ROCKWOOL Foundation Berlin
Institute for the Economy and the Future of Work (RFBerlin)

DISCUSSION PAPER SERIES 241/26

The evolution of UK income inequality: a GB2 distribution perspective

Stephen P. Jenkins

www.rfberlin.com SEPTEMBER 2026

The evolution of UK income inequality: a GB2 distribution perspective

Authors

Stephen P. Jenkins

Reference

JEL Codes: C16, C46, D31

Keywords: Income inequality, Generalized Beta distribution of the Second Kind, Singh-Maddala distribution, Dagum distribution

Recommended Citation: Stephen P. Jenkins (2026): The evolution of UK income inequality: a GB2 distribution perspective. RFBerlin Discussion Paper No. 241/26

Access

Disclaimer

Imprint

RFBerlin
ROCKWOOL Foundation Berlin –
Institute for the Economy
and the Future of Work

Gormannstrasse 22, 10119 Berlin
Tel: +49 (0) 151 143 444 67
E-mail: info@rfberlin.com
Web: www.rfberlin.com

The evolution of UK income inequality: a GB2 distribution perspective

Stephen P. Jenkins
(LSE, IZA, and RF Berlin; s.jenkins@lse.ac.uk)

5 September 2026

Abstract

This paper fits four-parameter GB2 distributions to UK unit record household survey data for 42 years from 1977 to 2018. Parameter estimates are used to describe key features of the household income distribution and its evolution over time, including inequality. The paper also discusses and illustrates issues of statistical inference and data quality, and the relative goodness of fit of distributions that are special cases of the GB2. Singh-Maddala distributions typically fit better than Dagum distributions, and the two-parameter Fisk distribution fits the UK income distribution well for almost every year in the 2003–2018 period.

Acknowledgements

This paper was prepared for a *Journal of Income Distribution* special issue in honour of J. B. McDonald, edited by José María Sarabia. Jim, I salute you. Thanks to Nicolas Hérault, my collaborator on multiple projects, for help with the UK ONS's ETB data and advice more generally.

Keywords:

Income inequality, Generalized Beta distribution of the Second Kind, Singh-Maddala distribution, Dagum distribution

JEL classification codes: C16, C46, D31

1. Introduction

Jim McDonald's (1984) paper on 'Some generalized functions for the size distribution of income' is a landmark contribution. It established the four-parameter Generalized Beta Distribution of the Second Kind (GB2) as the 'go to' parametric model of income distributions. The paper demonstrated how key distributional features such as location and dispersion (inequality), as well as moments more generally and the Gini coefficient, were functions of the GB2 parameters, and that many other commonly used distributions were special or limiting cases of the GB2. McDonald's empirical application to US family income data for 1970, 1975, and 1980 found that the GB2 was the best-fitting distribution of those considered, though the three-parameter Singh-Maddala (SM) distribution fitted the data better than the four-parameter Generalized Beta Distribution of the First Kind and all the other two- and three-parameter distributions considered.

McDonald's 1984 paper led to further papers using the GB2 and its special cases by him and collaborators and inspired contributions by many other authors including myself. Kleiber and Kotz (2003) is an encyclopaedic study of the GB2 and other parametric distributions. Chotikapanich et al. (2018) provide a useful review of GB2 distribution analytics, estimation, and applications, with many references.

The current paper is firmly within this research tradition – the GB2 distribution is the foundational building block – and motivated by three observations. First, virtually all GB2 empirical applications to date have fitted distributions to grouped (banded) income data, not unit record data. This matters because analysts must often use the income distribution definitions (receiving unit, 'income' and its components, equivalence scale) in published sources and this can reduce the comparability of distributions across countries or between years. Even when analysts have had access to unit record data and been able to create consistently defined distributions, they have typically grouped the data for estimation purposes. Examples include Bandourian, McDonald, and Turley (2003) and Jorda, Sarabia, and Jäntti (2021), both using data from the Luxembourg Cross-National Data Center. Moreover, estimation from unit record data can provide more efficient estimates than grouped data, as McDonald (1984, p. 655) observed. The only GB2 studies to date using unit record data to describe national distributions are Brachmann, Stich, and Trede (1996), Graf and Nedyalkova (2014), and Jenkins (2009).

A second observation is that, in recent years, the most common empirical applications of parametric models to income distributions have been to comparisons across countries, world regions, and the global distribution. (For examples, see the references in Chotikapanich et al.'s (2018) review article.) There are very few GB2-based studies analyzing national income distributions over time. It is well-known that inequality has changed markedly in many countries over recent decades, but we are missing analysis of trends from a parametric modelling perspective. McDonald (1984) is an obvious early exception; two more recent exceptions are Brachmann et al. (1996) for West Germany (1984–1993, 10 years) and Jenkins (2009) for the UK (1994/5 and 2004/5).

A third observation is that software availability has improved. It is now much easier to fit GB2 distributions and undertake post-estimation calculations of statistics of interest and sampling variability. Having to use a three-parameter distribution like SM or Dagum for tractability reasons is less of an issue nowadays.

This paper builds on these three observations. It fits GB2 distributions to unit record household survey data for the UK and compares their fit to SM and Dagum distributions. The 'ETB' data I use are the same as used by the Office for National Statistics to calculate the UK's official inequality statistics, with definitions of income distributions consistent with international standards and employed by international statistical agencies such as the OECD and Eurostat. There are 42 years of data, yearly from 1977 through 2018, including a low inequality period (the late-1970s), a period of rapid inequality increase (the 1980s), and fluctuating inequality around a high level thereafter.

Model fitting and post-estimation calculations for this paper use my Stata packages for unit record data **`gb2lfit`** and siblings **`gb2fit`**, **`smfit`**, and **`dagumfit`**, freely downloadable since 2007 (latest versions 2014).[1] The programs allow users to take account of survey design features (weights, clustering, and strata), and there are options to fit each parameter as a linear index of covariates. My Stata programs complement the publicly available R packages of Graf and Nedyalkova (2014) for unit record data and Jorda, Sarabia, and Jäntti (2021) for grouped data.

Section 2 of the paper briefly reviews the GB2 distribution including its special cases, the SM and Dagum distributions, and provides some details of my methods for estimation, goodness of fit checks, and post-estimation calculations of inequality indices and their

[1] Stata users can **`search`** on program name and download each program using **`ssc`**. Examples of use include Biewen and Jenkins (2003; multi-country poverty differences), Jenkins (2009; UK inequality trends), Jenkins et al. (2011; imputation of top-coded top incomes) with applications in Burkhauser et al. (2011, 2012), and Hérault and Jenkins (2026; calibration of bivariate income distributions for Monte Carlo analysis of the statistical performance of variance estimates of inequality indices).

standard errors. Section 3 introduces the ONS's ETB data and the definitions of the income distribution.

Sections 4 and 5 are the empirical heart of the paper. I fit GB2, SM, and Dagum distribution and derive the inequality indices implied by the estimated parameters, as well as their standard errors. I show that the GB2 parametric approach summarizes well the key features of UK income distribution trends that have been described to date (and reported in official statistics) using nonparametric approaches applied to the observed survey income data. I also find – echoing McDonald (1984) and Brachmann et al. (1996) – that the SM distribution fits better than the Dagum distribution for years 1977–2002. But for 2003–2018 (except 2010), I cannot reject the null hypothesis that the UK income distribution is described by the two-parameter Fisk distribution, a special case of both the SM and Dagum distributions.

Section 6 contains concluding remarks.

2. GB2, SM, and Dagum distributions: specification, estimation, and post-estimation

2.1 Specification

As McDonald (1984) wrote, the GB2 distribution has probability density function (PDF),

$$f(x) = \frac{ax^{ap-1}}{b^{ap}B(p,q)[1+(x/b)^a]^{p+q}} \tag{1}$$

where $B(u, v)$ is the Beta function $\Gamma(u)\Gamma(v)/\Gamma(u+v)$, $\Gamma(u)$ is the Gamma function, and income $x > 0$. The distribution function (CDF) is (Kleiber and Kotz, 2003)

$$F(x) = IB\left(p, q, \frac{(x/b)^a}{[1+(x/b)^a]}\right) \tag{2}$$

where $IB(p, q, r)$ is the Incomplete Beta function.

