and a simple function of interest (e.g., the population mean μ), so that the form of the confidence interval is known.

However, in the present context, we do not want to assume normality of the distribution. In addition, the functions of interest -poverty and income distribution indices- are non-

linear functions of income with the added complication that they are usually bounded. Thus, the form of the confidence interval is unknown.

There are three general types of solutions to this problem. The first solution is simply not to compute sampling variances and confidence intervals, assuming they will be small. However, Maasoumi [(1994), page 14][2] argues:

"[the argument that] in this area we often deal with large samples which do not justify too much concern for precision (sampling variance) … is occasionally contradicted by large standard errors, and it may be turned around in order to justify reporting even more statistical measures of precision and tests. This is because almost all of the useful statistical theory in this area is based on asymptotic approximations which are supposed to do well with large samples."

The second solution is to take advantage of the fact that we are working with large samples and to calculate asymptotic variances and confidence intervals. The asymptotic distribution provides an approximation to the true distribution. Asymptotic theory essentially assumes normality. Bishop et al. [1997], Rongve [1997], Bishop et al. [1995] and Kakwani [1993] have all used asymptotic methods to obtain consistent estimates of the variance-covariance structure of poverty measures in order to carry out statistical inference on the resulting estimates.

However, Mills and Zandvakili [(1997), page 134] point out:

"…the interval estimates available from asymptotic theory may not be accurate and the small sample properties of these intervals are not known. Further, all the decomposable inequality measures used in the literature are bounded (e.g., the Gini coefficient lies in the [0,1] interval], so the application of standard asymptotic results may lead to estimate intervals that extend beyond the theoretical bounds of a particular measure (e.g., a negative lower bound for Gini)".

The last solution to the problem is to carry out distribution-free inference. We have argued that classical inference assumes a normal distribution. As we do not possess any information about the distribution of the population, we would feel at best uneasy if we assumed normality. Distribution-free inference has the advantage that no prior knowledge about the distribution function of the population is required.

One such distribution-free method is bootstrapping and this is the method that we will use in this paper. The bootstrap is a computer-intensive method for estimating the standard error of a parameter. It will be described in greater detail in the next section. Following Sitter [(1992), page 136]: "Bootstrap methods reutilize the existing estimation system repeatedly, using computing power to avoid theoretical work". Mills and Zandvakili [1997] are using the bootstrap to carry out statistical inference for inequality measures.

[2] As quoted by Mills and Zandvakili [(1997), page 133].

Notice that when comparisons are the major focus of the analysis, we may reach valid conclusions about the direction of change in poverty and inequality by comparing the two distributions directly. Atkinson [1987] and Foster and Shorrocks [1988] have pioneered this approach. They advocate the use of dominance conditions in order to make inferences about changes in poverty over time or across countries.

The basic idea underlying the argument is that there is no widespread agreement as to what is the appropriate level for the poverty line. Thus, the poverty line may vary over a range $[z^-;z^+]$. First-order stochastic-dominance comparisons of income distributions over time or across countries then involves setting multiple poverty lines z in the range $[z^-; z^+]$ in order to determine whether we obtain the same poverty rankings for all the z's.

Examples of this kind of statistical inference are Blackburn [1994] for cross-country comparisons and Zheng et al. [1995] for inter-temporal comparisons. Anderson [1996] extends the theoretical literature in this field by proposing a non-parametric test of stochastic dominance in income distributions.

4. The bootstrap principle

Bootstrapping is based on re-sampling with replacement. Each bootstrap sample is an independent random sample of size n from the empirical distribution $\hat{F}$. The elements of the bootstrap sample are the same than those of the original data set. Some may appear only once in the bootstrap sample, some two or more times while some others may appear zero times. To each bootstrap sample corresponds a bootstrap replication of $\hat{\theta}$:

$$\hat{q}^* = s(x^*)$$

The bootstrap replication is the result of applying the same function *s(.)* to x^* as was applied to x. Following Efron and Tibshiriani [(1993), page 47] the bootstrap algorithm for estimating the standard error of a parameter is summarised by the following three steps:

Step 1: Select B independent bootstrap samples x^{*1}, x^{*2},….,x^{*B} , each consisting of n data values drawn with replacement from x.

