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

Philippe Van Kerm†

Eurostat and Statistics Finland International Conference on "Comparative EU Statistics on Income and Living Conditions: Issues and Challenges"

Helsinki, 6-7 November 2006

[Draft 0.2 of 2006-10-24]

Abstract

Social indicators (of poverty or inequality) are known to be potentially sensitive to the presence of data contamination and to extreme incomes at either or both tails of the income distribution. EU-SILC being an important source for the estimation of such indicators in Europe, it is important to assess how much indicators derived from it are sensitive to alternative adjustments meant to control the impact of extreme incomes. The paper presents the results of a large scale sensitivity analysis considering both simple, classical adjustments (trimming, winsorizing) and a more sophisticated approach based on modelling parametrically the tails of the income distribution. Reassuringly, ordinal comparisons of countries are found to be robust to varying data adjustment procedures. Cardinal comparisons may however be more markedly sensitive. Parametric tail modelling is found to be a potentially promising method.

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

---

*Comments by Alessio Fusco on an earlier draft are gratefully ackonwledged. 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 Differdange, Luxembourg. Institute for Social and Economic Research (ISER), University of Essex, Colchester, United Kingdom. Contact: philippe.vankerm@ceps.lu.

1 Introduction

EU-SILC, the *Community Statistics on Income and Living Conditions* is developed to become the reference source for income and social exclusion statistics across the European Union. One of its major objectives is to serve as a basis for the estimation of comparable welfare indicators (poverty, inequality) in EU countries. In particular, EU-SILC is identified as one of the reference data sources for estimating the common 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 containing comparable individual- and household-level data on income and living conditions, EU-SILC does indeed open up new possibilities for thorough distributional comparisons both over time and, perhaps more importantly, across European countries.

Given this objective, it is important to ensure that estimation of welfare indices from the EU-SILC instrument is as accurate as possible. Accuracy of estimates of poverty and inequality involves a bewildering array of issues, ranging from the mere definition of the underlying concept of economic "well-being", or the definition of "income", to the selection of appropriate summary welfare indicators, via the definition of the basic unit of analysis and the within-housheold income sharing assumptions, or considerations of sampling uncertaintly. The report of the so-called Canberra Group provides a thorough discussion of many of these issues (Canberra Group, 2001) and in-depth reflexions about the definition of social indicators are developed in Atkinson *et al.* (2002).

The present analysis focuses on the specific issue of the treatment of extreme income data and data contamination. Virtually all income data are "dirty". It would be naive to believe that all income data are collected without any error. And even if income data were collected without any errors, when assessing inequality or poverty, income data are not taken for themselves but as measures of economic well-being, of which they are only imperfect proxies. So, even if data were clean reflexions of a set of measured income sources, there are several reasons why they may still be dirty signals about economic well-being. Unfortunately, most welfare indicators are more sensitive to such data imperfections than, say, central tendency measures; in particular, if imperfections lead to the presence of very high or very low income data. There is therefore an issue about how much data imperfections, and differences thereof in cross-country comparisons, affect conclusions that are drawn from estimated welfare measures. It has also been realized that, irrespectively of data imperfections, many inequality and poverty measures are sensitive to the values taken by extreme (high or low) incomes. Besides potentially leading to large sampling error, it might be questioned whether it is acceptable that a few data points –a few responding households– have a large leverage on estimated national indicators.

In recognition of this, it is customary to follow prescriptions of robust statistical methods and to make adjustments to the datasets before estimating indicators, or to focus on indicators known to be more robust to the presence of extreme incomes in order to keep their influence under control. But how much do these adjustments affect the estimated indicators? In particular, how sensitive are cross-country comparisons to the choice of a specific adjustment? The objective of this study is to provide some responses to these questions by means of a large sensitivity analysis based on EU-SILC 2004 data (see Cowell *et al.* (1999) for a similar exercise).

Section 2 discusses briefly the issue of capturing well-being from income data and recalls why the latter, even correctly measured, can be an imperfect measure of the former. Section 3 reviews the models through which income data imperfections are usually considered. Section 4 identifies the magnitude and source of extreme incomes in the 2004 EU-SILC data. Section 5 describes data adjustment procedures in more detail and the lists the welfare indicators that are considered in the sensitivity analysis. Results are presented in Section 6. Section 7 summarizes the main lessons of the analysis and provides a final discussion.

2 Capturing well-being with household income data

Key social indicators are poverty and inequality statistics which inform about the dispersion of citizen's incomes and about the extend to which people are lagging behind societal standards. Note that income is not considered for itself in this context, but rather serves as a proxy for a person's living standard or level of welfare.

It is well appreciated that a person's "subjective well-being" or happiness can not be adequately captured by simple and practical indicators for it is the result of a combination of factors such as material situation, natural and social environment, psychological factors, *etc*. What one attempts to capture by income is people's "economic well-being" which is typically defined by people's access to goods and services (such as food, housing, health care, leisure, and so on). Comprehensive information about the latter is difficult to collect directly. Furthermore what people actually consume does not necessarily reflect the consumption "potential"; people have different preferences over their consumption levels and patterns, and some may discount future consumption differently, therefore may not consume all that is possible at a moment in time. There is therefore a need for indirect measurement.

In a market economy, the command over goods and services is determined by a person's financial resources, and these are themselves determined by income. Income flows increase today's access to goods and services, as well as possibly tomorrow's and yesterday's (through savings and borrowings). For this reason, income is considered as an adequate signal of people's command over goods and services. It must be appreciated that this is not a perfect measure, however. For example, the command over goods and services from a given income is mediated by the set of prices faced by individuals and by the availability of goods and services. Income determines access to goods and services through the market, but many services are publicly provided or do not go into the market (e.g. home produced goods). Also, because people may be able to save and borrow, the income flow over a certain period of time (e.g. one year) does not equate the command over goods and services during this same period. Typically, people in old age are able to consume wealth accumulated earlier in life, or young people can borrow money to benefit today from goods and services (e.g. of housing) that are drawn from income received later in life.

In addition to these imperfections of income to represent people's command over goods and services, there are difficulties associated to the measurement of income. The classical concept of income used for estimating inequality and poverty statistics is "household disposable income" which is computed by aggregating regular incomes received during a given period of time from a number of different sources (labour income, social transfers, capital income (e.g. rents, dividends, interests), *etc.*) minus direct taxes and social security contributions. However, some components are notably measured unreliably, e.g. self-employment income in particular (Eurostat, 2006c). Additionally, capital transfers and irregular incomes (e.g. windfall gains from lotteries or inheritance, capital gains from the selling of stocks or physical capital, *etc.*) are generally not incorporated on the ground that they are large, irregular amounts the benefits of which are distributed over a longer period of time than the income measurement period. However, it is obvious that they influence recipients' command over goods and services.

Finally, there is an additional important step which consists in the move from incomes measured at the *household* level to a measure of *individual* command over goods and services. This is typically achieved by aggregating incomes within households and assuming that all household members can equally draw from it so everyone's command over goods and services is identical. However, it is admitted that some goods or services are "public" within a household (such as heating, durable goods) and that there are economies of scale so that the command over goods and services from a given per capita amount is higher in larger households. This is reflected in the appplication of equivalence scales that convert the incomes of households of different sizes into "single-adult equivalent household incomes" ('equivalent income' for short). This mechanical conversion can potentially add further imprecision to the assessment economic well-being.

It is therefore clear that there are weaknesses in using a concept of annual income receipts from regular sources to build the household disposable income and the equivalent income measure which is finally used as indicative of economic well-being. These weaknesses are well appreciated; see Canberra Group (2001) or Eurostat (2005a, 2006b). It is generally admitted that, considering the difficulties and potential additional pitfalls in attempting to address some of these weaknesses, it is a reasonable compromise between (theoretical) accuracy in the measurement of people's command over goods and services and (practical) reliability of data collection. It remains that we must be aware that we are dealing with imperfect data which provide a potentially 'dirty' signal about the "economic well-being" that we are trying to capture when we engage in cross-national comparisons of inequality and poverty.

3 Treating income as an imperfect signal

Equivalent income can be seen as an imperfect signal about economic well-being, either because income data are imprecisely measured at data collection (respondent's error in surveys or under-reporting in fiscal records, or coding mistakes during the collection phase), or because collected incomes do not cover all relevant sources (e.g. income in-kind), or plainly because equivalent income is only an imperfect proxy for a person's command over resources for any of the reasons briefly exposed in the previous section (such as price differentials, inadequate equivalence scales, wrong assumption of equal-sharing of resources within households, *etc*.). Imputation procedures when data are missing are also source of potential error. The dirt in the signal will affect the precision of estimates of inequality or poverty, and, perhaps more disturbingly, in a cross-country comparison, the nature and structure of the dirt in each country can affect conclusions. There is therefore interest in trying to adopt measures that make cross-country comparisons less sensitive to noisy signals, and to attempt to measure the potential effect of these errors.

Two complementary frameworks have been used to study this model of data imperfection (Cowell, 2000, Bound *et al.*, 2001). The first is what is referred to as a "measurement error" model where it is assumed that what is measured is the true value plus some noise:

$$y = y^* + e$$

where y is measured income, y^* is economic well-being, and e is a random variable that capture all potential deviations. All the complexity of the modelling stands in the assumed distribution of e, which can be allowed to be correlated with y^*, which can have non-zero mean to reflect, e.g., systematic under-reporting, *etc*. Given specific distributional assumptions about e, one could determine the size and direction of biases in inequality and poverty indicators (see van Praag *et al.*, 1983, Gottschalk & Smeeding, 2000, Chesher & Schluter, 2002). However, it is difficult to make such distributional assumptions since, in practice, we have no way to validate these assumptions, and, in particular, to identify potential cross-country differences in these error distributions.

The second approach is referred to as a "data contamination" model where it is assumed that a (large) fraction of the data are accurate indicators of economic well-being (drawn from the true underlying distribution of economic well-being) and a (small) fraction of the data are gross errors that do not reflect closely the economic well-being of the respondent. What is measured can be written as

$$y = y^* \times (1 - P) + z \times P$$

where P is equal to 1 with probability ϵ (in which case what is observed is an erroneous, arbitrary value z) and the true value is observed with probability $1 - \epsilon$ when P is equal to 0 (Horowitz & Manski, 1995). Treatment of this model is typically done with reference to robust statistical methods (Hampel *et al.*, 1986). Given this model, it is possible to estimate the robustness properties of indicators of inequality or poverty to assess whether indicators are sensitive to the presence of contamination and whether they can be driven arbitrarily large (or small) for given values of the 'mistakes' z. Indicators are said to be robust if a single erroneous observation (e.g. an extremely large or small income) can not drive them arbitrarily large or small.

The robustness properties of a series of inequality and poverty measures are studied in detail in Cowell & Victoria-Feser (1996a,b, 2002) and Cowell & Flachaire (2004). The main lesson is that most commonly used indicators of inequality are not robust to data contamination. Most measures are sensitive to the presence of very large incomes, while others are also sensitive to very low incomes. Conversely, poverty indicators are generally robust (if the poverty line is exogenously determined, or is itself robustly estimated).

The robustness properties of inequality and poverty measures also give us indication of their sensitivity to valid, albeit extreme, observations. Indicators sensitive to high incomes will be biased by the presence of large errors, but will also be similarly “driven” by very large (or very small) correctly measured incomes. In general, it is not easy to identify whether an extreme datum is a correct value, is a moderately mis-measured value, or is a gross mistake. Careful data-processing and thorough data-checking obviously help tagging and discarding, or correcting, gross errors. However, errors may be missed at this stage, and very large incomes are often plausible given the recognized long upper-tails of income distributions. Similarly, very low and negative incomes are plausible given the definition of disposable income presented above, although they may also represent gross errors.[1] So, influential observations may be mistakes or genuine income data. But there are good reasons for trying to control their impact in both case: neither do we want mistakes to determine inequality or poverty estimates, nor, most probably, are we ready to accept that only very few correct observations drive our estimates. This is a classical concern in all the outlier detection and treatment literature (see e.g. Osborne & Overbay, 2004).

While adequate treatment for biases caused by measurement error are difficult to design and implement, adjustments for addressing data contamination are routinely implemented. They include trimming a fraction of the data –that is, discarding the top and/or bottom p percent of the data–, or the similar practice of recoding the top and/or bottom p percent of the data to the highest or smallest value in the inner $(1 - p)$ percent of the data. Trimming as a tool for making ‘robust’ welfare comparisons in distributional analysis is thoroughly discussed in Cowell & Victoria-Feser (2006a). More sophisticated approaches borrowed from the literature on robust statistics consist in modelling parametrically or semi-parametrically the income distribution and estimating the parameters with estimators that are robust to data contamination (Victoria-Feser & Ronchetti, 1994, Cowell & Victoria-Feser, 2006b). Section 5 presents such adjustments in more detail.

4 Extreme incomes in EU-SILC

Given the complexity of the disposable household income measure as a combination of different components, and potential cross-national differences in the prevalence and 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.

[1] Non-positive command over goods and services during a year is an implausible situation as it would imply complete starvation. However, there are several reasons for negative measured incomes (Eurostat, 2006d). First, 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, 2005b). Losses can be observed with self-employment income (Eurostat, 2006c). 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. So negative incomes do not directly translate into negative economic well-being and are therefore plausible. They may not be plainly tagged as “mistakes” in the sense of error of data collection. They are clear expressions of a mis-measurement of economic well-being, but the extent of the mis-measurement is not necessarily easier to assess than for other (positive) income observations.

Figures 1.1 to 1.5 provide depictions of the income data contained in the EU-SILC 2004 dataset (Eurostat, 2006a) in each of the fourteen country 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 constitutent income components are described along with the constructed variables "total household gross income" (where available) and "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 that emerge from the inspection of these figures are the following. Consider first the group of countries with incomes presented gross.

- 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 ('outlying') 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 any receipt of self-employment incomes (from 9 percent in BE up to 40 percent in DK). Besides actual differences, several well-known factors can be the cause of these discrepancies: differences in the mode of collection of the self-employment data (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 (2006c) and Eurostat (2006d).) Given the manifest cross-national variability in the resulting data, it seems important to further harmonize practice.

- Old-age benefits are the benefits with the largest range, and a few observations are recorded as "very high" in most 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 being observed 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. (Note however the non-negligible presence of negative amounts in LU.)

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).[2]

The prevalence of very high or very low incomes in the aggregated income variables of interest here, namely household disposable income and single-adult-equivalent household income is presented in Figure 2. 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 coloring 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 however well below one percent of the observations. But 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. Note that the picture is hardly affected if equivalent income is used rather than total disposable income. Extreme observations are present in all of the countries data, and there is

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

therefore interest in assessing how much comparisons of distributional indicators are affected by their presence.

5 Estimating inequality and poverty measures with extreme incomes

As mentioned in Section 3, welfare measures (inequality measures in particular) are sensitive to the presence of extreme incomes. Additionally, some measures are not even identified if there are non-positive incomes in the data (Atkinson indices of inequality, some Generalized Entropy measures, the variance of log-income, or the Watts poverty index). Recognizing this, several adjustments to the data are routinely applied before estimating inequality and poverty measures. Eurostat (2006d) discusses the particular problems posed by zero and negative incomes and suggests a set of data adjustments. However, high incomes are also known to be potentially influential. Such adjustments are generally deemed necessary, but their application is of an *ad hoc* nature and one rarely estimates the magniture of their impact on the estimated indicators. In particular, in cross-country comparisons, it is important to identify ranges of variation of the indicators under different types of adjustments to be able to check whether country orderings are sensitive to the nature of the adjustments. Inspection of the data suggests that the prevalence of extreme observations vary from country to country. So data adjustment may have different impacts. But different measures will be affected differently by alternative adjustements, so it is generally difficult to make accurate expectations about the magnitude of the impacts in different countries.

The present study addresses the issue by engaging in a thorough sensitivity analysis based on the EU-SILC 2004 data. Three families of adjustments are considered. The first two are simple standard methods: trimming fixed percentages of the data, and winsorizing the same fixed percentages of the data. The third approach makes use of a parametric model for the tail of the distribution which is then used to impute extreme incomes from the parametric tail. For completeness, following Cowell *et al.* (1999), an extreme adjustment is also considered which consists in removing data relying heavily on data sources known for their low reliability and with high prevalence of extreme values, namely self-employment incomes and incomes from interests, dividends and profits.

5.1 Data adjustment procedures

Trimming Trimming consists in removing from the dataset a given percentage of the highest and/or lowest incomes (Cowell & Victoria-Feser, 2006a). Estimations have been made with trimming percentages of 0.25%, 0.50%, 0.75%, 1% and 2%, both one-sided and two-sided. Trimming thresholds above and beyond which observations are discarded are therefore given by the quantiles $Q(1-0.01p)$ and $Q(0.01p)$ of the unadjusted dataset. (In practice, trimming precentages are often chosen in the range 0.5 to 1 in both tails.) 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 Winsorizing is a close relative to trimming. The difference is that the extreme data are not removed from the datasets but are replaced by the value of the trimming thresholds. The adjusted data can be expressed as follows:

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

Winsorizing can be seen as a particular form of income imputation. This is also referred to as 'top-coding' or 'bottom-coding' which is often applied with respect to data confidentiality issues.