Parameters $a > 0$, $p > 0$, and $q > 0$ are shape parameters. Larger values of a reduce the density in both left and right tails, and the relative sizes of p and q determine skewness. The product ap is a measure of left tail heaviness and aq is a measure of right tail heaviness; smaller values mean more tail heavy. The k^{th} moment of the GB2 distribution exists only if $-ap < k < aq$. Parameter $b > 0$ is a location parameter: mean income is a scalar multiple of b, with the multiple depending on the other three parameters. Relative (scale-independent) inequality indices are functions of a, p, and q, but not b. Whether one distribution Lorenz dominates another distribution depends on conditions relating to a, p, and q. Kleiber (1999) showed for two distributions A and B that, if $a_A \leq a_B$, $a_Ap_A \leq a_Bp_B$, and $a_Aq_A \leq a_Bq_B$, then A

Lorenz dominates B. Intuitively, one distribution is more equal than another if it has less density in the tails and is less tail heavy. Also, if A Lorenz dominates B, then $a_A p_A \leq a_B p_B$, and $a_A q_A \leq a_B q_B$ (Wilfling, 1996). As Kleiber (1999, p. 599) remarks, a complete characterization of the Lorenz order is not yet available, because his result does not cover the case $a_A \leq a_B$, $p_A \geq p_B$, $q_A \geq q_B$, but $a_A p_A \geq a_B p_B$, and $a_A q_A \geq a_{B q} q_B$.

Researchers have more commonly focused on specific scalar relative inequality indices. McDonald (1984) focused on the Gini coefficient – an index more sensitive to income differences in the middle of the distribution than in the tails ('middle sensitive') – showing it depends on 3F2 generalized hypergeometric functions of the parameters. These functions used to be unavailable in most statistical software, but that is either no longer the case, or the Gini can be calculated accurately by numerical integration instead. Jenkins (2009) derives expressions for the GB2 distribution for all Generalized Entropy indices of inequality.[2] The most commonly used members of this family are: GE(–1), a bottom-sensitive index; GE(0), the mean logarithmic deviation (MLD), a middle-sensitive index; GE(1), the Theil index (middle- to top-sensitive); and GE(2), half the squared coefficient of variation (HSCV) which is top-sensitive. Thus, researchers have available a portfolio of inequality indices encompassing the full range of sensitivities to the location of income differences. Since the conditions for Lorenz dominance may not hold, it is important to check whether inequality comparisons are robust to the choice of index.

The SM distribution, otherwise known as a Burr Type 12 distribution, is a GB2 distribution with $p = 1$, and the Dagum distribution, also known as a Burr Type 3 distribution, is a GB2 distribution with $q = 1$. The Fisk distribution arises when $p = q = 1$. These special cases have simpler expressions for their PDFS, CDFS, and inequality indices than the GB2 distribution, and Lorenz dominance conditions are fully characterized (Wilfling and Krämer, 1993; Kleiber, 1996).

2.2. Estimation

I estimate the GB2 parameters by maximizing the sample log pseudo-likelihood for each year's unit record data. The expression for the maximand is based on the PDF in eq. (1), suitably modified to account for survey design features such as weights, clustering, and

---

[2] There is an unfortunate typo in Jenkins's (2009, eqn. 4) expression for the Theil index. It should read "$I(1) = (v_1/\mu) - \log(\mu)$", where moment $v_1 = \int y \log y \, \mathrm{d}F(y)$. The calculations in **gb2lfit** and sibling programs are correct, however. They report the following inequality indices: Generalized Entropy GE(α) for $\alpha = -1, 0, 1, 2$; the Gini coefficient, the share of total income held by the top 10%, 5% or 1%, and the $p90/p10$ percentile ratio.

stratification where present.[3] Specifically, I use **`gb2lfit`** for maximization with a combination of Stata's BHHH and modified Newton-Raphson algorithms, also exploiting Stata's **`svy`** routines to account for survey design features. The maximization fits the GB2 parameters in a logarithmic metric to enforce positivity constraints (see above) and at the reporting stage back-transforms the fitted estimates to their natural metric with standard errors calculated using the delta method. I fit SM and Dagum models in the same way.

2.3. Post-estimation

Accounting for survey design features complicates assessment of model goodness of fit because it is a pseudo-likelihood that is maximized. Standard likelihood ratio tests are no longer valid, and nor are comparisons of AIC or BIC values. Consequently, I used several approaches for assessing goodness of fit. The first was visual inspection of P-P (probability) plots of the fitted CDF against the empirical CDF. The closer the plotted line to the 45° line, the better the fit. Graf and Nedyalkova (2014) also use P-P plots extensively. Unlike them, I do not also report my comparisons of parametrically estimated and kernel density-estimated PDFs because differences were too hard to discern in that metric. I also inspected Q-Q (quantile) plots – these provide a greater focus on the tails than P-P plots do – but do not report them because they provided similar conclusions to the P-P plots.

For comparisons of the goodness of fit of a GB2 distribution versus one of the GB2 special cases, I used a Wald test of the null hypotheses $\log(p) = 0$ for a SM distribution, and $\log(q) = 0$ for a Dagum distribution, and $\log(p) = \log(q) = 0$ for testing GB2 versus a Fisk distribution.

I calculate inequality indices and other distributional statistics using the expressions cited earlier and derive their standard errors using the delta method. Comparisons of the time series plot for each estimated index against its non-parametrically calculated index counterpart provide another perspective on the goodness of fit of the GB2 distribution compared to SM and Dagum distributions. An important caveat is that the non-parametrically calculated estimates may not provide an accurate benchmark if the relevant inequality index is particularly sensitive to outlier income values at the bottom or at the top of the distribution. The GE(–1) and GE(2) indices are most susceptible to this issue: see Cowell and Flachaire (2007) and Cowell and Victoria-Feser (1996).

---

[3] See Graf and Nedyalkova (2014, section 3) for elaboration.

3. The UK ONS's ETB data: 1977–2018[4]

My analysis is based on household survey data used by the UK Office for National Statistics (ONS) to prepare their annual articles about the 'Effects of taxes and benefits on household incomes' (ETB) until recently: see, e.g., ONS (2019) for 2018. The ONS data derive from the annual Living Costs and Food Survey (LCFS, from 2008) and its predecessor, the Family Expenditure Survey (FES, to 2007), each designed to be nationally representative of the UK private household population. The achieved sample size is approximately 6,500 households per year on average but ranging from around 7,500 in the 1980s to around 5,000 in the late 2010s. Survey years refer to financial years (12-month periods starting 5 April each year) from 1993/94 onwards and to calendar years before that. For brevity I label financial years by their first part: '2016' refers to financial year 2016/17, etc. ETB income data are not top-coded.

I use the ONS's definition of a household's income, namely 'equivalized household income', which is money income divided by the household's modified-OECD equivalence scale rate. Individuals are assumed to receive the equivalized income of the household to which they belong. Money income is net household income, i.e., household income from employment and self-employment, plus income from capital and government cash transfers, minus personal income tax payments, employee national insurance contributions, and local taxes such as council tax. Income variables are expressed in pounds per week pro rata and nominal incomes are not adjusted for inflation because relative inequality indices are scale invariant.

The ETB data include survey weights, but not cluster or strata information in the public use versions of the data to hand. Before 1996, a household's weight was simply the number of individuals in the household. From 1996 onwards, the weights also adjust for non-response and are calibrated to population totals. Prior to analysis, I trimmed the data lightly, dropping incomes that are (i) zero or negative, (ii) between zero and one, or (iii) top incomes that are more than twice the next highest income. For more details of the sample sizes and numbers dropped by trimming, see Hérault and Jenkins (2026, Appendix A). Many inequality indices are undefined for incomes less than or equal to zero and these values also raise questions about survey measurement error, hence rejection rule (i). Rules (ii) and (iii) reduce

[4] This section is taken almost verbatim from Hérault and Jenkins (2026).

the possibility that inequality index calculations produce estimates unduly affected by influential outliers, as mentioned above.

