The EU-SILC in comparative income distribution research: design and definitions in international perspective

Markus Jäntti
Economics and Statistics
Åbo Akademi University

Preliminary and incomplete draft
Not for citation
October 2006

1 Introduction

An important reason for creating microdatasets on income and other measures of living standards or well-being is to be able to make comparisons across countries. Cross-country comparisons are notoriously difficult to do in a way that does not leave some residual doubt (or doubters) as to the meaningfulness of such comparisons. For instance, influential academic economists, such as Martin Feldstein, have argued against the use of relative poverty as a criterion to compare poverty across countries.

Even if the meaningfulness of cross-country comparisons of typical measures of income inequality and poverty can be debated, there is less debate on the extent to which changes across time in living standards within a country are meaningful. While these, too, may suffer from problems that are similar to those that plague cross-country comparisons, especially if the time-span is long, studying change within a country is more widely recognised as being worthwhile.

To learn about progress (or a lack thereof), it is usually beneficial to have some perspective on what has happened in the recent past. Changes in income distributions, in particular, are often best understood in the context of what has happened in the past decade or so, at least, if not over longer periods. In examining the evidence based on the firstwaves of the EU-SILC, it will be helpful to use earlier data from other sources to place such evidence in context.

The three main sources of microdata-based evidence that predate the EU-SILC are the European Community Household Panel (ECHP), national data sources and data from the Luxembourg Income Study (LIS). Each of these sources have their own uses. For instance, the ECHP can be a valuable source for indicators of material deprivation and on longitudinal measures of well-being. National data sources, while by definition not comparable across countries can be useful for highlighting the longer-run trends in economic well-being, even though the comparability with EU-SILC-based estimates must be examined quite closely on a country-by-country basis.

The purpose of this paper is to examine the use of the third source, LIS, for comparisons with EU-SILC-based indicators. LIS is a widely used collection of harmonised data sets. It allows researchers to estimate within-country trends in some cases from the early 1970s at about 5-year intervals on income distribution and poverty statistics for several European as well as non-European countries. It is thus an important source for comparing recent developments in Europe with the changes that have occurred over the last couple of decades.

The paper is structured as follows. In the next section, I briefly discuss sample design issues in EU-SILC data and contrast them with the data in LIS. The section also compares income definitions and their measurement across the two sources. In Section **??**, I discuss national accounts aggregates comparisons. In Section 3, I take a closer look at inequality and poverty in the two sources.

2 Survey design and income measurement

In this section, I discuss differences in EU-SILC and LIS data sets in survey design and in the details of income definitions and measurement. The materials are collected from the user documention of both EU-SILC and LIS. This section also draws some lessons from a research project into various issues relating to the European Community Household Panel (ECHP), the CHINTEX project, which among other things used register data for Finland to explore issues such as the impact of using interview rather than register income components.

2.1 Survey design

The survey design in EU-SILC is regulated by the European directive governing the EU-SILC. The implementation is described in chapter 5 of Eurostat (2005). Since the EU-SILC attempts to gather both household- and person-level information, it has rules that apply to both types of units. Since part of the information is longitudinal in nature, there is also a substantial longitudinal component to the EU-SILC. However, only part of the data are longitudinal in nature. One deduces from this that the ECHP experiences may have led to this split in the design. What exactly those reasons are is not known to me.

The data sets in LIS have been provided by each member country and are typically the main source of income distirbution information in each country. In Wave 5 of LIS – those datasets that measure income centered on year 2000 – a substantial number of the European countries have contributed their ECHP data (Wave 8). This has the benefit of having provided LIS with data which have very similar content for a substantial subset of its members. The drawback, however, is that whatever the problems with ECHP, such as quite substantial panel attrition in some cases, those problems are also inherited by the LIS data.

Presumably some assessment of the desirability of continuing the ECHP was made within the European Union decision-making apparatus. Since the ECHP was discontinued, this indicates that the overall assessment was in the negative. However, I am not aware of any major problems which would definitively suggest that the ECHP datasets in LIS would have clearly superior and easily available alternative datasets. The ECHP data are used for cross-section purposes only and maybe getting the cross-sectional weights accurate enough works (LIS does not provide the longitudinal identifiers to link two datasets from the same country even if teh ECHP were available twice). However, by Wave 8 many ECHP datasets have suffered through quite substantial panel attrition and doubts remain as to how representative the remaining respondents are of the general population.

On the other hand, several long running panel surveys have learnt to live with things like panel attrition. The German Socio-Economic Panel, for instance, has over the year added new parts so as to keep it representative. Thus, that a longitudinal data source underlies the cross-sectional data in LIS is not in itself a problem.

Most of the datasets in LIS Wave 5 stem from complex sampling designs. In most cases, sampling is based on choosing clusters from geographical strata. Within the first-stage sampling units, a list of dwellings or buildings is then chosen from a register of such units. The Nordic countries tend to use registers of persons which are stratified according to income. In principle, the sampling design could be taken into account in estimation. However, the use of a longitudinal data source makes it close to impossible to use the correct sample structure when estimating complex statistics, such as inequality indices, from the data and all statistical inference will have to rely on approximations and simulation methods. Moreover, the exact information needed to accurately take the sampling design into account is not available in the LIS datasets.

The EU-SILC consists of national implementations of the common structure. From the descriptions of the individual sampling degisng in Annex 2 of the EU-SILC UDB description it seems that Wave 5 of LIS and the EU-SILC have in each country implemented a reasonabnly similar sampling design. The exceptions mainly consist of the countries that rely on late waves of panel data. But even in these (mainly ECHP) the original sample were drawn using similar methods. The EU-SILC user database includes in its first wave information on the strata, so

it seems to be at least in principle possible to take the true sampling structure into account in estimation.

Imputation and other data edits is an important issue in the analysis of empirical microdata. LIS does not make any imputations of its own and also keeps data editing to a minimum. In the main LIS relies on what country providers have done and attempts to inform users of those edits and imputations. The EU-SILC, by contrast, has rules about imputations that are to apply to all datasets even if (my understanding is) the actual imputation and editing is done in each country rather than centrally at Eurostat (see chapter 7 in Eurostat (2005)).

2.2 Measurement of disposable income

The definition of disposable household income inboth the LIS and the SILC is much influenced by the so-called Canberra group report, itself an update of the provisional UN guidelines.

Two important definitional and measurement issues cut across both the EU-SILC and the LIS datasets, namely whether or not incomes are recorded net of taxes and whether or not income are measured using interview or administrative registers of different sorts (or, indeed, some combination of the two types of sources). Countries tend to be either "net" or "gross", i.e., we have *either* net income sources or gross income sources. I have never quite understood the reason for the net/gross divide, but my understanding is that for some countries is difficult to overcome. Apparently serious efforts will be made in net countries to deliver gross data in later waves of the SILC.

Similarly, countries tend to either collect their income information mainly from registers (in particular, the Nordic countries) or from interviews (the rest). The countries that mainly rely on registers tend to be the countries which also have the lowest levels of inequality and of relative poverty, which in part motivates suspicions that maybe differences in sources may account for part of the measured differences.

