Pareto and the upper tail of the income distribution in the UK: 1799 to the present

A B Atkinson

Contents

CASE/198
October 2016

Centre for Analysis of Social Exclusion
London School of Economics
Houghton Street
London WC2A 2AE
CASE enquiries – tel: 020 7955 6679

Centre for Analysis of Social Exclusion

The Centre for the Analysis of Social Exclusion (CASE) is a multi-disciplinary research centre based at the London School of Economics and Political Science (LSE), within the Suntory and Toyota International Centres for Economics and Related Disciplines (STICERD). Our focus is on exploration of different dimensions of social disadvantage, particularly from longitudinal and neighbourhood perspectives, and examination of the impact of public policy.

In addition to our discussion paper series (CASEpapers), we produce occasional summaries of our research in CASEbriefs, and reports from various conferences and activities in CASEreports. All these publications are available to download free from our website. Limited printed copies are available on request.

For further information on the work of the Centre, please contact the Centre Manager, Jane Dickson, on:

Telephone: UK+20 7955 6679
Fax: UK+20 7955 6951
Email: j.dickson@lse.ac.uk
Web site: http://sticerd.lse.ac.uk/case

Editorial note

This paper was prepared for a special issue of *Economica* in honour of Frank Cowell, who has contributed so much to the success of the journal and with whom it has been a pleasure to work over five decades. The choice of subject recognizes his long-standing interest in the Pareto distribution and his major contribution to understanding its determinants (for example, Cowell, 1977 and Champernowne and Cowell, 1998). I am most grateful to Andrea Brandolini, of the Bank of Italy, and to the journal's referees for their constructive suggestions that have led to considerable improvements.

Tony Atkinson is an Honorary Fellow at Nuffield College, Oxford, Senior Research Fellow at INET at Oxford Martin School, and a Centennial Professor at London School of Economics. Email address for correspondence: tony.atkinson@nuffield.ox.ac.uk.

Abstract

The Pareto distribution has long been a source of fascination to economists, and the Pareto coefficient is widely used, in theoretical and empirical studies, as a summary of the degree of concentration of top incomes. This paper examines the empirical evidence from income tax data concerning top incomes in the UK, contrasting the dramatic changes that took place in the twentieth century, after 1918, with the much more modest changes in the preceding nineteenth century. Probing beneath the surface, it identifies a number of features of the evolution of the UK income inequality that warrant closer attention. These include the changing shape of the upper tail, where there is a link with Pareto's theory of elites, the need for a richer functional form to describe top incomes, and the limited evidence at the top of the distribution for a Kuznets curve in nineteenth century Britain.

Key words: Pareto, income, distribution, tail

JEL number: D63, I31, N33

Introduction: Pareto and the upper tail

The upper tail of the income distribution has long been a source of fascination to economists, and the Pareto curve has featured extensively in empirical and theoretical studies. Much of the literature on theoretical models of income distribution has been concerned with the generation of a thick upper tail of the Pareto form (for a recent review, see Benhabib and Bisin, 2016). This paper focuses on its empirical application to the upper tail, making use of historical data on incomes and earnings in the United Kingdom (UK) derived from the administration of the income tax from 1799 to the present day. As the title indicates, the point of departure is the Pareto coefficient, alpha, which is typically interpreted as an inverse measure of the concentration of top incomes (for a clear discussion of its relation to measures of inequality, see Chipman, 1974, and, earlier, Bresciani-Turroni, 1939): the lower the value of alpha, the more concentrated the distribution.

