

IRISS-C/I

*An Integrated Research
Infrastructure in the
Socio-economic Sciences*

**Extreme incomes and the estimation of poverty and
inequality indicators from EU-SILC**

by

Philippe Van Kerm





Extreme incomes and the estimation of poverty and inequality indicators from EU-SILC

Philippe Van Kerm
CEPS/INSTEAD

Abstract Micro-data estimates of welfare indices are known to be sensitive to observations from the tails of the income distribution. It is therefore customary to make adjustments to extreme data before estimating inequality and poverty statistics. This paper systematically evaluates the impact of such adjustments on indicators estimated from the EU-SILC (Community Statistics on Income and Living conditions) which is expected to become the reference source for comparative statistics on income distribution and social exclusion in the EU. Emphasis is put on the robustness of cross-country comparisons to alternative adjustments. Results from a sensitivity analysis considering both simple, classical adjustments and a more sophisticated approach based on modelling parametrically the tails of the income distribution are reported. Reassuringly, ordinal comparisons of countries are found to be robust to variants of data adjustment procedures. However, data adjustments are far from innocuous. Cardinal comparisons of countries reveal sensitive to the treatment of extreme incomes, even for seemingly small adjustments.

Reference IRISS Working Paper 2007-01, CEPS/INSTEAD, Differdange, Luxembourg

URL <http://ideas.repec.org/p/irs/iriswp/2007-01.html>

The views expressed in this paper are those of the author(s) and do not necessarily reflect views of CEPS/INSTEAD. IRISS Working Papers are not subject to any review process. Errors and omissions are the sole responsibility of the author(s).

Extreme incomes and the estimation of poverty and inequality indicators from EU-SILC*

Philippe Van Kerm†

February 2007

Abstract

Micro-data estimates of welfare indices are known to be sensitive to observations from the tails of the income distribution. It is therefore customary to make adjustments to extreme data before estimating inequality and poverty statistics. This paper systematically evaluates the impact of such adjustments on indicators estimated from the EU-SILC (Community Statistics on Income and Living conditions) which is expected to become the reference source for comparative statistics on income distribution and social exclusion in the EU. Emphasis is put on the robustness of cross-country comparisons to alternative adjustments. Results from a sensitivity analysis considering both simple, classical adjustments and a more sophisticated approach based on modelling parametrically the tails of the income distribution are reported. Reassuringly, ordinal comparisons of countries are found to be robust to variants of data adjustment procedures. However, data adjustments are far from innocuous. Cardinal comparisons of countries reveal sensitive to the treatment of extreme incomes, even for seemingly small adjustments.

Keywords: Social indicators; poverty; inequality; extreme incomes; parametric tail; EU-SILC

*Paper prepared for the Eurostat and Statistics Finland International Conference on “Comparative EU Statistics on Income and Living Conditions: Issues and Challenges”, Helsinki, 6-7 November 2006. Comments by Alessio Fusco, Stephen Jenkins and conference participants are gratefully acknowledged. An abridged version of this paper is appearing in Eurostat’s *Monographs of Official Statistics* (2007). Disclaimer: All views, errors and omissions are the author’s. Eurostat bears no responsibility for the results and conclusions reported in this paper.

†Centre d’Etudes de Populations, de Pauvreté et de Politiques Socio-Economiques/International Networks for Studies in Technology, Environment, Alternatives, Development (CEPS/INSTEAD), B.P. 48, L-4501 Diferdange, Luxembourg. Institute for Social and Economic Research (ISER), University of Essex, Colchester, United Kingdom. Contact: philippe.vankerm@ceps.lu.

1 Introduction

Regulation (EC) No. 1177/2003 of the European Parliament and of the Council the European Union of 16 June 2003 has given legal existence to EU-SILC, the *Community Statistics on Income and Living Conditions*, a new instrument aiming at collecting comparable micro-data about the income and living conditions in the European Union. EU-SILC has been created as one of the key data sources for estimating the statistical indicators for monitoring and reporting on social inclusion that were endorsed at the Laeken European Council in December 2001. By providing a common data source with comparable individual- and household-level micro-data on income and living conditions, EU-SILC is meant to foster distributional comparisons both over time and across European countries. EU-SILC is therefore expected to become a key instrument for benchmarking national performance with regard to redistributive and social policies, since it will allow the computation of poverty and inequality statistics consistently estimated across countries and over time throughout the European Union.

Given this anticipated usage, estimation of welfare indices from EU-SILC ought to be as accurate and consistent as possible. Accuracy of estimates of poverty and inequality involves a bewildering array of issues. It ranges from the mere definition of the underlying concept of “economic well-being” that one is trying to capture (typically, a person’s access to goods and services), or the definition of the relevant income components, to the selection of appropriate summary welfare indicators, via the definition of the basic unit of analysis and the within-household income sharing assumptions. These issues are well appreciated – the report of the *International Expert Group on Household Income Statistics* (Canberra Group, 2001) provides a thorough discussion of many of these issues.¹ Many of their recommendations were incorporated in the European regulations which define the official guiding framework for EU-SILC.²

It also is well-known, in particular since the work of Cowell & Victoria-Feser (1996a, 1996b, 2002), that welfare indicators estimated from micro-data can be very sensitive to the presence of a few extreme incomes. This is particularly problematic for many indicators of inequality which are not robust to the presence of data contamination at one or both ends of the distribution. This formally means that a single datum, provided it is sufficiently large (or small), can drive the estimated inequality indicator arbitrarily large (or small). By contrast, poverty indicators are considered robust provided the poverty line is exogenously determined (or is itself robustly estimated), but this generally holds under the assumption that income data are positive, or at least are bounded from below (that is, can not be arbitrarily small). Welfare indicators –inequality indicators in particular– are therefore potentially biased if the data are contaminated by ‘mistakes’ taking the form of very high or very low incomes.

Erroneous extreme observations can arise for various reasons. They can be gross mistakes, such as miscoding of a decimal separator or they can be due to severe reporting error by survey respondents. Mistaken extreme values can also arise even if income data are cor-

¹See also Atkinson *et al.* (2002) or van der Laan (2006) and Verma (2006) for a focus on EU-SILC.

²See, in particular, the Council and European Parliament regulation 1177/2003 and the Commission Regulation 1983/2003.

rectly collected because the measured annual income is not an error-free signal of a person's economic well-being. Think of extremely low (especially negative) incomes. Non-positive economic well-being (viewed as the access to goods and services) is an implausible situation as it would imply starvation. However, there are several reasons for observing very low, even negative, incomes in micro-data.³ Such observations may not be plainly tagged as “mistakes” in the sense of error of data collection but they are clear expressions of a mis-measurement of economic well-being that lead to extreme measured incomes.

Another implication of the sensitivity of welfare indicators to extreme incomes is that, even if there is no contamination in the data –extreme incomes are real, accurate measurements of people's well-being–, the sampling variability of welfare indicators can be large because of the sparseness of very high/very low incomes in the underlying population, thereby limiting the reliability of inferences made to the overall population of a country. The close link between robustness to contamination and sampling variability is formally clear from the *influence function* of the estimators which serves both as a tool for assessing the robustness properties of a statistic (see Hampel *et al.*, 1986) and as a component of their sampling variance in some linearization methods (see Deville, 1999).⁴ In addition, notwithstanding the large sampling error problem, one may question whether it is acceptable that a few data points –a few responding households– have a large leverage on estimated national indicators.

In recognition of these issues, it is customary to inspect the data and make some simple adjustments to extreme incomes prior to estimating indicators, such as eliminating or recoding a fraction of the data.⁵ The objective is to keep the influence of extreme incomes under control, thereby limiting the risk of making large (potentially unbounded) estimation errors and reducing the sampling variability of the estimates.⁶ However, while data adjustments have potential benefits and are often deemed necessary, in the vast majority of cases, their application is of an *ad hoc* nature and one rarely estimates the magnitude of their impact on the estimated indicators, or assesses the sensitivity of the resulting estimates to alternative adjustments (recoding rather than deleting data, for instance). In the context of EU-SILC which primarily involves cross-country comparisons, the problem is compounded because differences in the prevalence of extreme observations across countries is likely to lead to different impacts for different data adjustments. There is therefore an issue about the influence

³First, some elements are counted as income deductions. Negative incomes can arise because of taxes that have to be paid on incomes received in an earlier year (Eurostat, 2005). Losses can be observed with self-employment income (Eurostat, 2006b). Inter-household mandatory payments (alimonies) may also be a source of substantial income deductions. Second, several sources of income are not captured in standard definitions of disposable income (e.g. capital gains or home-production). Furthermore, income is measured during a limited time period but people can draw on past (and future) incomes to maintain their command over goods and services. See Eurostat (2006c) for further discussion.

⁴See also Osier (2006).

⁵Eurostat (2006c), for example, discusses the particular problems posed by zero and negative incomes and suggests a series of data adjustments; see *supra*.

⁶Some analysts have also tended to favour indicators thought to be more robust to the presence of extreme incomes (e.g. percentile ratios) over indicators with more attractive theoretical properties (such as (generalized) Gini, Generalized Entropy or Atkinson inequality indices).

of such practice in cross-country comparisons.

This paper attempts to shed light on this matter by presenting the results a sensitivity analysis based on EU-SILC 2004 data, the first official release of EU-SILC (Eurostat, 2006a). The exercise is meant to provide a systematic and comprehensive empirical assessment of the sensitivity of welfare indicators to extreme incomes and to assess the robustness of cross-country comparisons to alternative data adjustments.⁷ Both simple, classical adjustments and a more sophisticated approach based on modelling parametrically the tails of the income distribution are considered. Reassuringly, ordinal comparisons of countries appear to be generally robust to data adjustment procedures. However, the magnitude of the impact of seemingly small adjustments turns out to be such that cardinal comparisons of countries are surprisingly sensitive to the treatment of extreme incomes.

The rest of the paper is structured as follows. Section 2 sets the scene by describing the data and identifying the magnitude and source of extreme incomes in the 2004 EU-SILC data. Data adjustments for dealing with extreme incomes that are considered in the sensitivity analysis are described in Section 3. Results of the analysis are reported in Section 4. Section 5 summarizes the main lessons and provides a final discussion.

2 Extreme incomes in EU-SILC 2004

2.1 The EU-SILC 2004 data

The purpose of the EU-SILC is to ensure that EU countries possess the required statistical infrastructure to estimate reliable indicators reflecting the multi-dimensionality of poverty and social exclusion. It is expected to become the reference source for comparative statistics on income distribution and social exclusion in Europe. As indicated at the start of the paper, EU-SILC has a formal legal basis which makes its implementation compulsory in EU member states. The Council and European Parliament regulation 1177/2003 defines the scope of EU-SILC, provides definitions, time reference, data characteristics, sampling rules, sample sizes, *etc.* In addition, detailed technical aspects of implementation are defined in a set of Commission Regulations. The micro-data available in EU-SILC are expected to be representative of the population living in private households in each of the participating countries.

It is important to realize that EU-SILC is based on a common framework (that is, a common set of target variable definitions and rules), but it is not a common, European-wide survey. Distinct surveys are held in different countries with potentially different designs and implementation, within the framework provided by the regulations. Comparability is ensured by the common framework and definition of target variables, but flexibility is allowed in the collection of the data. Target variables on income, for example, are collected from household surveys in some countries while it is extracted from administrative sources in other countries.

The first official release of EU-SILC (Eurostat, 2006a), which is now available from Eurostat for research purposes and which is used in this paper, contains data collected in 2004

⁷See Cowell *et al.* (1999) for a similar exercise.

in 14 countries, namely Austria (AT), Belgium (BE), Denmark (DK), Estonia (EE), Spain (Es), Finland (FI), France (FR), Greece (GR), Ireland (IE), Italy (IT), Luxembourg (LU), Norway (NO), Portugal (PT), and Sweden (SE). Data from 25 Member states plus Norway and Iceland will be available from EU-SILC 2005 onwards (to be released in 2007).

The core of EU-SILC is collection of detailed data on income both at the household and individual level. The nature and definitions of the income components that are collected in EU-SILC are specified in detail in its legal framework. These specifications adhere as closely as possible to the recommendations of the International Expert Group on Household Income Statistics (Canberra Group, 2001). The income components measured are meant to allow the estimation of four main (household) aggregates: total disposable household income, total disposable household income before transfers (with and without old-age and survivor benefits) and total gross income. Differences in implementation of EU-SILC and transitory exceptions to the recommendations of the Canberra Group have led to two key distinctions in the EU-SILC 2004 income data. First, four countries have obtained income information from administrative sources (DK, FI, NO, SE) while the other ten countries have used surveys. Second, a number of countries have been temporarily allowed to deliver net incomes (or a mix of gross and net incomes) at the components level (rather gross income as required in the regulations).

Extensive documentation about the EU-SILC project and data is available in the aforementioned Council and European Parliament regulation and the associated Commission regulations, as well as in working group reports available from Eurostat. Clémenceau *et al.* (2006) also provide a detailed description of the EU-SILC project.

2.2 Income components and extreme incomes

Given the complexity of the household income aggregates as a combination of different components, and potential cross-national differences in the prevalence and way of collection of these incomes, it is useful to inspect the structure of the data in detail in order to identify the potential sources of extreme incomes.

Figures 1.1 to 1.5 provide depictions of the income data contained in the EU-SILC 2004 dataset in each of the fourteen countries for which data are available. For the eight countries presented in the first three figures, it is the gross income components that are presented. For the latter six countries, income components are normally reported net. All constituent income components are described along with the constructed aggregate “total household gross income” (where available) and “total disposable household income”. For each country, the plots are composed of three distinct panels. The first panel presents the income range within which 98 percent of the observations are contained. (All incomes are expressed relative to the mean disposable household income in the country.) The mean value of the income component (among all non-zero values) is marked by a vertical bar. Ranges that span to the left of the first vertical dashed line indicate that it contains negative values. Ranges that span to the right of the second vertical dashed line extend to higher incomes than household disposable income (for example, obviously, total gross incomes span to higher values than disposable incomes). The second panel gives the proportion of non-zero incomes. The proportion of

negative values is highlighted in light grey. The third panel gives the proportion of observations that are above the upper end of the range of disposable income (marked by the second vertical dashed line in the first panel).

The main observations emerging from inspection of these figures are the following. Consider first the group of countries with incomes presented gross in EU-SILC 2004.

Employee (cash) income is the most important income component. It has a wide range of variation and is received by a large fraction of households in all countries. Overall, cross-country variations are not very large (compared to other components). Expectedly, self-employment incomes signal themselves as potentially problematic in further analysis. First, they can turn negative in several countries (BE, DK, EE, LU, NO), and sometimes substantially so (see DK and NO). Second, it is generally the income component with the widest range of variation (see IE where the inner 98 percent of observations span from 0 to up to 8 times mean disposable income). Very high self-employment incomes are observed in all countries. Furthermore, in both aspects, substantial differences emerge across countries. Important cross-country differences also emerge in the proportion of households reporting receipt of self-employment incomes (from 9 percent in BE up to 40 percent in DK). Besides actual differences, several artefacts can be the cause of these discrepancies: differences in the mode of collection of the self-employment data which are allowed within the EU-SILC framework (e.g. drawings out of business, operating profits/losses or official tax declaration; deduction or not of losses), differences in patterns of under-reporting and non-response (which is substantial for self-employment income), differences in imputation rules. (See the discussion in Eurostat (2006b) and Eurostat (2006c).)

Old-age benefits are the benefits with the largest range, and some observations are recorded as “very high” in several countries. However, globally, benefits are unlikely to be problematic with regard to distribution tails. Most have a narrow range of variations. Cross-country differences also exist, but are less striking than differences in self-employment income. Rents from property or land also appear relatively innocuous, with very large observations found only in FI, IE and LU. Interests and dividends may be more problematic. First, the proportion of households recorded as receiving interests and dividends vary widely from 5 percent in EE up to 99 percent in NO. Second, the range of variation is generally small but there are striking exceptions in NO and FI. This leads to the presence of a number of outlying incomes in these two countries (as well as in BE and DK). The situation of DK is also peculiar with 58 percent of households receiving negative interests and dividends (however the range of variation is small). Inter-household transfers and taxes on wealth are unlikely to play any significant role with regard to income distribution tails and the generation of outlying observations. Finally, the overall picture of income taxes and social security contributions is remarkably similar across countries.

These observations are not much affected when considering the six countries reporting income components in net amounts (or partially net and partially gross amounts). Self-employment incomes still appear potentially problematic. They contribute many outlying observations in disposable income in the upper tail (along with employee income), and may contain substantial negative values. Rent from property or land now appear potentially large

as well. Note finally the very compressed income ranges observed in Sweden (overall and in most income components).⁸

2.3 The prevalence of extreme incomes in key income aggregates

The prevalence of very high or very low incomes in the aggregate income variables of interest here, namely household disposable income and single-adult-equivalent household income is presented in Figure 2.⁹ The bars to the left indicate the proportion of households which record an income below 10 percent of the mean. Negative incomes among those are highlighted in light grey. The bars to the right present the proportion of households recording incomes above 4 times the mean. Different colouring indicates the proportion of households with income between 4-8 times the mean and 8 times the mean and above.

All countries (except AT and PT) have households with negative incomes (with most prevalence in GR and EE). Their number is well below one percent of the observations. In many countries, the fraction of observations with incomes below 10 percent of the mean reaches one percent. Countries with the largest fractions of very low incomes also tend to have the highest fractions of very high incomes (EE and PT). Almost one percent of the observations in most countries are recorded above four times the average income in the country. The picture is hardly affected if equivalent income is used rather than total disposable income. Extreme observations are present in all of the countries datasets, and there is therefore interest in assessing how much comparisons of distributional indicators are affected by their presence.

3 Dealing with extreme incomes in inequality and poverty measurement

As explained in the Introduction, welfare measures are potentially sensitive to the presence of extreme incomes, large or small. Additionally, some measures are not even identified if there are non-positive incomes in the data. Adjustments to the data are therefore routinely applied before estimating inequality and poverty measures. Given the important role given to EU-SILC for monitoring progress in the EU with respect to poverty and inequality, it appears important to engage in a thorough sensitivity analysis in order to better understand the impact of such adjustments and re-assure ourselves that conclusions are not dramatically affected. For this purpose, a series of inequality and poverty indicators have been estimated using EU-SILC 2004, after applying a number of different, relatively standard, data adjustments meant to deal with extreme incomes. Six different types of adjustments, have been experimented

⁸A noticeable feature is the range of taxes on wealth in France. They, however, only concern a very small fraction of households.

⁹Welfare indicators are typically estimated from the single-adult equivalent household income which takes into account the size and age composition of households. Disposable income is divided by a scale which converts all incomes into a common standard equivalent to a single-adult household. See *supra*.

with. They are described in turn in this Section.

3.1 Trimming and Winsorizing

The first two types of data adjustment are common, naive methods: trimming and winsorizing a fixed percentage of the data at either or both ends of the income distribution.

Trimming the data is a simple strategy to prevent extreme incomes to influence estimated statistics. It consists in removing from the dataset a given number or a given percentage of the highest and/or lowest incomes (see, e.g., Bernstein & Mishel, 1997, Acemoglu, 2003). This implicitly considers that extreme observations contain no valid information about the economic well-being of the recipients and should not have any influence in the resulting estimates. Trimming as a tool for making ‘robust’ welfare comparisons in distributional analysis is thoroughly discussed in Cowell & Victoria-Feser (2006), while Hampel *et al.* (1986) discusses such an approach in contrast to more sophisticated procedures.¹⁰

Trimming thresholds above and beyond which observations are discarded are typically determined by quantiles $Q(1 - 0.01p)$ and $Q(0.01p)$ of the unadjusted dataset where p is the proportion of data to be discarded at each tail. In the sensitivity analysis, estimations have been made with trimming percentages of 0.25%, 0.50%, 0.75%, and 1%, both one-sided and two-sided. Additionally, to demonstrate the potentially large impact of just a few observations, estimation has been run by trimming only the single highest income, the top 5 incomes, and the top 10 incomes as well as the bottom 1, 5 and 10 incomes.

Winsorizing is a close relative to trimming. The difference is that the extreme data are not removed from the dataset but are replaced by the value of the trimming thresholds (see, e.g., Atkinson *et al.*, 1995, Gottschalk & Smeeding, 1997, Burtless, 1999, Gottschalk & Smeeding, 2000). While trimming drives the influence of extreme incomes to zero by eliminating them, winsorizing allows them to keep a high influence on the estimates, yet imposing a limit. Winsorizing can be seen as a particular form of income imputation. If fixed percentages of the data are adjusted, the adjusted data can be expressed as

$$y_i^a = \max(Q(0.01p), \min(y_i, Q(1 - 0.01p))).$$

Winsorizing is also referred to as ‘top-coding’ or ‘bottom-coding’ which is often applied with respect to data confidentiality issues. The impact on inequality of top-coding incomes in the American Current Population Survey data has been much researched and has been found to be significant (Fichtenbaum & Shahidi, 1988, Bishop *et al.*, 1994).

3.2 Parametric tail modelling and robustness weights

Trimming and, to a lesser extent, winsorizing are the most commonly adopted practices for making estimates robust to outlying observations. However, Hampel *et al.* (1986) emphasize that this practice can be viewed as overly conservative –especially trimming– in the sense

¹⁰Cowell & Victoria-Feser (2006) show formally that Lorenz curve ordinates (from which most inequality measures can be derived) become robust when trimming is applied.

of trading-off too much data information against robustness, thereby leading to a loss of efficiency. One more sophisticated approach to addressing robustness problems of inequality and poverty measures is to estimate a parametric model for the income distribution tails (e.g. the Pareto distribution function is a common choice) whose parameters are estimated using methods robust to outlying observations. The robust, parametrically estimated tails are then combined with the empirical distribution function for the bulk of the data to obtain a semi-parametrically estimated distribution of incomes from which inequality and poverty indicators can be estimated. This approach is detailed in Victoria-Feser & Ronchetti (1994) and Cowell & Victoria-Feser (2007). Brazauskas (2003) or Davidson & Flachaire (2004) contain illustrations of such methods. The good performance of this parametric tail modelling strategy has recently been demonstrated by Cowell & Flachaire (2007).

As such, this approach is not based on data adjustments. Tail distribution parameters are estimated from the available data and welfare indicators are estimated from the combined empirical distribution function and tail parametric distributions (using, e.g., numerical integration algorithms). It is however possible to use robust parametric models for devising data adjustment procedures. The idea is to impute extreme incomes by replacing the observed extreme values by random draws from the robustly estimated parametric tail distributions. Simulation uncertainty is introduced by the random draws, but this is easily controlled by simulating a set of replicate income data, estimating poverty and inequality on each of the replicate, just as one would do with the original micro-data, and taking as indicator the average over the replications –a practice familiar in multiple imputation procedures (Little & Rubin, 1987)–.

The robust estimation of a parametric tail model is obviously much more technically challenging than standard, naive adjustments, but there are expected gains to the exercise because this approach is meant to result in a more optimal trade-off between data information and robustness as emphasized by Hampel *et al.* (1986).¹¹

In the sensitivity analysis, we have implemented this semi-parametric approach using a Pareto distribution as the parametric tail model, as in Davidson & Flachaire (2004), Cowell & Flachaire (2007), or Cowell & Victoria-Feser (2007), with cumulative distribution function given by

$$F^H(y; \theta^h, y^h) = 1 - \left(\frac{y}{y^h} \right)^{-\theta^h}$$

for $y > y_h$. An inversed Pareto distribution was used for the lower tail:

$$F^L(y; \theta^l, y^l) = \left(\frac{2y^l - y}{y^l} \right)^{-\theta^l}$$

for $y < y_l$. There is a long tradition of using the Pareto distribution to fit the upper tail of income distributions (see Kleiber & Kotz, 2003), but the practice is less common for modelling

¹¹Note that, in practice, for EU-SILC, the technical difficulty can be circumvented if estimation of the tail distribution parameters is made centrally, e.g. by Eurostat, and multiply imputed values for extreme incomes are distributed along with the datasets. Estimation is then no more difficult than from the original data. Analysts would not be required to engage into the more difficult parameter estimation stage themselves.

the lower tail. The present exercise should be taken as illustrative for this model. Further investigation is called for to confirm the validity of this choice and to consider alternative specifications.¹²

Standard maximum likelihood methods could be used to estimate the parameters of these tail distributions. However, maximum likelihood estimates are not robust and are known to be themselves sensitive to the presence of extreme incomes. It is therefore useful to estimate the parameters with an algorithm that provide robust estimates of the Pareto distribution parameters.¹³ The procedure applied in this analysis is the so-called *optimal B-robust estimator* (OBRE) detailed in Victoria-Feser & Ronchetti (1994), Victoria-Feser (2000), or Cowell & Victoria-Feser (2007).

The cut-off points y^l and y^h have been determined respectively as

$$\begin{aligned} y^l &= \min(\max(0.3\mu, Q(0.02)), Q(0.03)) \\ y^h &= \max(\min(2.5\mu, Q(0.98)), Q(0.97)) \end{aligned}$$

where μ is mean income and $Q()$ are quantiles. This choice was selected by trial and error trying to achieve a good fit of the models while keeping a common rule applied to all countries. The models fitted well the upper tail of the distributions in all countries. Estimation for the lower tail was more problematic in two countries (Portugal and Sweden). The estimated tail models were then used to impute extreme incomes. The top and bottom 0.5% of the income data were multiply imputed by making random draws from the estimated tail distributions.¹⁴ Eight replicate values were drawn for each extreme income.