The most important limitation at this point of the SILC is that in initial waves, there are several exceptions to what exactly is recorded by the different countries.

2.3 Interview and register income: lessons from ECHP

There is a substantial literature on measurement errors in income, but much of this focusses on the errors that arise in interviews (CITATION). There is far less research on errors in register incomes, although of course studies of tax evasion and so on are quite relevant to these. Errors in register incomes can arise because of tax evasion or e.g. errors, omissions or undercoverage in the administrative processes that govern the income data in such registers. The kinds of errors that do occur in register income data probably vary substantially across countries depending on the precise nature of the underlying registers. However, it would probably be a

mistake to suppose that register data accurately measure "true" income.

We do have some evidence about the errors in income data obtained in a few cases where countries that traditionally take their income information from registers have also gathered interview data on all rather than just a limited number of income sources. One instance of such comparisons is the Finnish ECHP data, which were studied in an EU-funded research project called CHINTEX (Ehling & Rendtel, 2004).

The findings of those reports contain both good and bad news. Below, I include some results from Jäntti (2004) that examined these data. The income variable in use in this paper is disposable money income, which includes all cash income from labour and capital markets, private and public cash transfers, less direct taxes (but not e.g. imputed rents). There are a number of alternative measures, differences among which are one of the main objects of this study. There are three ways in which the different disposable incomes differ: whether it is based on information from registers or from the interviews, what time interval the variable refers to (monthly or annual), and finally on whether income is assumed to be shared within the household or the dwelling unit (which are based on interviews and registers, respectively).

Finnish income data in e.g. the IDS is based on income information gathered from registers, although that register income is then aggregated within households, as defined through interviews.

ECHP waves 3 and 7 gathered two types of disposable income information through interviews:

1. the household head was asked about the *current monthly income* of the household. If he/she could name an amount [Q 84], that amount was recorded [Q 85]. If not, he/she got to choose from a number of income ranges [Q 86]. The amount named or, if the income range is named, the class mid-point, adjusted to correspond to annual income, is taken to be *current household interview income*.[1]

2. in Waves 3 and 7, each household member was asked about all components of disposable money income (in the previous year) [H 137–H388]. These amounts are summed across components and then within households to get the *household interview income*.

For every person who is included in the population census, Statistics Finland has defined their personal disposable money income. I use this variable to construct two measures of disposable income based on register income:

[1]The class midpoint is undefined in the top interval (which is open). If the respondent indicates that current monthly income is in the top interval, I assign the household the average current monthly income among those who responded to Q 85, i.e., who could name an amount, and for whom the amount was in the top interval.

3. I take the disposable income of each *household* member in the previous year as it is recorded in the relevant registers. This is then aggregated within households to generate *household register income*.

4. I aggregated disposable income within *dwelling units* to generate *dwelling unit register income*.

The two register-based concepts of annual disposable money income are needed because I intend to examine both the effect of interview vs register income (2 vs 3) and how non-response and attrition affect income distribution statistics (which requires 4, because we do not know the household structure of non-respondents). We should also note that 3, *household* disposable register income mixes interview and register income since who belongs to a household is asked in interviews whereas 4, *dwelling unit* register disposable income is a purely based on registers. Differences between the two concepts may thus be due to differences in the two "household" concepts. While is is customary to assume that register incomes are a more accurate measure of income than interview income[2], there is no reason to assume that households are more accurately defined in registers than in interviews. I shall look into this issue also.

I use throughout the paper disposable equivalent money income, calculating a unit's number of equivalent adults using the so-called modified OECD scale, which assigns a weight of one to the first adult, a weight of 0.5 to each individual over the age of 14 and 0.3 to children who are less than 14 years old.(See Atkinson et al., 2002)

Table 1 shows the estimated mean, median and mean of log income for each of the income concepts described in above in the two waves of data along with the change (as measured by the difference) across waves. The two interview income concepts have lower estimated central tendencies that the two register-based incomes. Current interview income (the one measured by querying only of the interview person the household current monthly income) being lowest, while household register income is highest. The two register-based concepts are very close to each other.

Turning to the changes in central tendency across time, the interview-based definitions display larger increases, both in relative and in absolute terms, than the register-based incomes. The ordering of the income concepts is unchanged, however.

Selected percentiles of the income distributions are shown in Table 2. These suggest that the order displayed by the central tendencies holds (almost) throughout the distribution. The exception is dwelling-unit and household register income, for which in Wave 7 dwelling unit income overtakes household income by the 90th percentiles. Interestingly, the interview incomes increase much more across the waves than does register-based income. Indeed, the

[2]The implicit assumption is that grey income, i.e., income that is in fact received but is not for some reason included in the registers, such as in the case of tax evasion, is less than various errors that arsie in interviews.

Table 1 Central tendency of income variables

mean	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		12304		13966		1662
Household, interview, annual	13289		15138		1849	
Household, register, annual	15148		16512		1364	
Dwelling unit, register, annual	14970		16436		1466	

median	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		11553		12810		1257
Household, interview, annual	12328		13669		1342	
Household, register, annual	13913		15096		1183	
Dwelling unit, register, annual	13788		14806		1018	

logmean	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		9.34		9.45		0.107
Household, interview, annual	9.39		9.50		0.1050	
Household, register, annual	9.54		9.60		0.0640	
Dwelling unit, register, annual	9.52		9.58		0.0632	

Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

10th percentile of dwelling unit register income declines a little across the waves.

Selected income inequality statistics are shown in Tables 3 and 4. The 90/50, 90/10 and 50710 percentile ratios (measured as the difference in the log of the percentiles) shown in Table 3 suggest that, the 90/10 ratios of the interview based incomes are higher than for register incomes – e.g., 1.110 for current household interview income as opposed to 0.967 for household regsiter income. The breakdown of this difference into the difference in the ln of the 90th and 50th, and 50th and 10th percentiles suggest that this overall difference is due to differences below the median, The 90/10 ln difference is very close to 0.5 for all four income measures but is higher for interview incomes for the 50/10 difference.

Further light is shed on the differences across the distributions by inspection of the relative inequality indices in Table 4. Current household income inequality is in Wave 3 clearly the highest, with the other three income measures being very close to .23. By Wave 7, current household income inequality has risen only marginally and is at the same level as household register income inequality. Household interview income and dwelling unit register income inequality have risen much more, being now both at 0.270.

If we examine the squared coefficient of variation instead, the ordering of inequality by income type is different. The interview incomes are for this statistic lower than the register incomes in Wave 3, a contrast with the Gini coefficient that is most likely driven by the relative absence of very high income reports for interview income. To further muddy the waters, by Wave 7 the incomes are re-ordered with current household income showing the by far lowest level of inequality and dwelling unit register income the highest.

