This PDF is a selection from a published volume from the National Bureau of Economic Research

Volume Title: Measuring Distribution and Mobility of Income and Wealth

Volume Authors/Editors: Raj Chetty, John N. Friedman, Janet C. Gornick, Barry Johnson, and Arthur Kennickell, editors

Volume Publisher: University of Chicago Press

Volume ISBNs: 978-0-226-81603-6 (cloth), 978-0-226-81604-3 (electronic)

Volume URL:
https://www.nber.org/books-and-chapters/measuring-distribution-and-mobili ty-income-and-wealth

Conference Date: March 5-6, 2020

Publication Date: November 2022

Chapter Title: Using Tax Data to Better Capture Top Incomes in Official UK Income Inequality Statistics

Chapter Author(s): Dominic Webber, Richard Tonkin, Martin Shine

Chapter URL:
https://www.nber.org/books-and-chapters/measuring-distribution-and-mobility-income-and-wealth/using-tax-data-better-capture-top-incomes-official-uk-income-inequality-statistics

Chapter pages in book: p. 679 – 700

23

Using Tax Data to Better Capture Top Incomes in Official UK Income Inequality Statistics

Dominic Webber, Richard Tonkin, and Martin Shine

23.1 Introduction

The Office for National Statistics (ONS) and its predecessors have published statistics on the distribution and redistribution of household income since 1961, beginning with "The Incidence of Taxes and Social Service Benefits," which was one of the first publications in the world to give such a complete examination of these issues.

Throughout this time, ONS's statistics on income inequality have been based primarily on household surveys, in common with the majority of official statistics on the distribution of household finances globally. Data are currently derived from the Living Costs and Food Survey (LCFS), a voluntary sample survey of private households in the UK. While household surveys have several important benefits over relying solely on administrative records, there is a well-recognized challenge: they do not fully capture the

Dominic Webber is an economist at the UK Office for National Statistics.

Richard Tonkin is a statistician with the United Nations Statistics Division. He was Assistant Director of the Crime, Income and Wealth Division at the UK Office for National Statistics when this chapter was written.

Martin Shine is currently a Data Analyst at Butterfly Data, and was an Assistant Statistician in Household Income and Expenditure Analysis for the UK Office for National Statistics when this chapter was written.

The authors would like to thank Peter Matejic (formerly Department for Work and Pensions), Stephen Jenkins (London School of Economics), Steve Martin-Drury, Callum Clark, and Ozer Beha (all Office for National Statistics) for their incredibly helpful input throughout the development of this work. The UK Office for National Statistics bears no responsibility for the analyses and conclusions in this chapter, which are solely those of the authors. For acknowledgments, sources of research support, and disclosure of the authors' material financial relationships, if any, please see https://www.nber.org/books-and-chapters/measuring-distribution-and-mobility-income-and-wealth/using-tax-data-better-capture-top-incomes-official-uk-income-inequality-statistics.

incomes of the very richest individuals and households, particularly those among the so-called top 1 percent. There are several potential reasons for this (see, e.g,. Lustig 2018), the relative importance of which varies across countries and across surveys depending on the methods used. These include:

- frame or noncoverage error, where the frame used to select the sample for the survey does not fully cover the population of interest (in this case, households in the UK);
- unit nonresponse error, which may occur if individuals or households with higher incomes are less likely to participate in surveys than those in the rest of the income distribution;
- item nonresponse error, if those with higher incomes participating in surveys do not report all their sources of income;
- underreporting, where the levels of income received for some sources may be intentionally or unintentionally underreported by survey respondents;
- sparseness, where data on top incomes are limited due to the fewer number of observations within the dataset with very high incomes, making it difficult to estimate the true distribution.

Section 23.2 of this chapter looks at the nature and scale of undercoverage of top incomes in the UK, and in particular, in the LCFS data currently used to produce ONS's official statistics on the distribution of household income.

Section 23.3 considers the variety of approaches that have been used to address these issues in previous economic research, including those using survey data only (e.g., Ruiz and Woloszko 2016), those using tax data only (e.g., Alvaredo 2017; Atkinson and Ooms 2015), and those that combine survey and administrative data in some way (e.g., Burkhauser et al. 2018b; Jenkins 2017), before considering the most appropriate family of methods for potential application to ONS's statistics.

Section 23.4 introduces in more detail the methods developed first by the UK's Department of Work and Pensions (e.g., DWP 2015) and then expanded on by Burkhauser et al. (2018b), which make use of the so-called Survey of Personal Incomes (SPI), a microdata set containing taxable incomes, based on a sample of administrative records from UK taxpayers.

Section 23.5 examines two sets of methods which build upon this approach, in which survey-based mean incomes for quantile groups are replaced by equivalent figures from the SPI tax data. In the first, the mean gross income for each SPI quantile group is imputed onto individuals in the equivalent quantile groups in the survey data. In the second, means for each SPI quantile group are first imputed onto the survey data based on the monetary value of the boundary for each quantile group, before the dataset is reweighted to reflect control totals for each quantile group and the overall population. In addition, the analysis also examines the effects of using different levels of granularity in the quantile groups, along with those of different thresholds for applying the adjustments.

Finally, full SPI data for any year is not normally available at the time of producing official income distribution statistics, and to wait would introduce an unacceptable time lag for users of these estimates. This chapter therefore also assesses the impact of using projected SPI data versus waiting for final outturns on the estimates and the conclusions that might be drawn from them.

23.2 Survey under Coverage of Richest Households in the UK

The primary source of data for ONS's official statistics on household income inequality is the LCFS, a sample survey of private households in the UK, which collects detailed data on household income and expenditure, and currently covers approximately 5,000 households. The household income dataset produced from the LCFS is often known as the Effects of Taxes and Benefits (ETB) data, after one of the main publications which uses it, the *Effects of Taxes and Benefits on Household Income* (ONS 2019).

The response rate for the LCFS, following trends for social surveys internationally, has declined over recent years, falling from 62 percent in 2001–2 to 43 percent in 2017–18. Falling response rates are likely to impact the reliability of data, potentially across the whole of the distribution, and further strengthen the case for supporting survey data with other sources, as we describe in the next stages.

It is possible to make an assessment that the survey does not fully capture top incomes by making a comparison to the SPI, an individual-level dataset produced from tax records by Her Majesty's Revenue and Customs (HMRC), based a sample of individuals potentially liable for UK tax. Figure 23.1 displays the ratio of individual-level gross income in the SPI data to that in ETB, providing a clear demonstration of the issue of survey undercoverage of top incomes.