4. Estimates of GB2 parameters and inequality indices

4.1. *GB2 parameter estimates*

Figure 1 show time series of estimates for each of the four GB2 parameters accompanied by pointwise 95% confidence intervals (CIs). Estimates of *b* rise secularly over the 42-year period after the early 1980s recession, and with blips in the early 1990s and 2010s also corresponding to recession periods. Because my interest is in inequality, I focus discussion on *a*, *p*, and *q*. Two features are immediately apparent from the charts. First, sampling variability is non-negligible even though each annual survey sample size is around 5,000 households or more. Although each parameter is typically several times greater than its standard error (has a large 't-ratio'), the CIs are wide and typically overlap for pairs of years close to each other. Only for pairs of survey years from the start and end of the 1977–2018 period do CIs not overlap.

<Figure 1 near here>

Second, looking at the point estimates *a*, *p*, and *q*, there appear to be three subperiods. The first is from the late 1970s through to the early 1990s when *a* fell but *p* and *q* rose. Thus, the impact on the Lorenz curve is not immediately apparent, confirmed by Figure 2 which shows the time series of *ap* and *aq* estimates implied by the parameters. Although the density in the tails fell (a disequalizing influence), each of the two tail heaviness indices declined – skewness changed and both bottom and top tails became heavier. This means that different inequality indices may register these distributional changes differently.

<Figure 2 near here>

During the second subperiod, from the early 1990s up to 2003, the earlier trends reversed somewhat. The estimate of *a* increased whereas the estimates of *p* and *q* decreased, and the tail heaviness indices also fell albeit. However, it is important to note that, for both subperiods, these are general descriptions of temporal patterns, and there are year to year variations from the broad trends which will be reflected in scalar inequality indices, and the caveats about sampling variability remain relevant.

The third subperiod is from 2003 through to 2018 and is quite different from the earlier two. The estimate of *a* is broadly constant apart from an upward blip in 2011 (and note the wide CIs). More distinctly, the estimate of *p* hardly changed over the subperiod, and the fall in q is small, so the estimates of ap and aq also changed little. Consequently, one would expect inequality indices to change little. Since the subperiod estimates of p and q are each close to 1, the GB2 distribution looks like its special case, the Fisk distribution. I report tests of this hypothesis in Section 5.

4.2. *GB2 inequality index estimates*

Figure 3 shows time series of inequality index estimates and 95% CIs implied by the GB2 parameter estimates summarized in Figure 1 (and Appendix Table A1). Appendix Figure A1 shows the series for the bottom-sensitive GE(−1) index. All indices show the same patterns. UK income inequality was at its lowest level of the 42-period at the end of the 1970s, then increased through to around 1985, and increased much more rapidly in the following five years reaching a peak in 1990. After that, inequality levels fluctuated around the levels reached between 1985 and 1990. Income inequality in 2018 was about the same as inequality as in around 1989; e.g., the Gini coefficients are 0.323 and 0.324 respectively.

Again, one should also note the CIs around the estimates. The CIs are sufficiently large that finding a statistically significant difference between a pair of inequality indices is critically dependent on the pair of years being compared. Inequality differences between a year in the late 1970s and any year from the mid-1980s onwards are statistically significant but differences between pairs of years from around 1990 through 2018 rarely are (test statistics not shown for brevity). Another way of interpreting the patterns is to say that tests of inequality difference are under-powered because the sample sizes of the annual household budget survey used to compile the ETB income data are insufficiently small to reliably detect small inequality index differences.

The general remarks in this section have made no distinction between the inequality indices but there are some small differences related to how top-sensitive the indices are. Specifically, observe that the trend in the 20 years after 1990 in estimates of the middle to top sensitive Theil index is slightly flatter than for the middle sensitive Gini and MLD, and the estimates for the top sensitive HSCV index appear on a slight upward trend. This corresponds to the earlier finding of an increase in top tail heaviness over this period (the falling aq estimates shown in Figure 2).

<Figure 3 near here>

It is also of interest to compare the GB2 estimates of the time series of inequality indices with the usual narratives about UK inequality trends since the late 1970s. The main sources for information about these are the ONS's own estimates from ETB data (derived from the household budget survey) and the Department for Work and Pensions' estimates derived from their Family Resources Survey (FRS). Both focus on the Gini coefficient. The Institute for Fiscal Studies also provides an annual inequality commentary with accompanying spreadsheet containing inequality index time series derived from the household budget survey for 1961–1994 and the FRS since 1994/95 onwards (the financial year the FRS was launched): see, e.g., IFS (2026). Income distribution definitions are as described earlier in all three sources cited. The IFS spreadsheet contains estimates of the MLD, Theil, and HSCV, in addition to the Gini coefficient. Each of the three organisations calculates inequality indices non-parametrically directly from the unit record data and applying the relevant survey weights – what I label the 'empirical' method below.

The most important conclusion from a comparison of the ETB-based GB2 inequality index series with the series reported by the ONS, Department for Work and Pensions (DWP), and IFS, is that all four sets of estimates tell the same story about UK inequality trends since the late 1970s. In short, inequality rose rapidly during the 1980s, peaked around 1990, and has fluctuated around that high level for the following 35 years. Put differently, the good news is that a GB2-based approach to summarizing inequality trends does well relative to the official and IFS benchmarks.

The GB2-based approach has the advantage that it also provides standard errors straightforwardly. This is relevant because the official and IFS series do not report standard errors although, in principle, they could be calculated from the unit record data.[5] Hérault and Jenkins (2026) provide a detailed analysis of methods for statistical inference about trends in empirical estimates of UK income inequality. One of their headline results is that few of the inequality differences between pairs of years from 1990 onwards are statistically significant, and they point out that this is in part because the relevant tests are under-powered given the sample size of the household budget surveys underlying the ETB data – the point made earlier in the discussion of inference using GB2 distributions.

4.3. *A comparison of GB2 estimates from three UK household surveys*

It is interesting to compare the ETB-based estimates with those derived from other UK surveys given the remarks about sample size, also noting that the ETB data are derived from

[5] In part, this is because the DWP and ONS focus on non-sampling sources of survey error such as under-reporting at the top of the income distribution, and they report rounded estimates so that report readers do not conclude small differences are statistically significant. See Jenkins (2022) for a review of the UK's official income inequality statistics, including the top-income adjustments they now include.

annual household budget surveys collecting expenditure as well as income data, whereas until recently the UK had two other annual household surveys each of which focused on collecting income data. There is the FRS, mentioned above, which had a sample of up to 25,000 households over the period considered by this paper. The FRS is the source of the UK's official statistics on poverty ('households below average income'). The Survey of Living Conditions (SLC) was the UK survey component of the European Union Statistics on Income and Living Conditions database and had a sample size of around 12,000 households. (The SLC was discontinued in 2025.) Despite their different purposes and sample sizes, all three surveys provide data on a household income variable defined exactly as in the current paper (equivalised net household income among individuals), and all three aim to represent the UK population of individuals in private households. There are some differences in survey design features, but cluster and strata information are unavailable in the public-use files of the ETB and FRS data. For a comparative overview of the three surveys, see ONS (2024).

To maximize comparability, I consider only GB2 estimates derived from unit record data (though I am unaware of any grouped data estimates). What I have found is the FRS-based estimates reported by Jenkins (2009) for financial years 1994/95 and 2004/05, i.e., corresponding to '1994' and '2004' in the ETB shorthand used elsewhere in this paper, and the SLC-based estimates reported by Graf and Nedyalkova (2014, Table 3) for calendar year 2006 which I compare with ETB data for 2006/07. Graf and Nedyalkova report parameter estimates, but not standard errors or inequality indices. I calculated index estimates from the parameter estimates using the standard GB2 formulae.

Table 1 reports the three sets of comparisons of ETB-based estimates with FRS- and SLC-based estimates. There are some differences in parameter estimates, most notably for the two ETB-FRS comparisons. For example, for 1994/95, the estimates of a are around 2 for the ETB-based one and 3 for the FRS-based one; for p, they are 3.3 and 1; and for q, they are 2.7 and 1. These ETB estimates imply a distribution with greater density in the tails but also less top and bottom heaviness than the distribution implied by the FRS estimates. However, observe the differences in precision as well. Each ETB parameter estimate has a much larger standard error than its FRS counterpart, which is expected given FRS sample sizes are around four times larger than LCFS sample sizes. Nonetheless, formal tests of the equality of corresponding parameter estimates reject the null. The same is true for 2004/05 comparisons, though test statistics are smaller.

