the review of income and wealth

Review of Income and Wealth
Series 69, Number 3, September 2023
DOI: 10.1111/roiw.12591

TOP INCOME ADJUSTMENTS AND INEQUALITY: AN INVESTIGATION OF THE EU-SILC†

BY RAFAEL CARRANZA and BRIAN NOLAN

Institute for New Economic Thinking at the Oxford Martin School, The Department of Social Policy and Intervention, and Nuffield College, University of Oxford

AND

MARC MORGAN

University of Geneva, Geneva, Switzerland. World Inequality Lab, Paris, France

This paper provides a novel assessment of how the World Inequality Database (WID) top income adjustment applied by Blanchet, Chancel, and Gethin (2021) to European Union Statistics on Income and Living Conditions (EU-SILC) data for 26 countries over 2003–2017 for Distributional National Accounts purposes affects inequality in equivalized gross and disposable household income. On average, the Gini is increased by around 2.4 points for both gross and disposable income, with notable differences across countries but limited impact on trends. EU-SILC countries that rely on administrative register data see relatively small effects on inequality. Comparing with two other recent studies, differences in impacts on measured inequality depend less on the adjustment method and more on whether external data sources are used.

JEL Codes: D31, D63, N30

Keywords: inequality, reweighting, survey representativeness, top incomes

1. INTRODUCTION

The most striking finding from research on income inequality over the past couple of decades has been the growing share of total income going to people at the top of the distribution (Atkinson and Piketty, 2007; Atkinson and Piketty, 2010; Alvaredo *et al.*, 2017). Key to this burgeoning "top incomes" literature has been

†Note: This paper has greatly benefited from the suggestions of the editor and two anonymous referees. The authors thank Amory Gethin, Lucas Chancel, and Thomas Blanchet for sharing their data set, and also Charlotte Bartels and Maria Metzing for sharing their code and data. Helpful comments were provided by Facundo Alvaredo, Thomas Blanchet, Stephen Jenkins, and participants of the Second World Inequality Conference. This research was supported by the European Research Council Synergy Grant 75446 for project DINA—Towards a System of Distributional National Accounts.

*Correspondence to: Brian Nolan, Institute for New Economic Thinking at the Oxford Martin School, The Department of Social Policy and Intervention, and Nuffield College, University of Oxford, Oxford, UK (brian.nolan@spi.ox.ac.uk).

the use of administrative income tax data (see Atkinson *et al.* (2011) for an early overview). This literature has called into question the reliance on household surveys in much of the research and official monitoring of inequality, as they may fail to capture incomes at the top of the distribution. Such questioning reflects the difficulties surveys face in capturing a relatively small group in the population, together with specific issues with nonresponse and underreporting among this top group (Hlasny and Verme, 2018; Burkhauser *et al.*, 2018a; Hlasny and Verme, 2021; Blanchet *et al.*, 2022). As a consequence, there are very real concerns that household surveys may mis-measure both inequality levels and trends over time.

For example, in the US, Burkhauser *et al.* (2012) suggest that the Current Population Survey (CPS) closely tracks tax-based top shares up to the 99th percentile but not the top 1 percent income share. Moreover, Atkinson *et al.* (2011) find that a substantial share of the growth in inequality in the US as measured by the Gini coefficient may be "missed" by the CPS. This issue has grown in importance over time, with Yonzan *et al.* (2020) reporting that the gap between surveys and tax data at the very top has been growing in recent years. Similarly, the analysis by Morelli *et al.* (2015) suggests that conventional survey-based measures such as the Gini coefficient may increasingly misrepresent the actual extent of change in income inequality. To add to these concerns, there is every reason to believe that the extent of such bias varies across countries, and this may well be the case for a given country over time: country rankings in terms of inequality levels at a point in time or inequality change over time as seen in surveys may not be reliable. The impact of nonresponse and under-coverage among high-income earners—what Lustig (2019) calls the "missing rich" problem—clearly needs to be addressed in measuring and tracking inequality by official statistical agencies.

Recent research has investigated and employed various approaches to address this problem.[1] These approaches are typically grouped into two categories. First, one can replace the top α percent of the income distribution in the survey with observations drawn from a parametric distribution or an imputation method. Replacement methods assume that population shares (after base survey weights are applied) are correct and focus on issues of underreporting or under-sampling at the top. Second, assuming instead that the population shares in the survey are not correct, one can adjust the entire survey by reweighting, replacing the base weights with new weights that aim to reflect the heterogeneity in nonresponse rates. This is the approach followed, e.g., in Blanchet *et al.* (2022) or Muñoz and Morelli (2021). Reweighting methods are mostly focused on correcting for non-sampling issues such as low response rates among the very top, while replacement methods can also be used to address sparseness at the top, a particularly useful property when measuring income shares at the extreme right tail of the distribution (say, for the 0.1 or 0.01 percent of the population).

Jenkins (2017), in a study on the UK, suggests that fitting a parametric upper tail to the observed survey observations—without reference to external

[1]See the discussion on nonresponse bias and on modeling the top of the income distribution in Hlasny (2020); Hlasny (2021), and the extensive review of existing methods in Lustig (2019)

information—may not be an appropriate adjustment, as estimates may still fail to fully capture the true upper tail. This serves to motivate the use of external information, generally from income tax data, to implement a replacement approach favored by Jenkins (2017). Tax data can also provide information on the spread of incomes across much of the distribution that serves as a basis for reweighting well beyond the top. Yet, combining information from surveys and tax data is challenging in that the two sources mostly employ different income concepts and income recipient units. Survey microdata generally include sufficient information to allow them to match the income concepts employed in external sources, most often on taxable or "pre-tax" income among couples or individual adults (see Yonzan *et al.* (2020) for a detailed exercise of addressing this comparability issue). However, it has proven difficult to relate the results to the more standard definitions generally employed in the inequality literature, notably household income including cash transfers after direct taxes, equivalized and attributed to each person in the household (including children).

Notwithstanding such challenges, certain initiatives have made ground in combining tax-based information and household surveys. A pioneer in the developed world has been the Department of Work and Pensions in the UK, incorporating a top income adjustment to the survey from personal income tax records since 1992, which is reported annually in their *Households Below Average Income* (Burkhauser *et al.*, 2018b). The UK's Office for National Statistics (ONS) has since also revised its income distribution series to include a top-income adjusted series (see ONS (2019); ONS (2020)), influenced by the work of Jenkins (2017) and Burkhauser *et al.* (2018a); Burkhauser *et al.* (2018b).

More recently, efforts associated with the World Inequality Database (WID) have sought to combine data from tax sources, household surveys, and the national accounts to build Distributional National Accounts (DINA). The core aim of this approach is to allocate all the national income as measured in the national accounts to households. This means that items not included in the household income concept surveys seek to measure, such as the undistributed profits of corporations and the benefits from government spending on education and health services, are allocated, with the resultant series being consistent with macroeconomic growth series and accounting for the full distribution of national income (World Inequality Lab, 2020).

The recent study by Blanchet *et al.* (2021) presents DINA series for 38 European countries between 1980 and 2018, on a similar basis to the DINA estimates for the US produced by Piketty *et al.* (2017), for France by Garbinti *et al.* (2018), and for China by Piketty *et al.* (2019). One element in this complex exercise—before bringing undistributed profits, government spending on services, and other sources of income into the picture—is to combine data from tax data and surveys to produce an adjusted distribution of cash incomes. This is done in a two-step procedure that adjusts for both sampling errors, such as sparseness at the top of the distribution, and non-sampling errors, such as low response rates among high incomes, in a manner described in detail later. This serves as the initial building block in the larger exercise of distributing the entirety of national income, but the fact that it is based on distinctive income concepts and units designed for distributional national purposes—as outlined in detail below—makes it difficult to assess the implications of this initial survey adjustment procedure for the standard measures of inequality usually derived from household surveys for gross and disposable equivalized income among persons.

We address this gap using Blanchet *et al.* (2021)'s data set and the central element of their survey adjustment method (hereafter the WID-adjustment) for 26 countries covered by the European Union Statistics on Income and Living Conditions (EU-SILC) between 2003 and 2017. Their reweighting method allows us to construct inequality indicators of equivalized gross and disposable household income among individuals—the concepts most commonly employed for the analysis and tracking of income inequality. The comparison of unadjusted and adjusted income distributions using the standard concepts is a novel research question, which is not analyzed by Blanchet *et al.* (2021). We explore the reweighted data set in more depth, including its distribution across the population, as well as the impact of different units of observation and income concepts, two factors that Callan *et al.* (2020) showed to be crucial in explaining the differences in inequality trends in Ireland. We also compare the results we find with those of other recent studies attempting to adjust inequality measures from the EU-SILC, namely Hlasny and Verme (2018) and Bartels and Metzing (2019), to assess whether the choice of adjustment method really makes a difference. Our analysis allows us to address key concerns about the reliability of survey-based estimates of income inequality that have been highlighted by tax-based estimates of top income shares. To the best of our knowledge, no other paper has delved into the reliability of the EU-SILC to measure the income distribution in such detail.

Among our key findings is that the impact of the WID-adjustment on the Gini coefficient and top income shares for equivalized disposable income among individuals varies widely across the countries in the EU-SILC. The Gini is increased by up to 10 points for some countries but only very modestly for others, affecting country rankings in terms of inequality levels and the gaps between them. The scale of this impact also varies from one year to the next for some individual countries, thus affecting comparisons of trends and inequality rankings, although less substantially for the former. There are also some notable differences between the impacts of the WID-adjustment and those of Hlasny and Verme (2018) and Bartels and Metzing (2019) on the Gini coefficient, demonstrating that the adjustment approach employed does indeed matter—especially the choice between methods that rely on within-sample projection (Hlasny and Verme, 2018) versus those incorporating external information from tax data (Bartels and Metzing, 2019; Blanchet *et al.*, 2021).

The remainder of this paper is structured as follows. In Section 2, we elaborate on why surveys fail to capture top incomes and on the different approaches being employed to address this, and also highlight specific features of the EU-SILC that are relevant in this context, notably the varying use of data from administrative sources in addition to direct responses from those interviewed. In Section 3, we outline the WID-adjustment employed in Blanchet *et al.* (2021) and describe the distinctive income concepts on which their analysis is focused. In Section 4, we then employ the core of their adjustment to the EU-SILC microdata and see the impact this has on inequality in equivalized gross and disposable income among persons. This section also investigates the role of the income concept and unit of analysis employed, and digs deeper into the mechanics and impact of the reweighting process involved in the WID-adjustment. Section 5 compares the impact of these adjustments on the Gini coefficient for equivalized disposable income with those presented in Hlasny and Verme (2018) and Bartels and Metzing (2019). The final section concludes with a summary of the main findings and discussion of their implications.