I have also estimated for all income types a "robust" income statistic based on the in-

Table 2 Selected percentiles of the income distribution

p10	Wave 3		Wave 7		Change	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		6295		6855		560
Household, interview, annual	7032		7305		273.1	
Household, register, annual	8552		8560		8.1	
Dwelling unit, register, annual	8268		8244		−24.3	

p25	Wave 3		Wave 7		Change	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		8665		9657		992
Household, interview, annual	9487		9963		475	
Household, register, annual	10698		11266		568	
Dwelling unit, register, annual	10492		10870		378	

p50	Wave 3		Wave 7		Change	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		11553		12810		1257
Household, interview, annual	12328		13669		1342	
Household, register, annual	13913		15096		1183	
Dwelling unit, register, annual	13788		14806		1018	

p75	Wave 3		Wave 7		Change	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		14812		17245		2432
Household, interview, annual	16167		18207		2040	
Household, register, annual	18060		19518		1458	
Dwelling unit, register, annual	17903		19348		1444	

p90	Wave 3		Wave 7		Change	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		19108		21843		2736
Household, interview, annual	20080		23988		3908	
Household, register, annual	22494		24820		2325	
Dwelling unit, register, annual	22453		25140		2688	

Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

Table 3 Percentile ratios of the income variables

p90p10	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		1.11		1.16		0.0486
Household, interview, annual	1.049		1.19		0.1397	
Household, register, annual	0.967		1.06		0.0974	
Dwelling unit, register, annual	0.999		1.12		0.1160	

p90p50	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		0.503		0.534		0.0305
Household, interview, annual	0.488		0.562		0.0745	
Household, register, annual	0.480		0.497		0.0168	
Dwelling unit, register, annual	0.488		0.529		0.0418	

p50p10	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		0.607		0.625		0.0181
Household, interview, annual	0.561		0.627		0.0652	
Household, register, annual	0.487		0.567		0.0806	
Dwelling unit, register, annual	0.511		0.586		0.0742	

Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

Table 4 Relative inequality indices

gini	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		0.247		0.255		0.00782
Household, interview, annual	0.234		0.270		0.0361	
Household, register, annual	0.228		0.253		0.0252	
Dwelling unit, register, annual	0.234		0.270		0.0362	

cv2	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		0.226		0.25		0.024
Household, interview, annual	0.210		0.367		0.1570	
Household, register, annual	0.302		0.356		0.0539	
Dwelling unit, register, annual	0.313		0.524		0.2106	

iqrp50	**Wave 3**		**Wave 7**		**Change**	
	1995	1996	1999	2000	1999-95	2000-1996
Household, interview, monthly		0.532		0.592		0.0603
Household, interview, annual	0.542		0.603		0.0613	
Household, register, annual	0.529		0.547		0.0175	
Dwelling unit, register, annual	0.537		0.573		0.0351	

Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

terquártile range, standardized by the median. This statistic suggests inequality measured in all four income types is virtually the same and while not interview income inequality increases more across the two awaves, the levels recorded are still remarkably similar.

The differences across income inequality statistics reflect differences in where in the distribution the differences are largest. Since both the Gini coefficient and the squared coefficient of variation, CV^2, obey the Lorenz criterion, they generate different orderings only if Lorenz-curves cross. Visual inspection of the Lorenz curves, displayed in Figure 1, confirms this is the case. The graphs show the Lorenz curves less the population proportion to visually emphasize the differences across curves (this does not, of course, affect the ordering). In wave 3, it seems that all the curves cross, with the single exception that current household interview income and household register income do not appear to cross at any point. In Wave 3, even this exception is gone and none of the Lorenz curves either dominates or is dominate by another. We can therefore not say that inequality is unequivocally greater or less for any of the income sources against any of the others – -even absent considerations of statistical inference, which, while important, seem less interesting when Lorenz curves intersect than when there is dominance.

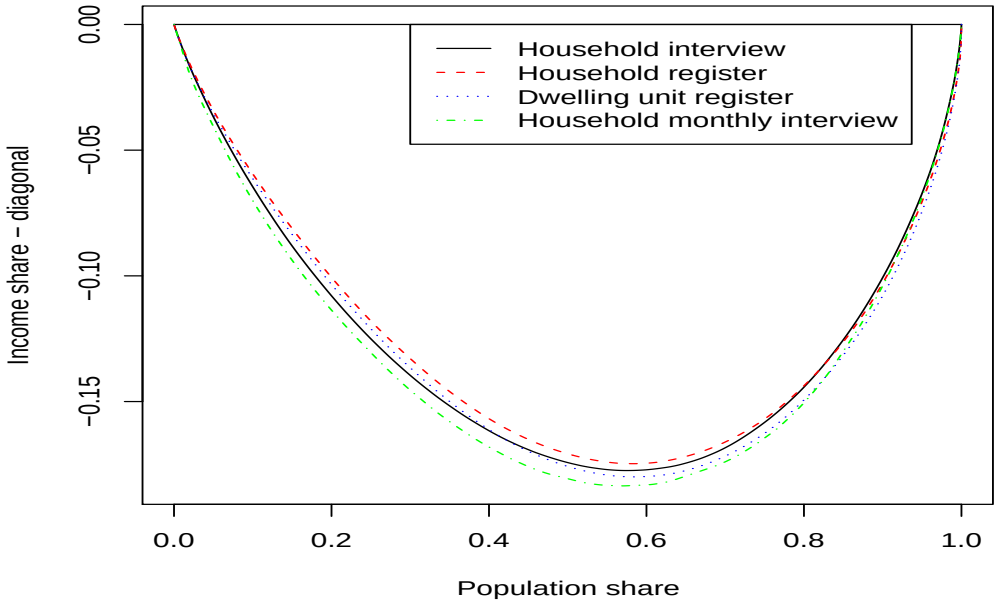
To examine the extent of poverty by each of the income sources, I use the TIP curve (Jenkins & Lambert, 1997). The TIP curve plots the cumulative average relative poverty gap against the (by income ordered) population proportion. At the point that it reaches its maximum, the horizontal axis shows the proportion who are poor and the vertical measures the poverty rate times the average relative poverty gap (see also Jäntti & Danziger, 2000). TIP curves generate poveryt orderings that are robust with respect variations in both poverty lines and poveryt indices, but generate partial orders (like Lorenz curves) in that when the TIP curves cross, further valuations need to be imposed.

The estimated TIP curves are shown in Figure 2 for Waves 3 and 7. In Wave 3, current household interview income displays most poverty and does not intersect any of the others. While household (annual) interview income is almost everywhere abnove the two remaining concepts, it intersects with dwelling unit register incomefairly close to the origin. While incomes are notoriously hard to measure, in the absence of statistical inference we can but conclude that these two income concepts can not be ordered. Household register income is, on the other hand, dominated by the three other concepts.