Drawing from parametric tails Trimming and, to a lesser extent, winsorizing are the most commonly adopted practice for making estimates robust to outlying observations. However, Hampel *et al.* (1986) emphasize that this practice can be viewed as overly conservative in the sense of trading-off too much data information against robustness. One more sophisticated approach to addressing robustness problems of inequality and poverty measures is to estimate a parametric model for the income distribution whose parameters are estimated using methods robust to outlying observations. Inequality and poverty indicators can then be estimated indirectly from the empirical distribution of the estimated distribution; see Victoria-Feser & Ronchetti (1994) or Cowell & Victoria-Feser (2006b) for a semi-parametric approach.

As such, this approach is not based on data adjustment and rely on the availability of accurate numerical integration algorithms to estimate inequality and poverty indices. This can reveal problematic for routine estimation. It is however possible to use robust parametric models for data adjustments. The idea is to impute extreme incomes by replacing the observed values by random draws from the robustly estimated parametric distribution. Simulation uncertainty is introduced by the random draws, but this can be easily controlled by simulating a set of replicate income data, estimating poverty and inequality on each of the replicate and taking as indicator the average over the replications of each indicator (Little & Rubin, 1987).

The robust estimation of a parametric tail model may appear much more technically challenging than standard 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. In this application, following Cowell & Flachaire (2004), Davidson & Flachaire (2004), Cowell & Victoria-Feser (2006b), a Pareto distribution is used as the parametric tail model, 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 inverse Pareto distribution is used for the lower tail:

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

for $y < y_l$. Standard maximum likelihood can be used to estimate the parameters θ_l and θ_h. However, maximum likelihood estimates are not robust and are known to be sensitive to the presence of extreme incomes. If extreme incomes are allowed to influence the parameter estimates, the simulated incomes will not be drawn from a robustly estimated distribution. It is therefore essential to estimate the parameters with an algorithm that provide robust estimates of the Pareto distribution parameters. The method applied here is the so-called *optimal B-robust estimator* (OBRE) detailed in Victoria-Feser & Ronchetti (1994) and Cowell & Victoria-Feser (2006b) which belongs to the family of M-estimators (Huber, 1972, 1981).

This procedure leads to an adjusted dataset such that

$$\begin{aligned} y_i^a = y_i^{sl} \quad &\text{if} \quad y_i < y_l \\ y_i^a = y_i \quad &\text{if} \quad y_l \leq y_i \leq y_s \\ y_i^a = y_i^{sh} \quad &\text{if} \quad y_i < y_h \end{aligned}$$

where y_i^{sl} is a random realization from $F^L(y;\theta^l, y^l)$ and y_i^{sh} is a random realization from $F^H(y;\theta^h, y^h)$. Multiple imputation was applied with eight replications. Note that the lower-tail Pareto distribution does not prevent negative incomes to arise, as these are considered plausible values. Remember, however, that the influence of large negative incomes on the estimated shape of the Pareto distribution is relatively low because of the application of the OBRE algorithm.[3]

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

$$\begin{aligned} y^l &= \min(\max(0.3\mu, Q(0.015)), Q(0.025)) \\ y^h &= \max(\min(2.5\mu, Q(0.98)), Q(0.97)) \end{aligned}$$

where μ is mean disposable income and $Q()$ are quantiles. This choice was selected by trial and error taking into account a trade-off between model fit and the need to keep the number of adjusted income data reasonably low, as well as the will to have a common rule applied to all countries datasets. The model 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) and the derived estimates ought to be considered provisional for these two countries.

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 to provide a set of weights that reflect how much "influential" the data are (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 are 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

[3] It is also conceivable to model the lower tail of the distribution with a model preventing non-positive values. It is unclear, however, how actual negative incomes would contribute to the estimation of such a model.

adjusted weights when computing poverty and inequality indicators partially offsets the effect of the largest and smallest observations but retaining them in the data.[4]

One illustrative way of comparing trimming, winsorizing and the parametric tail modelling is in terms of the assumed information content of the extreme data points. Trimming assumes that extreme observations contain no information about the economic well-being of the recipients. Winsorizing, on the contrary, assumes that high/low extreme incomes remain indicative of high/low economic well-being, but that its magnitude is unknown and is capped by an upper/lower limit. In parametric tail modelling finally, extreme data points are also considered as representing high/low economic well-being, and it is further assumed that the observed extreme points are indicative of the true underlying distribution of extreme data, albeit with potential contamination and outlying values.

Dropping unreliable income recipients 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 subpopulations 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).

5.2 Inequality and poverty measures

The sensitivity analysis reported in the next section considered the impact of the different procedures on the following set of indicators.

- *Central tendency indicators*: mean and median equivalent income. It is expected that only the mean will be affected by the adjustments. Note that both can be used to determine poverty lines and therefore their sensitivity gives us indication about the sensitivity of the determination of the poverty line.

- *Inequality indicators*: 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 Gini coefficient; and 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

[4]Note 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.

of non-positive incomes (with the exception of GE(2)). Note that the Gini coefficient and the S80/S20 are both in the list of Laeken indicators.

- *Poverty indicators*: 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) and 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. The at-risk-of-poverty rate and the relative median at-risk-of-poverty gap are Laeken indicators. All poverty indices were also estimated for households with dependent children.

All indicators are estimated from the single-adult equivalent household income estimated at the household level. Data are weighted by the household size times the household sample weight in order to depict the distribution of income among individuals.

6 Results of the sensitivity analysis with the EU-SILC 2004 data

Full results of the sensitivity analysis are collected in Figures 3.1 to 3.22. Each figure presents the results for one particular indicator. Estimates for the 14 countries are reported as columns of points. Each column corresponds to a particular data adjustment for extreme values.

- The first column presents the estimates obtained from the full sample, that is, without any adjustment to the sample. (These benchmark points are repeated in the last column for easier reference.)

- The next set of (nine) columns (labelled “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 incomes are trimmed. The next three columns trim the bottom 1, 5 and 10 incomes. The last three columns show results of trimming the 1, 5 and 10 extreme incomes at both tails.

- The second set of (fifteen) columns (labelled “Trimming”) shows results of applying a systematic trimming of a fixed percentage of the data, respectively 0.25%, 0.5%, 0.75%, 1% and 2%, either at the top only, the bottom only, and trimming the same percentage on both tails.

- The third set of (fifteen) columns (labelled “Winsorizing”) is equivalent to the previous results except that the fixed percetnage of the data are winsorized rather than trimmed.

- The fourth set of (nine) columns (labelled “Model-based”) presents the results based on applying a parametric-tail model. Both estimates obtained from a standard maximum likelihood estimation of the Pareto parameters and from the OBRE algorithm are presented. Again results are presented for modelling the upper tail only, the lower tail only, and both tails. The last three columns of the set are the estimates obtained by keeping the full original sample but re-weighting observations with the leverage weights delivered after application of the OBRE algorithm.

- The last set of (three) columns (labelled "Sources") finally 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 obtained in the different countries. The main observations are summarized in the concluding section.

Consider first, briefly, the central tendency measures. Expectedly, the median is hardly affected by any adjustment for extreme values. Interestingly, the median remains 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 some support to the case of making reference to median income to set poverty lines.

Inequality indices are potentially the most problematic indicators as many of them are known to be sensitive to extreme values. The indices considered here are presented in increasing order of expected sensitivity to extreme incomes.

Quantile ratios are influenced by trimming at either tail of the distribution. Even though quantiles are robust statistics, the exclusion of even a small fraction of the data can have a noticable influence, but the effect is common enough in all countries not to lead to rank-reversals of countries. Expectedly, imputation-type adjustments leave quantile ratios unaffected. Removal of self-employment incomes mostly affect the P80/P20 results for southern European countries (Greece, Italy and Spain) with a marked inequality reduction, but otherwise this impact is limited. The effect is much stronger on the P90/P10 measure, although with limited reversals of countries.

Income share ratios reveal themselves more sensitive to data adjustments. This had to be expected because these are non-robust statistics potentially driven arbitrarily large by either very large or very small incomes. Adjustments in both tails have an impact. It is striking to notice that the removal of just 10 observations can lead to large changes in the estimated indicators. Cardinal differences between countries are now markedly affected. Winsorizing also lead to variations, and some rank reversals are observed in the S90/S10 case (compare e.g. EE and PT or GR and IT). Model-based indicators however remain close to the full sample estimates.

The Gini coefficient is varying but is more substantially affected by adjustements to top incomes. Again winsorizing keeps estimates more stable, but the changes are much larger than with the model-based approach. The Gini coefficient and the inequality indicators discussed so far are well-defined in the presence of non-positive incomes. The picture changes dramatically with the next set of indicators: Generalized Entropy measures and Atkinson measures. In Figures 3.8 to 3.13, only estimates that could be estimated because the underlying adjusted data only contained positive values are plotted. (GE(2) is an exception as it is defined even in the presence of negative incomes.) The full sample could not be used in any country. Only winsorizing and trimming at least one percent of the low income data lead to the estimation of the indicators for the full set of countries. The model-based approach allows negative incomes to belong to the data and, therefore, does not allow any estimation of such measures. GE(2) although defined with negative incomes appears very sensitive to the presence of, and therefore the treatment applied to, high incomes (see the striking impact of a single income datum on the inequality estimate for Norway). This leads to some substantial country reversals whenever upper-tail adjustement is applied compared to full sample estimates.

Most poverty measures have been reported to be robust to data contamination provided the poverty line is exogenously defined (or is itself robustly defined) and as long as incomes are bounded from below (Cowell & Victoria-Feser, 1996a). The approach here does not assume that there is a lower bound on incomes and this theoretical argument may not be valid. There is therefore interest in checking the behaviour of the poverty indicators to the treatment of extreme incomes, in particular low incomes.

Unsuprizingly, only trimming has an impact on the headcount ratio, and the impact remains relatively low anyway. The picture varies substantially for all the other measures. Trimming and, to a lesser extent, winsorizing at the lower tail lead to marked estimated poverty reductions. Somewhat more reversals in the ranking of countries are observed than for inequality measures, in particular if trimming is applied. Removing the self-employed also makes a critical difference to the results.

The most striking effect is visible for the average squared poverty gap ratio in Greece which is dramatically affected by the treatement of a few very low incomes. For such a measure, substantial reranking is observed with trimming or with the elimination of the self-employed. Model-based adjustment lead to much more limited adjustements than trimming or winsorizing, especially for high poverty countries.

Interestingly, trimming substantially affects the results for the median poverty gap ratio among the poor (one of the Laeken indicators) whereas, by construction, indicators are hardly affected by winsorizing or model-based adjustments.

Finally, if we consider the impact of the various adjustments on poverty in a sub-population of interest (households with dependent children), similar observations emerge. Trimming the lower tail can seriously modify the conclusions that one might draw, especially for indices more sophisticated than the headcount ratio.

7 Summary and discussion

The main lesson that emerges from the exercise is probably the following

Lesson 1 *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 suprizing result, but it is certainly a reassuring baseline. However, this result must be carefully qualified.

Lesson 2 *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 *magnitude* of cross-country differences can vary substantially, even with relatively small data adjustements. Care is therefore called for, and it is recommended to check the sensitivity of one's cardinal comparisons to different data adjustements before making strong statements about it. This is true for most of both inequality and poverty indicators.

Lesson 3 *Poverty measures are not notably less sensitive to the treatment of extreme incomes than inequality measures. However, it is mostly extreme low incomes that matter for the former.*

Lesson 4 *Theoretically sound inequality indices such as Generalized Entropy measures and Atkinson inequality measures suffer from either estimation impossibility with non-positive values or from very high sensitivity to extreme incomes. The routine data adjustements considered here do not appear most appropriate for their estimation.*

Arguably, most of these results were anticipated given our knowledge of the robustness properties of poverty and inequality indicators, but there is virtue, however, in examining the *magnitude* of the results in EU-SILC.

The comparison of the implications of adopting different data adjustment methods provide interesting guidance for further analysis. Different procedures can lead to markedly different results.

Lesson 5 *Adjustments that modify/impute the extreme data without removing them from the dataset lead to results that are markedly more stable than trimming procedures.*

Again, even if ordinal comparisons are generally preserved under alternative adjustment methods, cardinal comparisons may be affected. It is important to emphasize, however, that 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 for various reasons 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. But note that careful examination of the sensitivity analysis suggests that, provided a common procedure is adopted (e.g. trimming percentages or winsorizing or parametric modelling), adopting different parameters (such as different trim percentages in a "sensible range") 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 (if at all for quantile-based measures) to the sample percentage that is "imputed". A similar argument can be put forward for model-based imputation. In addition, parametric-tail modelling 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. But even then, goodness-of-fit of the Pareto distribution gives indication about where to select the cut-off. This is a clear advantage.

Lesson 6 *Data adjustment based on parametric-tail imputation perform well, and may be seen as preferable to trimming.*

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 "innner 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 undisputably 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. As far as intepretation is concerned, it does not modify the underlying distribution but merely assume 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 adjustements. Again this is a plausible fact considering the extensive prior data cleaning of the dataset.

Lesson 7 *Self-employment income remains a source of concern: (a) it is a major source of extreme incomes on both tails of the distribution, and (b) 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.

The position taken in this analysis is not to consider negative and zero incomes as different from the rest of the data on *a priori* grounds. The main reason is that given the definition of household disposable income, non-positive incomes are plausible. 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 aforementioned 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 clearly signal themselves as weak indicators of the command over goods and services of the household, but no positive amount is necessarily more accurate. We considered sound to treat them no differently from the rest of the data. Nevertheless, the difference is more a matter of principles since, in practice, we end up correcting these values to the extend that they are treated as extreme values by the data adjustment procedures in the lower tail. An important exception is the parametric modelling approach, which models the distribution of low (and negative incomes) rather than discard or recode them. This leads to data adjustments with negative values and prevent the estimation of some inequality measures.

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 observed directly. All adjustments are simply meant to keep the magnitude of potential errors under control, balancing robustness and data information. It is worth repeating the evidence that 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 important to re-assure ourselves that important conclusions are not dramatically affected by extreme incomes and they way the are handled.

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 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.

References

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

Bound, J., Brown, C. & Mathiowetz, N. (2001), 'Measurement error in survey data', in J. J. Heckman & E. Leamer (eds.), *Handbook of Econometrics, Vol. 5*, North Holland, Amsterdam, The Netherlands, 3705–43.

Canberra Group (2001), 'Expert group on household income statistics: Final report and recommendations', The Canberra Group, Ottawa.

Chesher, A. & Schluter, C. (2002), 'Welfare measurement and measurement error', *Review of Economic Studies*, **69**:357–78.

Cowell, F. A. (2000), 'Measurement of inequality', in A. B. Atkinson & F. Bourguignon (eds.), *Handbook of Income Distribution*, North Holland, Amsterdam, The Netherlands, 87–166.

Cowell, F. A. & Flachaire, E. (2004), 'Income distribution and inequality measurement: The problem of extreme values', Cahiers de la Maison des Sciences Economiques (Série verte) 2004-101, Université Paris 1.

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.

— (2006a), 'Distributional dominance with trimmed data', *Journal of Business and Economic Statistics*, **24**(3):291–300.

— (2006b), 'Robust stochastic dominance: A semi-parametric approach', Forthcoming in *Journal of Economic Inequality*.

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.

Eurostat (2004), 'Imputation procedures', EU-SILC Documents 136/04, European Commission, Eurostat.

— (2005a), 'Continuity of indicators between end-ECHP and start-SILC', European Commission, Eurostat.

— (2005b), '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' data base', Release June 2006, European Commission, Eurostat.

— (2006b), 'EU-SILC user database description', European Commission, Eurostat.

— (2006c), 'Self-employment income', EU-SILC Documents TFMC-02/06, European Commission, Eurostat.

— (2006d), 'Some proposals on the treatment of negative incomes', EU-SILC Documents TFMC-15/06, European Commission, Eurostat.

Gottschalk, P. & Smeeding, T. M. (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.

Horowitz, J. L. & Manski, C. F. (1995), 'Identification and robustness with contaminated and corrupted data', *Econometrica*, **63**(2):281–302.

Huber, P. J. (1972), 'Robust statistics: A review', *Annals of Mathematical Statistics*, **43**(4):1041–67.

— (1981), *Robust statistics*, John Wiley, New York.

Little, R. J. A. & Rubin, D. B. (1987), *Statistical analysis with missing data*, John Wiley, New York.

Osborne, J. W. & Overbay, A. (2004), 'The power of outliers (and why researchers should always check for them)', *Practical Assessment, Research and Evaluation*, **9**(6).

van Praag, B., Hagenaars, A. & van Eck, W. (1983), 'The influence of classification and observation errors on the measurement of income inequality', *Econometrica*, **51**(4):1093–1108.

Victoria-Feser, M.-P. & Ronchetti, E. (1994), 'Robust methods for personal income distribution models', *Canadian Journal of Statistics*, **22**(2):247–58.