2. Surveys and the Coverage of Top Incomes

Surveys fail to capture the top of the income distribution for different reasons. These reasons are typically grouped into sampling and non-sampling issues. The former reflects problems with the original design of the survey, e.g., how a small sample size can result in sparseness of certain population groups. The latter reflects heterogeneous response rates, e.g., when those at the top of the distribution decline to respond to the survey, after being included in the sample, more than those among the rest of the distribution. As a result of these problems, the gap between surveys and administrative data is particularly large at the top (see, e.g. Burkhauser *et al.* (2018a) for the UK, Blanchet *et al.* (2022) for France, the UK, Norway, Brazil, and Chile, or Lustig (2019) for Uruguay) and it has been growing over time for countries like Ireland (Callan *et al.*, 2020) and the US (Yonzan *et al.*, 2020). Administrative data like income tax records are not without their own limits. Despite the data not being based on samples, but on a universe of (often third party reported) declarations, the share of the total population covered can vary substantially by country. Moreover, the degree to which tax evasion affects the quality and reliability of tax data should also be expected to vary by country. However, in all cases, tax data can be at least considered a reliable lower bound of the upper tail of the distribution.

Sampling issues are mostly associated with non-coverage error and with sampling error. Non-coverage errors happen when, by design, individuals have zero probability of being selected into the sample. Most statistical institutes design the sampling strategy so as to avoid non-coverage errors, e.g., by replacing the population that cannot be covered, and it is not usually a major issue. Sparseness, on the contrary, means that there is insufficient density at the top of the income distribution and therefore very few observations for that group. This might not necessarily bias inequality estimates, but will reduce their reliability. This creates a problem when estimating top income shares, particularly those at the very top (the top 0.1 or 0.001 percent). These issues can be resolved at the design stage, e.g., by over-sampling the relevant population, or subsequently by replacing the top of the distribution with estimates from a parametric model or from linked administrative data.[2]

Non-sampling issues reflect differences in behavior among the surveyed or in choices taken by survey administrators. It includes both unit and item nonresponse, as well as underreporting and top coding. Unit nonresponse happens when

[2]The replacement approach expands the sample for the top α percent while keeping everything else consistent, including weight totals by replacing the original weight w for w/n where n is the number of times that observation was expanded.

individuals in the potential sample do not respond. Similarly, item nonresponse happens when respondents opt to not answer income questions. This is often addressed by hot-deck imputations, which replace an individual's missing value with the value observed for people with similar observable characteristics. If the population not responding is not missing at random, income imputations may worsen the situation, potentially mirroring the situation when they do answer but underreport the amount. Finally, top coding happens when incomes are censored (or truncated) above a certain threshold, usually to protect the anonymity of very high income respondents. Under non-sampling issues the final sample will differ from the original sample design and if the difference is associated with income—e.g., if high income earners are more likely to not answer the survey—inequality estimates in the survey will be biased.

Several solutions have been proposed to address these issues. Some solutions focus on adjusting inequality estimates, while others aim to adjust the survey itself. The former approach combines an inequality estimate (say, the Gini index) for the poorest $1 - p$ percent derived from the survey with another inequality estimate for the richest p percent which can be estimated in various ways, e.g., using random draws of a Pareto distribution, estimated using either survey or tax data, resulting in a semi-parametric estimate (Jenkins, 2017). The two Gini estimates are combined following the approach of Atkinson (2007), later extended by Alvaredo (2011). As a result, this first solution provides an inequality estimate that addresses both sampling and non-sampling issues.

The second solution is to adjust the survey itself, either by replacing the top of the distribution or by reweighting the survey. Replacement, as the name suggests, seeks to replace the top of the distribution with a more representative distribution of top incomes. This could be using cell-based means drawn from tax data, as in Burkhauser *et al.* (2018a), or random draws from a parametrized Pareto distribution, as in Bartels and Metzing (2019). While the replacement does not modify or alter the rest of the distribution, a reweighting approach by contrast adjusts the whole distribution. The reweighting approach adjusts the survey weights to address nonresponse rates, so that the new weights match a certain reference point. (Korinek *et al.*, 2005; Korinek *et al.*, 2007) use the data on average response rates by groups such as geographic areas, as does Hlasny and Verme (2018). Alternatively, Blanchet *et al.* (2022) and Blanchet *et al.* (2021) use external data on top incomes to address both nonresponse and underreporting of incomes via a combination of reweighting and replacement. The resulting outcome of the replacement and reweighting approaches is an adjusted survey, including individuals and households, such that one can estimate different inequality indexes such as top income shares or the Gini index. While the two approaches can result in similar outcomes (i.e., inequality levels), there are differences between the two that might be relevant, depending on what the researcher wants to estimate. For example, the reweighting approach does not modify the maximum income in the survey, while the replacement can do so. Conversely, the reweighting does not modify the number of respondents nor their individual characteristics (to the extent that no observation is assigned a weight of zero).

In the specific context of top income adjustments applied to data from EU-SILC, substantial variation across countries in survey sampling, implementation, and how the income data are produced must be kept to the fore. EU-SILC was launched in 2003 and extended to all EU member states and some associated countries over time so that by 2020 it was implemented in 37 countries, i.e., the 27 EU countries, Iceland, Norway, Switzerland, the UK, Albania, Kosovo, Montenegro, Northern Macedonia, Serbia, and Turkey. Crucially, EU-SILC is based on a common "framework," as opposed to being a common "survey." This framework consists of common procedures, concepts, and classifications, including harmonized lists of target variables to be transmitted to Eurostat, which are made available to researchers for analysis in the form of microdata subject to restrictions and conditions. Data are collected from probability samples of the population residing in private households within the country (irrespective of nationality or legal residence status), with sampling frame and methods of sample selection differing across countries but aiming to ensure that every individual and household in the target population is assigned a known nonzero probability of selection.

Measuring income is a central aim of EU-SILC, and this is done in terms of a substantial set of specific components of income, mostly at the individual level, but some at household level. The income reference period is generally the previous calendar year.[3] The EU-SILC framework encourages the use of existing sources and/or administrative data, and it is key here to distinguish three different situations:

1. In the countries commonly referred to as "register" countries, information on income (as well as some demographic variables) is obtained through accessing administrative registers, while other personal variables are obtained according to the "selected respondent model" where only one member of the household answers the detailed questionnaire.
2. Some other countries have moved over the course of EU-SILC to retrieving at least some income information from registers, but without moving to the selected respondent model.
3. In the remaining countries, all the information on income is obtained by means of survey responses.

One would expect, in general, that drawing on administrative information—for the most part from income tax and social security records—would improve accuracy in the measurement of income, and this has been validated in general terms with respect to EU-SILC in, e.g., Törmälehto *et al.* (2017). For example, Burricand (2013) finds that between 2007 and 2008, when France started using register data, the average disposable income increased by 15 percent and the Gini increased by five points, particularly due to differences in real estate and asset income. However, a degree of complexity, and indeed uncertainty, arises when it comes to assigning participating countries to these categories.

The first category is generally referred to in the literature on EU-SILC as the "old register countries." This includes not only the Nordic countries—Denmark, Finland, Iceland, Norway, and Sweden—which traditionally rely for many purposes on comprehensive population registers incorporating data from a variety of

[3]Two exceptions are Ireland and the UK. In the former, the income reference period is 12 months before the month of the survey, while in the latter the current income is annualized and aims to refer the current calendar year, i.e., weekly income is multiplied by 52, and monthly by 12.

sources and thus are generally termed register countries, but also the Netherlands and Slovenia. Complications arise principally with respect to the second category, sometimes—though perhaps somewhat misleadingly—referred to as "new register countries." Over the life of EU-SILC, an increasing number of countries have been combining survey data with some administrative/register data. The extent and nature of the use of administrative data vary widely, and for individual countries may differ across income components and change over time, in a fashion that sometimes cannot be traced satisfactorily from the information provided by Eurostat or national statistics offices. A particularly unclear issue is whether investment income data from tax records is also used alongside administrative data on employee and self-employed income. This all means that the correct categorization of certain countries depends on the specific year being considered, and more generally has given rise to some confusion and variation across studies as to which countries belong in which category.

To try to clarify this insofar as possible, we draw on various studies including Törmälehto *et al.* (2013), Törmälehto *et al.* (2017), and Goedemé and Trindade (2020), from which the following information can be collated:

- For Austria, the transition to using register data was fully implemented in EU-SILC 2012, and Statistics Austria subsequently revised the EU-SILC data sets for 2008–2011.
- For France, the transition to register-based income data took place in 2008.
- For Italy, some register information has been used since 2004.
- For Spain, the use of administrative data was implemented in EU-SILC 2013.
- For Switzerland, register data were being used in 2011, but the timing of introduction is not clear.
- For Cyprus, Estonia, Latvia, and Malta, some register data were being used in 2018, but the timing of introduction is not clear.
- For Belgium, the transition to using tax and transfer data from administrative sources has been implemented from EU-SILC 2019 onward.
- For Ireland, some income data from social transfer sources have been drawn on from the start of EU-SILC where survey respondents agree, but the extent of use of register data has increased substantially over time, especially from around 2010 when administrative data on employee and self-employed income started to be drawn on.

On this basis, it seems that Italy can be assigned to the second category throughout, as can Austria and France from 2008, Spain from 2013, and Belgium from 2019. Switzerland, Cyprus, Estonia, Ireland, Latvia, and Malta can also be assigned to that category for recent years, but their situation in earlier EU-SILC years is unclear. Moreover, the heterogeneity among countries (or country-years) included in the second category in terms of how administrative data is actually drawn on must be emphasized. For the remaining countries—Croatia, Czechia, Germany, Great Britain, Greece, Hungary, Lithuania, Luxembourg, Poland, Portugal, Romania, Serbia, and Slovakia—it appears that the use of administrative data on incomes is minimal or nonexistent, although it is not always possible to be sure about this from the available documentation. The nuances of timing and variations in the nature of the use of register data will be important when we consider the extent to which the impact of the top income adjustment on EU-SILC data depends on the extent to which register data are drawn upon.

3. The WID-Adjustment to the EU-SILC

Having outlined the issues that arise in seeking to address "the missing top" in surveys, we now turn to the WID-adjustment to the EU-SILC that provides an appealing way to do so. We begin by describing this adjustment process and the assumptions it involves. We then outline the differences between the DINA context in which Blanchet *et al.* (2021) apply this WID-adjustment and the more long-standing "traditional" research literature on income inequality, particularly with respect to the income concepts and units of observation employed. Blanchet *et al.* (2021) construct DINA for 38 European countries between 1980 and 2017, drawing on multiple data sources and methods to construct comparable series over time, including EU-SILC for the period that covers. To create these DINA, the exercise presented in Blanchet *et al.* (2021) can be divided into two major steps. The first is to adjust the survey so that the top 1, 5, and 10 percent income shares as measured in the survey match those reported using tax data. The second step is to add in and allocate the other income components required to go from adjusted survey income to national income, such as imputed rents, government spending, and undistributed corporate profits. For our purposes it is only part of the top income adjustment that is relevant (concretely, the reweighting adjustment), but the broader context is also essential to understand how that adjustment is employed by Blanchet *et al.* (2021) versus here.