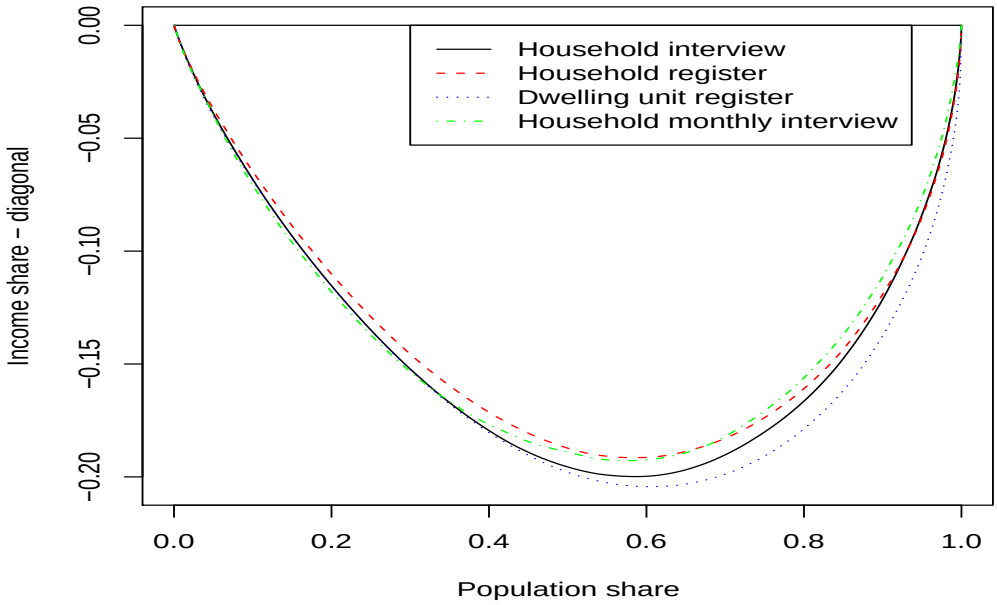
By Wave 7, things have changed, although household register income is still poverty dominated by all others. Household interview income is now dominated by both dwelling unit register income and by current household income, and the latter two intersect. Thus, not even for the (by and large) same population in two different years do these different concepts generate the same patterns.

Figure 1 Income inequality for different income concepts – Lorenz curves

Wave 3



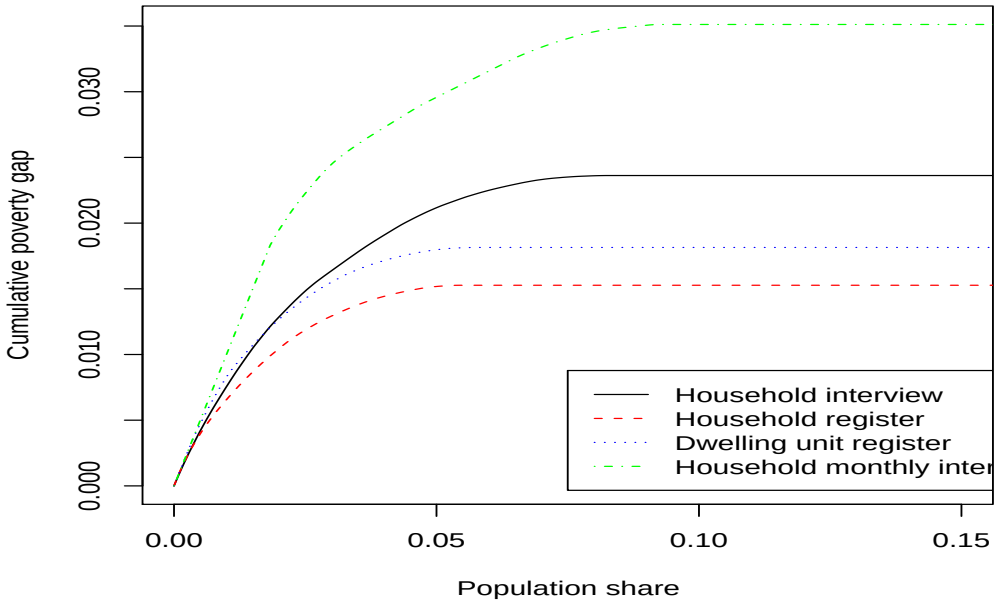
Wave 7



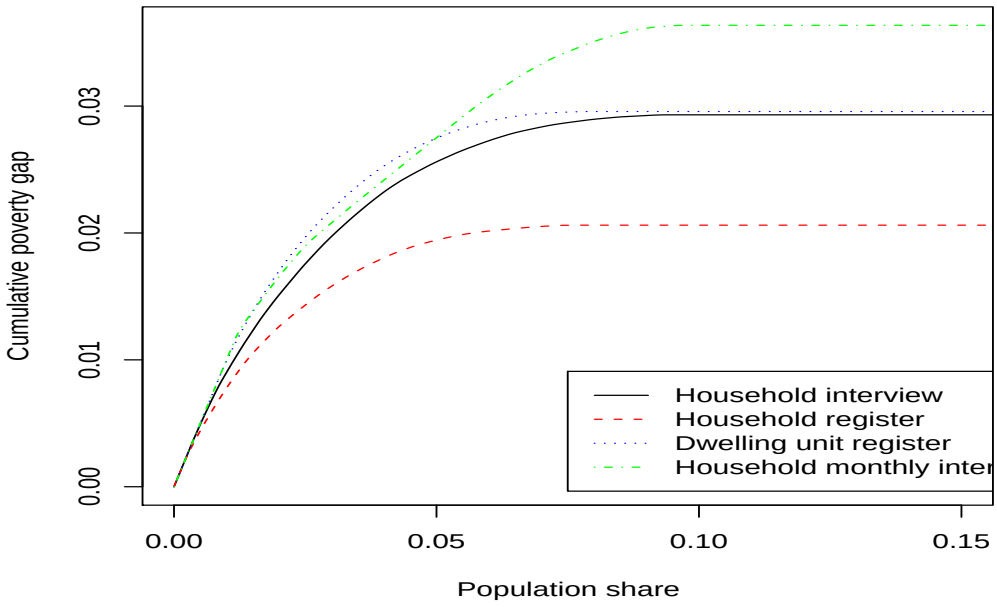
Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

Figure 2 Poverty for different income concepts – ECHP Wave 3

Wave 3



Wave 7



Note: The numbers refer to 2001 euros of disposable equivalent money income using the income source indicated, estimated for the responding ECHP sample in each wave.
Source: Jäntti (2004)

In closing To conclude this section, I would suggest that the underlying data in LIS and the EU-SILC are reasonably similarly structured, except that many of the European datasets in LIS in Wave 5 are based on late waves of the ECHP. Those datasets that provide mainly net income variables are pretty much the same in the two sources. Also, the datasets that mix registers and interviews for their incomes sources are the same in the two sources. I would also argue, based on Jäntti (2004) and Ehling & Rendtel (2004) that while income distribution statistics tend to be larger in interview than in register sources, the differences are unlikely to be large enough to significantly affect country rankings.

3 Income inequality and poverty

This section examines both income inequality and poverty across the two data sources based on the full (covered) population and selected population breakdowns, such as household type, age and other socio-economic characteristics. To keep things focused, I only include the adult population (18+) and mainly examine equivalent household disposable income. I discard observations with negative disposable income from the inequality and poverty statistics and use the square root of household size as the equivalence scale.

The section starts by looking at the economic context that characterises the each country in the income year of each survey source.

A comparison of income distribution statistics from the SILC with LIS data is problematic, since differences are likely to reflect changes in the underlying distribution of income rather than differences in the survey instruments. The comparisons below should thus be thought of as being illustrative. To give some idea of the extent to which one should expect income distributions to have been, Table 5 shows macroeconomic data for the countries at the year 2000 and 2003, the relevant years for each data source.