The OBRE algorithm used to estimate the Pareto parameters robustly is an iterative algorithm which involves determining iteratively robustness weights to all the data points. Therefore, a by-product of the algorithm is a set of weights that reflect how much “influential” each datum is (Hampel *et al.*, 1986, Victoria-Feser & Ronchetti, 1994). Data with a weight of 1 are considered non-outlying according to the model, whereas deviating observations end up with a weight between 0 and 1 that reflects the degree of “deviation”. These weights were exploited to devise yet another possible approach to handle extreme incomes consisting in keeping all income data unaffected, but multiplying the sample weights by the “influence weights” returned by the application of the OBRE algorithm. Application of these adjusted weights when computing poverty and inequality indicators partially offsets the effect of the largest and smallest observations but retain them in the dataset.¹⁵

¹²In particular, it is conceivable to model the lower tail of the distribution with a model preventing non-positive values, although this may not fit the observed datasets which actually contain such data.

¹³Cowell & Flachaire (2007) note, however, that more simple, non-robust estimation of tail distribution parameters already lead to significantly improved estimation and inference.

¹⁴This means that between 2 and 3 percent of the data are used to estimate the parameters of the distribution tail, but only the highest/lowest 0.5 percent are actually imputed. The objective is to obtain stable tail-parameter estimates but to adjust only the most extreme data. Cowell & Flachaire (2007) proceed similarly.

¹⁵Note that trimming can be seen as a particular case of data re-weighting where observations receive either a weight of 1 or a weight of 0 if they fall below/above the trimming thresholds. Other re-weighting procedures that down-weight extreme data could also be considered. See, for example, Fields & Smith (1994).

3.3 Adjustments for zero and negative incomes

Eurostat (2006c) proposes a number of simple adjustments specific to the treatment of zero and negative incomes found in EU-SILC. To complement the analysis based on the adjustments described in the previous sub-sections, three of these suggestions have also been included in the sensitivity analysis –although one must bear in mind that they are meant to handle negative incomes only, not extreme incomes in general.

The first adjustment consists in removing zero and negative incomes. This is a form of one-sided trimming where the threshold is determined by a fixed value (zero), not by a given quantile. The second adjustment is a winsorizing procedure which consists in lifting up all negative incomes to zero. As an alternative, all zero and negative incomes were also replaced by a value equal to 10% of median income and 25% of median income. The third adjustment consists in imputing zero and negative incomes by randomly drawing income values from a set of observed positive income data. We experienced with drawing positive incomes from the lower decile group of observed incomes, and drawing from the lower quartile group. This ensures that negative and zero incomes are replaced by low, but positive, incomes.

3.4 Dropping recipients of unreliable income sources

As a final check, a drastic data adjustment was applied to assess the impact and the sensitivity of cross-country comparisons to the exclusion of incomes notably unreliably measured, namely self-employment income and income from interests, dividends and profits. Observations were discarded if the considered income source represented more than a quarter of either gross household income or disposable household income. This procedure should obviously not be taken as a standard for estimating inequality and poverty as people relying on self-employment incomes represent a substantial population. However as we compare inequality or poverty for sub-populations which, it can be argued, report their income more reliably, it provides a benchmark to assess the potential influence of unreliable income sources on country rankings in terms of welfare indicators (for a similar exercise, see Cowell *et al.*, 1999).

4 Are cross-country comparisons of inequality and poverty measures sensitive to extreme income adjustments?

Summary measures of inequality and poverty are numerous. Because indices may have different sensitivity to extreme incomes, it was decided to include a wide array of alternative indicators in the analysis in order to provide a comprehensive record. In total, twenty-five welfare indicators have been estimated: two central tendency indicators, fourteen inequality measures and nine poverty indicators.

The central tendency indicators are the mean and median equivalent income. It is expected that only the mean is affected by the adjustments. Note that both can be used to determine poverty lines and therefore their sensitivity gives indication about the sensitivity of the determination of the poverty line.

The inequality indicators considered are the following: two percentile ratios (P80/P20 and P90/P10) which are robust in the sense that arbitrarily set income values can not make the ratio arbitrarily large; two income share ratios (S80/S20 and S90/S10) which are non-robust statistics; the coefficient of variation; the standard deviation of log-incomes; the Gini coefficient; the relative mean deviation (also known as Schutz coefficient); a set of Generalized Entropy measures (GE(0), GE(1), GE(2)) and Atkinson inequality measures (A(0.5), A(1), A(2)) which are known to be non-robust and potentially very sensitive to extreme incomes as well as undefined in the presence of non-positive incomes (with the exception of GE(2)). The Gini coefficient and the S80/S20 income share ratio are both in the list of “Laeken indicators”, which officially identifies key indicators for monitoring social cohesion in EU Member States. For discussion and definitions of these inequality measures, see e.g. Jenkins (1991), Cowell (1995) or Cowell (2000).

The poverty indicators considered are: three Foster-Greer-Thorbecke with parameters 0 (a.k.a. the headcount ratio or at-risk-of-poverty rate), 1 (a.k.a. the average poverty gap ratio) and 2 (a.k.a. the average squared poverty gap ratio); the Watts index; the Sen-Shorrocks-Thon index; the median poverty gap ratio among the poor (a.k.a. the relative median at-risk-of-poverty gap). The poverty line is set at 60 percent of the median equivalent income. The headcount ratio is also estimated with a line set at 50 percent of the median. Finally FGT(0) and FGT(1) are also estimated for households with dependent children only. The at-risk-of-poverty rate and the relative median at-risk-of-poverty gap are both Laeken indicators. See Zheng (1997) for a survey of these (and other) poverty measures.

All indicators were estimated for each of the fourteen countries from the single-adult equivalent disposable household income. Data were weighted by the household size times the household sample weight in order to depict the distribution of income among individuals.

Full results are collected in Figures 3.1 to 3.25. Each figure presents the estimates for one particular indicator. Estimates for the 14 countries are reported as columns of points (each country corresponding to a specific symbol). Each column of points corresponds to estimates obtained after a particular data adjustment for extreme values. The picture in the top panel reports the values of the estimated indicators. The bottom panel reports the relative difference between the indicator estimated without any adjustment for extreme values and the indicator obtained after the data adjustment (relative reductions are plotted, hence a positive value indicates that the estimated indicator after adjustment is smaller than the estimator obtained without any adjustment). Note that the bottom panel is empty whenever indicators can not be estimated from the unadjusted datasets. The estimates presented in the figures are as follows:

- The first column presents estimates obtained from the raw, full sample, that is, without any adjustment. (These benchmark points are repeated in the last column for easier reference.)
- The next set of columns (labelled “A Trim obs.”) helps assessing the sensitivity of estimates to the removal of just a few extreme incomes. In the first three columns, the top 1, 5 and 10 observations are removed from the data. In the next three columns the bottom 1, 5 and 10 incomes are trimmed. The last three columns show results of trimming the 1, 5 and 10 extreme incomes at both tails.

- The second set of columns (labelled “B Trimming %”) shows results of applying a systematic trimming of a fixed percentage of the data, respectively 0.25%, 0.5%, 0.75%, and 1%, either at the top only, the bottom only, and trimming the same percentage on both tails.
- The third set of columns (labelled “C Winsorizing”) is equivalent to the previous results except that the fixed percentages of the data are winsorized rather than trimmed.
- The fourth set of columns (labelled “D Model-based”) presents the results based on applying a parametric-tail model. The first column presents estimates obtained by imputing the top 0.5% of the data by drawing from a Pareto distribution (with parameters estimated from the OBRE algorithm). The second column presents estimates based on re-weighted data with robustness weights obtained from the OBRE algorithm for the upper tail Pareto distribution. Subsequent columns present the same results applying the Pareto modelling to the lower tail only, and to both the lower and upper tails.
- The fifth set of columns (labelled “E Adjust ≤ 0 ”) presents estimates obtained by applying specific adjustments for zero and negative incomes, namely dropping zero and negative incomes, lifting these incomes to either 0, a tenth of the median, or a quarter of the median, and imputing zero and negative incomes by randomly drawing income values from the bottom decile group or the bottom quartile group.
- The last set of columns (labelled “F Sources”) presents estimates obtained after removing households relying on self-employment income, or on interests, dividends and profits, or both.

All results are reported in this paper for the record and completeness. For the sake of brevity, only the main tendencies as well as a few observations are discussed here, however. Focus is put on differences in the pictures provided under different data adjustments, it is not the purpose of the exercise to make any substantive interpretation about the particular values taken by the indicators in the different countries.

Consider first, briefly, the central tendency measures. Expectedly, the median is hardly affected by any adjustment for extreme values. Interestingly, the median remains largely unaffected if we drop households living on self-employment income or interest, dividends and profits. This suggests that these households are found in similar proportions on both the upper half and the lower half of income distributions. Mean income is somewhat more affected, especially by adjustments to the upper tail of the distribution, but not so much as to lead to any reversals in the ranking of countries. These observations provide support to the case of making reference to median income to set poverty lines.

Inequality indices are potentially the most problematic indicators because many of them are known to be sensitive to extreme values and/or require adjustment for extreme low (zero and negative) incomes for identification. Many of the twelve indicators reported do indeed reveal themselves sensitive to data adjustments. In most cases, it is adjustment for top incomes that is most influential. However, for indices that do not allow non-positive incomes

(SD of log-incomes, Atkinson indices or most Generalized Entropy measures), adjustments for low incomes are required and are therefore relevant too.

Quantile ratios reveal themselves as the most robust indicators of inequality. However, this is at the cost of giving in substantive relevance because such indicators do not take into account what happens beyond the quantile points considered. Estimates of quantile ratio are slightly influenced by trimming at either tail of the distribution, but the effect is small and common enough in all countries not to lead to any rank-reversals of countries. Expectedly, imputation-type adjustments leave quantile ratios unaffected. Removal of self-employment incomes mostly affects results for southern European countries (Greece, Italy and Spain) with a marked inequality reduction (of up to 10 percent for P90/P10), but otherwise this impact is limited.

Income share ratios reveal themselves much more sensitive to data adjustments. This had to be expected because, as opposed to quantile ratios, they are non-robust statistics potentially driven arbitrarily large or small by either very large or very small incomes. The S80/S20 indicator being a Laeken indicator, it is worth paying particular attention to this measure (Figure 3.5). Adjustments in both tails have an impact. It is striking to notice that the removal of just 10 observations at the top can lead to marked changes in the estimated indicators (often by more than 5 percent). Trimming one percent of the data at each tail generally leads to inequality reductions in the range of 10-15 percent. Imputation-based approaches, especially Winsorizing or model-based imputation, have a more modest impact -generally well below 10 percent. The ordering of countries tends to be preserved by all approaches. A marked exception is the comparison of Greece and Italy: while lower tail adjustments hardly affect estimates for Italy, their impact on estimates for Greece is such that the ranking of the countries depend on the adjustment. More generally, cardinal differences between countries are now markedly affected (compare also Spain and Ireland, for example). Expectedly, these remarks hold for the S90/S10 ratio but with impacts of more important magnitudes. Adjustments often lead to apparent inequality reductions well beyond 10 percent (up to 30 percent for trimming) and reranking of countries are more numerous.

Consider now the second Laeken indicator among the inequality measures, namely the Gini coefficient (Figure 3.9). The behaviour of the Gini coefficient is very similar to the S80/S20 index, with the distinction that the Gini coefficient is less affected by adjustments to lower incomes. Upper tail adjustments are more influential, but their impact remain relatively limited like for the S80/S20 index. Again, ordinal ranking of countries is largely preserved (compare Spain and Ireland or Italy and Greece, however) but cardinal comparisons are affected, especially by adjustments for top incomes.

The impact of data adjustments on the Gini or the S80/S20 are nothing compared to their impact on inequality coefficients such as the coefficient of variation (Figure 3.7) or the GE(2) index (Figure 3.13) which are very sensitive to the treatment of high incomes. Reductions of estimated indices by more than 50 percent are common. The example of Norway is the most striking: elimination of a single top income reduces the GE(2) index by almost 60 percent. On the contrary, the relative mean deviation (Figure 3.10) turns out to be relatively stable.

Other indices are more difficult to assess because they are not identified on the full sample which contain zero or negative incomes. Only adjustments that eliminate these values can

be considered. This, by itself, is a source of concern. Winsorizing or trimming at least one percent of the low income data lead to the estimation of the indicators for (almost) all countries. The model-based approach does not prevent negative incomes and, therefore, often does not allow estimation of such measures.

Most poverty measures have been reported to be robust to data contamination when the poverty line is exogenously defined (or is itself robustly defined), provided incomes are bounded from below (Cowell & Victoria-Feser, 1996a). Given the presence of negative incomes in the data –therefore of no explicit lower bound for incomes–, there is interest in checking the behaviour of the poverty indicators to the treatment of extreme incomes, in particular low incomes.

Unsurprisingly, besides removing self-employment income recipients, only trimming has a somewhat marked impact on the headcount ratio. The impact remains relatively low anyway, at least for a poverty line set at 60% of the median income (in general less than 5 percent); see Figures 3.17 and 3.23.

The picture varies substantially for the other poverty measures. The magnitude of variation of the average poverty gap ratio (FGT(1)) and the Sen-Shorrocks-Thon index are similar (mostly below 20 percent); see Figures 3.18 and 3.21. Unsurprisingly, the average squared poverty gap ratio reveals itself very sensitive to extreme incomes with variations often in the range 20-80 percent, irrespective of the data adjustment procedure (Figure 3.19). Ordinal comparisons for this index are sensitive to data adjustments (see, e.g., the cases of Greece and Belgium). The Watts index is both sensitive to data adjustment and requires that all non-positive incomes are adjusted. Lower tail adjustments –trimming in particular, but not exclusively– lead to marked estimated poverty reductions with these poverty measures. Somewhat more reversals in the ranking of countries are observed than for inequality measures, in particular if trimming is applied. Winsorizing, as well as the specific adjustments for zero and negative incomes, seem to be more influential than model-based approaches for poverty measures. Finally, note that the median poverty gap ratio among the poor –another Laeken indicator– is affected by trimming but is unaffected by Winsorizing or the model-based approaches (Figure 3.22). Similar observations emerge if we consider the impact of the various adjustments on poverty in a sub-population of interest (households with dependent children).

5 Summary and discussion

The main lesson that emerges from the exercise is probably that ordinality in cross-country comparisons is generally preserved, irrespectively of the data adjustment procedure applied. Marked rank reversals are rarely observed because of the treatment of extreme incomes: e.g. high/low inequality or poverty countries remain identified as such in all scenarii. Admittedly, this is not a surprising result, but it is certainly a reassuring baseline.

This result must be carefully qualified. Cardinal comparisons of countries *are* sensitive to data adjustments made to control for extreme income data. Even if the relative ranks of countries are rarely affected by the treatment applied to extreme incomes, the apparent *mag-*

nitude of cross-country differences can vary substantially, even with relatively small data adjustments. This will also be relevant in the time-series dimension when additional waves of EU-SILC will become available to track the evolution over time of indicators. Care is therefore called for, and it is recommended to check the sensitivity of one's cardinal comparisons to different data adjustments before making strong statements about it. This is true for most of both inequality and poverty indicators. Note that poverty measures are not notably less sensitive to the treatment of extreme incomes than inequality measures (as soon as one considers indices more sophisticated than the headcount ratio). It is mostly extreme low incomes, and how they are handled, that matter for poverty indicators whereas inequality indices are more sensitive to extreme high incomes. Theoretically sound inequality indices such as Generalized Entropy measures and Atkinson inequality measures are particularly problematic because they suffer from either estimation impossibility with non-positive values or from high sensitivity to extreme incomes. On the contrary, it is reassuring to note that indices which are part of the list of Laeken indicators (Gini, S80/S20, headcount ratio, median poverty gap ratio among the poor) are among the most stable measures.

Different data adjustment procedures can lead to different results. Adjustments that modify/impute the extreme data without removing them from the dataset lead to results that are markedly more stable than trimming procedures. In particular, the trimming proportions can matter a lot. On the contrary, the proportion of data winsorized, or the limit above which a parametric model is applied tends to be less determining. Again, even if ordinal comparisons are generally preserved under alternative adjustment methods, cardinal comparisons may be affected. For example, both the estimated Gini and S80/S20 inequality indicators fall by about 5 to 20 percent if the top and bottom one percent of the income data are trimmed. Winsorizing the same sample fractions leads to falls of about 3 to 10 percent.

Common data adjustment procedures have been applied to all countries. While country-specific adjustments are hard to defend in such an exercise, one may argue that the amount of data contamination may vary from country to country and that "optimal" adjustments should be tailored for each country. Although this is arguably valid, it is hard to come up with objective arguments for this tailorization. Careful examination of the sensitivity analysis suggests that, provided a common procedure is adopted (e.g. trimming percentages or winsorizing or parametric modeling), adopting different parameters (such as different trim percentages) is unlikely to lead to complete changes in the ordinality of cross-country comparisons for most of the measures. Winsorizing has an edge over trimming in this respect as it tends to be less sensitive to the sample percentage that is "imputed". A similar argument can be put forward for model-based imputation. In addition, parametric-tail modeling is *de facto* selecting country-specific parameters (the parameters of the Pareto distributions) that lead to the best fit to the hypothesized Pareto distribution, although the fraction of the data which are imputed has yet to be decided by the analyst.

The advantage of trimming is the ease of implementation, its effectiveness in discarding the impact of extreme values, its long tradition, and the possible interpretation of the results as depicting what happens to the "inner p percent" of the population, even in the absence of data contamination. However, results show that its effectiveness is at the cost of substantially affecting the estimated indicators and being sensitive to the trimming percentage.

This is consistent with the claim found in Hampel *et al.* (1986) that trimming is trading-off too much valid data information against robustness. This may be particularly true in the EU-SILC since the dataset can be considered as a “clean” dataset. The available EU-SILC user database has undergone substantial pre-processing, and grossly outlying observations have been scrutinized and possibly adjusted already (Eurostat, 2004). Winsorizing is also straightforward and leads to more stable estimates. It suffers however from a lack of natural interpretation; what does the adjusted sample represent if there were no contamination? The imputation approach based on a parametric tail model seems a promising possibility. It has theoretical advantages in terms of ‘optimal’ trade-off between robustness and conservation of data information. Its good performance apparent here confirms what has been demonstrated recently by Cowell & Flachaire (2007) from a series of Monte Carlo studies. As far as interpretation is concerned, it does not modify the underlying distribution but merely assumes that the tails follow a parametric distribution. The observation that parametric tail adjustments lead to much smaller modifications of the results may be indicative that trimming and winsorizing are making excessive data adjustments. Again this is a plausible fact considering the extensive prior data cleaning of the dataset.¹⁶

Self-employment income remains a source of concern: (i) it is a major source of extreme incomes on both tails of the distribution, and (ii) it may substantially affect cross-country comparisons (especially when southern European countries are concerned). Self-employment income is traditionally difficult to collect and often the least reliable among the major income sources. Inspection indeed revealed that it may have leverage on social indicators as a source of extreme incomes. Some important differences across countries have emerged and it can be conjectured that the different practices in the collection of data within EU-SILC is cause of concern. Given the manifest cross-national variability in the resulting data, it would seem important to further harmonize practice in the future of EU-SILC.

The position taken in the analysis was not to consider negative and zero incomes as different from the rest of the data on *a priori* grounds (as in, e.g., Gottschalk & Danziger, 1985, O’Higgins *et al.*, 1989), but rather treat them as extreme incomes. The reason is that, given the definition of household disposable income, non-positive incomes are plausible in EU-SILC’s main household income aggregates. Even if we agree that a household’s command of goods and services can not be below a certain minimal amount to secure the survival of its members, given the limitations of the income measure as an indicator of economic well-being, we can not rule out the presence of ‘true’ small or negative amounts (‘true’ in the sense that they are not the result of errors or mis-reporting in any of the income components collected). They were therefore not treated differently from other extreme data. These values were adjusted according to data adjustment procedures for the lower tail, similarly to adjustments for the upper tail.

Needless to say, the analysis presented in this paper does not help identifying a single all-in-one approach, nor does it help identifying the adjustment that makes the indicators the most accurately representative of the true value of the indicator if economic well-being were

¹⁶The parametric tail approach has also been reported by Davidson & Flachaire (2004) and Cowell & Flachaire (2007) to have virtues with respect to (resampling-based) variance estimation.

observed directly. This would require knowledge of the true, underlying indicator, as well as information about the degree of data contamination in the dataset. All adjustments are simply meant to keep the magnitude of potential errors under control, balancing robustness and data information. No single adjustment guarantees to lead to estimates closer to the “true” underlying welfare indicator. Adjustments ought to be considered in light of the resulting stability of the estimates, and, more importantly, sensitivity checks are useful to re-assure ourselves that conclusions are not dramatically affected by extreme incomes and they way they are handled. Part of the analysis is also exemplary. In particular, the parametric tail approach implemented here would deserve further testing and fine-tuning, especially with regard to the lower tail.¹⁷ Additionally, the analysis has focused on the impact of adjustments on point estimates. A complementary analysis of the variance stabilization achieved by the various adjustments could further help selection of a specific procedure.

One can argue that data adjustments are always hazardous in the absence of objective information about the validity of the measured data. Subsidiary information about sources of economic well-being (people’s accumulated physical or financial assets in particular) would come in useful to assess the reliability of income data, even if they are not themselves incorporated in the income concept. Perhaps more practically, over time, the longitudinal dimension of EU-SILC will become relevant for making reliability assessment of the recorded incomes, both by allowing the estimation of social indicators based on income flows received over longer periods, and by serving as potential checks of household’s reporting of income.

¹⁷Combined approaches could also be considered by, say, trimming negative incomes and applying a parametric tail model to small positive values only. It is also conceivable to adopt an asymmetric strategy with different procedures to handle extreme high and extreme low incomes, although one must bear in mind that both extremes can cause serious trouble.

References

- Acemoglu, D. (2003), 'Cross-country inequality trends', *Economic Journal*, **113**(485):F121–F149.
- Atkinson, A. B., Cantillon, B., Marlier, E. & Nolan, B. (2002), *Social Indicators. The EU and Social Inclusion*, Oxford University Press, Oxford, UK.
- Atkinson, A. B., Rainwater, L. & Smeeding, T. M. (1995), *Income Distribution in OECD Countries: Evidence from the Luxembourg Income Study (LIS)*, no. 18 in Social Policy Studies, OECD, Paris.
- Bernstein, J. & Mishel, L. (1997), 'Has wage inequality stopped growing?', *Monthly Labor Review*, **120**(12):3–16.
- Bishop, J. A., Chiou, J.-R. & Formby, J. P. (1994), 'Truncation bias and the ordinal evaluation of income inequality', *Journal of Business and Economic Statistics*, **12**(1):123–27.
- Brazauskas, V. (2003), 'Influence functions of empirical nonparametric estimators of net reinsurance premiums', *Insurance: Mathematics and Economics*, **32**:115–133.
- Burtless, G. (1999), 'Effects of growing wage disparities and changing family composition on the U.S. income distribution', *European Economic Review*, **43**(4–6):853–865.
- Canberra Group (2001), *International Expert Group on Household Income Statistics: Final Report and Recommendations*, The Canberra Group, Ottawa.
- Clémenceau, A., Museux, J.-M. & Bauer, M. (2006), 'EU-SILC (Community statistics on Income and Living Conditions): issues and challenges', Paper presented at the Eurostat and Statistics Finland International Conference on *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges*, Helsinki.
- Cowell, F. A. & Flachaire, E. (2007), 'Income distribution and inequality measurement: The problem of extreme values', *Journal of Econometrics*, doi:10.1016/j.jeconom.2007.01.001 (forthcoming).
- Cowell, F. A., Litchfield, J. & Mercader-Prats, M. (1999), 'Income inequality comparisons with dirty data: The UK and Spain during the 1980s', DARP Papers 45, Suntory and Toyota International Centres for Economics and Related Disciplines, London School of Economics.
- Cowell, F. A. & Victoria-Feser, M.-P. (1996a), 'Poverty measurement with contaminated data: A robust approach', *European Economic Review*, **40**(9):1761–1771.
- (1996b), 'Robustness properties of inequality measures', *Econometrica*, **64**(1):77–101.
- (2002), 'Welfare rankings in the presence of contaminated data', *Econometrica*, **70**(3):1221–33.