As employed in Blanchet *et al.* (2021), this top income adjustment combines reweighting and replacement of the top of the income distribution, using methods such as the one developed and illustrated in Blanchet *et al.* (2022).[4] The replacement step is applied after the reweighting step to gain more precision for indicators that are sensitive to small groups at the top of the income distribution, such as the top 0.1 or 0.01 percent income share.[5] In that step, the income distribution from tax data is replaced into the reweighted survey distribution, thereby increasing the number of observations at the top of the distribution. Because we wish to work with the original survey sample size, and do not focus on very small top income groups but rather on synthetic indicators like the Gini and at most the top 1 percent income share, our analysis focuses only on the reweighting step of the WID-adjustment. As shown in Blanchet *et al.* (2022), this step accounts for the overwhelming majority of the overall adjusted population, which in European countries tends to be very small (e.g., less than 1 percent of the population is absent from the top in surveys when

[4]Blanchet *et al.* (2022) develop a top income adjustment method that uses both reweighting and replacing to reconcile household surveys and tax data for people at the top of the income distribution, providing an application for France, the UK, Norway, Brazil, and Chile.

[5]Statistically, the survey after reweighting should be indistinguishable from the tax data in terms of averages and variance for that part of the distribution as we are exactly replicating the tax data. See Section 2 in Blanchet *et al.* (2022) for details.

compared to tax data). While only using the reweighting step of the adjustment, we continue to employ the "WID-adjustment" labeling for convenience.

3.1. *Survey Reweighting*

The reweighting process calibrates the EU-SILC weights so that the top 1, 5, and 10 percent income shares match those previously estimated using administrative tax data. These shares are available in the WID, complemented with additional top income share estimates from newly collected data. (See the https://wid.world/document/technical-appendix-to-why-is-europe-more-equal-than-the-united-states-world-inequality-lab-wp-2020-19/extended online appendix of Blanchet *et al.* (2021) for a detailed description of the country-by-country adjustments.) The purpose of this exercise is to maintain the survey structure intact while simultaneously adjusting the income distribution, allowing for any type of analysis that relies on survey data. Note the reweighting approach on its own would not allow for the survey to match top income groups at the very top of the income distribution such as the top 0.01 or 0.0001 percent, as these groups might not have been captured by a survey across every country or year.

The reweighting aims to address the gap between the survey top income estimates and the tax-based estimates. Blanchet *et al.* (2021) model under-coverage rates as a linear function with kinks at each relevant threshold (top 10 percent, top 5 percent, etc.), such that rates increase with income, allowing for much larger rates at the very top of the distribution. The authors interpret their model as addressing both sampling errors and non-sampling errors. In practice, they are indistinguishable unless one has access to individually matched income data across the two sources, which is quite rare. The authors note that the estimated under-coverage profile is mostly flat except at the very top, with the top 0.1 percent being underrepresented by a factor of 3 on average.

Concretely, the new weights are the solution to a constrained optimization problem: minimizing the distance between the new and original weights, subject to the top income shares being equal to those reported in the WID and to original age and gender compositions being respected. The solution to this problem can be interpreted as a nonresponse model, where nonresponse rates are a function of the relevant statistics (say, the top income shares) and the Lagrange multipliers attached to each specific constraint. This calibration approach helps in reducing both variance and bias on survey data. However, standard calibration methods do not apply in this case because inequality estimates are not linear, which is why they apply linearization methods and the inclusion of nuisance parameters (see Section I.B. in Blanchet *et al.* (2021) for additional details into this). The resulting weights address the fact that high-income earners are assumed to have higher nonresponse and higher underreporting rates than those lower down the distribution. (This assumption may be violated in cases where nonresponse or mis-reporting is also increasing at the very bottom of the distribution. However, without access to administrative data covering bottom incomes to corroborate this, little else can be done.) With this method the authors are able to preserve all survey covariates under the assumption of no re-ranking of survey observations. This is a necessary assumption given that they cannot assess income underreporting with the available anonymous tax data.

In Section 4, we use these new weights to assess the impact of the WID top income adjustment on conventional measures of inequality in gross and disposable equivalized income among persons. First, though, we explain how this differs from and relates to the series presented in Blanchet *et al.* (2021).

3.2. *Diverging Motives and Diverging Concepts*

Before producing and presenting these figures, we need to explain why this cannot be simply seen from Blanchet *et al.* (2021). The reason is because their paper has a different core objective, namely the construction of DINA. Thus, all the income variables produced from the EU-SILC, including the adjustments implemented, are framed to fit with that objective. The consequence is that they differ from the income measures employed in the standard inequality literature in significant ways, in terms of both the unit of analysis and treatment of the household, as well as the income components included or excluded. Indeed, DINA measures go beyond household income, adding up to national income. These choices and the rationale for them are discussed in detail in World Inequality Lab (2020) and summarized in Blanchet *et al.* (2021). Here we only provide a brief summary.

For DINA purposes the benchmark unit is "equal-split adults," whereby all the income of a household is distributed equally within couples or adults in the household. This definition is employed both for conceptual and data availability reasons, namely to align definitions with the way tax data—a key ingredient in the process—are originally structured. Individual adults then constitute the unit of analysis, with children not being included in this benchmark. (The DINA Guidelines in (World Inequality Lab, 2020, p. 20) note that alongside this benchmark "it also makes sense to distribute it across the whole population (including children) to study the distribution of how much people can consume, which can be a better proxy for standards of living"; this alternative is not included in Blanchet *et al.* (2021).) Furthermore, household income is not adjusted to account for the needs of children via equivalization, nor is equivalization employed to reflect economies of scale in consumption among adults. The main reason for not using equivalization is that the sum of total income no longer matches national income (World Inequality Lab, 2020). Having inequality measures consistent with macroeconomic aggregates is the core goal of the DINA project.

With regard to the income measures employed, the DINA framework assigns a central role to the following income variables:

1. **Pre-tax national income:** the sum of all factor income flows, before considering the operation of the tax and transfer system, but after considering the operation of the pension and unemployment insurance systems;
2. **Post-tax national income:** pre-tax income after subtracting all taxes and adding all forms of government spending.

The pre-tax national income variable is distinctive in deducting all social contributions and adding all social insurance benefits. It also adds the undistributed profits of corporations to household incomes. In standard survey-based measures of pre-tax income that would not be the case, though the treatment of both employer and social insurance contributions and benefits with respect to pensions in particular is debated and varies across studies. Instead, this concept is closer to what is usually labeled "gross" household income in survey-based inequality studies, which includes all cash social transfers whether social insurance-based or social assistance, but excludes undistributed corporate income.

The post-tax national income measure subtracts not only the direct taxes that would be deducted in arriving at the standard disposable income measures from surveys but also indirect and other taxes. Furthermore, it adds to household income the (assumed) benefits to households from all other elements of government spending like in-kind transfers related to health, education, and public infrastructure. The DINA guidelines also describe an intermediate "post-tax disposable income" measure, in which corporate-retained earnings are still distributed to individuals but government spending other than cash transfers are not included. Unlike with pre-tax income, the WID definition of post-tax income (before including government spending and other sources of National income) is similar to that in EU-SILC.

These income variables, like the unit of analysis and non-equivalization of income, are framed in light of the core objective of the DINA exercise to allocate all of national income to individuals. This means, however, that results published using the DINA framework cannot be taken to apply to the conventional equivalized household income measures, nor can they be directly compared with previous estimates of inequality. Our aim is to fill this gap. From Blanchet *et al.* (2021) we take the reweighting of the survey, while everything else is consistent with "traditional" measures of economic inequality. Compared with the DINA measures, our measures look at household income as reported in the survey, define pre-tax and post-tax income differently, equivalize income, attribute this equivalized income to all household members, and count children as well as adults in the analysis (by weighting each household by the number of persons in it). These differences allow us to provide top-income adjusted measures of economic inequality—namely the Gini index and the top 1 percent share—that can be directly compared with previous estimates, e.g., to those reported by Eurostat.

Blanchet *et al.* (2021) did not have full top income shares series right up to 2017 from fiscal data for a significant number of countries and had to base their adjustment of EU-SILC survey data on extrapolated estimates. In what follows, we confine our analysis to those country-years for which both EU-SILC and top income share estimates based on actual tax data are available, to avoid the additional complications introduced by extrapolated top fiscal income shares. This means that for most countries our analysis does not go up as far as 2017, and for many it covers only up to EU-SILC 2012 or 2013.

4. Inequality Estimates after Reweighing

We now present our results for the impact of the WID-adjustment via reweighting on conventional measures of inequality in gross and disposable equivalized income among persons. We first present an overview of two summary inequality measures in EU-SILC without and with that adjustment, and then assess the impact of the adjustment on inequality levels and trends in more depth.

4.1. *Inequality Using EU-SILC and WID-Adjusted Weights*

We compare the level of inequality as measured by the Gini index and the top 1 percent income share under two income measures that we produce from the EU-SILC, equivalizing and counting all individuals throughout. These are the widely used EU-SILC income measures for gross and disposable income (variables hy010 and hy020, respectively). We present inequality estimates using both the standard EU-SILC weights and the calibrated WID weights, which Blanchet *et al.* (2021) employ to adjust the EU-SILC data. We report these series for the Gini index in Figure 1 and for the top 1 percent income share in Figure 2.

Figure 1 shows the Gini index for equivalized gross and disposable income with the standard EU-SILC weights as continuous lines and the corresponding series with the WID-adjusted weights as dashed lines, with dark lines for gross income and gray lines for disposable income. We see that the impact of the WID-adjustment is heterogeneous, with some countries experiencing almost no change in their inequality levels, while others see substantial increases. Germany and Poland are the two countries where the WID-adjustment has most impact, while Belgium, Luxembourg, Switzerland, and the UK also see relatively large increases in inequality.[6] With the exception of Switzerland, these are all "survey" countries or countries that do not rely on register data in the sample period covered. Countries where the Gini hardly changes as a result of the WID-adjustment, on the contrary, are Denmark, Greece, Hungary, Ireland, Italy, and Sweden; of these only Greece and Hungary do not draw substantially on register data. These findings clearly suggest that the extent of use of register data is highly relevant to the capture of top incomes in the surveys.

Figure 2 reports the original (continuous lines) and adjusted (dashed lines) top 1 percent income shares in equivalized incomes among persons, with the dark lines once again showing the figures for gross income and the gray lines those for disposable income. This also shows a great deal of variation in the impact of the adjustment. Germany and Poland are again the countries with the highest increase in the top 1 percent share as a result of the adjustment, Switzerland and the UK also have relatively high increases and so too now do Romania and Serbia. Once again, countries that rely heavily or significantly on register data such as Italy, Denmark, Ireland, and the Netherlands have relatively small differences in income shares, though this is also the case once again for Greece and Hungary which do not.