Step 2: Evaluate the bootstrap replication corresponding to each bootstrap sample,

$$\hat{q}^*(b) = s\left(x^{*b}\right) \qquad b = 1,2,\ldots,B.$$

Step 3: Estimate the standard error by the sample standard deviation of the B replications,

$$\hat{se}_B = \left\{ \sum_{b=1}^{B} \left[\hat{\theta}^*(b) - \hat{\theta}^*(.) \right]^2 / (B-1) \right\}^{1/2},$$

$$where \quad \hat{\theta}^*(.) = \sum_{b=1}^{B} \hat{\theta}^*(b) / B.$$

It can be shown that the following result holds:

$$\lim_{B \to \infty} \hat{se}_B = se_{\hat{F}} = se_{\hat{F}}(\hat{\theta}^*)$$

i.e., the empirical standard deviation approaches the population standard deviation as the number of bootstrap replications grows large.

How many bootstrap replications are necessary in order to obtain a robust estimate of the standard error? There is a total of $\binom{2n-1}{n}$ distinct bootstrap samples on which the function *s(.)* can be evaluated. When we are dealing with very small samples, we may be able to compute *s(.)* for all the distinct bootstrap samples. Notice however that a sample as small as *n = 10* already yields 92.378 distinct bootstrap samples. For samples of the sizes with which we are working here, the total number of distinct bootstrap samples is very large indeed.

In fact, the real constraint on the number *B* of bootstrap replications is computer time, which increases linearly with *B*. Based on their experience, Efron and Tibshiriani [1993] propose the following set of rules of thumb: even a small number of bootstrap replications *(B = 25)* is usually informative. *B = 50* is often enough to yield a good estimate of $se_F(\hat{q})$. Very rarely are more than *B = 200* replications necessary. However, much bigger values *(B > 1000)* are required if we want to obtain bootstrap confidence intervals. For the purposes of this paper, we are using *B = 2000*.

Like asymptotic methods, bootstrapping is also an approximate method. However, and unlike asymptotic methods, bootstrapping attempts to obtain small sample results. In practice, bootstrapping seems to work very well i.e., it yields a correct confidence interval. However, theory is in its infancy and the justification for the good performance of the bootstrap is asymptotic.

Notice that independence of observations is a *sine qua non* condition for the bootstrap to be valid. If this condition is violated, difficulties arise. This precludes the use of the simple bootstrap for statistical inference using complex survey data. Variants of the bootstrap method may be used in those cases. The Rescaling Method, the Mirror-Match Method and the Without-Replacement Bootstrap (BWO) are discussed by Sitter [1992] who also proposes extensions of the BWO method which makes the method applicable to data

obtained through stratified sampling, two-stage cluster sampling and unequal-probability sampling.

Finally, the bootstrap can be used for hypothesis testing. In fact, it turns out to be a rather powerful tool of analysis in that respect. As Mills and Zandvakili [(1997), page 134] point out: "Further, since bootstrap intervals computed using the percentile method have a clear Bayesian interpretation, they provide a straightforward solution to the Behrens-Fisher problem of comparing means from two distributions".

The bootstrap confidence intervals reported in this paper were computed using the percentile method. As the name already suggests, this procedure is based on the percentiles of the histogram of bootstrap replications. A fuller exposition of the percentile method is provided in appendix A1.

5. Bootstrapping poverty and inequality indices for Hungary, Luxembourg and the United Kingdom in 1992

In this section, we will use 1992 cross-sectional Panel Comparability Project (PACO) household files in a cross-country study of poverty and inequality in Hungary, Luxembourg and the United Kingdom. The PACO is a comparative cross-national and longitudinal data base. It contains harmonised and consistent variables and identical data structures for each country included.

Poverty and income distribution are analysed at the household level. A household is defined to be all persons living under the same roof, sharing income and expenditures. We have chosen two types of households for this study: households with a head who is retired and household with a head who is working.

Poverty is measured in terms of disposable income per equivalent adult household member[3]. Much has been written on the question whether poverty is best comprehended in terms of income or consumption. The arguments in this debate are of a philosophical rather than an economic nature. Is poverty, for instance, the result of inequality in opportunities or inequalities in outcomes? In the former case income is a more appropriate proxy of welfare while in the latter case, consumption should be used.

The criterion which made us chose income rather than consumption to approximate the standard of living of households was a pragmatic one: information on household consumption is not available in the PACO data base.

In Table 1, we report bootstrap standard errors and confidence intervals for poverty and inequality indices.

[3] The particular income concept retained is the PACO variable hxx053 (total gross household income), after suitable standardisation using equivalence scales.

Sachs [(1984), page 273] argues that: "A comparison of two parameters is possible in terms of their confidence intervals: (1) If the confidence intervals intersect, it does not necessarily follow that the parameters do not differ significantly. (2) If the confidence intervals do not intersect, there is at the given significance level a genuine difference between the parameters".