Examining the three most recent years where full SPI datasets are available, figure 23.1 highlights that at around the 97th percentile, average personal income as reported in the SPI is higher than that reported in ETB. This shows that survey undercoverage is an issue for ETB, and therefore measured estimates of income inequality are potentially lower than they should be. These findings are similar to those presented by Burkhauser et al (2018a, 2018b) where they examined the issue of survey under coverage of top incomes in the Household's below Average Income (HBAI) dataset, produced by Department for Work and Pensions (DWP). That the ratio becomes close to 1 below the top few percentiles of the distribution provides evidence to suggest that the largest challenge affecting top incomes in UK survey data is that of underreporting by survey respondents rather than

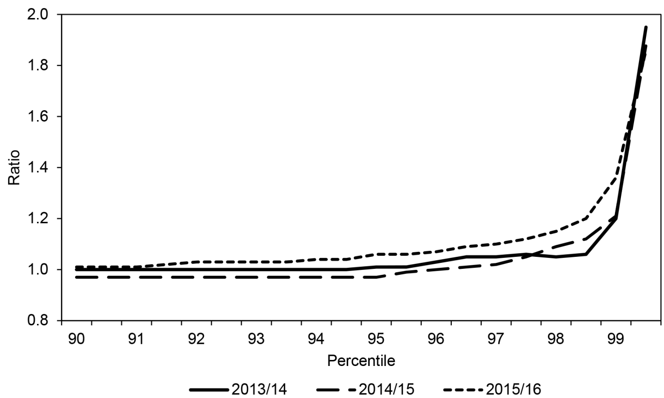

Fig. 23.1 Ratio of gross income of tax data to survey data, by quantile, UK, 2013/14 to 2015/16

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

lower survey participation. If the primary issue was that of unit nonresponse, it might be expected that the ratio would remain above 1 considerably further down due to those with the highest incomes being absent from the survey, thus misaligning the two distributions.

23.3 Approaches to Addressing Survey under Coverage of Top Incomes

Economic research has employed a variety of methods to address the issues outlined in the previous section. While a more detailed taxonomy is provided in Lustig (2018), broadly speaking these can be divided into three groups: methods that use survey data only, those that use tax data only, and those that combine survey and administrative data in some way.

In the first of these approaches, income estimates are calculated directly from the survey data, for all but the very richest. To derive an estimate of overall inequality, these are combined with estimates of inequality among the very rich calculated by approximating the tail of the distribution by a Pareto distribution (for example, see Ruis and Woloszko 2016). However, Jenkins (2017) notes that such an approach may be unreliable, due to undercoverage, resulting in downward bias, particularly where sparseness in the survey data is an issue.

Sources, such as World Inequality Database (WID.world), do not use survey data at all for their UK estimates. Instead, they use HMRC data about personal incomes subject to tax, supported by population and income control totals from the midyear population estimates and national accounts, respectively (e.g., Alvaredo 2017; Atkinson and Ooms 2015). Approaches based on tax data also have their limitations, however. For example, while such an approach can provide estimates of measures such as top-income shares, it does not provide microdata that allow analysis of the full income distribution. Also, such methods can typically provide measures of only individual income rather than household income, and exclude several important income sources such as interhousehold transfers and individual savings account (ISA) interest.

For these reasons, in trying to address the issue of measuring top earners' income, we have focused on approaches that combine both survey and administrative sources. There is a wide variety of potential methods within this category (e.g., Burkhauser et al. 2018b; Campos and Lustig 2017; Jenkins 2017; Medeiros, Galvão, and Nazareno 2018), but the common feature is that such methods allow one to draw on the relative strengths of each relative strengths of each source.

Within this broad category of methods which combine survey and administrative sources, we have considered several other criteria to select our preferred approach for adjusting ONS's ETB statistics. First, recognizing our role as producers of official statistics, our approach needs to be methodologically robust, based on academic research and existing best practice, as well as being relatively transparent and understandable by users. In addition, the value of ETB microdata to academics and researchers needs to be reflected, with any approach continuing the enable the replication of headline measures using the data. This means the method needs to be applied to the underlying microdata, rather than the headline measures themselves.

The adjusted data should also enable the reporting of income on a household rather than an individual basis, reflecting the greater insight this measure provides due to, for instance, intrahousehold sharing of resources. Finally, the selected approach necessarily needs to be feasible to achieve considering the current availability of source data.

It is for this final reason that we have not focused on the direct use of linked survey and administrative data in this chapter, though in the slightly longer term this is our ambition as more record-level administrative data become available within ONS under the Digital Economy Act. This will help to improve the quality of estimates of income at both the bottom and the top of the income distribution, while maintaining the detailed information about people and households (such as intrahousehold relationships, spending, health status, etc.) that is not so readily available in administrative data. However, there remains a clear need to develop and introduce a method of bringing together the survey and administrative data that does not rely on direct record linkage, that can be applied now, and that allows for the production of historical data.

These criteria have led us to the set of methods first implemented in the UK by the DWP for the HBAI, which is based on the Family Resources Survey (FRS), and later adapted by Stephen Jenkins and colleagues (Burkhauser et al. 2018a, 2018b). These methods replace the highest incomes in the survey with cell-mean imputations based on corresponding observations in tax return data. In the UK context, this adjustment is often referred to as the "SPI adjustment," because it uses HMRC's SPI data (with the Burkhauser et al. 2018b modification referred to as the "SPI2").

23.4 The SPI and SPI2 Adjustments

The DWP introduced the pioneering SPI adjustment to their HBAI statistics, based on the FRS, during the early 1990s. The adjustment itself was developed to correct for both levels of and volatility in the highest incomes captured in the survey.

While the approach has been modified over recent years, the current SPI methodology is as follows:

1. Estimate personal total taxable income for individuals on survey data.
2. Use ONS population data to estimate the number of people equivalent to the top 0.32 percent of nonpensioner adults, separately for Great Britain and Northern Ireland.
3. Estimate the income threshold on the SPI dataset above which the number of nonpensioners is equal to that estimated in stage 2.
4. Flag all nonpensioners on HBAI data with personal total taxable income above this threshold.
5. All flagged individuals have their income replaced with the mean average of SPI income above the threshold.
6. All flagged individuals are reweighted so that the population total estimated in step 2 is achieved.
7. All nonflagged people are reweighted so that population totals are maintained.

The same methodology is applied for pensioners, but based on the richest 1.16 percent.

As the SPI data are not usually available until a number of years after the end of the reference period, the adjustments that DWP applies to its statistics need to rely on projected information supplied by HMRC. The impact of adjustments that are derived using projected rather than final data will be explored below.

Figure 23.2 shows the overall impact of the SPI adjustment, increasing reported levels of income inequality while dampening the volatility of the overall series. Between 1994/95 and 2017/18, the SPI adjustment increases the Gini coefficient by an average of 1 percentage point per year, while the average absolute annual change decreased from 1.0 to 0.5 percentage points.

