High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)

V.-M. TÖRMÄLEHTO

2017 edition

STATISTICAL WORKING PAPERS

eurostat

High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)

V.-M. TÖRMÄLEHTO

2017 edition

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*)**:**
00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the Internet (http://europa.eu).

Luxembourg: Publications Office of the European Union, 2017

ISBN 978-92-79-64150-3
ISSN 2315-0807
doi: 10.2785/611164
Cat. No: KS-TC-16-027-EN-N

Theme: Population and social conditions
Collection: Statistical working papers

Preface

Eurostat is the Statistical Office of the European Union (EU). Its mission is to provide high-quality statistics on Europe. To that end, it gathers and analyses data from the National Statistical Institutes (NSIs) across Europe and provides comparable and harmonised data for the EU to use in the definition, implementation and analysis of EU policies. Its statistical products and services are also of great value to Europe's business community, professional organisations, academics, librarians, NGOs, the media and citizens.

In the field of income, poverty, social exclusion and living conditions, the EU Statistics on Income and Living Conditions (EU-SILC) is the main source for statistical data at European level.

Over the last years, important progress has been achieved in EU-SILC as a result of the coordinated work of Eurostat and NSIs.

In June 2010, the European Council adopted a social inclusion target as part of the Europe 2020 Strategy: to lift at least 20 million people in the EU from the risk of poverty and exclusion by 2020. To monitor progress towards this target, the 'Employment, Social Policy, Health and Consumer Affairs' (EPSCO) EU Council of Ministers agreed on an 'at risk of poverty or social exclusion' indicator. To reflect the multidimensional nature of poverty and social exclusion, this indicator consists of three sub-indicators: i) at-risk-of-poverty (i.e. low income); ii) severe material deprivation; and iii) (quasi-)joblessness.

In this context, the Second Network for the Analysis of EU-SILC (Net-SILC2) is bringing together NSIs and academic expertise at international level in order to carry out in-depth methodological work and socio-economic analysis, to develop common production tools for the whole European Statistical System (ESS) as well as to ensure the overall scientific organisation of the third and fourth EU-SILC conferences.

It should be stressed that this methodological paper does not in any way represent the views of Eurostat, the European Commission or the European Union. This is independent research which the author has contributed in a strictly personal capacity and not as representatives of any Government or official body. Thus he has been free to express his own views and to take full responsibility both for the judgments made about past and current policy and for the recommendations for future policy.

This document is part of Eurostat's Methodologies and working papers collection, which are technical publications for statistical experts working in a particular field. These publications are downloadable free of charge in PDF format from the Eurostat website: http://ec.europa.eu/eurostat/en/web/products-statistical-working-papers.

Eurostat databases are also available at this address, as are tables with the most frequently used and requested short- and long-term indicators.

Abstract[1]

This paper examines the top tail of the income distributions in the 2012 EU-SILC data. First, it discusses issues related to data quality, including under-estimation of top incomes. Then, the data are used as they are to compute several income-based measures of affluence. Finally, the link between non-income information and high incomes is analysed. The paper shows that EU-SILC is a useful complementary source on high incomes, in particular when the aim is to measure the size of the economically very well-off group. It also shows that identifying the affluent only on the basis of relative incomes is not sufficient. In a number of countries, many households in the upper tail of the income distribution report having difficulties in making ends meet.

(1) Statistics Finland. This is a revised version of a paper prepared for the International Conference on Comparative EU Statistics on Income and Living Conditions, Lisbon, 15-17 October 2014. The author wishes to thank Rolf Aaberge, Sir Tony Atkinson, Anne-Catherine Guio and Eric Marlier for valuable comments and suggestions. All errors remain strictly the author's responsibility. This work has been supported by the second Network for the analysis of EU-SILC (Net-SILC2), funded by Eurostat. The European Commission bears no responsibility for the analyses and conclusions, which are solely those of the author. Email address for correspondence: veli-matti.tormalehto@stat.fi.

Table of contents

1. Introduction

This paper describes the top tails of the income distributions in Europe, based on EU Statistics on Income and Living Conditions (EU-SILC). Using the most recent cross-sectional data, the paper explores issues related to measurement of the right tail in sample surveys, reviews measures of income-based affluence, and briefly addresses non-income dimensions of affluence.

Top incomes have raised considerable debate recently, based on new estimates derived from tax data (Atkinson, Piketty and Saez, 2011; Piketty, 2014). From a data point of view, it is interesting to see whether a set of household surveys with a reasonable degree of comparability could offer something to the debate, although sample surveys often are considered to have low accuracy in the top tail. This needs not always be the case, in particular with register-based measurement of incomes coupled with appropriate sampling designs and reweighting schemes. Consequently, with an increasing number of countries relying on register-based income data, the EU-SILC evidence on those with high incomes deserves to be examined. Moreover, EU-SILC is mandated to be the reference source for comparative statistics on EU income distribution, which obviously includes also the top tail, in addition to poverty and social inclusion statistics.

In addition to data specific concerns, there are several other motives to study the right tail of the income distribution. In many OECD countries where income inequality has grown, the high income households have increased their share of the total (OECD, 2014). There is also a high correlation of overall inequality and measures of "richness", and understanding the "rich" may contribute to a better understanding the poor and the redistributional policies directed towards them (Leigh, 2009). Although the tendency recently has been to reduce top personal income tax rates (OECD, 2014), the top tail still pays disproportionally direct taxes to finance the redistribution of income. Finally, there is some evidence that the relationship and mechanism between inequality and growth may be different in the top compared to the bottom and middle parts of the income distribution (Cingano, 2014).

The paper is structured as follows. The first part deals with measurement issues, and begins with a description of the data and concepts, followed by a discussion of errors of representation and errors of measurement when analysing EU-SILC data and top of the distribution. To assess sensitivity of the results, a semi-parametric approach with different under-estimation scenarios is used. We also compare results for selected countries with the World Top Incomes Database, which includes top income share estimates derived from tax data (Atkinson and Piketty, 2007, 2010). A concise analysis of outlying values and the impact of various adjustments on top income shares and whole income distribution is also conducted.

The second part goes through a range of income-based affluence measures, i.e. measures of "richness" proposed in the literature. Rather than fixing a certain high income threshold, a dominance approach is followed over the range of 200% of median upwards. The aim is for affluence orderings, and the affluence measures include headcounts, top income shares, transfer-sensitive measures and average excess incomes or "affluence gaps".

The third part briefly turns to non-income evidence of high economic wellbeing, including an attempt to construct a multidimensional affluence indicator, making use of the income and non-income dimensions that can be jointly measured with EU-SILC. This is important given the weaknesses of the relative income approach when comparing countries with very different income levels. The approach is that of multi-dimensional counting, with binary affluence indicators defining a positive achievement matrix. It turns out that non-income EU-SILC data to identify the affluent is quite limited. We also find that in a number of countries those in the upper tail of the distribution report having difficulties in making ends meet.

2. Definitions and data source

2.1 The data and definitions

Our data derive from the EU-SILC cross-sectional users' database 2012 (UDB, version August 2014). Only results from the most recent available data are reported, in order to cover more "register" countries. Important for this paper is whether or not also property income (dividends, rents etc.) are based on registers. This appears to be the case in the Nordic countries, the Netherlands, Slovenia, France, and Switzerland while register data on earnings and transfers are used in Austria, and on transfers in Latvia[(2)]. Registers as a part of mixed methods are used in some other countries (e.g. Italy). The income distribution refers to the distribution of the standard modified-OECD equivalent household disposable income allocated equally to household members. Incomes are normalised to median, and individual is the unit of analysis.

The identification of the well-off depends on the unit of analysis and the income definition. The affluent defined from the distribution of taxable income of persons or tax units or identified in the "rich lists" can be quite different from the distribution of equivalent household income among persons, which is the definition used in this paper. The family members of high-income persons can be lifted into the affluent group even if they have no personal incomes, and those who have to support family with just own earnings can move down in the income distribution. The issue of units and income definitions is elaborated further later when comparing EU-SILC top income shares to the World Top Incomes Database.

It is important to note that capital gains are not included as income in EU-SILC. Capital gains can be considered as income. They are concentrated to the very top, are quite volatile and can increase top income shares markedly. Because of the concentration to the very top, they are not likely to affect headcount measures or rank-based measures significantly. Capital gains are typically measured from registers, which means that what is measured are taxable realised capital gains. What is taxable in a country is very important, and serious comparability issues may rise. For instance, in the height of the internet bubble in 2000 capital gains added more than 5 percentage points to Gini coefficient in Sweden, whilst in Finland the increase was around 2 percentage points and around 1 percentage point in Norway (Törmälehto, 2006). This was due to more extensive definition of capital gains in Sweden.

Top coding or other censoring or truncation of top incomes in the micro data would be a problem for the analyses. In general, we consider the UDB data as not top-coded, at least in the sense of top-coding typically applied in this context. For instance, the Luxembourg Income Study top-codes the data at 10 times the median of household non-equivalised income. Likewise, the income data in the U.S Census Bureau's Current Population Survey is subject to top-coding (Armour, Burkahuser & Larrimore, 2014). In the computation of Eurostat indicators, to our knowledge, no systematic top- or bottom-coding has been used[(3)].

The anonymised UDB data set have been subject to certain measures to ensure disclosure control and confidentiality, both at Eurostat and country level. According to the EU-SILC quality reports and documentation, there are no common rules on top-coding of income data that would affect all countries. Some countries have instructed Eurostat to apply "specific rules" to protect confidentiality of income data. These countries are Estonia, Finland, Slovenia, and the UK.

In Estonia, households with three highest gross incomes (HY010) are first selected. For these households, the values of gross income, disposable income, dividends, and taxes are replaced by their weighted means, and other related income components are proportionally adjusted. In Finland, income data for some households are perturbed (not top-coded) to protect confidentiality. In Slovenia, income variables are coded in classes and the values replaced by the center of the class. The values in the highest class are replaced by the mean of the values above the lower threshold. In the UK, all household and personal income variables are rounded to the nearest 50 euros.

The number of cases at maximum income value in the micro data indicate possible top-coding (see Eriksson, 2012, for analysis with the LIS data). This was examined at household level (although top coding may occur for personal variables), and for total gross income (HY010), disposable income (HY020), rents (HY040), and interest and dividends (HY090). The method detected the top-coding applied in Estonia and the UK[(4)], but not in Finland, Slovenia, or any other country.

(2) However, the quality report of Switzerland was not available and the quality report of France was only in French. In Finland, interest received is based on interviews and in Denmark measured as net interest (interest received-paid).

(3) In the income distribution indicators, the problem of negative values seems to be a more prominent problem - for instance, Eurostat publishes Gini coefficents for net disposable income before transfers (HY023), which are heavily influenced by negative values.

(4) In these countries, the three largest values of disposable income were exactly the same.

This paper is mostly concerned with non-sampling errors and uncertainty relating to e.g. richness lines, but standard errors of top 5% income shares for selected countries are reported when comparing results to the World Top Incomes Database. These were approximated using the Rao-Wu rescaling bootstrap method based on 1 000 replicates (Rao, Wu & Yue, 1992). Unfortunately, the EU-SILC UDB still does not contain the necessary design information for variance estimation, since primary sampling units and stratification variables are not included or are incomplete. The lack of design variables was partially circumvented using the pseudo-design variables created by Goedemé (2013). Calibration to margins could not be taken into account, which could result in too wide confidence intervals, but this depends on the auxiliary variables used in the calibration. The estimated standard errors are indicative of the true sampling variances, although they are likely to be biased due to imperfections of the users' database.

2.2 Household surveys and top end of the distribution

Household sample surveys are often expected to perform poorly in the tails of the distributions. In particular those with very high incomes and their self-reported incomes are thought not to be well represented in sample surveys. The sample may not cover the very well-off units, or only to a limited extent, leading to sampling bias and variance. The measured data may contain errors, and also errors of estimation may arise when generalizing the results to the population level.

Regarding ***representation of units,*** estimating characteristics of rare domains (sub-populations), such as top 1%, may require specific sampling methods. It may be that dual frames and/or highly stratified samples would be needed to adequately reach the very well-off (e.g. Kennickell, 2007). Over-sampling of high income households would also imply lower sampling weights and more precise estimates. Understandably, such oversampling designs generally are not used in EU-SILC[5].

The number of households in the samples above 200, 250,300 and 500% of median are shown in Table 1 in the annex. There is significant variation across countries, resulting from actual thickness of the tails and the sampling designs. Some EU-SILC -specific limits for feasible high income thresholds can be derived from the minimum sample size requirements. To avoid flagging the results, the sample size in the group having high incomes should exceed 50 households in all of the countries[6]. In the 2012 data, the minimum threshold with at least 50 households in every country was 280% of median. Requiring at least 100 households would lower the threshold further to 230%. Given this, the sample size restrictions imply that the upper affluence for practical purposes could be limited to 250% of median. To derive the lower bound, we will later make use of non-income information. The computations for this paper were mostly done from 200% of median upwards.

Further to sampling ratios, a common concern is unit non-response; the households in the tails are not under-sampled but they participate poorly. A working hypothesis is that unit non-response is correlated with income level, and that those in the top of the distribution have high unit non-response rates. This differential unit non-response may cause bias which is not easily compensated with weighting and calibration (see Vermeulen, 2014). A more general consequence is that the achieved net sample size is smaller, which increases standard errors of the estimates.

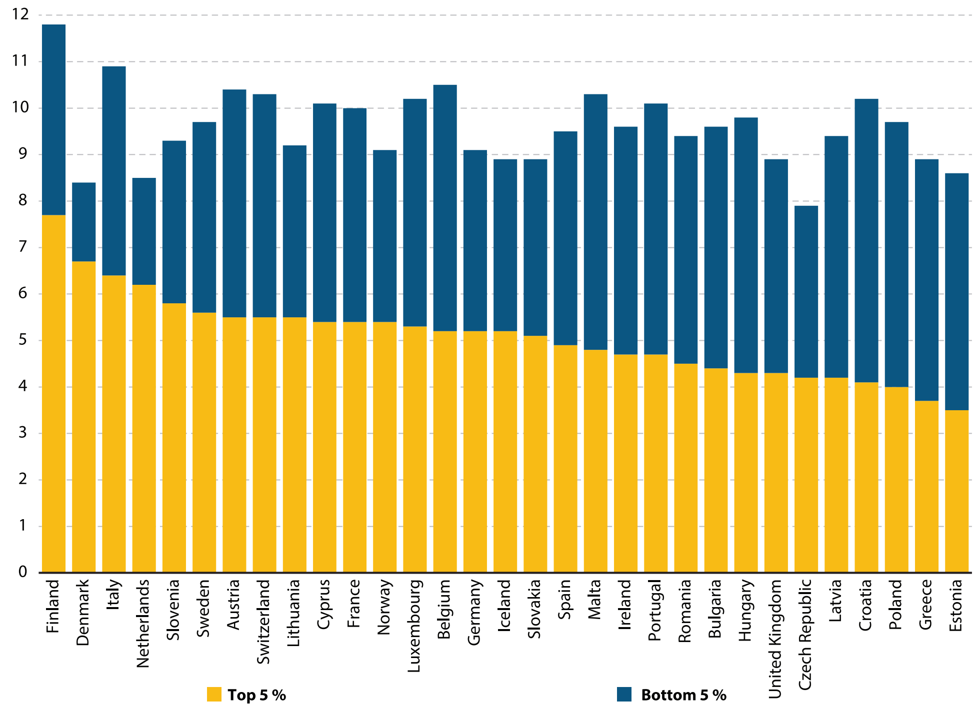
Without auxiliary data, we cannot examine how well the sample represents the top of the distribution. Neither can we observe response propensities by income groups from the UDB data. Some insight on sample representation can be gained by looking at the allocation of sample observations in the tails of the (estimated) income distribution (Figure 2.2.1). One would expect to have similar shares in the sample and in the (estimated) population if the sample was drawn randomly from the population. A disproportionate sample allocation may result from sampling design (e.g. stratification) and/or from differential unit non-response. If households in the tails of the distribution are less likely to respond, the share of the sample may be lower than the share of the population.

As shown in Figure 2.2.1, there's more than 5% of the sample in the top 5% in about half of the countries, and these countries generally have less than 5% of the sample in the bottom 5%. The two countries with the highest shares in the top are Finland and Denmark, which both sample persons but have otherwise different sampling strategies: Finland stratifies the sample to heavily over-sample high income households, while Denmark uses simple random sampling. The lowest achieved sampling rates in the top are in Estonia and Greece. Somewhat worrying for the aims of EU-SILC on poverty measurement is that there is close to or less than 5% of the sample in the bottom vingtile in many countries.