- (2006), ‘Distributional dominance with trimmed data’, *Journal of Business and Economic Statistics*, **24**(3):291–300.
- (2007), ‘Robust stochastic dominance: A semi-parametric approach’, *Journal of Economic Inequality*, **5**(1):21–37.
- Cowell, F. A. (1995), *Measuring Inequality*, Harvester Wheatsheaf, Hemel Hempstead, 2nd edn.
- (2000), ‘Measurement of inequality’, in A. B. Atkinson & F. Bourguignon (eds.), *Handbook of Income Distribution*, North Holland, Amsterdam, The Netherlands, 87–166.
- Davidson, R. & Flachaire, E. (2004), ‘Asymptotic and bootstrap inference for inequality and poverty measures’, GREQAM Working Paper 2004-20, Groupement de Recherche en Economie Quantitative d’Aix-Marseille.
- Deville, J.-C. (1999), ‘Variance estimation for complex statistics and estimators: Linearization and residual techniques’, *Survey Methodology*, **25**(2):193–203.
- Eurostat (2004), ‘Imputation procedures’, EU-SILC Documents 136/04, European Commission, Eurostat.
- (2005), ‘Current treatment of taxes and their implication on negative income and on comparability between countries’, EU-SILC Documents TFMC-1, European Commission, Eurostat.
- (2006a), ‘Cross-sectional EU-SILC 2004 users’ database’, Release June 2006, European Commission, Eurostat.
- (2006b), ‘Self-employment income’, EU-SILC Documents TFMC-02/06, European Commission, Eurostat.
- (2006c), ‘Some proposals on the treatment of negative incomes’, EU-SILC Documents TFMC-15/06, European Commission, Eurostat.
- Fichtenbaum, R. & Shahidi, H. (1988), ‘Truncation bias and the measurement of income inequality’, *Journal of Business and Economic Statistics*, **6**(3):335–37.
- Fields, C. A. & Smith, B. (1994), ‘Robust estimation – a weighted maximum likelihood approach’, *International Statistical Review*, **62**(3):405–24.
- Gottschalk, P. & Danziger, S. (1985), ‘A framework for evaluating the effects of economic growth and transfers on poverty’, *American Economic Review*, **75**(1):153–61.
- Gottschalk, P. & Smeeding, T. M. (1997), ‘Cross-national comparisons of earnings and income inequality’, *Journal of Economic Literature*, **35**(2):633–87.

- (2000), ‘Empirical evidence on income inequality in industrialized countries’, in A. B. Atkinson & F. Bourguignon (eds.), *Handbook of Income Distribution*, North Holland, Amsterdam, The Netherlands, 261–307.
- Hampel, F. R., Ronchetti, E. M., Rousseeuw, P. J. & Stahel, W. A. (1986), *Robust statistics: The approach based on influence functions*, John Wiley, New York.
- Jenkins, S. P. (1991), ‘The measurement of income inequality’, in L. Osberg (ed.), *Economic Inequality and Poverty: International Perspectives*, M.E. Sharpe, New York, 3–38.
- Kleiber, C. & Kotz, S. (2003), *Statistical size distributions in Economics and Actuarial Sciences*, John Wiley and Sons, Inc., New Jersey.
- Little, R. J. A. & Rubin, D. B. (1987), *Statistical analysis with missing data*, John Wiley, New York.
- O’Higgins, M., Schmaus, G. & Stephenson, G. (1989), ‘Income distribution and redistribution: A microdata analysis for seven countries’, *Review of Income and Wealth*, **35**(2):107–131.
- Osier, G. (2006), ‘Variance estimation: The linearization approach applied by Eurostat to the 2004 SILC operation’, Paper presented at the Eurostat and Statistics Finland Methodological Workshop on EU-SILC, Helsinki.
- van der Laan, P. (2006), ‘The income concept in EU-SILC: Relevance, feasibility, challenges’, Paper presented at the Eurostat and Statistics Finland International Conference on *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges*, Helsinki.
- Verma, V. (2006), ‘Issues in data quality and comparability in EU-SILC’, Paper presented at the Eurostat and Statistics Finland International Conference on *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges*, Helsinki.
- Victoria-Feser, M.-P. & Ronchetti, E. (1994), ‘Robust methods for personal income distribution models’, *Canadian Journal of Statistics*, **22**(2):247–58.
- Victoria-Feser, M.-P. (2000), ‘Robust methods for the analysis of income distribution, inequality and poverty’, *International Statistical Review*, **68**(3):277–93.
- Zheng, B. (1997), ‘Aggregate poverty measures’, *Journal of Economic Surveys*, **11**(2):123–62.

Figure 1.1: Income components in Austria, Belgium and Denmark: income ranges, fraction of non-zero observations, and proportion of extreme positive values (gross components)

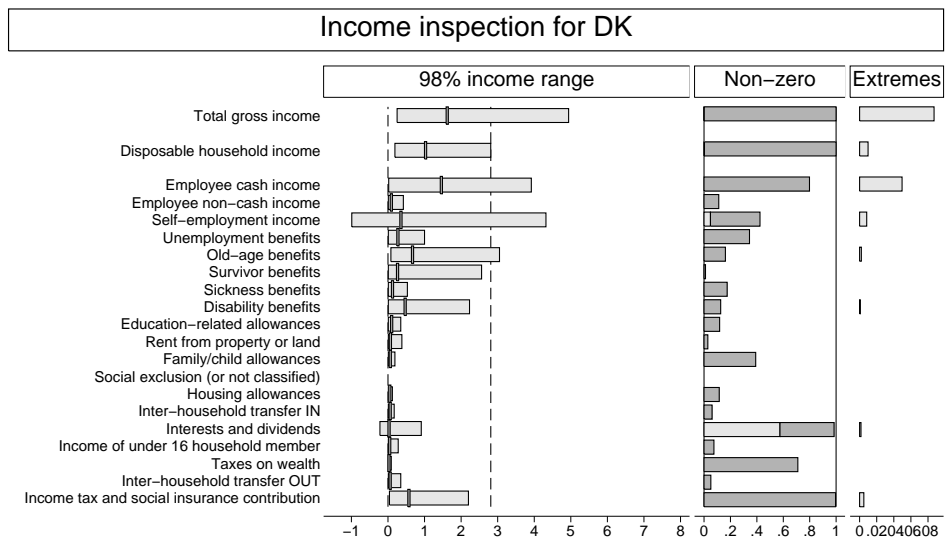
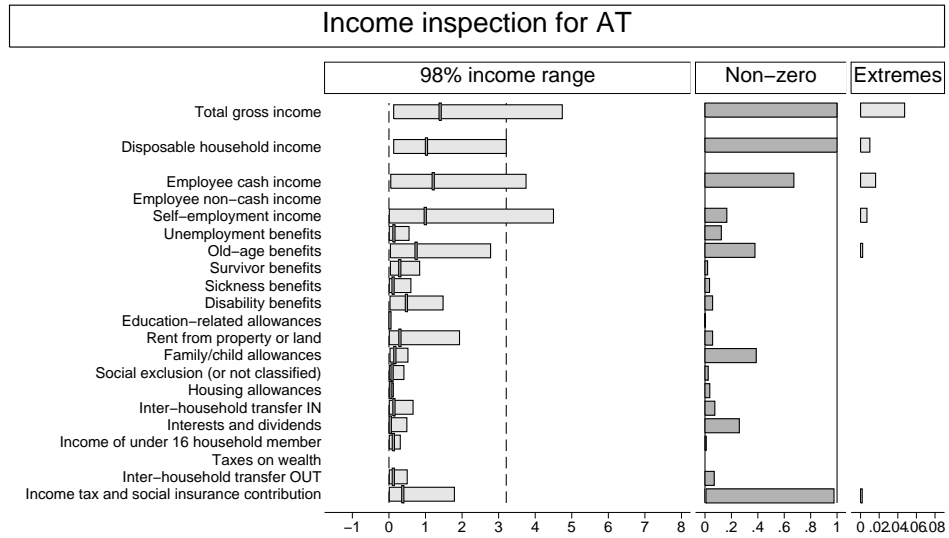


Figure 1.2: Income components in Estonia, Finland and Ireland: income ranges, fraction of non-zero observations, and proportion of extreme positive values (gross components)

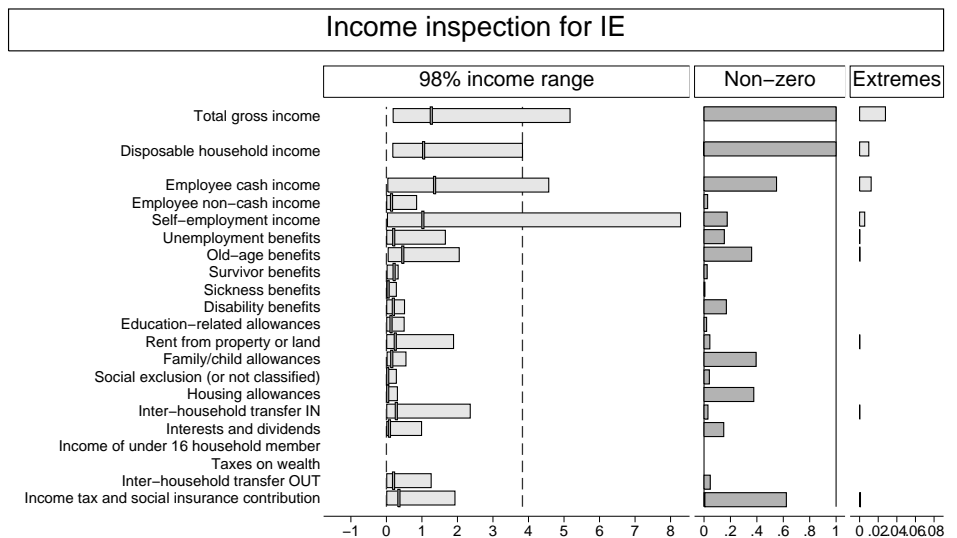
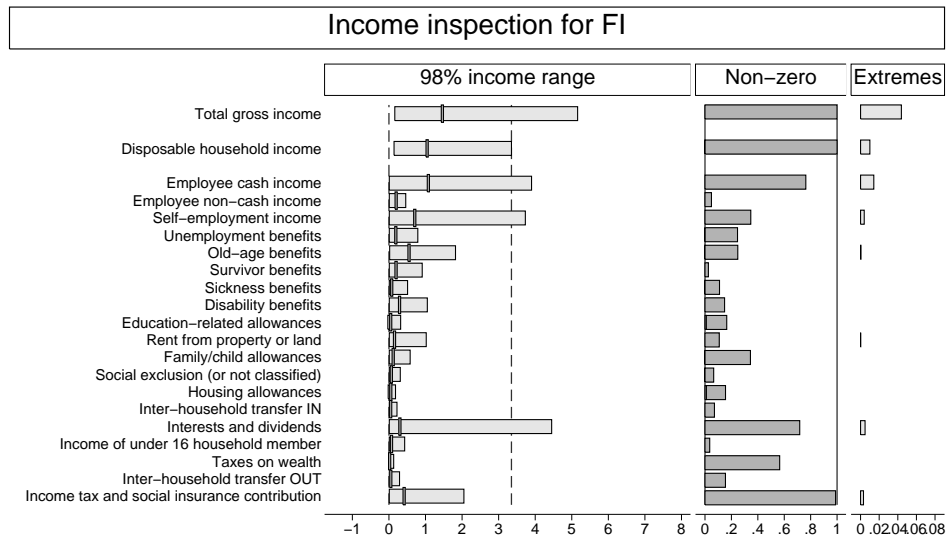
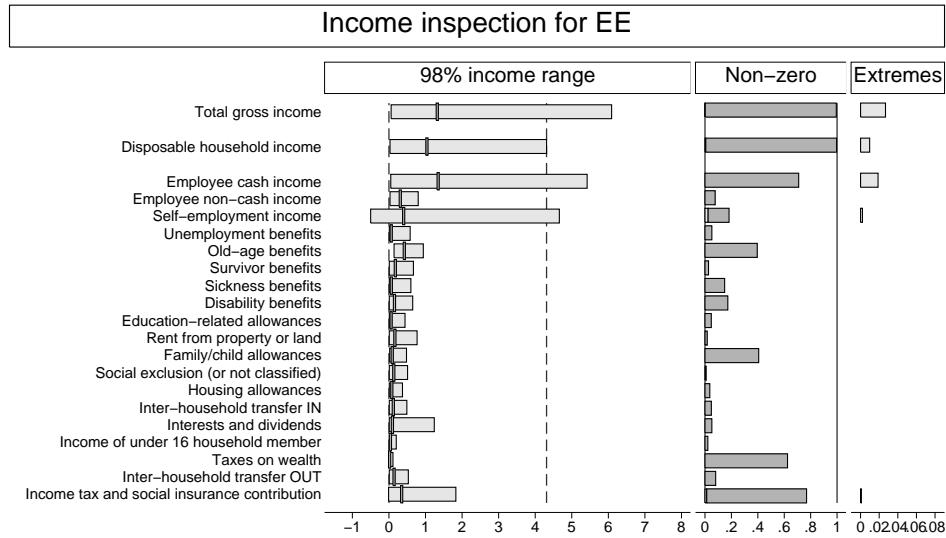


Figure 1.3: Income components in Luxembourg and Norway: income ranges, fraction of non-zero observations, and proportion of extreme positive values (gross components)

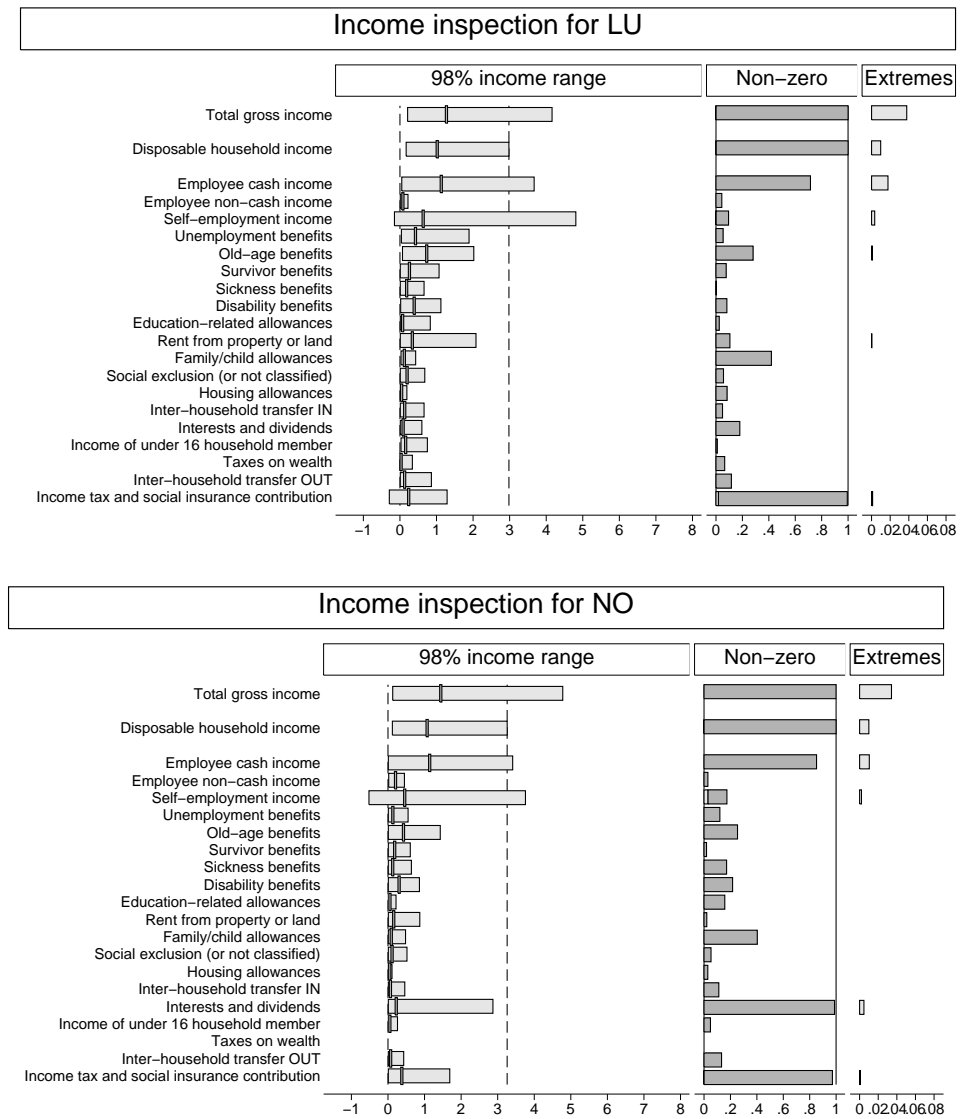


Figure 1.4: Income components in Spain, France and Greece: income ranges, fraction of non-zero observations, and proportion of extreme positive values (net and/or gross components)

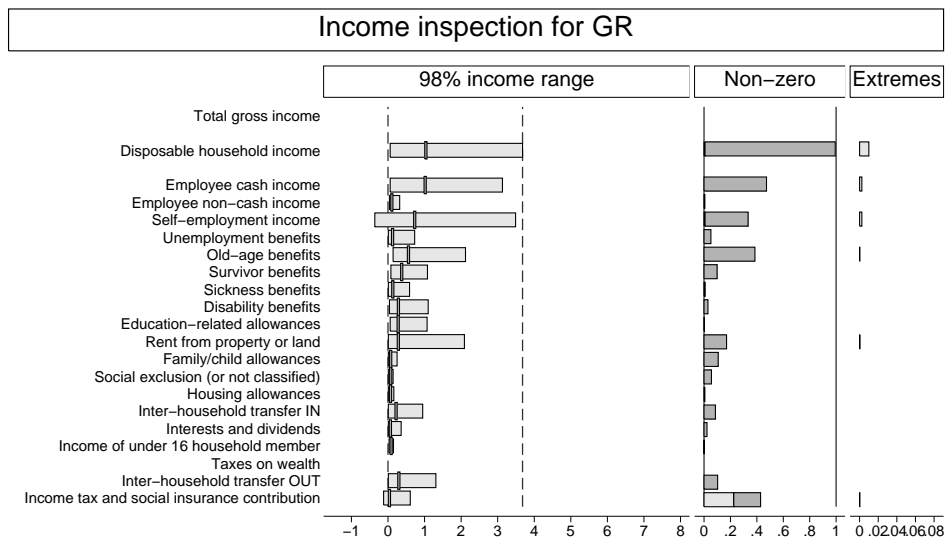
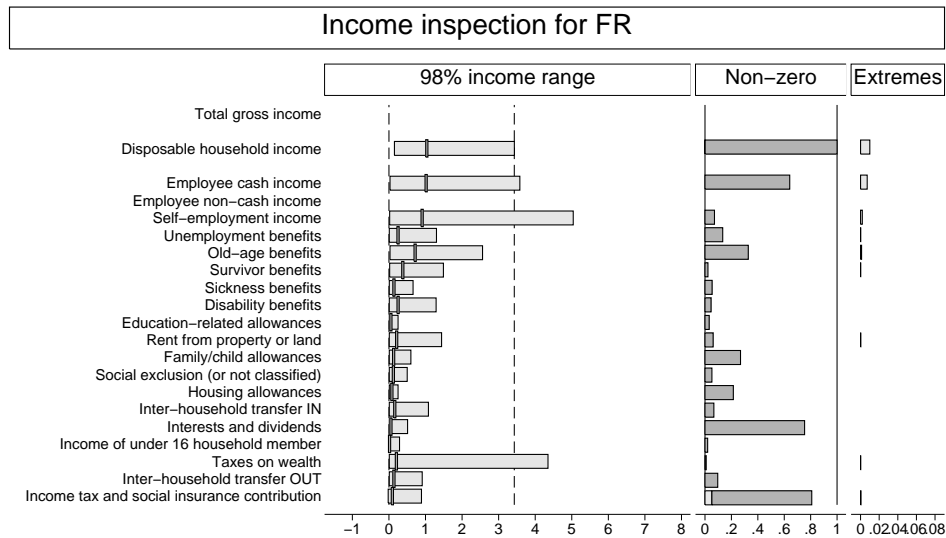
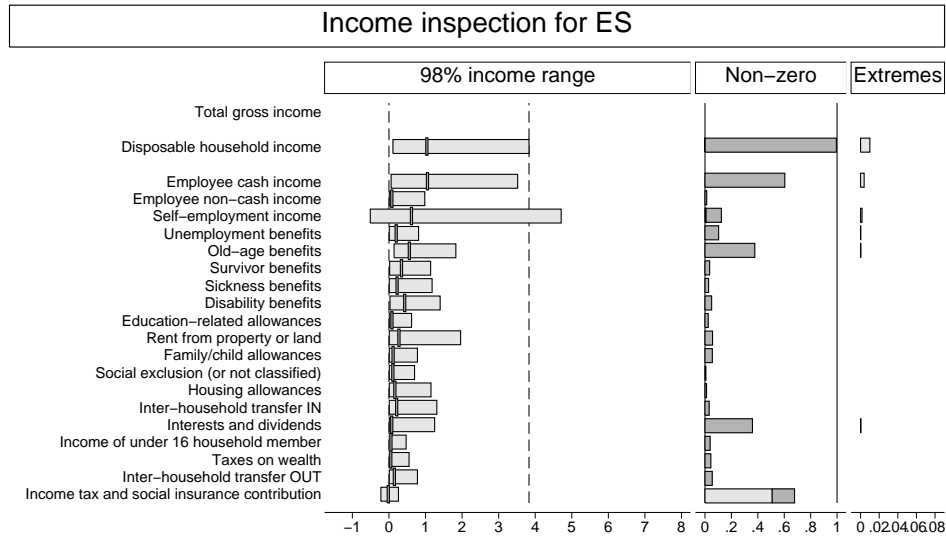


Figure 1.5: Income components in Italy, Portugal and Sweden: income ranges, fraction of non-zero observations, and proportion of extreme positive values (net and/or gross components)

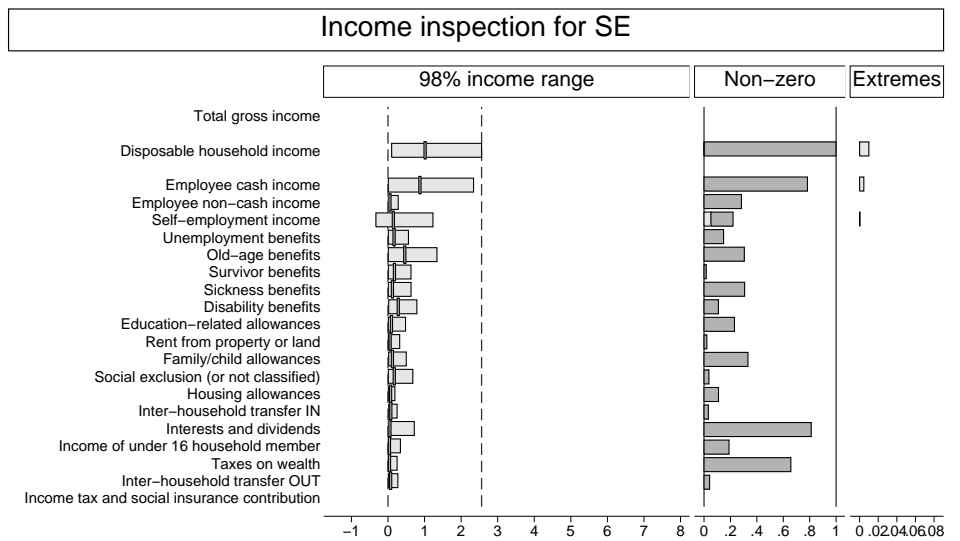
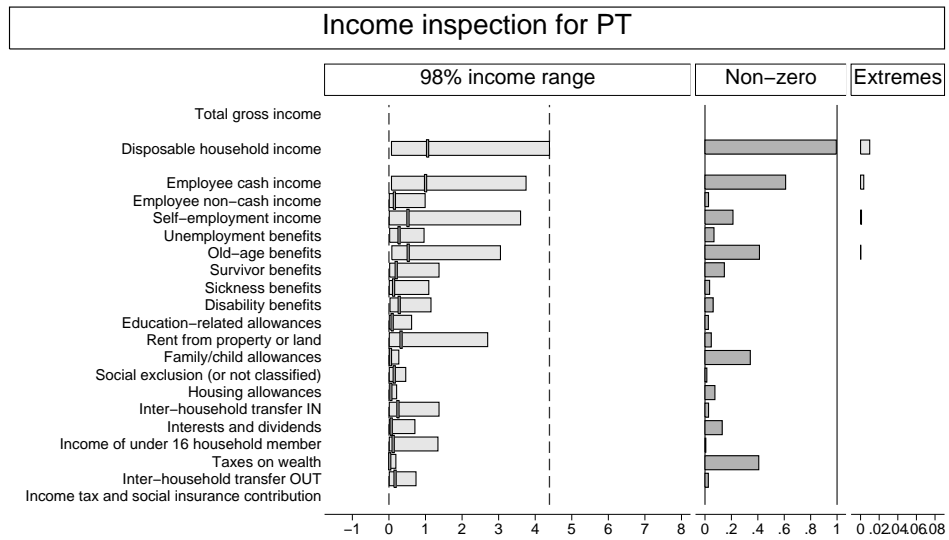
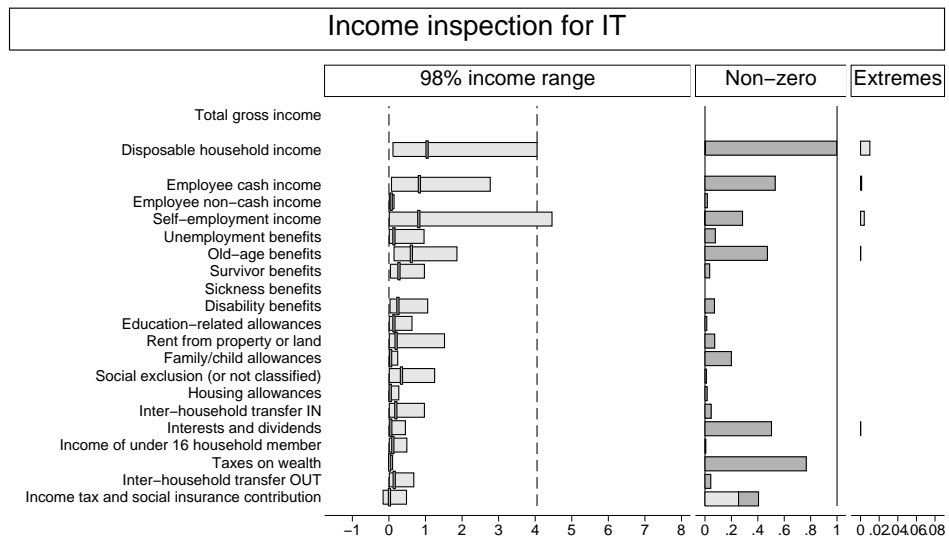


Figure 2: Proportion of very low and very high incomes in disposable income and in equivalent income. Low incomes are incomes below 10 percent of the mean (of which negative incomes are identified in light grey), high incomes are incomes above 4 times the mean (two groups are depicted 4-8 times the mean and 8 times mean income and above).