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)

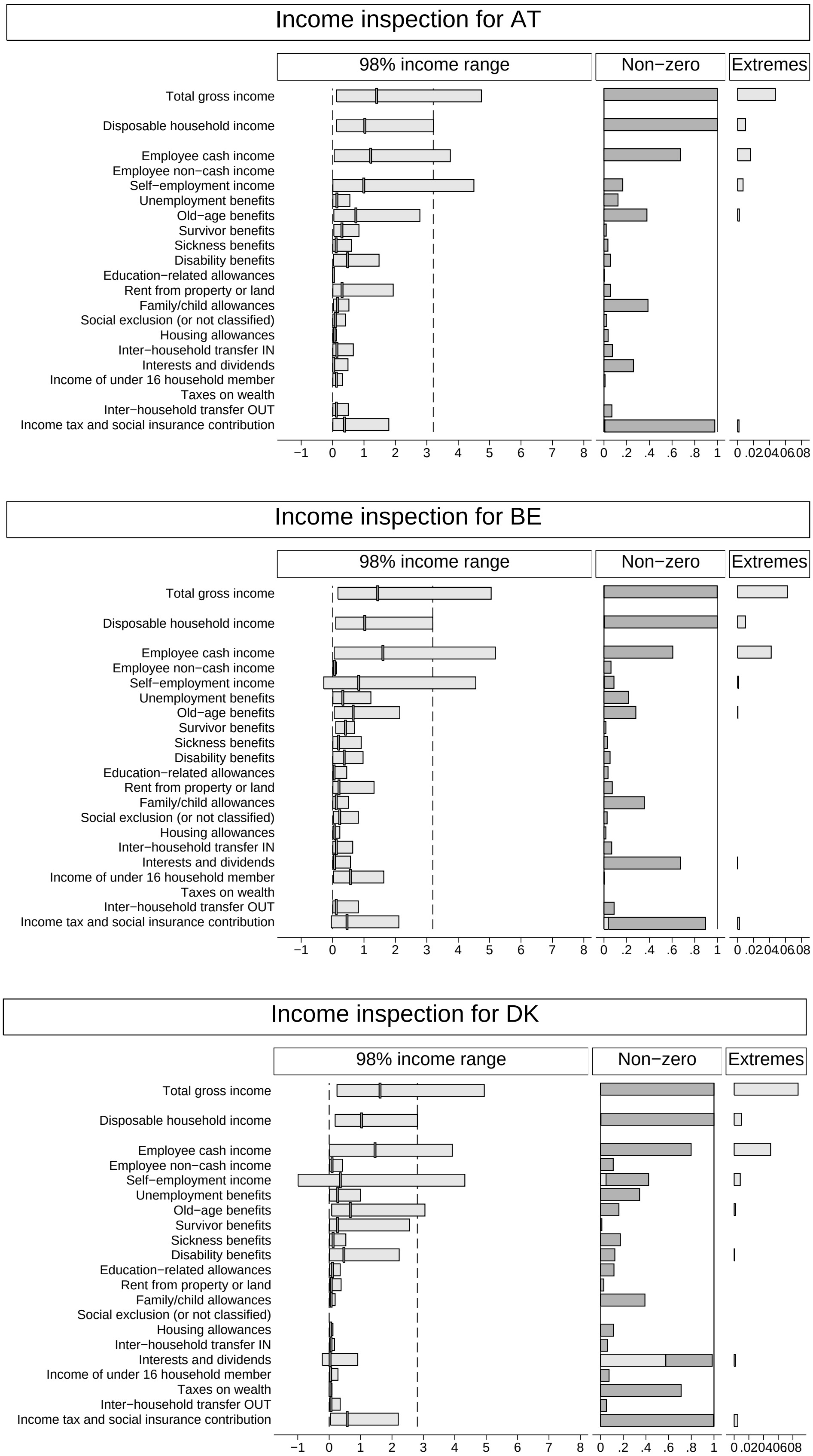


Figure 1.2: Income components in Estonia, Finland and Ireland: income ranges, fraction of non-zero observations, and proportion of extreme positive values (gross components)

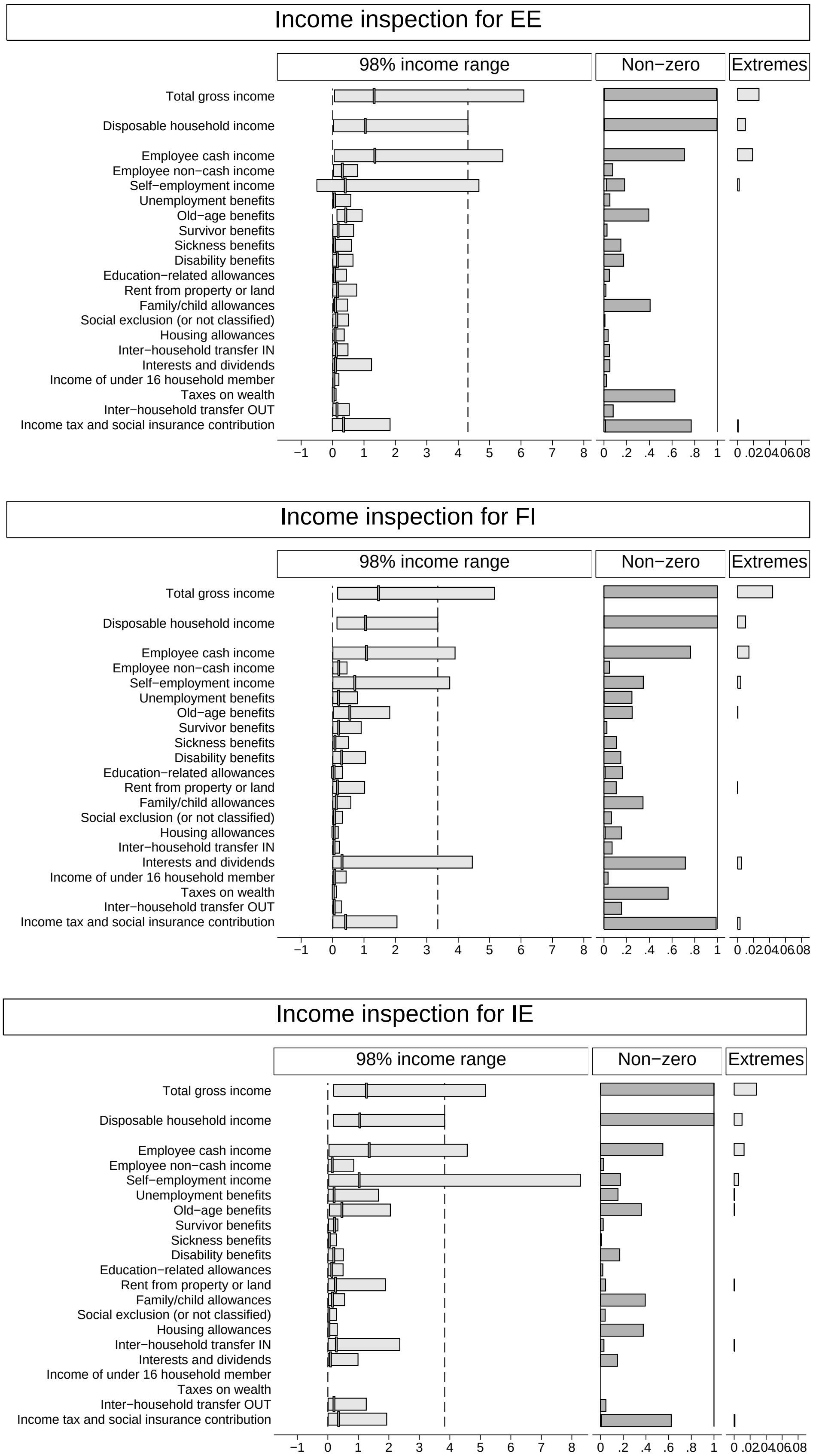
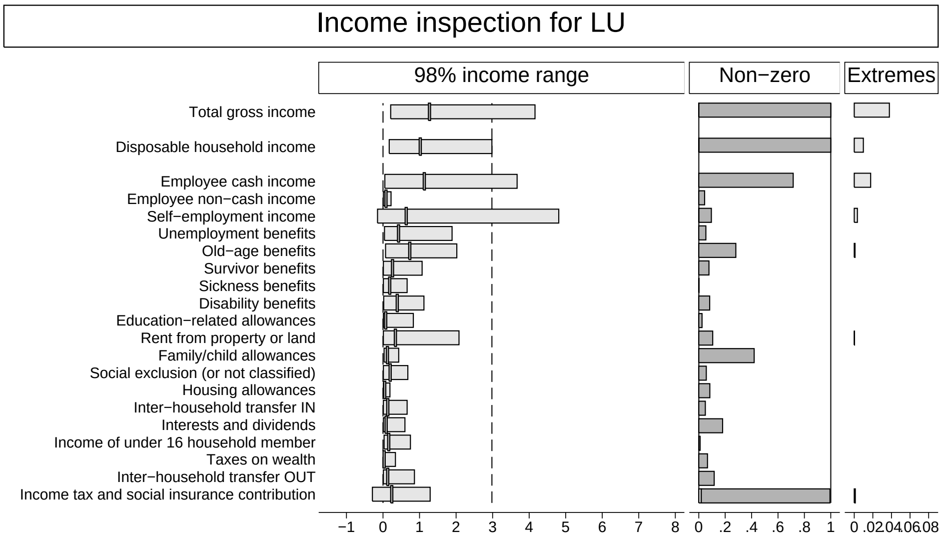


Figure 1.3: Income components in Luxembourg and Norway: income ranges, fraction of non-zero observations, and proportion of extreme positive values (gross components)



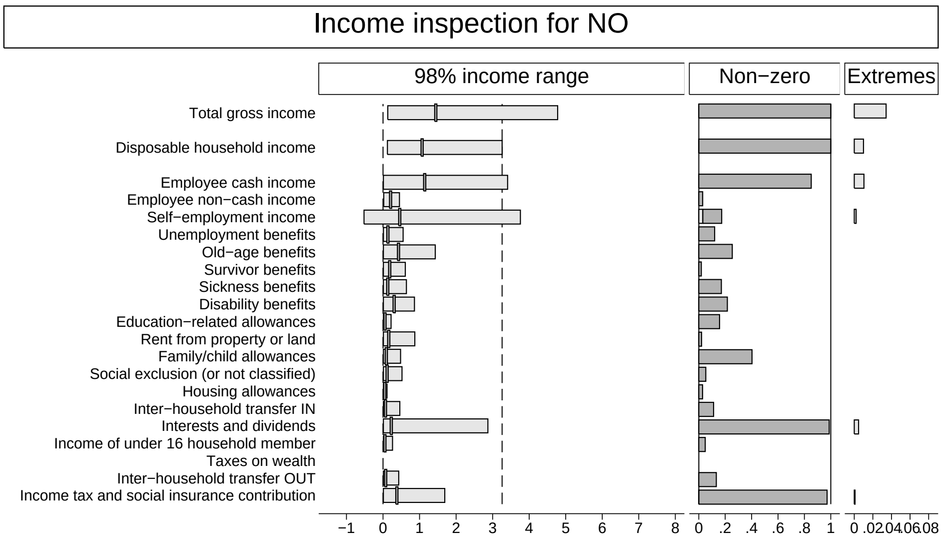
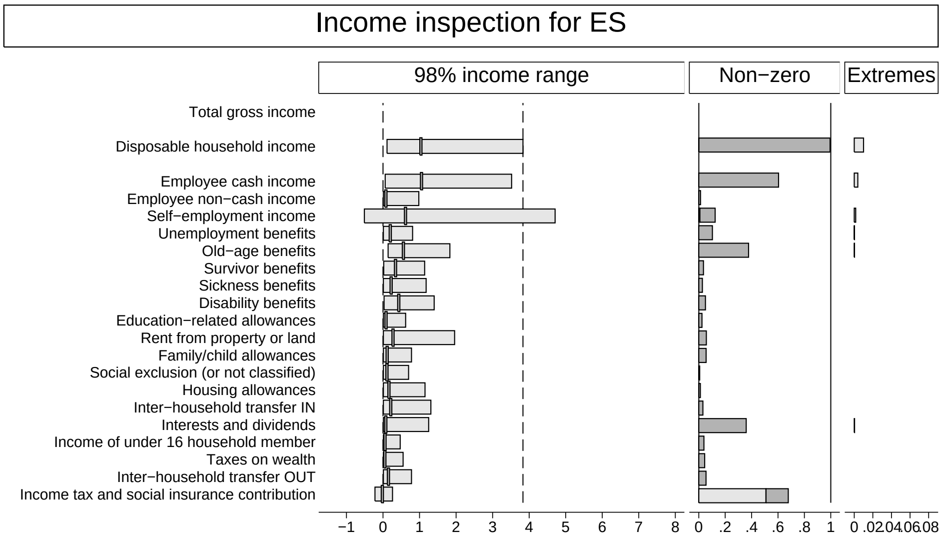
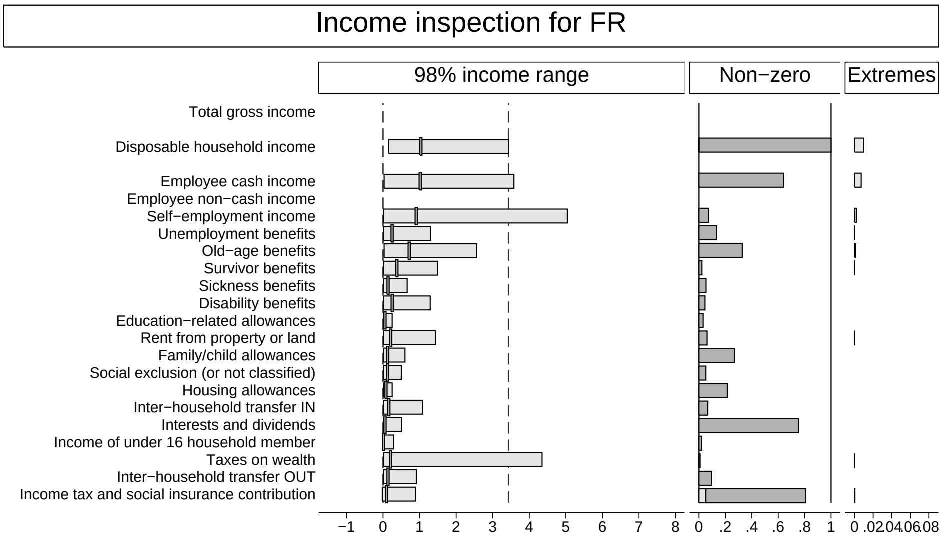


Figure 1.4: Income components in Spain, France and Greece: income ranges, fraction of non-zero observations, and proportion of extreme positive values (net and/or gross components)





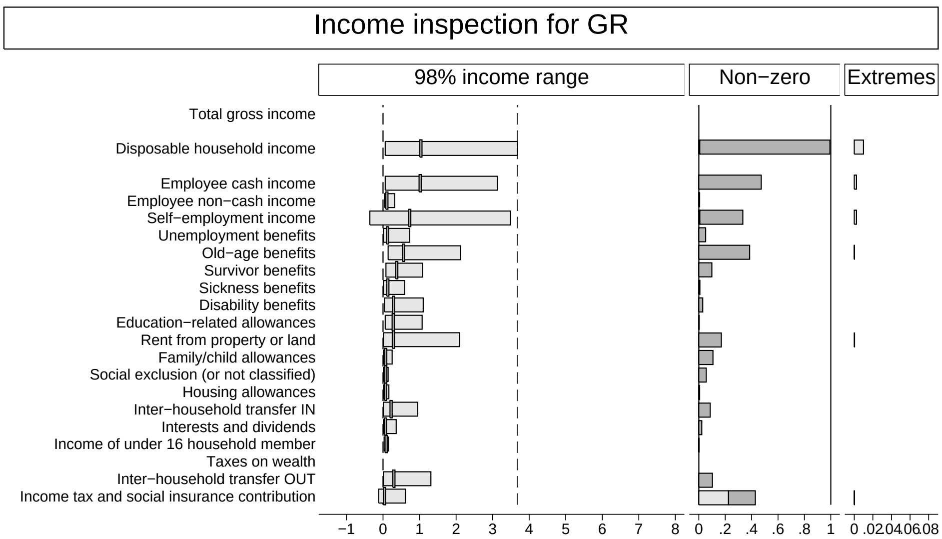


Figure 1.5: Income components in Italy, Portugal and Sweden: income ranges, fraction of non-zero observations, and proportion of extreme positive values (net and/or gross components)