(5) An exception is Finland which heavily over-samples high income households in EU-SILC by stratifying based on source and type of income.

(6) According to guidelines for publication based on EU-SILC UDB, cross-sectional results based on less than 20 and 20-49 observations should be marked separately.

Figure 2.2.1: Sample allocation in the tails: unweighted proportion of people in the sample belonging to the weighted top/bottom 5% of the population, Income Year 2011/ Survey Year 2012
(% of persons in the sample)



Reading note: EU-SILC net sample size in Denmark was 5 355 households and 13 352 persons. Of the 13 352 persons in the sample, 901 or 6.7% were in the estimated top 5% of the population and 224 or 1.7% in the bottom 5%. The vingtiles were constructed on the basis of equivalent disposable income and weighted by sampling weight (DB090) multiplied by household size (HX040).

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Whilst sampling bias is difficult to measure, *sampling variance* of an estimate can be quantified by estimating it from an observed sample. Since the sample sizes in the tails are small and the distributions skewed, some of the indicators (e.g. top income shares) could have low precision and wide confidence intervals. EU-SILC samples are often complex multi-stage stratified samples, and weights are calibrated to auxiliary data. These would have to be accounted for in the variance estimation, in particular as there are important differences in the sampling designs and calibration models across the countries. Stratification (over-sampling) and calibration could lead to more precise estimates while clustering reduces effective sample size and increases standard errors.

Some confidence intervals of the top 5% shares are reported later alongside the comparison to the World Top Income Database. Some rules of thumb are likely to be needed, since variance estimation for many different indicators and affluence lines is in practise not feasible. In our estimations, the estimated relative standard errors of top 5% income shares tended to be around 2-3% and no higher than 3.6%, implying that safety margin of around 7% could be used as a rule of thumb to control for sampling error. That is, if the top 5% income share is 15%, it could be assumed that the 95% confidence interval is not likely to be wider than +/- one percentage points. With income share of 20%, the margin would be no higher than 1.5 percentage points. Moving up the distribution, more margin would be needed, in relative terms, because sample size decreases.

A distinct problem are *measurement errors*. The observed household income in EU-SILC contains measurement error, and the measurement error may be positively correlated with (true) income level. Moreover, the measurement errors are likely to be more severe in the "survey" countries which collect income data via interviews, due to higher item non-response or misreporting compared to register data[7]. Measurement errors may be more serious with non-regular income components, such as dividends and other property income, and may result in severe under-estimation of property income totals. Comparisons with the national accounts aggregates often show more severe under-estimation of property and self-employment income totals (Mattonetti, 2013; Alkemade and Endeweld, 2014; Törmälehto, 2006). This may result in under-estimation of the proportion of income attributable to the top of the distribution.

[7] In self-reported tax data, tax evasion and coverage problems may be a problem as well. It is reasonable to assume that in most countries register data contain less measurement error than interview-based data, particularly when the data are reported by a third party to the register authorities.

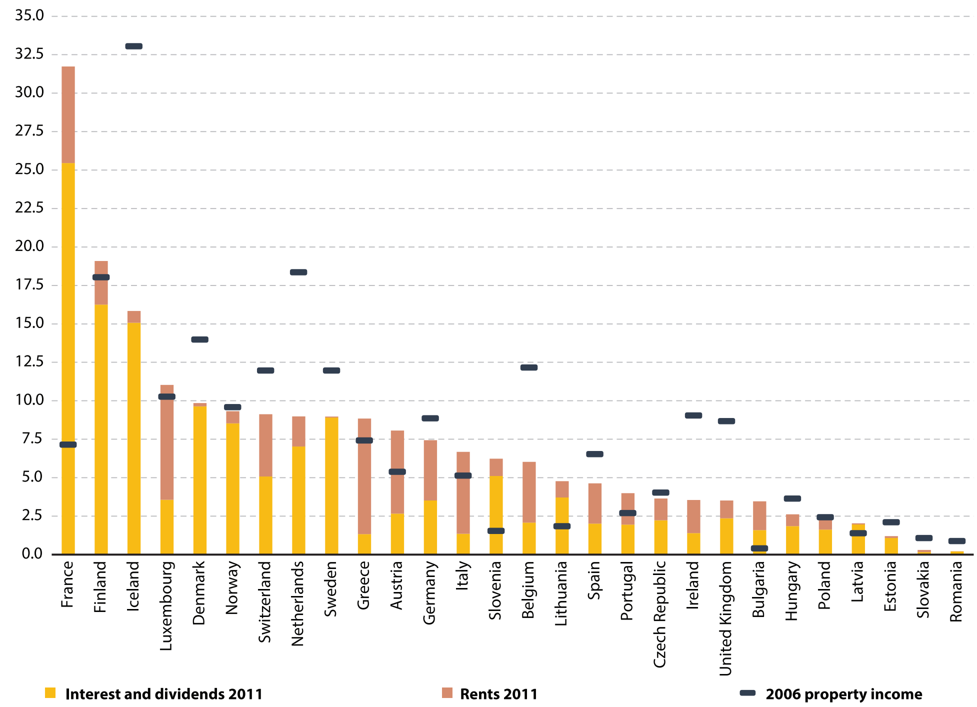
Moreover, registers can be used in calibration of survey weights, and consequently ***errors in estimation*** (bias and variance) in the top tail can be much less severe in the register-based SILC implementations. There is variation among the register countries in their use of auxiliary register data in estimation, and this may potentially explain some of the observed differences.

Given this, Figure 2.2.2 shows the shares of property (capital) income in the top 5%. Capital income consist of rents, interest, dividends and profit sharing as a sleeping partner, but excludes capital gains. Register countries have higher shares than survey countries, but apart from France, Finland, and Iceland the differences are perhaps not as large as one would expect. There is very significant variation among the register countries. France stands out as having by far the highest share of capital income in the top 5%, followed by Finland and Iceland. This can be contrasted with Sweden and Denmark, with lower shares and where capital income consists mostly of interest and dividends[8]. Luxembourg and Greece have high share of rental income.

Labour income dominates the upper part of the EU-SILC distribution, and this holds also for the top 1% in all countries (not reported here). The EU-SILC samples are not sufficient to detect a point where capital incomes would become the main income source, which one would expect to be the case with the truly affluent even in the absence of capital gains. This would often mean going to the very top, such as the richest 0.1% of the population.

Although capital income shares are above 5% in nearly all register countries, this cannot be taken as an evidence of under-reporting in the survey countries. The figure also shows the share of property income in 2006. The changes in most cases reflect the effects of the financial crisis, but also changes in measurement. A striking example is France, which changed from interviews to register-based incomes in EU-SILC 2008. There is a conspicuous increase in the share of property income: for the top 5%, it jumped from 7.1 to 32.6% from 2006 to 2007 (income reference year).

Figure 2.2.2: Property income, top 5%, income 2011 (EU-SILC 2012) and income 2006 (EU-SILC 2007)
(% of total gross income)



Reading note: In the top 5% of France, the share of interest, dividends and rents was 31.7% of pre-tax household income (gross income) in 2011. It was 7.1% in 2006.

Source: Authors' elaborations from the EU-SILC users' database 2007 and 2012 (version August 2014).

(8) Property income in Denmark is not fully comparable to others, because Denmark measures net interest (received-paid).

2.3 Extreme values

Outliers are often defined in terms of multiples of central tendency, such as those exceeding five or ten times the median income (e.g. Neri et. al., 2009; Luxembourg Income Study Key Figures). We next look at the prevalence of extreme outliers in the data set, following a semi-parametric Pareto modelling methodology for detecting and handling non-representative outliers proposed by Alfons and Templ (2013). In semi-parametric estimation, the basic idea is to use the empirical distribution function below a certain threshold, while replacing the values above the threshold with parametric estimates (ibid.; Van Kerm, 2007; Vermeulen, 2014). A strand of literature has also replaced tails to overcome effect of top-coding in the source data (Armour, Burkhauser & Larrimore, 2014).

Inequality measures computed with semi-parametric estimation may be much less sensitive to extreme values (Cowell & Victoria-Freser, 2007). Van Kerm (2007) evaluated the impact of adjustments of extreme data on EU-SILC indicators, with emphasis on the robustness of cross-country comparisons to alternative adjustments. He considered simple adjustments (trimming, winsorising) as well as model-based parametric adjustments of the tails. Ordinal comparisons were usually robust, but cardinal comparisons were more sensitive.

The distribution of income (and wealth) in the upper tail is often assumed to follow a power law distribution, such as Pareto distribution, although this view has been recently empirically contested by Brezinski (2014). We nevertheless proceed with the Pareto hypothesis. Apart from crude visual methods, verifying the power law hypothesis with EU-SILC data is well beyond the scope of this paper (see Clauset et. al, 2009).

The (complementary) cumulative Pareto distribution function of income ***y*** is the following:

(1) $1-F(y)=({}^{k}/_{y})^{\alpha}$, where $\alpha>1$, $k>0$

where k is the scale (threshold) parameter above which the power law is assumed to hold, and α is the shape parameter, which measures the heaviness of the right tail. Lower α implies fatter upper tail, and it is an inequality measure in itself. In this paper, we try to fit Pareto distribution to EU-SILC data in order to identify extreme outliers that deviate from the (assumed) Pareto model and to have a maximum likelihood estimate of shape parameter.

The outlier detecting method proposed by Alfons and Temple (2013) is the following. First, a theoretical Pareto-model is fitted to the values above a threshold. Second, outliers are identified as observations that deviate from the Pareto model. Third, the influence of the outliers can be reduced by replacing the outliers with values drawn from the fitted distribution(9).

There are many ways to estimate the shape parameter of the Pareto distribution, and different methods may yield quite different results. Some methods are suitable for tabulated data, but in this paper the parameters are estimated from micro data. Since EU-SILC sample sizes are small in the tails, robust modelling methods that take into account sampling weights are called for. The estimations here are based on the freely available R-package "laeken", implemented by Alfons and Temple (2013). There are various estimators implemented in the package, of which we eventually chose to use the conventional Hill maximum-likelihood estimator with sampling weights(10). The estimated Pareto tail indexes ranged from about 2.5 to 4.5 on the 300% of median threshold (see Table 2 in Annex).

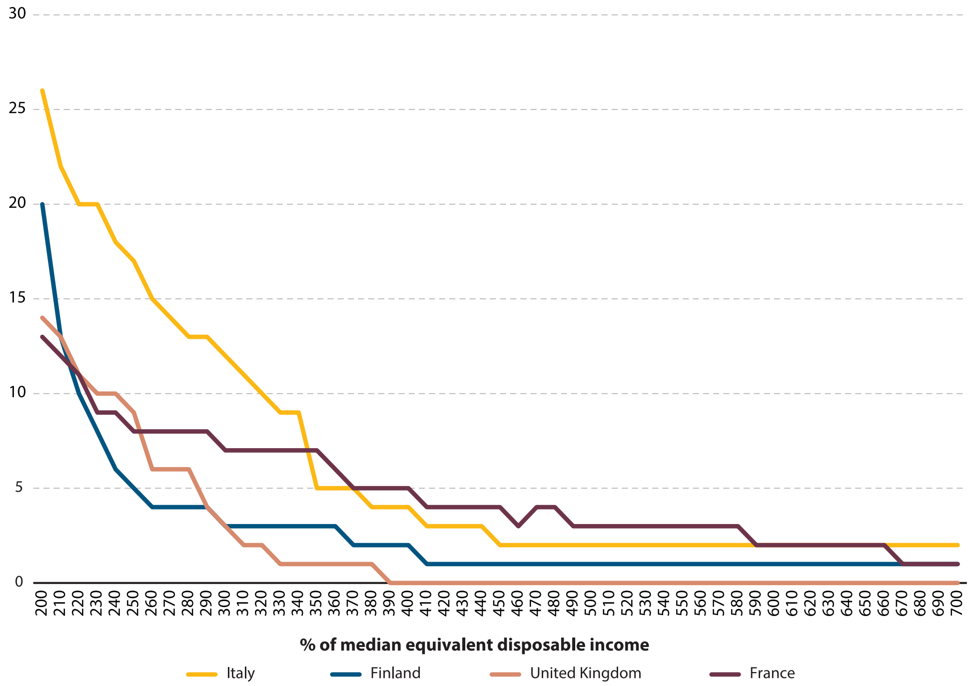
Regarding the threshold parameter k, it is often simply assumed to be e.g. two or three times the median or a certain quantile. Clauset (2009) suggested that power law distributions should be fitted to the data for various thresholds with maximum likelihood methods, choosing the threshold with the lowest value of Kolmogorov-Smirnov test statistic. We do not estimate or assume the threshold; rather, we go through a range of thresholds from 200% of median upwards.

The estimated shape parameters and the number of outliers detected for selected thresholds are reported in the annex. Figure 2.3.1 illustrates the results for three countries. Both the absolute numbers and shares of outlying values are low, in particular with the higher thresholds. If we take thresholds such as 5 times the median as the benchmark, as proposed by some authors, there number of outliers is zero in most countries. This suggest that extreme, non-representative outliers may not be prevalent in the data. Moreover, there is no guarantee that the values picked by the method are non-representative. Repeating the procedure with Finnish register data covering the whole population, true and representative extreme values were detected in the upper tail.

(9) Two other methods suggested by Alfons and Templ, reweighting and winsorizing ("shrinkage"), are not used here. Reweighting, i.e. putting weights of assumed population uniques in the sample to one and re-calibrating, is a method sometimes used by statistical offices to treat extreme outliers.

(10) We did not use the package's default partial density component estimator, on the basis of stability of shape parameter estimates along the thresholds in EU-SILC (Hill plots), weights and sample sizes, as well as the discussion on robustness of Pareto estimators in Brzezinski (2013). Detailed results of different estimation methods over the range of 200 to 700% of median are available upon request from the author.

Figure 2.3.1: Number of outlying observations based on Pareto fitting, selected countries, threshold 200 to 700% of median, income 2011 (EU-SILC 2012)
(number of households in the sample)



Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

2.4 Adjustments of the tail for outliers

Our next step is to go through the thresholds and adjust the data by replacing outlying values with values drawn from an estimated Pareto distribution. Two indicators were computed from the new distributions: Gini coefficient and the share of income accruing to the population above the threshold.

The values above the threshold can also be replaced with constants, i.e. by top-coding(11). For instance, the outlying values could be replaced with sampling-weight weighted mean of incomes above the threshold, which is a practice of some statistical institutes to protect confidentiality, with the advantage that aggregate amounts and income shares are not affected. We report two alternatives, which do affect the top income shares. The first is simple top-coding to ten times the median of non-equivalised income, which is the method used by the Luxembourg Income Study in its "*Key Figures*". The second is "fixed multiple Pareto imputation", which is based on the relationship between mean incomes over threshold and the Pareto coefficient. The theoretical expected value of Pareto distribution with threshold parameter k and shape parameter α is:

(2) $$E(x)=\frac{\alpha k}{(\alpha-1)}, \quad \alpha>1$$

By equalling this to sample mean over the threshold k, the mean over threshold is:

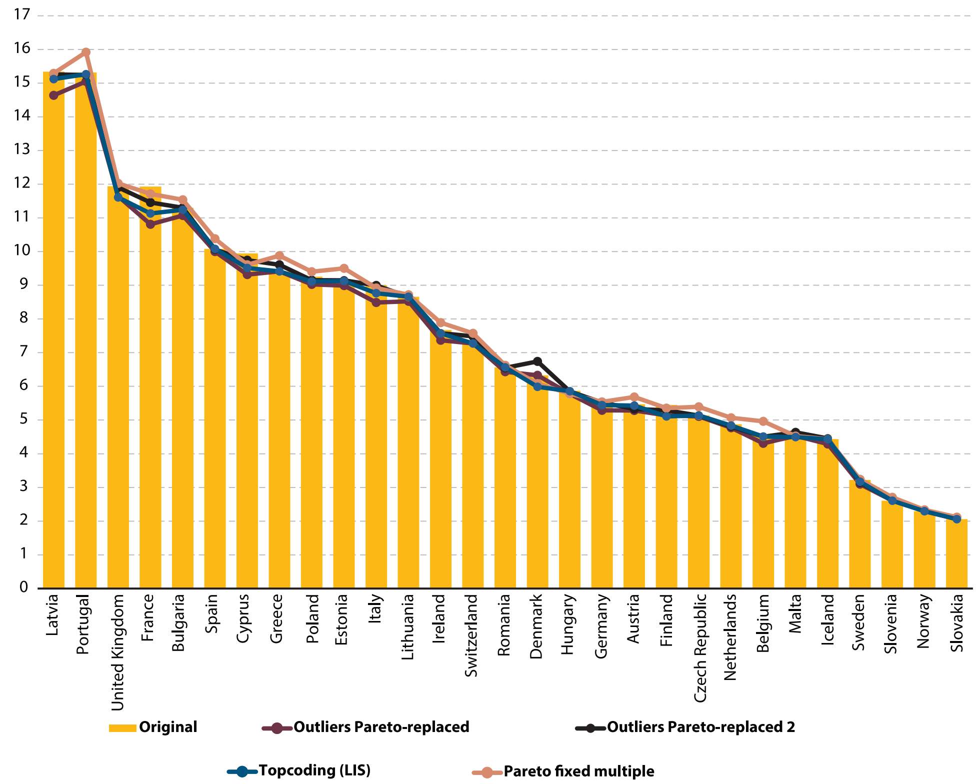
(3) $$\mu=\left(\frac{\alpha}{\alpha-1}\right)*k$$

(11) Trimming is also sometimes used, but deletion of extreme observations is clearly inferior to any other method, and was not tested here.