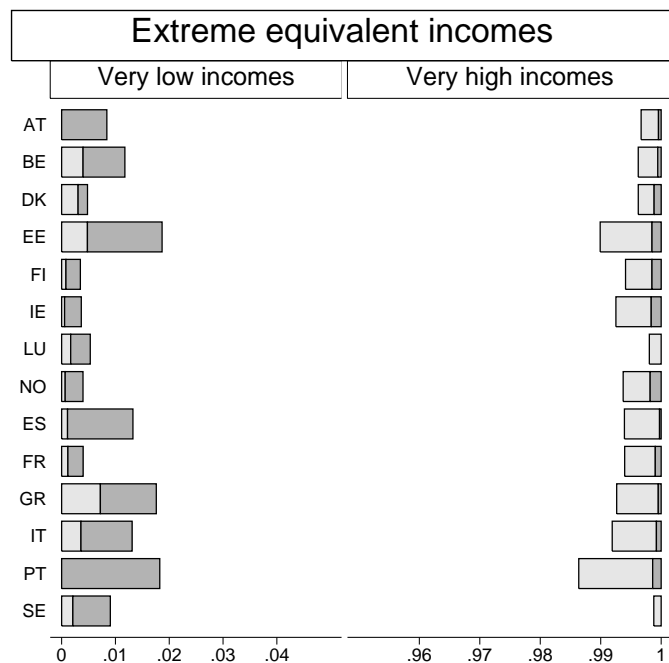
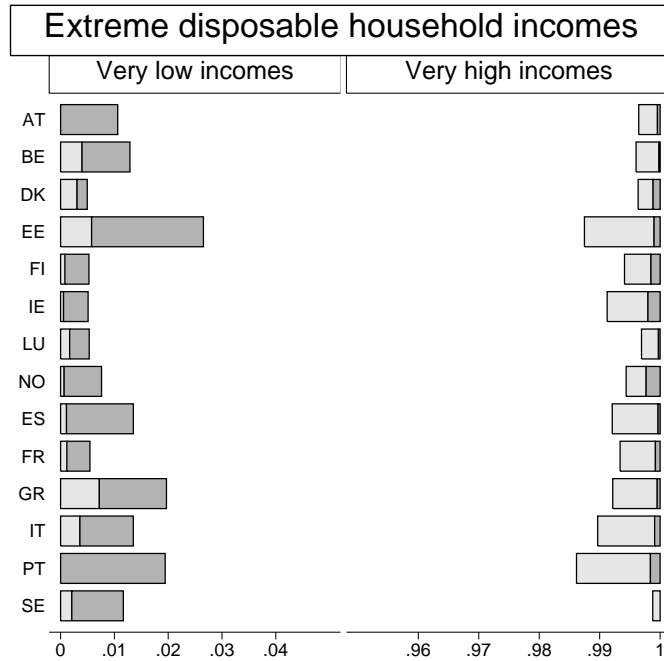
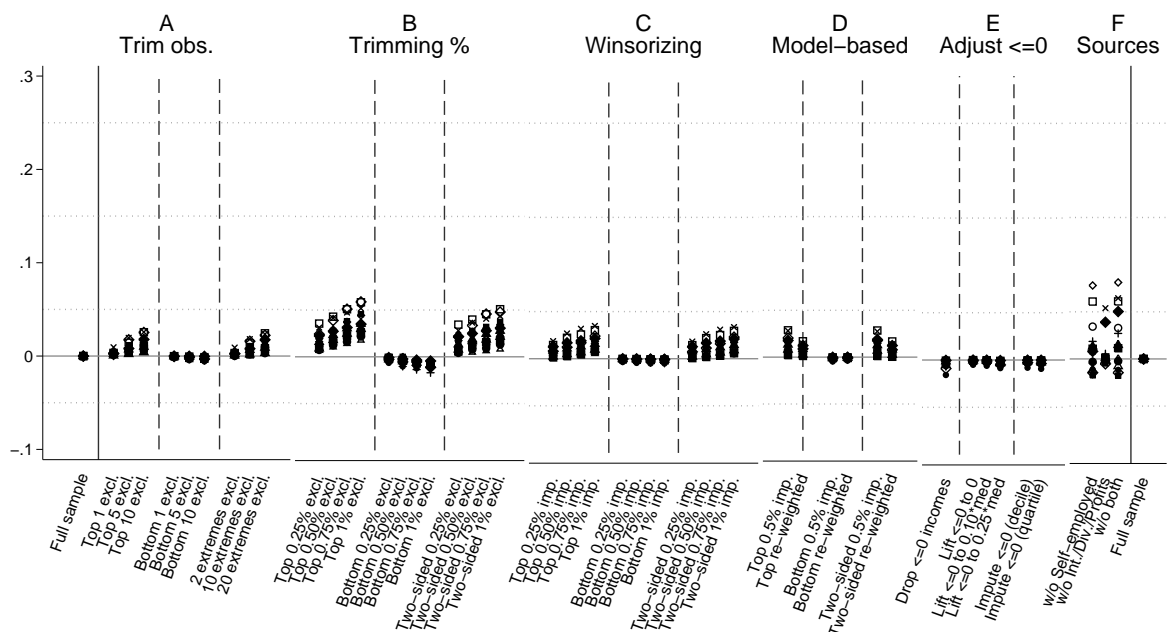
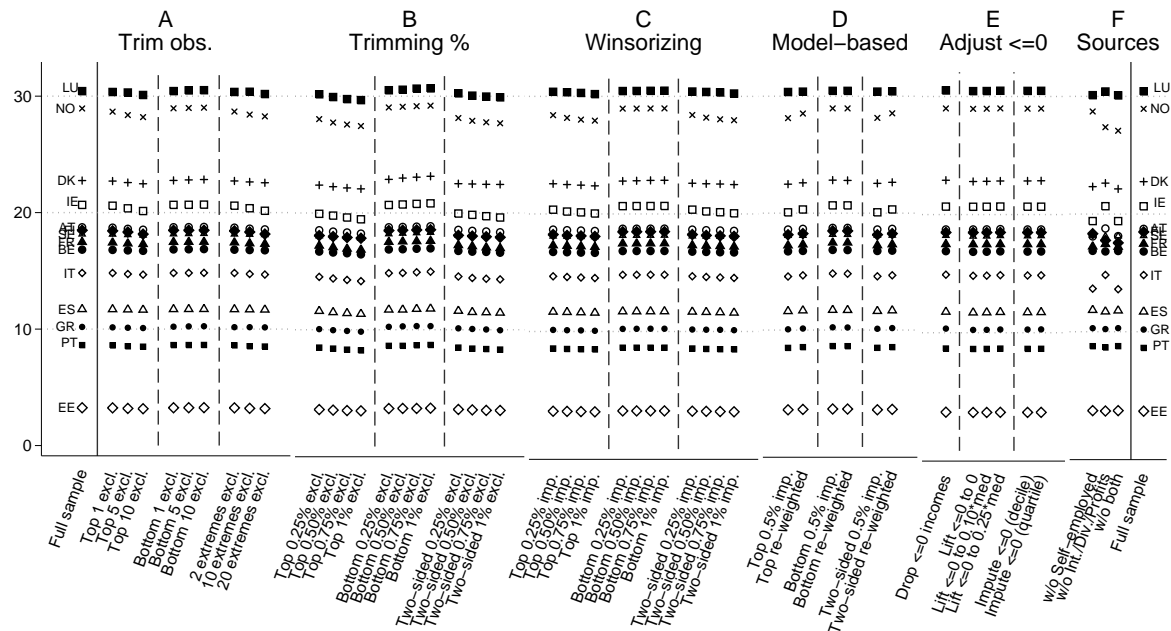
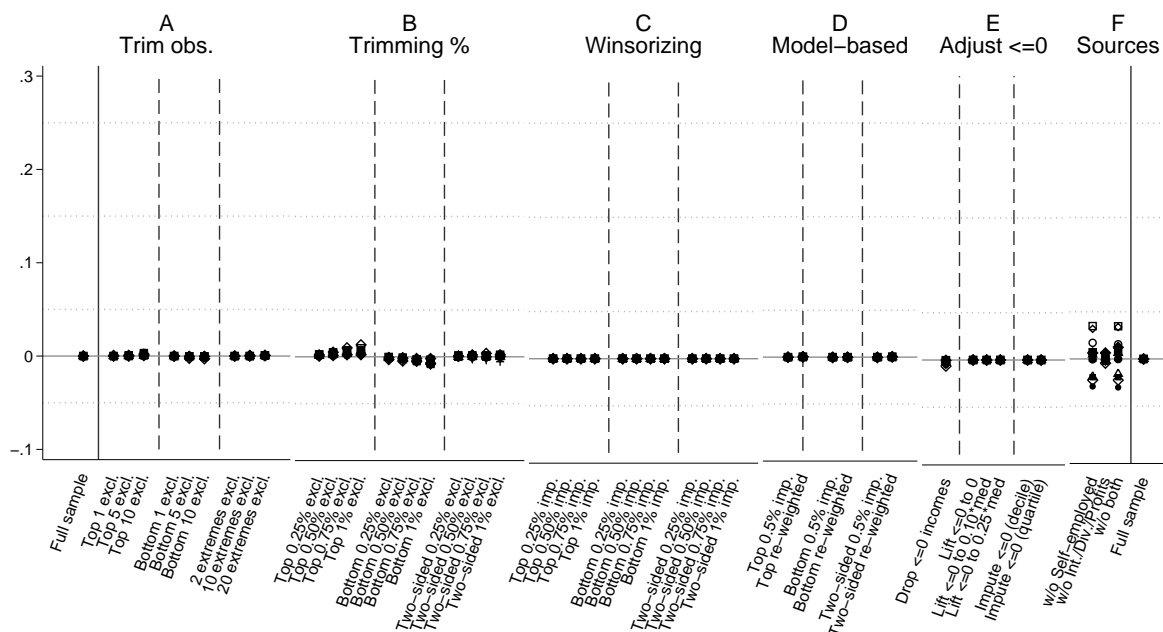
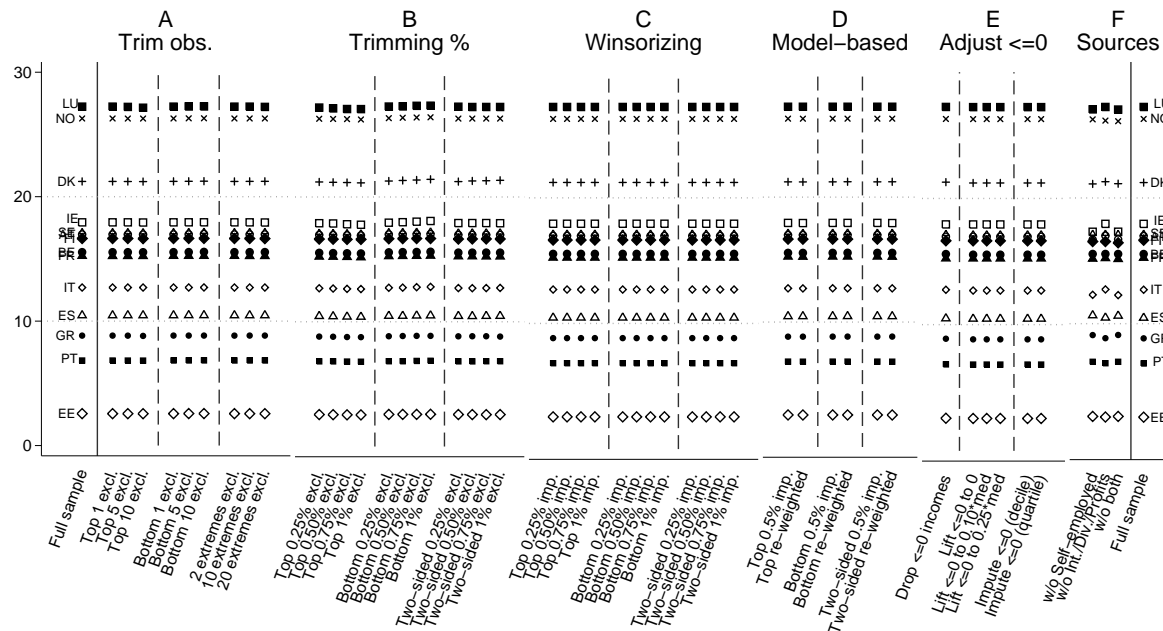


Figure 3.1: Central tendency measures: Estimates of mean equivalent income (in '000 euros) in 14 countries under alternative treatment of extreme incomes



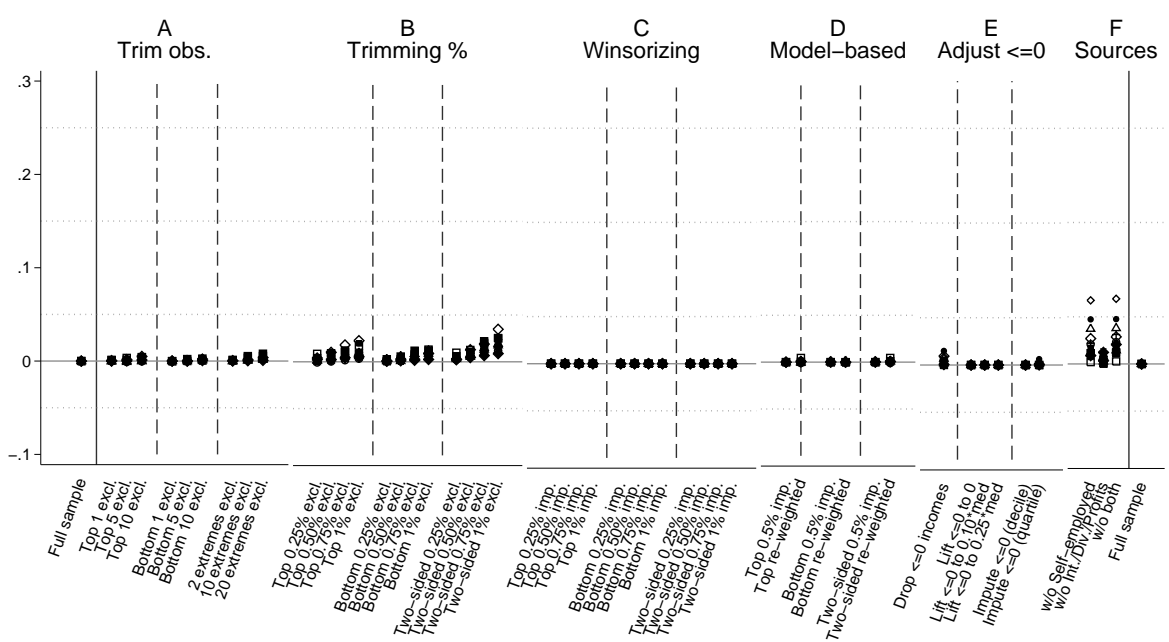
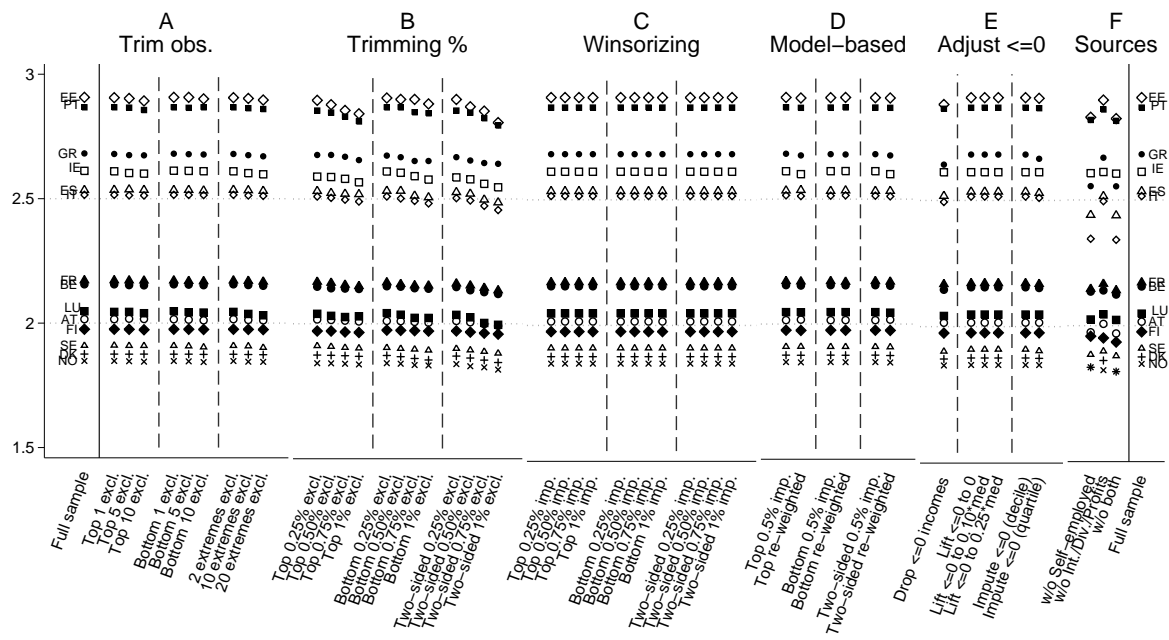
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.2: Central tendency measures: Estimates of median equivalent disposable income (in '000 euros) in 14 countries under alternative treatment of extreme incomes



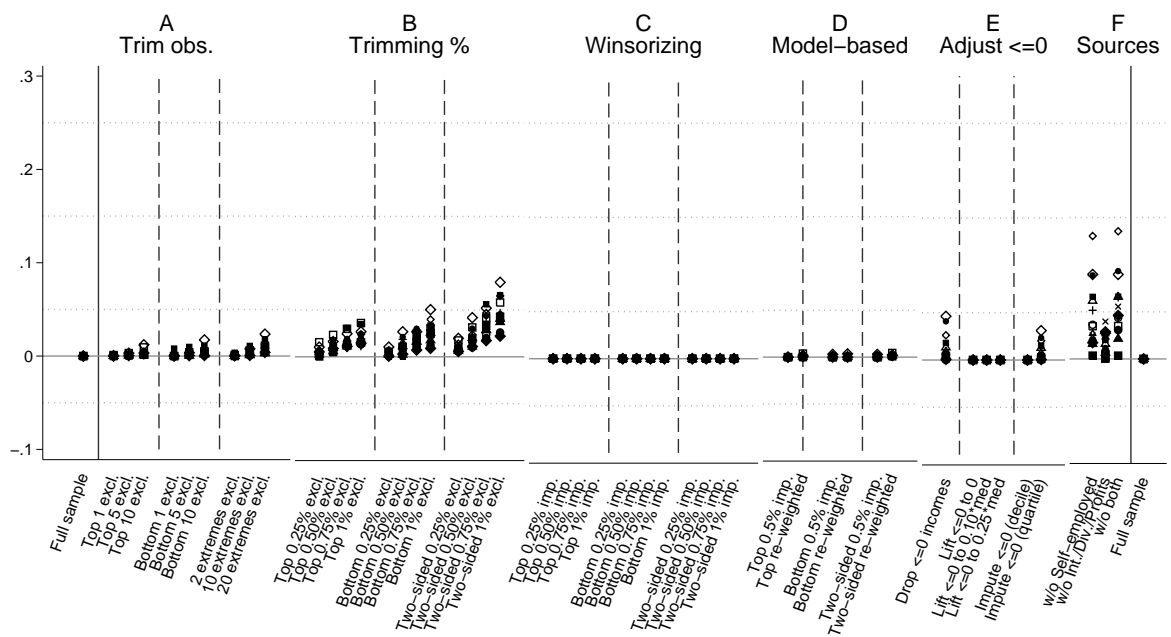
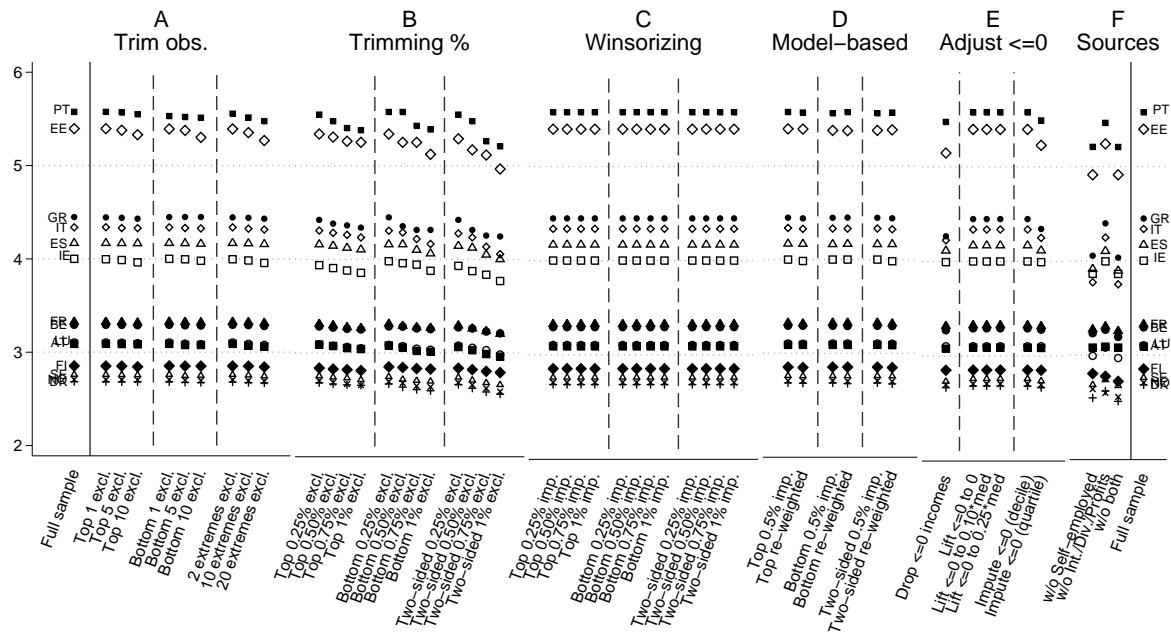
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.3: Inequality: Estimates of the P80/P20 quantile ratio in 14 countries under alternative treatment of extreme incomes