Looking at this in more detail, figure 23.2 highlights the respective role of each stage of the adjustment. It shows that the impact of adjusting the income of only the richest individuals acts primarily to smooth the series, without affecting the levels substantially compared with the unadjusted

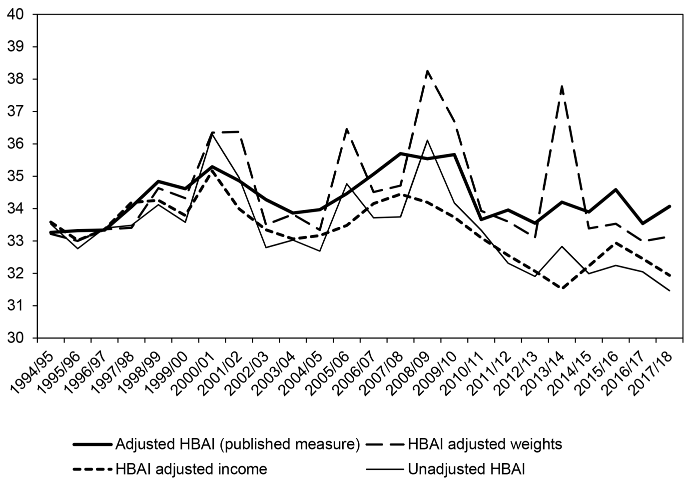

Fig. 23.2 Relative impacts of the different phases of the SPI adjustment on Households below Average Income statistics, 1994/95 to 2017/18

Source: Department for Work and Pensions, Family Resources Survey; HM Revenue and Customs, Survey of Personal Incomes.

series. This demonstrates the amount of volatility there is in the HBAI data, with the number of very rich individuals surveyed throughout fluctuating over the period. The impact of adjusting only the weights, so that the number of very rich individuals on HBAI is consistent with the SPI data, acts to increase measured inequality with little effect on survey volatility.

The groundbreaking work of the original SPI adjustment is recognized by Burkhauser et al. (2018a), but they go on to argue that with the increasing focus on income inequality and the income shares of specific groups, such as the so-called top 1 percent, the SPI approach requires "new scrutiny." They outline several recommendations for optimizing the SPI adjustment, setting out a so-called SPI2 methodology.

First, they demonstrate that survey undercoverage of top incomes in HBAI data tends to become more of an issue from around the 95th percentile upward, becoming particularly acute from the top 2 percent. As highlighted in figure 23.1, similar results are demonstrated in ETB data, making a case for adjusting incomes at a lower threshold than that currently set by the current SPI1 adjustment.

Burkhauser et al. (2018a) also compare the ratios of adjusted HBAI data with SPI data at different quantile groups toward the top of the distribution. They highlight that the gap between the mean incomes of HBAI and SPI quantile groups is reduced for the top 2 percent to 1 percent group, as well as for the top 1 percent to 0.5 percent group. However, they further highlight that the correspondence between adjusted HBAI data and SPI remains low toward the very top of the distribution (in the top 0.5 percent to 0.1 percent group, and the top 0.1 percent). They argue that this stems from the fact that cell means from the original SPI adjustment are calculated from a wide range of incomes. Therefore, the adjustment tends to impute incomes that are too low for the 0.5–0.1 percentile group, and too small for the top 0.1 per cent. They conclude that more granular adjustments could lead to improved measures of income inequality for the very top incomes.

Aside from applying a lower threshold and increased granularity, the SPI2 adjustment differs in two important ways from the original SPI methodology. First, the SPI2 methodology contains no stratification for pensioner and nonpensioner individuals, or for Great Britain and Northern Ireland. Second, the SPI2 does not involve reweighting of the data, instead simply replacing survey incomes for each quantile group with the SPI mean for the same quantile group.

This chapter therefore builds on the work of both DWP and Burkhauser et al. by exploring different methodological choices with the aim of identifying a perceived optimum variant for use with ONS's household income statistics, considering the various constraints that exist.

23.5 New Approaches: Quantile and Reweighting Methods

In determining the optimal approach for adjusting for undercoverage of top incomes in ETB statistics, two underlying methods are used and tested. The first, which we term the "quantile approach," closely resembles the so-called SPI2 approach developed by Burkhauser et al. (2018b). The second—"reweighting"—brings together elements of both the SPI2 and the original SPI adjustment adopted for HBAI statistics described earlier.

Under the quantile method, the mean gross income for each SPI quantile group is imputed onto individuals in the equivalent quantile groups in the survey data. More specifically, the process is as follows:

1. Estimate personal taxable income for individuals on ETB data.
2. Add a dummy case to the SPI data to account for individuals who do not pay tax. Their personal taxable income is set to zero and their weight reflects the difference in population totals between the ETB and SPI data sets.
3. Rank individuals in ETB and SPI data by personal taxable income.
4. Allocate individuals at the top of both the ETB and SPI distributions to quantile groups, depending on the threshold and granularity selected. For instance, at the 97th percentile and 0.5 percent levels of granularity, there will be six groups of individuals at the top, each representing 1/200th of the population.
5. Calculate the mean personal taxable income for each quantile group in the SPI data.

6. Replace the income of each case within the ETB quantile groups with the mean SPI income from the corresponding group.
7. Add back several income components to the ETB cases not represented in SPI data, such as ISAs and intrahousehold transfers.
8. Recalculate income tax and national insurance contributions for the adjusted ETB cases based on new estimates of personal pretax income.
9. Aggregate personal-level income across household members to estimate adjusted household disposable income.

By contrast, the reweighting methodology replaces steps 4–6 with the following:

4a. Allocate individuals at the top of the SPI distributions to quantile groups, depending on the threshold and granularity selected. For instance, at the 97th percentile and 0.5 percent levels of granularity, there will be six groups of individuals at the top, each representing 1/200th of the population.
4b. Calculate the lower-income boundaries for each of these quantile groups on the SPI data. Create bands in the ETB data using these boundaries.
5. Calculate the mean personal taxable income for each quantile group in the SPI data and impute this onto individuals in the equivalent survey bands.
6a. Reweight the ETB bands so that their weights are the same as the SPI quantiles.
6b. Reweight the unadjusted ETB data so that overall population totals for each weighting variable are maintained.

Where the primary challenge affecting top incomes is that of underreporting rather than lower survey participation of very rich households, the effects of the two methods should be largely equivalent in practice. However, where lower participation also has an important impact, the second "reweighting" method should prove more effective.

The combination of two different SPI adjustment methods (reweighting and quantile), many possibilities in both the threshold and granularity of adjustments, and the decision whether to adjust separately for pensioners and nonpensioner means that there are many choices to be made in selecting a preferred approach. In determining this, we have sought to address the following questions:

1. Should the richest pensioner and nonpensioners *p* be adjusted separately?
2. How low should the threshold be?
3. How granular should the adjustment be?
4. Should the quantile or reweighting method be chosen?
5. Should estimates be revised once final outturn data is available?