After estimating the parameter α, the values above the threshold are replaced by (3). That is, if α is estimated to be 3 and the threshold is 100 000 euro, then all values above that would be top-coded to 150 000 euro.

All adjustments were done for thresholds 200 to 700% of median, but Figure 2.4.1 below reports the results only for the threshold 300% of median. Rankings of countries are not much affected with any of the methods. The large maximum values observed in some countries (France, Italy, Finland) decrease significantly, and the largest income relative to median drops from 108 to 18 times the median. Two variants of the Pareto-replacement strategy are shown, although there is not much difference between them. In the first, the outlying values are replaced with values drawn from Pareto-distribution with ***k*** set at 300% of median and α estimated from the micro data (see annex). This means that the outliers may be replaced with values below the smallest outlier value. In the second variant, the replacement values must not be smaller than the minimum of the outlying values. Using Pareto fixed multiple imputation, the income shares in some cases increase.

Figure 2.4.1: Income share of population above 300% of median under various adjustments of outliers, income 2011 (EU-SILC 2012)
(share of equivalent disposable income, %)



Reading note: Income share of those above 300 percent of median in Italy is 9%. Replacing 12 outliers with values drawn from Pareto distribution reduces the income share to 8.5%.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

2.5 Adjustments of the tail for measurement errors

The preceding section indicates that there are no serious outlier problems in the data sets. In contrast, it may be that this results from not measuring the top incomes correctly. Under-estimation of top incomes could be adjusted for with parametric tail adjustments, using external benchmark to assess the size of the measurement errors. For instance, Vermeulen (2014) used the Forbes list of extremely wealthy to improve estimates of wealth survey micro data. Lakner and Milanovic (2013) have proxied the missing top incomes with discrepancy between survey and National Accounts consumption data, and allocated this to the top using Pareto fitting.

While similar adjustments could be conceived with the EU-SILC data, there is no Forbes list of very high incomes available, and National Accounts based adjustments are not straightforward. There are register-based income statistics in some countries, which could serve as the benchmark, but such an exercise for all countries is beyond the scope of this paper. For the Nordic countries at least, top income shares can be computed from register-based sources with almost the same definitions as in EU-SILC. If the survey weights are properly calibrated and income measured from registers, the top income share estimates should not be far from census-type sources. For instance, the use of registers and appropriate calibration should ensure that SILC definition-adjusted top income share estimates in Denmark and Finland are close to their population values (Quitzau, 2013; Statistics Finland).

Some idea of the sensitivity of the results in EU-SILC can be gained by replacing the whole tail instead of the outliers with estimated Pareto distributions, using hypothetical Pareto coefficients. The strategy is to fatten the tail by going through a range of Pareto coefficients, in the absence of external data on incomes exceeding EU-SILC maximum values. Otherwise the method is in principle the same as in Vermeulen (2014), who combined observations from the Forbes listed with wealth surveys, estimated the Pareto coefficient, and then used this to draw observations from Pareto distribution and replaced survey values with the fitted values.

Some justification for the assumed tail indexes in terms of under-estimation of incomes in the top can be derived from equation (3) by expressing the parameter α in terms of the mean and the threshold as:

(4) $$\alpha = \frac{\mu}{(\mu\text{-k})'}$$

where μ is the sample mean of those above the threshold *k*. That is, the empirical estimate of tail index is the ratio of mean above threshold to the difference between mean and threshold. Fixing the threshold, we can guess about the ratio of under-estimation in the top and adjust the mean incomes to derive empirical alphas. Table 2.5.1 illustrates the relationship to under-estimation of incomes.

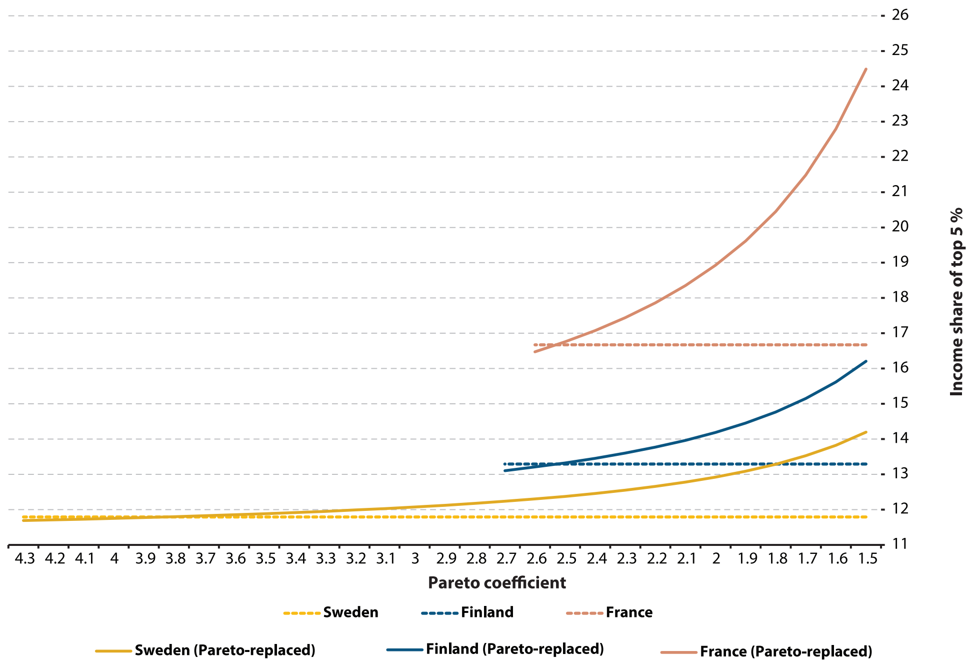
Table 2.5.1: Example of assumed Pareto tail indexes under different top income under-estimation scenarios

Threshold	Assumed ratio of under-estimation above the threshold	Mean above threshold	Empirical Pareto index
30 000	0	40 000	4.0
30 000	20 %	50 000	2.5
30 000	35 %	61 538	2
30 000	50 %	80 000	1.6

In what follows, the threshold is fixed to 300% of median, and the distribution above that is replaced by drawing values from Pareto distribution with different shape parameters(12). Top income shares and Gini coefficients are then computed from the adjusted data. The estimations were done on the range of Pareto coefficients from 5 to 1.5 by 0.1, but the results here are reported only for selected under-estimation scenarios. For most countries, we are drawing from a heavier-tailed distribution than is estimated from the data itself. The top income shares above the 300% threshold of course increase by the same amount as mean incomes, so the focus here is on top 5% of the distribution.

Figure 2.5.1 provides a three-country illustration of the procedure. The estimated top 5% income share in Sweden is 11.8%, in Finland 13.3%, and in France 16.7%, shown by the horizontal dotted lines. The estimated Pareto tail coefficients are 4.3, 2.8 and 2.6, respectively, above the 300% threshold. The simulated top income shares are roughly the same as the original estimates with these Pareto coefficients, but increase with lower coefficients. Even with simulated values based on alpha = 1.5, corresponding to close to 40% of under-reporting, the Swedish and Finnish estimates are lower than any of the top shares in France. This suggest that the results are quite robust to under-estimation of income. The Swedish Pareto-replaced curve crosses the Finnish original estimate at around alpha = 1.8. This implies that results are robust up to 30% of under-estimation in the Swedish incomes above 300% of median(13). In fact, since all countries measure incomes from registers, the under-estimation due to measurement errors should not be large.

Figure 2.5.1: Top 5% income shares in Sweden, Finland, and France, income 2011 (EU-SILC 2012), EU-SILC original estimate & estimates from semi-parametric Pareto simulations.
(coefficient and %)



Reading note: The top 5% share in Sweden is 11.8%, as shown by the dotted line. The Pareto index computed from the data is 4.2. The solid line shows the hypothetical top 5% shares when actual data above 300 x median are replaced with values drawn from heavier-tailed Pareto distributions. Lower Pareto coefficient implies hypothesis of heavier upper tail than is actually observed, assuming more under-estimation of top incomes. Even with very severe under-estimation (Pareto coefficient of 1.5), the top 5% share in Sweden remains lower than in France (both actual and simulated values).

Source : Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

(12) This was also carried with the top 1%, but the number of observations was quite low in some countries.

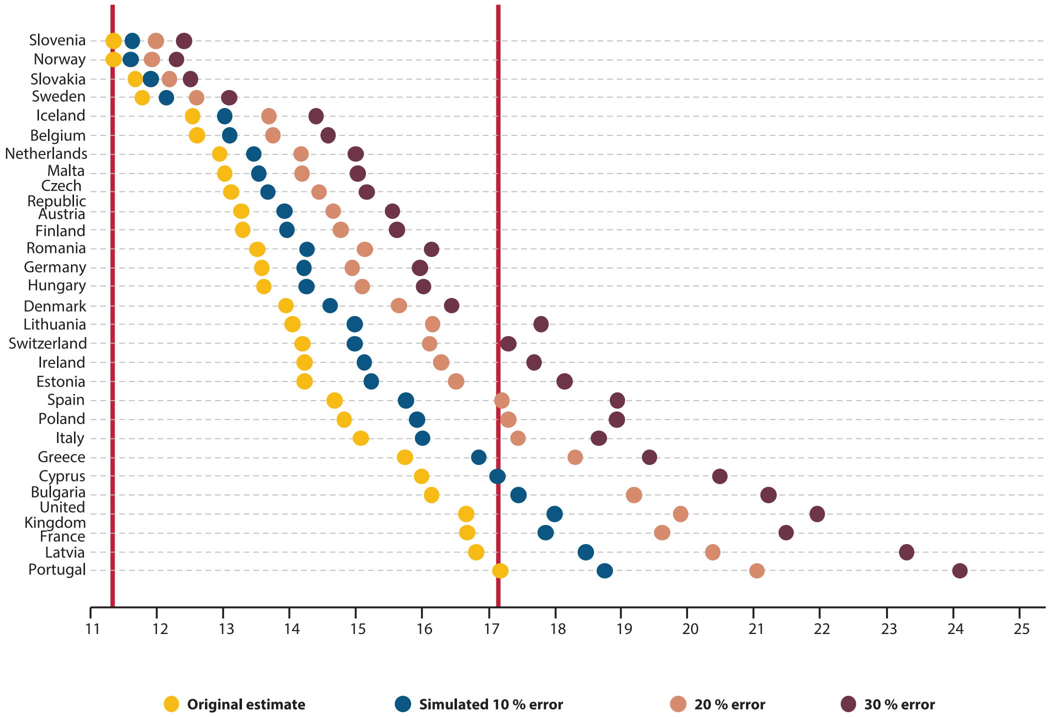
(13) The threshold in Sweden was 74 161 and mean above threshold 96 245 euro in the original distribution. Replacing the tail with values drawn Pareto with alpha = 1.8 would result in top 5% share of 13.3%, i.e. the same as in Finland. The mean above threshold would be 140 438 euro, so that the original mean above threshold is 68% of the Pareto-fitted value, reflecting the assumption of around 30% under-estimation.

Figure 2.5.2 reports the original top 5% income shares estimated from the data as well as the income shares based on simulated values corresponding to 10 to 30% under-estimation of top incomes. The original values range from 11-12% in Slovenia and Norway to around 17% in Portugal, Latvia, France, and the UK. The original estimate is always the lowest because the incomes of the top of the distribution where replaced with values drawn from Pareto distributions with lower Pareto coefficient.

Three other income shares are shown, based on different Pareto coefficients but corresponding to increase of 10, 20 and 30% to the mean above the 300% of median threshold. The reference lines show the income share of Sweden assuming 30% under-estimation of top incomes and the share of Portugal assuming no measurement error (i.e. the original value). The 30% assumption generally means Pareto coefficients of around 2 or even lower, which seems rather low for equivalent disposable income.

Reading off the chart, one could construct different scenarios. For instance, it could be assumed that register countries have no measurement error (first dots) and that survey countries have 10 or 20% of measurement error (second or third dot). This would imply some re-ranking, and the countries with low overall inequality and using registers would have even lower shares. France is an exception as a register country, and for instance the UK would have higher share assuming that its interview-based incomes are under-estimated. The figure is indicative only, but it seems that the Pareto-replacement could be a viable tool for sensitivity analyses. The hard task for further work would be to have some reasonable assessment of the extent of the actual under-estimation of top incomes across the countries.

Figure 2.5.2: Top 5% income shares, original EU-SILC estimates & based on Pareto-replaced values over the 300% of median threshold, income 2011 (EU-SILC 2012)
(%)



Reading note: In Sweden, the top 5% share was 11.8%. The Pareto index computed from the data was 4.2. Assuming that the mean above 300 x median is under-estimated by 20% would correspond to Pareto index of 2.3 in Sweden. If the actual values above 300 x median are replaced with hypothetical values drawn from a Pareto distribution with shape parameter 2.3, the share of top 5% would be 12.6%.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

2.6 Comparison with the World Top Incomes Database

In general, direct comparison of SILC top incomes to external benchmarks is not feasible, aside from some country-specific register sources and limited comparisons with the World Top Incomes Database (WTID)[14]. The WTID provides tax-based estimates of the shares of fixed quantile groups of pre-tax incomes in selected countries over a very long period of time. Tax data are used because of the assumption that household surveys do not capture well the top of the distribution, and because of the fact that surveys do not cover long periods of time.

The WTID income concepts, units, populations and estimation methods differ from EU-SILC, since the latter aims to measure the distribution of economic welfare rather than personal incomes. Therefore, a direct comparison is not meaningful. Below, we compare the adjusted top 5% income shares of EU-SILC to WTID, for income year 2009, by modifying the EU-SILC income concept and using personal rather than household incomes to the extent possible. The WTID estimate for Finland 2009[15] is in fact based on the national micro data of SILC, which means that the differences to EU-SILC UDB-based estimates are solely due to different definitions of income and income receiving unit, and sampling and non-sampling errors play no role. Therefore Finland serves as the benchmark in the comparison.

Figure 2.6.1 illustrates how the WTID estimates differ from SILC estimates because of different definitions in Finland 2009. The WTID estimate for Finland is 20.7%, which is the share of taxable income of persons over 14 years of age (Jäntti et. al., 2010). While the WTID only reports the top shares, the whole distribution is shown in the figure. Many of those who are over 14 years of age do not have taxable income, implied by the zero income shares in the bottom of the distribution. The concept of taxable incomes in Finland excludes tax-free incomes, which include many social transfers received in the very bottom (e.g. housing allowances) but it also excludes for instance tax-free dividends which accrue to the very top. In contrast, it includes realised capital gains.

The EU-SILC income concept captures tax-free incomes but excludes capital gains. The figure also shows adjusted personal incomes from SILC, defined as the share of all personal *pre-tax* incomes plus household pre-tax property income divided by persons aged 16 and over[16]. The top 5% income share of 17.9 is closer to the WTID than the equivalent net household income distribution used in this paper. The remaining difference is partly due to the better coverage of incomes in the bottom of the distribution.