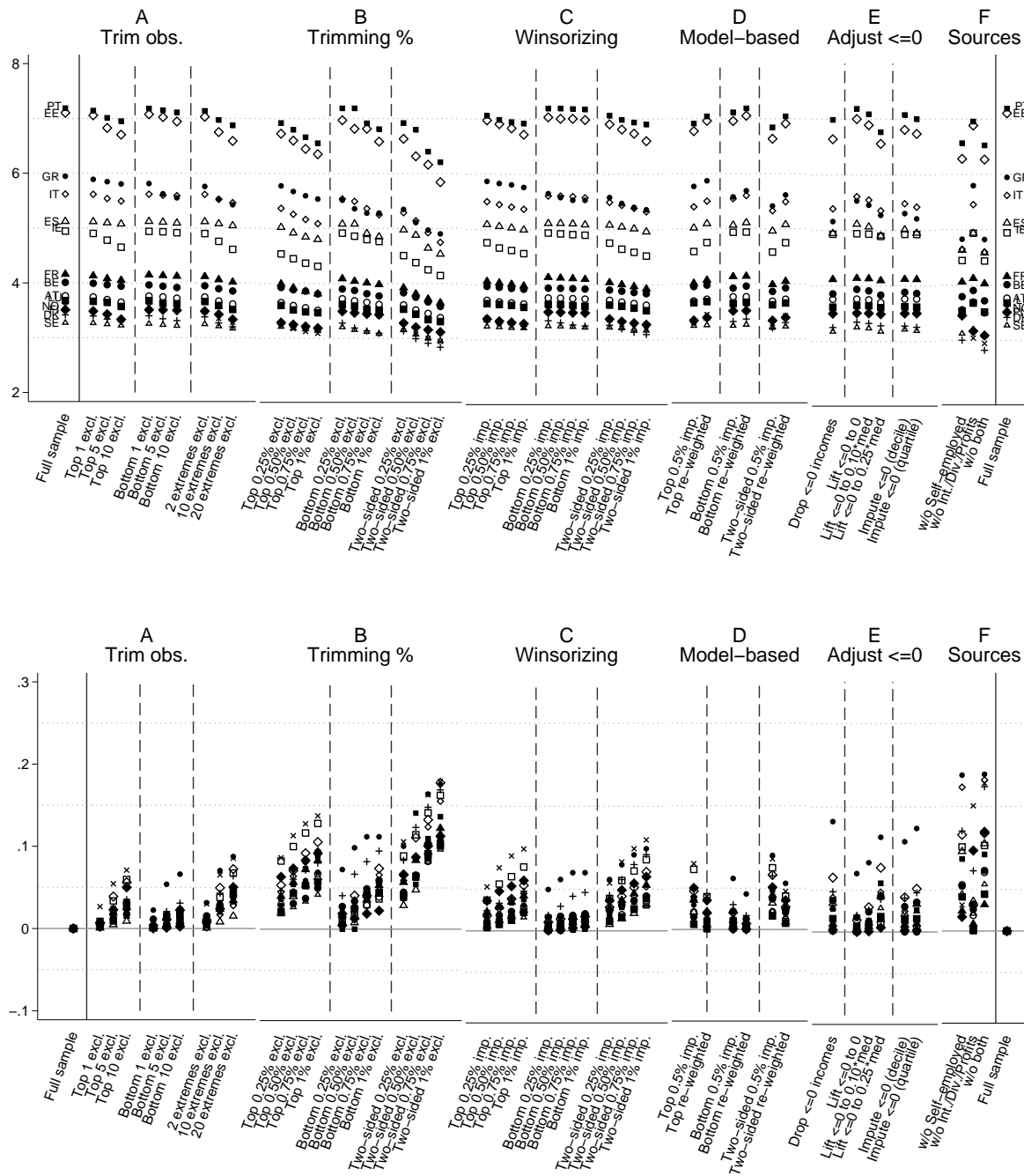
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.4: Inequality: Estimates of the P90/P10 quantile ratio in 14 countries under alternative treatment of extreme incomes



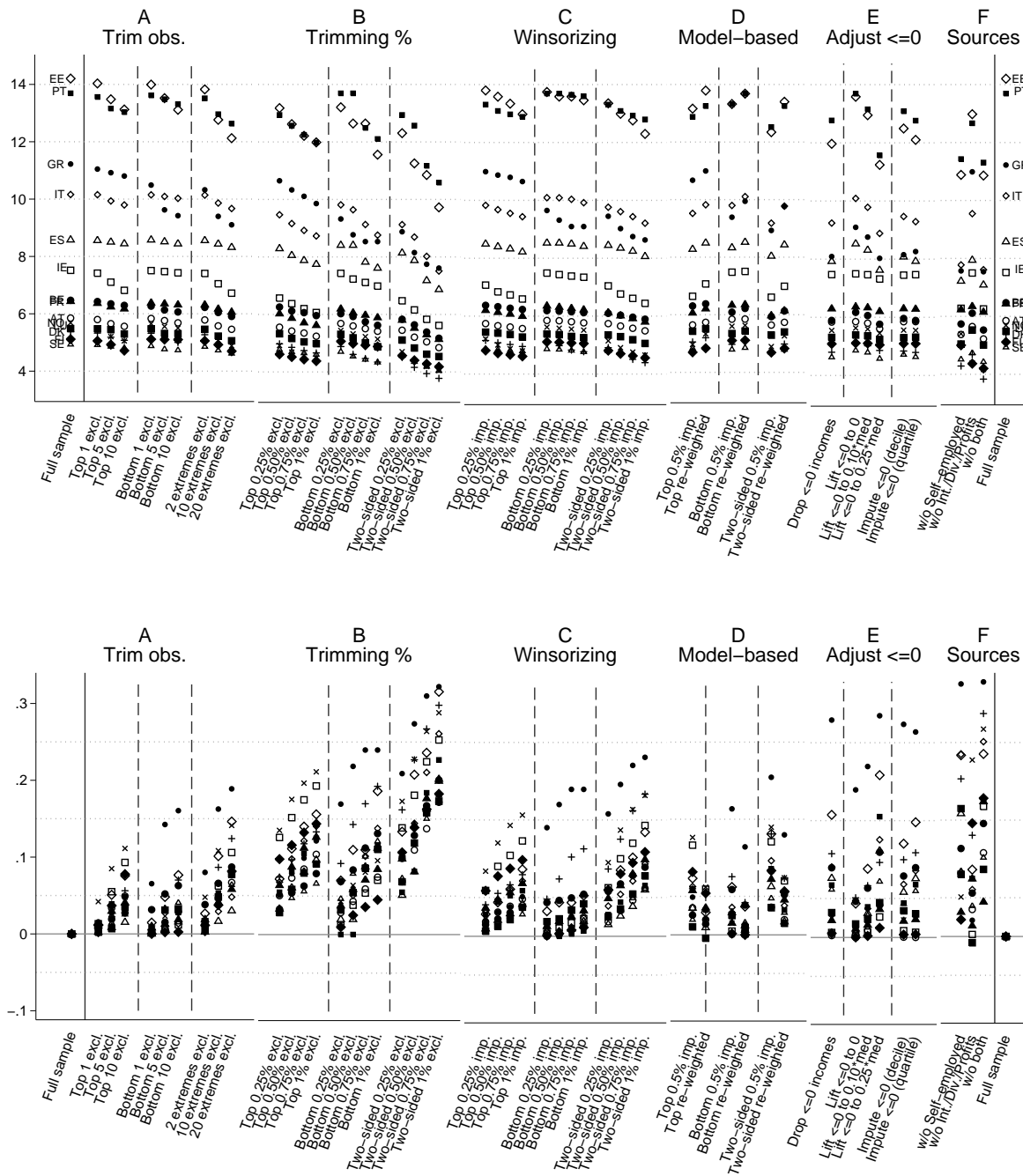
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.5: Inequality: Estimates of the S80/S20 income share ratio in 14 countries under alternative treatment of extreme incomes



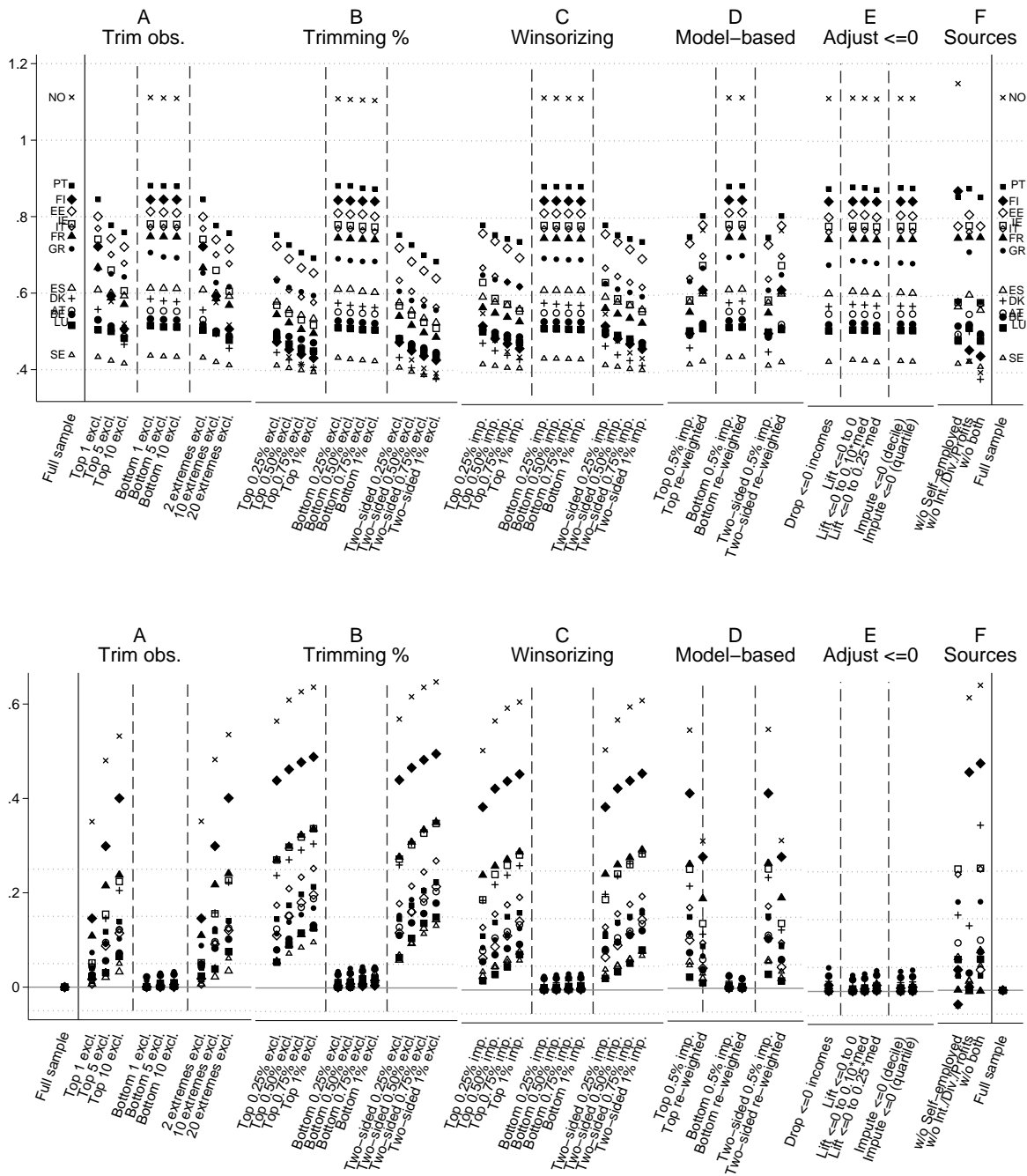
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.6: Inequality: Estimates of the S90/S10 income share ratio in 14 countries under alternative treatment of extreme incomes



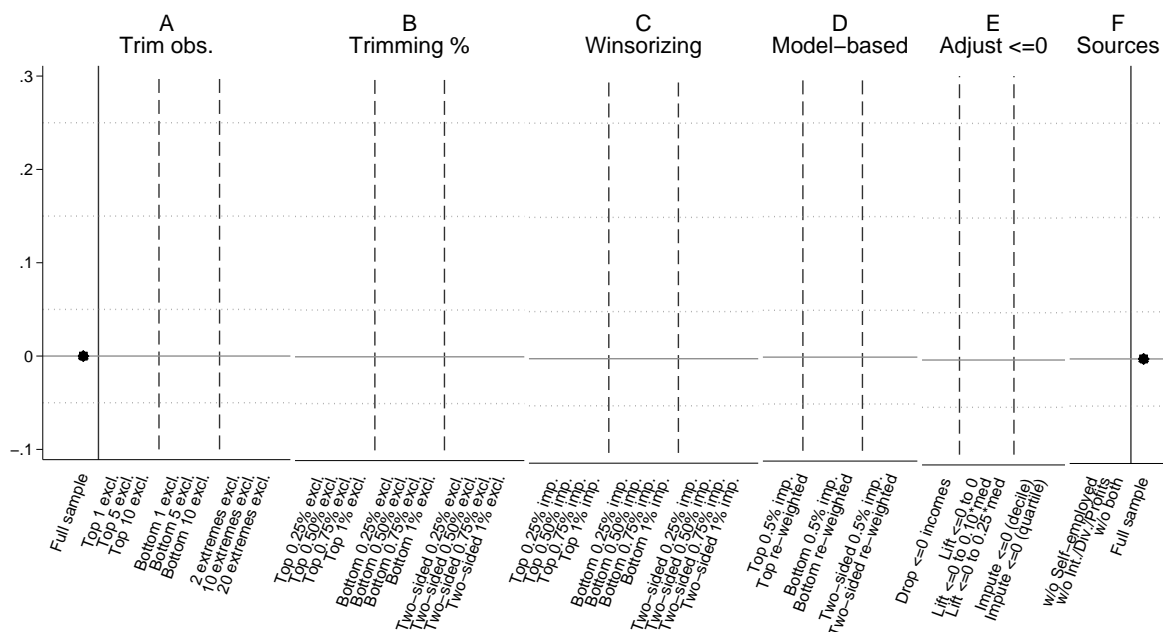
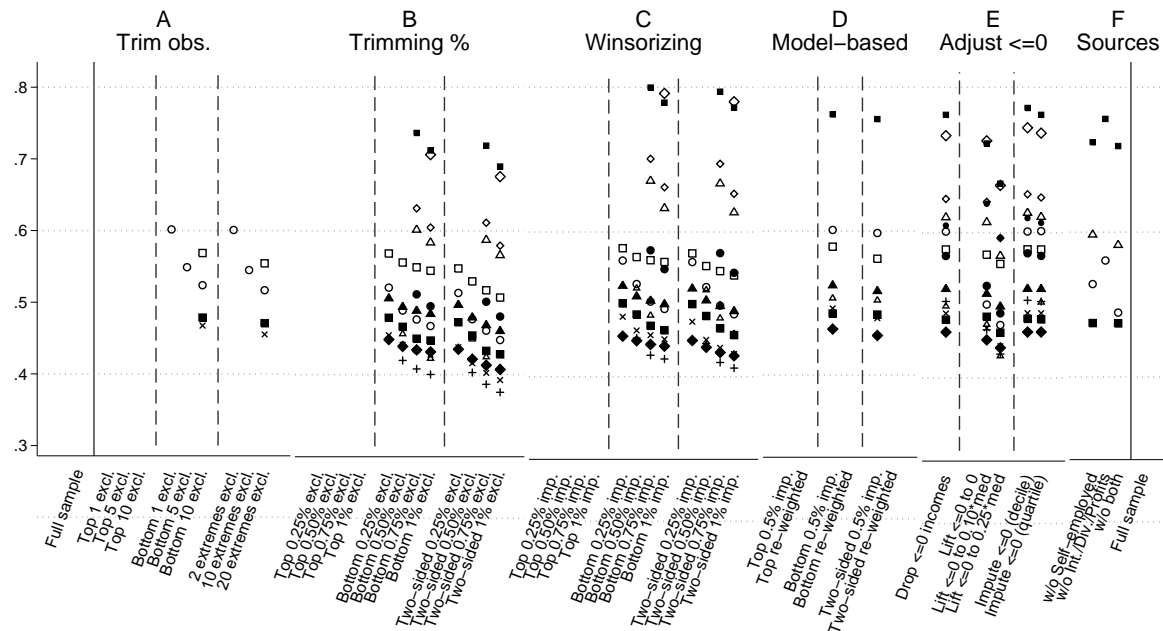
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.7: Inequality: Estimates of the Coefficient of Variation in 14 countries under alternative treatment of extreme incomes



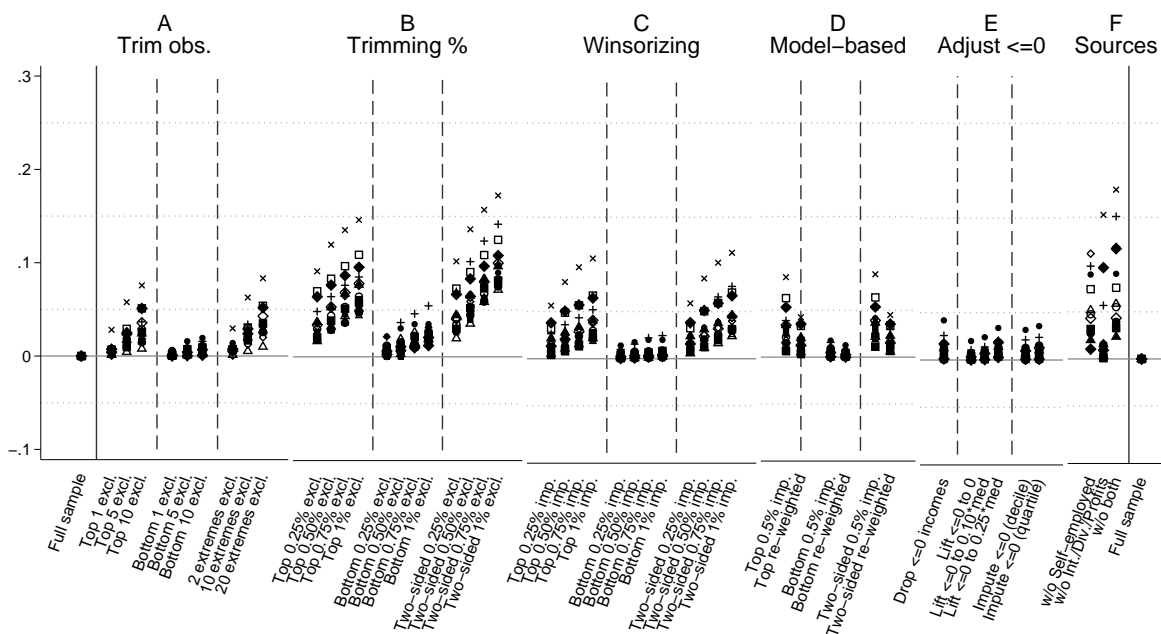
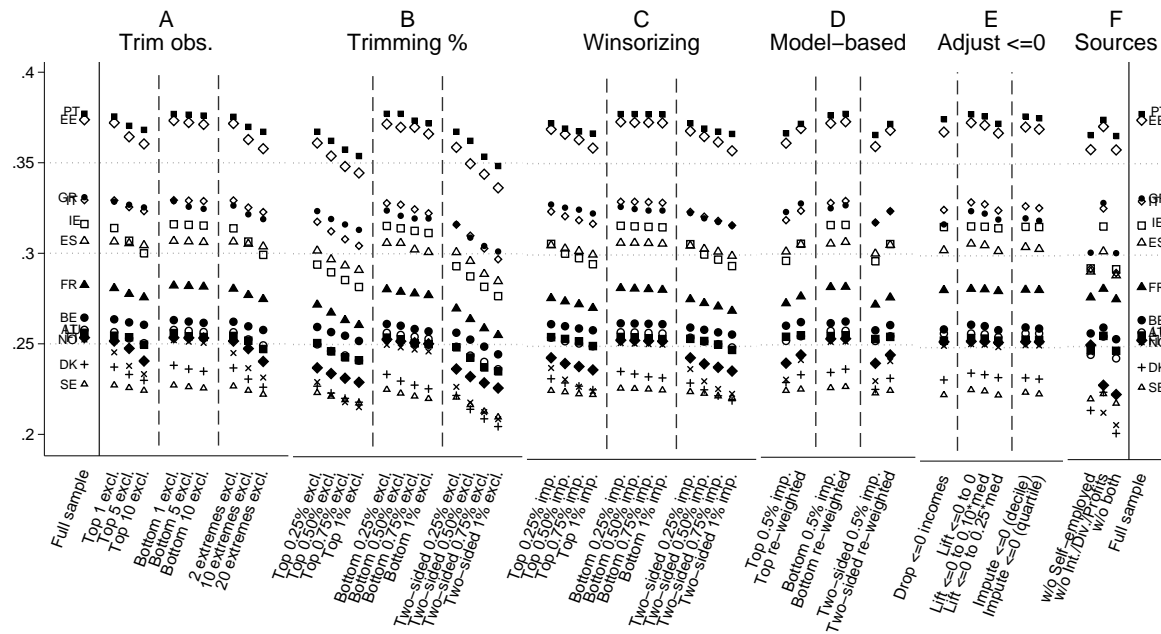
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.8: Inequality: Estimates of the standard deviation of log-incomes in 14 countries under alternative treatment of extreme incomes



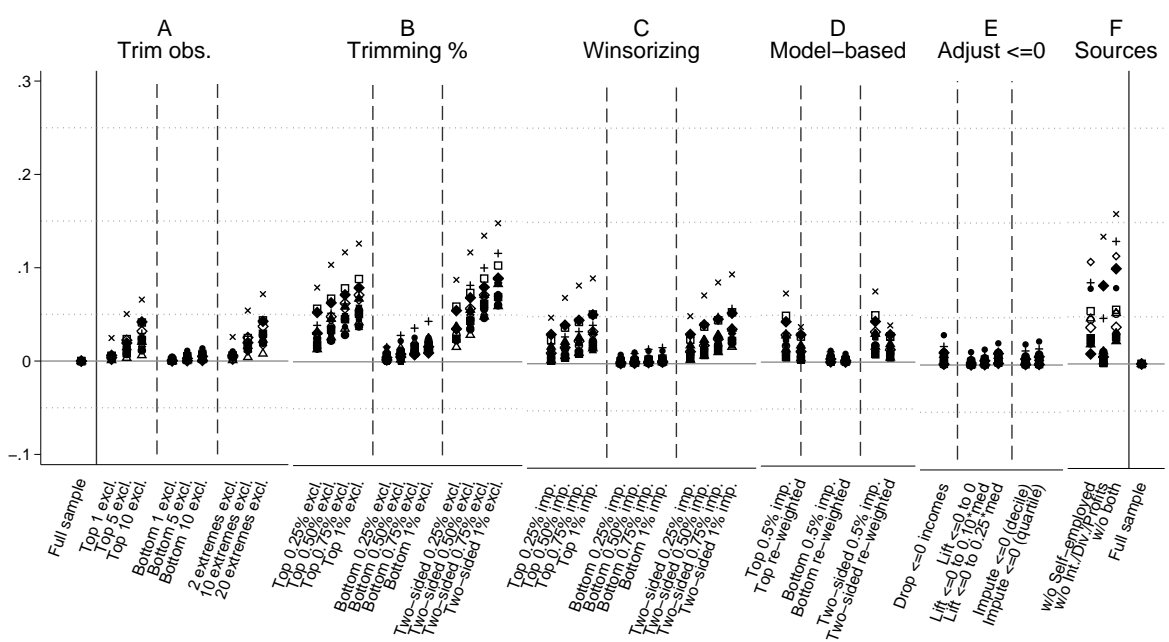
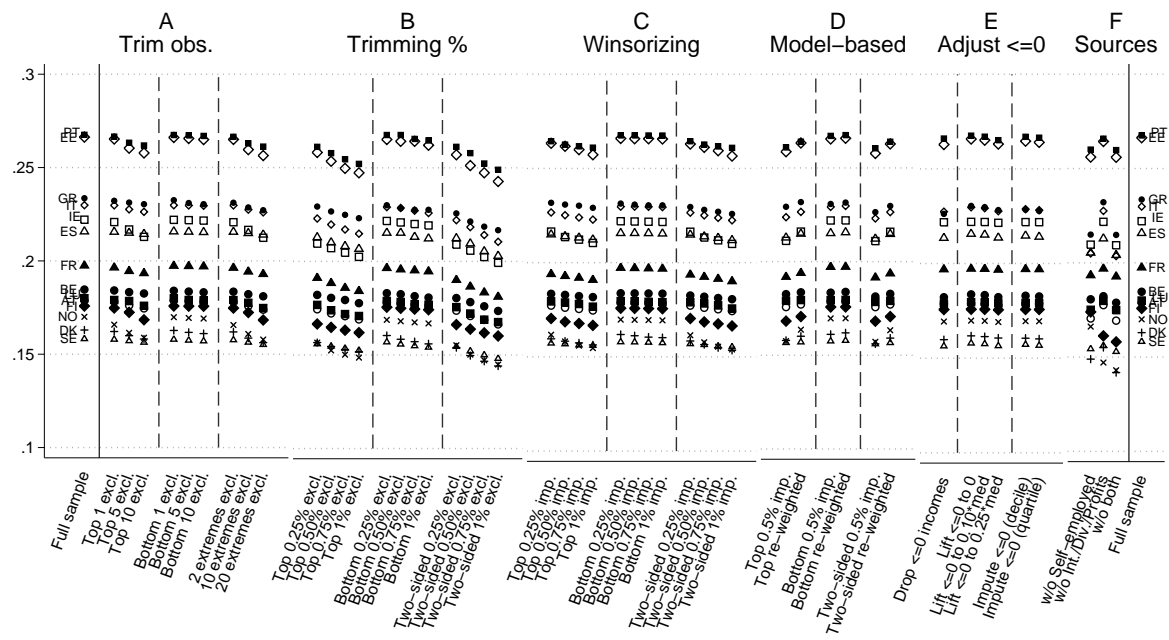
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.9: Inequality: Estimates of the Gini coefficient in 14 countries under alternative treatment of extreme incomes