The large spikes in inequality for Norway in 2005 and Iceland in 2007 when using the WID-adjusted weights merit discussion. These appear with both income concepts for both the Gini index and the top 1 percent income share, and can also be seen in the original top income shares reported on https://wid.worldwid.world). For the Norwegian case, Aaberge and Atkinson (2010) trace this to a tax reform that began to tax dividends from 2006 onwards, giving strong incentives for higher-than-normal dividend payouts in 2005. Ólafsson and Kristjánsson (2013) attribute the spike in Iceland to the speculative bubble before the Great Recession, which reached its peak in 2007. These spikes reveal how using taxable income as

[6]We exclude Iceland from this group, as its high average is driven by exceptional increase in financial earnings, as explained later in this section.

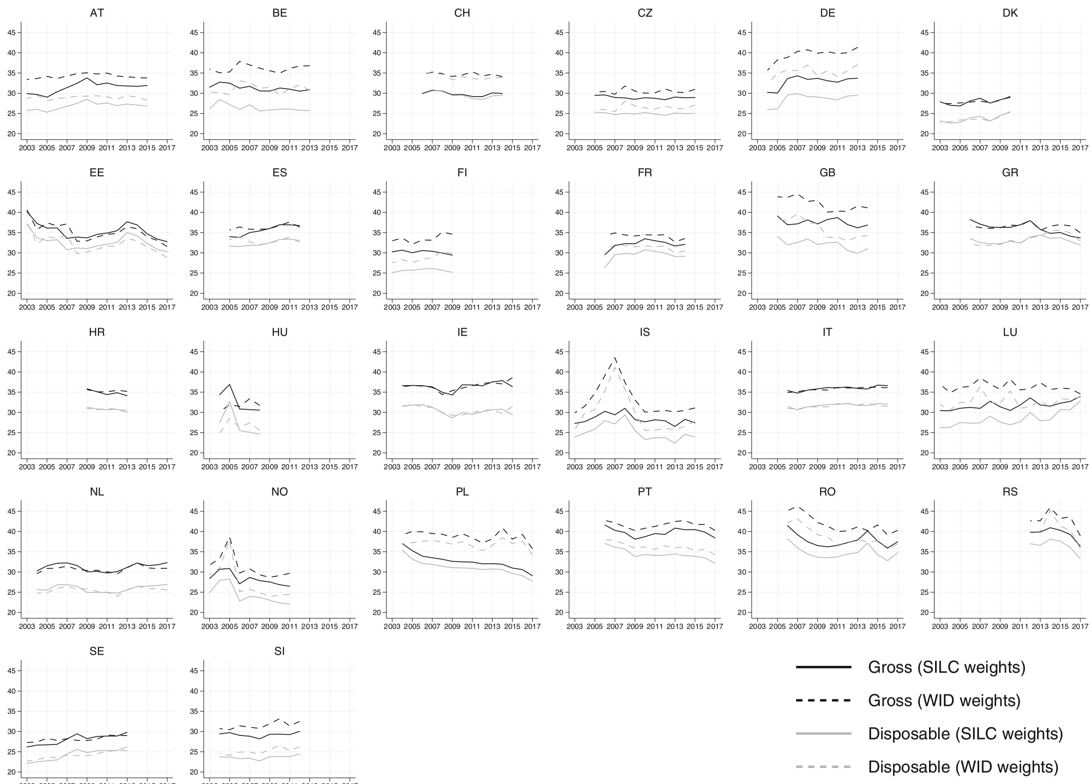

FIGURE 1. Gini Index Before and After the Adjustment.

Notes: Gini estimates for the SILC income concepts household gross and household disposable income (hy010 and hy020). Household totals divided by the OECD equivalence scale. SILC and WID-adjusted weights. Old register countries: DK, FI, IS, NL, NO, SE, SI. New register countries: AT, BE, CH, EE, ES, FR, IE, IT, LU. Survey countries: CZ, DE, GB, GR, HR, HU, PL, PT, RO.

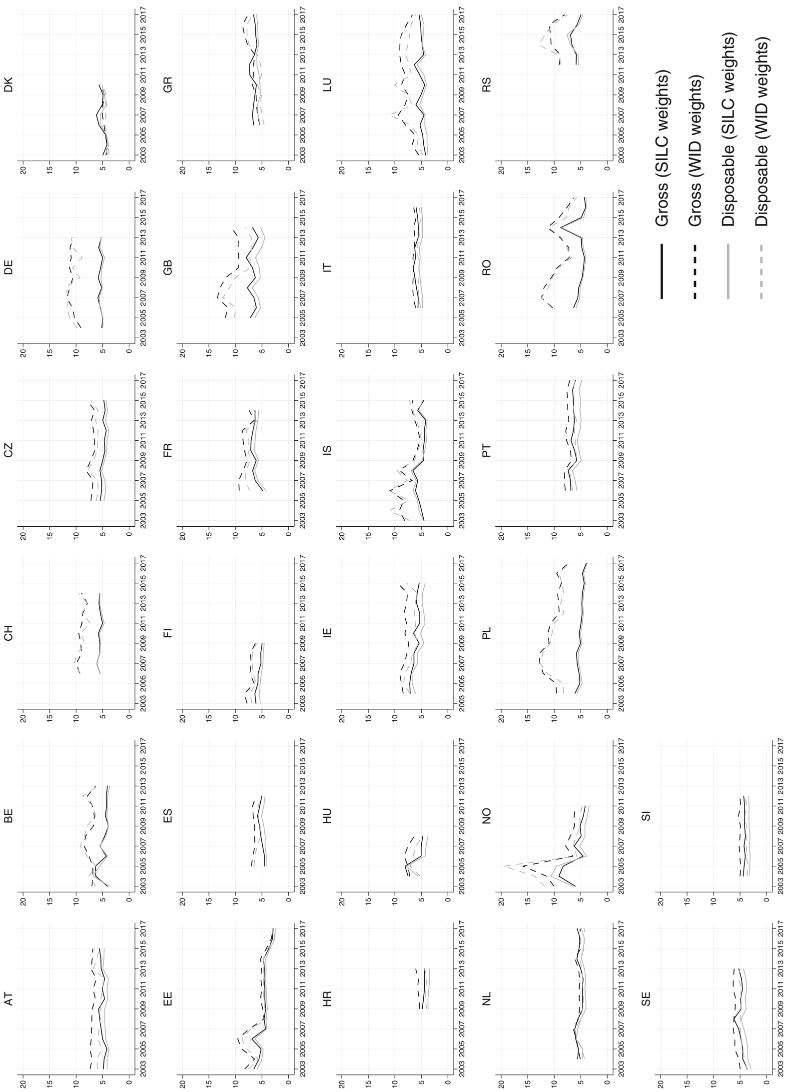

FIGURE 2. Top 1 Percent Income Share Before and After the Adjustment.

Notes: Top 1 percent income share for the SILC income concepts household gross and household disposable income (hy010 and hy020). Household totals divided by the OECD equivalence scale. SILC and WID-adjusted weights. Old register countries: DK, FI, IS, NL, NO, SE, SI. New register countries: AT, BE, CH, EE, ES, FR, IE, IT, LU. Survey countries: CZ, DE, GB, GR, HR, HU, PL, PT, RO.

the income concept—particularly in the context of tax policy changes—can impact the extent of these adjustments, and the importance of considering these issues when using income tax data in isolation.

4.2. *Changes in Inequality and Concentration Estimates*

In this section, we explore the impact of the WID-adjustment on inequality in gross and disposable equivalized income in more depth. We look at differences in levels, by quantifying the change in the Gini index after the WID weights are applied, and then look at the differences in trends, comparing the evolution of inequality over time when using the original EU-SILC versus the WID weights.

Details on the absolute change in the Gini index expressed in "Gini points" (e.g., where the Gini is shown as 35.1 rather than 0.351) as a result of the WID-adjustment are presented in Appendix A1 (Tables A3 and A4 and Figure A1). In summary, these impacts are quite similar for gross and disposable income, except for a few exceptions such as Austria, the UK, and Slovenia, where the change for gross income is slightly higher than disposable income. The countries with the largest impacts are Germany and Poland, which see an increase of about 6 Gini points, while Italy, Hungary, and Denmark have the smallest changes, of less than one-third of a point on the Gini index. On average, and for our available sample, the Gini index increases by some 2.4 points for gross income and 2.3 points for disposable income (with median change of about 2). After 2013, where we have a clearer picture of what countries rely on register data, non-register countries see higher adjustments to the Gini on average, at 2.8 points, compared to 0.9 points in register countries. (If we look at the entire sample—from 2003 to 2017—we see that register countries experience an increase of 1.7 points of the Gini for gross and 1.5 for disposable income; survey countries, on average, see an increase of 2.8 and 2.7 points for gross and disposable income, respectively.)

The WID-adjustment calibrates the survey using external income shares for the top 1, 5, and 10 percent based on tax data. We can therefore expect that the extent of the adjustment will be larger among countries with high top income shares. Such a positive correlation is seen between the percentage point change in the Gini (after reweighting the survey) and the top 1 percent income share based on fiscal data as reported by WID in Figure 3. An increase in the top 1 percent share of 1 percentage point is associated with an increase in 0.7 points of the Gini index.

Most countries see an increase in average income after the adjustment (see Figure A2 in the Appendix). However, inequality does not respond in the same way among these countries. Czechia and Germany see a large increase in both average income and in the Gini as a result of the WID-adjustment. On the contrary, Italy and the Netherlands see a large increase in average income but no change in the Gini. Moreover, for the few countries where the WID-adjustment decreases the average income, like Portugal and Serbia, we still observe an increase in inequality. Ultimately, whether accounting for the "missing rich" increases inequality or not depends on the interaction between the change in average income and the change in the variance of income after the adjustment.