The Table 2.6.1 then reports top 5% incomes shares for those EU-SILC countries, which had estimates available in the WTID for year 2009. Except for Finland, the WTID estimates are based on tax data. The cross-country comparability of the WTID depends on the definitions of taxable incomes and target populations as well as the estimation method, which may differ. For instance, the WTID estimate for Finland relates to persons aged 15 and over whilst that for Spain relates to persons aged 20 and over. For the EU-SILC estimates, approximate confidence limits are provided to control for sampling variance. Standard errors were estimated with Rao-Wu rescaling bootstrap (1 000 replicates) and pseudo-design information (see discussion earlier in this paper).

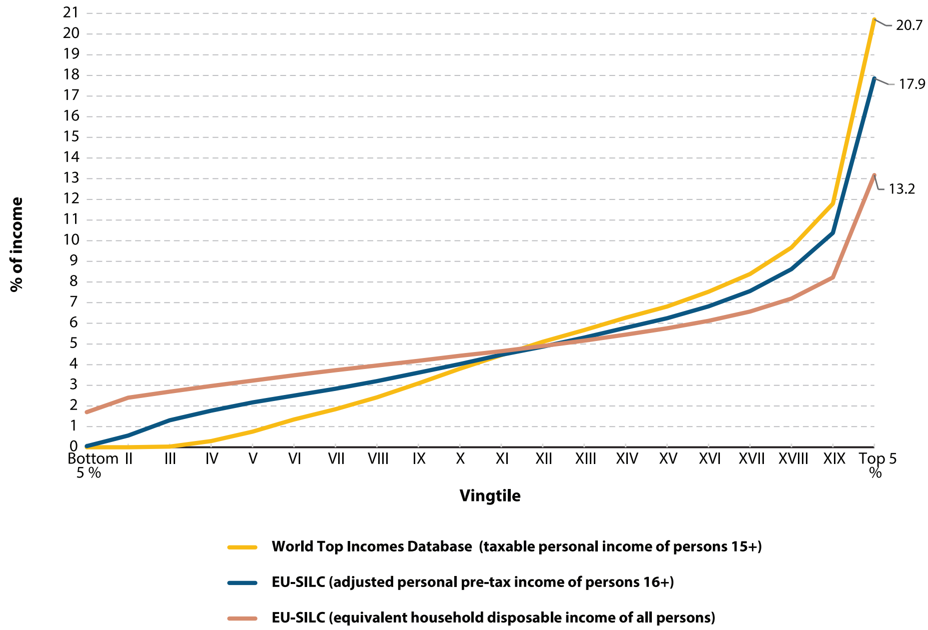
Despite the caveats of the comparison, Table 2.6.1 provides useful insight to the quality of EU-SILC estimates of top income shares. The results suggest that the EU-SILC estimates are not that incoherent with the WTID. For instance, after the adjustments the top 5% income shares seem to be at the same level in France and Spain, and lower than in Italy or Portugal in both sources. Sampling error cannot be ruled out, though. Finland appears to have higher top 5 shares than Norway and Sweden, and Switzerland higher shares than the Nordic countries. The bottom row shows that top 5 shares based on equivalent person-weighted disposable incomes, i.e. the standard EU-SILC income definition. This gives further evidence on the importance of the income concept, income receiving unit and target populations as potential sources of differences.

[14] Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty and Emmanuel Saez, The World Top Incomes Database, http://topincomes.g-mond.parisschoolofeconomics.eu/.

[15] This corresponds to Top income shares –IDS series in the WTID database (IDS stands for Income Distribution Statistics, under which the results from EU-SILC are published in Finland).

[16] This is the definition closest to the WTID-definition that could be constructed from the EU-SILC UDB for Finland. From EU-SILC, we can only look at personal incomes above 16 years of age. Using 16 years as the age threshold instead of 14 years would in itself decrease the top 5 share of taxable income in Finland from 20.7% to 20.5%.

Figure 2.6.1: **Income shares in vingtiles, Finland 2009: World Top Incomes Database, adjusted EU-SILC and standard EU-SILC definitions and units.**
(vingtiles and %)



Reading note: The top 5% income share in Finland is 20.7%in the World Top Incomes Database (IDS series, data retrieved August 2014). The shares in the bottom 15% were 0%. Incomes based on the EU-SILC definitions are more equally distributed.

Source: Authors' elaborations from the Finnish Income Distribution Statistics 2009 micro data (WTID) and EU-SILC users' database 2010 (version August 2014).

Table 2.6.1: **Comparison of the World Top Incomes Database top 5% pre-tax income shares with the EU-SILC estimates (income year 2009 / survey year 2010).**
(% and percentage points)

	Finland	Norway	Sweden	Netherlands	France	Spain	Switzerland	Italy
WTID 2009*	20.7	18	18	19.1	21.4	21.4	22.9	23.2
Adjusted EU-SILC estimate, pre-tax personal incomes, 16+ persons	17.9	16.9	16.1	19.1	20.2	20.1	19.7	21.3
95% confidence limit, +/-, percentage points	*0.6*	*1.0*	*0.9*	*0.6*	*1.0*	*0.5*	*1.1*	*0.7*
Difference (pp), WTID – EU-SILC	2.8	1.1	1.9	0	1.2	1.3	3.2	1.9
Original EU-SILC estimate (2009 incomes), equivalent household DPI, person weighted	13.2	12.2	11.7	12.8	15.8	14.5	15.1	14.5
95% confidence limit, +/-, percentage points	*0.6*	*0.8*	*0.7*	*0.5*	*1.0*	*0.5*	*0.9*	*0.6*

Reading note: In 2009, the top 5% share of taxable income of 14+ persons was 20.7% in Finland in the World Top Incomes Database. Using the EU-SILC adjusted pre-tax personal incomes the share was 17.9% (+/-0.6 pp) and EU-SILC equivalent incomes 13.2% (+/- 0.6 pp). In Finland, all estimates are based on the same sample survey. In the other countries, the WTID estimates are estimated from different sources (tax data). Estimated half 95% confidence intervals in separate rows.

Sources: World Top Incomes Database (retrieved 19.8.2014) and authors' elaborations from the EU-SILC users' database 2010 (version August 2014).

3. Measures of richness and affluence

The measurement of top incomes and "richness" has evolved significantly in recent years, and new measures that go beyond simple headcounts have been introduced (see Medeiros, 2014, for a review). In this section, we examine how some of these measures would look on the basis of EU-SILC. The measures that are covered are headcounts, transfer-sensitive richness indices, affluence gaps and redistribution-based measures, and also top income shares.

3.1 The line of richness

To identify the affluent or the rich based on income (or wealth or other resources), a richness line must be specified. The cut-off can be based on absolute values or on rank in the income distribution counting from the top, i.e. counting the "millionaires" or creating "rich lists". These are not suitable approaches for sample surveys. The threshold could be fixed to a specific quantile (e.g. 99th percentile) or defined as a distance from a reference level of income (e.g. twice the median), as discussed earlier and mostly used in this paper. Index-method based methods have been proposed, such as defining threshold on the basis of marginal increase in Gini or other inequality measures. Some methods define threshold on the basis of redistributional effect, such as eradication of poverty (Medeiros, 2006).

Regarding non-income dimensions, consideration of wealth as a resource would be essential, but EU-SILC has only very limited information on ownership and values of assets. Wealth-based thresholds that have been proposed include e.g. having net worth more than 30 times the median income (Atkinson & Brandolini, 2013). Other alternatives would include reference groups or clustering based on non-income information, for instance defining middle and upper classes on the basis of tastes, habits, and occupational status. This is the very field of household surveys, but EU-SILC is not that strong in the kinds of variables that would separate upper class from middle class households[17].

It is also possible to leave the question of affluence threshold open, by ordering income distributions by a given index for all or a subset of affluence lines (Bose, Chakravarty and d'Ambrosio, 2014; Aaberge and Atkinson, 2013; Michelangeli, Peluso and Trannoy, 2010). This implies using distribution functions, and seeking for first or second order dominance of the distributions. In general, this paper aims for affluence orderings of the distributions by following the dominance approach, i.e. measures are evaluated over a range of income thresholds. Whether those above, say, 250% of median are affluent or have very high economic well-being, cannot be determined only on the basis of the rank in the distribution. For this, the level of income, other resources such as wealth, and non-income information such as subjective experiences are needed.

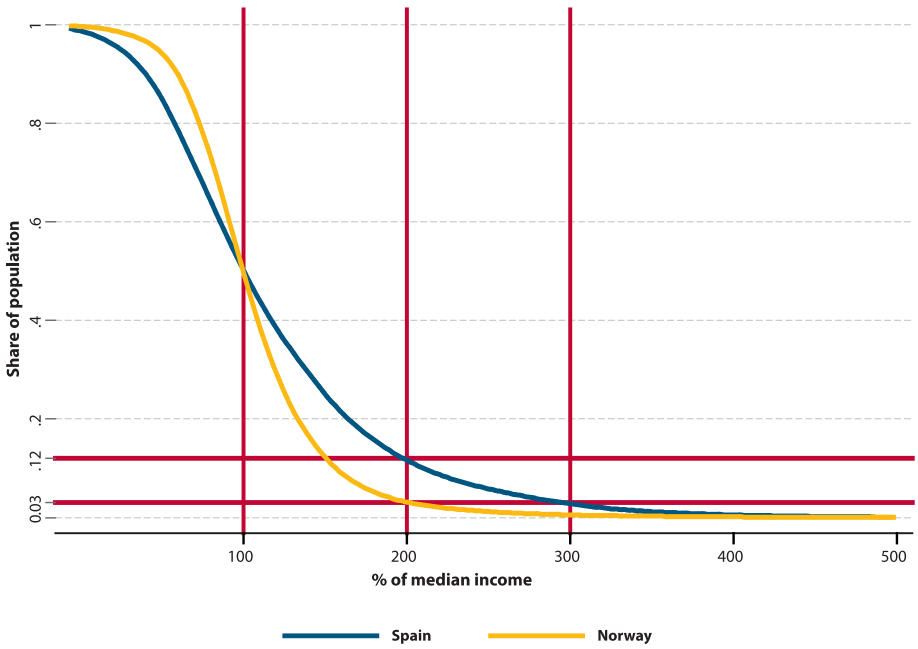
When the affluence threshold is determined relative to the distance from the median income, the size of the affluent group can be measured. This "income space" approach also relates to income poverty as well as to certain definitions of middle-income households as those with incomes between 75 and 125% of median (Atkinson & Brandolini, 2013). A number of authors have defined affluent as those exceeding a certain threshold above the median. Common limits appear to be two or three times the median.

An alternative would be to choose "people space" and use quantiles of income (deciles, percentiles) as richness lines. This would fix the population shares of the affluent and put focus on their resources (e.g. income shares of top 1%). Table 3 in the annex shows the relationship between median-normalized incomes and 95th and 99th percentiles for all EU-SILC countries in income year 2011 (EU-SILC 2012). Averaging across countries, the 95th percentile corresponds roughly to 200-250% of median, ranging from 182% in Norway to 288% in Portugal. The top 1% threshold corresponds to more than three times the median, ranging from 269% in Norway to 480% in Portugal. Both the fixed quantile and the multiplier thresholds are relative to income levels of the country, and thus neglect the differences in living standards in Europe. The 95th percentile was 73 000 euro in Norway (182% of median) and 7 000 euro in Bulgaria (410% of median).

Figure 3.1.1 further provides a two-country example of how quantiles relate to income distribution which is normalised to median. The lines show the fraction of population above the median-normalized income levels in Norway and Spain. Spain has much more unequal distribution than Norway. In Norway, the fraction of population with income above 200% of median was 3%, whilst in Spain the corresponding fraction was 12%.

[17] While not reported in this paper, an attempt was made to define upper class as persons with a university degree and working in a managerial position, living in a good neighbourhood (no noise, pollution, grime, other environmental problems, crime, violence or vandalism in the area) and in households making ends meet easily or very easily and having have capacity to finance unexpected expenses.

Figure 3.1.1: **Complementary cumulative distribution functions of income, expressed as % of median, Norway and Spain, income 2011 (EU-SILC 2012)**
(coefficient and %)



Reading note: Curves show the fraction of population above the income level expressed as % of median income. In Spain, 12% of persons had incomes more than twice the median income.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

3.2 Headcount measures based on multiplier thresholds

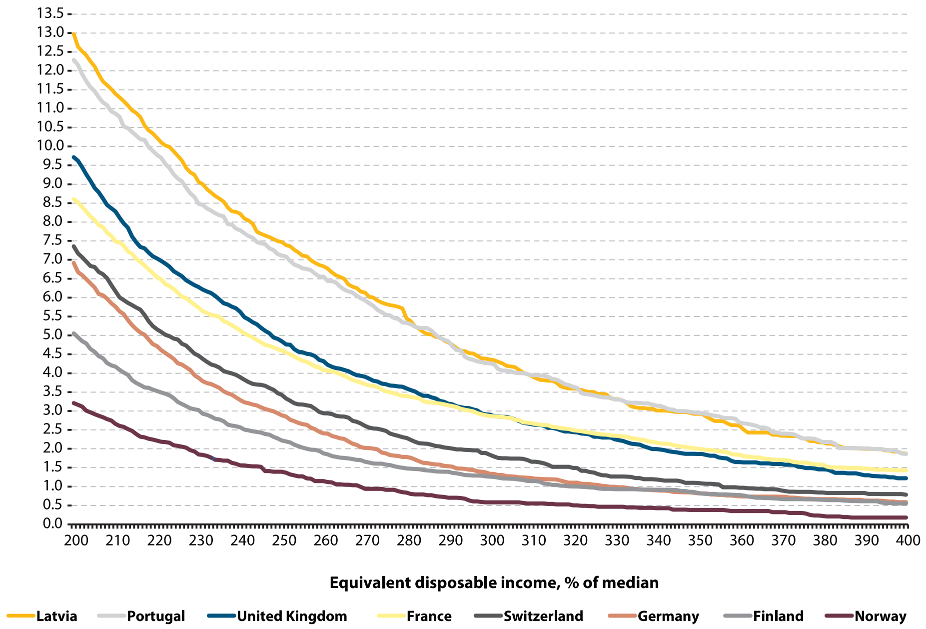
The simplest measure of richness is the share of population exceeding a high income threshold. As noted, we follow the dominance approach, and look at headcount affluence curves that show the proportion of people above a range of richness lines, determined in relation to median income. If the proportion is higher over a range of thresholds in distribution A compared to distribution B, then distribution A is the more affluent one if measured by headcount and relative to country median. In practical terms, we compute the share of population exceeding a threshold over a range of threshold, and check whether the population shares are higher at least for one affluence line and never lower. This conforms to first order stochastic "affluence" dominance (ibid.).

The headcount affluence shares were computed over the range of 200 to 400% of median income, and results for all countries for selected thresholds are reported in the annex. The results in the annex suggest that headcount rates are fairly robust to different affluence lines, and that Latvia and Portugal have the highest share of relatively high income households and Nordic countries and Slovenia the lowest shares. Figure 3.2.1 below compares headcount affluence curves for selected countries over the whole range. The figure shows that Portugal and Latvia "affluence-dominate" the other countries, but the curves of these countries cross at some points. Thus, the headcount richness is more prevalent in these countries with all thresholds, but dominance cannot be established between Portugal and Latvia. However, since no standard errors are provided, nothing conclusive about dominance can be said.

The figure also shows that the share of high income people in Norway is always significantly below of Finland, Germany and other countries. It also seems that France affluence dominates Switzerland, with both countries using register-based incomes, and also a pure survey country Germany. Aside from sampling errors, also non-sampling errors such as different data sources may affect the results. Headcount measures are based on empirical complementary cumulative distribution functions, which should be robust to rank-preserving measurement errors or other non-sampling errors, and extreme outliers.

Headcount measure suffers from the same drawbacks as relative poverty, which are even more pronounced in the European setting. Affluence headcount does not capture the differences in average levels of income, nor in the distribution among the rich. If we were to pick 250% as the threshold for a "richness" indicator, the conclusion would be that 1.4% of the Norwegians are rich compared to 7.5% in Latvia. Whilst welfare comparisons in countries with very unequal levels of income are not appropriate, comparing for instance Norway and Switzerland seems sensible, in particular because both measure incomes from registers. In Figure 3.2.1 Switzerland affluence-dominates Norway throughout the range, indicating that there are more relatively rich people in Switzerland.