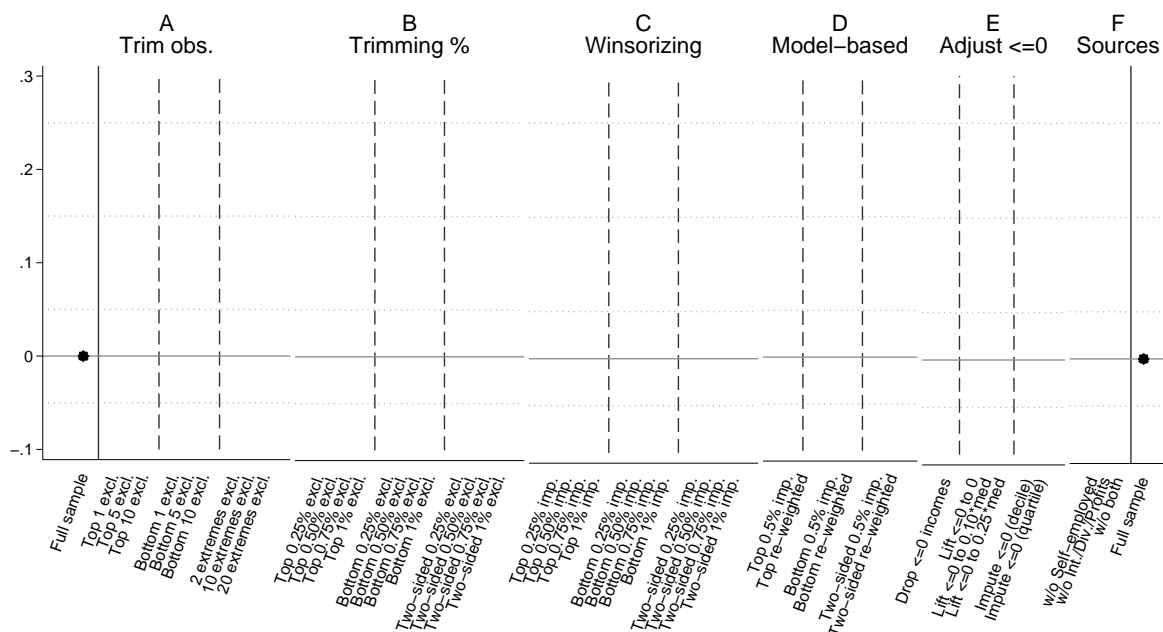
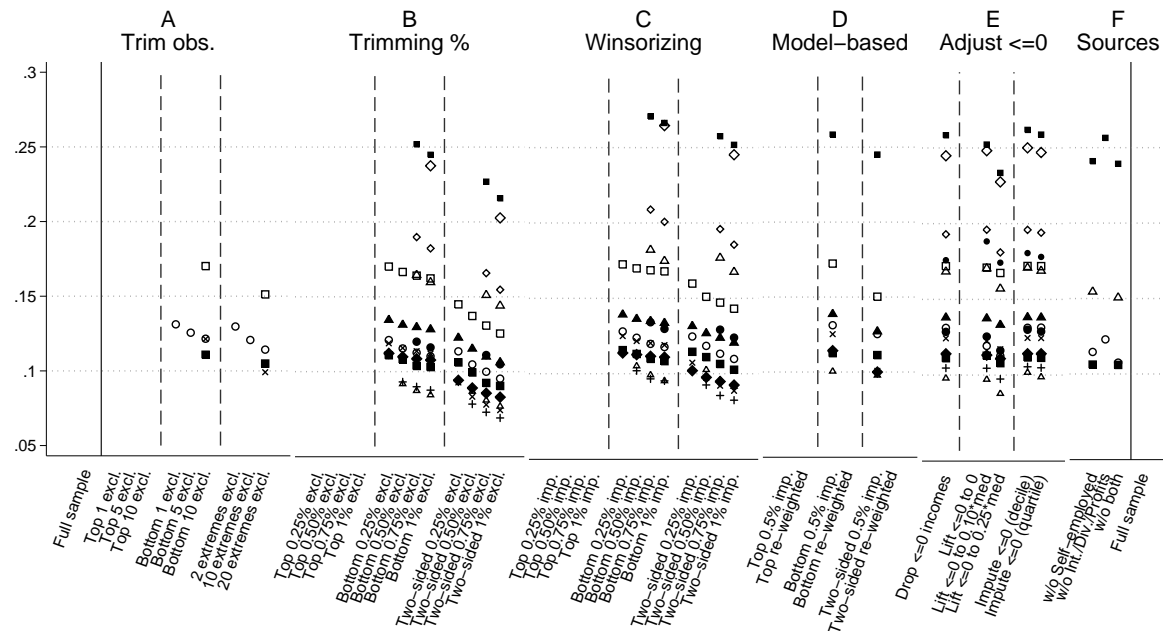
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.10: Inequality: Estimates of the Relative Mean Deviation (Schutz coefficient) in 14 countries under alternative treatment of extreme incomes



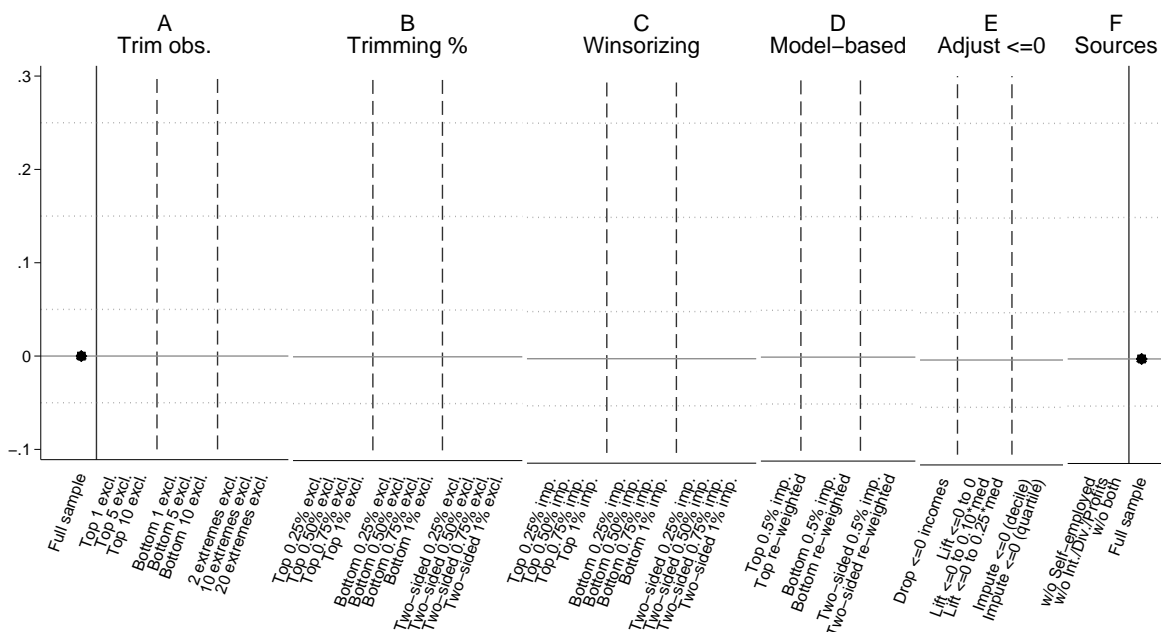
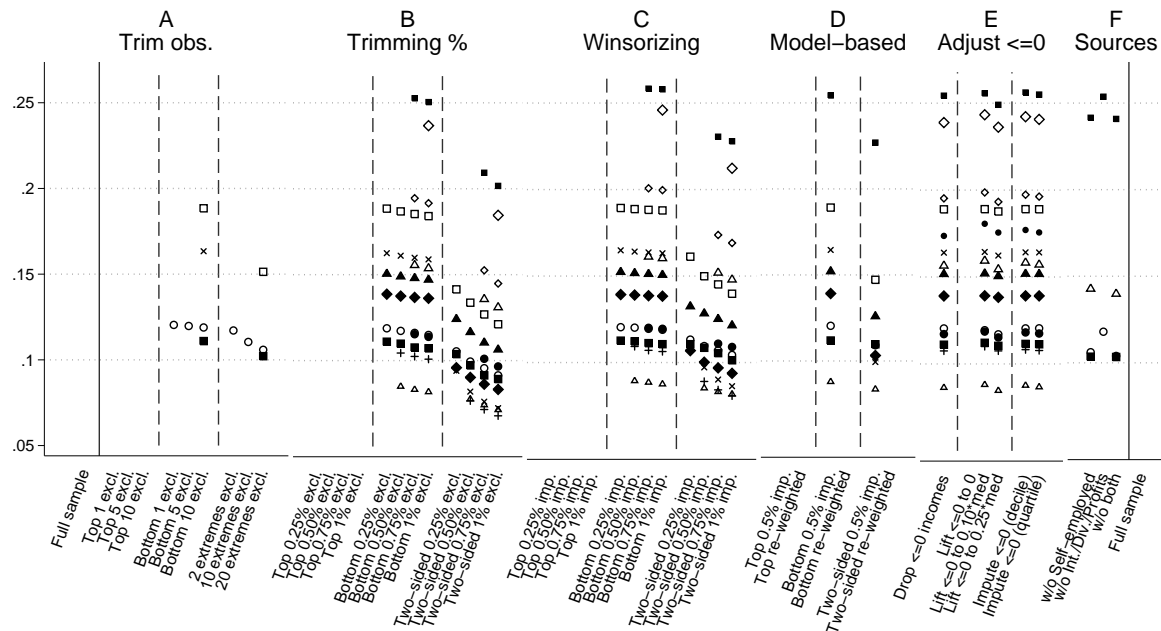
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.11: Inequality: Estimates of the Generalized Entropy measure with parameter 0 in 14 countries under alternative treatment of extreme incomes



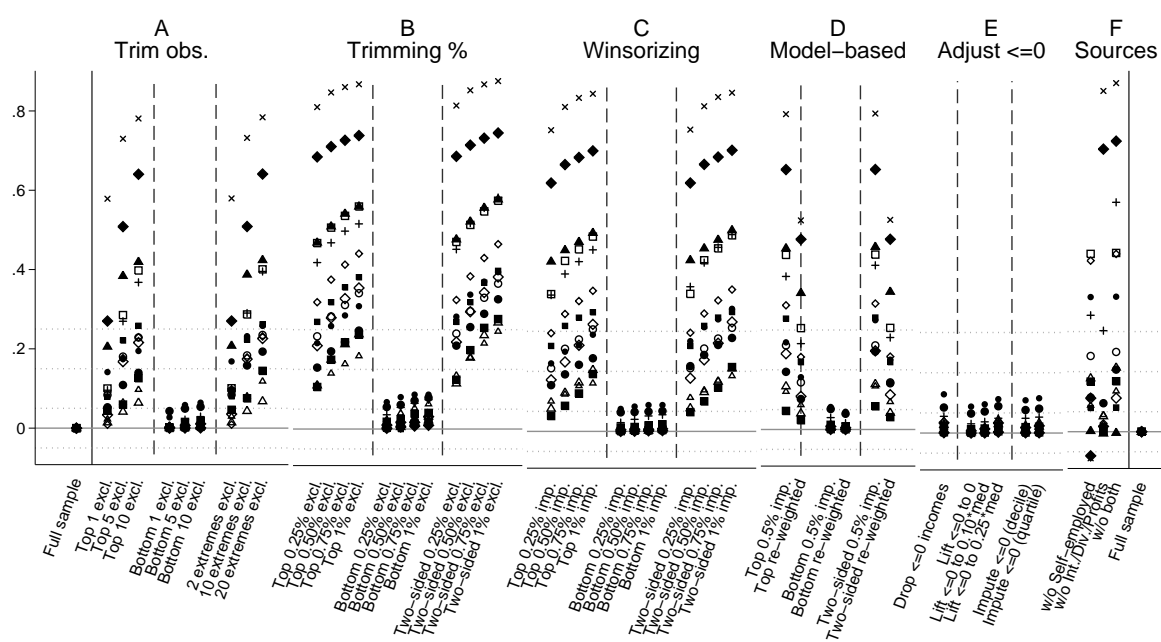
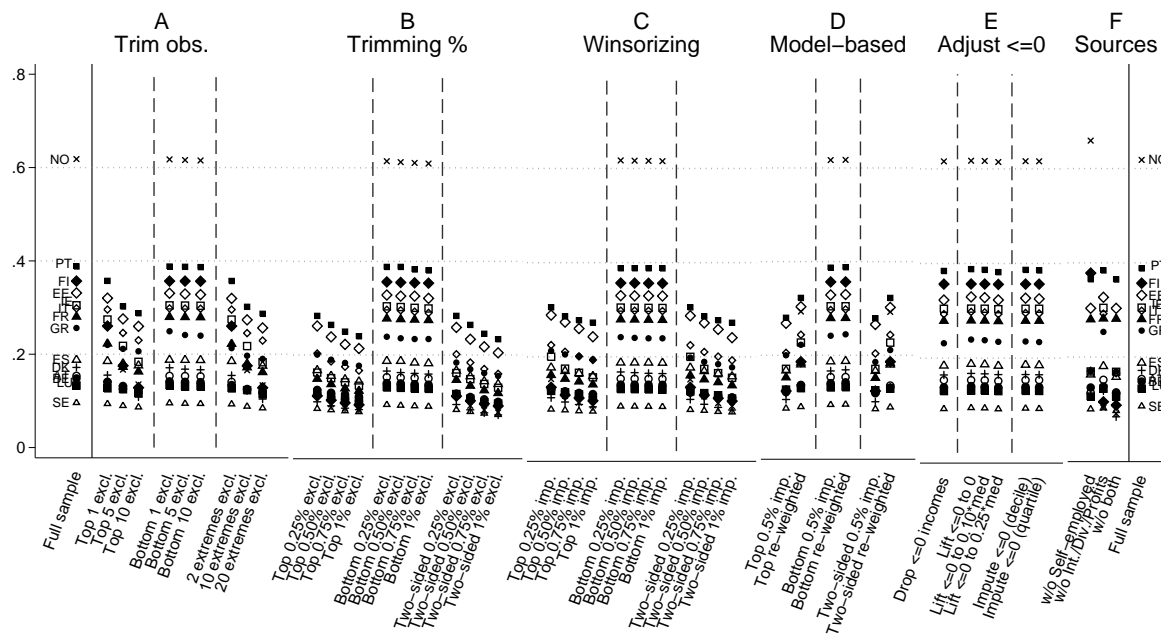
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.12: Inequality: Estimates of the Generalized Entropy measure with parameter 1 in 14 countries under alternative treatment of extreme incomes



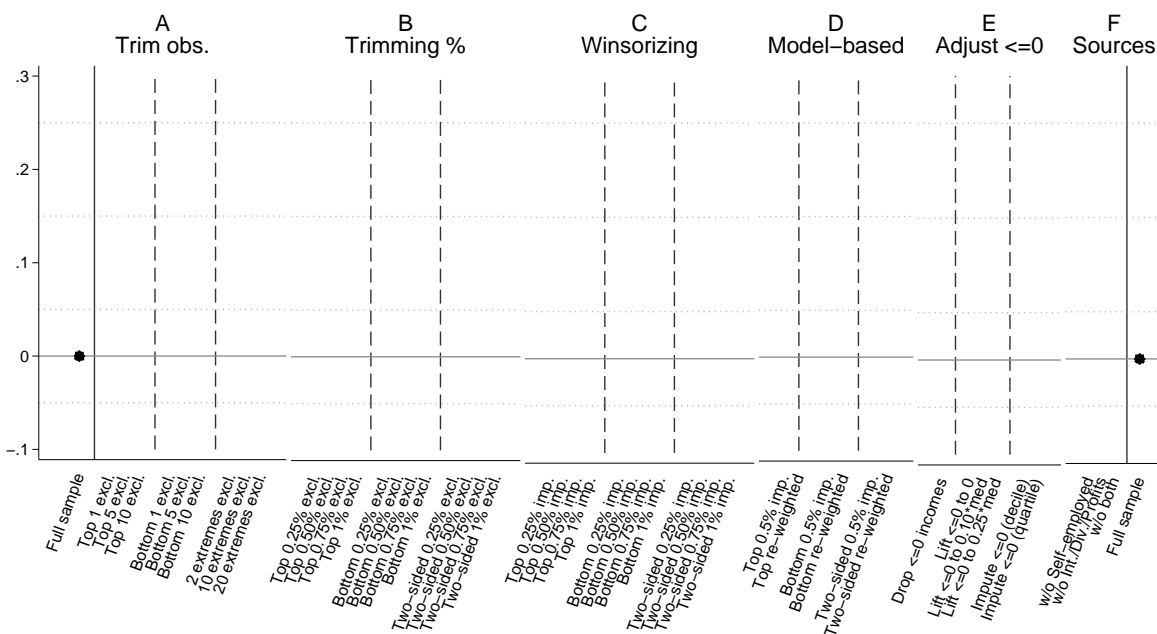
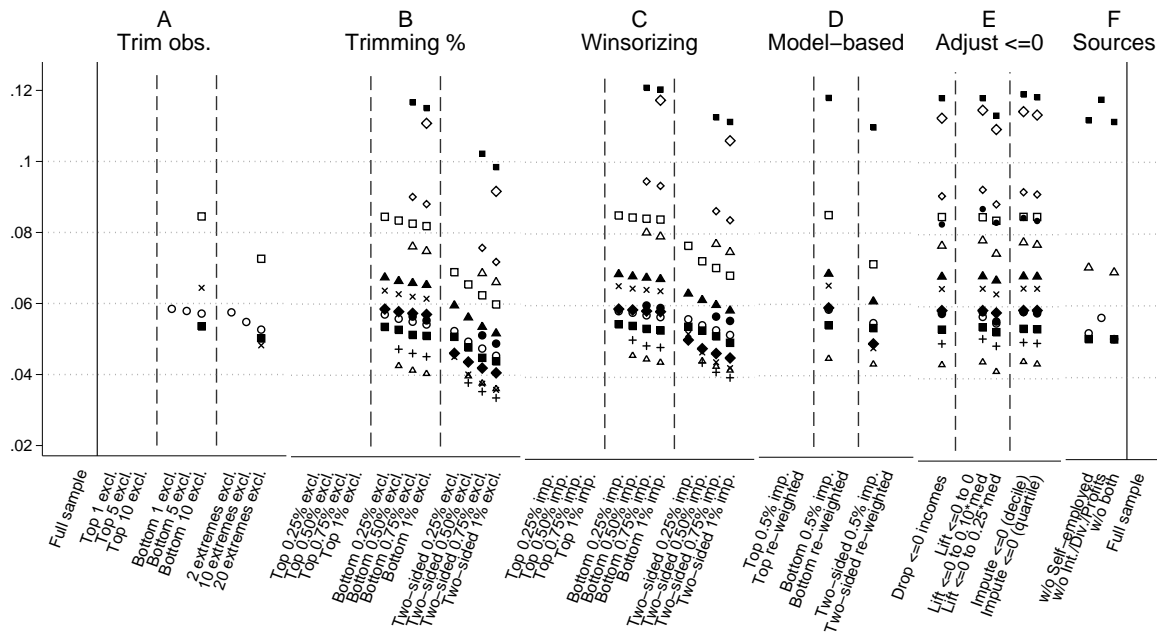
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.13: Inequality: Estimates of the Generalized Entropy measure with parameter 2 in 14 countries under alternative treatment of extreme incomes



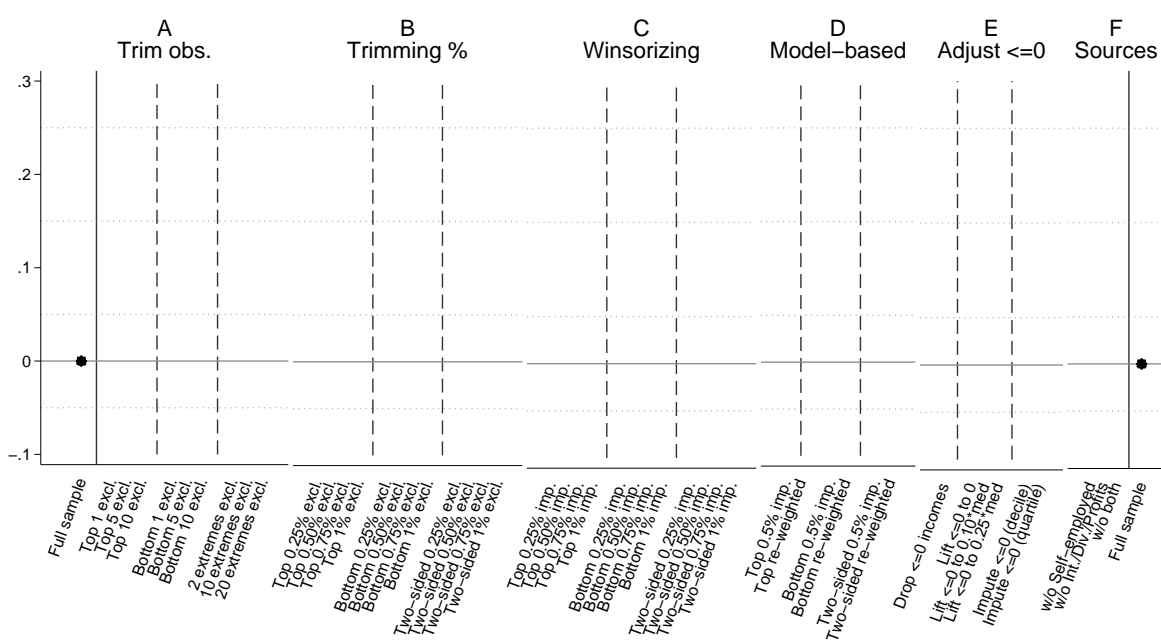
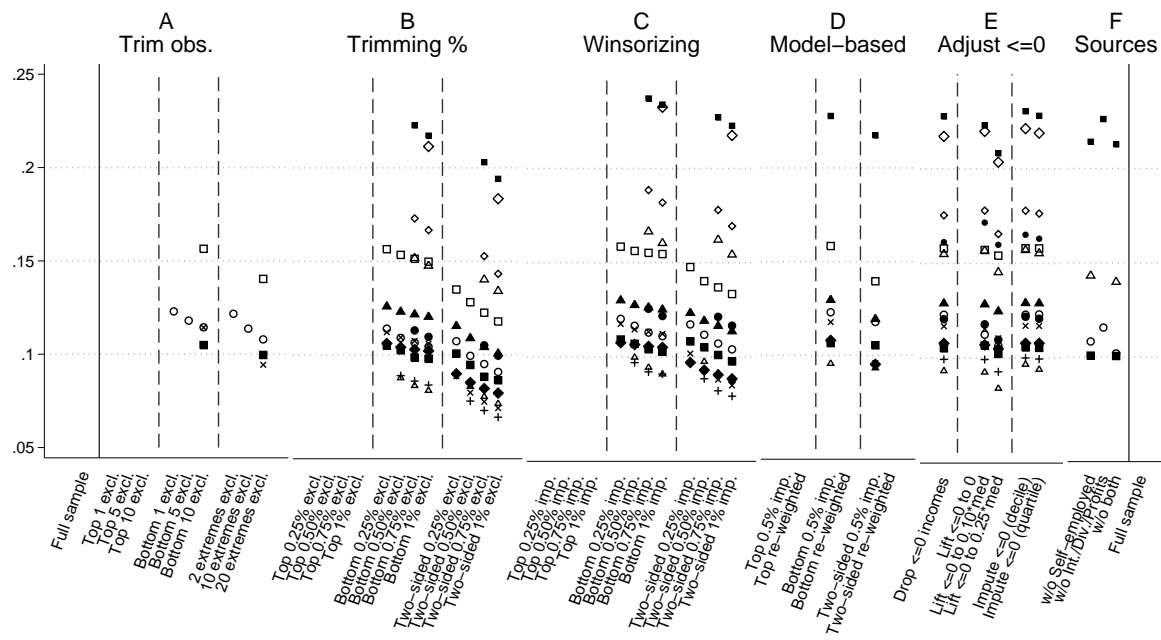
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.14: Inequality: Estimates of the Atkinson measure with parameter 0.5 in 14 countries under alternative treatment of extreme incomes