The interaction between changes in average income and its variance—and the overall change in inequality—appears to be mediated using register data. We

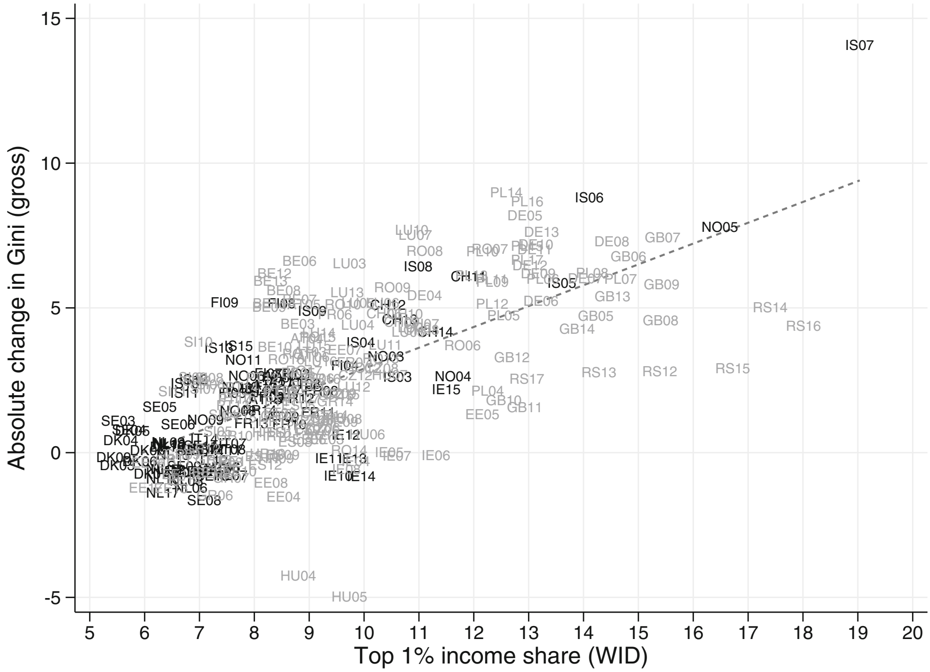

FIGURE 3. Correlation Between Size of Adjustment and the Top 1 Percent Income Share.
Notes: Gini estimates for equivalized household gross income (hy010) using the OECD scale. Countries in gray rely solely on survey data, while countries in black use register data. Top income shares from the World Inequality Database plus the updated series in Blanchet *et al.* (2021). Dashed line shows the linear fit.

have noted the larger increases in inequality among non-register countries (such as Czechia, Germany, Portugal, and Serbia), as opposed to countries that draw on register data (such as Italy and the Netherlands). This is to be expected as the estimates of top income shares in the WIL database which are central to the WID-adjustment are based in part or in full on the tax/register data that are drawn on by surveys in register countries. However, the way those top income share estimates are produced and the information employed in doing so (generally a combination of tax data and national accounts aggregates) varies from one country to the next, in ways that are not always entirely transparent. The use of matched register data at the individual level more accurately captures the variance of income within the survey support to sufficiently mitigate underestimation of average incomes (we return to this in the following section). This is consistent with the findings in Blanchet *et al.* (2022), which show that the most important part of a tax-based adjustment to surveys comes from reweighting inside the survey support rather than the addition of incomes beyond the support.

Trends in the Gini for both equivalized gross income and equivalized disposable income using the EU-SILC weights and the WID-adjusted weights are shown in Appendix Figure A3, revealing some interesting features. One example is France, where the impact of the WID-adjustment almost disappears in 2008, the year it started using register data. We also see a similar decrease in Spain around 2009, before it started using register data. Finland and Poland see an increase in the gap between both measures over time, while Romania and Luxembourg (albeit with a noisier trend) experience decreases over time. The trend is mixed in the UK, where the impact of the calibration decreased until 2011, followed by an increase for the following two years. The series also show a few 1-year spikes, such as Iceland in 2007 and Norway in 2005, as previously noted. These series suggest that the impact of the calibration is not fixed over time, and changes in the distribution of taxable income (and therefore on the top income shares) play an important role in determining its extent.

Overall, trends appear to be quite similar across all four inequality estimates. The most salient exception is Poland, where inequality measured using EU-SILC weights showed a constant decrease, being around 20 percent lower in 2017 compared to 2004. However, when looking at the WID-adjusted inequality series, we see that inequality remained constant across the period. We see the opposite in France and Luxembourg, where WID-adjusted inequality remained constant, while the EU-SILC series shows an increase over time. Germany, the other country besides Poland with the largest impact of the WID-adjustment, shows very similar trends across all inequality estimates. On the contrary, countries with small changes following the WID-adjustment can still see differences in trends, e.g., in Sweden or Greece. The conclusion we can draw from Appendix Figure A3 is that large adjustments do not necessarily translate into different trends over time.

4.3. *The Distribution of the Reweighting and Its Impact on Average Incomes*

In this subsection, we discuss the reweighting approach in more detail. We report how survey weights change across the income distribution following the WID-adjustment, and how those changes affect average incomes. We then discuss the relationship between these changes and changes in a summary measure of inequality such as the Gini index.

To capture the reweighting structure, we focus on the ratio between the new weight, modified using the WID-adjustment, and the original weight provided by EU-SILC. A ratio above 1 means that an individual or household sees their weight increase as a result of the adjustment (the converse is true for a ratio below 1), while a ratio equal to 1 means that their weight has not changed. Note that, by construction, the overall magnitude of the increase in weights among the top of the distribution must be matched by the decrease in weights among the rest of the distribution for the total population to remain unchanged (which is what is desired in these methods). We can interpret this ratio in terms of a nonresponse model, where income percentiles with a ratio above 1 were underrepresented in the original survey and receive a higher weight to correct for this.

Figure 4 shows the ratio between the new and old weights across the income distribution for the last available year in each country, shown in percentiles. It is in the top 10 percent, and particularly in the top 1 percent, where survey weights are adjusted upwards, anywhere from 3 to 35 times, with the exception of the Netherlands, Denmark, and Estonia, where the top 1 percent is adjusted downwards and where the Gini index barely changes. The rest of the distribution sees a relatively uniform change in the survey weights to maintain the population totals. (Note that

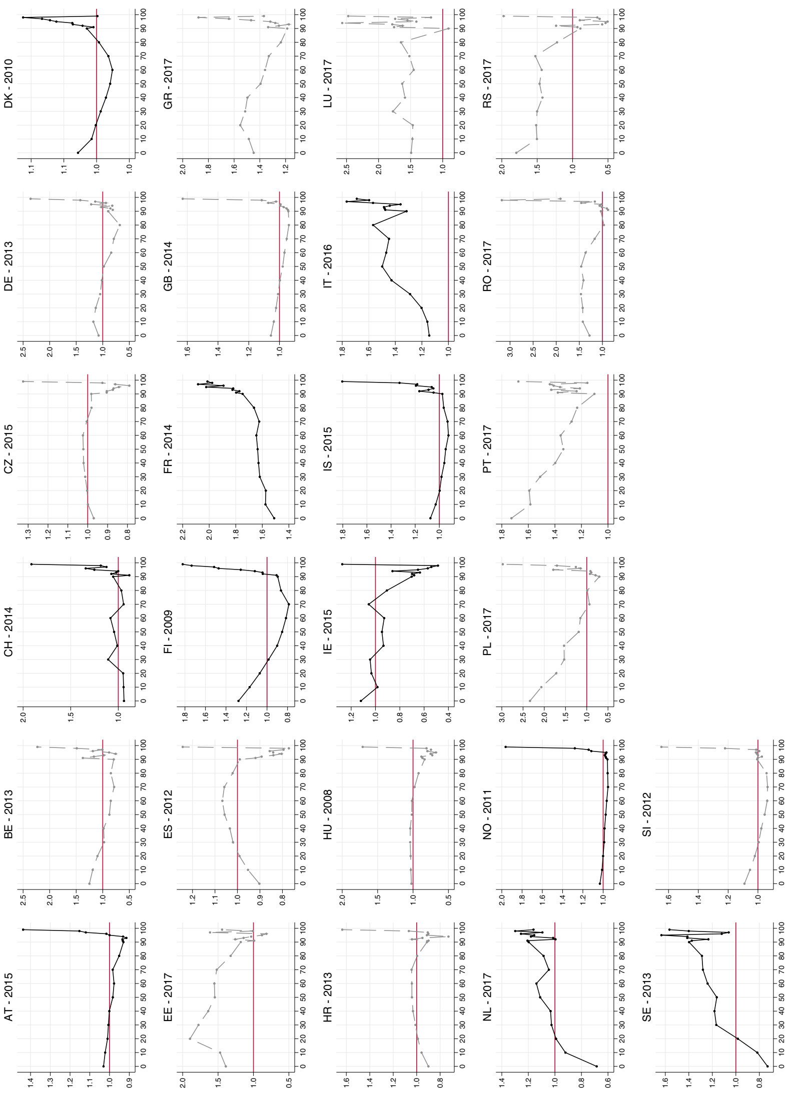

FIGURE 4. Reweighting Across the Income Distribution.

Notes: Ratio between the calibrated WID weight and the original SILC weight across years. Last available year. Black continuous lines for register countries, gray dashed lines for survey countries. The horizontal line is set at one, where the weights do not change. The *Y*-axis differs across countries to show the extent of the reweighting. Household disposable income, equivalized using the OECD scale. *X*-axis reports income deciles, from 1 to 9, and income percentiles from 91 to 100 [Colour figure can be viewed at wileyonlinelibrary.com].

a few countries, such as France and Italy, show ratios above 1 across the entire distribution; we are computing the average ratio for each percentile, where the ratio is always above zero but has no upper bound.)

Figure 5 shows how average incomes change across the distribution. In all cases, most of the distribution (i.e. percentiles 1–90) remain relatively unchanged. As expected, the largest changes happen at the top of the distribution for all countries, where average incomes increase for most countries following increases in the population weights. As the WID-adjustment changes the denominator (number of observations) as opposed to the numerator (income levels by observation), changes in Figure 4 outside the top should be mirrored in the opposite direction in Figure 5. For example, the weight of the bottom 20 percent is increased in Finland, reflected in a fall of their average income. In the top of the distribution, the relationship is mainly positive, indicating that changes in total incomes coming from the increase in population weights are proportionally greater than the increase in weights.

Whether these changes translate into changes in the overall inequality level is not straightforward, as Figure 6 shows. Average changes in income appear to be weakly correlated with changes in the Gini, while the correlation with changes in average incomes of the top 10 percent of the distribution is much stronger. In other words, the Gini tends to increase more when income at the top of the distribution goes up. However, there are clear outliers. Among countries where the Gini barely changes, we see that Italy and Sweden see relatively large increases in average income. Judging from Figure 4, it appears these countries follow a similar distribution of reweighting, with a clear upward trend in the degree to which weights are positively increased, at least until percentiles 80 or 90. This positive sloping line explains the relatively large increases in average income, as more weight is given to observations around the median, rather than the bottom, while the change in weights at the very top is not as steep as in other countries. The Netherlands also follows this pattern, like Italy and Sweden, a register country with minimal changes in inequality but notable increases in its average income; France, which is also a register country, has a similar pattern, but unlike those three has a higher increase in inequality due to a greater change in the average income of the top decile (Figure 6), driven by higher increases in the weight of its very top groups (Figure 4). Overall, the correlation between income changes and changes in the Gini depends on what part of the distribution we are looking at.