Figure 3.2.1: Relative affluence headcounts over the range of 200% to 400% of median income in selected countries, income 2011 (EU-SILC 2012)
(% of persons)



Source: Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

3.3 Transfer-sensitive measures of richness

The headcount measure of number of persons above an affluence line is insensitive to the incomes of the affluent. Peichl, Schaefer and Scheicher (2010) proposed a class of indices which react also to transfers among the rich. There are two versions of the indices, concave and convex, with important underlying normative differences. The convex version is transfer sensitive in a sense that it decreases when a rank-preserving progressive transfer between two rich persons takes place, e.g. when a billionaire gives money to a millionaire. Moreover, it increases when income of a rich person increases (monotonicity axiom) and is independent of the incomes of the non-rich (focus axiom), although not on the number of non-rich. The convex index resembles both the FGT family of poverty indices, and also the excess function averaged over all units, not just over those exceeding the affluence line:

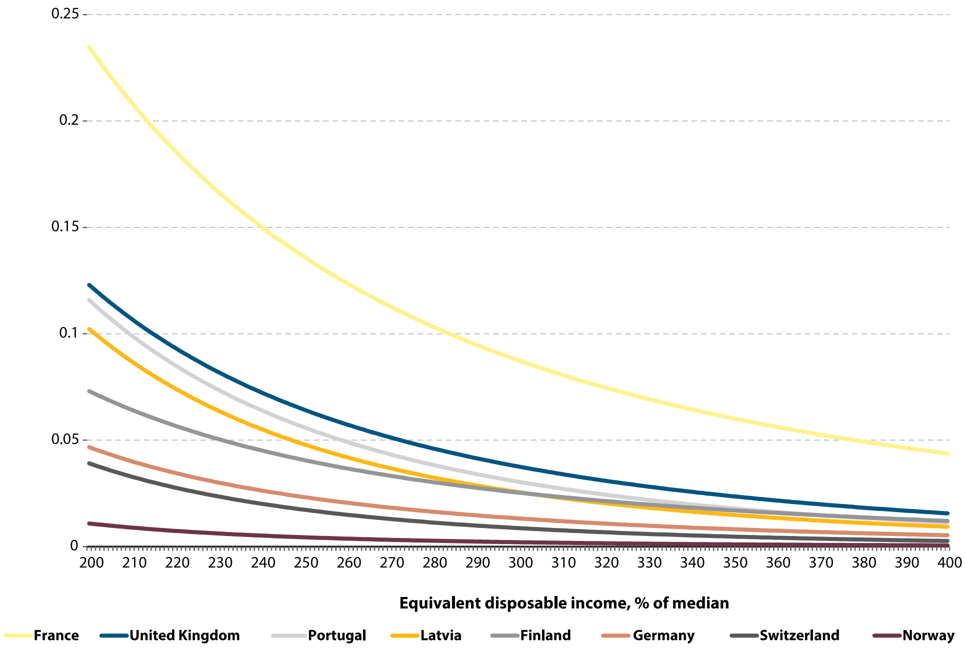
$$RT2_{\alpha}(Y_i,k) = \frac{1}{n}\sum_{i=1}^{n}\left(\left(\frac{Y_i-k}{k}\right)I_{[Y_i>k]}\right)^{\alpha}, \alpha > 1 \quad (3.4.1)$$

Figure 3.3.1 shows the convex index over the 200-400% of median range, with the normalized excesses over threshold squared, a richness measure analogous to FGT(2) poverty measure. Through the range of thresholds, the index is highest in France, second highest in the UK, and lowest in Norway. Finland dominates Germany and Switzerland and Germany Switzerland, reversing the results with the headcount measure. Thus, compared to the headcount, there are significant changes in the country rankings.

The convex index takes the maximum value when inequality among the rich is maximized, by giving all the excess to one person with all the other rich being at the richness line. The concave version of the transfer-sensitive richness index increases with more equal distribution among the rich, i.e. with more homogeneity in the top (Peichl, Schaefer and Scheicher, 2010, p. 8)[18]. The concave version increases when a rank-preserving progressive transfer takes place among the rich, i.e. when a billionaire gives money to a millionaire. This could imply a more homogenous group of rich, with more shared interests and visibility in advancing their concerns (e.g. lowering marginal tax rates). The interpretation of what is "affluence" is therefore quite different with the concave index. It also is inconsistent with a key element of distributional analysis, the Pigou-Dalton principle of transfers. Technically, the concave version is likely to be more robust to extreme values (see Brzezinski, 2010).

The results for the class of richness indices proposed by Peichl et. al. are reported in the annex for 2011 (EU-SILC 2012), based on 2.5 times the median income richness line. The convex version is indeed quite sensitive to thickness of the tail, in particular with higher values of parameter α. Figure 3.3.2 illustrates the indices for 2011 incomes (EU-SILC 2012), with countries sorted according to the headcount affluence rate. Despite the attractive theoretical properties of the transfer sensitive indices, they are not well suited to become indicators of richness for a wider audience. The interpretation is not straightforward and the choice of parameters are not easy to communicate. Headcount measures or top income shares see much more attractive from a practical point of view.

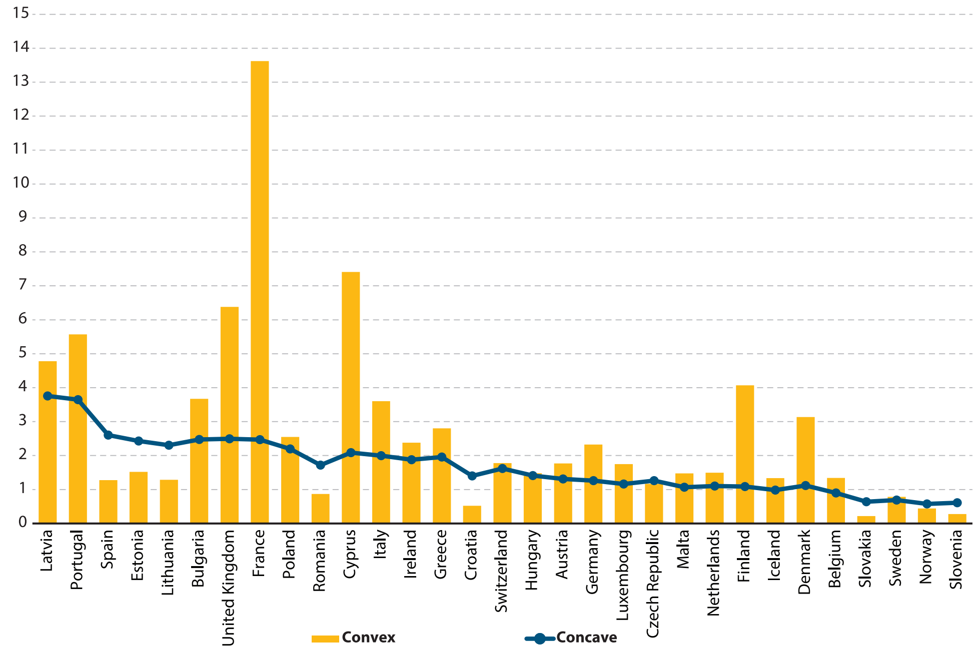
Figure 3.3.1: Convex richness index of Peichl, Schaefer & Scheicher, income 2011 (EU-SILC 2012), alpha=2
(convex RFGT_T2 index, alpha = 2)



Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

(18) The concave version is defined as: $R_{\beta}^{Cha}(Y_i,k)=\frac{1}{n}\sum_{i=1}^{n}\left(1-\left(\frac{k}{Y_i}\right)^{\beta}\right) I_{[Y_i>k]},\ \beta>0$

Figure 3.3.2: **Transfer-sensitive convex (α = 2) and concave richness (β = 2) indices (Peichl et. al.), income 2011 (EU-SILC 2012), income threshold 2.5 times the median of equivalent disposable income. Countries sorted by the headcount share of affluent.**
(index)



Reading note: Richness would be highest in France and Cyprus with the convex index (alpha=2), which increases with more unequal distribution of income among the rich. With the concave index (beta=3), richness would be highest in Latvia and Portugal. The concave index is clearly less sensitive to outliers, and it decreases with more inequality among the rich.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

3.4 Top income shares

The complementary cumulative distribution functions such as the "headcount affluence curves" do not tell about the resources available to the well-off group. Therefore, the income shares of those above a selected quantile (90th, 95th, 99th) are often used as an indicator of relative richness, or even income inequality (Piketty, 2014). The top 1%, for instance, can be considered as a small elite group with fixed size, assumed to have economic and political power beyond its 1% population share because income and other resources are concentrated to this group.

The top income shares are piecewise Lorenz-ordinates of the top of the distribution. Therefore, they often are highly correlated with inequality measures based on the whole distributions, such as Gini coefficients or decile or share ratios (e.g S80/S20). The fixed quantile approach is well suited for examining the characteristics of the well-off group or their share of total resources.

Eurostat publishes in their web database quite detailed results from EU-SILC, including income shares of each of the top 5 percentiles. Table 3.4.1 reports the shares of income accruing to those in the top 5%[19]. The dominance approach here reflects uncertainty about the proper high incomes cut-off (top 5%, top 1%). The 95th and 99th percentiles vary across countries; annex 1 provides the mapping of the quantiles to median-normalized incomes. The 95th quantile is in most countries above twice the median.

(19) The table on top income shares is based on Eurostat web database (table ilc_di010). The results for some countries differ somewhat from our own estimates from the UDB used elsewhere in this paper.

Over the range from top 5 to top 1%, Slovenia and Slovakia never have higher top income shares than other countries. Reading from the bottom of the table, dominance holds until Finland, after which we find crossings of the piecewise Lorenz-curves in this range up to Denmark. Register countries have differences. Among them, there is (point-estimate based) affluence-dominance of France-Denmark-Finland-Netherlands-Iceland-Sweden and Norway. France appears to have a much thicker tail than the "old" register countries, and it also affluence dominates Switzerland. Slovenia is tied with Norway, and comparison between Denmark and Switzerland is inconclusive.

Table 3.4.1: Income shares over the 95th to 99th quantiles in 2011 (EU-SILC 2012)
(% of total equivalent disposable income)

Country	Top 5%	p96	p97	p98	Top 1%	Country	Top 5%	p96	p97	p98	Top 1%
Portugal	17.3	14.9	12.3	9.3	5.8	**Hungary**	13.7	11.7	9.6	7.1	4.4
Latvia	16.9	14.5	11.8	8.8	5.4	**Romania**	13.6	11.4	9.0	6.5	3.8
United Kingdom	16.7	14.5	12.2	9.6	6.3	**Germany**	13.6	11.6	9.5	7.2	4.6
France	16.7	14.6	12.2	9.5	6.2	**Luxembourg**	13.4	11.4	9.3	7.0	4.5
Bulgaria	16.1	13.9	11.4	8.6	5.5	**Austria**	13.4	11.3	9.2	6.9	4.2
Cyprus	16.1	13.9	11.7	9.3	6.4	**Finland**	13.3	11.5	9.5	7.3	4.7
Greece	15.8	13.6	11.3	8.8	5.7	**Malta**	13.3	11.4	9.3	7.1	4.3
Italy	15.1	12.9	10.6	8.0	5.0	**Czech Republic**	13.2	11.2	9.1	6.8	4.2
Poland	14.8	12.7	10.4	7.8	4.7	**Netherlands**	12.9	11.0	9.0	6.8	4.2
Spain	14.7	12.3	9.8	7.1	4.2	**Iceland**	12.6	10.7	8.7	6.6	4.0
Estonia	14.3	12.1	9.7	7.1	4.2	**Sweden**	11.8	10.0	8.1	6.0	3.6
Lithuania	14.3	11.9	9.7	7.0	4.1	**Norway**	11.4	9.6	7.8	5.7	3.4
Switzerland	14.2	12.1	9.9	7.5	4.7	**Slovenia**	11.4	9.6	7.7	5.6	3.3
Denmark	14.0	12.1	10.2	7.9	5.3	**Slovakia**	11.4	9.6	7.7	5.6	3.3

Reading note: Sweden affluence-dominates Norway as it has higher income shares than Norway for all percentiles. Finland has lower share of top 5% but higher share of top 1% than Spain, so dominance cannot be established.

Source: Eurostat (table ilc_di010, retrieved August 2014).

The potential bias of these estimates was already discussed earlier. Even if unbiased, they are not likely to be precise because they are based on small numbers of observations. Regarding sample sizes, the top 1% typically has less than one hundred households in the sample. The UK estimate is, for instance, based on 79 households and the French estimate on 123 households in the sample. Consequently, it may be better to use top 5% rather than top 1% as the fixed quantile definition. Arguably, the strength of a collection of cross-national sample surveys is not accurate estimation of income shares at the very top, but having comparable data on the characteristics of those in the top quantiles.

Another use for the fixed quantiles would be to define absolute affluence thresholds from a supranational distribution. Fixing the threshold to quantiles of the European income distribution would yield absolute affluence thresholds ranging from 34 000 (90th) to 42 600 (95th) to 70 900 (99th) euro, without correcting for differences in price levels. As shown in Table 3.4.2, 44% of Swiss and 43% of Norwegians would have incomes above the 95th European percentile, while half of those exceeding the threshold would come from France, the United Kingdom, and Germany.

Table 3.4.2: Fraction of population above the 95th percentile (42 529 euro) of the European income distribution, income 2011 (EU-SILC 2012)
(%)

	Switzerland	Norway	Luxembourg	Denmark	France	Austria	Sweden	United Kingdom	Finland	Germany	Netherlands	Cyprus	Iceland	Italy	Spain
% of country population	44	43	30	12	8	8	7	7	7	5	5	4	3	3	1
% of European top 5 %	13	9	1	3	19	3	3	17	1	16	3	0	0	8	2

Reading note: in France, 8% of persons lived in households with equivalent disposable cash income above 42 529 euro (without purchasing power adjustment). These persons comprised 19% of the estimated population of nearly 25 million above the 95th percentile in Europe.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

3.5 Affluence gaps and redistribution-based measures

The headcount measures, transfer-sensitive measures and income shares ignore the levels of income among the group of affluent. Thus, the low overall relative income inequality is behind the low share of the affluent and their incomes in Norway, but this tells nothing about how rich or well-off those in the top are. We next look at measures that reflect the absolute levels of top incomes.

3.5.1 Excess functions

The concept of excess is common both to evaluating the fit of the Pareto distribution and to some affluence measures, and appeared in the convex version of the transfer-sensitive measure in equation 3.4.1 Empirical mean excess function (or mean residual life function) can be defined as the average of excess incomes above a certain threshold:

$$(3.5.1) \quad \frac{\sum_{i=1}^{n} (Y_i - k) I_{[Y_i > k]}}{\sum_{i=1}^{n} I_{[Y_i > k]}}, \; k >= 0$$

where Y_i is income of unit i and k is the high income threshold. One useful property of the empirical mean excess function is that it can be used a graphical tool to check whether the distribution conforms to a Pareto distribution, in which case the empirical mean excess plot has a positive linear trend[20]. The excess can also be called affluence gap, and it is an absolute measure of richness in itself. Consequently, we continue here with a more substantiated analysis of mean excess functions. The nominator can be the total population instead of those exceeding a threshold, resulting in excesses per capita. The numerator can be divided by the threshold resulting in normalized excesses as in the convex richness measure. In the following, we look at non-normalized mean excesses per capita (instead of per affluent).