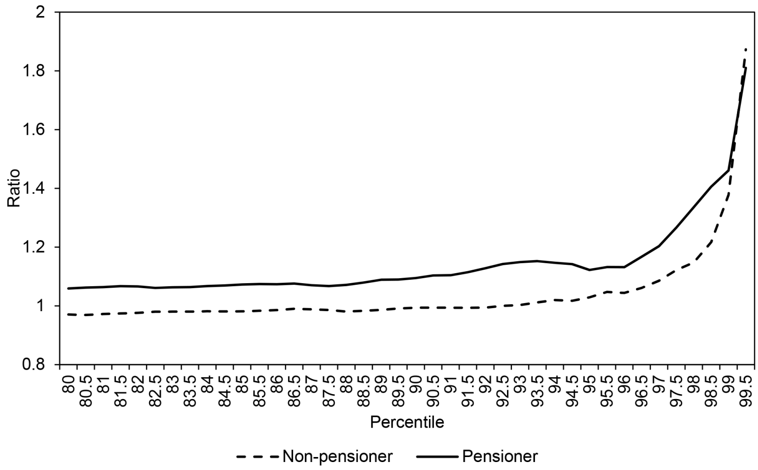

Fig. 23.3 Ratio of gross income measured using SPI and ETB data, by quantile, and pensioner status, UK, 2015/16

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

The following sections will present analysis, examining each of these questions in turn, in order to arrive at an evidence-based rationale for deciding on the method to be used in future official statistics.

23.5.1 Should the Richest Pensioner and Nonpensioners Be Adjusted Separately?

The SPI adjustment currently implemented in HBAI statistics involves separately adjusting the income of the top pensioner and nonpensioners. Considering this approach, Burkhauser et al. (2018a) questions whether there is clear rationale for doing so. Exploring these issues in more detail, figure 23.3 presents the ratio of average (mean) personal taxable income by quantile group reported on ETB and SPI, for both pensioner and nonpensioner distributions.

The ratio of taxable income measured on ETB and SPI for working age people closely resembles the whole population average shown in figure 23.1, hovering around 1.0 before sharply increasing at the 96th percentile. We see a similar observation in the distribution of pensioners, where the ratio also increases at the 96th percentile. In contrast to the non-pensioner distribution, the ratio remains above 1 during the entirety of the portion of the distribution shown in this chart, suggesting that the income distribution for pensioners is affected by both underreporting and unit nonresponse.

These findings indicate that survey undercoverage of top incomes is one which affects both the non-pensioner and pensioner distributions. Only 1.7 percent of pensioners have a personal taxable income high income high enough to feature in the top 5 percent of the overall income distribution. This means that an adjustment applied just to the overall distribution would be unlikely to fully adjust for undercoverage of the incomes of pensioners.

Providing statistics on pensioner incomes is an important breakdown for users and, given the issues with underreporting presented here, there is a clear rationale for these statistics to stratify by pensioners and nonpensioners, as is currently done by DWP. In addition, pensioners and nonpensioners typically have very different sources of income, which can mean that government policy can impact on these groups in different ways, and so it is important for the LCFS data to best reflect the differences these different groups.

23.5.2 Which Income Threshold?

In considering the threshold to use, there is a balance that needs to be struck. Too high, and there is a risk that the adjustment does not fully account for survey undercoverage. Too low, and survey data are being unnecessarily discarded in exchange for averages from the SPI.

As demonstrated in figure 23.1, undercoverage of top incomes in the LCFS begins to become an issue at around the 96th percentile. Based on this, we explored and tested thresholds ranging from the 95th to 99th percentile, using both the quantile and reweighting methods. In testing these different thresholds, the quantile group sizes were held constant at 0.5 percent.

In general, these variations in the thresholds were found to have a relatively small impact on measures of average income for the top decile and inequality. This is demonstrated in figures 23.4 and 23.5, which present the average disposable income of the top decile and Gini coefficients, respectively, comparing these measures under a range of different thresholds to the unadjusted estimate. By far the largest difference is that between having any adjustment (with a threshold between the 95th and 99th percentiles) and not having one at all.

Over the period considered, differences between the various adjustments, based on different thresholds, are relatively small under both the quantile and reweighting methods. For instance, figure 23.4 looks at the mean of the richest 10 percent, across the five adjustments, and over the period 2001/02 to 2017/18. It shows that under the quantile method, the average absolute deviation of the five adjustments from their mean is 0.4 percent, compared with the 14.2 percent average difference between each adjusted estimate and the unadjusted amount. Similarly, under the reweighting approach, these figures are 0.5 percent and 16.2 percent, respectively.

The same is also true when examining the Gini coefficient (figure 23.5). Under the quantile method, the average absolute deviation of the five adjustments from their mean is 0.1 percentage points, compared with the 1.8 percentage point average absolute difference between each adjusted estimate and the unadjusted amount. Similarly, under the reweighting approach, these figures are 0.1 and 2.0 percentage points, respectively.

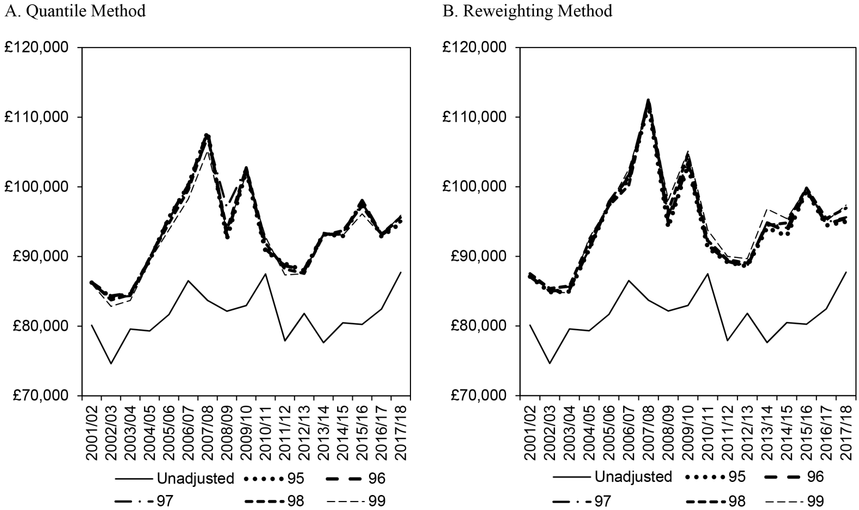

Fig. 23.4 Mean equivalized household disposable income of the richest 10 percent of people, with varying thresholds, 0.5 percent granularity, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

The gap between the adjusted and unadjusted was greatest during the period between 2005/06 and 2009/10 when, according to the reweighting method (based on 97th percentile threshold and 0.25 percent granularity), the average income of the richest 10 percent of people increased by 28.5 percent between 2001/02 and 2007/08, before falling 20.8 percent by 2012/13. This compares with the unadjusted data which was much more stable over this period.

The trends of adjusted data compared with unadjusted data over time are broadly similar. The most notable exception is during the four years from 2005/06, where income inequality increased sharply as measured using adjusted data, before falling back to similar levels observed in the unadjusted data. Also, between 2012/13 and 2015/16, there was a larger rise in inequality in the adjusted data compared with the unadjusted data. However, the gap between adjusted and unadjusted data narrowed between 2015/16 and 2017/18 due to a larger rise in the inequality levels seen in the unadjusted data.