5. DOES THE TOP INCOME ADJUSTMENT METHOD MATTER?

Having seen the results of applying the WID reweighting adjustment for the Gini and top income shares, an obvious question is whether other top income adjustment methods give similar results, does the method employed matter for measured inequality levels and trends? We can assess this to an extent by comparing our results from the WID-adjustment with two other top income adjustments made to inequality estimates from EU-SILC, reported in Hlasny and Verme (2018) and Bartels and Metzing (2019). Both of these studies use EU-SILC income concepts and the Gini index to measure inequality, but differ in their approaches. To briefly summarize, Hlasny and Verme (2018) assess the relative impacts of reweighting and

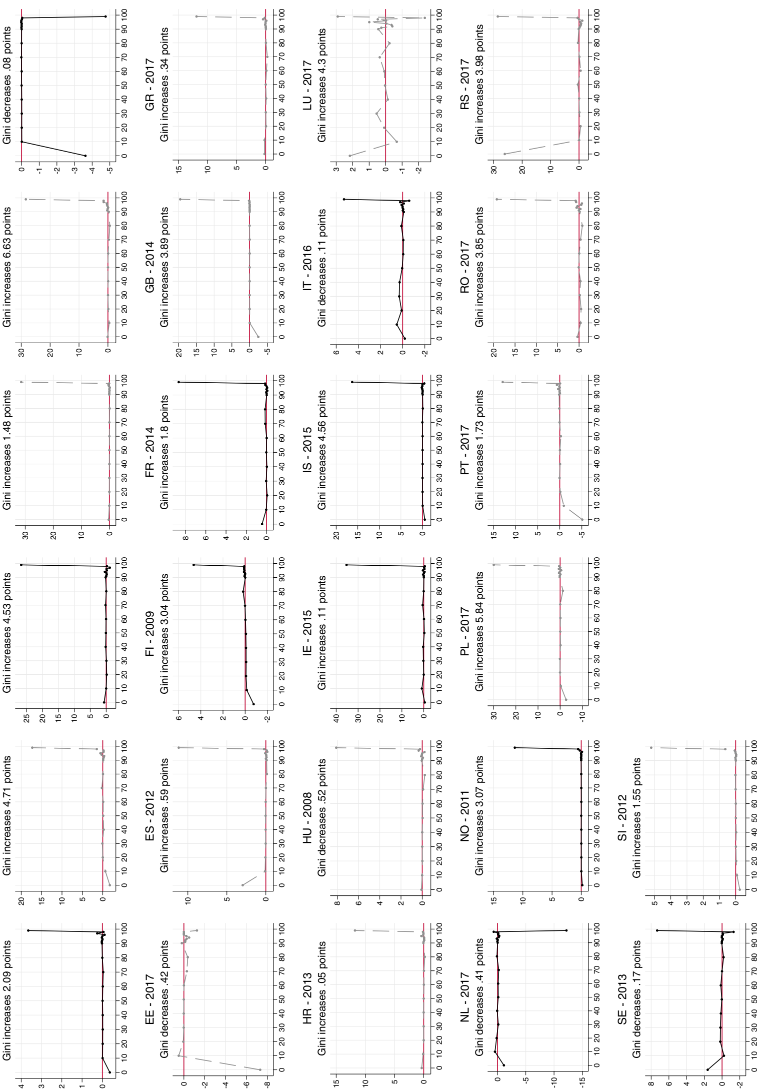

FIGURE 5. Change in Mean Income Across the Income Distribution.

Notes: Ratio between the calibrated WID weight and the original SILC weight across years. Black continuous lines for register countries, gray dashed lines for survey countries. The horizontal line is set at zero, where the average income does not change. The *Y*-axis differs across countries to show the extent of the change in average incomes. Household disposable income, equivalized using the OECD scale. *X*-axis reports income deciles, from 1 to 9, and income percentiles from 91 to 100 [Colour figure can be viewed at wileyonlinelibrary.com].

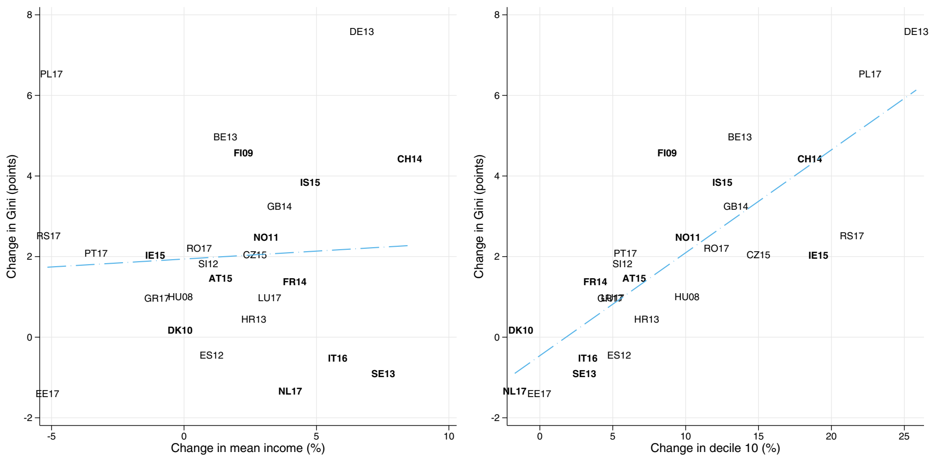

FIGURE 6. Change in Gini vs. Change in Average Incomes.
Notes: Last available year. Register countries in bold. *Y*-axis: Differences in the Gini index after the WID-adjustment. *X*-axis: Change in incomes on average (left) and at the top 10 percent (right). Household disposable income, equivalized using the OECD scale [Colour figure can be viewed at wileyonlinelibrary.com].

replacement only using survey information. Bartels and Metzing (2019) exclusively use the replacement method so as to match the top 1 percent income share from WID. Here, we focus on the reweighting adjustment of Hlasny and Verme (2018) and on the replacement adjustment of Bartels and Metzing (2019), so as to gauge the importance of (1) the use of external sources of data and (2) the adjustment approach.

Hlasny and Verme (2018) use reweighting and replacing methods without recourse to external administrative data on incomes to adjust inequality measures for top income biases. They use the EU-SILC for 2011 and the Gini index to measure inequality across 31 countries. Their outcome of interest is equivalized disposable income (hy020 divided by hx050). Following Mistiaen and Ravallion (2003) and (Korinek *et al.*, 2005; Korinek *et al.*, 2007), their reweighting approach uses information on regional nonresponse rates as well as information about the within-region distribution. Adjusting via reweighting, the Gini index increases by 3.2 points on average (with a median increase of 2.1 points).

The "integrated approach" of Bartels and Metzing (2019) replaces the top 1 percent of the survey income distribution with Pareto-imputed incomes, estimated using top income shares from WID. They opt for the 1 percent as the cutoff as they find this is the point at which survey and tax data report a significant difference in Germany (using the German SOEP). They get the complete distribution of gross income and predicted equivalized net income (obtained through a tax-benefit system transformation) for an unbalanced panel of 11 European countries between 2003 and 2013. Inequality increases among countries that exclusively rely on EU-SILC survey data (1.7 points on average for Great Britain and Germany), as opposed to countries that link register data to their surveys, where the adjustment is negligible.

We first compare the results we have derived from the WID-adjustment described in Section 3.1 with the Hlasny and Verme (2018) adjustment, both of which use reweighting, thus allowing for an evaluation of the importance of using external data. This can only be done for the single year which the latter employed, namely 2011. We then compare the WID-adjustment with the adjustment in Bartels and Metzing (2019), both of which rely on the same external source of data for their computations over numerous years, thus allowing the impact of the adjustments on trends as well as levels to be compared.[7]

5.1. *The Importance of Using External Data Sources*

The first column of Figure 7 (upper panel) shows that the Hlasny and Verme (2018) adjustment results in much higher estimates for the Gini than the WID-adjustment for Great Britain and France, and higher levels for Ireland, Italy, Greece, Norway, and Sweden. The WID-adjustment produces markedly higher estimates than those produced by Hlasny and Verme (2018) only for Romania, while also being higher for Germany and Poland. For the common sample of 22 countries, the Hlasny and Verme (2018) adjustment increases the Gini index by 3.3 points on average (median of 2.2), while the WID-adjustment increases the Gini by only 1.8 points (median of 1.4). The first column of Figure 7 (lower panel) shows that there is almost no correlation between the size of the two adjustments. This strongly suggests that the use of external data for the calibration in the latter is playing an important role. The magnitude of the adjustment in Hlasny and Verme (2018) is greatest for the UK, which is among the countries relying entirely on survey data, but it is also very large for France, which incorporates register data (in the year they examine). The WID-adjustment is much lower for these countries, and generally quite limited for “old register” countries. The Hlasny and Verme (2018) adjustment is also much larger for two “old register countries,” Norway and Sweden, than the WID-adjustment, again questioning the reliability of methods that do not use external administrative information.

5.2. *The Importance of the Method*

Turning to the comparison between the WID-adjustment and that of Bartels and Metzing (2019) in the second column of Figure 7, the upper panel compares the levels of the adjusted Gini estimates for each of the individual years for which both are available and the lower panel compares the size of these adjustments in each country-year. We see that the WID-adjustment produces estimates that are generally higher, especially for Great Britain and Germany, as well as for Norway in 2005 (which was anomalous in that dividend payments were exceptionally large for tax reasons), though lower in countries like Spain. However, unlike with Hlasny and Verme (2018), the size of the adjustment in Bartels and Metzing (2019) is positively correlated with the WID-adjustment. Bartels and Metzing (2019) highlight that their adjustments are notably higher in the two countries covered that rely

[7]There are slight differences in the unadjusted Gini estimates across the three approaches, reflecting differences in methodological choices affecting the sample employed.

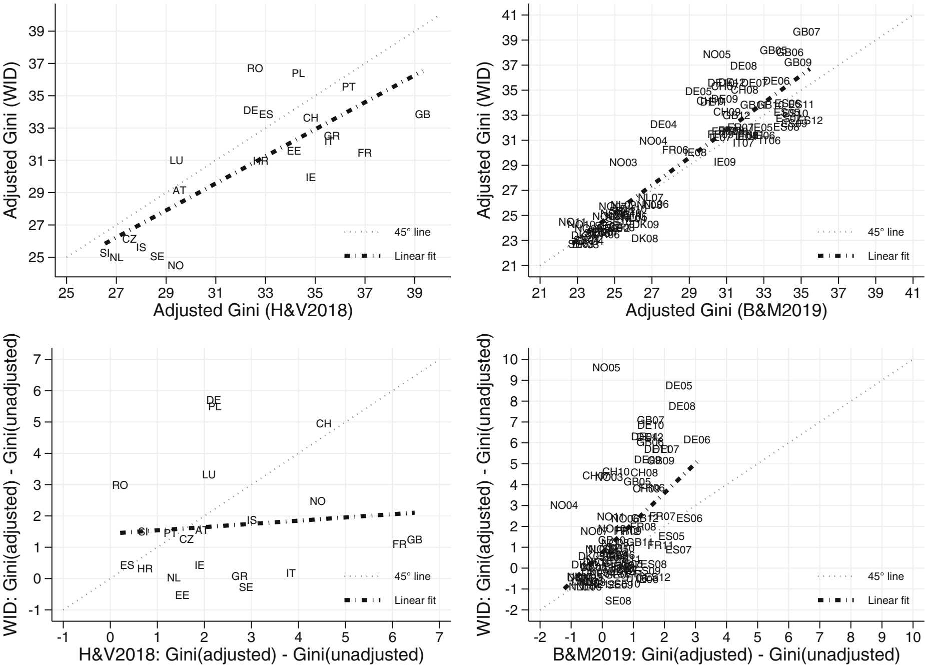

FIGURE 7. Gini Estimates—WID Data, Hlasny and Verme (2018) and Bartels and Metzing (2019).