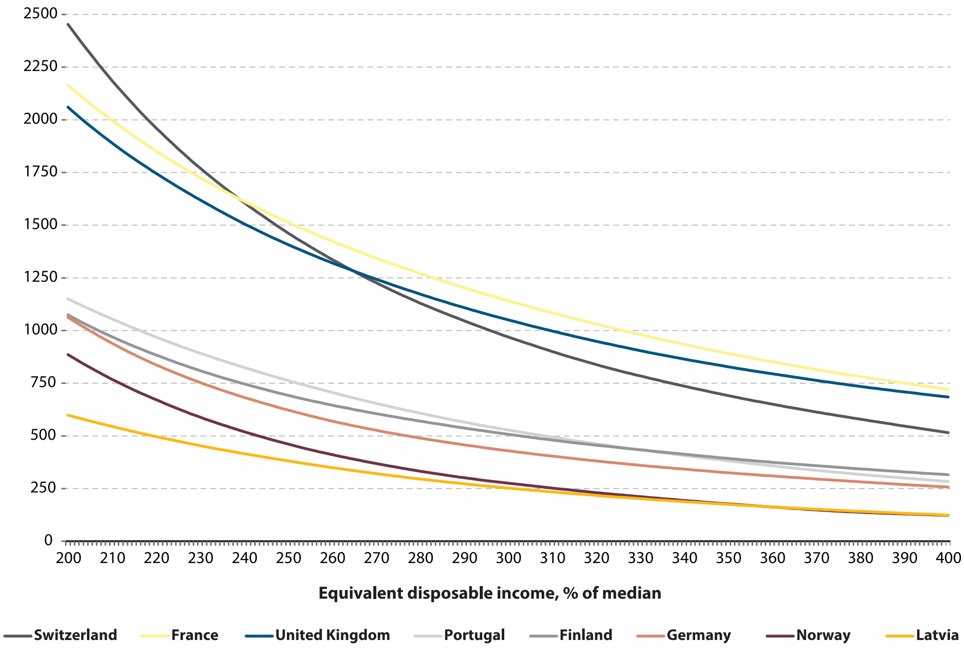
Excess function is also a tool for evaluating second-degree affluence dominance, as discussed by Michelangeli et. al. (2010), Bose et. al. (2014) and Aaberge and Atkinson (2013). Mean excess is the average of differences from the high income threshold, and interpreted as the affluence gap, it is the right tail counterpart of mean poverty risk gap, i.e. the average shortfall from the poverty risk line. Second-order stochastic dominance conforms to computing average distances from a threshold over the whole range of thresholds. In second-order affluence dominance, the evaluation starts from the highest income downwards and evaluates mean excess income for each level of income.

[20] In practise, it turned out not to be very useful in this respect when computed from the EU-SILC datasets.

Figure 3.5.1 plots average excesses for selected countries, with excess divided by total population (as in the FGT poverty measures) rather than the number of person above the threshold (as in the mean excess plot). This implies that average of excess incomes decreases with the income threshold. Whether the nominator is total population or the number of affluent makes a difference and ranks the countries differently. The curves for Switzerland, France and the UK affluence-dominate the other countries, but cross with each other. Throughout the range, the per capita affluence gaps in relatively high-income countries Norway, Finland and Germany are much lower than for instance in the UK.

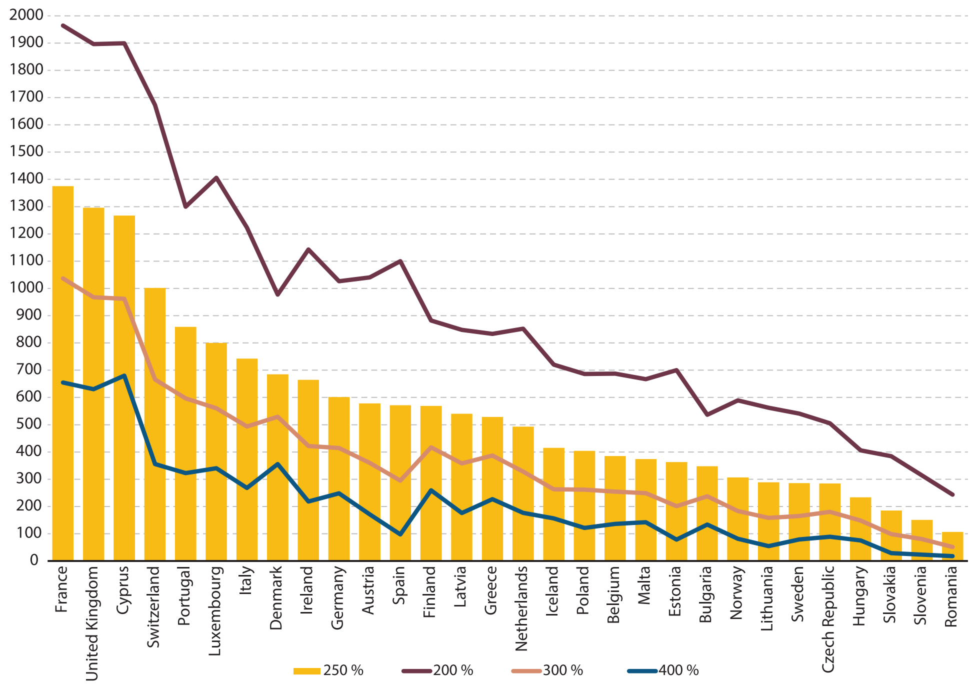
Figure 3.5.2 shows the average per capita affluence gaps for all countries based on five thresholds, ranking the countries by the 2.5 times the median threshold (see appendix for the numbers). Since the average affluence gap is an absolute measure, the values in Figure 3.5.2 are expressed in purchasing power standards, taking into account differences in consumer price levels between the countries. France, the UK and Cyprus have the highest per capita excess income, followed by Switzerland, Portugal, Luxembourg and Italy. Eastern Europe has the lowest per capita excesses, but otherwise there is no clear pattern, not even among the register countries.

Figure 3.5.1: Excess incomes per capita over the range of 200% to 400% of median income in selected countries, 2011 (survey year 2012)
(excess incomes per capita, euro, equivalent DPI)



Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Figure 3.5.2: Excess income per capita (average affluence gaps), selected high income thresholds, income 2011 (survey year 2012)
(excess incomes per capita, PPS, equivalent DPI)



Reading note: In France, the sum of excess incomes above 250% of median divided by total population was 1 375 PPS (purchasing power standards).

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

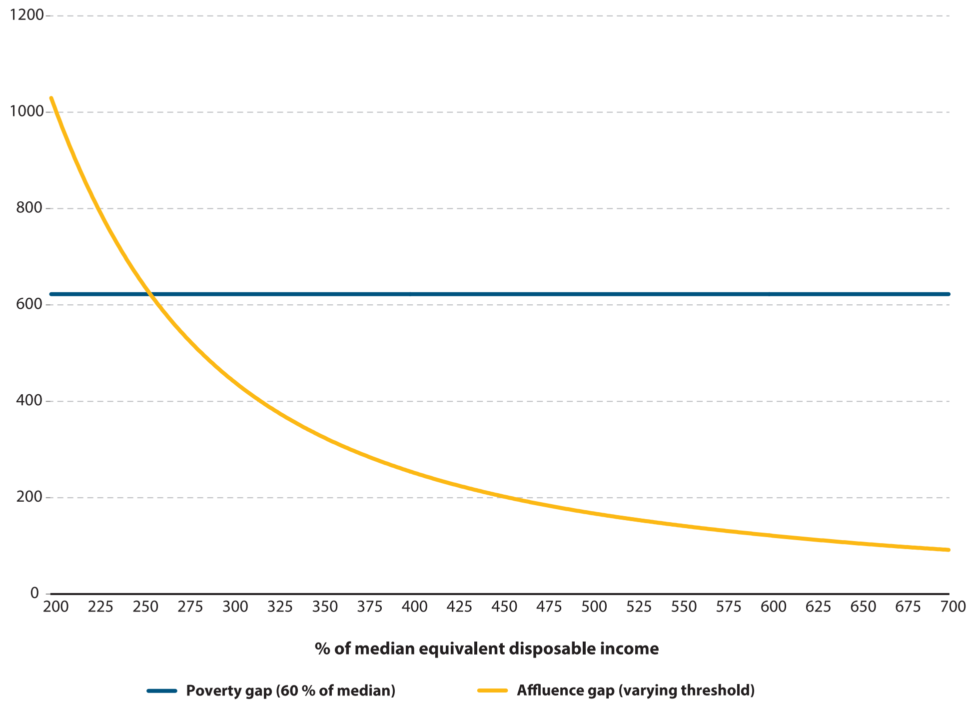
3.5.2 Redistribution-based affluence line

A variant of richness measures are based on the idea of redistribution of funds from the rich to non-rich so that the rich do not fall below the richness line. That is, the redistribution is financed by reducing the excesses above a richness threshold. From a policy perspective, this could take place by raising top marginal tax rates or progressive capital income taxes. At the extreme, all the rich persons could be put to richness line and their excess incomes channelled to some socially beneficial investment (Bose, Chakravartry & Ambrosio, 2014).

Medeiros (2006) proposed that the affluence threshold could be defined as the value where the sum of excesses would be sufficient to get all poor at the poverty risk line, i.e. when the sum of excess incomes (affluence gaps) would be equal to the sum of poverty risk gaps (for an application, see Brezinski, 2010). Hypothetically, if the value of affluence gap is transferred to the bottom of the distribution, poverty risk could be eradicated, ignoring all higher order effects. The Medeiros affluence line depends on the tail distributions and the poverty risk line, but is not sensitive to transfers within the tails.

Figure 3.5.3 uses the pooled EU-SILC data to illustrate how the excess function works when evaluating the redistributive affluence line, using the method of Medeiros. The Medeiros affluence line would set conveniently at 256% of median when determined from the supranational estimate of European distribution of equivalent disposable income. In other words, transferring the excess incomes from those having more than 256% of median income (39 000 euro) to those below 60% of median income (9 150 euro) would - purely technically of course and ignoring all higher order effects - eradicate relative income poverty risk from the supranational entity of Europe.

Figure 3.5.3: Affluence line determined from affluence and poverty gaps- Supranational EU -estimate 2011 (survey year 2012)
(billion euro, equivalent DPI)



Reading note: Poverty line is fixed at 60% of median. Sum of poverty gaps (in billion euros) is one fixed number. Excess over incomes above the threshold are shown for all income levels from 200 to 700% of median. The Medeiros affluence line is where the lines cross (256%), which is when the sum of affluence gaps is equal to the sum of poverty gaps.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

As a further example, in Finland 2012 the headcount poverty stood at 13.2% or 704 000 persons, and the poverty line was 13 619 euro. Mean poverty gap was 20.2%, and the average of absolute gaps therefore 2 750 euro and their sum close to two billion euro. Counting from the top, the sum of excess income which would yield two billion euro but leave all people above the threshold at the threshold value would be reached at 367% of median income. The affluence line would be 83 305 euro, with 0.7% of persons above this threshold[21].

At country level, this method would yield quite different affluence lines, as reported in Table 3.5.1 below. Let us note however, that in all countries the lines are above 200% of median, with minimum of 206 in Sweden. The affluence lines reflect the thickness of the tails, and are above 600% in Cyprus and France. The Medeiros affluence line for France stands at 607% of median (125 000 euro) while for the UK it was much lower at 450% (85 500 euro), again indicating that the uppermost tail in France is heavier.

The redistribution-based lines are dependent on the policy goals and the levels and distributions in both tails, and can be sensitive to data imperfections. They do not obey certain axioms of richness indexes, in particular the focus axiom, i.e .that a richness index should be independent of the incomes of the non-rich (Peichl et. al., 2010, p.601). For this reason, it is not likely to be suitable for analysis of richness in cross-country settings.

[21] The average income in this group was 135 500 euro, which is 1.67 times the threshold. The 1.67 ratio to threshold of the average income above the threshold is in fact empirical inverted Pareto-Lorenz coefficient. The average of excesses was 52 198 euro among the 37 026 affluent, and the sum of excesses by definition two billion.

Table 3.5.1: Redistribution-based affluence lines of Medeiros, % of median, 60% poverty line, income 2011 (EU-SILC 2012)
(%)

Cyprus	637	**Luxembourg**	303	**Malta**	278	**Belgium**	238
France	607	**Switzerland**	302	**Estonia**	276	**Romania**	216
United Kingdom	450	**Netherlands**	301	**Lithuania**	272	**Norway**	211
Portugal	417	**Ireland**	301	**Italy**	272	**Slovenia**	211
Finland	367	**Hungary**	298	**Austria**	261	**Slovakia**	211
Latvia	352	**Bulgaria**	290	**Denmark**	252	**Sweden**	205
Czech Republic	331	**Iceland**	285	**Greece**	251		
Poland	314	**Germany**	283	**Spain**	238		

Reading note: In Ireland, all those below 60% of median would have income of 60% of median if incomes in excess of 301% of median would be transferred to them, leaving those above 301% of median with that income.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

4. High incomes and other dimensions of affluence

The measures of richness based on income do not take into account all available resources (wealth, in particular), and mostly neglect the differences in living standards among the European countries. The relative distance from a national median, even if very high, does not necessarily guarantee high economic well-being. We next complement the evidence on high incomes with other dimensions of affluence that could be possibly identified from EU-SILC data.

While EU-SILC does not contain much information on wealth, it does have a wealth of information on non-monetary shortfalls and subjective economic well-being. Unfortunately, there is not much that can help in distinguishing the very well-off. Nevertheless, two questions are quite useful: the first is whether the household has difficulties in making ends meet and the second whether the household can finance an unexpected expense without borrowing or other help. We combine these into one measure by restricting to households who can make ends meet fairly easily, easily or very easily (HS120) and have capacity to finance unexpected expenses of 1/12th of the annual poverty line from own resources (HS060). The combination of the variables leaves out usually less than 5% of those who make ends meet very easily.

This dichotomous variable aims to combine household's perception of adequacy of income in relation to consumption as well as emergency-funding type of wealth. If a household finds it easy to pay for its usual necessary expenses, the ratio of its income to its necessary expenses should be quite high. This may follow from its resources or consumption preferences. All three categories of "easy" are included (instead of just "very easily"), since the responses may reflect some personal and cultural differences[22]. Regarding the second condition, if a household can afford an unexpected required expense without borrowing or asking for help, this indicates having some buffer savings in the form of liquid financial wealth (Morrone et. al., 2011; Törmälehto et. al., 2013).

Figure 4.1 shows the proportion of people who make ends meet easily and have capacity to finance unexpected expenses, of those who are above three high income thresholds(23). Nearly all of those above the 200% income threshold make ends meet easily in Sweden, and more than 85% all the Nordic countries, Luxembourg, the Netherlands, Switzerland, Austria, and the UK. In heavy-tailed France, the ratio is a bit lower, and in Eastern and Southern Europe much lower. In particular, only around one third of the relatively high income households make ends meet easily in Lithuania, Greece, Bulgaria, and Latvia. Differences of this magnitude cannot be explained by the inconsistent reference times of EU-SILC; income data are from previous year while the subjective variables are from the time of the interviews.

In most countries, the share of subjective material well-being increases with higher income thresholds. The differences among the countries reduce, and the figure brings some support to having 250 or 300% as the high income threshold instead of 200% of median. It seems evident that using only income to identify the affluent is far from satisfactory; for instance, the comparatively large 17% share of total income going to the top 5% in Latvia does not translate to uniformly high subjective economic well-being of this group.

Given the large disparities in average living standards in Europe, affluence may be better measured with absolute measures or multidimensional affluence indicators. Our next step is to construct a multidimensional affluence indicator, which combines non-income indicators with high income indicator. The aggregation is based on multidimensional counting, following a dual cut-off approach developed for multidimensional poverty measurement (Alkire and Foster 2011). Peich and Pestel (2011) used the same approach to create an affluence indicator combining income and wealth. The method is an extension of the Foster-Greer-Thorbecker poverty index to multidimensional non-income space.

The first step in this method is to define the dimensions and the dimension-specific cut-offs. For income, we use the cut-off discussed earlier, with affluence line set at 250% of national median income. With non-income dimensions, we combine absence of financial hardship into one variable, and thus effectively just exclude those with financial difficulties from the high income group. To this end, not being in arrears was added to the making ends meet/capacity variable discussed above. In another variant, income and financial non-deprivation variables were kept as separate dimensions, and capacity to afford one week holiday was included as well. This method did not prove very useful, and the results are only reported in the annex.