Specifically, the Pareto distribution holds where the proportion of the population with incomes of y or higher, referred to as the "survivor function", S(y), is given by

$$S(y) = A\ y^{-\alpha} \qquad (1)$$

where A is a constant. (S(y) is more commonly written as 1-F(y), where F is the cumulative distribution, but S(y) is a more economical notation here.) With the data employed in this study, in 2013/14, S(£100,000) = 0.014 means that 1.4 per cent of the population aged 15 and over received a gross income in that year in excess of £100,000. Equation (1) can be re-written as

$$\log_e[S] = \log_e A - \alpha \log_e[y] \qquad (2a)$$

In other fields of economics, such as the distribution of firm size, equation (2a) with $\alpha = 1$ is known as Zipf's Law, after the work of Zipf (1932 and 1949) on word frequency (see Gabaix, 2009, for a review). Equation (2a) may be re-written in terms of Pen's Parade (Cowell, 1977, p.19) giving (in logarithmic form) the income corresponding to different ranks in the distribution (as measured by 1/S)

$$\log_e[y] = (1/\alpha)\ \{\log_e[A] + \log_e[1/S]\} \qquad (3)$$

Plotting $\log_e[y]$ against $\log_e[1/S]$, there is an upward linear relationship, with slope $1/\alpha$. The smaller is α, the steeper the gradient.

The original idea of Pareto was that he had identified in the constancy of alpha "the law of total incomes, and have found it was almost the same for very different countries" (2003 (1896), p.472), but this was soon rejected as untenable. Contemporary commentators noted that Pareto's own estimates of alpha in his Table 3 range from 1.35 (England 1879-80) to 1.73 (Prussia 1881). A half a century later, Clark, who assembled no fewer than 152 estimates of the Pareto coefficient covering 25 countries, stated clearly that "Pareto was mistaken in thinking that there was a high degree of uniformity between the value of his

coefficients in different times and places" (1951, p.538). Indeed, the interest in the Pareto coefficient stems largely from the fact that it varies over time and across countries. It is variation over time in the UK that is the focus here.

A second, and less discussed, reason for examining the Pareto coefficient relates to another of Pareto's manifold interests: as an indicator of the shape of the income elite. In his original article, he argues that the structure of incomes in society is "not that of a pyramid, but rather that of an arrow with a very pointed head and a broad base" (2003 (1896), p.467). As he notes, observe an iron arrow "with a magnifying glass and you will see that it actually has a very complex form" (2003 (1896), p.467). It is therefore ironic that the Pareto distribution itself imposes a particular, pyramidic form: from equation (3), at any point in the distribution a person sees the upper tail stretching ahead with a constant logarithmic slope $1/\alpha$. But, as Pareto's parallel with the arrow indicates, the slope may not be constant. The shape of the elite may take different forms of departure from the Pareto formula. The degree of elite income differentiation may decline as one reaches a higher rank, as illustrated in the left hand diagram in Figure 1, a situation described here as "baronial" in that distinctions among those at the top become progressively less evident. Or the degree of differentiation may be accentuated, as in the right hand income diagram, a situation described as "regal" in that the very top increasingly stands out. (As indicated below the diagrams, real-world arrows may also take diverse shapes; that on the left is probably more useful for indicating direction than for imposing harm.)