Notes: Estimates using equivalized household disposable income using the OECD scale. The first row compares the adjusted Gini estimates using adjusted weights. The second row reports the difference between the adjusted Gini and the Unadjusted Gini. WID estimates for equivalized household disposable income (hy020) using the OECD scale. Figures for Hlasny and Verme (2018) exclude Belgium, as their adjustment increases its Gini by over 20 points. Bartels and Metzing (2019) estimates use predicted equivalent net household income based on the imputed (i.e., adjusted) gross household income, using an approximation of the tax-benefit system introduced by Feldstein (1969). Bartels and Metzing (2019) replace the top 1 percent of observations and impute synthetic values from a Pareto distribution estimated using WID top 1 and 0.1 percent shares. Register countries: DK, NO, SE, NL, IE, ES (since 2008), FR (since 2008), IT (since 2011), and CH (since 2007). Survey countries: DE and UK. For Ireland and the Netherlands, the Pareto α is calculated with the income share ratios of top 1 percent and top 0.5 percent, as the income share of the top 0.1 percent is currently not available in WID. The dotted line is the 45° line and the thick dashed line is the linear fit.

exclusively on survey data, namely Germany and the UK, which is also the case for the WID-adjustment.

Figure 8 compares the inequality trends in the adjusted Gini series between the WID-adjustment and Bartels and Metzing (2019) method. We see that the gaps between the two adjusted series for Germany and the UK are wider in some years than in others, with the UK series being much closer together in later than earlier years. For Switzerland both levels and trends diverge. For Norway in 2005, the tax-related spike is reflected in quite different adjustments as already noted, but the series are also much closer to each other in later than earlier years. This suggests that dividend payouts were also important outside of the top 1 percent in Norway in 2005, and more broadly it reveals the limits to taking a limited cutoff point (i.e., the top 1 percent as opposed to a larger group) when adjusting the top of the distribution.

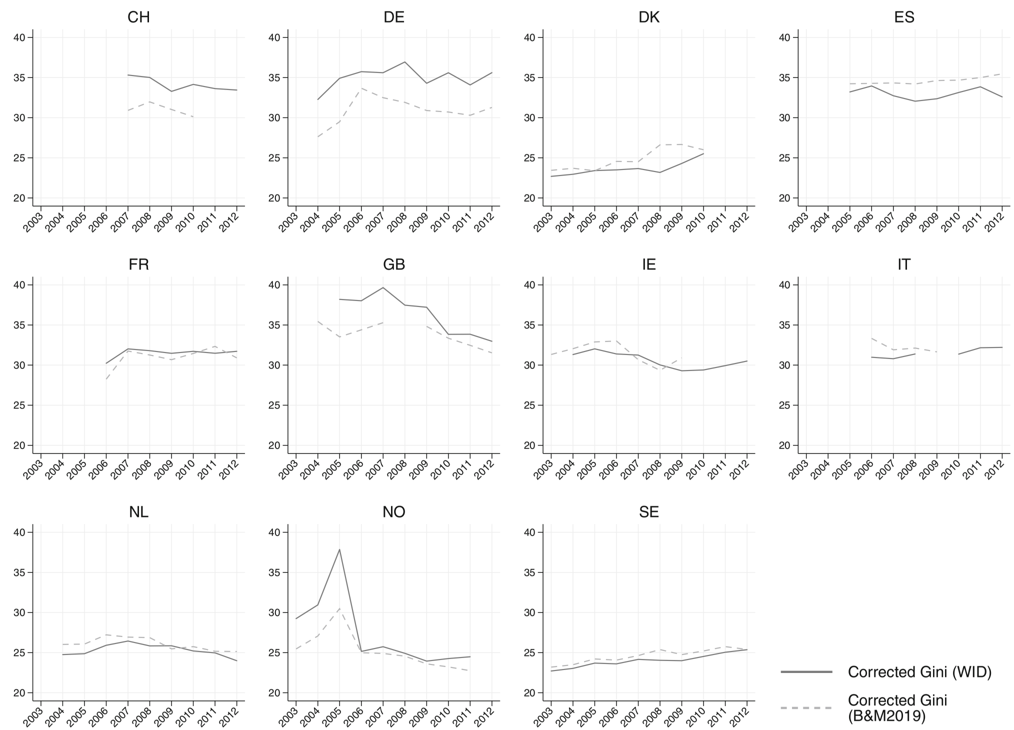

FIGURE 8. Gini Trends—WID vs. Bartels and Metzing (2019).

Notes: Adjusted Gini estimates. WID estimates for equivalized household disposable income (hy020) using the OECD scale. Bartels and Metzing (2019) estimates use predicted equivalent net household income based on the imputed (i.e., adjusted) gross household income, using an approximation of the tax-benefit system introduced by Feldstein (1969). Bartels and Metzing (2019) replace the top 1 percent of observations and impute synthetic values from a Pareto distribution estimated using WID top 1 percent and 0.1 percent shares. Register countries: DK, NO, SE, NL, IE, ES (since 2008), FR (since 2008), IT (since 2011), and CH (since 2007). Survey countries: DE and UK. For Ireland and the Netherlands, the Pareto α is calculated with the income share ratios of top 1 percent and top 0.5 percent, as the income share of the top 0.1 percent is currently not available in WID.

TABLE 1
COMPARISON OF ALL TOP INCOME ADJUSTMENT METHODS

	WID-adjustment			Bartels and Metzing (2019)			Hlasny and Verme (2018)		
	Gini	Increase (Level)	Increase (%)	Gini	Increase (Level)	Increase (%)	Gini	Increase (Level)	Increase (%)
DE	34.1	5.7	20.1%	30.3	1.8	6.2%	32.4	2.2	7.3%
ES	33.9	0.4	1.3%	35.0	0.8	2.4%	33.0	0.4	1.1%
FR	31.5	1.1	3.6%	32.3	1.9	6.2%	37.0	6.2	19.9%
GB	33.8	1.2	3.8%	32.5	1.2	3.9%	39.3	6.5	19.7%
NL	25.0	0.0	0.1%	25.2	−0.2	−0.8%	27.0	1.4	5.3%
NO	24.5	2.5	11.2%	22.7	0.3	1.2%	29.4	4.4	17.7%
SE	25.1	−0.3	−1.1%	25.8	0.9	3.7%	28.7	2.9	11.2%

Notes: Comparison of the adjusted Gini index from Hlasny and Verme (2018), Bartels and Metzing (2019), and the WID-adjustment based on Blanchet *et al.* (2021), all using household disposable income (hy020), equivalized using the OECD scale (Bartels and Metzing (2019) use a predicted value as their income measure, derived as a function of gross income). It includes all countries for which there is data in 2011.

Table 1 includes all countries in 2011 with estimates from all three methodologies. On average, it is Hlasny and Verme (2018) that shows the largest increments in the Gini index after adjustment, where France, Great Britain, Norway, and Sweden have increments of over 10 percent. The other two methods not only show lower increments in general, but the higher increments happen in different countries to those in Hlasny and Verme (2018). Both the Bartels and Metzing (2019) method and the WID-adjustment find that Germany has the biggest increment, of 6.2 and 20.1 percent, respectively, while Sweden has one of the smallest increments. Conversely, the WID-adjustment produces a high increment for Norway, while Bartels and Metzing (2019) does not. Overall the changes are most similar (both in level and direction) between the WID-adjustment and Bartels and Metzing (2019), suggesting that external data play a bigger role in explaining the impact of the adjustment than the method itself. Regarding the size of the changes, the WID-adjustment reports larger increases than Bartels and Metzing (2019) but somewhat below those of Hlasny and Verme (2018), suggesting that reweighting approaches have larger impact on the Gini index than replacing the top of the distribution.[8]

6. Conclusion

The increasing availability of estimates of top income shares derived from tax data has called into question the reliance of much inequality research and official

[8]Comparing the replacement approach in Hlasny and Verme (2018) bears similar results: reweighting produces larger inequality estimates, especially when the replacement happens at the very top of the distribution, say the top 1 percent, instead of the top 5 or 8 percent. This is largely to be expected, given that reweighting is a more interventionist approach that relies on changing weights along the entire distribution, thus impacting a composite index like the Gini more than solely modifying the very right tail of the distribution. The extent of the difference depends in part on the amount of mass beyond the survey's support, i.e., beyond the maximum income reported in the survey.

monitoring on income distribution data from household surveys which struggle to capture incomes at the very top. Research has investigated and employed various approaches to adjust survey data to address this problem. The estimates of top income shares included as "fiscal income" series in the WID for some time now represent an extremely valuable resource in this context. More recently, the WID has turned to the production of DINA series, combining data from tax sources, household surveys, and the national accounts to allocate all of national income to households. An initial step in that complex exercise is to combine micro-level data on incomes from surveys and administrative records to produce a "corrected" distribution of cash incomes. However, the distinctive income concepts and units employed make it difficult to assess the implications of this adjustment for standard measures of inequality usually derived from household surveys in the traditional inequality literature.

We have addressed this gap by employing a micro-level top income adjustments based on Blanchet *et al.* (2021) for 26 EU-SILC countries (what we have called the WID-adjustment) in a novel fashion, namely to produce adjusted inequality indicators for equivalized gross and disposable cash income among individuals. The WID-adjustment reweights the survey so that the top 1, 5, and 10 percent income shares as measured in the survey match those measured using tax data. The extent to which the impact of this adjustment on those indicators varied across countries and how that related to whether the EU-SILC surveys themselves draw on administrative/tax data was explored. Finally, we compared the results of applying the WID adjustment procedure with two other recent studies attempting to adjust inequality measures from EU-SILC, namely Hlasny and Verme (2018) and Bartels and Metzing (2019).

Our key findings are that the impact of the WID-adjustment on the Gini coefficient and top income shares for equivalized gross and disposable income among individuals varies widely across the countries in the EU-SILC. The Gini for disposable income increased by about 2.3 points on average, but by up to six points for some countries and only very modestly for others. The scale of this impact also varies from one year to the next for some individual countries, thus affecting comparisons of trends, though less substantially.

We also explored how the WID-adjustment modifies the sample weights and average incomes across the distribution. Most of the action occurs at the top of the distribution, where individual weights are increased, generally in a smaller proportion than the increase in total resulting incomes, implying an increase in the mean income for this group in the population. The larger the increase in the average income of top groups, the larger the increase in the measured Gini index.

Finally, we found some notable differences between the impacts of the WID-adjustment on the Gini coefficient and those of Hlasny and Verme (2018) and Bartels and Metzing (2019), demonstrating that the adjustment method employed does indeed matter, but only to the extent that it relies on external information from tax data as opposed to within-sample projection.

These findings underline the value of combining data from surveys with information from administrative sources to better capture top incomes. How much a post-survey adjustment will affect inequality measures depends significantly on the extent to which the survey itself draws on administrative data. In general, we find that adjustment with external data in EU-SILC has less impact on inequality measures for "register countries," where the Gini of disposable equivalized income only increases by 0.9 points in recent years. However, the use of administrative data matched to survey respondents in EU-SILC varies considerably across countries and over time, and while detailed information is available for some individual countries, it would be very helpful to have a clearer picture overall.