But the bootstrap procedure also allows us to obtain tail probability values for hypothesis tests directly from the bootstrap distribution. In particular, we are interested in paired comparisons of poverty and inequality between Hungary, Luxembourg and the United Kingdom. In Tables 2, 3 and 4, we provide estimates of the standard errors of the observed cross-country differences and we compute the probability that the differences in poverty and inequality indices are significantly different from zero.

Finally, in Table 5, we report results for similar within-country comparisons between households with a retired and households with a working head.

As we are using a relative poverty line as the low-income cut-off point, income inequality and poverty are intrinsically linked. For instance, in the case of an highly unequal distribution of incomes, outliers at the top end of the income distribution will exert upward pressure on the mean income and henceforth push more people into poverty, relatively speaking. More generally, a generalised improvement in living conditions that is shared equally by all income groups leaves poverty unchanged. Likewise, a general decline in living standards does not lead to additional poverty if the relative income positions remain unaffected. Hence, we will also report inequality of incomes indices in our result tables. The formulas used to compute the poverty and inequality indices are summarised in appendix A2.

Furthermore, poverty indices are very sensitive to the choice of poverty line. In appendix A3, we report point estimates for the headcount, FGT and Sen poverty indices using poverty lines ranging from two thirds to 30 percent of equivalised median income. In the core of the text, we will focus exclusively on the results obtained for the 50 percent of median income poverty line.

Table 1: Bootstrap standard errors and confidence intervals for poverty and inequality measures (1992

	Hungary		Luxembourg		United Kingdom	
	retired	*working*	*retired*	*working*	*retired*	*working*
Headcount index	0.124	0.029	0.099	0.021	0.344	0.040
Standard deviation	0.012	0.005	0.011	0.004	0.014	0.004
95% $_{lower}$	0.101	0.019	0.078	0.013	0.316	0.033
95% $_{upper}$	0.148	0.040	0.121	0.029	0.372	0.047
FGT index	0.132	0.071	0.032	0.011	0.370	0.054
Standard deviation	0.032	0.023	0.009	0.004	0.036	0.010
95% $_{lower}$	0.074	0.033	0.016	0.004	0.303	0.037
95% $_{upper}$	0.202	0.123	0.053	0.019	0.445	0.074
Sen index	0.026	0.010	0.013	0.004	0.086	0.011
Standard deviation	0.005	0.003	0.002	0.001	0.005	0.002
95% $_{lower}$	0.018	0.005	0.009	0.002	0.076	0.009
95% $_{upper}$	0.036	0.016	0.018	0.006	0.097	0.015
Gini coefficient	0.244	0.265	0.243	0.245	0.356	0.307
Standard deviation	0.008	0.010	0.008	0.007	0.012	0.005
95% $_{lower}$	0.227	0.247	0.228	0.232	0.335	0.297
95% $_{upper}$	0.260	0.285	0.258	0.259	0.380	0.317
Theil index	0.046	0.057	0.043	0.046	0.106	0.069
Standard deviation	0.003	0.005	0.003	0.003	0.011	0.003
95% $_{lower}$	0.039	0.048	0.037	0.039	0.088	0.064
95% $_{upper}$	0.052	0.068	0.049	0.053	0.128	0.075
Rel. mean deviation	0.172	0.184	0.172	0.173	0.260	0.217
Standard deviation	0.006	0.007	0.006	0.005	0.009	0.004
95% $_{lower}$	0.160	0.172	0.162	0.163	0.243	0.210
95% $_{upper}$	0.184	0.198	0.183	0.183	0.278	0.224

The results in Table 1 reveal a very high incidence of relative poverty among retirees' households in the UK in 1992. On the basis of the headcount poverty index, we find that between 31.6 and 37.2 percent of these households had incomes below the poverty line in 1992. This is 20-25 percent higher than the corresponding rates for Luxembourg and the transition economy Hungary.

The high incidence of poverty among pensioners' households in the UK is confirmed by the FGT and Sen indices which take into account not only the number of poor people but also the average shortfall from the poverty line. The UK 95 percent confidence intervals for the three poverty indices exhibit no overlap with the intervals obtained for Hungary and Luxembourg. This is very strong evidence against the hypothesis of no significant difference in poverty rates between the UK and the other two countries, a result which is confirmed by the hypothesis tests in Tables 3 and 4.

While poverty among pensioners in the UK clearly exceeds that in Hungary and Luxembourg, greater care has to be taken when comparing the latter two countries. On the basis of the headcount point estimate, poverty among Hungarian retirees' households is only 2.5 percent higher than that among their Luxembourg counterparts and the confidence intervals overlap considerably. However, the hypothesis test in Table 2 shows that this 2.5 percent difference between headcount ratios is statistically significant.