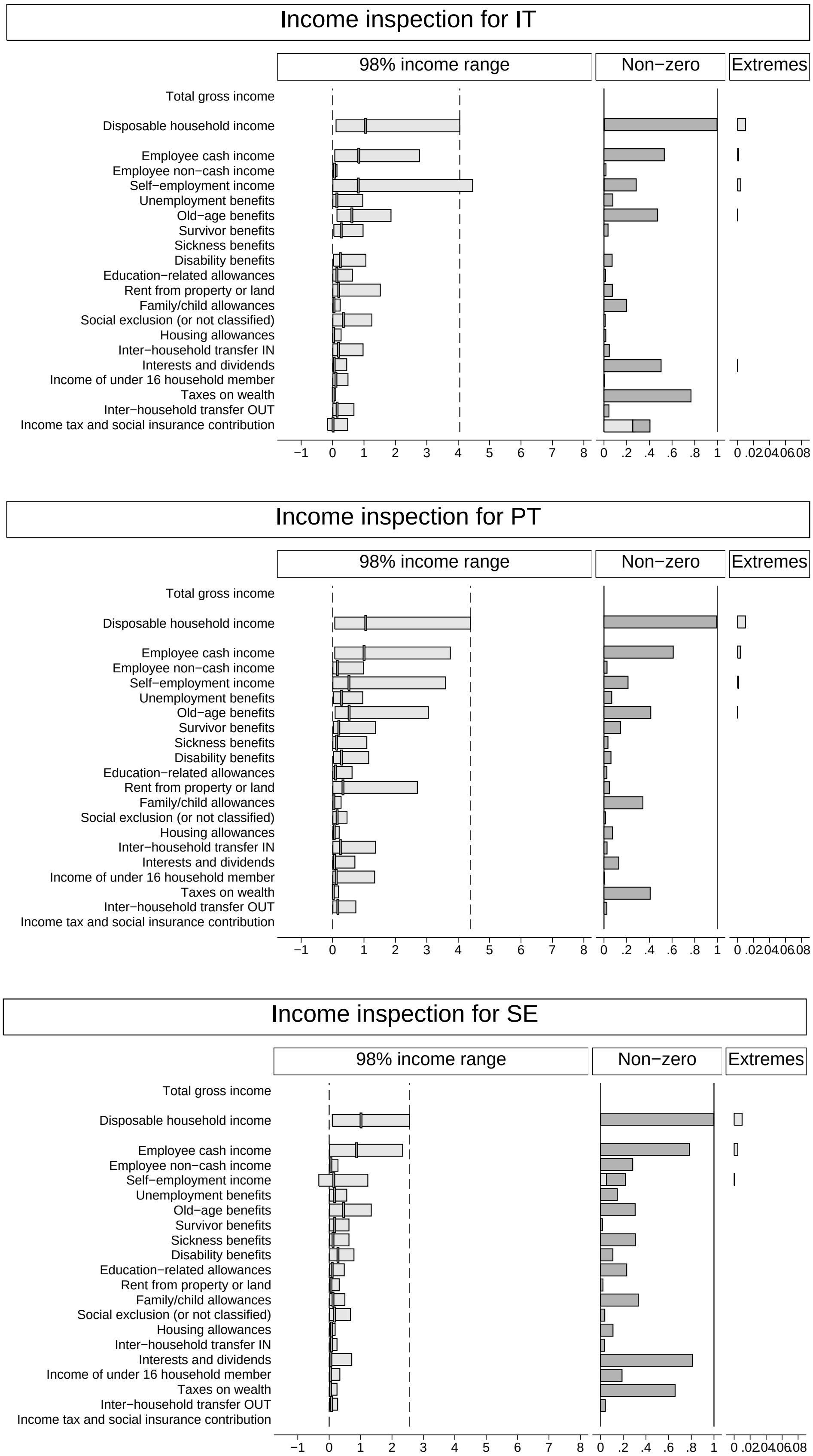


Figure 2: Proportion of very low and very high incomes in disposable income and in equivalent income. Low incomes are incomes below 10 percent of the mean (of which negative incomes are identified in light grey), high incomes are incomes above 4 times the mean (two groups are depicted 4-8 times the mean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

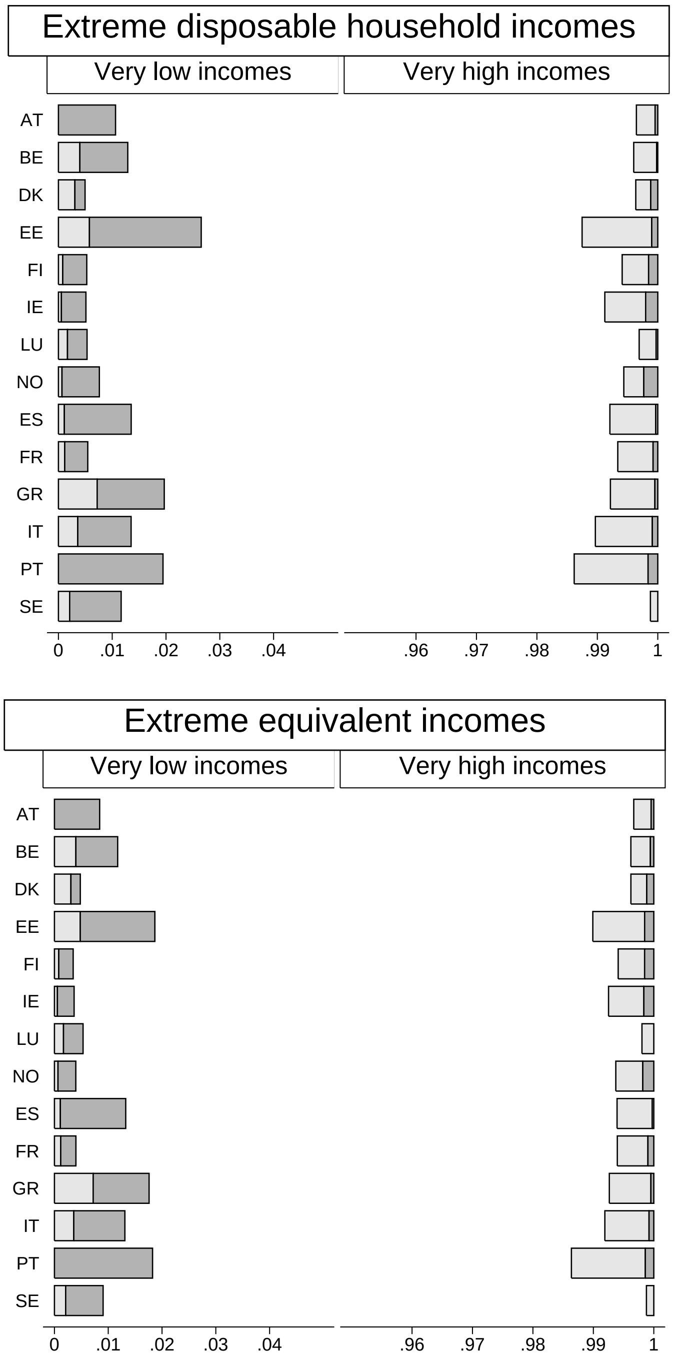


Figure 3.1: Central tendency measures: Estimates of mean equivalent income in 14 countries under alternative treatment of extreme incomes

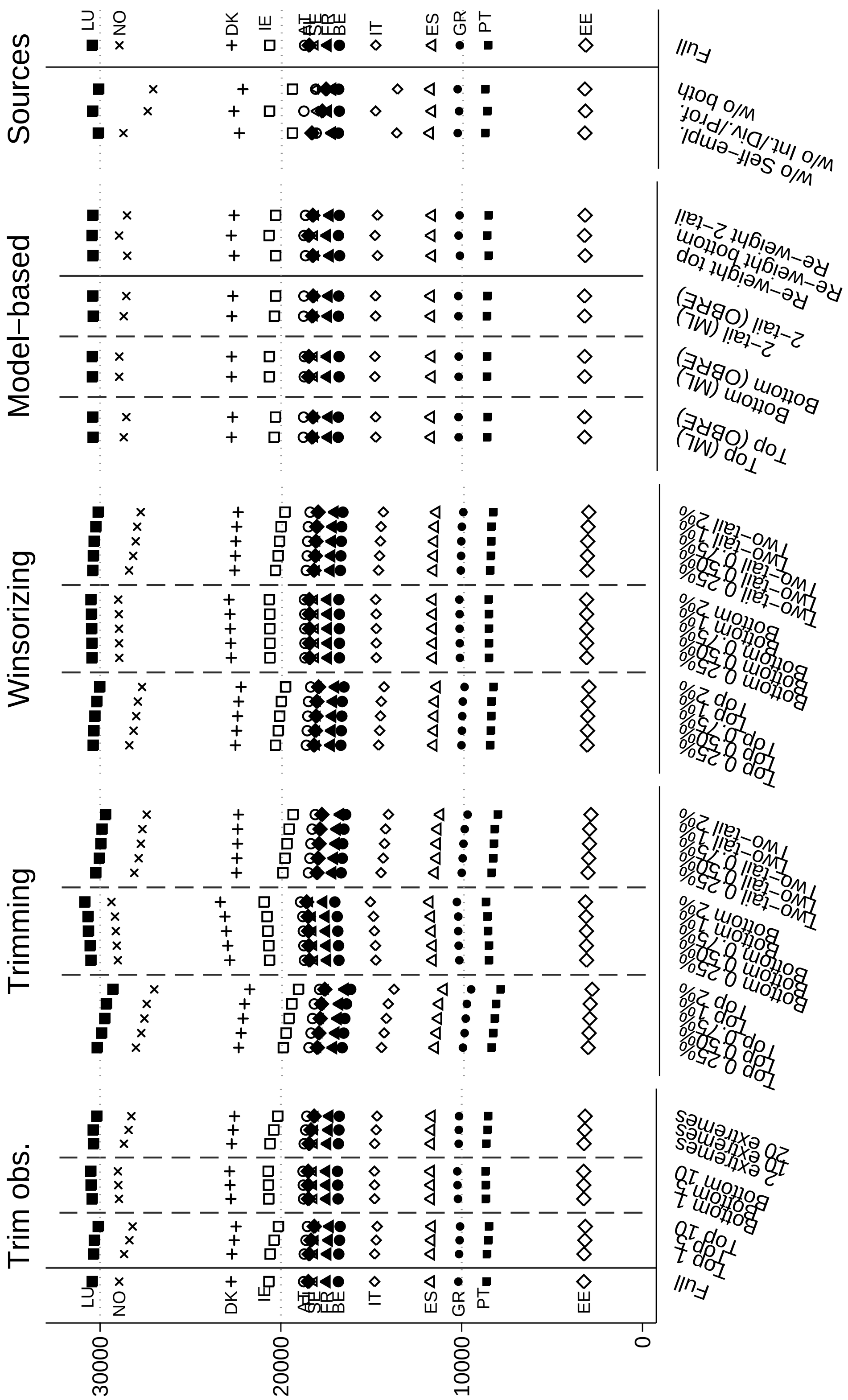


Figure 3.2: Central tendency measures: Estimates of median equivalent disposable income in 14 countries under alternative treatment of extreme incomes

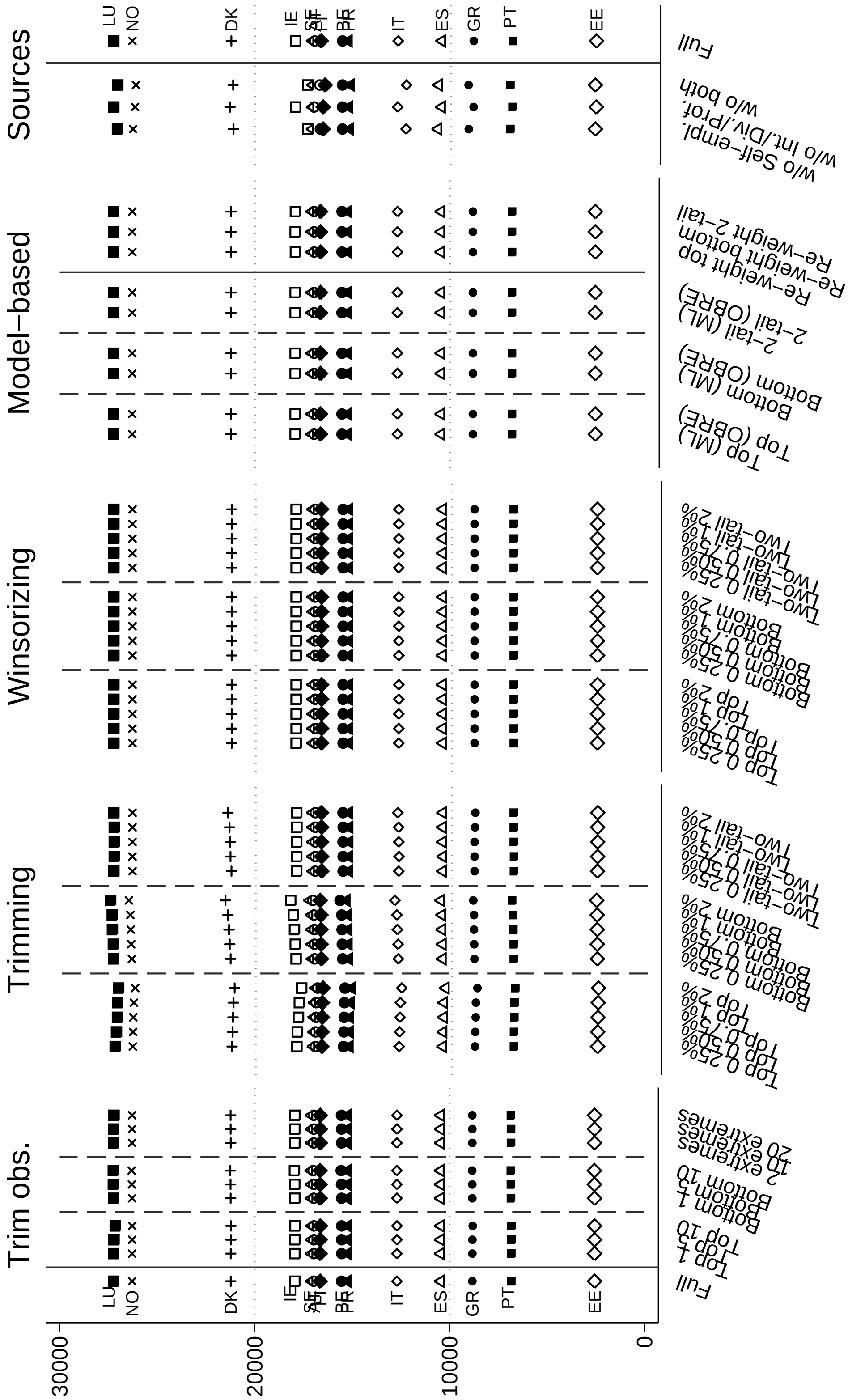


Figure 3.3: Inequality: Estimates of the P80/P20 quantile ratio in 14 countries under alternative treatment of extreme incomes

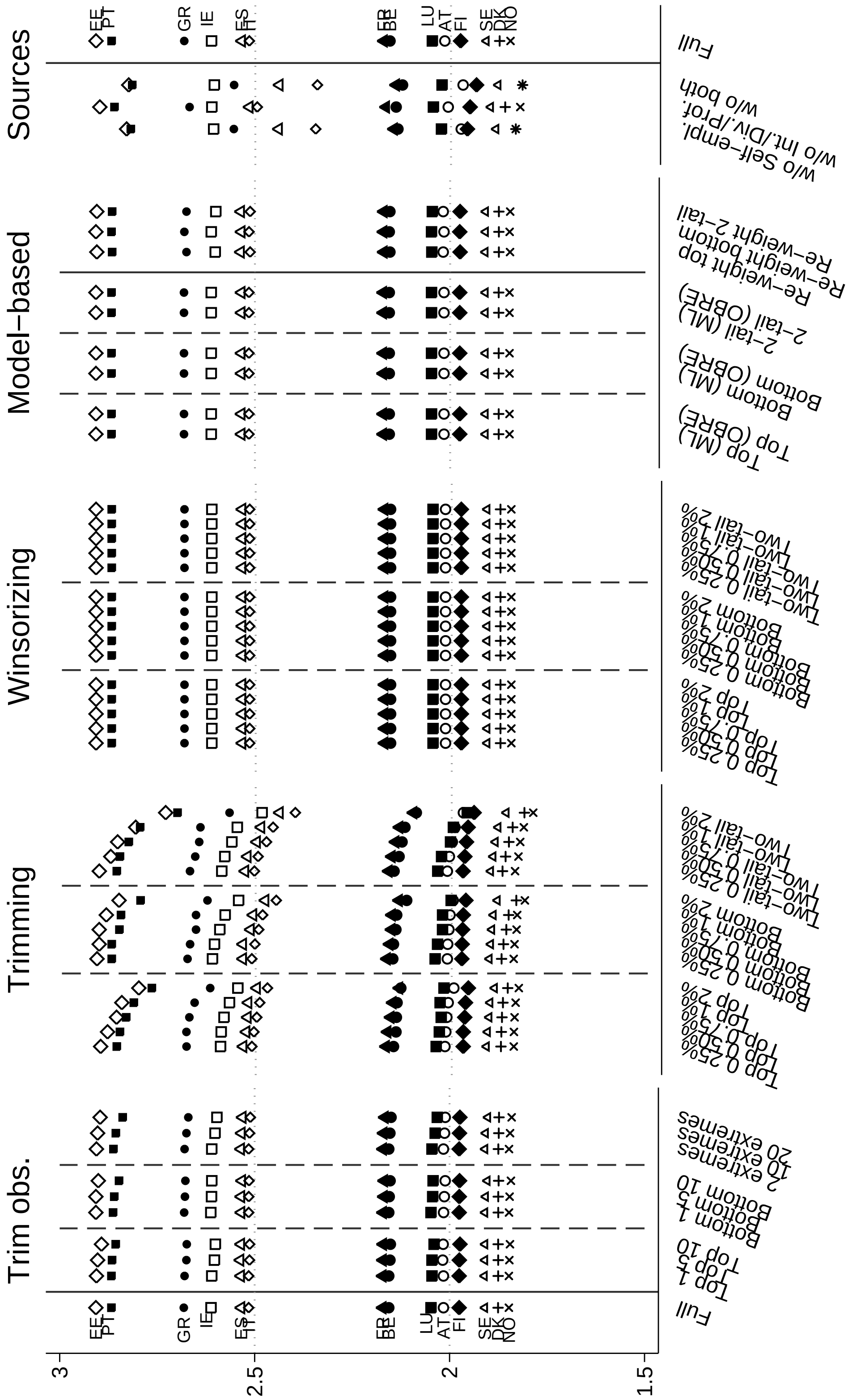


Figure 3.4: Inequality: Estimates of the P90/P10 quantile ratio in 14 countries under alternative treatment of extreme incomes