The findings presented here suggest that more comprehensive inclusion of administrative data on incomes should be a priority for European statistics on income and living conditions. This has the potential to substantially improve the assessment of distributional changes within and between countries, as seen from the existing biases we have shown in monitoring and ranking country inequality levels and trends, particularly among countries that make little or no use of income data from administrative sources in their surveys. An option worth serious consideration is for Eurostat to provide statistics incorporating ex-post adjustments to the surveys' income distribution following a standard methodology. The ONS in the UK has recently embarked on a procedure to provide top-income adjusted series, following the footsteps of the UK's Department of Work and Pensions (see ONS (2020)). At the same time, countries could be encouraged and supported to systematically link their surveys to register data at the point of data collection. Insofar as possible this would cover all income sources, matching survey respondents to data from income tax, social security, and other administrative sources. For countries whose register data do not cover the full adult population or are not rich in other sociodemographic information, the matching of available register data into the SILC survey could go alongside efforts to expand both the coverage and content of register data. It must be recognized that such linking with administrative registers is neither straightforward nor without error, even in the relatively straightforward case of labor earnings, as research for Denmark (Bingley and Martinello, 2017), Sweden (Kapteyn and Ypma, 2007), and the UK (Jenkins and Rios-Avila, 2021) has shown.

Further investigation will be needed to determine the most satisfactory method for Eurostat to carry out ex-post top income adjustments along the lines we suggest. In doing so, the reweighting methodology used in this paper from Blanchet *et al.* (2021), influenced by Blanchet *et al.* (2022), as well as the ones already employed by statistical offices such as the ONS in the UK (ONS, 2020) will be helpful examples.

References

Aaberge, R. and A. B. Atkinson, "Top Incomes in Norway, Chapter 9," in A. B. Atkinson, and T. Piketty (eds), *Top Incomes: A Global Perspective*, 1st ed., Oxford University Press, Oxford, 2010.

Alvaredo, F., "A Note on the Relationship between Top Income Shares and the Gini Coefficient," *Economics Letters*, 110(3), 274–7, 2011.

Alvaredo, F., L. Chancel, T. Piketty, E. Saez and G. Zucman, "Global Inequality Dynamics: New Findings from WID.world," *American Economic Review*, 107(5), 404–9, 2017.

Atkinson, A. B., "Measuring Top Incomes: Methodological Issues, Chapter 2," in A. B. Atkinson, and T. Piketty (eds), *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, 1st ed., Oxford University Press, Oxford, 2007.

Atkinson, A. B. and T. Piketty, *Top Incomes Over the Twentieth Century: A Contrast Between Continental European and English-Speaking Countries*, First ed., Oxford University Press, Oxford, 2007.

Atkinson, A. B. and T. Piketty, *Top Incomes: A Global Perspective*, 1st ed., Oxford University Press, Oxford, 2010.

Atkinson, A. B., T. Piketty and E. Saez, "Top Incomes in the Long Run of History," *Journal of Economic Literature*, 49(1), 3–71, 2011.

Bartels, C. and M. Metzing, "An Integrated Approach for a Top-Corrected Income Distribution," *Journal of Economic Inequality*, 17(2), 125–143, 2019.

Bingley, P. and A. Martinello, "Measurement Error in Income and Schooling and the Bias of Linear Estimators," *Journal of Labor Economics*, 35(4), 1117–1148, 2017.

Blanchet, T., L. Chancel and A. Gethin, "Why Is Europe More Equal Than the United States?" *American Economic Journal: Applied Economics*, 2021 (Forthcoming). https://www.aeaweb.org/articles?id=10.1257/app.20200703

Blanchet, T., I. Flores and M. Morgan, "The Weight of the Rich: Improving Surveys Using Tax Data," *Journal of Economic Inequality*, 2022 (Forthcoming). https://link.springer.com/article/10.1007/s10888-021-09509-3

Burkhauser, R. V., S. Feng, S. P. Jenkins and J. Larrimore, "Recent Trends in Top Income Shares in the United States: Reconciling Estimates from March CPS and IRS Tax Return Data," *Review of Economics and Statistics*, 94(2), 371–388, 2012.

Burkhauser, R. V., N. Hérault, S. P. Jenkins and R. Wilkins, "Survey Under-Coverage of Top Incomes and Estimation of Inequality: What is the Role of the UK'S Spi Adjustment?" *Fiscal Studies*, 39(2), 213–240, 2018b.

Burkhauser, R. V., N. Hérault, S. P. Jenkins and R. Wilkins, "Top Incomes and Inequality in the UK: Reconciling Estimates from Household Survey and Tax Return Data," *Oxford Economic Papers*, 70(2), 301–326, 2018a.

Burricand, C., "Transition from Survey Data to Registers in the French SILC Survey, Chapter 8," in V.-M. Törmälehto, M. Jäntti, and E. Marlier (eds), *The Use of Registers in the Context of EU-SILC: Challenges and Opportunities*, Eurostat, Luxembourg, 2013.

Callan, T., K. Doorley, and A. McTague, "Top Incomes in Ireland: Reconciling Evidence from Tax Records and Household Survey Data," IZA Discussion Paper Series 13585, 2020, 1–17.

Cowell, F. A. and M. Mercader-Prats, "Equivalence Scales and Inequality, Chapter 15," in J. Silber (ed), *Handbook of Income Inequality Measurement, Volume 71 of Recent Economic Thought Series*, Springer Netherlands, Dordrecht, 405–435, 1999.

Feldstein, M. S., "The Effects of Taxation on Risk Taking," *Journal of Political Economy*, 77(5), 755–764, 1969.

Garbinti, B., J. Goupille-Lebret and T. Piketty, "Income inequality in France, 1900–2014: Evidence from Distributional National Accounts (DINA)," *Journal of Public Economics*, 162, 63–77, 2018.

Goedemé, T. and L. Z. Trindade, "MetaSILC 2015: A Database on the Contents and Comparability of the EU-SILC Income Variables," INET Working Paper 2020-1, 2020, 1–285.

Hlasny, V., "Nonresponse Bias in Inequality Measurement: Cross-Country Analysis Using Luxembourg Income Study Surveys," *Social Science Quarterly*, 101(2), 712–731, 2020.

Hlasny, V., "Parametric Representation of the Top of Income Distributions: Options, Historical Evidence and Model Selection," *Journal of Economic Surveys*, 35, 1217–1256, 2021.

Hlasny, V. and P. Verme, "Top Incomes and Inequality Measurement: A Comparative Analysis of Correction Methods Using the EU SILC Data," *Econometrics*, 6(2), 1–21, 2018.

Hlasny, V. and P. Verme, "The Impact of Top Incomes Biases on the Measurement of Inequality in the United States," *Oxford Bulletin of Economics and Statistics*, 2021 Forthcoming. https://onlinelibrary.wiley.com/doi/10.1111/obes.12472

Jenkins, S. P., "Pareto Models, Top Incomes and Recent Trends in UK Income Inequality," *Economica*, 84(334), 261–289, 2017.

Jenkins, S. P. and F. Rios-Avila, "Reconciling Reports: Modelling Employment Earnings and Measurement Errors using Linked Survey and Administrative Data," IZA Discussion Papers 14405, 2021, 1–41.

Kapteyn, A. and J. Y. Ypma, "Measurement Error and Misclassification: A Comparison of Survey and Administrative Data," *Journal of Labor Economics*, 25(3), 513–551, 2007.

Korinek, A., J. A. Mistiaen and M. Ravallion, "Survey Nonresponse and the Distribution of Income," *Journal of Economic Inequality*, 4(1), 33–55, 2005.

Korinek, A., J. A. Mistiaen and M. Ravallion, "An Econometric Method of Correcting for Unit Nonresponse Bias in Surveys," *Journal of Econometrics*, 136(1), 213–235, 2007.

Lustig, N., "The "Missing Rich" in Household Surveys: Causes and Correction Approaches," Commitment to Equity (CEQ) Working Paper Series 75, 2019, 1–40.

Mistiaen, J. A. and M. Ravallion, "Survey Compliance and the Distribution of Income," World Bank Policy Research Working Paper 2956, 2003.

Morelli, S., T. Smeeding and J. Thompson, "Post-1970 Trends in Within-Country Inequality and Poverty: Rich and Middle-Income Countries, Chapter 8," in A. B. Atkinson, and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol 2, Elsevier, Amsterdam, 593–696, 2015.
Muñoz, E. and S. Morelli, "kmr: A Command to Correct Survey Weights for Unit Nonresponse using Groups' Response Rates," *The Stata Journal*, 21(1), 206–219, 2021.
Ólafsson, S. and A. S. Kristjánsson, "Income Inequality in Boom and Bust: A Tale from Iceland's Bubble Economy, Chapter 15," in J. C. Gornick, and M. Jäntti (eds), *Income Inequality: Economic Disparities and the Middle Class in Affluent Countries*, 1st ed., Stanford University Press, Palo Alto, CA, 416–438, 2013.
ONS, "Using Tax Data to Better Capture Top Earners in Household Income Inequality Statistics," *Office for National Statistics*, 1–15, 2019.
ONS, "Top Income Adjustment in Effects of Taxes and Benefits Data: Methodology," *Office for National Statistics*, 1–17, 2020.
Piketty, T., "Income Inequality in France, 1901–1998," *Journal of Political Economy*, 111(5), 1004–1042, 2003.
Piketty, T. and E. Saez, "Income Inequality in the United States, 1913-1998," *Quarterly Journal of Economics*, 118(1), 1–41, 2003.
Piketty, T., E. Saez and G. Zucman, "Distributional National Accounts: Methods and Estimates for the United States," *Quarterly Journal of Economics*, 133(2), 553–609, 2017.
Piketty, T., L. Yang and G. Zucman, "Capital Accumulation, Private Property, and Rising Inequality in China, 1978–2015," *American Economic Review*, 109(7), 2469–2496, 2019.
Törmälehto, V.-M., *High Income and Affluence: Evidence from the European Union Statistics on Income and Living Conditions (EU-SILC)*, Publications Office of the European Union, Luxembourg, 2017.
Törmälehto, V.-M., M. Jäntti and E. Marlier, "The Use of Registers in the Context of EU-SILC: Challenges and Opportunities, Chapter 28," in A. B. Atkinson, A.-C. Guio, and E. Marlier (eds), *Monitoring Social Inclusion in Europe*, 2017th ed., Eurostat, Luxembourg, 2017.
Törmälehto, V.-M., O. Kannas and M. Säylä, "Integrated Measurement of Household-Level Income, Wealth and Non-monetary Well-being in Finland," Statistics Finland Working Papers 1/2013, 2013, 1–36.
World Inequality Lab, Distributional National Accounts Guidelines: Methods and Concepts Used in the World Inequality Database, 2020.
Yonzan, N., B. Milanovic, S. Morelli, and J. Gornick, "Drawing a Line: Comparing the Estimation of Top Incomes Between Tax Data and Household Survey Data," Stone Center Working Paper Series 27, 2020, 1–50.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Appendix S1 Supplementary Material