Table 2: Hypotheses tests for poverty and income distribution measures, cross-country comparison between Luxembourg and Hungary, 1992

		retired	working
Headcount index	*DIFF*	-0.025	-0.008
	standard deviation	0.001	0.001
	p-value	0.000	0.000
FGT index	*DIFF*	-0.100	-0.061
	standard deviation	0.023	0.019
	p-value	0.000	0.000
Sen index	*DIFF*	-0.013	-0.006
	standard deviation	0.002	0.002
	p-value	0.000	0.000
Gini coefficient	*DIFF*	-0.012	-0.020
	standard deviation	0.000	0.003
	p-value	0.018	0.000
Theil index	*DIFF*	-0.003	-0.012
	standard deviation	0.000	0.002
	p-value	0.000	0.000
Rel. mean deviation	*DIFF*	0.001	-0.012
	standard deviation	0.001	0.002
	p-value	0.141	0.000

NOTE: if *DIFF<0*, poverty/inequality are higher in Hungary than in Luxembourg

A clearer picture emerges, however, when we look at the FGT and Sen poverty indices. The 95 percent confidence intervals for Luxembourg and Hungary do not overlap for these two indices, indicating that Hungary has a significantly higher incidence of poverty among pensioners' households than Luxembourg. It also suggests that poverty, when it occurs, is "deeper" in Hungary. The results in Tables 2-4 confirm that we can safely reject the hypothesis of no difference between the poverty indices of the three countries.

We can see from Table 1 that the UK does not only exhibit a higher incidence of poverty among its pensioners' households but the distribution of incomes is also much more unequal. For all three inequality measures, there is no overlap between confidence intervals estimated for the UK and the other two countries, suggesting that we can safely reject the hypothesis of no difference between the income distributions. The income distributions for Luxembourg and Hungary, on the other hand, are looking very similar. Point estimates are very close to each other and confidence intervals overlap substantially. The results in Table 2 suggests that the cross-country differences for the Gini and Theil measures are statistically significant while the difference for the relative mean deviation is not significantly different from 0.

Table 3: Hypotheses tests for poverty and income distribution measures, cross-country comparison between Luxembourg and the United Kingdom, 1992

		retired	working
Headcount index	*DIFF*	-0.245	-0.019
	standard deviation	0.003	0.001
	p-value	0.000	0.000
FGT index	*DIFF*	-0.338	-0.043
	standard deviation	0.026	0.006
	p-value	0.000	0.000
Sen index	*DIFF*	-0.073	-0.007
	standard deviation	0.003	0.001
	p-value	0.000	0.000
Gini coefficient	*DIFF*	-0.113	-0.062
	standard deviation	0.004	0.002
	p-value	0.000	0.000
Theil index	*DIFF*	-0.063	-0.024
	standard deviation	0.008	0.001
	p-value	0.000	0.000
Rel. mean deviation	*DIFF*	-0.088	-0.045
	standard deviation	0.003	0.001
	p-value	0.000	0.000

NOTE: if *DIFF<0,* poverty/inequality are higher in the United Kingdom than in Luxembourg

The relative mean deviation of incomes has a straightforward interpretation in economic terms. It represents the income transfer from households above the mean income to households below the mean income necessary to achieve perfect equality of incomes. In Luxembourg and Hungary this transfer amounts to 17 percent while in the UK, this transfer amounts to 26 percent.

A very different picture of poverty -yet not inequality- emerges when we analyse poverty among households headed by a head who is working. Poverty rates, as measured by the headcount index, are very low and not very different from each other. The 95 percent confidence intervals overlap for Hungary and Luxembourg but results from hypothesis tests suggest that the difference is statistically significant. The United Kingdom has again a higher incidence of poverty but the welfare gap between UK households and households from the other two countries is much smaller than for households with a retired head. This offers a very stark contrast to the situation of pensioners' households in the United Kingdom.

Table 4: Hypotheses tests for poverty and income distribution measures, cross-country comparison between Hungary and the United Kingdom, 1992

		retired	working
Headcount index	*DIFF*	-0.220	-0.011
	standard deviation	0.002	0.002
	p-value	0.000	0.000
FGT index	*DIFF*	-0.238	0.017
	standard deviation	0.004	0.013
	p-value	0.000	0.087
Sen index	*DIFF*	-0.060	-0.001
	standard deviation	0.001	0.001
	p-value	0.000	0.179
Gini coefficient	*DIFF*	-0.112	-0.042
	standard deviation	0.003	0.005
	p-value	0.000	0.000
Theil index	*DIFF*	-0.060	-0.012
	standard deviation	0.008	0.002
	p-value	0.000	0.000
Rel. mean deviation	*DIFF*	-0.088	-0.033
	standard deviation	0.003	0.003
	p-value	0.000	0.000