In 2010/11, there was little difference between the Gini coefficients for the adjusted and unadjusted data. This most likely reflects the introduction of a 50 percent top tax rate in 2011/12, with evidence to suggest that this led to people forestalling their income (HMRC 2012), which has resulted in closer similarity between income for top earners as reported in the SPI data and in the LCFS.

Looking across the time series, the differences between the adjustments

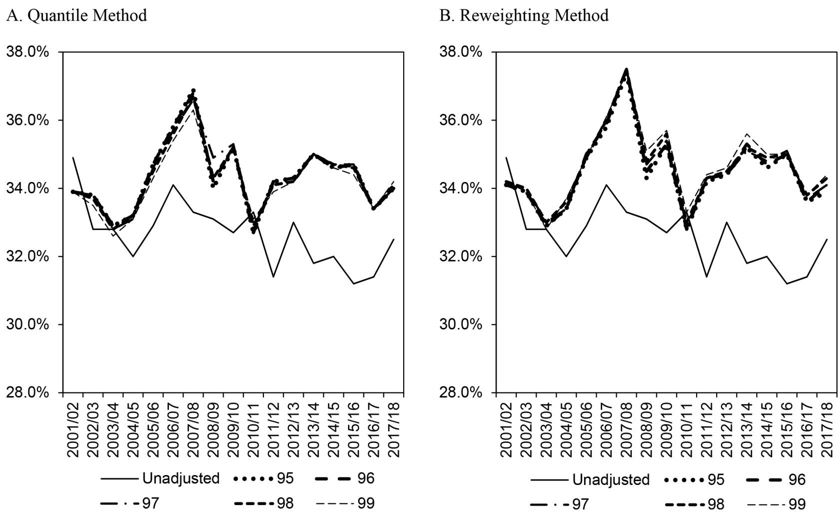

Fig. 23.5 Gini coefficient of disposable income, with varying thresholds, 0.5 percent granularity, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

based on different thresholds are largest between adjustments based thresholds at the 99th percentile and 98th percentile, suggesting that the former may not necessarily fully addressing the survey undercoverage (though these differences are very small in comparison to the difference between having an adjustment and not). Given that figure 23.1 highlights that survey undercoverage becomes more apparent at the 97th percentile, it seems that this be a sensible long-term position for the threshold to be set.

23.5.3 What Size Should the Quantile Bands Be?

Another consideration when applying a methodology for adjusting top incomes is the width of quantile bands. As explained earlier, while the current SPI adjustment in place for HBAI statistics using a single quantile band, Jenkins and colleagues (Burkhauser et al. 2018a, 2018b) demonstrate that more precise estimates of the incomes of the very richest individuals may be achieved by introducing more granular band. There is, of course, a trade-off: while smaller quantile groups may provide more granularity to the adjusted data—potentially allowing for a closer representation of the upper tail of the income distribution—we risk finding ourselves with very few, maybe zero, cases within bands in the survey data.[1]

1. This can be found only under the reweighting approach. By construction there will always be cases on the survey data in which bands are formed on the basis of quantiles, as is the case under the quantile approach, rather than the income thresholds used for the reweighting approach.

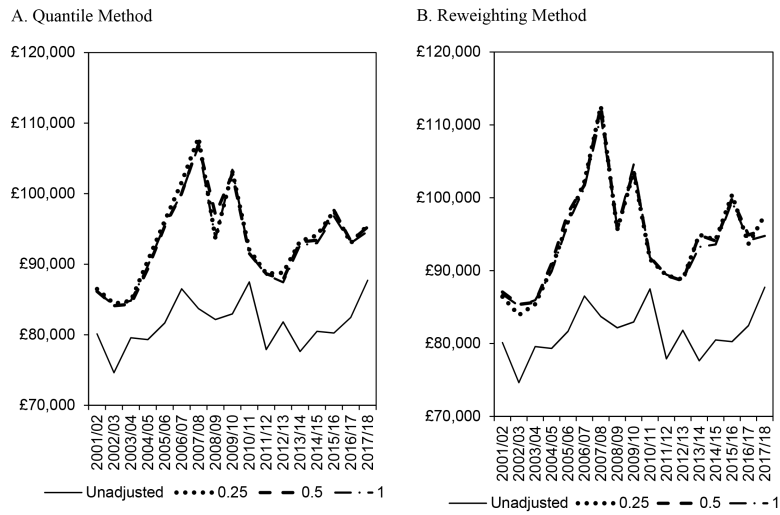

Fig. 23.6 Mean equivalized household disposable income of the richest 10 percent of people, with varying granularities, 97th threshold, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

This section explores the impact of varying quantile band sizes, looking again at the incomes of the richest 10 percent of people, and the Gini coefficient, under different adjustments. This time, it is the threshold that is kept constant (97 percent), so that the impact of changing the quantile group size between 0.25 percent, 0.5 percent, and 1 percent is clearly seen.

Whichever quantile group size is used, the top-income adjustment has a similar effect on both the average income of the richest 10 percent (figure 23.6) and the Gini coefficient (figure 23.7). In both cases, the differences between adjustments are much smaller than any differences between adjusted and unadjusted data (excluding 2010/11, as discussed in section 23.5.2).

Across all years, the average change in income of the richest 10 percent, compared with the unadjusted data, is 14.5 percent for the quantile method, and 15.9 percent for the reweighting method (figure 23.6). The different trends in the adjusted and unadjusted data, for the income of the richest 10 percent over time, are broadly similar to those discussed in section 23.5.2. While the differences between adjustments based on different quantile bands are small under the quantile method, they are slightly more pronounced under the reweighting method. For instance, while the average absolute difference between adjustments based on 0.25 percent and 1 percent quantile bands is 0.6 percent under the quantile method, it is 0.8 percent using the reweighting approach.

Looking at income inequality as measured by the Gini coefficient, over the

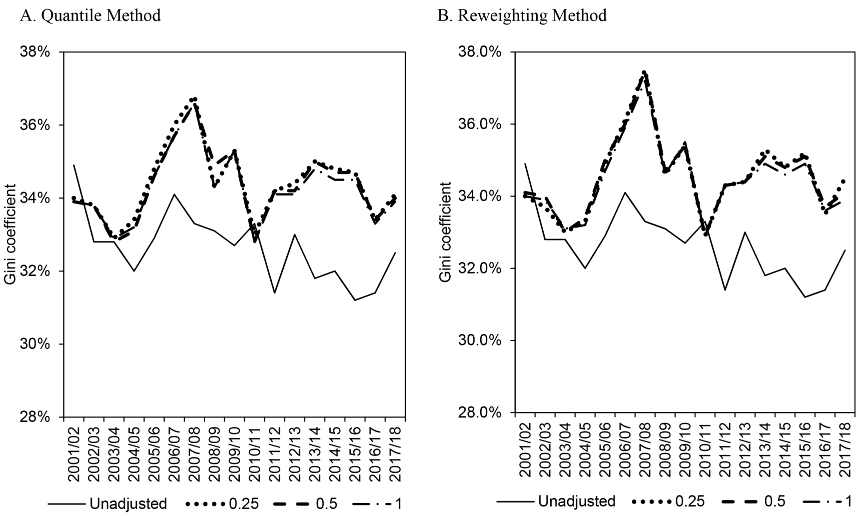

Fig. 23.7 Gini coefficient of disposable income with varying granularities, 97th percentile threshold, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

period analyzed, the average difference between the unadjusted and adjusted data is 1.7 and 1.8 percentage points for the quantile and reweighting methods, respectively, with the differences between the different quantile sizes being considerably smaller. Notably, the difference between the 1 percent and 0.25 percent quantile bands in the reweighting approach is not so pronounced when measuring the Gini coefficient, compared with measures for the average income of the top 10 percent.