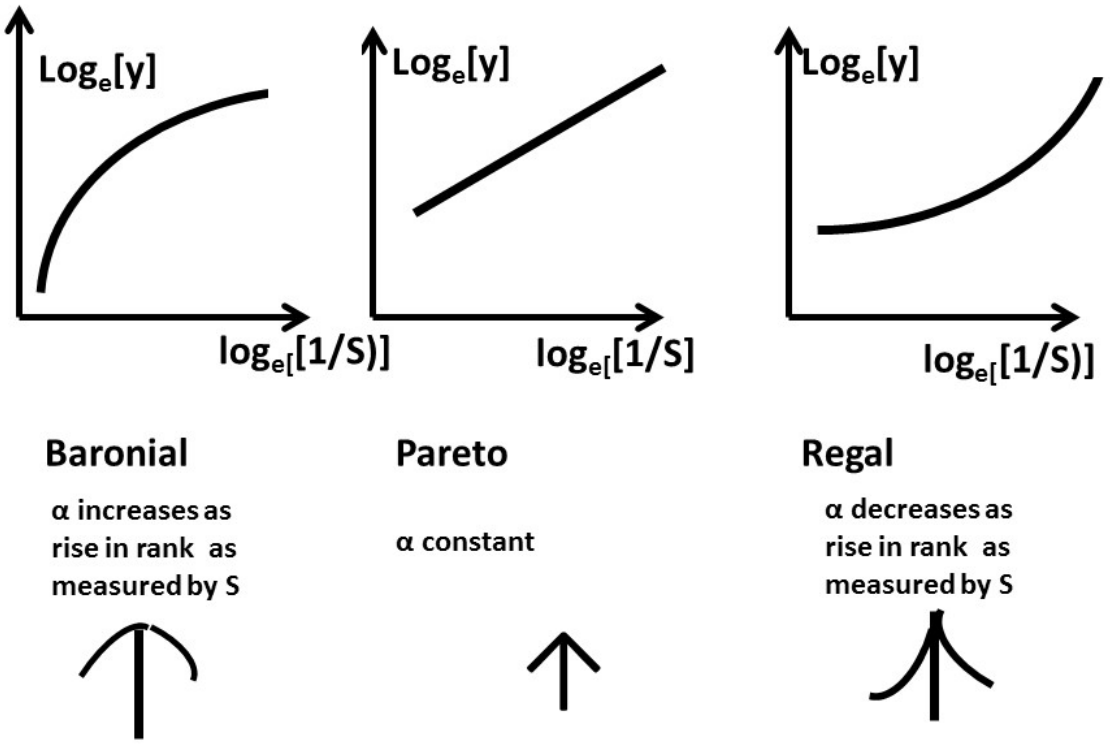
It is in the pyramidic form, however, that Pareto's work is known to economists. Using tabulated income data for 1843 and 1879/1880, Pareto estimated the linear relation (2a), adding an error term and applying the method of ordinary least squares, obtaining an estimate of the coefficient α, and this has become standard practice (see for example Cowell, 1977, Chapter 5). Figure 2 shows the results obtained if we apply the same method as Pareto to the relevant data on the gross (before tax) incomes of the top 5 per cent,[1] obtained since 1949-50 from the Survey of Personal Incomes and earlier from comparable sources (these sources, and their reliability, are discussed later in the paper). The results take us back to the beginning of the income tax in 1799, and the limited evidence for the nineteenth century suggests that incomes at the top were then more concentrated (the Pareto coefficient was lower than today). In the twentieth century, there was less concentration, but also much more volatility. From 1919 to 1979, the estimated Pareto coefficient rose from 1.46 to 2.96, indicating a major reduction in the concentration of incomes at the top. This was subsequently reversed in an equally dramatic fashion: the period after 1979 saw the alpha coefficient fall back to around 1.75.[2] It is in the twentieth century that most of the recorded action takes place.

---

1 Strictly, the estimates are based on the ranges of tabulated data down to the first that contains at least 5 per cent. This definition of the field is discussed further below.

2 In addition to being affected by the recession, the statistics for 2009-10 and subsequent years need to be interpreted in the light of the fact that the 2009-10 returns included a sizeable amount of