NOTE: if *DIFF<0*, poverty/inequality are higher in the United Kingdom than in Hungary

When we take into account "depth" of poverty by looking at the FGT and Sen indices, it becomes clear that Luxembourg not only has fewer poor households but that these poor

households are quite significantly better off than those in the other two countries. Confidence intervals for the FGT index for Luxembourg do not overlap with those for Hungary and the United Kingdom but there is some overlap between the Sen confidence intervals for Luxembourg and Hungary. A hypothesis test performed on the difference between Sen indices shows, however, that this difference is statistically significant. Again, these figures suggest that poverty, when it occurs in Hungary and particularly in the United Kingdom is associated with much larger shortfalls from the poverty line than in Luxembourg and that poverty is henceforth much harder to escape.

While the conclusions with respect to poverty are similar for Hungary, Luxembourg and the UK, this is not the case for inequality. In fact, in Hungary, retirees' households exhibit a more equal distribution of incomes than households with a head who is working. The opposite is true in the United Kingdom: households with a retired head face a more unequal distribution of incomes. The case of Luxembourg is less clear-cut. Point estimates suggest that retirees' households' incomes are distributed more equally but for one inequality index, this difference is not statistically significant.

In Table 5, we compare directly poverty and inequality between households with a retired and households with a working head for each individual country. Our results show that in all three countries, households headed by a retired person are poorer than households headed by a person that is working.

This result is not surprising. A person who retires receives a retirement income which is a certain percentage of his or her final income and possibly also his or her average lifetime income. Almost always is the retirement income lower than the final income and hence we would expect pensioners to be, on average, worse off than workers or employees. Tullock [(1984), page 121] explains why the retired receive lower incomes: "If we go back to the period before social security and before the significant drive on the part of the government to get retirement ages down, we find that people tended to have declining incomes towards the end of their lives. Those who had some source of income other than work […] eventually reach the point where their preferred that other source of income, together with leisure, to working. Thus, as a general rule, people retired, their income went down, and they chose retirement at a lower income than their final working income because they preferred leisure. Thus, the custom that older people had lower incomes than people in the active phase of their lives became well established. There does not, however, seem to be any other reason for it".

Thus, one may wonder whether the observed differences between retired and active households reflect actual differences in their standards of living. Retired persons have more leisure at their disposal and a purely monetary welfare proxy does not take into account the value of leisure. However, a one-to-one relationship between income foregone and additional leisure can only be established in the case where retirement is based on a voluntary decision. As the retirement decision typically entails a discontinuous reduction of hours worked to zero -rather than a gradual reduction of hours- it is very difficult to know how much of the additional leisure is voluntary and how much is not.

Table 5: Hypotheses tests for poverty and income distribution measures, within-country comparisons b a retired and households with a working head

		Hungary	Luxembourg
Headcount index	*DIFF*	-0.098	-0.078
	Standard deviation	0.007	0.007
	p-value	0.000	0.000
FGT index	*DIFF*	-0.060	-0.021
	Standard deviation	0.010	0.006
	p-value	0.000	0.000
Sen index	*DIFF*	-0.016	-0.009
	Standard deviation	0.002	0.001
	p-value	0.000	0.000
Gini coefficient	*DIFF*	0.021	0.002
	Standard deviation	0.002	0.001
	p-value	0.000	0.010
Theil index	*DIFF*	0.011	0.003
	Standard deviation	0.002	0.000
	p-value	0.000	0.000
Relative mean deviation	*DIFF*	0.013	0.001
	Standard deviation	0.001	0.001
	p-value	0.000	0.284

NOTE: if *DIFF<0*, poverty/inequality are higher among households with a retired head than among households with a w

This raises a more general measurement issue. Young and old households may differ in terms of their needs and wants and an approach to poverty measurement which only uses money-based proxies of the standard of living may be inappropriate to carry out such comparisons.

But what about inequality? It is much more difficult to formulate any *a priori* expectations in this case. All three countries operate progressive tax systems which are supposed to redistribute incomes from the rich to the poor. However, this redistribution does not necessarily take the form of a pure money transfer from rich to poor. Furthermore, in all three countries, public pension schemes are supposed to redistribute income from high lifetime earners to low lifetime earners. Whether this means that pensioners should face a more equal distribution of incomes than non-pensioners is unclear. Tullock [(1984), page 107, emphasis added] argues that: "What we now think of as a welfare state was invented by Bismarck for the specific purpose of providing a politically viable way of beating off socialism. The early programs in Germany (the programs that everyone else has copied over the years) had rather diffuse and complicated distributional effects. *It is certain, however, that they were not intended to transfer funds from the middle class to the poor*".