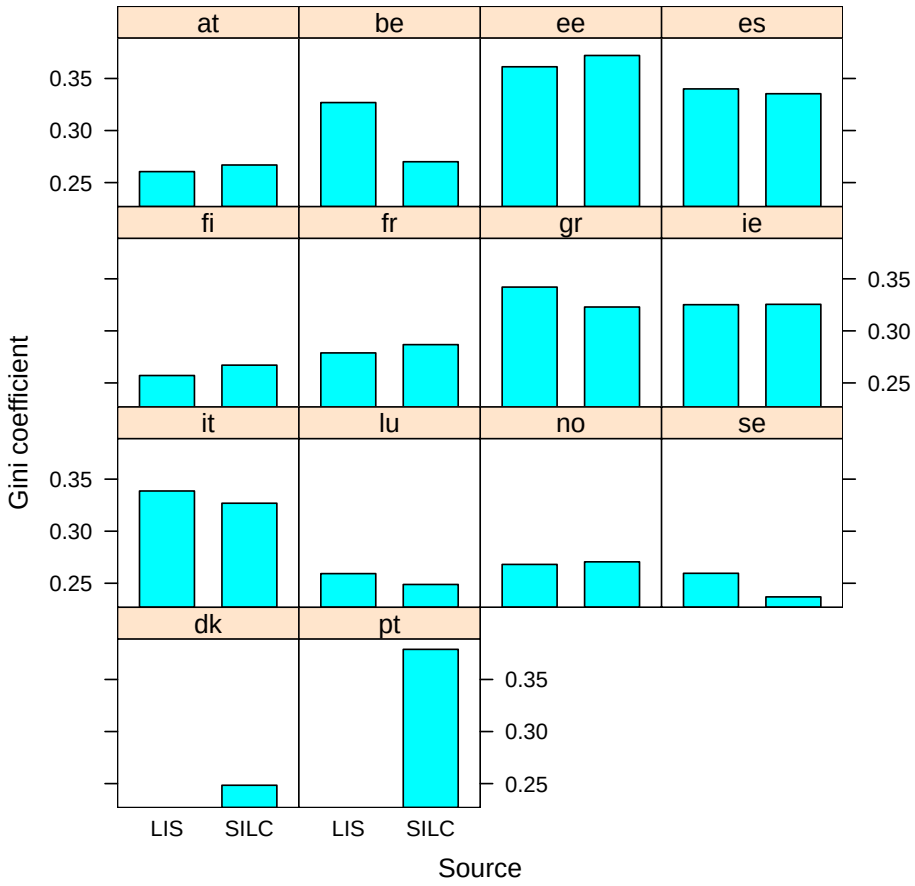
The connection between income distribution and macroeconomic indicators is not clear cut (see Parker, 2000). While it could bex expected that unemployment would lead to a lowering of the incomes of the low end of the distribution, the evidence on that is in fact quite mixed. It is also not clear what to expect across countries to happen when income growth slows down – it might well be that the incomes of those whose main income sources are fron transfers would be insulated from cyclical turn downs and this their relative income might increase.

It is noteworthy that GDP growth was slower in 2003 than in 2000 in every country except Greece. The difference was in some cases quite substantial. For instance, the growth rate of Luxemboutg declined from 8.4 per cent to only 1.3. In most countries, the growth rate of GDP was between 2 and 4 percentage points lower in 2003 than in 2000. For unemployment, the difference is not as systematic. Unemployment rates decreases in Estonia, Finland, Italy and Greece, remained unchanged in Spain and Sweden and increased, but by only about 0.5 to 1.3

Table 5 The economic context

	LIS (ca. 2000			EU-SILC (ca. 2003)		
Country	GDP growth	Unemployment	Inflation	GDP growth	Unemployment	Inflation
at	3.4	3.6	2.0	1.1	4.3	1.3
be	3.9	6.9	2.7	0.9	8.2	1.5
ee	10.8	12.8	3.9	7.1	10.0	1.4
es	5.0	11.1	3.5	3.0	11.1	3.1
fi	5.0	9.8	2.9	1.8	9.0	1.3
fr	4.0	9.1	1.8	1.1	9.5	2.2
gr	4.5	11.3	2.9	4.8	9.7	3.4
ie	9.4	4.3	5.3	4.3	4.7	4.0
it	3.6	10.1	2.6	0.0	8.4	2.8
lu	8.4	2.3	3.8	1.3	3.7	2.5
no	2.8	3.4	3.0	1.1	4.5	2.0
se	4.3	5.6	1.3	1.7	5.6	2.3

Figure 3 Income inequality as measured by the Gini coefficient of disposable income in EU-SILC and LIS



percentage points in the rest. The inflation rate was also lower in most (but not all) cases in 2003 than in 2000.

Conjecturing that slower growth may be associated with lower inequality, I would expect inequality to have declined. In the cases where unemployment also declined, I would again expect this effect to be quite strong, while where unemployment increased, on would expect a less pronounced impact. On the other hand, it is possible that no impact from macroeconomic fluctuations is to be expected.

3.1 Income inequality

The Gini coefficients for each country, shown in Figure 3, suggest that for the most part, changes were quite modest. The most substantial difference is for Belgium, with a measured decline of around .05 points in the Gini coefficient. Estonia, Spain, Greece, Italy and Ireland have quite high Gini coefficients (as does Portugal based only on the SILC), and the Northern European countries lower. A noteworthy difference in these data is that Sweden and Denmark

(SILC only) have substantiually lower inequality than Finland and Norwary.

Next I show the Lorenz curves (drawn as the difference between the Lorenz curve and the main diagonal) for each country in the two data sources (see Figure 4). For a little added clarity, Table 6 shows the Lorenz dominance of each pairwise comparison. As the table shows, there are many crossings of Lorenz curves, making inequality orderings depend on the specific measure chosen.

Even Estonia, which from the figure appears to have quite substantially higher inequality than most other countries turns out to have a Lorenz curve that crosses than of many other countries. Estonia has unambiguously more inequality than Belgium, France and Sweden (LIS data) and Austria, Spain and Luxembourg (SILC data). It would also appear from the table that the orderings are not necessarily very consistent across the years. For instance, Luxembourg Lorenz dominates many countries in the SILC data but none in LIS. Sweden also turns out to have quite different patterns of dominance in the two cases.

It is important to recall that changes in these orderings can be due to both changing differences in underlying inequality across the years and in changes to the measurement instruments. It is not possible, based on the comparison of these two sources, to conclude that inequality orderings really are very different across the years.