<Table 1 near here>

Differences between ETB- and FRS-based inequality estimates and standard errors are less apparent than differences in parameter estimates. For both 1994/95 and 2004/05, the difference in Gini estimates is not statistically significant. But all four other inequality index estimates for 1994/95 are statistically significant, with the test statistic larger the more top sensitive the index. For 2004/05, there is a similar association with top sensitivity, but test statistics are a bit smaller and the GE(–1) difference is not significant.

Perhaps of greater interest is whether the ETB and FRS data provide different conclusions about GB2-calculated inequality changes between 1994/95 and 2004/05. Jenkins (2009) reported that the rise in inequality between the two years was statistically significant only if one used the top sensitive HSCV. However, if one uses the ETB data rather than the FRS data, Table1 shows this conclusion does not survive. This is another demonstration of the point that the ETB data are under-powered for reliable detection of small inequality index differences.

Turning finally to the ETB-SLC comparison for 2006, parameter estimates are similar and differences unlikely to be statistically significant. The ETB inequality index estimates are each slightly larger than their SLC counterparts and some of the differences are likely to be statistically significant. For example, if SLC standard errors were the same as ETB ones – a plausible upper bound given the larger SLC sample size – the Gini and Theil index differences would be significant.

All in all, these cross-survey comparisons provide a salutary reminder that the choice of dataset matters for the conclusions drawn. Even if survey definitions are fully comparable, differences in sample size and thence in sampling variability can have a non-trivial effect on assessments of the statistical significance of inequality index differences.

5. Goodness of fit: GB2 versus SM, Dagum, and Fisk distributions

I now turn to consider the goodness of fit of the GB2 distribution relative to its special cases, especially SM and Dagum distributions. The Fisk distribution is also relevant: see the discussion of Figure 1, and also note that, according to FRS data, a Fisk distribution fitted the UK income distribution well in 1994/95 (Table 1).

Figure 4 shows P-P plots for GB2, SM, and Dagum distributions. Recall that the closer the plot to the 45° line, the better the fit to the empirical distribution. For brevity I focus on 1977, 1990, and 2018. Inspection of the charts for all 42 years shows that the patterns for these three years are representative of the start of the period (when inequality was low), the period around 1990 (high inequality), and the years leading up to and including 2018 (high but fluctuating inequality) respectively.

<Figure 4 near here>

All three distributions, GB2 and the two leading special cases, fit the data well for the three years shown, but do least well for 1990 when inequality was at a peak. There is a tendency for the fitted parametric distributions to slightly over-estimate concentration in the poorest tenth of the income range and slightly under-estimate concentration thereafter up to around the 25th percentile, and these differences are more pronounced for 1990.

For 1977 and 1990, the GB2 distribution fits the empirical data better than both SM and Dagum distributions, with little visible difference between SM and Dagum fits. Observe the additional differences in plots between the median and 75th percentiles of the empirical distribution that are apparent for 1990 in particular. For 2018, the three P-P plots are remarkably similar, and all fit the data well, better than for 1977.

Figure 5 provides additional evidence about the relative fit of the parametric models to the empirical data, showing time series of inequality indices comparing series derived using the non-parametric empirical estimator and using the GB2, SM, and Dagum parameter estimates. (The GB2 point estimates are the same as those shown in Figure 3 earlier.)

For the period as a whole, the GB2 distribution provides the series that is closest to the empirical estimate series benchmarks, most obviously for the Gini, MLD, and Theil indices but also for the top-sensitive HSCV. The SM distribution performs better than the Dagum distribution, again for the four indices shown in Figure 4, especially for HSCV. This likely reflects the fact that the SM distribution allows greater flexibility in top tail heaviness than does the Dagum distribution. Correspondingly, the Dagum distribution performs at least as well as the SM distribution for the bottom sensitive GE(–1) index (see Appendix Figure A2).[6]

Superior performance of the SM distribution relative to the Dagum distribution is also reported by Brachmann et al. (1996) in their analysis of unit record data for Western Germany. It is at variance with Kleiber's remarks that "[a] look at the table 1 from McDonald and Mantrala (1995) reveals that, with their data, the Dagum performs almost as well as the

[6] Figure A2 also shows that none of the three distributions track the empirical GE(–1) series well between 2001 and 2018; the empirical series shows higher levels as well as marked temporal fluctuations. This likely reflects a relatively high prevalence of low-income outliers for some years during this period that remained after the data trimming described earlier.

GB2 (and much better than the Singh-Maddala!)” (1999, p. 267). However, it should be remembered that McDonald and Mantrala’s estimates refer to grouped US data for 1990. Bandourian, McDonald, and Turley (2003) compare the goodness of fit of many distributions using 82 datasets from the Luxembourg Cross-National Data Center, and report that the Dagum distribution typically provided the best fit of all the three-parameter distributions considered, including the SM. However, again, the data were grouped prior to estimation, the income definition was gross household income (income before the deduction of taxes or inclusion of government cash transfers among households, unequivalized), and derivations of vigintile shares and boundaries apparently did not use survey weights. The relative goodness of fit of SM and Dagum distributions may well depend on the income definitions used.

An additional feature of the inequality series shown in Figure 4 is that all three distributions provide almost identical estimates for the inequality indices for the 2003–2018 period, which is what we’d expect from Figure 1 showing estimates of p and q both equal to approximately 1 over this period.

I have undertaken tests of null hypotheses that the estimates of these two parameters are separately or jointly equal to one. Appendix Table A2 shows the p-values from these tests for each of the 42 years. With 5% as the statistical significance threshold, I reject the SM and Dagum distributions in favour of the GB2 distributions for all years between 1977 and 2002 (the SM distribution in 2001 is an exception). However, for each year from 2003 to 2018 (except 2010), I cannot reject the null hypothesis that the estimates of p and q are both equal to one. Thus, over this period, the two-parameter Fisk distribution appears sufficient to describe the UK income distribution. The flexibility provided by the four-parameter GB2 distribution is not required.

6. Concluding remarks

GB2 and related distributions are rarely fitted to unit record data or used to study trends in income distribution over time. This paper goes some way to filling that gap using 42 years of data. These are the same data as used by the ONS to compile the UK’s official statistics on income inequality, for which they calculate indices non-parametrically from the empirical distributions. I have shown how one can use the GB2 distribution and its special cases to summarise key distributional features such as location, skewness and tail heaviness, and inequality, and their evolution over time, while also taking sampling variability into account.

I have also shown how the parametric and empirical approaches are both contingent on the nature of the survey data available. In addition to the need for comparable definitions of income variables across datasets, I have illustrated out how the sample sizes of well-known household surveys may constrain analysts' ability to infer whether inequality index differences are statistically significant.

In short, using a GB2-based parametric approach, one can capture almost all the features of the UK's income distribution, and its trends, that the more conventional empirical approach is usually used to summarize. Jim McDonald's foundational research on GB2 and related distributions made this possible.