However, what should be clear is that inequality among pensioners is much harder to tackle than inequality among economically active persons or households as pensioners' incomes are typically based on past earnings which are, once you retire, a constant rather than a variable.

6. Conclusions and policy implications

Panel Comparability Project (PACO) data was used to compute bootstrap standard errors and confidence intervals for poverty and inequality measures. We also used the bootstrap to perform hypothesis tests in order to establish the statistical significance of the observed within-country and cross-country differences for poverty and inequality indices.

We are comparing relative income positions for Hungary, Luxembourg and the United Kingdom in 1992 for households headed by retired and economically active heads. Our results suggest that the welfare state in the United Kingdom fails to provide adequate income insurance for the retired. More than one in three households headed by a retired person receives an income below the poverty line. By all accounts, poverty among retirees' households in the United Kingdom is 3 times higher than in Luxembourg and in Hungary.

Luxembourg and Hungary are relatively successful at protecting the retired from poverty. Headcount poverty rates in the two countries are in the 10 percent region. However, the average shortfall from the poverty line for the poor is much lower in Luxembourg than it is in Hungary. This result also holds for households headed by a working head. High FGT and Sen indices for Hungary and particularly the UK are worrying as they suggest that

poverty is deeply rooted. This makes it much harder to escape from poverty and poverty spells are therefore much longer.

The results we obtain for households with a working head suggest that the best way to escape from poverty is through an active involvement in the labour market. In fact, headcount poverty rates are less than 5 percent for these households in all three countries. Policymakers and politicians become increasingly aware of this fact. In the United Kingdom, for instance, the government contemplates far reaching reforms to the welfare state while emphasising that the best way out of poverty is work. However, such reforms are of little use to the retired whose involvement in the labour market has obviously ceased. Our results quite clearly show that this group of the population is at a particular risk from poverty. This is an important fact to bear in mind when discussing options for old-age pension reforms and reforms of the welfare state in general.

References

ANDERSON, Gordon (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions", *Econometrica*, vol. 64 no. 5 (September), pp. 1183-1193

ATKINSON, Anthony B. (1987), "On the Measurement of Poverty", *Econometrica*, vol. 55, no. 4, pp. 749-764

BISHOP, John A., John P. FORMBY and Buhong ZHENG (1997), "Statistical Inference and the Sen Poverty Index", *International Economic Review,* vol. 38, no. 2, pp. 381-388

BISHOP, John A., K. Victor CHOW and Buhong ZHENG (1995), "Statistical Inference and Decomposable Poverty Measures", *Bulletin of Economic Research*, vol. 47, no.4, pp. 329-340

BLACKBURN, McKinley L. (1994), "International Comparisons of Poverty", *American Economic Review (Papers and Proceedings)*, vol. 84, no. 2, pp. 371-374

BUHMANN, Brigitte, Lee RAINWATER, Günther SCHMAUS and Timothy M. SMEEDING (1988), "Equivalence Scales, Well-Being, Inequality, and Poverty: Sensitivity Estimates Across Ten Countries Using the Luxembourg Income Study (LIS) Database", *Review of Income and Wealth* 34, pp. 115-143

BURKHAUSER, Richard V., Timothy M. SMEEDING and Joachim MERZ (1996), "Relative Inequality and Poverty in Germany and the United States Using Alternative Equivalence Scales", *Review of Income and Wealth*, vol. 42, no. 4, pp. 381-400

DEATON, Angus and John MUELLBAUER (1980), *Economics and Consumer Behaviour*, Cambridge University Press

CALLAN, Tim and Brian NOLAN (1991), "Concepts of Poverty and the Poverty Line", *Journal of Economic Surveys* vol. 5, no.3, pp. 243-261

EFRON, Bradley and Robert J. TIBSHIRIANI (1993), *An Introduction to the Bootstrap*, Monographs on Statistics and Applied Probability 57, Chapman & Hall, New York

FOSTER, James, Joel GREER and Eric THORBECKE (1984), "A Class of Decomposable Poverty Measures", *Econometrica* 52, pp. 761-766

FOSTER, James E. and Anthony F. SHORROCKS (1988), "Poverty Orderings", *Econometrica*, vol. 56, no. 1, pp. 173-177

GOEDHART, T., V. HALBERSTAD A. KAPTEYN and B. VAN PRAAG (1977), "The Poverty Line: Concept and Measurement", *Journal of Human Resources* 12, pp. 503-520