23.5.3.1 Income Share of the Top 1 Percent

While the impact of changing the granularity is modest on estimates of income of the richest 10 percent, and the Gini coefficient, there are much more substantial differences between quantile band sizes when examining an alternative measure of income inequality: the household income share of the richest 1 percent of individuals. The income of the so-called top 1 percent is a topic of considerable focus that has not historically been reported in ONS's household income releases due to the reported issues with the top end of survey data. These adjustments provide the opportunity to examine these analyses of the top 1 percent share on ONS's survey data for the first time.

Figure 23.8 highlights that estimates for the income share of the top 1 percent are slightly higher with a 0.25 percent quantile band size, compared with 0.5 percent, which in turn gives considerably higher estimates in most years than 1 percent. For example, the average share of income for the top 1 percent is 6.1 percent between 2010/11 and 2017/18, based on quantile

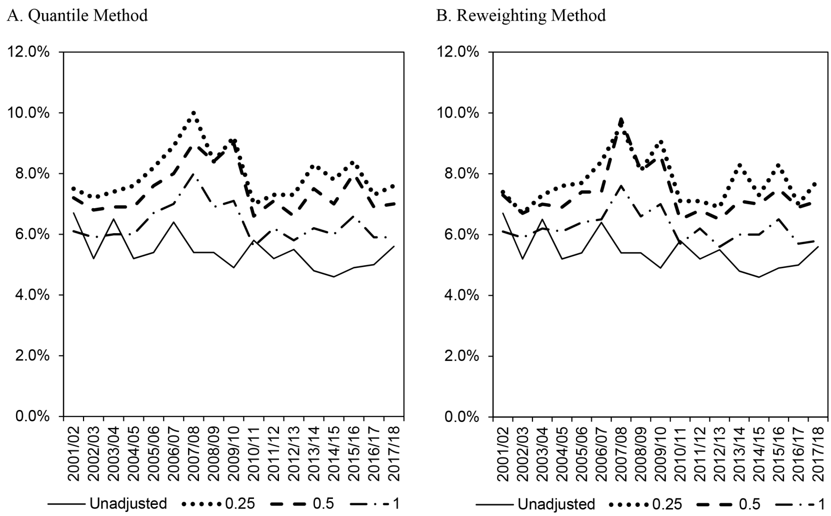

Fig. 23.8 Share of household equivalized disposable income received by the richest 1 percent of individuals, with varying granularities, 97th percentile threshold, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

band sizes of 1 percent, compared with 5.1 percent using that data unadjusted. This average increases to 7.1 percent and 7.7 percent for 0.5 percent and 0.25 percent quantile band sizes, respectively.

These differences arise as a result of the composition of households at the top of the distribution, reflecting that most households have multiple occupants, but typically only one person will have a personal income high enough to warrant being adjusted. For instance, in the 2017/18 dataset, just under half of the people in the top 1 percent based on personal income are in the top 1 percent based on household income. This means that most of those in the top 1 percent of household income, using a 0.25 percent quantile band size adjustment, will contain those whose incomes have been replaced with an average of the 0.5 percent richest people as reported in the SPI; whereas people in the top 1 percent of household income, adjusted using 1 percent quantile band sizes, will include people who have had their income replaced with the lower mean derived from the top 1 percent of the SPI data, and hence they have a lower estimated income share.

Given the increasing focus on measures such as the income share of the top 1 percent, these findings suggest that 1 percent quantile bands are too broad. However, the trade-off—when applying the reweighting method—is the increasing likelihood of empty bands on the survey data for smaller quantile band sizes, resulting in adjusted survey data that is not as representative of the tax data. For example, in the hypothetical situation where

Table 23.1 **Number of survey cases in the top 20 0.25 percent quantiles of pensioners**

	Cases in top 20 0.25% quantiles of pensioners (20 = top)																			
Year	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2001/02	3	8	3	4	4	7	3	3	4	4	2	2	6	3	7	2	7	6	2	6
2002/03	6	4	4	1	9	8	2	7	5	5	3	5	4	6	1	3	3	5	7	**0**
2003/04	4	2	6	2	1	3	4	10	4	3	3	2	5	2	1	6	1	1	4	5
2004/05	7	4	6	3	5	4	3	3	4	6	7	2	2	2	3	1	5	3	**0**	2
2005/06	2	8	2	6	5	4	2	7	3	4	4	4	8	4	3	1	4	4	1	3
2006/07	1	7	3	1	6	6	4	3	2	5	2	7	5	2	2	3	2	5	2	2
2007/08	5	2	5	6	1	3	4	2	4	3	2	2	3	8	3	4	2	**0**	1	2
2008/09	1	2	4	1	3	**0**	4	3	6	3	3	2	4	3	3	6	2	3	5	1
2009/10	6	7	2	1	1	3	3	5	2	4	3	5	2	1	4	4	2	1	3	1
2010/11	4	3	2	**0**	3	4	4	5	6	5	6	5	4	1	5	2	3	5	4	1
2011/12	3	6	7	6	3	3	5	7	2	7	4	6	1	6	8	7	3	3	3	2
2012/13	6	3	3	1	5	4	2	6	3	5	5	7	4	7	4	2	4	3	4	3
2013/14	1	3	2	6	5	8	4	7	3	1	2	3	6	9	7	5	2	7	7	1
2014/15	1	5	8	7	3	5	4	4	2	4	6	10	5	2	9	3	2	5	5	2
2015/16	6	5	7	6	7	2	6	4	3	2	7	2	2	3	2	4	**0**	6	4	2
2016/17	3	5	5	7	1	1	6	5	9	5	6	6	8	8	8	5	3	4	**0**	2
2017/18	3	7	11	4	3	7	8	7	6	3	8	3	10	6	7	6	2	2	2	3

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

the richest three 0.25 percent bands of pensioners are empty, adjusted data will not reflect the average income of the richest 0.75 percent of this group, resulting in less precise measures of income inequality, and greater volatility at the top of the distribution.

Table 23.1 explores this in more detail, counting the number of cases in the pensioner distribution within the top 20 0.25 percent quantile groups (as defined in the SPI data). It highlights that over the period 2001/02 to 2017/18 only seven bands were empty. This highlights that while the issue of empty bands can occur, it is not at a scale considered large enough to be considered a major issue.