References

Bandourian, R., McDonald, J. B., and Turley, R. S. (2003). A comparison of parametric models of income distribution across countries and over time. *Estadística*, 55 (164–165), 135–152. (No registered DOI) Preprint copy available as LIS Working Paper 305, 2002. https://www.lisdatacenter.org/wps/liswps/305.pdf

Biewen, M. and Jenkins, S. P. (2005). A framework for the decomposition of poverty differences with an application to poverty differences between countries. *Empirical Economics*, 30, 331–358. https://doi.org/10.1007/s00181-004-0229-1

Brachmann, K., Stich, A., & Trede, M. (1996). Evaluating parametric income distribution models. *Allgemeines Statistisches Archiv*, 80, 285–298. (No registered DOI)

Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2011). Estimating trends in US income inequality using the Current Population Survey: the importance of controlling for censoring. *Journal of Economic Inequality*, 9 (3), 393–415. https://doi.org/10.1007/s10888-010-9131-6

Burkhauser, R. V., Feng, S., Jenkins, S. P., and Larrimore, J. (2012). Recent trends in top income shares in the USA: reconciling estimates from March CPS and IRS tax return data. *Review of Economics and Statistics*, 94 (2), 371–388. https://doi.org/10.1162/REST_a_00200

Chotikapanich, D., Griffiths, W. E., Hajargasht, G., Karunarathne, W., and Rao, D. S. P. (2018). Using the GB2 income distribution, *Econometrics*, 6, 21. https://doi.org/10.3390/econometrics6020021

Cowell, F. A. & Flachaire, E. (2007). Income distribution and inequality measurement: the problem of extreme values. *Journal of Econometrics*, 141, 1044–1072. https://doi.org/10.1016/j.jeconom.2007.01.001

Cowell, F. A. & Victoria-Feser, M.-P. (1996). Robustness properties of inequality measures. *Econometrica*, 64, 77–101. https://doi.org/10.2307/2171925

Graf, M. and Nedyalkova, D. (2014). Modeling of income and indicators of poverty and social exclusion using the generalized beta distribution of the second kind. *Review of Income and Wealth*, 60, 821–842. https://doi.org/10.1111/roiw.12031

Hérault, N and Jenkins, S. P. (2026). Assessing the statistical significance of pairwise differences in income inequality indices using weighted survey data. Paper presented at the 39th General Conference of the International Association for Research on Income and Wealth, Brussels. https://virtual.oxfordabstracts.com/event/75057/submission/137

Institute for Fiscal Studies (2026). Living Standards, Poverty and Inequality in the UK, https://www.ifs.org.uk/tools_and_resources/ incomes_in_uk. London: Institute for Fiscal Studies.

Jenkins, S. P. (2009). Distributionally-sensitive inequality indices and the GB2 income distribution. *Review of Income and Wealth*, 55 (2), 392–398. https://doi.org/10.1111/j.1475-4991.2009.00318.x

Jenkins, S.P. (2022). Top-income adjustments and official statistics on income distribution: the case of the UK. *Journal of Economic Inequality*, 20, 151–168. https://doi.org/10.1007/s10888-022-09532-y

Jenkins, S. P., Burkhauser, R. V., Feng, S., and Larrimore, J. (2011). Measuring inequality using censored data: a multiple-imputation approach to estimation and inference. *Journal of the Royal Statistical Society (Series A)*, 174 (1), 63–81. https://doi.org/10.1111/j.1467-985X.2010.00655.x

Jorda, V., Sarabia, J. M., and Jäntti, M. (2021). Inequality measurement with grouped data: parametric and non-parametric methods, *Journal of the Royal Statistical Society (Series A)*, 184, 964–984. https://doi.org/10.1111/rssa.12702

Kleiber, C. (1999). On the Lorenz order within parametric families of income distributions. *Sankhyā*, B61, 514–517. https://scispace.com/pdf/on-the-lorenz-order-within-parametric-families-of-income-53j8g8jpmu.pdf

McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica*, 52 (3), 647–665. https://doi.org/10.2307/1913469

McDonald, J. B. and Mantrala, A. (1995). The distribution of personal income: revisited. *Journal of Applied Econometrics*, 10, 201–204. https://doi.org/10.1002/jae.3950100208

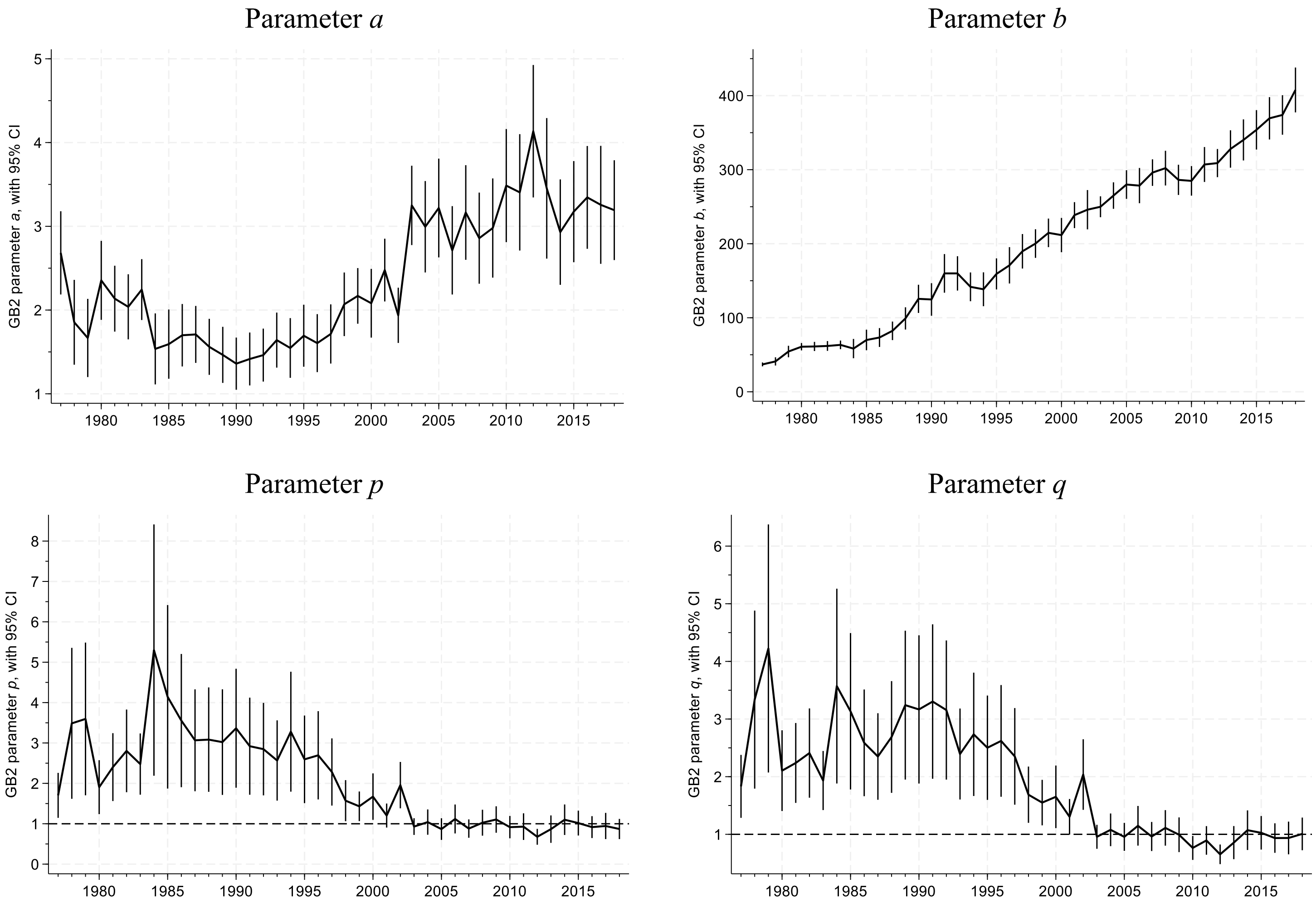
Office for National Statistics (2019). *Effects of taxes and benefits on UK household income: financial year ending 2018*. London: Office for National Statistics. https://www.ons.gov.uk/peoplepopulationandcommunity/personalandhouseholdfinances/incomeandwealth/bulletins/theeffectsoftaxesandbenefitsonhouseholdincome/financialyearending2018

Office for National Statistics (2024). *Income statistics: coherence and comparison information*. London: Office for National Statistics. https://www.ons.gov.uk/file?uri=/employmentandlabourmarket/peopleinwork/earningsandworkinghours/methodologies/aguidetosourcesofdataonearningsandincome/incomestatisticscoherenceandcomparisonreport.pdf

Wilfling, B. (1996). Lorenz ordering of Generalized Beta-II income distributions. *Journal of Econometrics*, 71, 381–388. https://doi.org/10.1016/0304-4076(94)01710-7

Wilfling, B. and Krämer, W. (1993). The Lorenz ordering of Singh-Maddala income distributions. *Economics Letters*, 43, 53–57. https://doi.org/10.1016/0165-1765(93)90134-X

Figure 1. GB2 parameter estimates, with 95% CIs, by year



Note. Estimates and their standard errors are reported in Table A1.