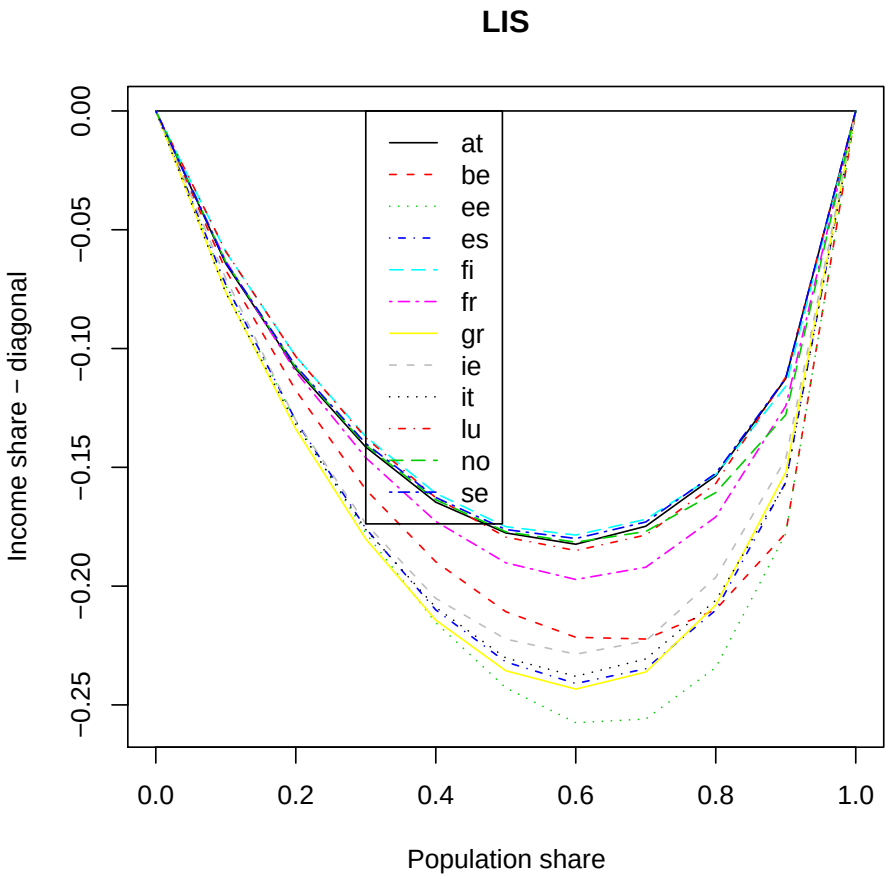
3.2 Poverty

I next turn to examine poverty, defined as having income les than one half of the national median.

The differences in poverty as measured by the head-count ratio appear to be quite large in contrast with those for inequality (at least as measured by the Gini coefficient). Ireland and Greecem both of which have substantially higher poverty in the SILC than in the LIS data, and Spain, Italy and Estonia have high relative poverty. The broad groups of countries are similar but the ordering does shift around a little but once we take into account also the poverty gap – by using the FGT index with $\alpha = 1$ (Foster et al., 1984). For instance, poverty now appears to be a worse problem in Ireland than in Italy. The Nordic countries seem to have the lowest poverty rates. This contrasts with the ordering based on the Gini coefficient, where the Nordic and northern European countries had reasonably similary levels.

However, the issue of interest here is whether LIS and SILC lead to different results on poverty orderings of countries. On this score, there is reasonably little to say. Poverty is lower in the SILC than in the LIS data in some of the high poverty countries, such as Ireland, Italy and Greece. The lower level of these countries' overall poverty rate in SILC allows for Estonia and Spain to change their ranking relative to these countries. Among the countries with fairly low levels of poverty, again, it appears that the increase in relative poverty in Luxembourg changes its relative position.

Figure 4 Lorenz curves for LIS and SILC data



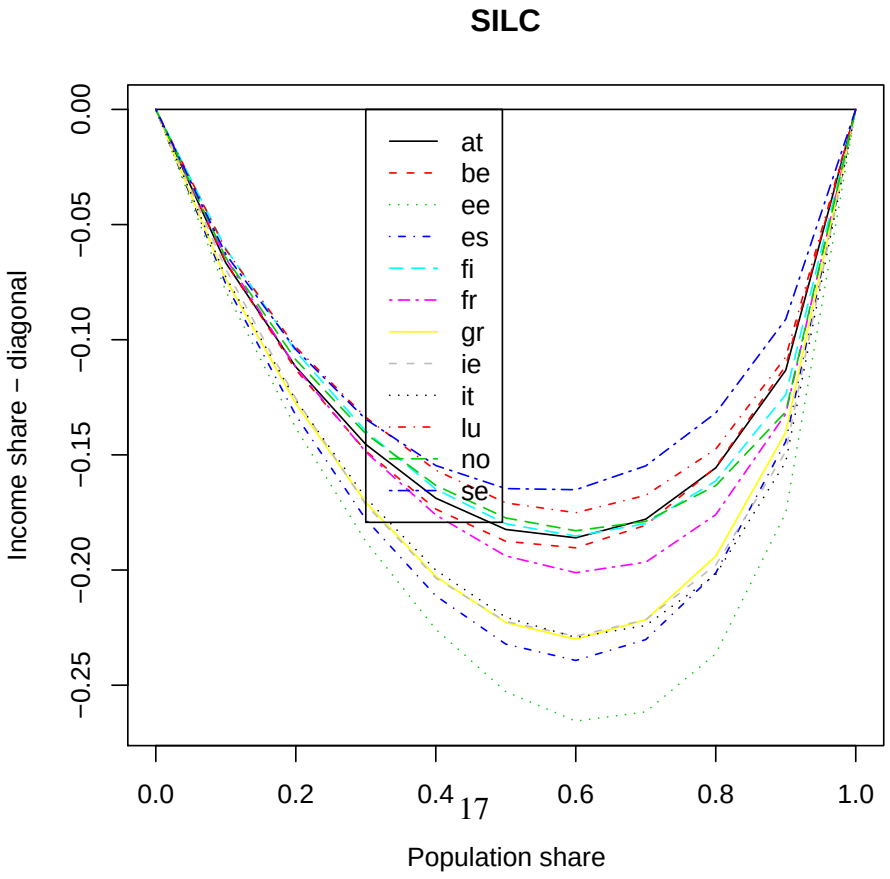


Table 6 Lorenz dominance in LIS and SILC data

A. LIS

	at	be	ee	es	fi	fr	gr	ie	it	lu	no	se
at		0	0	0	0	0	0	0	1	0	0	−1
be			1	0	0	−1	0	0	0	0	0	−1
ee				0	0	−1	0	0	0	0	0	−1
es					0	−1	0	−1	0	0	0	−1
fi						0	0	0	0	0	1	0
fr							1	1	1	0	0	0
gr								−1	0	0	0	−1
ie									1	0	0	−1
it										0	0	−1
lu											0	0
no												0
se												

B. SILC

	at	be	ee	es	fi	fr	gr	ie	it	lu	no	se
at		0	1	0	0	0	1	1	1	−1	0	0
be			0	0	0	0	1	1	0	−1	0	0
ee				−1	0	0	0	0	0	−1	0	0
es					0	0	0	0	0	−1	0	0
fi						0	0	0	0	0	0	0
fr							1	1	0	−1	0	0
gr								0	0	−1	−1	−1
ie									0	−1	0	0
it										−1	0	0
lu											1	0
no												−1
se												

Note: A "1" indicates the row country lorenz dominates the column country, a "-1 that the column Lorenz dominate the row and a "0" indicates crossing Lorenz curves.