In summary, figures 23.6, 23.7, and 23.8 highlight the impact of different granularities on measures such as the average income of the richest 10 percent of people, the Gini coefficient, and the income share of the top 1 percent. The income share of the top 1 percent and, to a lesser extent, the Gini coefficients highlight larger deviations in adjustments based on 1 percent quantile bands, compared with 0.5 percent and 0.25 percent bands. The conclusion is that 1 percent bands are potentially too broad, which inclines us toward smaller band sizes.

In deciding between 0.25 percent and 0.5 percent quantile bands, table 23.1 is instructive in that it highlights that the former is not overly affected by the issue of zero survey observations within quantile groups, even when

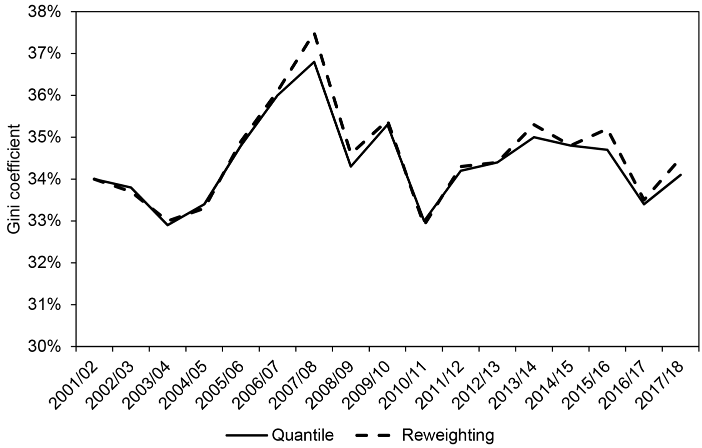

Fig. 23.9 Gini coefficient of disposable income based on quantile and reweighting adjustment, 97th percentile threshold and 0.5 percent quantile bands, UK, 2001/02 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

looking at the income distribution of pensioners. For these reasons, we conclude that adjustments based on 0.25 percent quantile bands are most likely to be optimal, with the greater granularity offered ensuring a more realistic approximation of the upper tail of the income distribution.

23.5.4 Quantile or Reweighting Method?

Figure 23.9 shows the differences in the Gini for disposable income under the quantile and reweighting approaches based on a fixed threshold (97th percentile) and granularity (0.5 percent). It shows that, while broadly similar, since 2011/12, the Gini under the reweighting approach has been marginally higher than for the quantile approach in most years.

Figure 23.1 suggested that the primary reason for the undercoverage of top incomes in the LCFS was underreporting rather than unit nonresponse. However, if that were entirely the case, it might be expected that the quantile and reweighting approaches would be essentially comparable, with the distributions under the threshold being the same for each. That the Gini is marginally higher under the reweighting approach suggests that nonresponse at the top of the distribution does play some role, indicating that, although more complex, the reweighting approach is to be preferred.

Another reason for adopting the reweighting approach comes from figure 23.3, which highlighted that although nonresponse may be a lesser concern for the overall income distribution (mirroring the findings of Burkhauser et al. 2018a), there is evidence to suggest it may be more noticeable in the distribution of pensioners' incomes.

A further important consideration is coherence. The reweighting

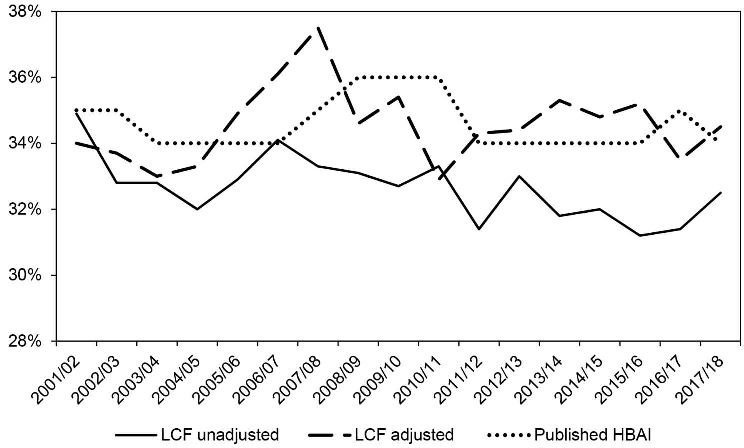

Fig. 23.10 Gini coefficients of published Households below Average Income data compared with unadjusted ONS data and adjusted ONS data (adjusted using the re-weighting method, on the 97th percentile threshold in 0.25 percent quantile groups), UK, 2001/02 to 2016/17

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes; Department for Work and Pensions, Family Resources Survey.

approach is closest in methodological terms to the original SPI adjustment currently used by DWP's HBAI statistics. Adopting this approach therefore ensures greatest coherence in terms of methods across the UK statistics (figure 23.10).

23.5.5 Should Estimates Be Revised Once Final Outturn Data Is Available?

As previously discussed, one of the challenges in implementing the approaches discussed in this chapter is their reliance on SPI data, which is not typically made available to researchers until at least two years after the end of the income reference period. The dataset, for example, covering 2016/17 was released on the UK Data Service (UKDS) in November 2019. To ensure that detailed analysis of household income is published in a timely manner, it is necessary to use estimates provided by HMRC which are based from projections from historical SPI datasets.

While Burkhauser et al. (2018a) have demonstrated that there can be notable differences between the projected data and the published outturn data, the key questions are whether these lead to significant impacts on headline measures of income inequality, and then (assuming the projected data are deemed suitable to adjusting top incomes) whether there should be a revision once final data are published.

In order to address these questions, we have compared estimates of inequality based on projected and outturn data. In this analysis, we used

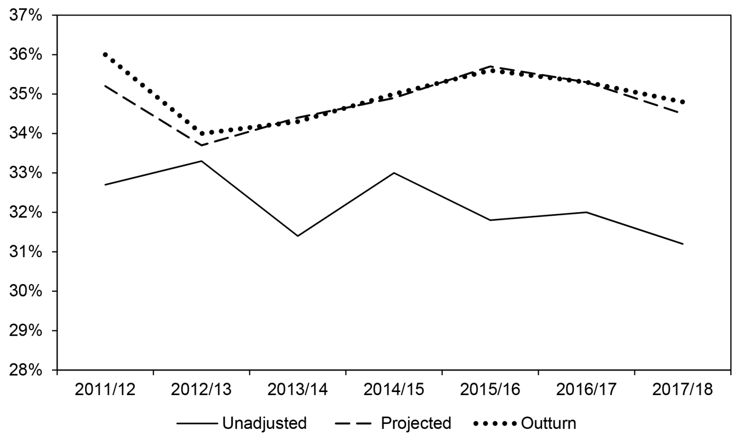

Fig. 23.11 Comparison on measures of inequality using ETB data that is unadjusted, adjusted using projected SPI data, and adjusted using final published SPI data, UK, 2011/12 to 2017/18

Source: Office for National Statistics, Living Costs and Food Survey; HM Revenue and Customs, Survey of Personal Incomes.

projected SPI estimates, which were originally supplied to DWP in the production of its HBAI statistics, to adjust ETB statistics. These results are then compared to estimates using the same adjustment, but using final outturn data which have since been published.