Figure 2. GB2 tail indices, with 95% CIs, by year

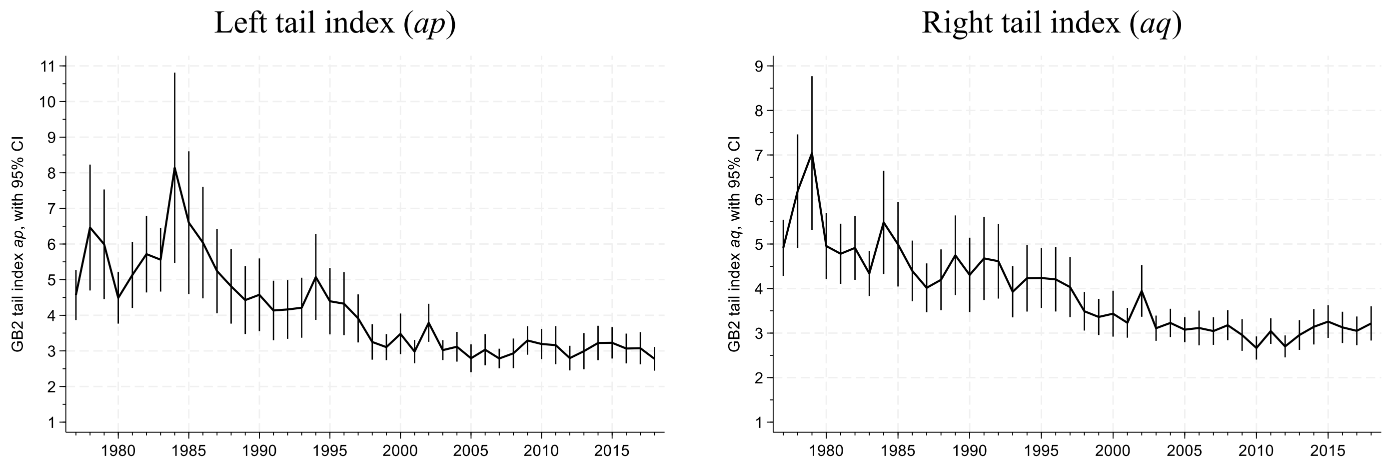


Figure 3. GB2-estimated inequality index estimates, with 95% CIs, by year

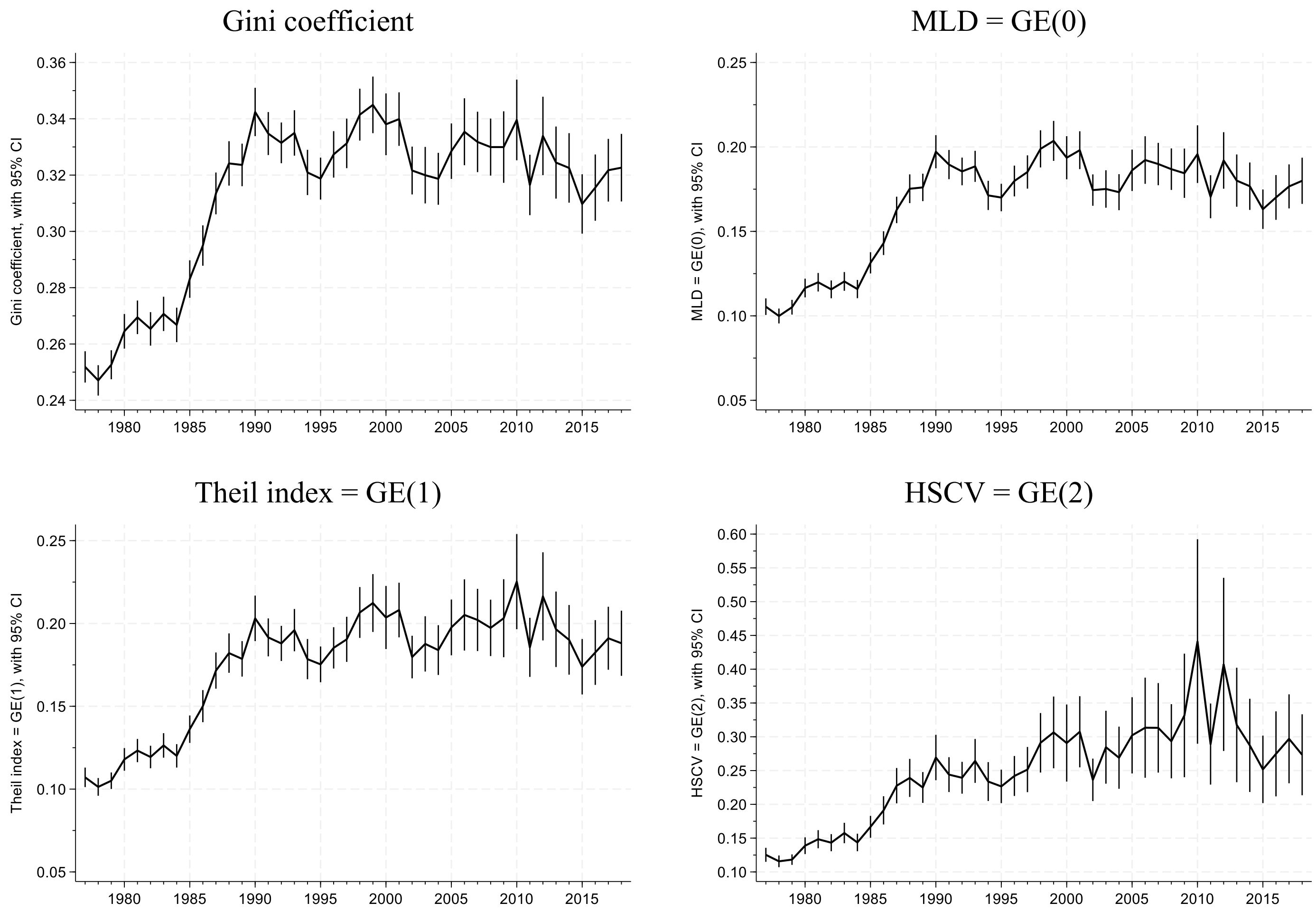


Figure 4. GB2 P-P plots, for 1977, 1990, and 2018

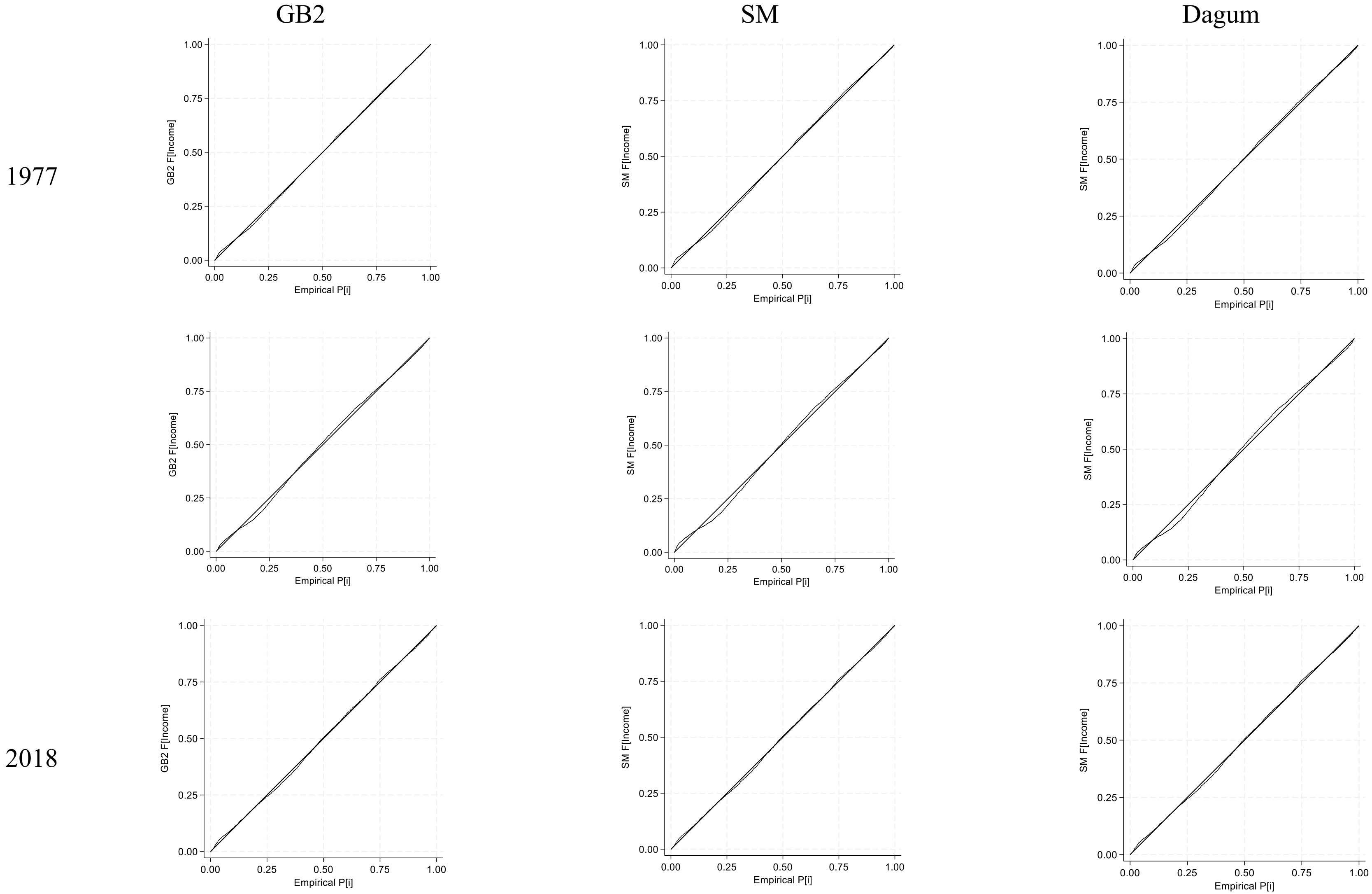


Figure 5. Inequality index estimates: empirical, GB2, SM, and Dagum, by year

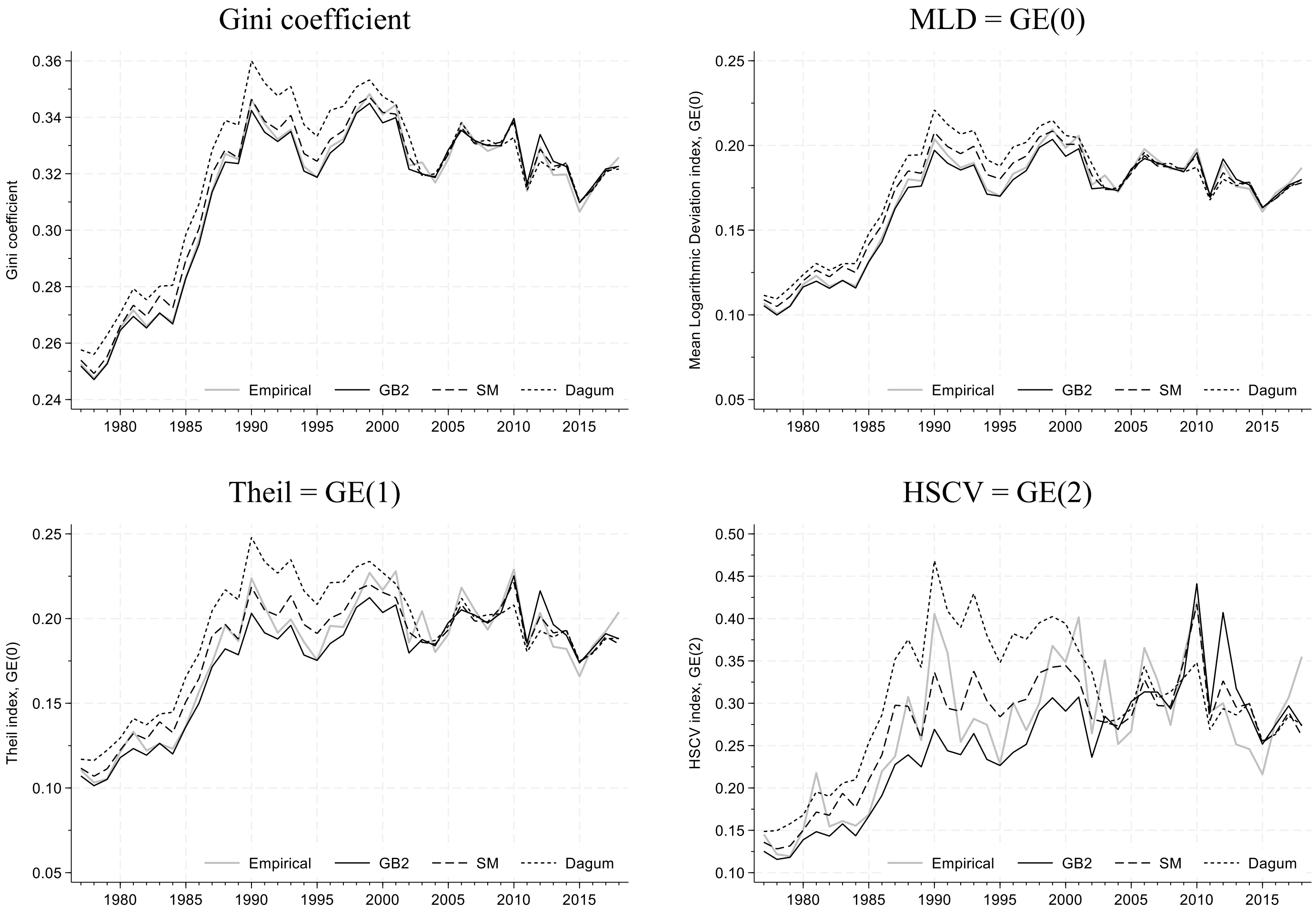
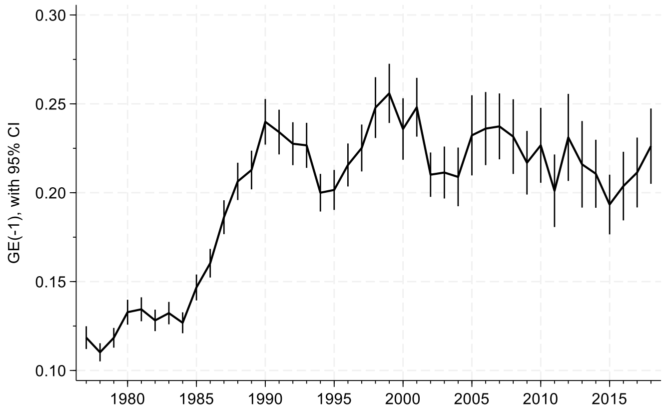


Table 1. GB2 parameters and inequality indices: ETB data estimates compared with estimates from other UK surveys

	1994/95		2004/05		2006	
	ETB	FRS	ETB	FRS	ETB	SLC
a	1.947	2.994	2.994	4.257	2.713	2.803
	(0.182)	(0.057)	(0.278)	(0.179)	(0.269)	
p	3.279	1.063	1.041	0.682	1.118	0.976
	(0.758)	(0.072)	(0.160)	(0.042)	(0.183)	
q	2.735	1.015	1.078	0.635	1.148	1.329
	(0.546)	(0.058)	(0.145)	(0.037)	(0.175)	
Inequality indices						
Gini	0.321	0.327	0.319	0.329	0.335	0.316
	(0.004)	(0.003)	(0.005)	(0.003)	(0.006)	
GE(–1)	0.200	0.217	0.209	0.220	0.236	0.224
	(0.005)	(0.005)	(0.008)	(0.005)	(0.010)	
MLD = GE(0)	0.171	0.182	0.173	0.186	0.192	0.173
	(0.004)	(0.003)	(0.005)	(0.004)	(0.007)	
Theil = GE(1)	0.178	0.198	0.184	0.211	0.205	0.173
	(0.006)	(0.005)	(0.007)	(0.006)	(0.011)	
HSCV= GE(2)	0.234	0.310	0.269	0.397	0.313	0.224
	(0.015)	(0.017)	(0.023)	(0.030)	(0.038)	

Notes. ETB estimates are described in current paper (see also Appendix Table A1, estimates for 1994, 2004, 2006). FRS estimates are based on Family Resources Survey unit record data, and taken from Jenkins (2009, Table 1). The SLC estimates are from Graf and Nedyalkova (2014, Table 3, 'PML full' estimates; no SEs reported). All three studies use the same definition of income (modified-OECD-scale-equivalized household net income among individuals). All estimates shown in the table above are derived using survey weights. (Only the SLC estimates use cluster and strata information.) Estimates for *b* are available from each study but not reported here. The current author calculated SLC-based inequality indices using the parameters reported by Graf and Nedyalkova (2014, Table 3).