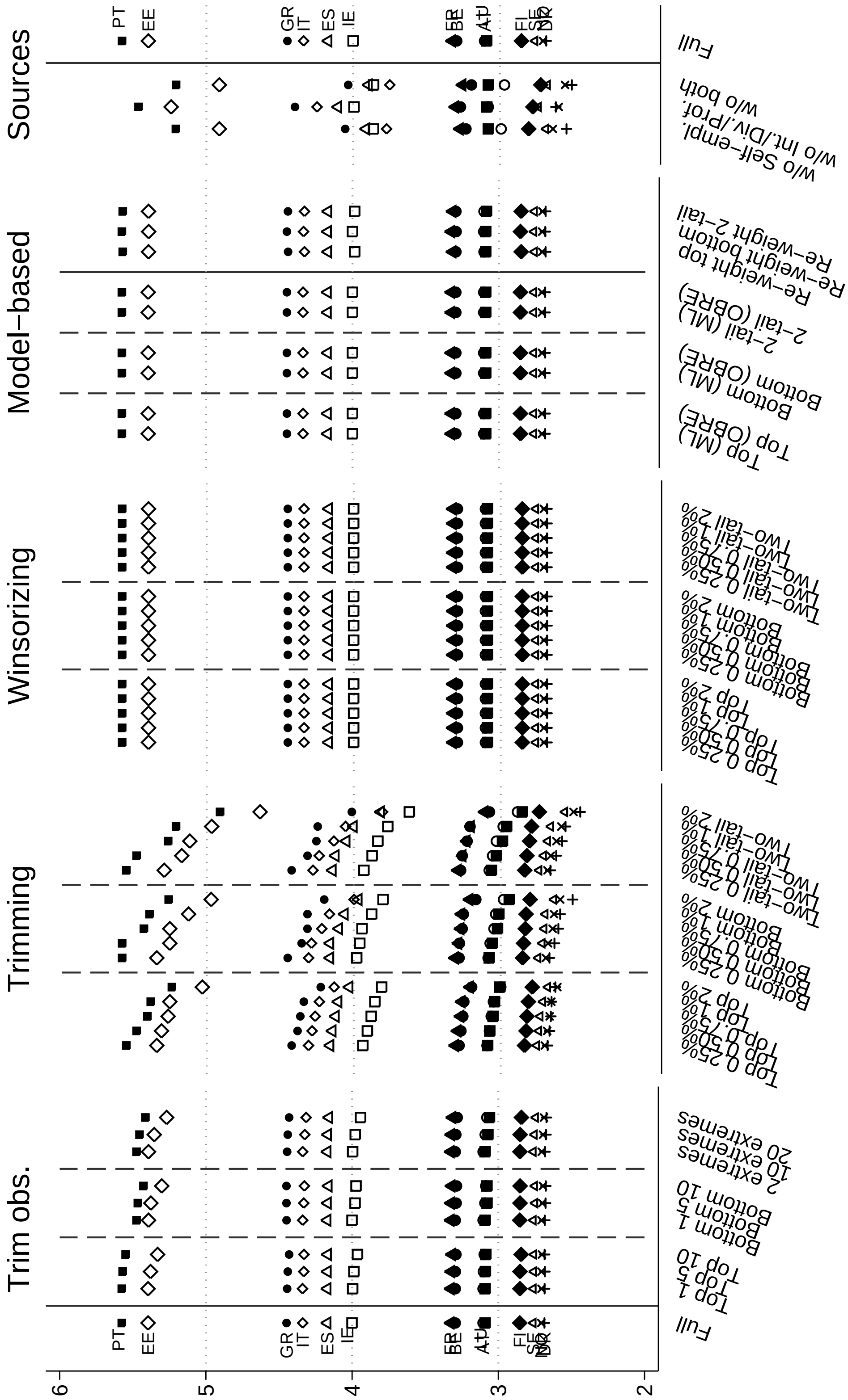


Figure 3.5: Inequality: Estimates of the S80/S20 income share ratio in 14 countries under alternative treatment of extreme incomes

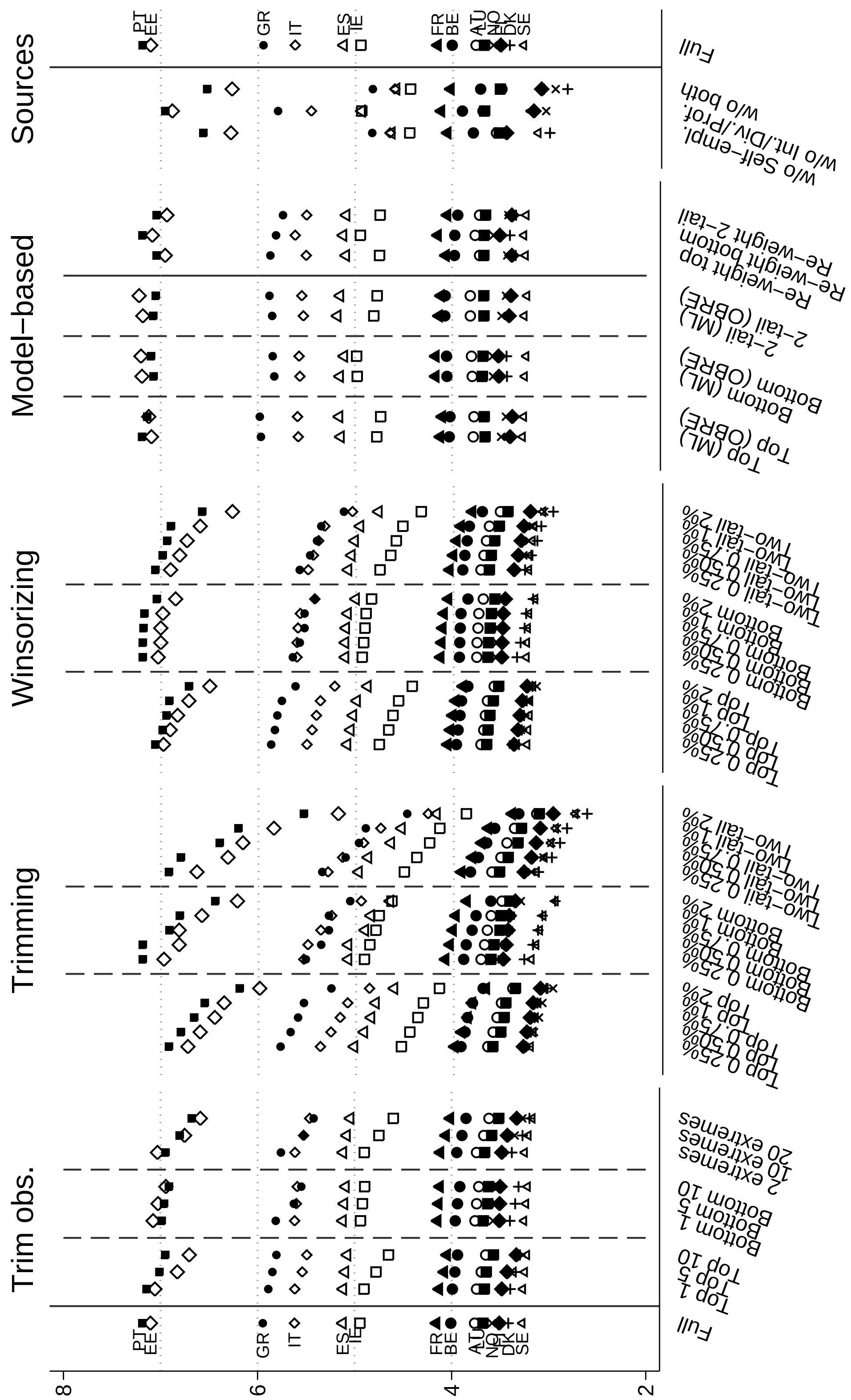


Figure 3.6: Inequality: Estimates of the S90/S10 income share ratio in 14 countries under alternative treatment of extreme incomes

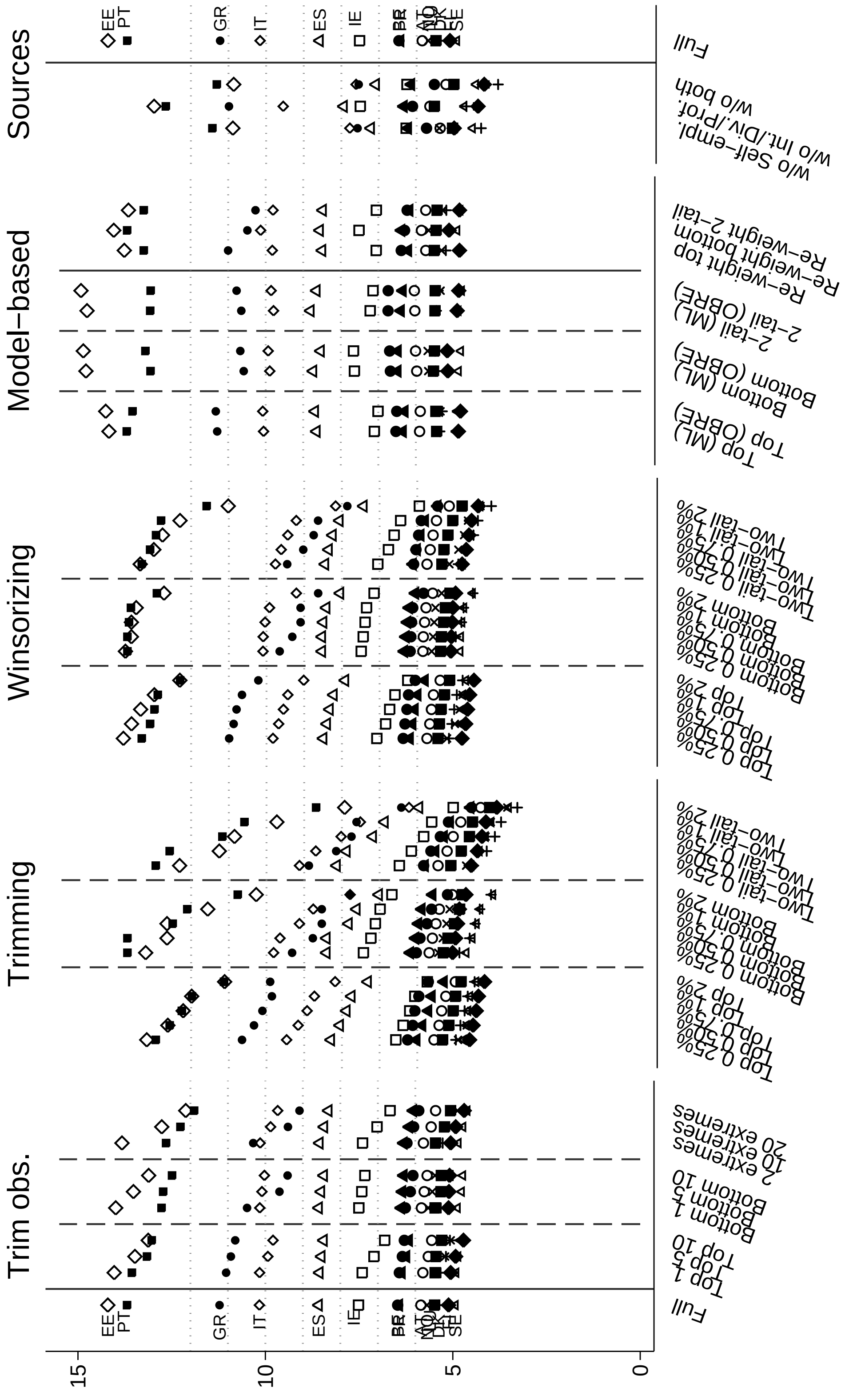


Figure 3.7: Inequality: Estimates of the Gini coefficient in 14 countries under alternative treatment of extreme incomes

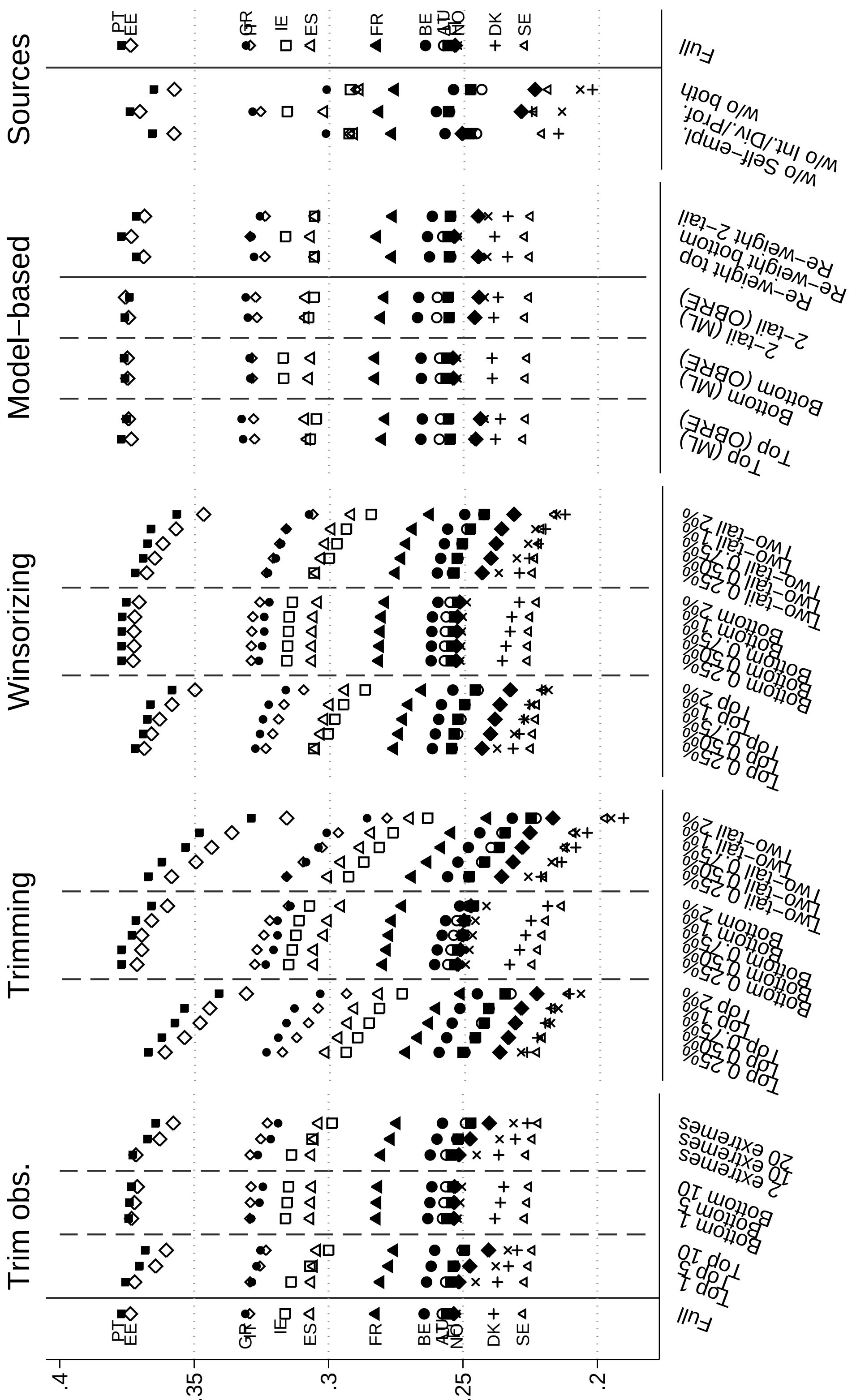


Figure 3.8: Inequality: Estimates of the Generalized Entropy measure with parameter 0 in 14 countries under alternative treatment of extreme incomes

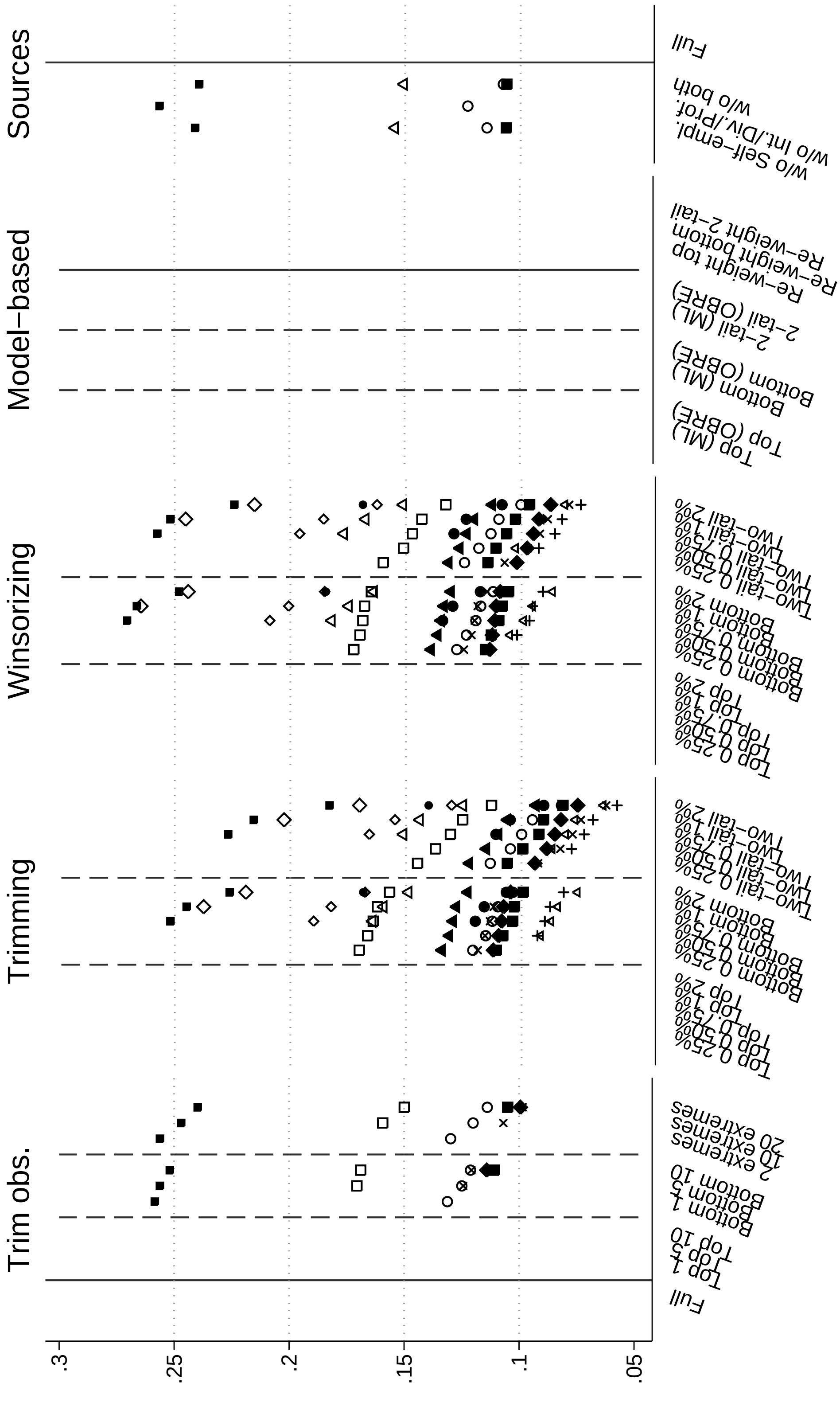


Figure 3.9: Inequality: Estimates of the Generalized Entropy measure with parameter 1 in 14 countries under alternative treatment of extreme incomes