The analysis in figure 23.11 demonstrates that the impact of moving from projected to final data leads to, on average, a 0.2 percentage point revision of the Gini coefficient, without a systematic bias either way. Given that the revisions are small, coupled with the effort involved in revising data two years or more after their initial publication, we feel there is not a compelling case for routinely revising measures of Gini coefficients once final SPI data are made available, at this stage. Therefore, we do not propose a policy of regular revision to our household income statistics, following the adoption of the new top income adjustment, to take account of the outturn SPI data when they becomes available. However, this will need to be closely monitored, initially once the 2017/18 SPI data are released later in 2020, and for a few years thereafter, to determine whether this revision policy needs reevaluating.

23.6 Conclusion

This chapter has sought to develop, test, and decide on a methodology for addressing survey undercoverage of top incomes in ONS's official statistics based on the ETB data. Building on the SPI adjustment developed by the DWP and more recent work by Burkhauser et al. (2018a, 2018b), we have tested two separate adjustment methods—the quantile and reweighting approaches—and under each we have explored different variations, including the thresholds above which incomes are adjusted, and different size quantile bands.

The analysis has highlighted that the issue of survey undercoverage affects both the pensioner and nonpensioner populations, which, given the importance of such breakdowns within ONS's analysis, provides a clear case for adjusting separately for pensioners and nonpensioners as DWP's HBAI statistics currently do.

The findings of this analysis also provide evidence to suggest that the "reweighting" approach may be preferable to the "quantile" approach. The former helps ensure that undercoverage due to both underreporting and unit nonresponse is adequately covered, and again helps maximize coherence with DWP's HBAI statistics.

In exploring the impact of varying the threshold, this chapter finds that by far the key difference is between having an adjustment and not, and that thresholds between the 95th and 99th percentiles have relatively little impact. However, as survey coverage of top incomes starts to become most problematic above the 97th percentile, there is a rationale for making adjustments above this level.

In contrast to differences due to varying the threshold, differences between measures based on different quantile sizes are much more visible. This is particularly the case when examining the income share of the richest 1 percent of the population, with band sizes of 0.25 percent leading to the top 1 percent share being 1.5 percentage point higher on average than for estimates based on 1 percent quantile band sizes. We conclude that for these types of measures, 1 percent quantile bands are too broad, not sufficiently reflecting the steep tail at the richest end of the income distribution. We conclude that 0.25 percent quantile bands are able to benefit from increased granularity, while being sufficiently broad to ensure that having cells to impute in the survey based on zero cases does not become a problem.

Finally, we discuss the issue of revisions, reflecting that the SPI data we use as the basis for adjusting the highest earners in our surveys is not usually available until at least two years after the end of the income reference period. This means that, to ensure statistics that are timely and that also don't suffer from the survey undercoverage of top incomes, it is necessary to use projected SPI data. Our analysis determines that, in terms of headline measures such as the Gini coefficient, the impact of projected against outturn data is marginal. This leads us to conclude that, having adopted the adjustment, we will not plan to regularly revise once full SPI data become available, though this decision will continue to be reviewed.

This preferred top income adjustment (reweighting approach, stratified by pensioners and nonpensioners, 97th percentile threshold and 0.25 percent quantile groups) will be introduced from ONS's 2018/19 income distribution statistics onward, to be published in March 2020. At this stage, the adjusted

ETB time series will go back to 2001/02, reflecting the availability of SPI data from the UKDS. However, we intend to explore options for extending the time series back further, ideally to 1977, reflecting the start of the ETB data currently available. We will also make the top income adjusted ETB series itself available to researchers via the UKDS.

References

Alvaredo, Facundo. 2017. "UK Estimates of Top Income Shares 2013–2014 and 2014–15: Note on Methods." WID.World Technical Note 2017/2. Paris: World Inequality Lab.

Atkinson, Anthony B., and Tahnee Ooms. 2015. "UK Estimates of Top Income Shares 2012–2013: Note on Methods." WID.World Technical Note No. 2015/4. Paris: World Inequality Lab.

Burkhauser, Richard V., Nicolas Hérault, Stephen P. Jenkins, and Roger Wilkins. 2018a. "Survey Under-Coverage of Top Incomes and Estimation of Inequality: What Is the Role of the UK's SPI Adjustment?" *Fiscal Studies* 39 (2): 213–40.

Burkhauser, Richard V., Nicolas Hérault, Stephen P. Jenkins, and Roger Wilkins. 2018b. "Top Incomes and Inequality in the UK: Reconciling Estimates from Household Survey and Tax Return Data." *Oxford Economic Papers* 70 (2): 301–26.

Campos-Vazquez, Raymundo M., and Nora Lustig. 2017. "Labour Income Inequality in Mexico: Puzzles Solved and Unsolved." UNU-WIDER Working Paper No. 2017/186. Helsinki: United Nations University World Institute for Development Economics Research.

Department of Work and Pensions (DWP). 2015. *Households below Average Income: An Analysis of the Income Distribution 1994/95 to 2013/14.* London: Department for Work and Pensions.

Her Majesty's Revenue and Customs (HMRC). 2012. *The Exchequer Effect of the 50 per cent Additional Rate of Income Tax*. London: The National Archives.

Jenkins, Stephen P. 2017. "Pareto Models, Top Incomes and Recent Trends in UK Income Inequality." *Economica* 84 (334): 261–89.

Lustig, Nora. 2018. "Measuring the Distribution of Household Income, Consumption and Wealth." In *For Good Measure: Advancing Research on Well-Being Metrics beyond GDP*, edited by Joseph E. Stiglitz, Jean-Paul Fitoussi, and Martine Durand, 49–83. Paris: OECD.

Medeiros, Marcelo, Juliana de Castro Galvão, and Luísa de Azevedo Nazareno. 2018. "Correcting the Underestimation of Top Incomes: Combining Data from Income Tax Reports and the Brazilian 2010 Census." *Social Indicators Research* 135: 233–44.

Office for National Statistics (ONS). 2019. *Effects of Taxes and Benefits on UK Household Income: Financial Year Ending 2018*. Newport: Office for National Statistics.

Ruiz, Nicolas, and Nicolas Woloszko. 2016. "What Do Household Surveys Suggest about the Top 1% Incomes and Inequality in OECD Countries?" OECD Economics Department Working Paper No. 1265. Paris: OECD.

Shine, Martin, Callum Clark, Richard P. Tonkin, and Dominic Webber. 2019. *Using Tax Data to Better Capture Top Earners in Household Income Inequality Statistics*. Newport: Office for National Statistics.