Figure 5 Poverty comparisons – head count poverty ratio in EU-SILC and LIS data

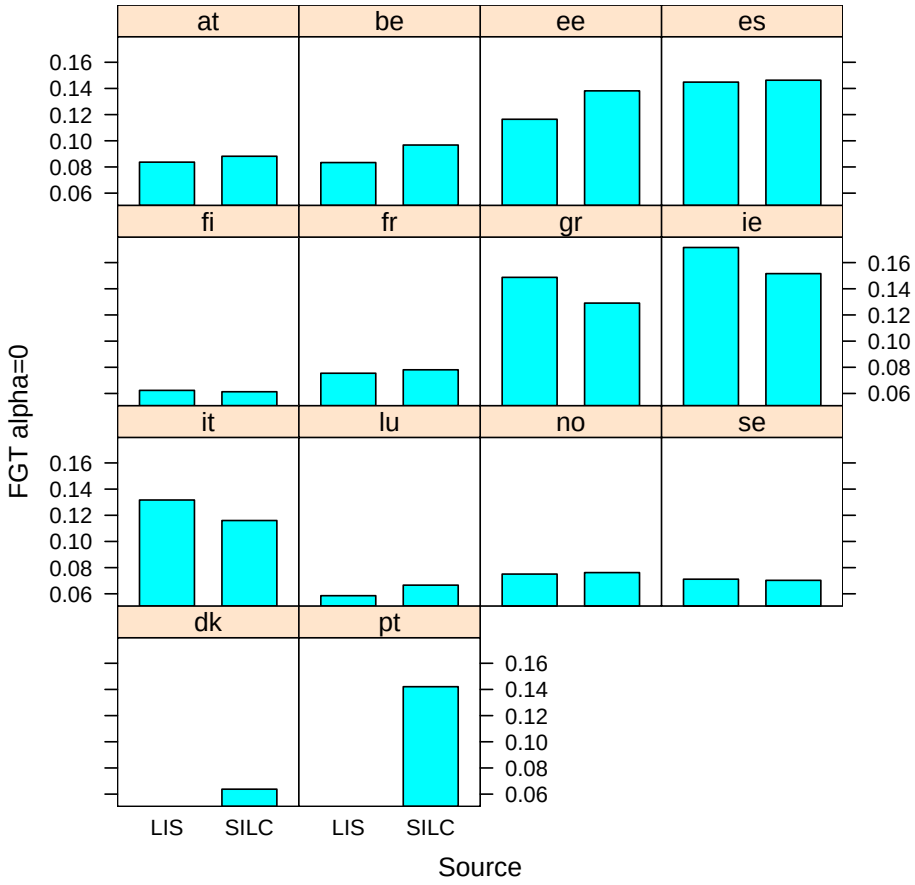


Figure 6 Poverty comparisons – the poverty gap ratio in EU-SILC and LIS data

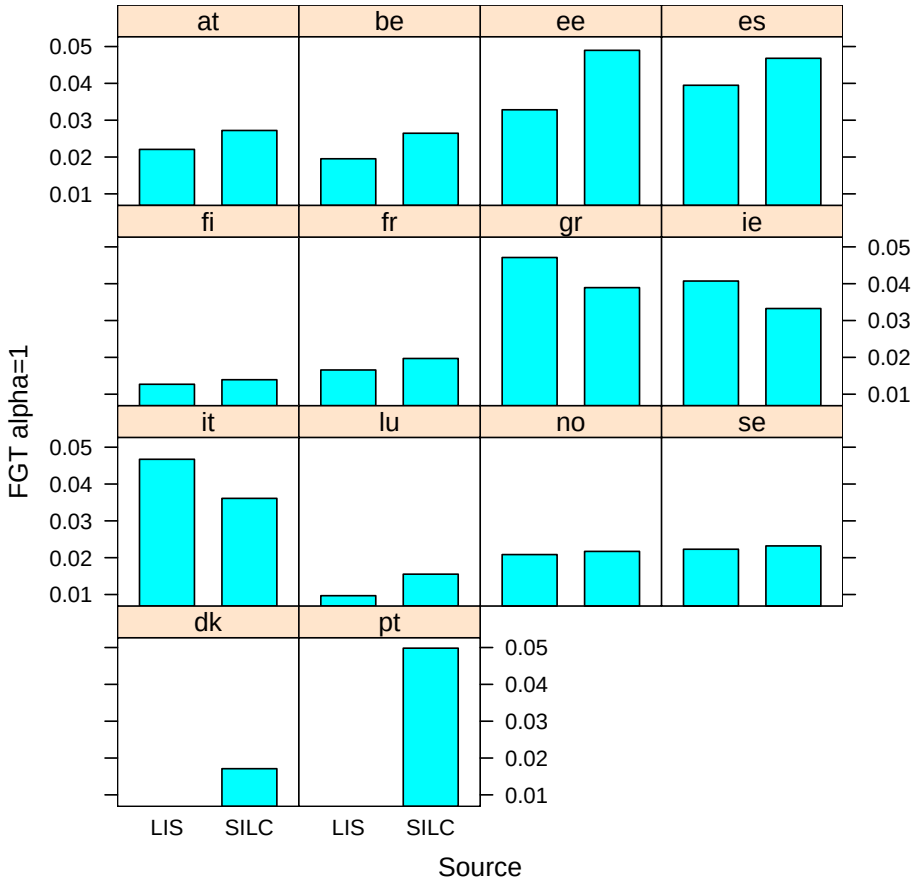
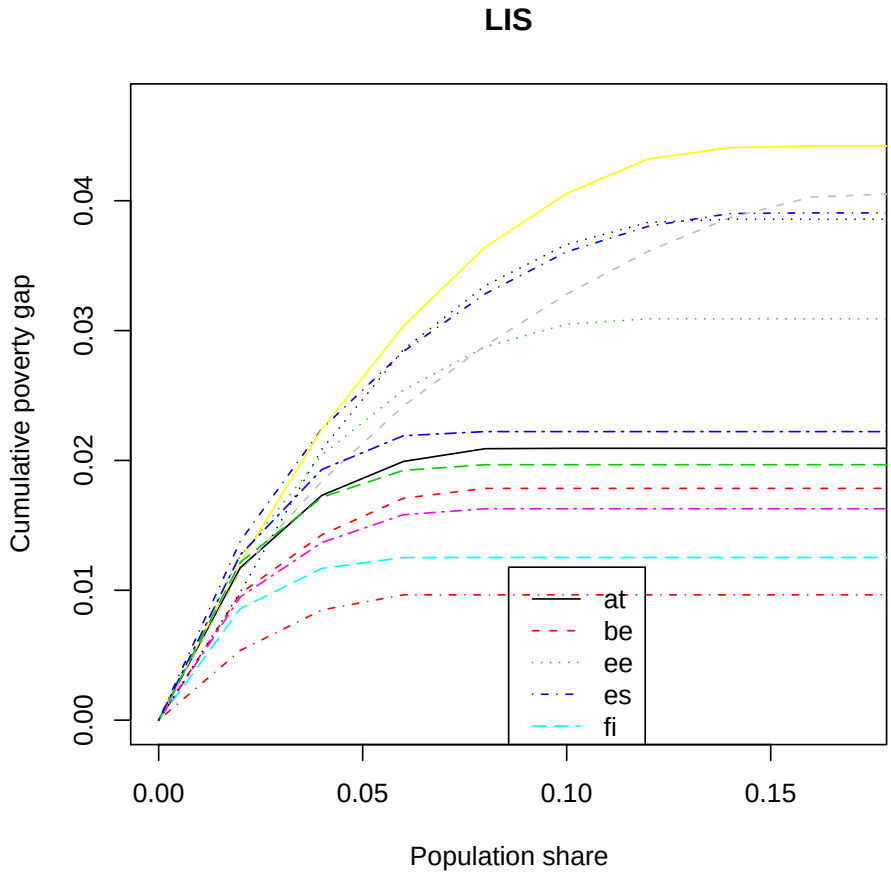