Figure A1. GB2-estimated inequality index GE(–1), with 95% CIs, by year



Note. Estimates for other inequality indices are shown in Figure 3 of the main text.

Figure A2. Inequality index GE(–1) estimates: empirical, GB2, SM, and Dagum, by year

Note. Estimates for other inequality indices are shown in Figure 5 of the main text.

Table A1. GB2 parameter estimates and SEs

Year	*a*	SE(*a*)	*b*	SE(*b*)	*p*	SE(*p*)	*q*	SE(*q*)	*ap*	SE(*ap*)	*aq*	SE(*aq*)	p-value of test		
													log(*p*) = 0	log(*q*) = 0	log(*p*) = log(*q*) = 0
1977	2.683	0.254	36.968	1.328	1.703	0.284	1.832	0.279	4.569	0.359	4.915	0.322	0.001	0.000	0.000
1978	1.854	0.258	41.023	2.826	3.486	0.954	3.337	0.788	6.463	0.902	6.186	0.651	0.000	0.000	0.000
1979	1.667	0.238	54.383	3.953	3.596	0.965	4.224	1.098	5.993	0.785	7.041	0.882	0.000	0.000	0.000
1980	2.354	0.240	60.880	2.539	1.906	0.340	2.104	0.358	4.489	0.369	4.955	0.378	0.000	0.000	0.000
1981	2.136	0.201	61.221	3.161	2.402	0.429	2.239	0.354	5.132	0.472	4.782	0.344	0.000	0.000	0.000
1982	2.039	0.198	61.898	3.360	2.804	0.522	2.410	0.395	5.717	0.549	4.913	0.365	0.000	0.000	0.000
1983	2.244	0.185	63.373	2.976	2.478	0.387	1.933	0.261	5.561	0.457	4.339	0.258	0.000	0.000	0.000
1984	1.536	0.216	58.348	6.642	5.300	1.588	3.572	0.862	8.141	1.363	5.486	0.592	0.000	0.000	0.000
1985	1.593	0.211	69.977	7.097	4.143	1.159	3.135	0.692	6.600	1.022	4.994	0.483	0.000	0.000	0.000
1986	1.699	0.190	73.322	6.493	3.555	0.842	2.587	0.472	6.040	0.799	4.396	0.348	0.000	0.000	0.000
1987	1.709	0.173	82.340	6.386	3.067	0.644	2.350	0.383	5.241	0.605	4.017	0.280	0.000	0.000	0.000
1988	1.561	0.171	99.110	7.673	3.083	0.660	2.689	0.494	4.813	0.534	4.197	0.348	0.000	0.000	0.000
1989	1.465	0.171	125.550	9.714	3.022	0.667	3.241	0.659	4.426	0.486	4.748	0.456	0.000	0.000	0.000
1990	1.360	0.159	124.723	11.216	3.365	0.751	3.167	0.656	4.576	0.520	4.307	0.427	0.000	0.000	0.000
1991	1.416	0.161	159.940	13.316	2.921	0.613	3.304	0.683	4.135	0.427	4.677	0.477	0.000	0.000	0.000
1992	1.462	0.161	159.927	11.837	2.848	0.584	3.156	0.616	4.163	0.421	4.615	0.429	0.000	0.000	0.000
1993	1.641	0.168	141.774	9.917	2.566	0.507	2.393	0.403	4.212	0.430	3.928	0.293	0.000	0.000	0.000
1994	1.547	0.182	138.484	11.670	3.279	0.758	2.735	0.546	5.074	0.614	4.232	0.382	0.000	0.000	0.000
1995	1.693	0.189	159.205	10.709	2.595	0.553	2.502	0.462	4.395	0.475	4.237	0.343	0.000	0.000	0.000
1996	1.605	0.177	170.839	12.531	2.695	0.558	2.621	0.494	4.326	0.450	4.206	0.368	0.000	0.000	0.000
1997	1.715	0.180	189.790	11.888	2.282	0.424	2.352	0.428	3.912	0.344	4.033	0.344	0.000	0.000	0.000
1998	2.069	0.194	200.238	9.864	1.572	0.259	1.687	0.249	3.251	0.254	3.491	0.221	0.006	0.000	0.001
1999	2.169	0.169	214.694	9.845	1.432	0.187	1.551	0.201	3.106	0.187	3.363	0.207	0.006	0.001	0.003
2000	2.081	0.209	211.695	11.805	1.672	0.293	1.651	0.277	3.479	0.291	3.436	0.263	0.003	0.003	0.009
2001	2.477	0.191	238.625	8.940	1.203	0.152	1.304	0.157	2.981	0.167	3.231	0.171	0.142	0.028	0.061
2002	1.937	0.168	245.984	13.542	1.956	0.293	2.037	0.312	3.789	0.273	3.946	0.295	0.000	0.000	0.000

2003	3.249	0.242	249.944	7.210	0.930	0.105	0.957	0.106	3.022	0.142	3.110	0.144	0.522	0.692		0.775
2004	2.994	0.278	265.041	9.137	1.041	0.160	1.078	0.145	3.117	0.213	3.229	0.162	0.795	0.574		0.734
2005	3.219	0.301	279.955	9.885	0.868	0.137	0.956	0.124	2.793	0.199	3.077	0.143	0.368	0.728		0.399
2006	2.713	0.269	278.511	12.145	1.118	0.183	1.148	0.175	3.034	0.223	3.115	0.199	0.494	0.366		0.649
2007	3.165	0.288	296.068	9.145	0.881	0.116	0.962	0.129	2.787	0.140	3.046	0.158	0.335	0.774		0.264
2008	2.858	0.278	302.111	11.966	1.026	0.164	1.112	0.157	2.932	0.212	3.177	0.172	0.874	0.455		0.426
2009	2.979	0.302	286.294	10.409	1.105	0.167	0.993	0.154	3.293	0.204	2.958	0.181	0.508	0.962		0.298
2010	3.486	0.344	284.909	10.190	0.916	0.141	0.765	0.106	3.194	0.216	2.665	0.133	0.570	0.053		0.009
2011	3.405	0.355	307.020	12.030	0.928	0.168	0.894	0.127	3.162	0.273	3.045	0.146	0.682	0.430		0.591
2012	4.135	0.403	308.916	9.705	0.677	0.100	0.653	0.087	2.798	0.176	2.701	0.126	0.008	0.001		0.006
2013	3.453	0.428	327.986	12.873	0.867	0.173	0.856	0.147	2.993	0.259	2.957	0.170	0.474	0.365		0.641
2014	2.931	0.321	340.235	14.155	1.100	0.193	1.072	0.175	3.223	0.246	3.141	0.203	0.588	0.671		0.857
2015	3.175	0.308	353.843	13.560	1.017	0.157	1.026	0.148	3.227	0.225	3.258	0.187	0.915	0.858		0.979
2016	3.345	0.313	369.426	14.586	0.917	0.138	0.935	0.131	3.067	0.213	3.129	0.179	0.565	0.632		0.847
2017	3.256	0.359	373.959	13.609	0.944	0.166	0.937	0.145	3.076	0.230	3.051	0.164	0.745	0.673		0.906
2018	3.193	0.304	407.731	15.464	0.870	0.128	1.007	0.145	2.778	0.170	3.217	0.196	0.342	0.959		0.141