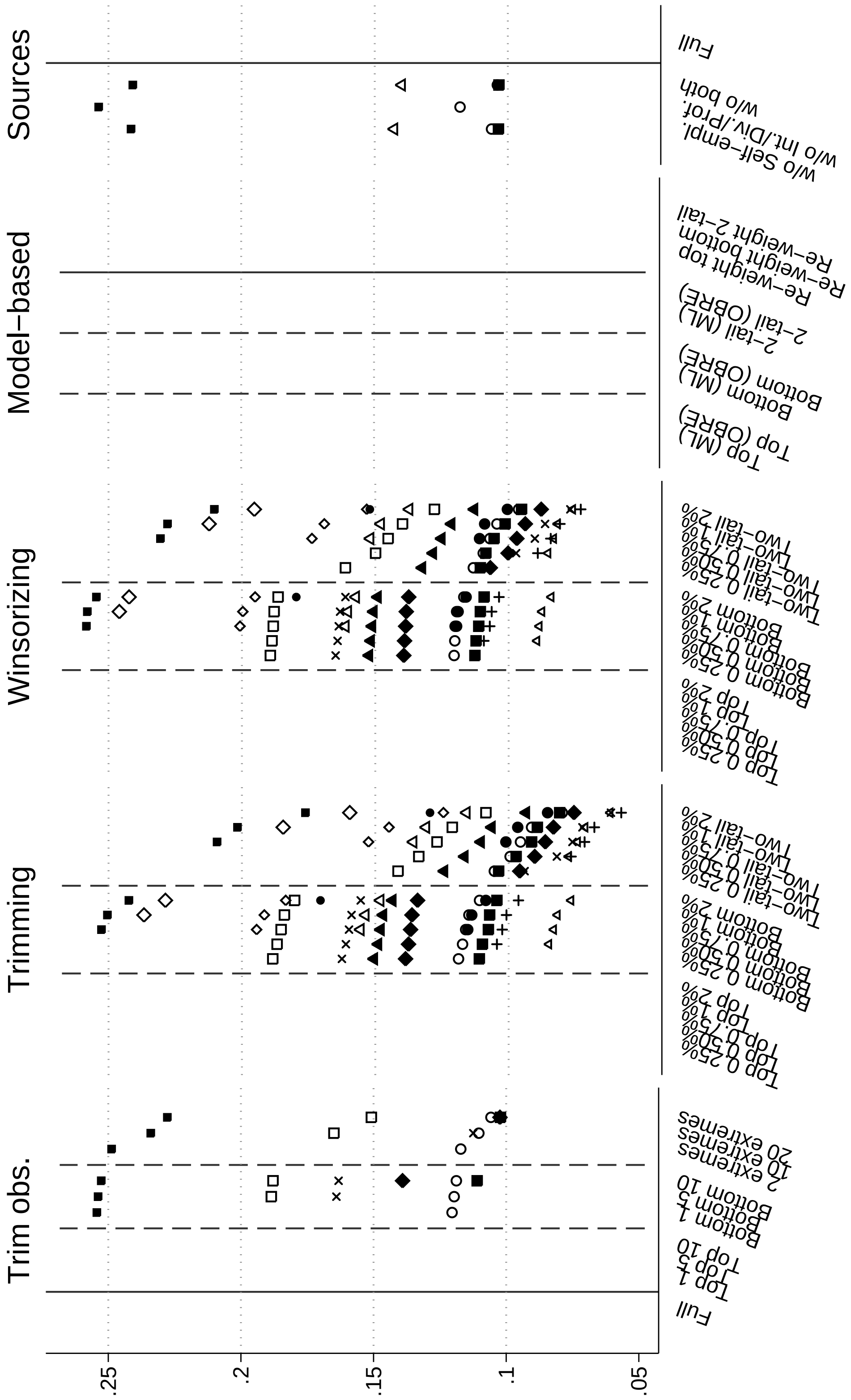


Figure 3.10: Inequality: Estimates of the Generalized Entropy measure with parameter 2 in 14 countries under alternative treatment of extreme incomes

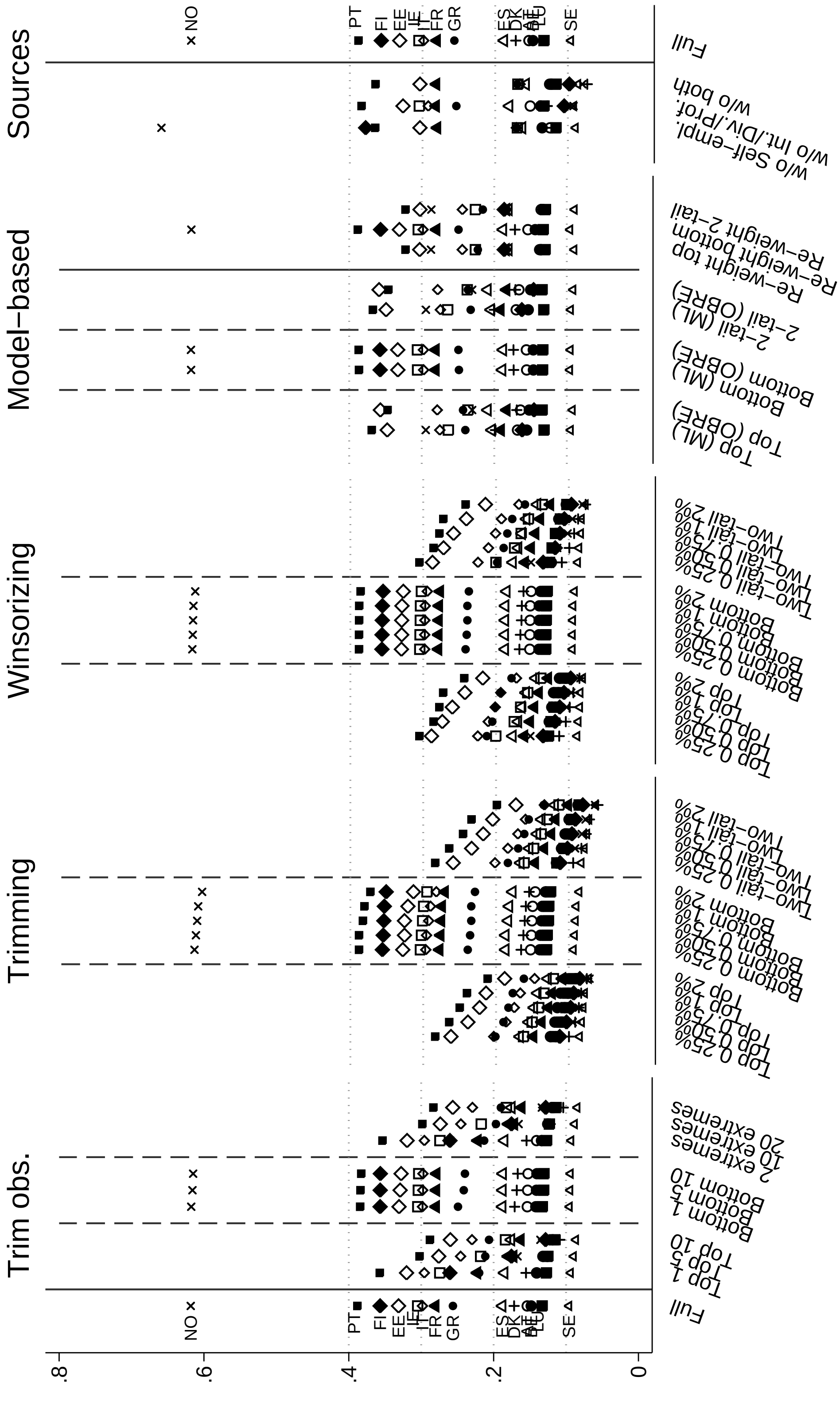


Figure 3.11: Inequality: Estimates of the Atkinson measure with parameter 0.5 in 14 countries under alternative treatment of extreme incomes

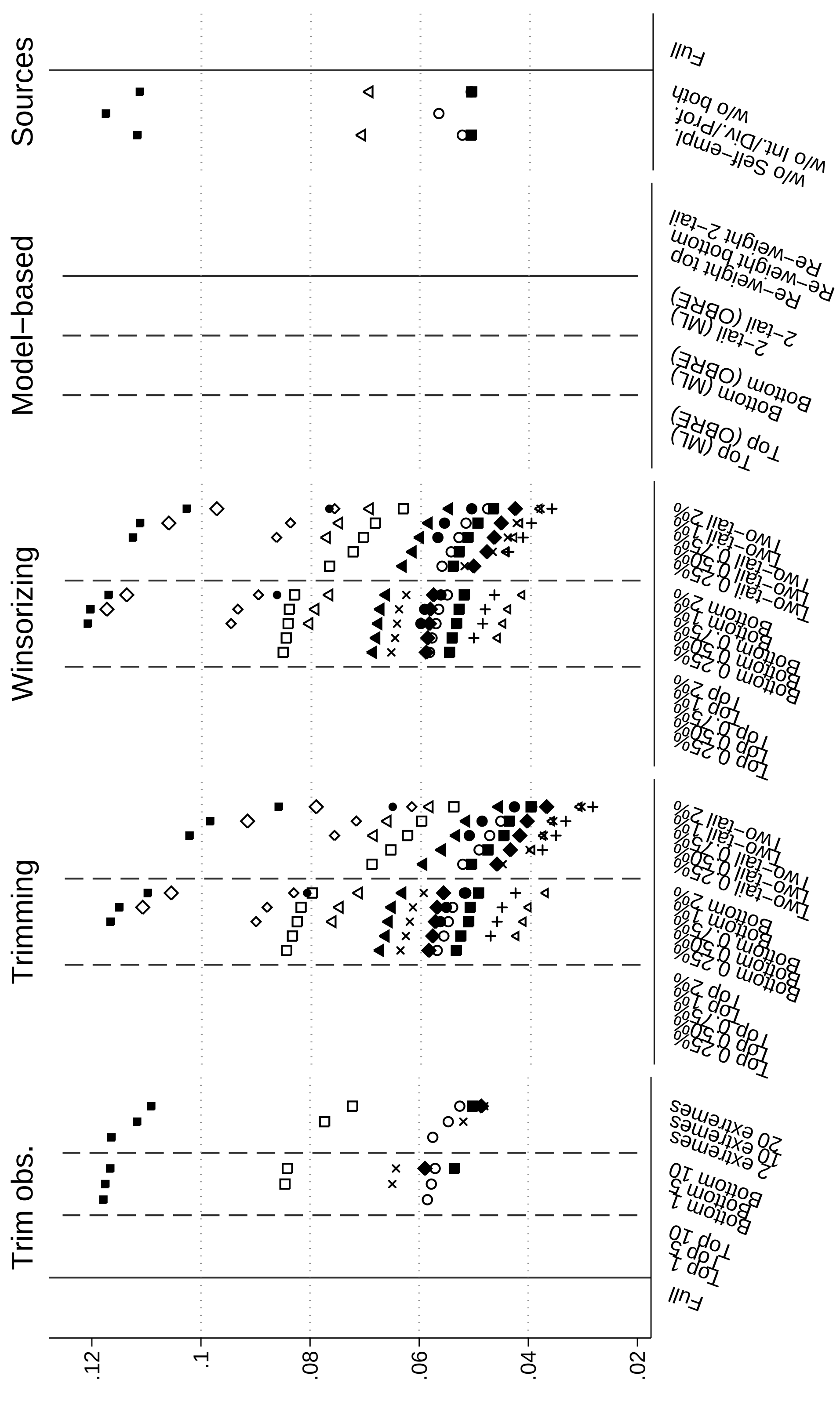


Figure 3.12: Inequality: Estimates of the Atkinson measure with parameter 1 in 14 countries under alternative treatment of extreme incomes

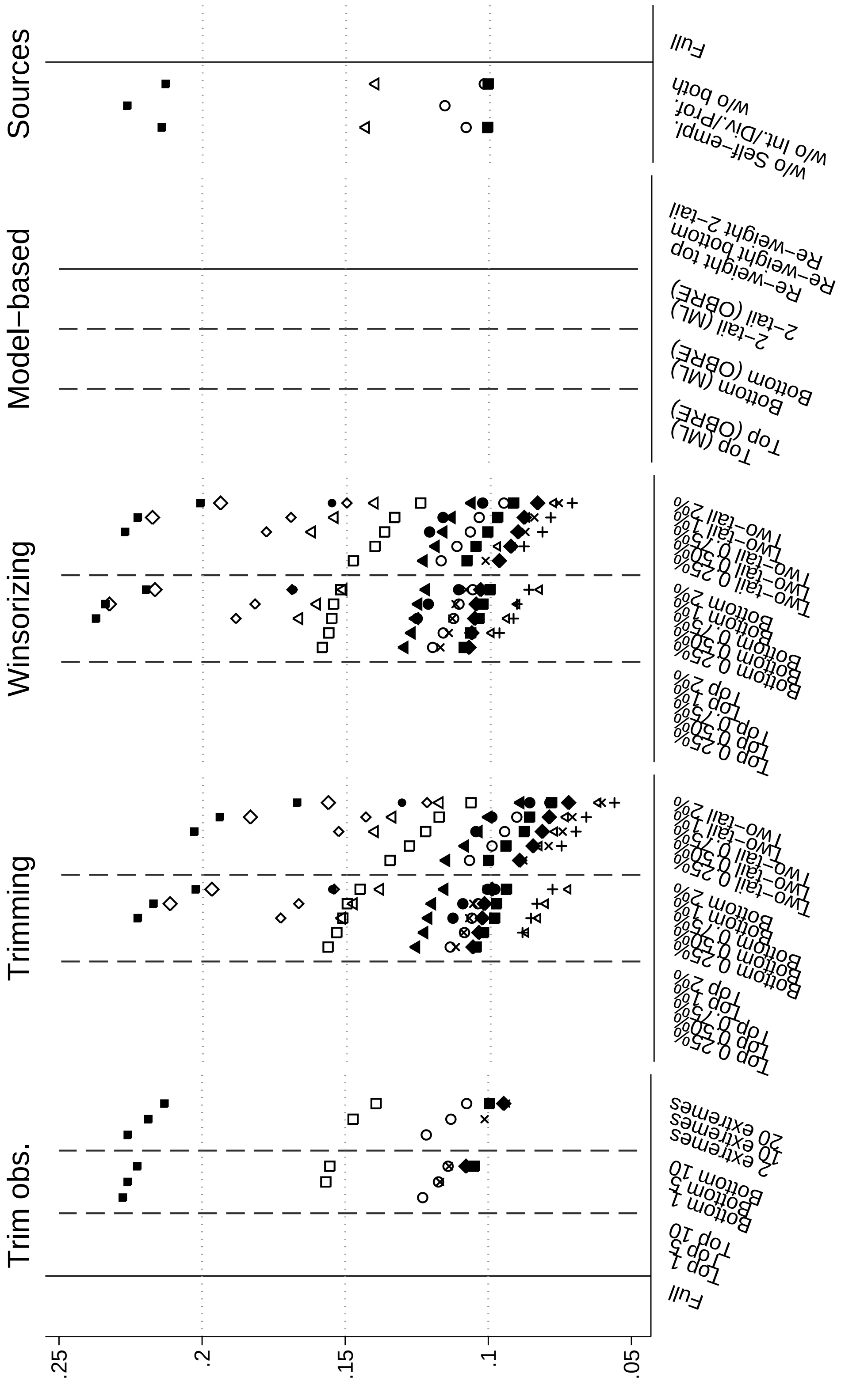


Figure 3.13: Inequality: Estimates of the Atkinson measure with parameter 2 in 14 countries under alternative treatment of extreme incomes