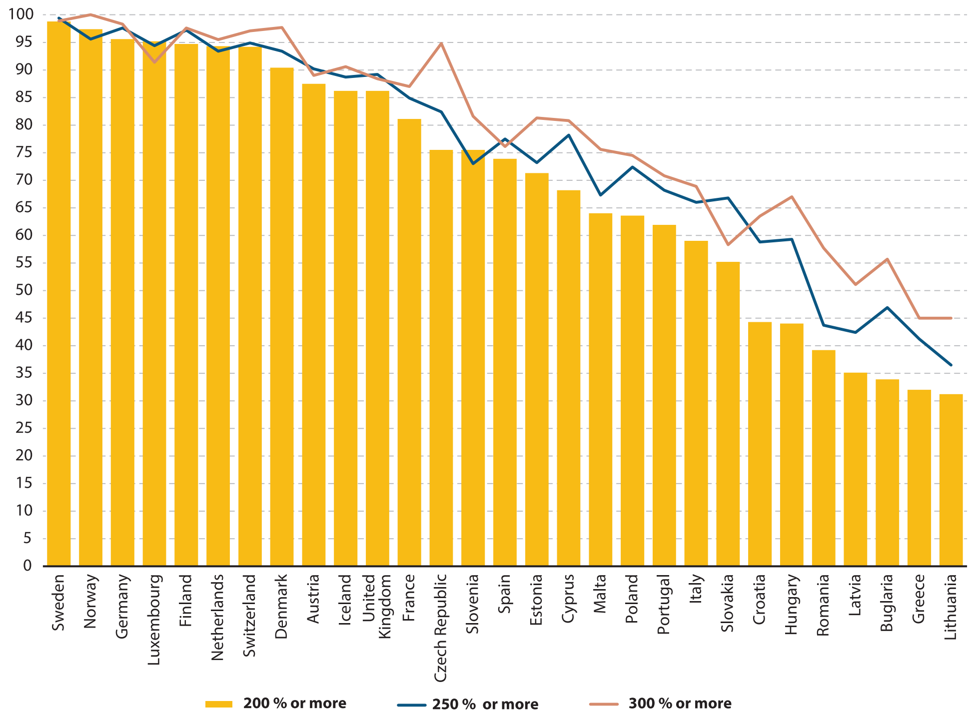
The second step is to determine in how many dimensions (k) the person is required to be affluent (or have positive achievements). In a union approach, the cut-off is set at k=1, with for instance high incomes being a sufficient condition for being affluent. In

(22) This is based on author's notions with the Finnish data: even very wealthy households may respond making ends meet just fairly easily and not very easily, but they rarely report having difficulties.

(23) Repeating figure 4.1 for top 5 and top 1% does not bring much additional insight.

another extreme, all conditions must be met to be determined affluent. The results reported here collapse to simple dual-condition indicator of high income and financial non-deprivation, since we set the k=2.

Figure 4.1: Share of persons who make ends meet easily and have capacity to face unexpected expenses in households above 200/250/300% of median, income year 2011 /survey year 2012
(% of persons)



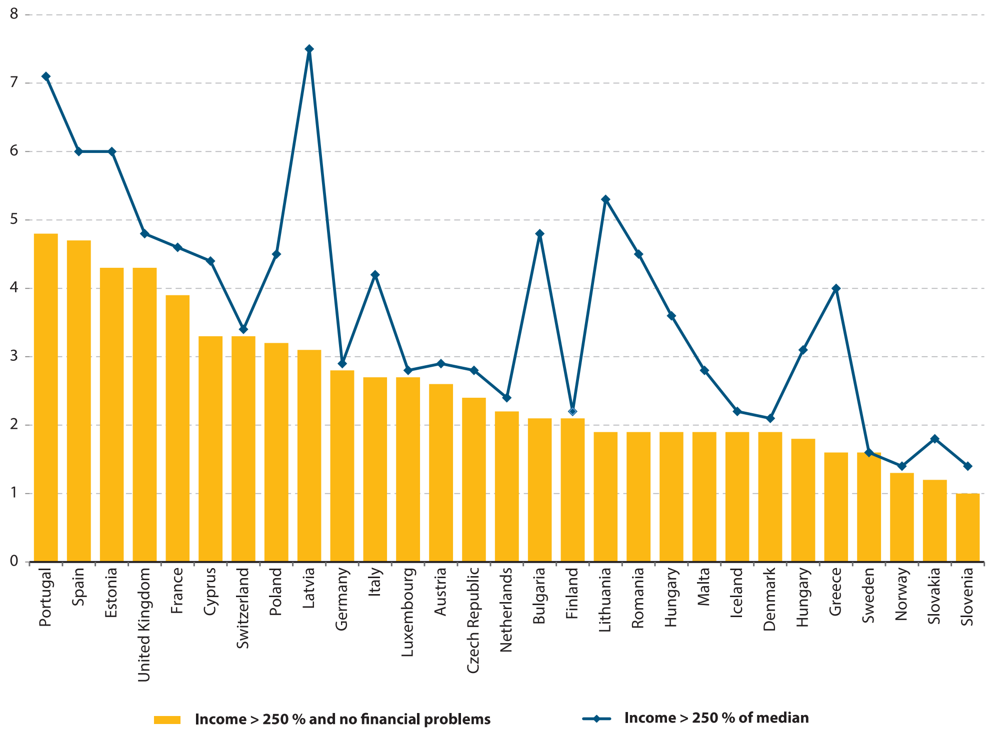
Reading note: In Sweden, 99% of persons above 200% of median lived in households making ends meet easily and having capacity to finance unexpected expenses without borrowing or asking for help. In Lithuania, the corresponding share was 35%.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version April 2014).

Figure 4.2 shows the proportion of persons who are both above the 250% threshold and do not report financial deprivation in the sense that they make ends meet easily, have capacity to finance unexpected expenses, and are not in arrears. This definition of affluence reduces the rates significantly, in particular in Latvia, Lithuania, Romania, and Greece. The ranking of the countries also changes somewhat, but Portugal, Spain and Estonia remain countries with high relative affluence.

The method would allow adjusting the headcounts with the proportion of positive achievements of the affluent, but this is not relevant with the two dimensions and k = 2. An index with more variables was also tested, and the results for an eight-dimensional index are shown in the annex. The dimensions could be weighted differentially as well. In addition to poverty measurement, multidimensional counting could be a useful tool for creating affluence indexes. Development of such an index would require considering having EU-SILC variables which could distinguish better the preferences and habits of the upper classes. This could be considered in the contents of the likely every six-year module on wealth and consumption in the revised EU-SILC. Just to give some examples, ownership and values of asset types such as cars, boats and yachts, and valuables could potentially be useful in affluence measurement.

Figure 4.2: Share of persons with income above 250% of median and not financially deprived, income 2011 (EU-SILC 2012)
(% of persons)



Reading note: In Latvia, 7.5% of persons lived in households with income above 2.5 times the national median but only 3.1% if requirement of making ends meet easily, having capacity to face unexpected expenses and not being in arrears is added.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

5. Summary and conclusions

Although sample surveys are not generally regarded as good sources on top incomes, EU-SILC is potentially a useful data source as long as its weaknesses and strengths are recognised. The main drawback is the low number of observations in the top end in many countries, which limits the analysis considerably. In contrast, the top tail is not seriously affected by top-coding or extreme outliers. The representation of the sample, possible unit non-response bias, and likely under-estimation are difficult to examine. This paper addressed the data quality indirectly, by simulating measurement errors via the same tool as in outlier detection, i.e. semi-parametric modelling (Pareto-fitting). It allows for instance sensitivity analysis of more serious under-estimation in the survey countries. Tentatively, it seems that the estimates of top incomes seem not to be overly sensitive to reasonable assumptions about under-estimation of top incomes.

The split to survey and register countries is important, but the differences between the register countries deserve attention as well. The impact on top incomes in France when changing to register data was very significant and shows up in a sizable increase in property incomes. While labour income still dominated the incomes of the top 5% in all countries, the share of property income in France was much higher than in other countries. The analysis of the change to register data in France and its effect on comparability should be further examined. We also find differences among the Nordic register countries. The possible impact of different calibration models should be examined; for instance, Denmark and Finland seem to use much more income data in their re-weighting schemes and seem to have fatter tail than the other Nordics.

Regarding the results, relative measures such as headcount shares and top income shares are highly correlated with overall inequality measures. A somewhat different country rankings emerge with transfer-sensitive measures and average affluence gaps, which take into account the distribution and/or the absolute levels of income in the top tail. For instance, the share of persons above 2.5 times the median is highest in Latvia, Portugal and Spain while with average affluence gaps France, Switzerland and the UK would be ranked to be the most affluent However, combining affluence headcounts with absence of financial deprivation suggest that identifying the affluent only on the basis of income is far from satisfactory.

While the paper followed the dominance approach, in many cases the threshold need to be fixed. Given the sample size restrictions, the upper limit for affluence threshold should not be higher than 2.5 times the median if all countries are to be analysed. This could be the recommended default choice of high income threshold instead of the conventional two or three times the median. The Pareto-fitting and non-income information also give some support to this choice.

Technically, semi-parametric modelling seems a potentially useful tool, both for sensitivity analyses and for outlier detection and adjustments. More work remains to be done on the robust estimation of Pareto coefficient in small samples with complex sampling designs. Further work in general is needed on this type of analysis, as well as on measuring the size of actual under-estimation in the top tail using external data. Comparisons to national external benchmarks and other case studies, focusing e.g. on register countries only, would be one way to partially validate the EU-SILC results.

Regarding outliers, the data validation routines of the National Statistical Institutes may have improved over the years so that very extreme values or highly erroneous or implausible values are not present. Therefore, the recommendation is to take the data as they are, and only in very dubious cases consider altering the data. If this is necessary, semi-parametric modelling is a method that is quite easy to implement and preserves the distributional information of the data. Simple measures such as top-coding to a threshold should be avoided.

The measurement of "richness" based on EU-SILC is best seen as related to measurement of relative income poverty and non-monetary deprivation as well as studies of middle-income households and middle class. The value added that EU-SILC can bring to top incomes debate is more on the size and composition of the economically very well-off group rather than concentration of income to the very rich. Sample surveys do have difficulties in capturing top incomes, and sampling errors of population shares of the affluent are likely to be less worrying than sampling errors of top income shares. In terms of the choice between the "income space" and the "people" space, the former would be where EU-SILC-based richness measures have more to add. In other words, the focus could be put on the distances of the "mildly" affluent from the average individual, the size of the affluent group, and their living conditions.

6. References

Aaberge, R., Atkinson, A.B. (2013), "The Median as Watershed", *Discussion Paper* n°749, Statistics Norway.

Alfons, A., Templ, M. (2013), "Estimation of social exclusion indicators from complex surveys: The R package Laeken", *Journal of Statistical Software*, 54(15): 1-25.

Alkemade, P., Endeweld, M (2014), "LIS Micro-Data and National Accounts Macro-Data Comparison: Findings from wave I - wave VIII", *Technical working papers series* n°7, Luxembourg Income Study.

Alkire, S., Foster, J. E. (2011), "Counting and Multidimensional Poverty Measurement", *Journal of Public Economics*, 95(7-8): 476-487.

Armour, P., Burkhauser, R.V. and Larrimore, J. (2014), *Using the Pareto Distribution to Improve Estimates of Topcoded Earnings*, US Census Bureau Center for Economic Studies Paper n°CES-WP-14-21.

Atkinson, A. B., Piketty, T. (2010), *Top Incomes - a Global Perspective*, Oxford University Press.

Atkinson, A. B., Piketty, T. (2007),*Top Incomes over the Twentieth Century*, Oxford University Press.

Atkinson, A.B., Brandolini, A. (2013), "On the Identification of the Middle Class", in J. Gornick and M. Jäntti (eds.) *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*, Stanford University Press: 77-99.

Atkinson, A.B., Piketty, T., Saez, E. (2011), "Top incomes in the long run of history", *Journal of Economic Literature*, 49(1): 3-71.

Bose, A., Chakravarty, S. R. and D'Ambrosio, C. (2014), "Richness Orderings", *The Journal of Economic Inequality*, March 2014, 12(1): 5-22.

Brezinski, M. (2014), "Do wealth distributions follow power laws", *Physica A: Statistical Mechanics and its Applications*, 406: 155-162.

Brezinski, M. (2013), "Robust Estimation of the Pareto Index: A Monte Carlo Analysis", University of Warsaw, Faculty of Economic Sciences, *Working Papers* n°32/2013(117).

Brezinski, M. (2010), "Income affluence in Poland", *Social Indicators Research*, 99(2): 285-299.

Cingano, F. (2014), "Trends in Income Inequality and its Impact on Economic Growth", *OECD Social, Employment and Migration Working Papers* n°163. Available at: http://dx.doi.org/10.1787/5jxrjncwxv6j-en [Accessed: 20 April 2016]

Clauset, A., Shalizi, C.R. and Newman, M.E.J. (2009), "Power-Law Distributions in Empirical Data", *SIAM Review*, 51(4): 661-703.

Cowell, F. A., Victoria-Freser, M-P. (2007), "Modelling Lorenz Curves: robust and semi-parametric issues", London School of Economics and Université de Genève, *DARP* n°91.

Eriksson, K.H. (2011), "Censoring and Top-Coding in LIS Data", *LIS Technical Working Paper Series* n°4.

Goedemé, T. (2013), "How much confidence can we have in EU-SILC? Complex Sample Designs and the Standard Error of the Europe 2020 Poverty Indicators", *Social Indicators Research*, 110(1): 89-110.

Jäntti, M., Riihelä, M., Sullström, R. and Tuomala, M. (2010), "Trends in top Income shares in Finland", in A. B. Atkinson and T. Piketty (editors) *Top Incomes: A Global Perspective*, Oxford University Press: 372-447.

Kennickell, A. (2007), "The role of over-sampling of the wealthy in the Survey of Consumer Finances", Federal Reserve Board working papers, Washington.

Lakner, C., Milanovic, B. (2013), "Global Income Distribution: from the Fall of Berlin Wall to the Great Recession", World Bank Policy Research Working Paper n°6719.

Leigh, Andrew (2009), "Top Incomes", in W. Salverda, B. Nolan and T.M. Smeeding (editors) The Oxford Handbook of Economic Inequality, Oxford University Press.

Mattonetti, M.L. (2013), *European household income by groups of households*, Eurostat Methodologies and Working Papers, Publications Office of the European Union, Luxembourg.

Medeiros, M. (2006), "The rich and the poor: the Construction of an Affluence Line from the Poverty Line", *Social Indicators Research*, 78(1): 1-18.

Medeiros, M., Ferreira de Souza, Pedro H. G. (2014), "The rich, the affluent and the top Incomes: a Literature Review", *IRLE Working Paper* n°105-14, IRLE – Institute for Research on Labor and Employment, Berkeley.

Michelangeli, A., Peluso, E. and Trannoy, A. (2010), "Detecting a Change in Wealth Concentration Without the Knowledge of the Wealth Distribution", *The Journal of Economic Inequality*, 9(3): 373-391.

Morrone, A., Scrivens, K., Smith, C. and Balestra, C. (2011), *Measuring vulnerability and resilience in OECD countries*, Paper presented at the 2011 IARIW-OECD Conference on Economic Insecurity Measurement: Measurement, Causes, and Policy Implications, November 22-23, 2011, Paris.

Neri, L., Gagliardi, F., Ciampalini, G.,Verma, V and Betti, G. (2009), "Outliers at upper end of income distribution (EU-SILC 2007)", *Working Paper* n°86, Department of Economics and Statistics, Universita di Siena.

OECD (2014), "Focus on Top Incomes and taxation in OECD Countries: Was the crisis a game changer?", *Focus,* May 2014, OECD.

Peichl, A., Schaefer, T. and Scheicher, C. (2010), "Measuring richness and poverty: a micro data application to Europe and Germany", *Review of Income and Wealth*, 56(3): 597-619.

Peichl, A., Pestel, N. (2011), "Multidimensional Affluence: Theory and Applications to Germany and the US", *IZA Discussion Paper* n°5926.

Piketty, T. (2014), *Capital in the twenty-first century,* Harvard University Press, Cambridge.

Quitzau, J. (2013), "Om usikkerhed i EU-SILCs målinger af indkomstfordelingen", *Memo* 9, December 2013, Danmarks Statistik.

Rao, J.N.K., Wu, C.F.J. and Yue, K. (1992), "Some recent work on resampling methods for complex surveys», *Survey Methodology*, 18: 209-217.

Statistics Finland, *Total statistics on income distribution*, Available at: http://www.stat.fi/til/tjkt/index_en.html [Accessed: 20 April 2016]

Törmälehto, V-M (2006), *Measurement of property income in EU-SILC*. Proceedings of the EU-SILC Conference, Helsinki, 6-8 November 2006. Eurostat, Methodologies and Working Papers 2007 edition.

Törmälehto, V-M, Kannas, O. and Säylä, M. (2013), "Integrated Measurement of Household-level Income, Wealth and Non-monetary Well-being in Finland", *Statistics Finland Working Papers* n°1/2013.

Van Kerm, P. (2007), *Extreme incomes and the estimation of poverty and inequality indicators from EU-SILC. Comparative EU Statistics on Income and Living Conditions: Issues and Challenges.* Proceedings of the EU-SILC-Conference, Helsinki, 6-8 November 2006.