KAKWANI, Nanak (1993), "Statistical Inference and the Measurement of Poverty", *Review of Economics and Statistics*, vol. 75, no. 4, pp. 632-639

MAASOUMI, E. (1994), "Empirical analysis o inequality and welfare" in P. SCHMIDT and H. PESARAN (eds.), *Handbook of Applied Microeconomics*

MILLS, Jeffrey A. and Sourushe ZANDVAKILI (1997), "Statistical Inference via Bootstrapping for Measures of Inequality", *Journal of Applied Econometrics*, vol. 12, pp. 133-150

NELSON, Julie A. (1993), "Household Equivalence Scales: Theory versus Policy?", *Journal of Labour Economics*, vol. 11, no. 3, pp. 471-493

PRYKE, Richard (1995), *Taking the Measure of Poverty. A Critique of Low-Income Statistics: Alternative Estimates and Policy Implications*, Institute of Economic Affairs, Research Monograph 51, London

RONGVE, Ian (1997), "Statistical Inference for Poverty Indices with Fixed Poverty Lines", *Applied Economics*, vol. 29, no. 3 (March), pp. 387-392

SACHS, Lothar (1984), *Applied Statistics. A Handbook of Techniques (second edition)*, Springer Verlag, New York

SEN, Amartya (1976), "Poverty: An Ordinal Approach to Measurement", *Econometrica*, vol. 44 no.2, pp. 219-231

SEN, Amartya (1973), *On Economic Inequality, Oxford*, Clarendon Press

SHAO, Jun and Dongsheng TU (1995), *The Jackknife and Bootstrap*, Springer Verlag, New York

SITTER, R.R. (1992), "Comparing three bootstrap methods for survey data", *The Canadian Journal of Statistics*, vol. 20, no.2, pp. 135-154

TSAKLOGLOU, Panos (1996), "Elderly and Non-Elderly in the European Union: A Comparison of Living Standards", *The Review of Income and Wealth*, Series 42, no. 3 (September), pp. 271-291

TULLOCK, Gordon (1984), *Economics of Income Redistribution* (Kluwer-Nijhoff studies in human issues), Kluwer-Nijhoff Publishing

ZHENG, Buhong, Brian J. CUSHING and Victor K. CHOW (1995), "Statistical Test of Changes in U.S. Poverty, 1975-1990", *Southern Economic Journal*, vol. 62, no. 2, pp. 334-347

Appendix A1: Confidence intervals based on the bootstrap percentiles

In this appendix, we briefly describe how confidence intervals based on the bootstrap percentiles are computed. The appendix is drawing on Efron and Tibshiriani [(1993), chapter 13].

Consider $\hat{\theta}$, an estimate of the parameter q and $s\hat{e}$ is the estimated standard error. The standard normal confidence interval is given by:

$$\left[\hat{\theta} - z^{(1-\alpha)} s\hat{e}; \hat{\theta} - z^{(\alpha)} s\hat{e}\right]$$

If $\hat{\theta}^*$ is a random variable drawn from a normal distribution, the endpoints of the confidence interval can also be described as follows:

$\hat{\theta}_{lower} = \hat{\theta}^{*(a)} = 100\ a^{th}$ *percentile of* $\hat{\theta}^*$*'s distribution*
$\hat{\theta}_{upper} = \hat{\theta}^{*(1-a)} = 100\ (1\text{-}a)^{th}$ *percentile of* $\hat{\theta}^*$*'s distribution*

Assume that the bootstrap data set x^* is generated according to $\hat{F} \rightarrow x^*$ and that the bootstrap replications $\hat{\theta}^* = s(x^*)$ are computed. Let $\hat{G}$ be the cumulative distribution function of $\hat{\theta}^*$. The *(1-2a)* percentile interval is defined by the a and *1-a* percentiles of $\hat{G}$:

$$\left[\hat{\theta}_{\%lo}; \hat{\theta}_{\%up}\right] = \left[\hat{G}^{-1}(\alpha); \hat{G}^{-1}(1-\alpha)\right]$$

By definition, $\hat{G}^{-1}(\alpha) = \hat{\theta}^{*(\alpha)}$, the *100 a^{th}* percentile of the bootstrap distribution. Thus:

$$\left[\hat{\theta}_{\%lo};\hat{\theta}_{\%up}\right] = \left[\hat{\theta}^{*(\alpha)};\hat{\theta}^{*(1-\alpha)}\right]$$

This confidence interval refers to the ideal bootstrap situation i.e., $B = ¥$. The approximate *(1-2a)* percentile confidence interval is:

$$\left[\hat{\theta}_{\%lo};\hat{\theta}_{\%up}\right] \approx \left[\hat{\theta}_{B}^{*(\alpha)};\hat{\theta}_{B}^{*(1-\alpha)}\right]$$

The central limit theorem tells us that as B ® ¥, the bootstrap histogram will become normally shaped.