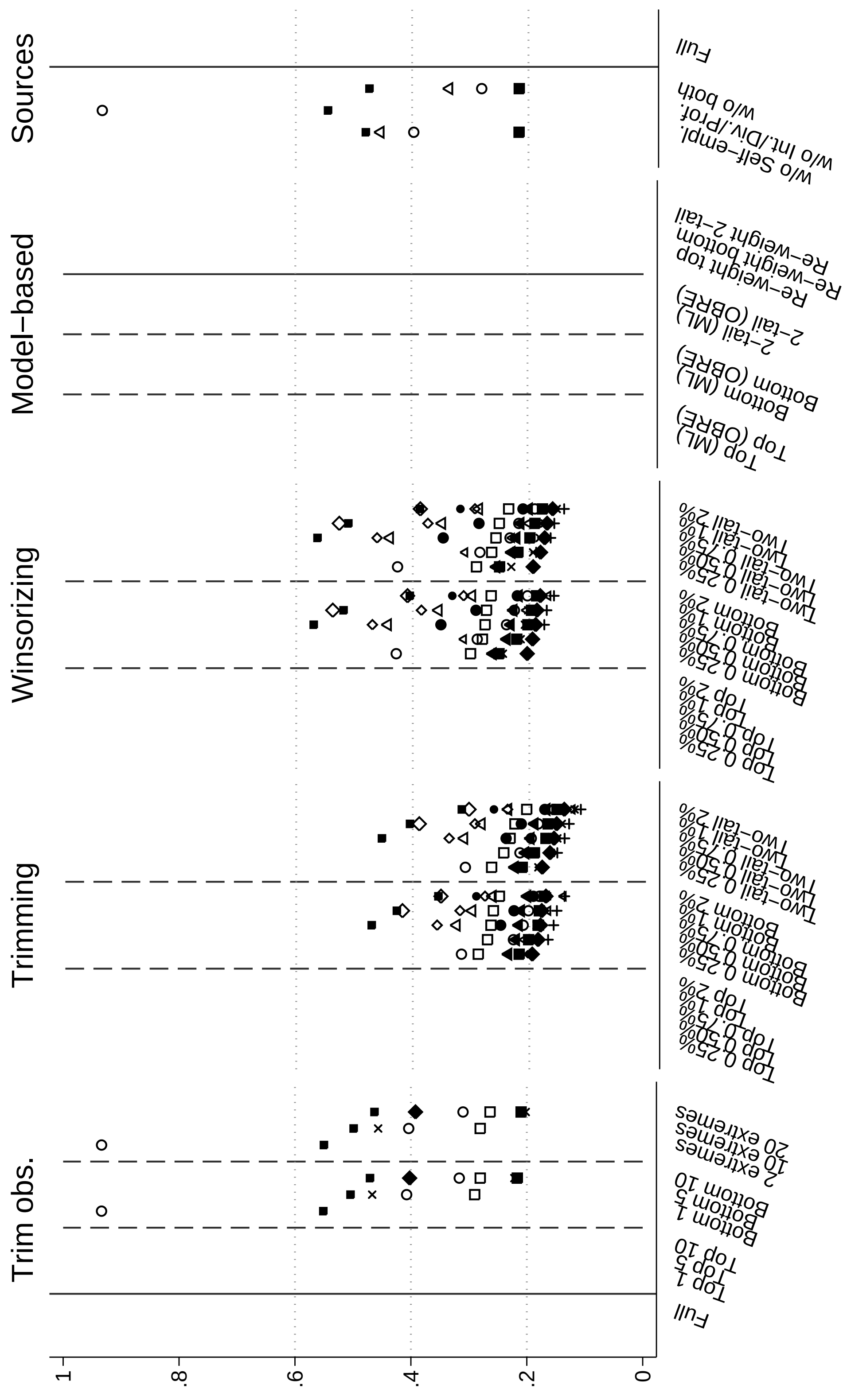


Figure 3.14: Poverty: Estimates of the headcount ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

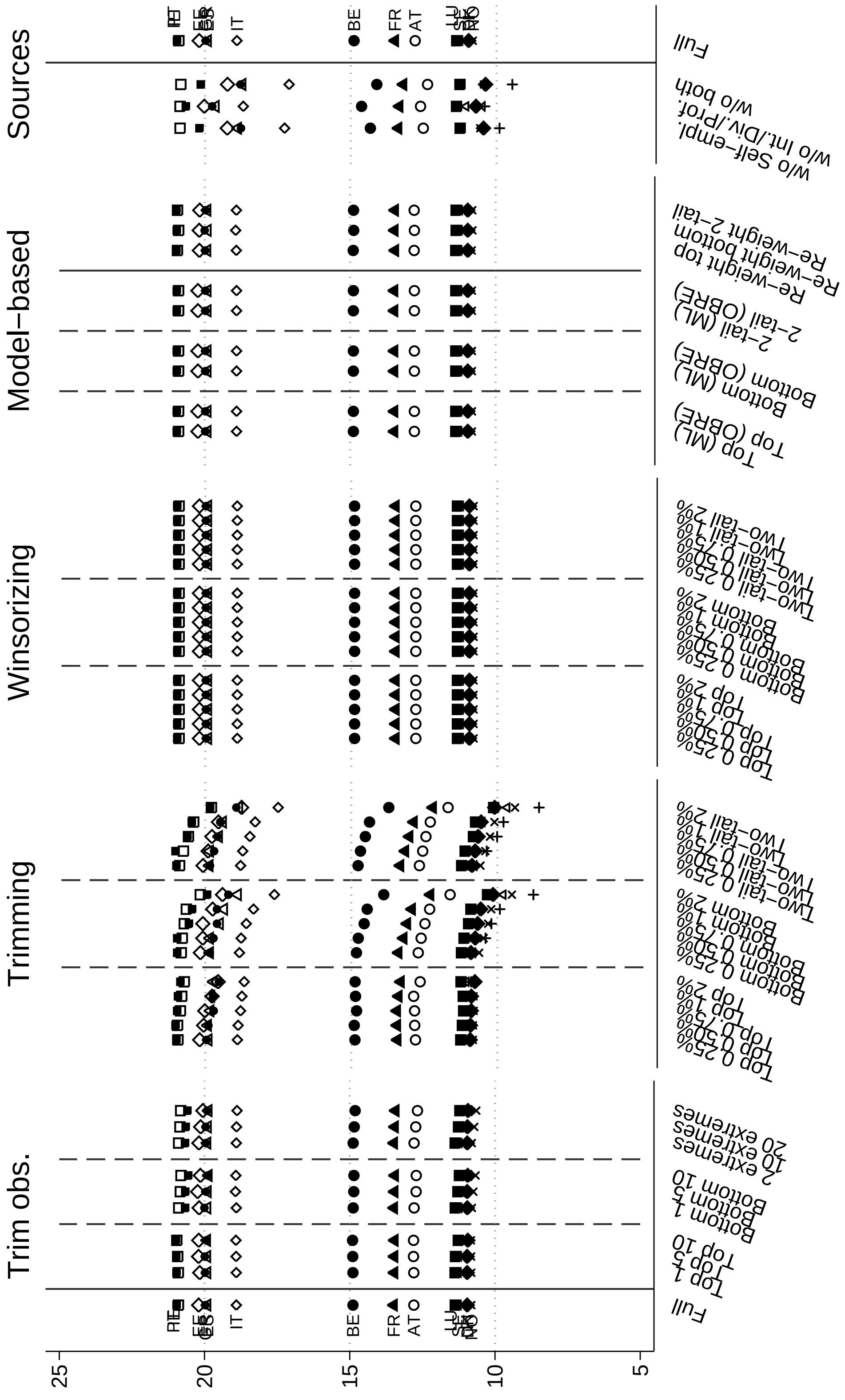


Figure 3.15: Poverty: Estimates of the average poverty gap ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

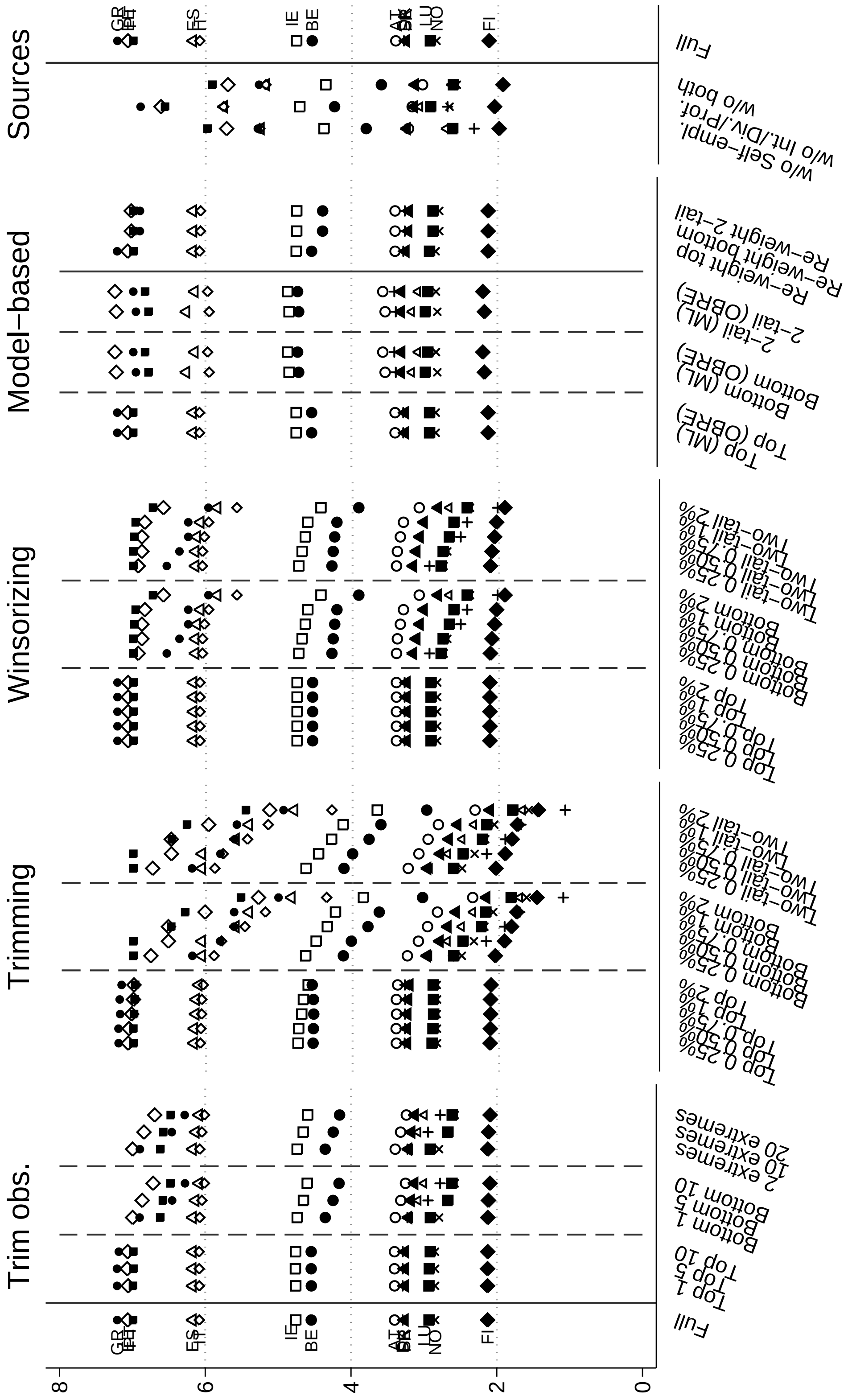


Figure 3.16: Poverty: Estimates of the average squared poverty gap ratio based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

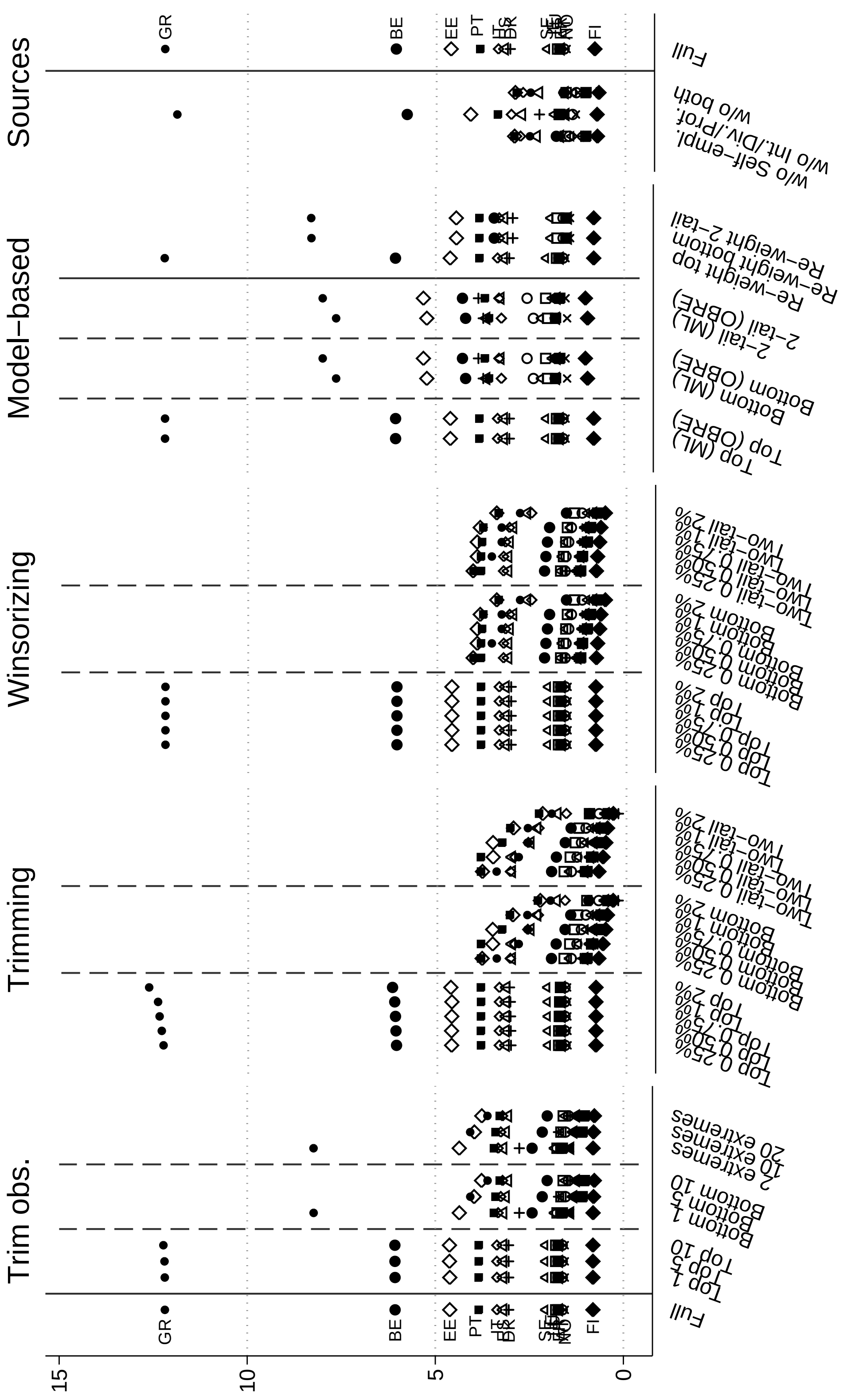


Figure 3.17: Poverty: Estimates of the median poverty gap ratio among the poor based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