Vermeulen, P. (2014), "How fat is the top tail of the wealth Distribution?", *ECB Working Paper Series* n°1692, European Central Bank, Frankfurt am Main.

Table 1: Number of households in the sample and affluence headcount ratios; thresholds 200% to 500% of median equivalent disposable income.EU-SILC 2012 (income year 2011)
(number of households)

Sample size (households)					Affluence headcount ratio				
	200%	250%	300%	500%		200%	250%	300%	500%
Latvia (LV)	679	382	219	39	LV	13	7.5	4.4	0.7
Portugal (PT)	786	457	286	59	PT	12.4	7	4.2	0.9
Spain (ES)	1546	786	412	40	ES	11.9	6	3.1	0.3
Estonia (EE)	461	218	107	9	EE	11.8	6	2.8	0.3
Romania (RO)	747	303	146	10	RO	11.6	4.5	2	0.1
Lithuania (LT)	603	308	168	16	LT	11.5	5.3	2.6	0.2
United Kingdom (UK)	871	419	229	65	UK	9.8	4.9	2.9	0.9
Cyprus (CY)	468	227	124	38	CY	9.7	4.5	2.3	0.7
Poland (PL)	1140	558	311	48	PL	9.7	4.5	2.6	0.4
Bulgaria (BG)	410	223	137	30	BG	9.3	4.9	3	0.7
Ireland (IE)	419	188	100	20	IE	9.2	4.1	2.1	0.3
Greece (EL)	353	160	86	27	EL	8.6	4	2.2	0.9
France (FR)	1181	638	388	93	FR	8.6	4.6	2.9	0.7
Italy (IT)	2144	1131	677	147	IT	8.5	4.2	2.4	0.4
Switzerland (CH)	617	289	161	36	CH	7.5	3.5	1.9	0.5
Austria (AT)	480	219	111	19	AT	7	3	1.5	0.2
Germany (DE)	934	402	191	39	DE	7	2.9	1.4	0.3
Malta (MT)	264	119	54	11	MT	6.6	2.6	1.1	0.3
Hungary (HU)	636	291	142	22	HU	6.4	3.1	1.5	0.3
Czech Republic (CZ)	428	197	94	18	CZ	6	2.9	1.4	0.3
Netherlands (NL)	773	320	164	35	NL	5.7	2.4	1.2	0.3
Slovakia (SK)	282	98	36	3	SK	5.6	1.9	0.6	0.1
Belgium (BE)	308	125	72	13	BE	5.3	1.9	1.2	0.2
Iceland (IS)	174	76	40	6	IS	5.1	2.2	1.2	0.2
Finland (FI)	836	394	214	57	FI	5.1	2.3	1.3	0.3
Denmark (DK)	354	145	76	15	DK	4.7	2.1	1.4	0.4
Sweden (SE)	299	120	60	6	SE	3.8	1.6	0.9	0.1
Slovenia (SI)	406	160	89	6	SI	3.7	1.4	0.7	0.1
Norway (NO)	221	86	38	7	NO	3.3	1.4	0.6	0.1

Note: Sample size is the un-weighted number of households above the threshold. Headcount is the weighted % of persons above the threshold. Countries are sorted by headcount ratio at 200% threshold.

Note: Orange cells are cells with less than 50 observations

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Table 2: Estimated Pareto coefficients and number of outliers; thresholds 200 to 500% of median income. EU-SILC 2012 (income year 2011)

Estimated Pareto shape parameter					Outliers				
	200%	250%	300%	500%		200%	250%	300%	500%
LV	2.8	3.1	3.3	3.6	LV	4	3	2	0
PT	2.6	2.8	2.8	2.7	PT	6	4	2	0
ES	3.4	3.9	4.4	5.1	ES	6	3	3	0
EE	3.5	4.0	3.8	2.5	EE	2	2	1	0
RO	4.2	4.7	5.1	3.5	RO	6	4	4	0
LT	3.5	3.8	4.3	3.7	LT	5	3	3	0
UK	2.9	2.7	2.6	2.4	UK	14	9	3	0
CY	3.1	2.9	2.5	1.8	CY	13	6	3	0
PL	3.1	3.2	3.3	3.5	PL	7	4	2	1
BG	2.8	2.9	3.0	2.9	BG	6	3	2	0
IE	3.5	3.4	3.2	2.6	IE	10	4	3	0
EL	3.1	3.0	2.5	4.3	EL	3	2	0	0
FR	2.7	2.6	2.6	2.8	FR	13	8	7	3
IT	3.1	3.1	3.3	3.2	IT	26	17	12	2
CH	3.3	3.1	3.0	3.7	CH	4	2	0	0
AT	3.6	3.4	3.2	2.5	AT	6	1	1	0
DE	3.6	3.4	2.9	2.8	DE	13	9	3	0
MT	4.0	4.0	3.0	2.5	MT	7	6	0	0
HU	3.4	3.6	3.5	3.3	HU	5	2	1	0
CZ	3.3	3.4	3.1	3.8	CZ	3	2	1	0
NL	3.5	3.2	2.9	3.4	NL	9	1	1	1
SK	4.6	5.1	4.6	5.8	SK	2	1	0	0
BE	3.5	2.9	3.0	2.1	BE	7	2	2	0
IS	3.6	3.6	3.5	2.6	IS	3	2	1	0
FI	3.5	3.1	2.8	2.7	FI	20	5	3	1
DK	3.3	2.7	2.6	1.9	DK	5	1	0	0
SE	4.0	3.9	4.2	2.2	SE	5	2	1	0
SI	4.0	3.7	4.3	7.6	SI	2	0	0	0
NO	4.1	3.9	3.6	5.4	NO	4	1	0	0

Note: ML-estimate of the shape parameter of Pareto distribution. Outliers is the number of households who deviate from the hypothesized Pareto-model. Countries are sorted by headcount ratio at 200% threshold.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Table 3: 95th and 99th quantiles of person-weighted distribution of household equivalent disposable income in 2011 (EU-SILC 2012)
(% of median & euro)

	% of median			Euro	
	p95	p99		p95e	p99e
PT	286	480	PT	23 793	39 947
LV	286	464	LV	12 679	20 544
ES	266	383	ES	31 807	45 825
EE	262	391	EE	15 686	23 422
LT	257	380	LT	11 153	16 471
UK	247	450	UK	47 021	85 433
BG	246	411	BG	7 048	11 789
RO	245	332	RO	5 165	6 993
PL	244	402	PL	12 326	20 365
FR	242	455	FR	49 852	93 793
CY	241	417	CY	40 862	70 648
IT	238	381	IT	38 117	61 114
EL	234	442	EL	22 250	42 034
IE	233	370	IE	44 515	70 511
CH	223	360	CH	87 773	141 730
HU	217	347	HU	10 319	16 505
DE	217	330	DE	42 578	64 579
AT	215	344	AT	46 965	74 998
CZ	212	331	CZ	16 548	25 788
MT	210	317	MT	24 087	36 310
NL	207	319	NL	42 593	65 694
SK	206	274	SK	14 252	19 008
BE	202	312	BE	40 533	62 589
FI	201	321	FI	45 564	72 971
IS	201	306	IS	38 994	59 253
DK	197	346	DK	52 360	92 087
SE	189	288	SE	46 663	71 173
SI	189	273	SI	22 846	33 057
NO	182	269	NO	73 048	107 805

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Table 4: **Transfers-sensitive richness measures (Peichl, Schaefer and Scheicher, 2008), income 2011 (EU-SILC 2012), , affluence line 2.5 x median**

	Threshold	Headcount	Concave		Convex	
	Euro	% of persons	Cha(0,3)	Cha(3)	FGTT2(1)	FGTT2(2)
LV	11 102	7.5	0.7	3.8	3.4	4.8
PT	20 809	7.0	0.7	3.6	3.6	5.6
ES	30 056	5.9	0.4	2.6	1.8	1.3
EE	14 988	5.9	0.4	2.4	1.8	1.5
LT	10 863	5.2	0.4	2.3	1.7	1.3
BG	7 166	4.9	0.4	2.5	2.4	3.7
UK	47 738	4.8	0.5	2.5	2.9	6.4
FR	51 517	4.6	0.5	2.5	2.9	13.6
PL	12 649	4.5	0.4	2.2	1.9	2.6
RO	5 267	4.5	0.3	1.7	1.2	0.9
CY	42 318	4.4	0.4	2.1	2.7	7.4
IT	40 094	4.2	0.4	2.0	1.9	3.6
IE	47 694	4.1	0.3	1.9	1.6	2.4
IE	47 694	4.1	0.3	1.9	1.6	2.4
EL	23 926	3.9	0.4	2.0	2.1	2.8
HR	13 511	3.6	0.2	1.4	0.9	0.5
CH	98 647	3.4	0.3	1.6	1.5	1.8
HU	11 886	3.1	0.2	1.4	1.2	1.5
AT	54 518	2.9	0.2	1.3	1.1	1.8
DE	49 025	2.9	0.2	1.3	1.3	2.3
LU	82 077	2.8	0.2	1.2	1.2	1.7
CZ	19 478	2.8	0.2	1.3	1.1	1.2
MT	28 656	2.6	0.2	1.1	1.0	1.5
NL	51 512	2.3	0.2	1.1	1.0	1.5
FI	56 754	2.2	0.2	1.1	1.2	4.1
IS	48 445	2.1	0.2	1.0	0.9	1.3
DK	67 042	2.0	0.2	1.1	1.4	3.1
BE	50 281	1.9	0.2	0.9	0.8	1.3
BE	50 281	1.9	0.2	0.9	0.8	1.3
SK	17 318	1.8	0.1	0.6	0.4	0.2
SE	61 849	1.6	0.1	0.7	0.6	0.8
NO	100 126	1.4	0.1	0.6	0.5	0.4
SI	30 265	1.4	0.1	0.6	0.4	0.3

Note: Values in %, except the threshold , which is equivalent disposable income in euro, person weighted. Countries sorted by the headcount rate.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Table 5: **Average excess income per capita, selected high income thresholds as % of median. EU-SILC 2012 (income year 2011)**
(PPS)

	Threshold, % of median equivalent DPI						
	200%	**225%**	**250%**	**275%**	**300%**	**350%**	**400%**
FR	1.964	1.623	*1.375*	1.186	1.037	810	655
UK	1.896	1.548	*1.296*	1.112	967	762	630
CY	1.899	1.520	*1.267*	1.088	962	797	680
CH	1.673	1.273	*1.002*	807	666	476	356
PT	1.300	1.048	*859*	711	597	432	323
LU	1.406	1.034	*800*	655	561	433	341
IT	1.223	941	*742*	601	494	353	268
DK	977	800	*685*	599	529	425	356
IE	1.143	858	*665*	519	422	291	218
DE	1.026	766	*601*	490	415	314	248
AT	1.041	762	*578*	454	360	239	171
ES	1.100	791	*571*	410	295	164	97
FI	882	694	*569*	483	417	323	260
LV	848	673	*540*	436	358	248	176
EL	833	652	*529*	446	387	292	227
NL	852	632	*493*	399	328	235	177
IS	720	537	*415*	328	263	196	157
PL	686	518	*404*	324	262	176	122
BE	688	498	*385*	311	254	179	136
MT	667	487	*374*	297	249	181	142
EE	700	504	*363*	267	201	123	78
BG	536	427	*348*	286	237	173	134
NO	589	418	*307*	233	183	118	82
LT	562	398	*289*	211	158	92	55
SE	541	386	*286*	217	165	106	79
CZ	505	375	*284*	223	180	125	89
HR	406	304	*234*	184	149	104	76
SK	384	262	*185*	134	99	53	29
SI	314	214	*151*	110	81	43	24
RO	244	161	*107*	74	52	28	18

Note: Countries are sorted by column 250 % of median. Values in PPS (purchasing power standards).

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

Table 6: Experimental multidimensional affluence indicator, 8 dimensions. EU-SILC 2012 (income year 2011)

	Share of persons in each dimension (%)								Share of persons by the number of required dimensions (k)		
	2.5 x median	Very easy	Capacity	Burden1	Burden2	noarrerars	holiday	Wealth proxy	4	5	6
LV	7.5	0.3	26.3	11.8	4.2	75.6	36.9	1.2	10.9	3.4	1
PT	7.1	0.7	64.1	11.7	5.1	90.3	43.9	5.9	16.7	4.8	1
EE	6	1.4	55.2	17.6	8.1	86.9	51.5	1.7	18.3	5.3	1.2
ES	6	0.8	57.9	3	0.7	89.1	53.4	5.3	10.4	1.7	0.3
LT	5.3	0	39.6	8.6	1.2	86.8	50.1	2.8	10.7	2.2	0.3
UK	4.8	7.4	56.3	25.5	15.8	88.9	60.9	6.3	29.5	13.1	4.5
BG	4.8	0.3	31.4	4.5	0.5	68.4	26	4.1	5.6	1.4	0.4
FR	4.6	1.9	66.7	48.2	21.5	89.8	71.9	20.9	49.2	23.3	5.8
PL	4.5	0.9	45.8	5	3.6	84.6	37.8	2	8.6	2.1	0.4
RO	4.5	0.4	46.9	3.8	1.1	68.2	25.1	0.8	5	0.9	0.2
CY	4.4	1.3	49.5	7.7	1.5	68.5	47.7	8.8	11.2	2.9	0.8
IT	4.2	0.4	57.5	1	0.4	86.5	49.2	11.1	10.8	1.3	0.1
EL	4	0.4	59.6	4.2	1.8	61.1	47.5	11.7	11.5	2.9	0.5
HR	3.6	0.3	32.7	3	0.7	70.3	28.9	3.2	5.9	1.2	0.1
CH	3.4	13.2	82.3	13.7	1.7	95.1	92.4	8.3	27.9	8.4	1.9
HU	3.1	0.3	25.7	6.9	2.3	73.6	33	2.1	6.8	2.1	0.5
AT	2.9	10.7	77.8	29.4	4.3	93.5	79.6	5.7	33.6	11.7	2.4
DE	2.9	8.2	66.2	21.8	4.8	95.1	77.8	8.3	27.1	9.2	2.5
LU	2.8	9.5	75.1	13.5	8.3	95.8	85.1	9.4	26.4	10	3.4
CZ	2.8	0.9	57.6	7.4	1.1	94	56.5	2.4	9.9	2.3	0.3
MT	2.8	0.7	75.5	6.6	1.4	89.9	46.8	9.5	12.2	2	0.3
NL	2.4	16.3	77.3	43	6.3	95	82.7	6.4	46.1	17.2	3.2
FI	2.2	16.8	71.3	23.7	19.8	89.1	84.2	10.1	39.3	17.6	5.4
IS	2.2	7.7	64	20.6	14.2	81.9	91	12.4	30.5	11.7	3.1
DK	2.1	16.2	71.5	60.3	17.7	92.4	85.7	5.5	56.3	23.5	5.2
SK	1.8	0.6	63.9	7.6	1.4	91.7	49.7	0.7	8.4	1.4	0.1
SE	1.6	21.3	79.7	56.9	25.6	93.5	89	8	61.2	32.4	8.9
NO	1.4	23.5	89.4	59.1	19.9	92.2	93.3	9.8	65.7	34.6	8.6
SI	1.4	1.2	54.3	9.5	3.3	79	70.4	3.8	11.8	2.9	0.5

Note: Very easy = making ends meet very easily. Capacity= can afford to finance unexpected expenses. Burden1 = housing costs not at all a burden. Burden 2 = others costs not at all a burden. Noarrears = not in arrears. Holiday = can afford one week holiday away from home. Wealth proxy = owner-occupier and property income more than 5% of gross income.

Source: Authors' elaborations from the EU-SILC users' database 2012 (version August 2014).

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or
  calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

KS-TC-16-027-EN-N

High income and affluence: Evidence from the European Union statistics on income and living conditions (EU-SILC)

This paper examines the top tail of the income distributions in the 2012 EU-SILC data. First, it discusses issues related to data quality, including under-estimation of top incomes. Then, the data are used as they are to compute several income-based measures of affluence. Finally, the link between non-income information and high incomes is analysed. The paper shows that EU-SILC is a useful complementary source on high incomes, in particular when the aim is to measure the size of the economically very well-off group. It also shows that identifying the affluent only on the basis of relative incomes is not sufficient. In a number of countries, many households in the upper tail of the income distribution report having difficulties in making ends meet.

For more information
http://ec.europa.eu/eurostat/

Publications Office

ISBN 978-92-79-64150-3