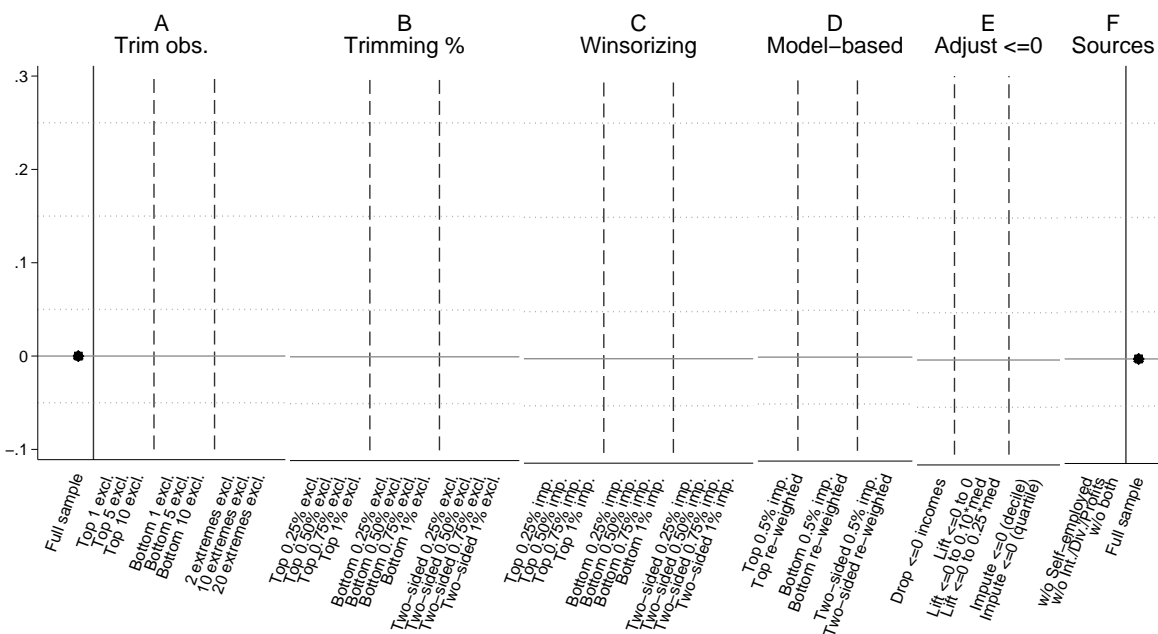
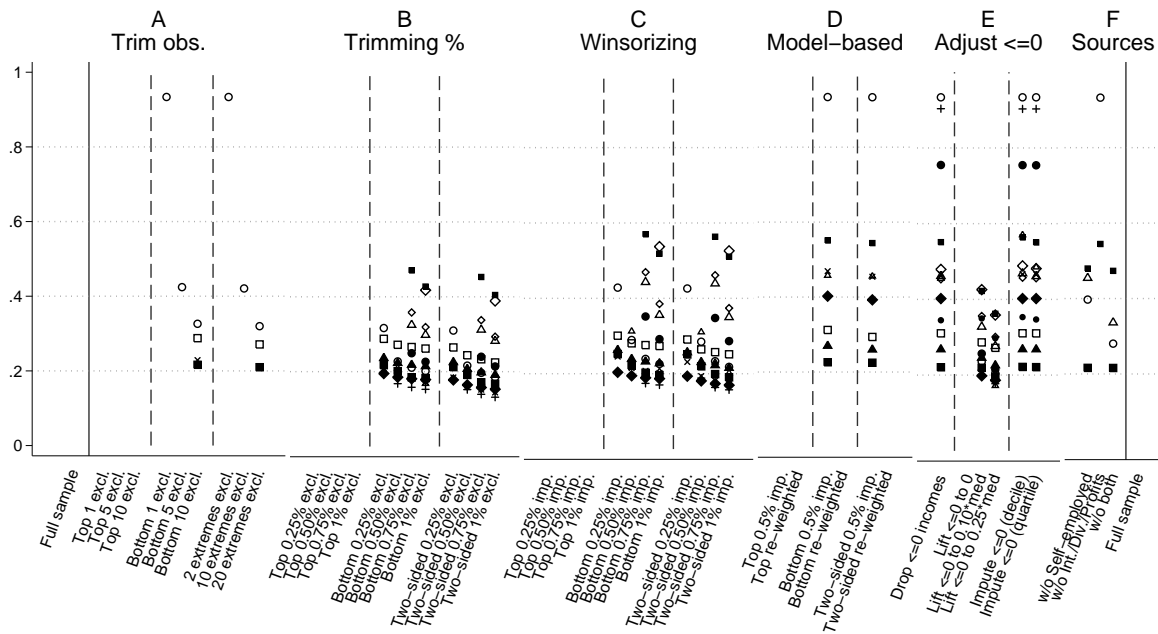
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.15: Inequality: Estimates of the Atkinson measure with parameter 1 in 14 countries under alternative treatment of extreme incomes



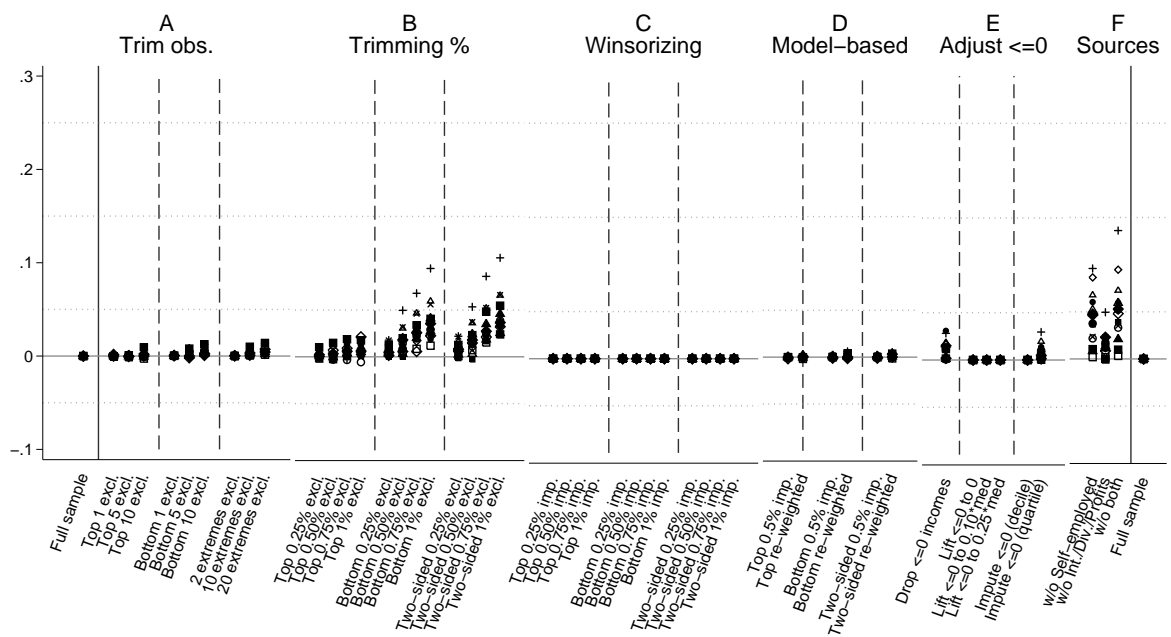
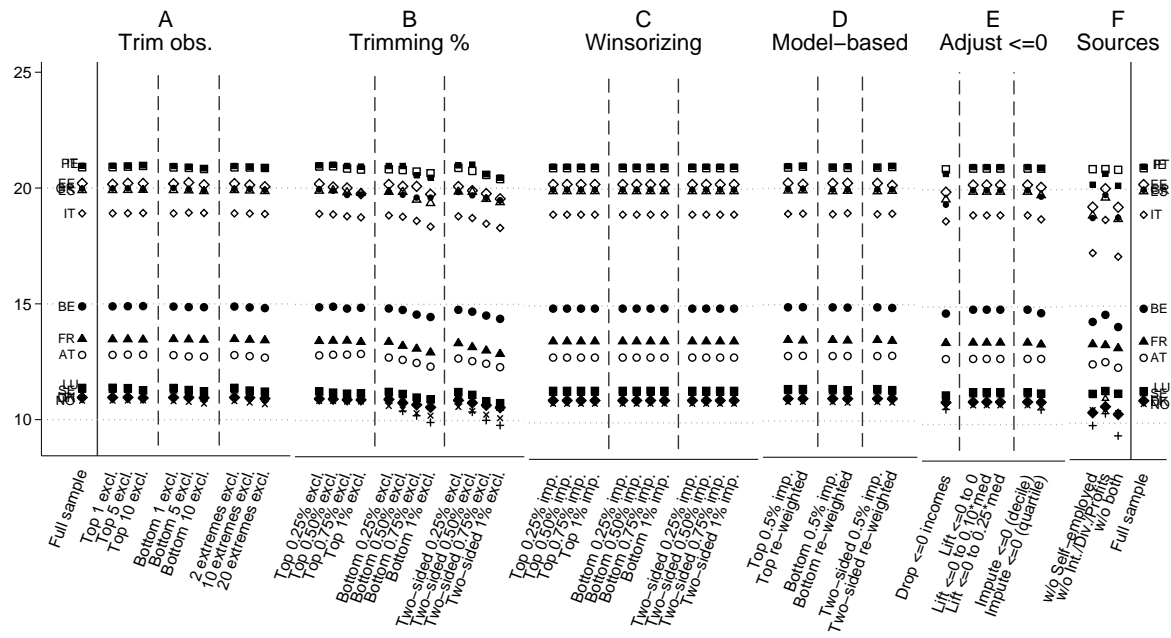
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.16: Inequality: Estimates of the Atkinson measure with parameter 2 in 14 countries under alternative treatment of extreme incomes



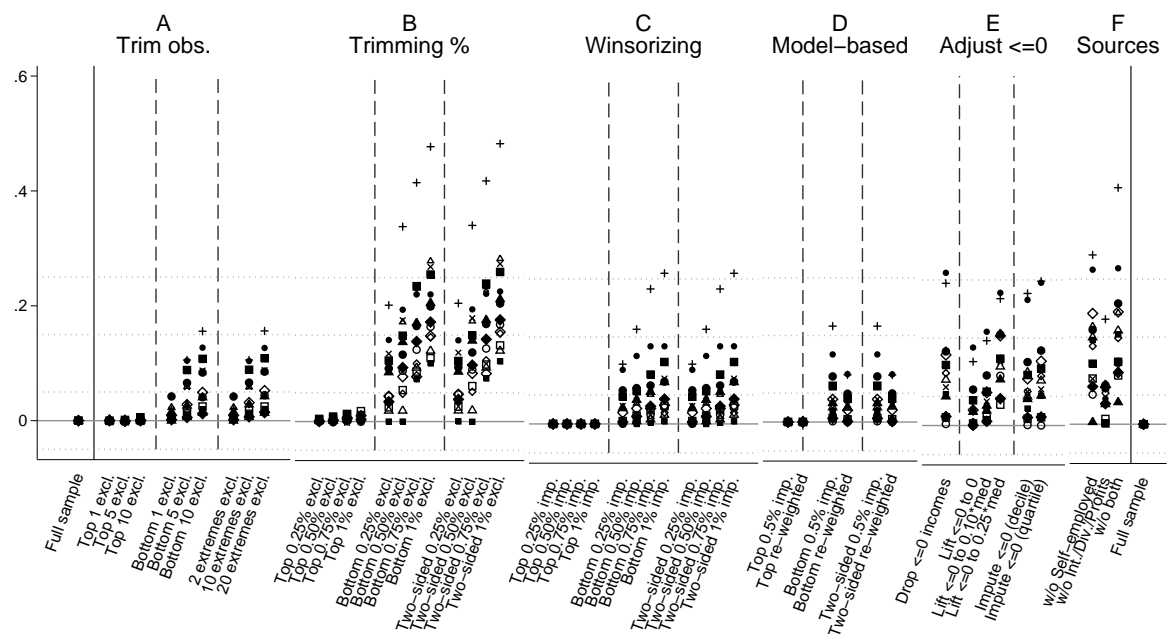
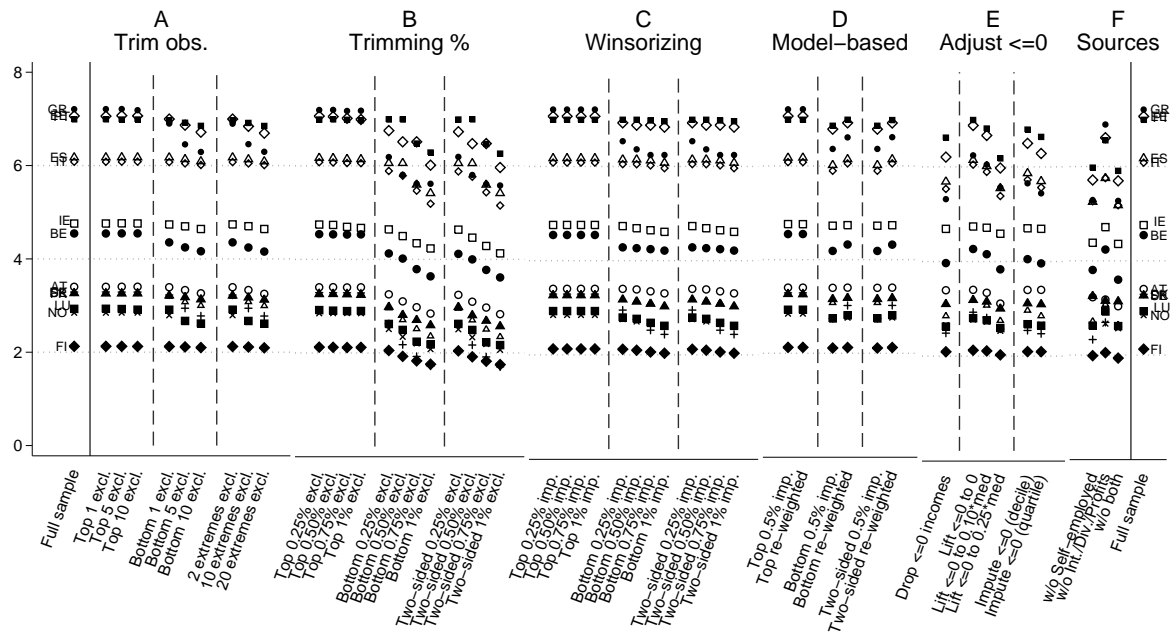
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.17: Poverty: Estimates of the headcount ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



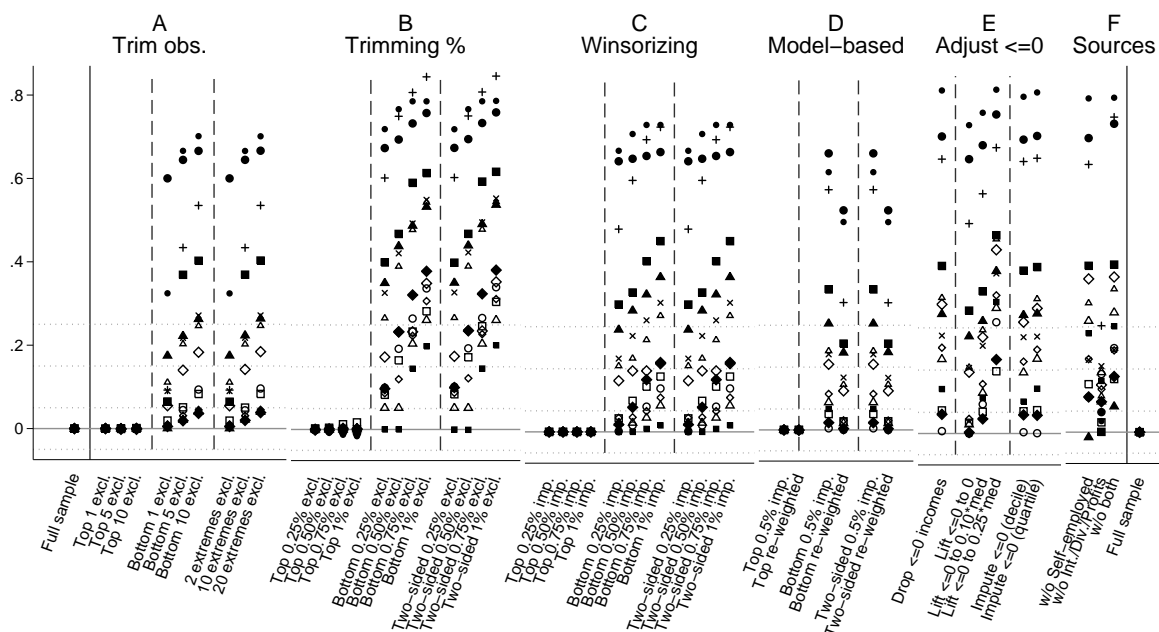
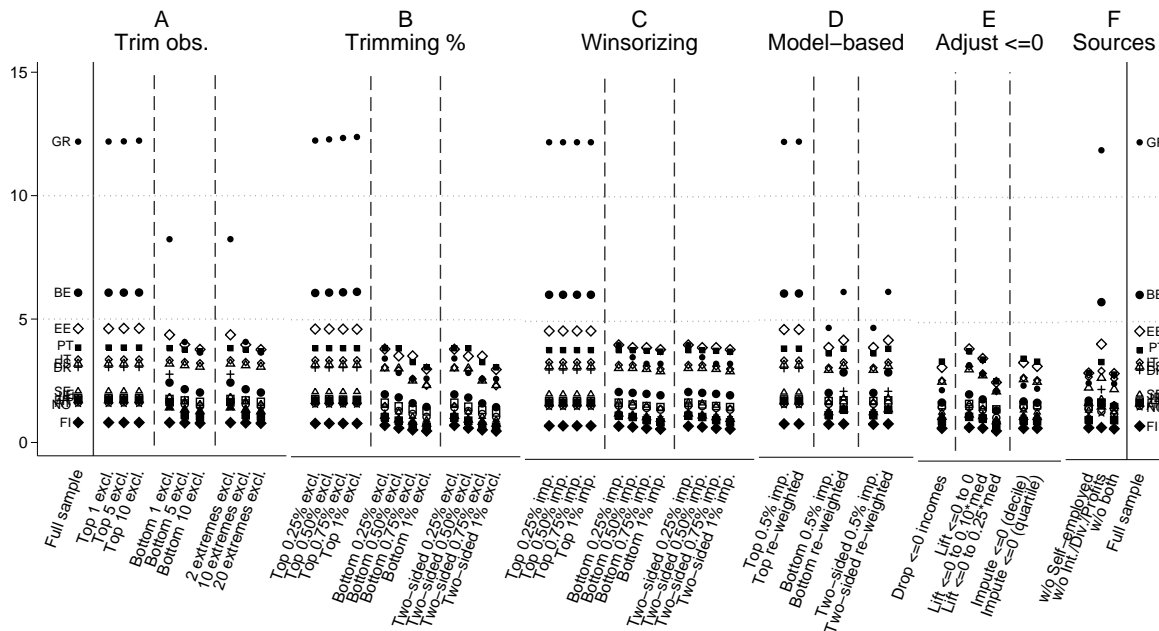
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.18: Poverty: Estimates of the average poverty gap ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



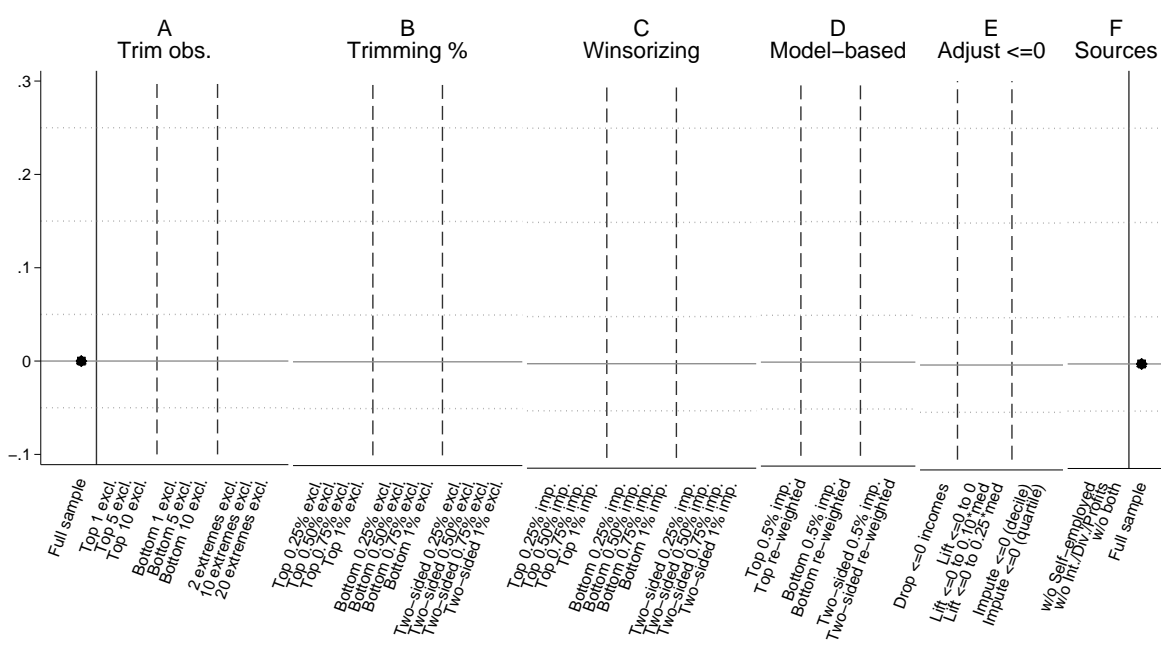
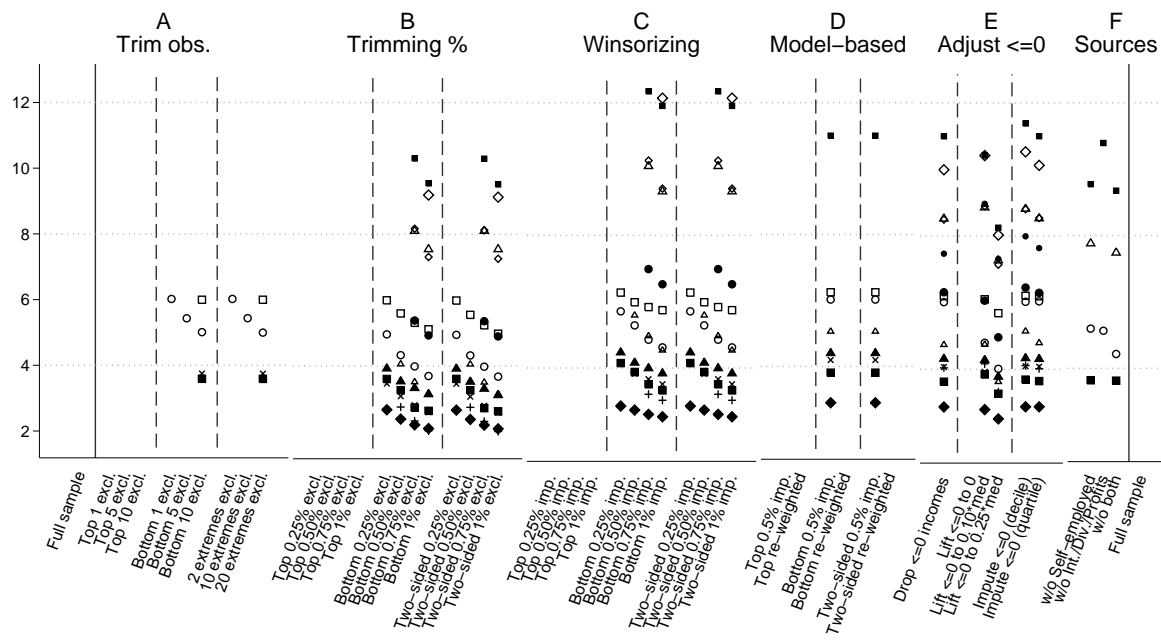
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.19: Poverty: Estimates of the average squared poverty gap ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



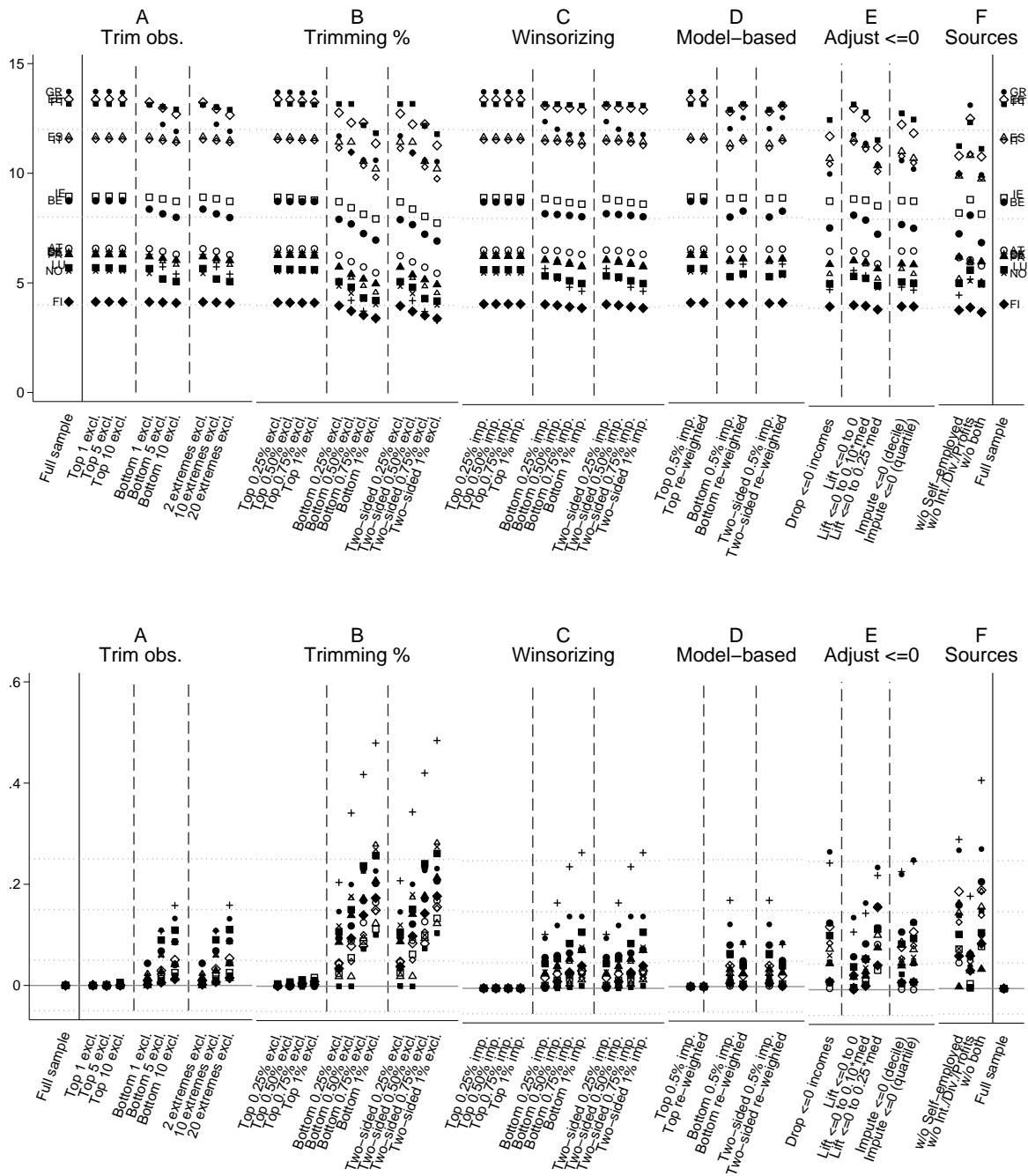
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.20: Poverty: Estimates of the Watts index based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