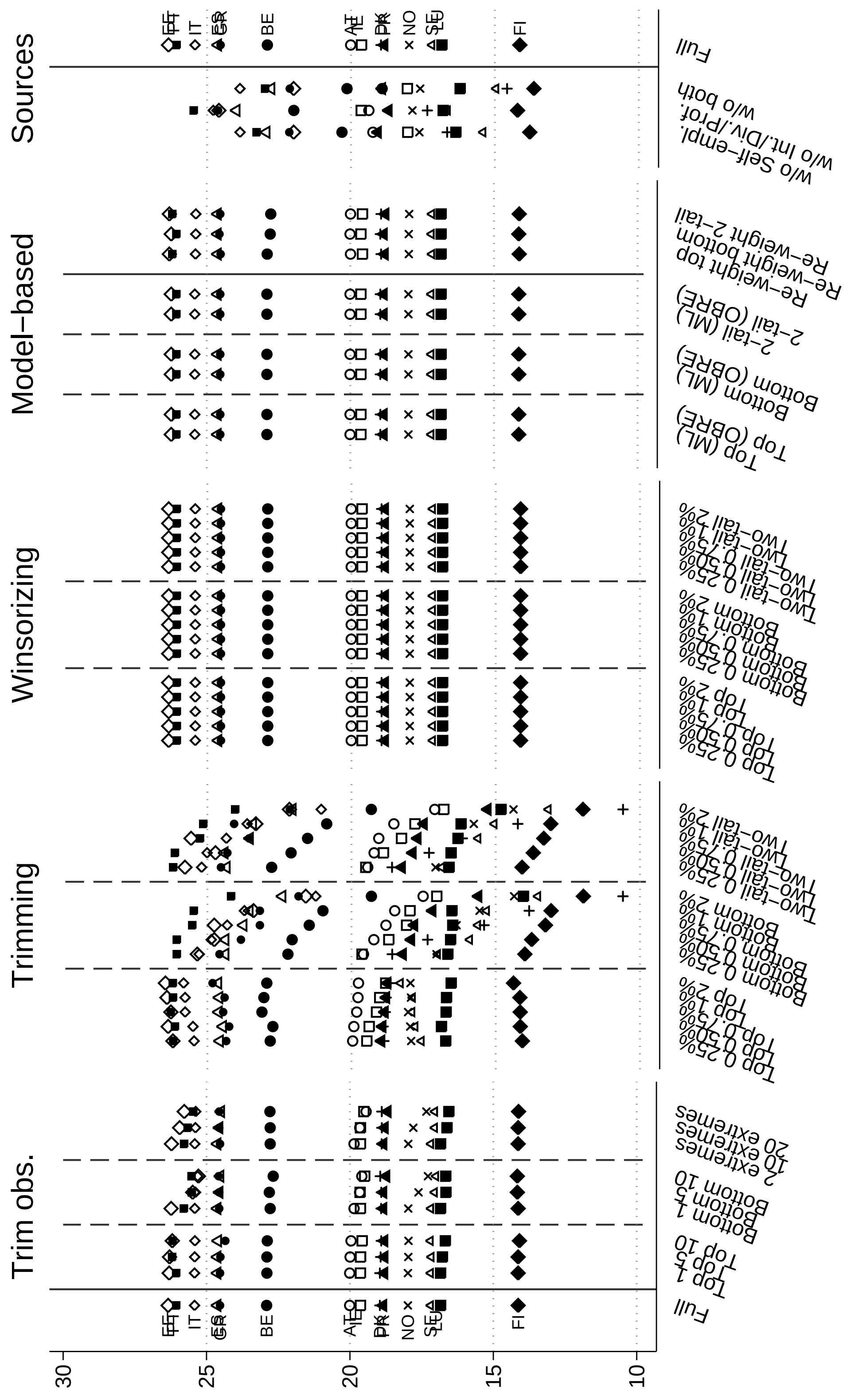


Figure 3.18: Poverty: Estimates of the headcount ratio based on a line at 50% of median income in 14 countries under alternative treatment of extreme incomes

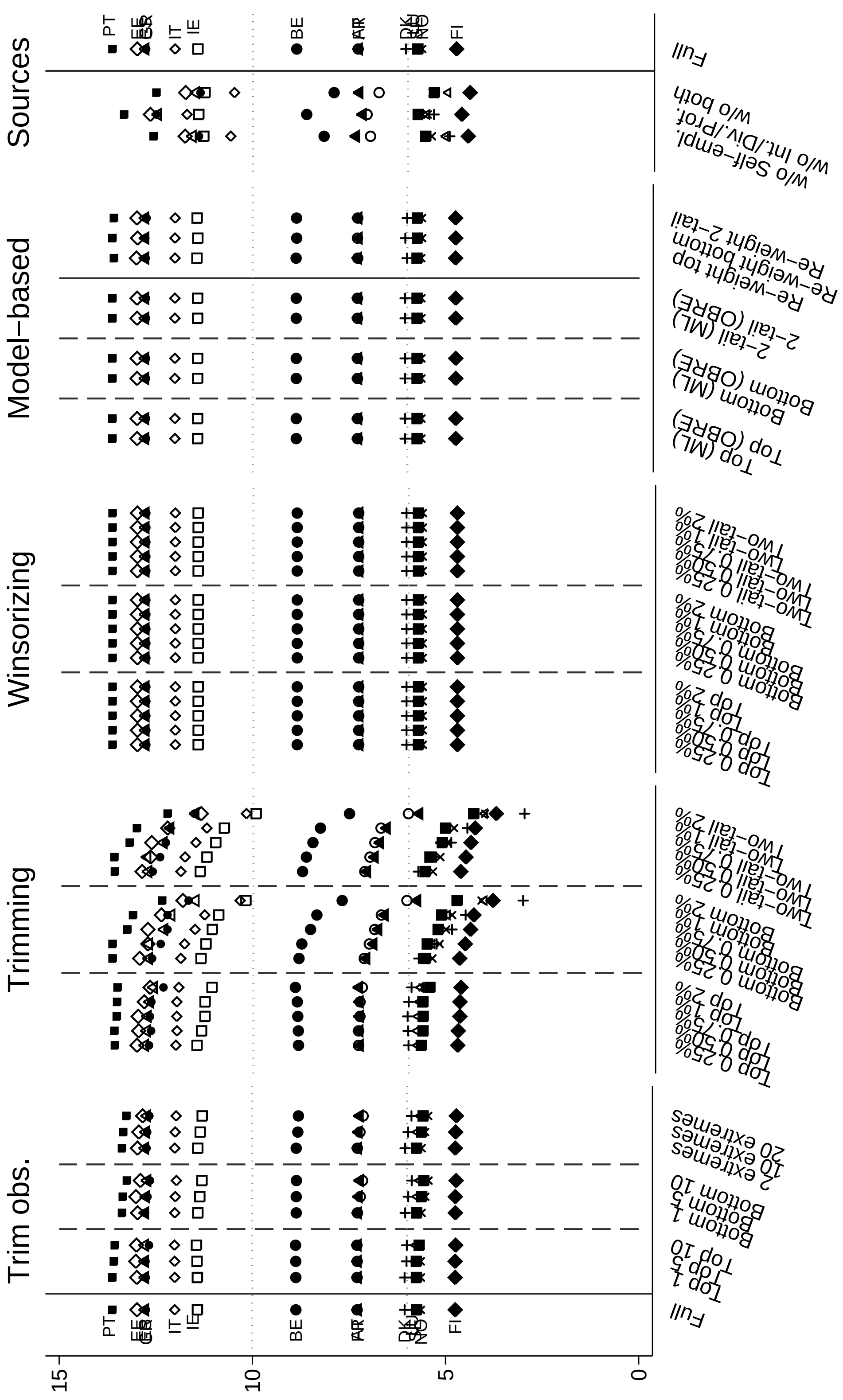


Figure 3.19: Poverty: Estimates of the headcount ratio for households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

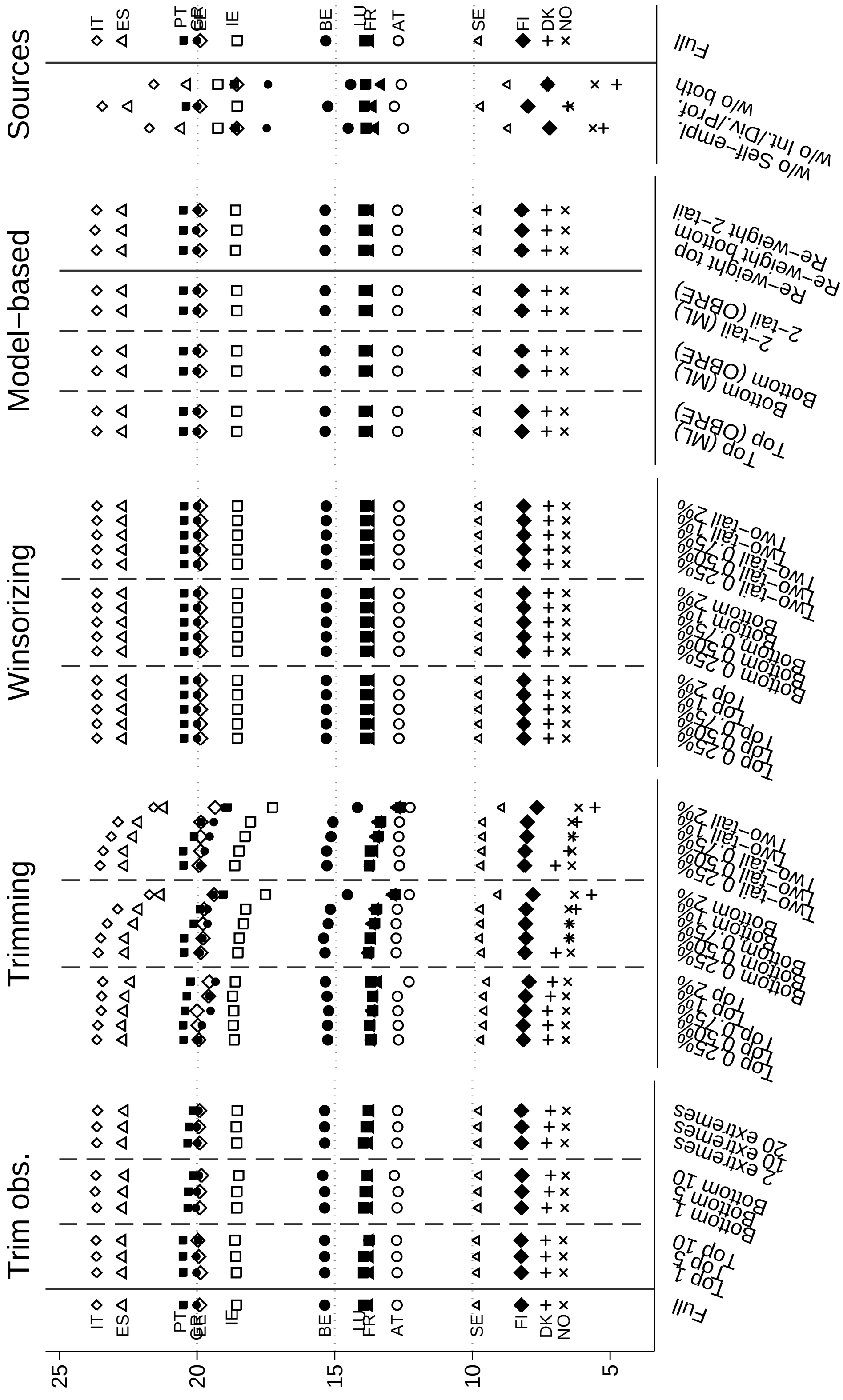


Figure 3.20: Poverty: Estimates of the average poverty gap ratio for households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

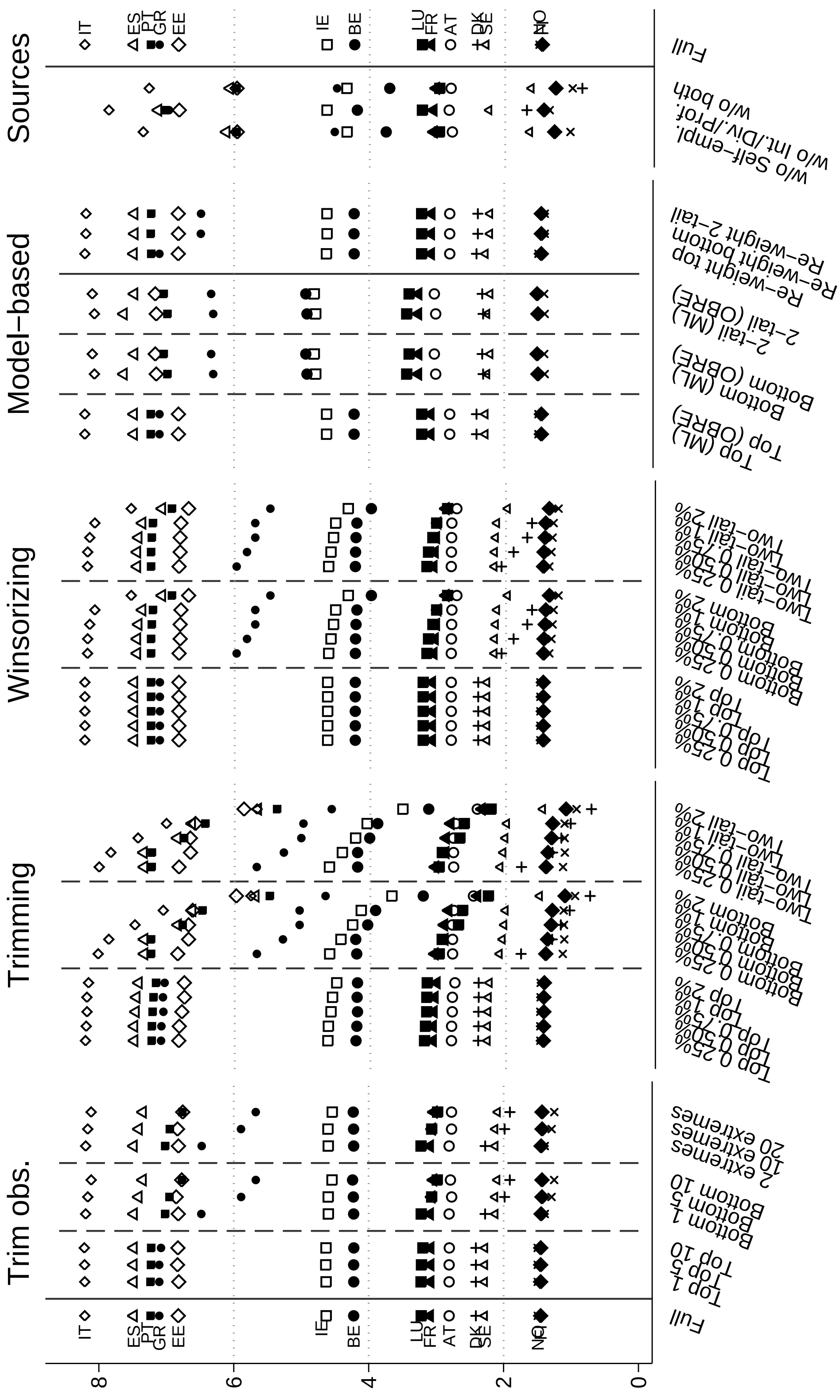


Figure 3.21: Poverty: Estimates of the average squared poverty gap ratio for households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes

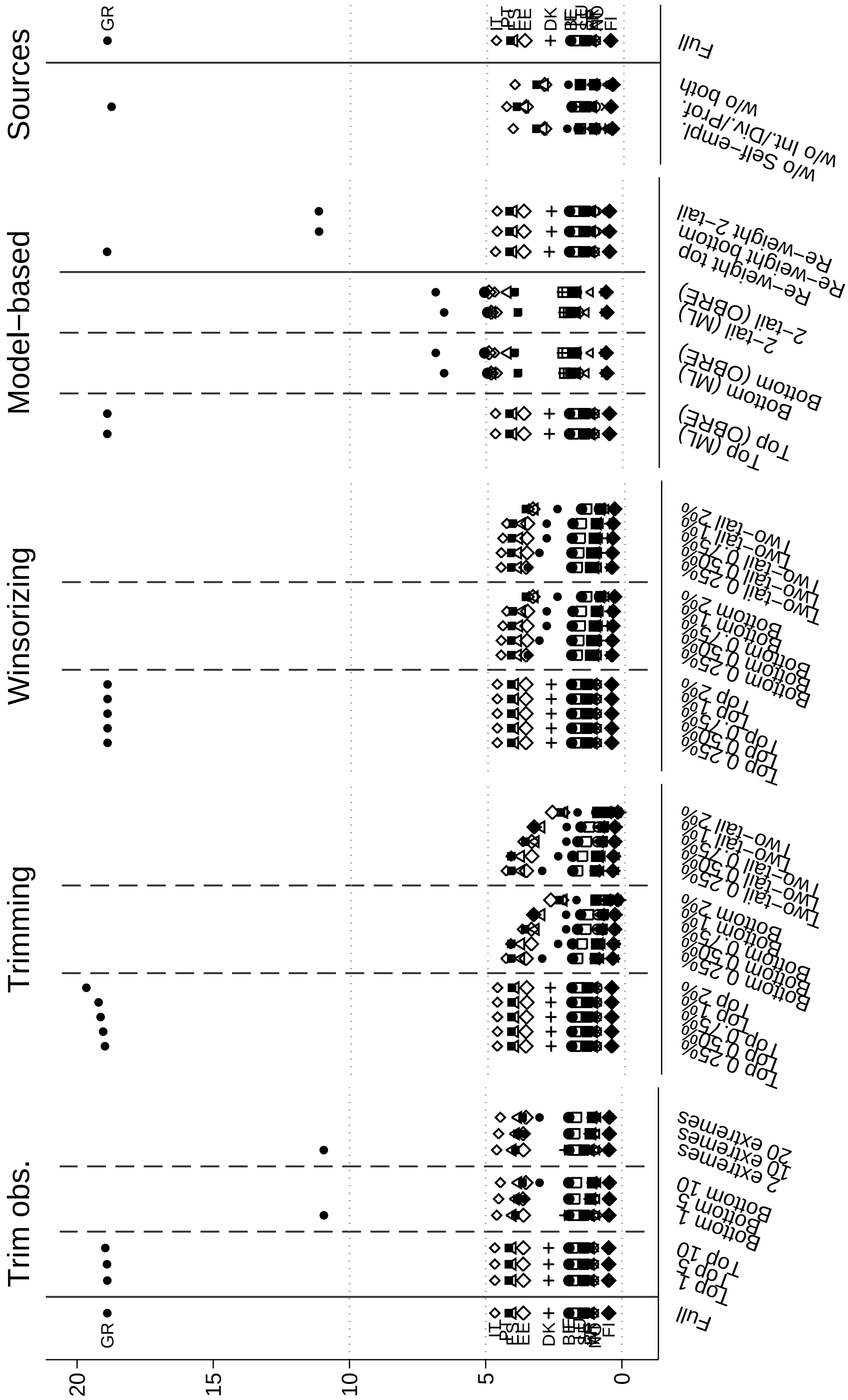


Figure 3.22: Poverty: Estimates of the median poverty gap ratio among poor households with dependent children based on a line at 60% of median income in 14 countries under alternative treatment of extreme incomes