Figure 7 Poverty comparisons – the three I's of poverty in LIS and SILC data



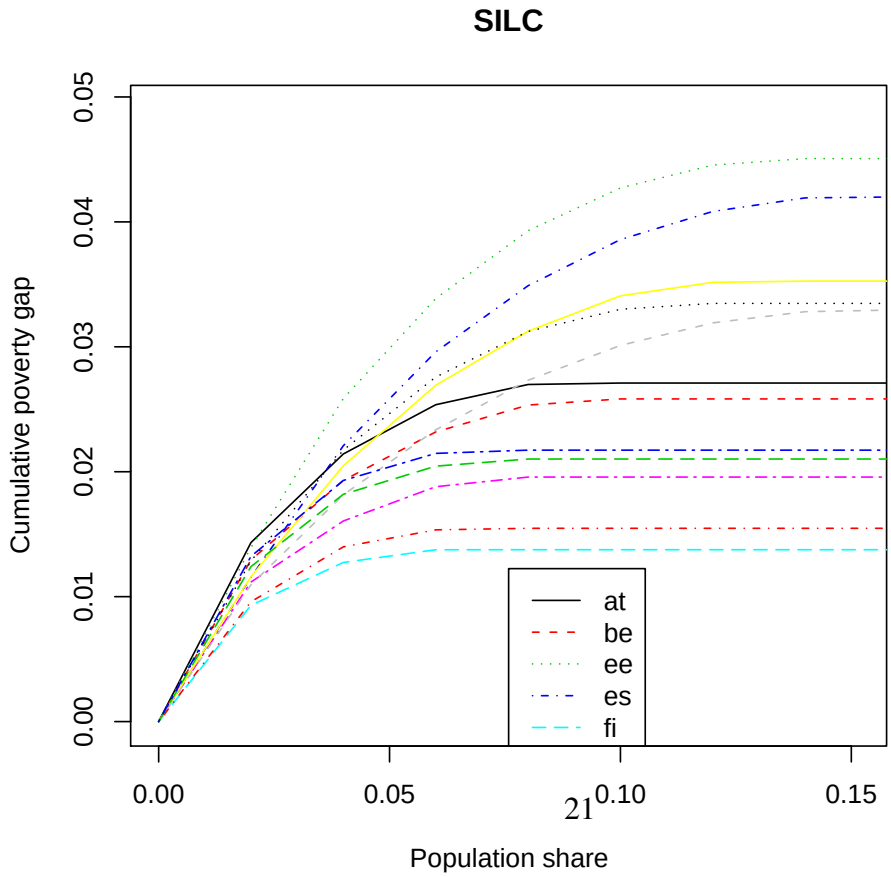


Table 7 "TIP" dominance in LIS and SILC data

A. LIS

	at	be	ee	es	fi	fr	gr	ie	it	lu	no	se
at		−1	1	1	−1	−1	1	0	0	−1	0	1
be			1	1	−1	−1	1	1	1	−1	1	1
ee				1	−1	−1	0	0	0	−1	−1	0
es					−1	−1	0	0	0	−1	−1	−1
fi						1	1	1	1	−1	1	1
fr							1	1	1	−1	1	1
gr								−1	−1	−1	0	0
ie									0	−1	0	0
it										−1	0	0
lu											1	1
no												1
se												

B. SILC

	at	be	ee	es	fi	fr	gr	ie	it	lu	no	se
at		−1	0	0	−1	−1	0	0	0	−1	−1	−1
be			1	0	−1	−1	0	0	0	−1	−1	0
ee				−1	−1	−1	−1	−1	−1	−1	−1	−1
es					−1	−1	0	−1	0	−1	0	0
fi						1	1	1	1	1	1	1
fr							1	0	1	−1	1	1
gr								−1	0	−1	0	0
ie									1	−1	0	0
it										−1	−1	0
lu											1	1
no												1
se												

Note: A "1" indicates the row country "tip" dominates the column country, a "−1 that the column "tip" dominates the row and a "0" indicates crossing TIP curves.

To examine the extent of poverty by each of the survey sources, I use the "Three I's of Poverty" – incidence, intensity and inequality – or "TIP" curve (Jenkins & Lambert, 1997). The TIP curve plots the cumulative average relative poverty gap against the (by income ordered) population proportion. At the point that it reaches its maximum, the horizontal axis shows the proportion who are poor and the vertical measures the poverty rate times the average relative poverty gap (see also Jäntti & Danziger, 2000). TIP curves generate poverty orderings that are robust with respect variations in both poverty lines and poverty indices, but generate partial orders (like Lorenz curves) in that when the TIP curves cross, further valuations need to be imposed.

The TIP curves in each of the two survey sources are shown in Figure 7 and the corresponding matrix of "TIP" dominances is shown in Table 7. Here, perhaps somewhat surprisingly, many dominance relations do arise. The patterns appear to be reasonably similar across time, even if a number of re-rankings do occur.

4 Concluding remarks

To be added.

References

Atkinson, A., Cantillon, B., Marlier, E., & Nolan, B. (2002). *Social Indicators: The EU and Social Inclusion*. Oxford: Oxford University Press.

Ehling, M. & Rendtel, U. (Eds.). (2004). *Harmonisation of panel surveys and data quality*. Wiesbaden: Statistisches Bundesamt. CHINTEX: The Change from Input Harmonization to Ex-post Harmonization in National Samples of the European Community Household Panel – Implications on Data Quality.

Eurostat (2005). Eu-silc user database description. EU-EILC/BB D(2005) 24-05-06, Eurostat. Version 2004-1.

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

Jenkins, S. P. & Lambert, P. J. (1997). Three 'i's of poverty curves, with an analysis of U.K. poverty trends. *Oxford Economic Papers*, *49*(3), 317–327.

Jäntti, M. (2004). The effect of measurement errors, non-response and attrition on income inequality, poverty and mobility. In Ehling & Rendtel (2004). CHINTEX: The Change

from Input Harmonization to Ex-post Harmonization in National Samples of the European Community Household Panel – Implications on Data Quality.

Jäntti, M. & Danziger, S. (2000). Poverty in Advanced Countries. In Anthony B Atkinson & François Bourguignon (Eds.), *Handbook of Income Distribution* chapter 9, (pp. 309–378). Amsterdam: North-Holland.

Parker, S. C. (2000). Opening a can of worms: the pitfalls of time-series regression analyses of income inequality. *Applied Economics*, *32*(2), 221–230.