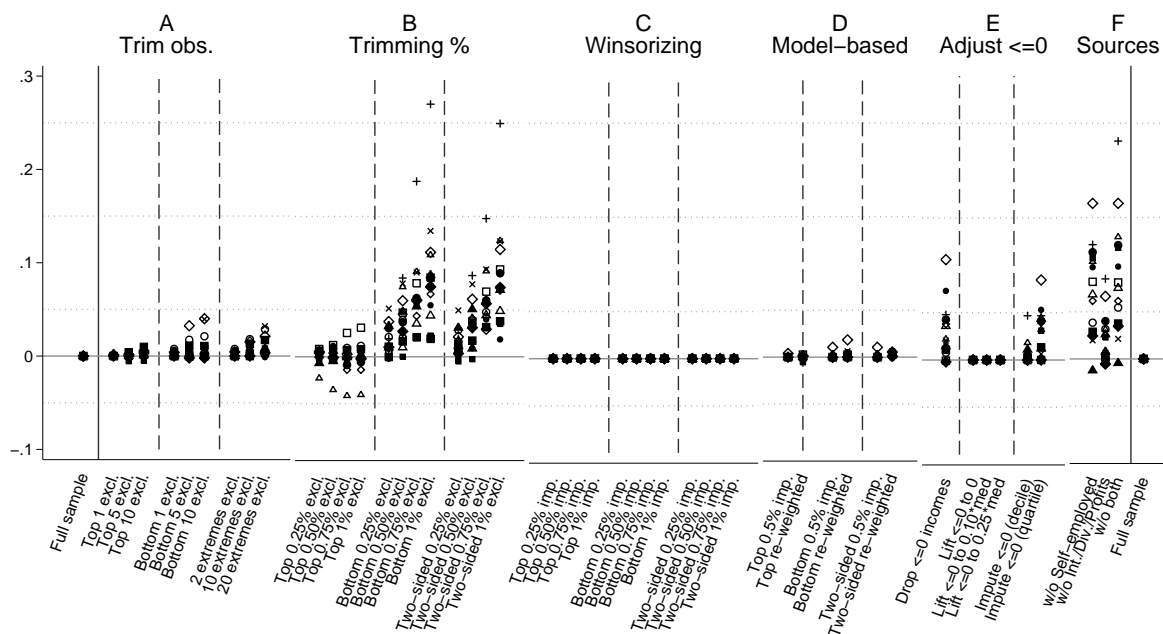
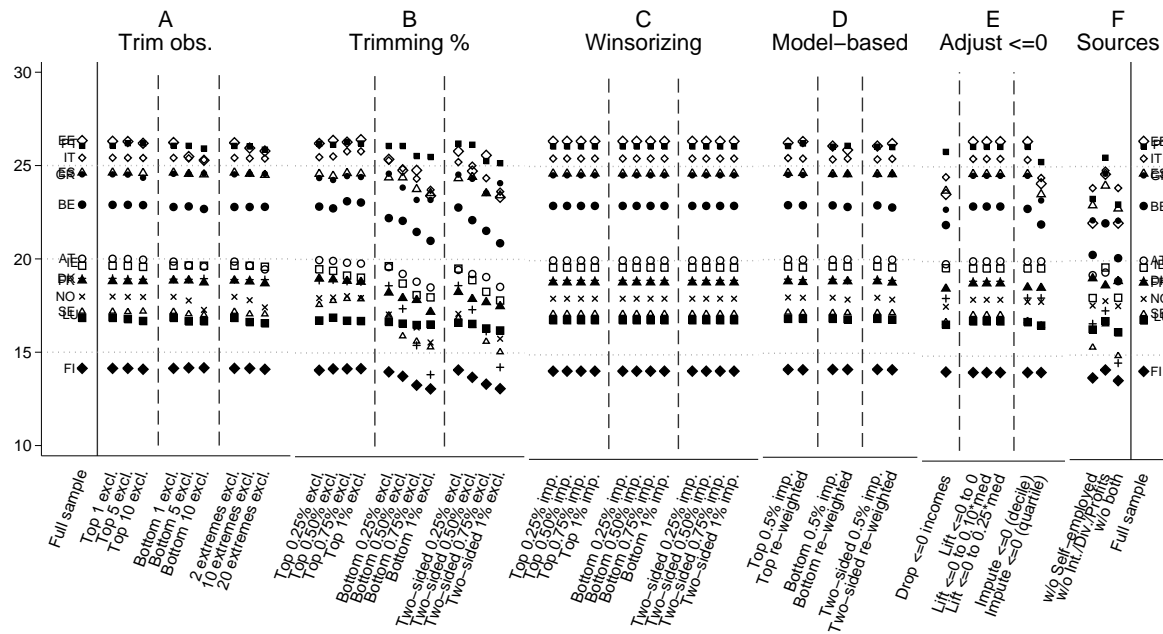
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.21: Poverty: Estimates of the Sen-Shorrocks-Thon index based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



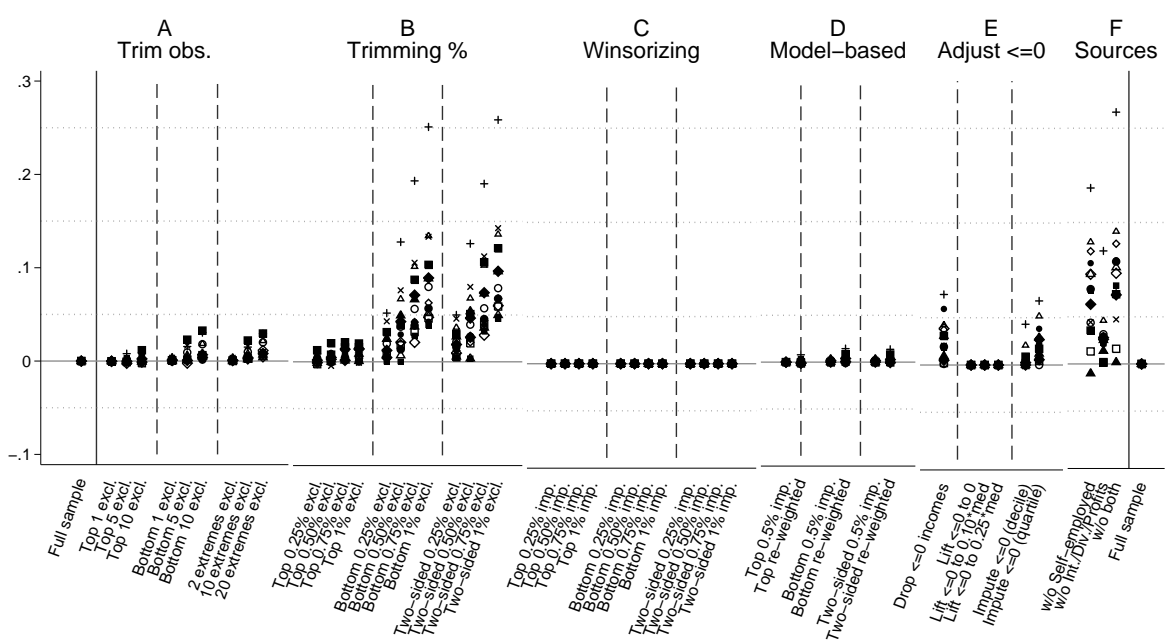
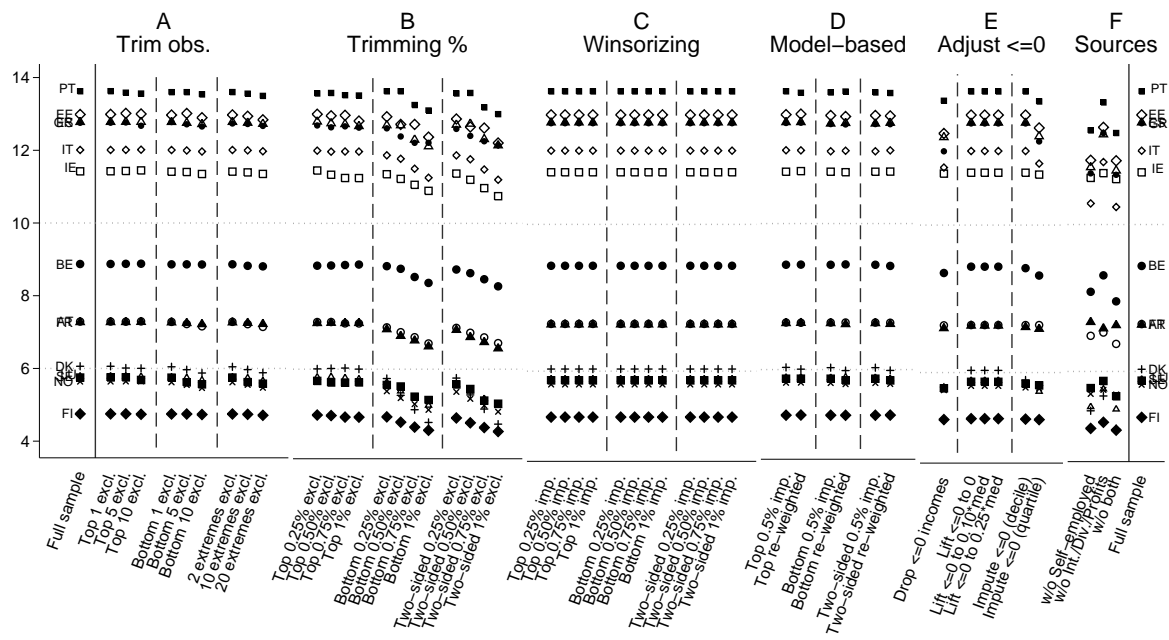
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.22: Poverty: Estimates of the median poverty gap ratio among the poor based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



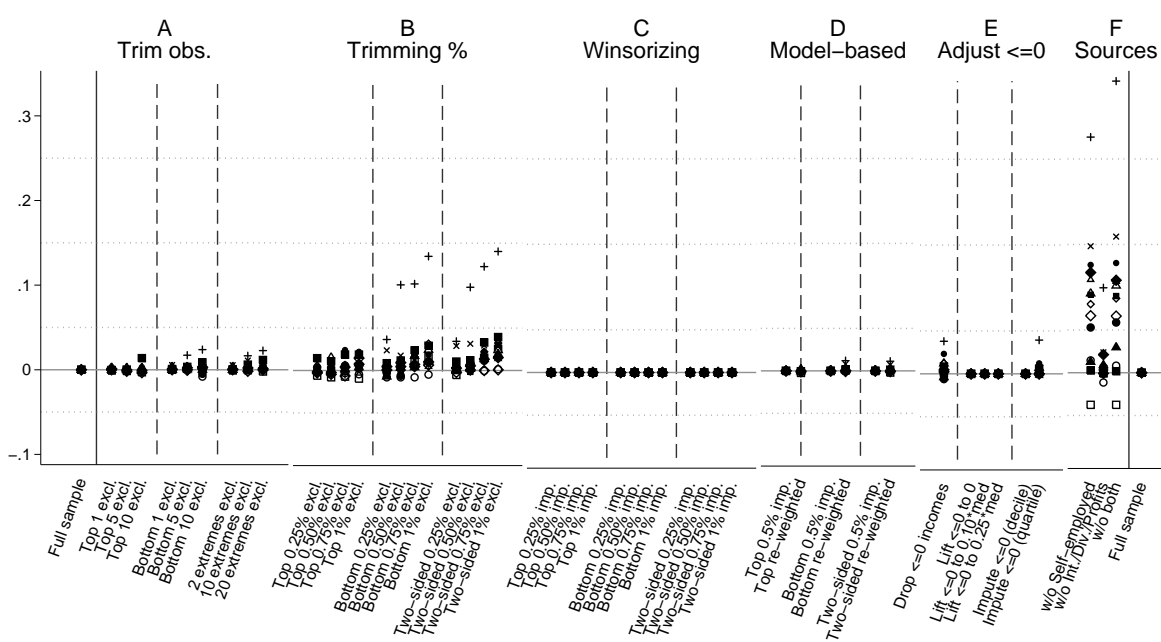
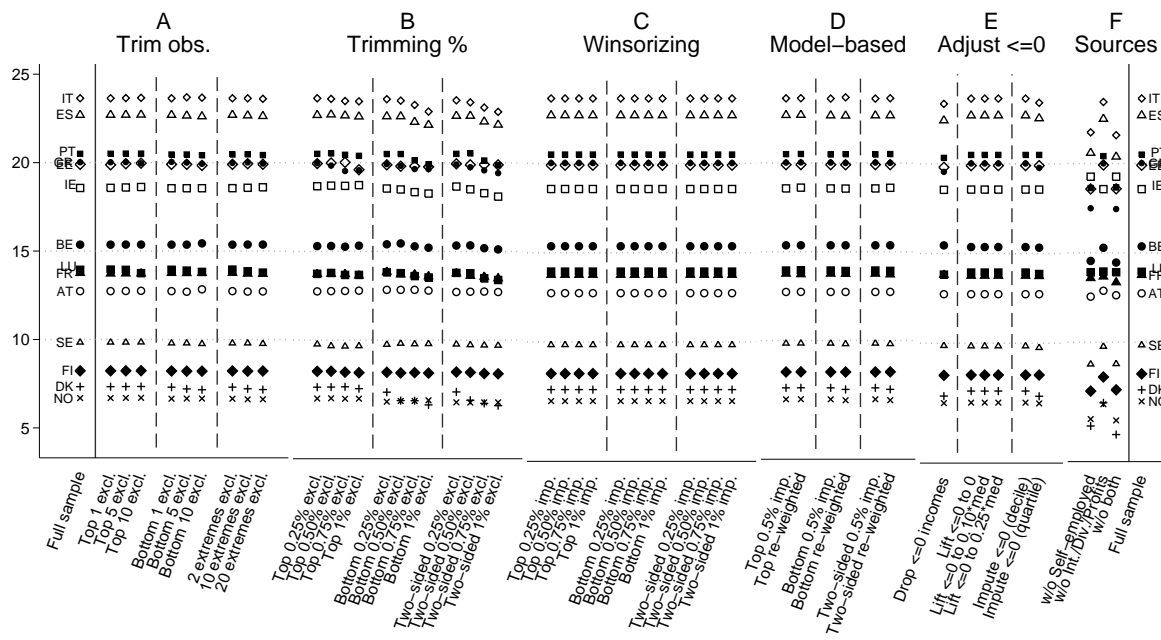
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.23: Poverty: Estimates of the headcount ratio based on a line at 50% of median income in 14 countries under alternative treatment of extreme incomes



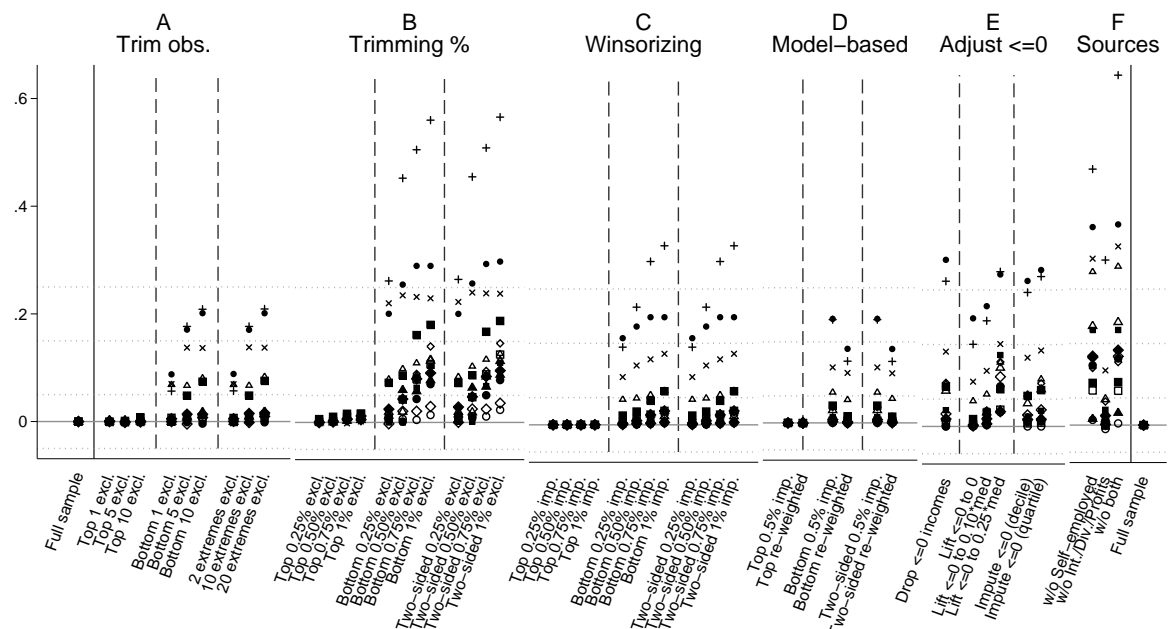
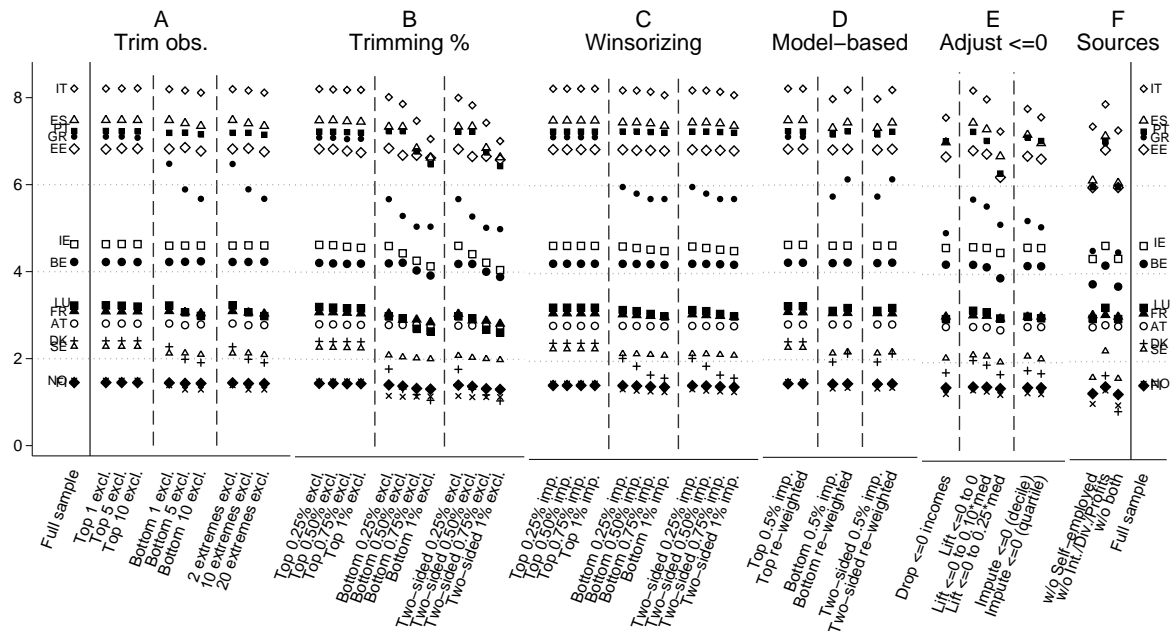
Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.24: Poverty: Estimates of the headcount ratio for households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



Bottom panel depicts indicator reduction relative to unadjusted estimate.

Figure 3.25: Poverty: Estimates of the average poverty gap ratio for households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes



Bottom panel depicts indicator reduction relative to unadjusted estimate.



IRISS Working Papers

The IRISS Working Paper Series has been created in 1999 to ensure a timely dissemination of the research outcome from the IRISS-C/I programme. They are meant to stimulate discussion and feedback. The working papers are contributed both by CEPS/INSTEAD resident staff, research associates and visiting researchers.

The fifteen most recent papers

Fusco A., 'Les mesures monétaires et directes de la pauvreté sont-elles substituables ? Investigations sur base de la courbe du ROC', IRISS WP 2006-13, December 2006.

Sirovatka T. & Valentova M., 'The Legitimacy of Redistribution: the Czech Republic in International Comparison', IRISS WP 2006-12, November 2006.

Voicu M., Voicu B. & Strapcova K., 'Housework and gender inequality across Europe', IRISS WP 2006-11, November 2006.

Makdissi P. & Mussard S., 'Between-Group Transfers and Poverty-Reducing Tax Reforms', IRISS WP 2006-10, October 2006.

Makdissi P. & Mussard S., 'Decomposition of s-Concentration Curves', IRISS WP 2006-09, October 2006.

Mussard S. & Philippe B., 'Okun's Law, Creation of Money and the Decomposition of the Rate of Unemployment', IRISS WP 2006-08, October 2006.

Williams D., 'The Economic Returns to Multiple Language Usage in Western Europe', IRISS WP 2006-07, October 2006.

Chiappero-Martinetti E. & Civardi M., 'Measuring poverty within and between population subgroups', IRISS WP 2006-06, October 2006.

Mussard S. & Philippe B., 'De la relation épargne/investissement à l'évolution du taux de chômage', IRISS WP 2006-05, September 2006.

Tanova C., 'Using Job Embeddedness Factors to Explain Voluntary Turnover in Five European Countries', IRISS WP 2006-04, July 2006.

Van Kerm P., 'Comparisons of income mobility profiles', IRISS WP 2006-03, July 2006.

Fusco A. & Dickes P., 'Rasch Model and Multidimensional Poverty Measurement', IRISS WP 2006-02, May 2006.

Verbelen B., 'Is Taking a Pill a Day Good for Health Expenditures? Evidence from a Cross Section Time Series Analysis of 19 OECD Countries from 1970 2000', IRISS WP 2006-01, May 2006.

Kwiatkowska-Ciotucha D. & Zaluska U., 'Job Satisfaction as an Assessment Criterion of Labor Market Policy Efficiency. Lesson for Poland from International Experience', IRISS WP 2005-04, May 2005.

Heffernan C., 'Gender, Cohabitation and Martial Dissolution: Are changes in Irish family composition typical of European countries?', IRISS WP 2005-03, March 2005.

Electronic versions

Electronic versions of all IRISS Working Papers are available for download at
<http://www.ceps.lu/iriss/wps.cfm>



IRISS-C/I is a visiting researchers programme at CEPS/INSTEAD, a socio-economic policy and research centre based in Luxembourg. It finances and organises short visits of researchers willing to undertake empirical research in economics and other social sciences using the archive of micro-data available at the Centre.

What is offered?

In 1998, CEPS/INSTEAD has been identified by the European Commission as one of the few *Large Scale Facilities* in the social sciences, and, since then, offers researchers (both junior and senior) the opportunity to spend time carrying out their own research using the local research facilities. This programme is currently sponsored by the European Community's 6th Framework Programme. Grants cover travel expenses and on-site accommodation. The expected duration of visits is in the range of 2 to 12 weeks.

Topics

The major resource offered to visitors is access to a series of internationally comparable longitudinal surveys on living conditions at the household and individual level. The anonymised micro-data provide information on wages and income, health, education, employment and professional activities, accommodation, social relations,... Comparable micro-data are available for EU countries, Central European countries, as well as the USA. These data offer opportunities to carry out research in fields such as *survey and panel data methodology, income distribution and welfare, income and poverty dynamics, multi-dimensional indicators of poverty and deprivation, gender, ethnic and social inequality, unemployment and labour supply behaviour, education and training, social protection and redistributive policies, fertility and family structures, new information technologies in households and firms, ...*

Who may apply?

All individuals (doctoral students as well as experienced academics) conducting research in an institution within the EU-25 or an FP6 Associated State. IRISS-C/I can be meeting place for groups of researchers working on a joint project. We therefore encourage joint proposals by two or more researchers.

For more detailed information and application form, please consult our website: <http://www.ceps.lu/iriss> or contact us at

IRISS-C/I, CEPS/INSTEAD
BP 48, L-4501 Differdange, G.-D. Luxembourg
Tel: +352 585855 610; Fax: +352 585588
E-mail: iriss@ceps.lu