Appendix A2: Poverty and inequality indices

The poverty and inequality indices used in this paper are widely used in the literature. In this appendix, we just present the formulas that we have used to compute them.

a. Poverty indices

The *headcount index* simply gives the proportion of poor people in the total population. It is defined as follows:

$$HC = \frac{p}{n}$$

where p is the number of poor and n is the number of persons in the population.

The index proposed by *Foster, Greer and Thorbecke* [1984] takes into account the number of poor people in the total population as well as their average shortfall from the poverty line. The Foster-Greer-Thorbecke *(FGT)* index is defined as follows:

$$FGT = \frac{1}{n}\sum_{i=1}^{p}\left(\frac{z - y_i}{z}\right)^{e}$$

where z is the poverty line, y_i is the income of poor household i and e is a poverty aversion parameter. In our calculations, we use the value $e = 2$ suggested by Foster et al. [1984]. As the FGT index is typically a small number, we have multiplied our estimates by 10.

Sen [1976] proposes a poverty index that combines the headcount index (measuring the number of poor), the poverty gap ratio (measuring "depth" of poverty) and the Gini

coefficient (measuring income inequality) into a single measure of poverty. *Sen's poverty index* is defined as follows:

$$SEN = \frac{2}{(p+1)nz}\sum_{i=1}^{p}(z - y_i)(p+1-i)$$

All three poverty indices lie between 0 and 1. They are equal to 0 if every household has income greater than the poverty line. The headcount index is equal to 1 if all households have incomes below the poverty line. The FGT and Sen indices are equal to 1 if every household has zero income.

b. Inequality measures

The *Gini coefficient* is based on the Lorenz curve of income distribution. It can be interpreted as the expected income gap between two individuals randomly selected from the population:

$$Gini = \frac{2}{n^2\bar{y}}\sum_{i=1}^{n} i\left(y_i - \bar{y}\right)$$

where n is the population size and $\bar{y}$ is the mean income. The incomes y_i are ordered in ascending order. The index varies between 0 (perfect equality) and 1 (perfect inequality).

Theil's measure of entropy is defined as:

$$Theil = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{\bar{y}}\log\frac{y_i}{\bar{y}}$$

Theil's index also reaches the value of 1 in the case of perfect inequality.

Finally, the *relative mean deviation of incomes* is defined as follows:

$$RMD = \frac{\sum_{i=1}^{n}\left|y_i - \bar{y}\right|}{2n\bar{y}}$$

A higher value of the index implies greater inequality. The index can be interpreted as the average percentage income transfer from those above the mean to those below the mean necessary to achieve perfect equality.

Appendix A3: Sensitivity of the poverty indices to changes in the poverty line (1992, point estimates only)

a. headcount index

	Hungary		Luxembourg		United Kingdom	
poverty line as percentage of median equivalent income	*retired*	*working*	*retired*	*working*	*retired*	*working*
66.67 %	0.392	0.064	0.251	0.123	0.597	0.069
60.00 %	0.271	0.045	0.184	0.075	0.526	0.057
50.00 %	0.124	0.029	0.099	0.021	0.344	0.040
40.00 %	0.041	0.018	0.018	0.007	0.187	0.025
30.00 %	0.019	0.012	0.004	0.003	0.052	0.009

b. FGT index

	Hungary		Luxembourg		United Kingdom	
poverty line as percentage of median equivalent income	*retired*	*working*	*retired*	*working*	*retired*	*working*
66.67 %	0.299	0.107	0.156	0.050	0.816	0.105
60.00 %	0.211	0.091	0.092	0.027	0.614	0.083
50.00 %	0.132	0.071	0.032	0.011	0.370	0.054
40.00 %	0.097	0.058	0.012	0.004	0.211	0.034
30.00 %	0.077	0.047	0.003	0.000	0.137	0.022

c. Sen index

	Hungary		Luxembourg		United Kingdom	
poverty line as percentage of median equivalent income	*retired*	*working*	*retired*	*working*	*retired*	*working*
66.67 %	0.081	0.019	0.050	0.020	0.185	0.022
60.00 %	0.055	0.016	0.034	0.011	0.149	0.017
50.00 %	0.026	0.010	0.013	0.004	0.086	0.011
40.00 %	0.016	0.008	0.004	0.002	0.043	0.006
30.00 %	0.010	0.007	0.001	0.000	0.022	0.004