Figure 1 Different shapes of the income elite



Note: a rise in α corresponds to a fall in concentration

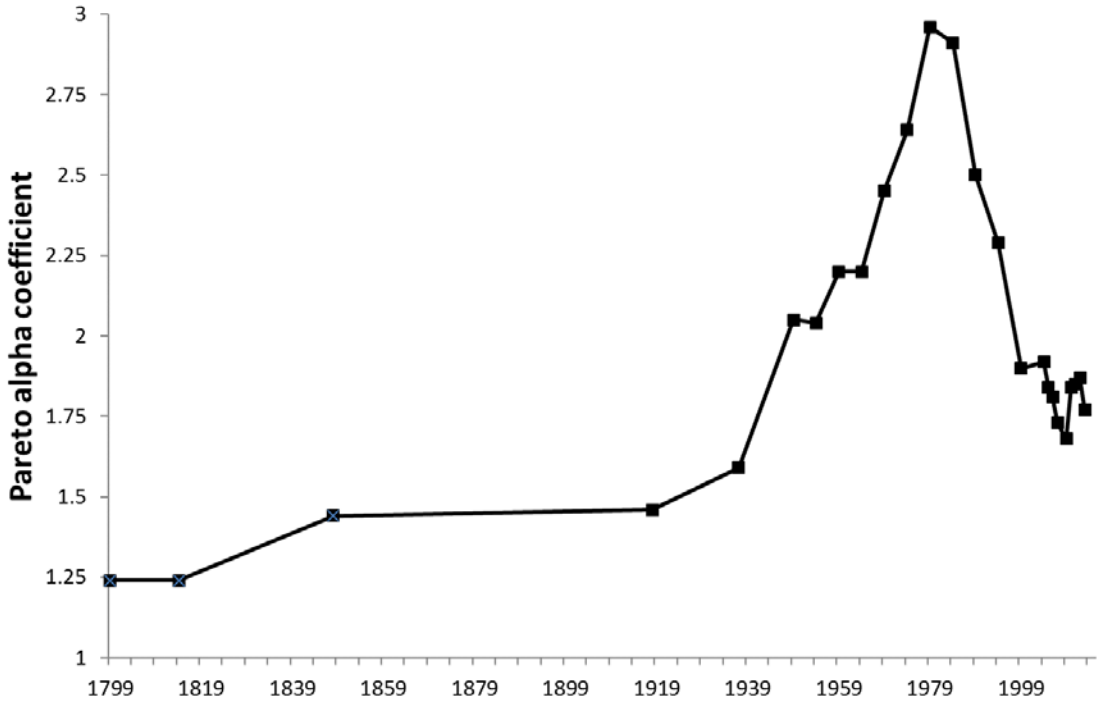
The evidence about top incomes in Figure 2 serves, however, to raise questions as much as to give answers, and two of these questions are the concern of the rest of the paper. The first question concerns the shape of the upper tail and the doubts expressed by Pareto himself, as well as contemporary critics, about the adequacy of the pyramidic form. Can we rely, as in Figure 2, on the Pareto coefficient to summarise the shape of the upper tail? This is the subject of Section I. Is it the case that, as Shirras concluded in 1935, "there is indeed no Pareto law. It is time that it should be entirely discarded in studies on the distribution of income" (1935, p.680)? Was Schumpeter right to say in his obituary of Pareto, that his 'Law' was "path-breaking in the literal sense even though in the end nothing whatever is left of its particular form" (Schumpeter, 1949, p.156)?

The second question concerns the nineteenth century, where few points are shown in Figure 2. In fact neither of the years for which data were used by Pareto (1843 and 1879/1880) appears in the graph. As is explained in Section II, these data are not what Pareto assumed

income brought forward from 2010-11 in order to avoid the 50 per cent top rate of tax introduced with effect from April 2010 (HMRC, 2012). Later years were affected by the reversal of this effect, and by action taken by taxpayers in advance of the reduction in the top rate to 45 per cent from April 2013.

them to be. The reasons for this paucity of observations, and a partial, incomplete, attempt to provide new evidence covering the UK upper tail of earned incomes in this important period are the subject of Section II.

Figure 2: Estimates of the Pareto alpha coefficient for top incomes in the UK from 1799 to the present



Source: 1918/19 to 2013/14 calculated from tabulated SPI data (Appendix 1). The Pareto coefficient is estimated by method (a) over the range of incomes that includes the top 5 per cent of tax units or, since 1990, the top 5 per cent of individuals. The sources of the estimates before 1918 are described in Section II.

I The Changing Shape of the Upper Tail 1918 to the Present

The basic sources used here are tabulated data from the published income tax reports.[3] The income tax data have many evident limitations, reflecting the form of the tax and the efforts of taxpayers to avoid or evade its reach, but, as Pareto wrote in 1896, "income tax furnishes

[3] The paper is based throughout on tabulated data; micro-data on income tax returns are only available for the most recent years: the UK Public Use Tapes provide data for 1985-86 and 1995-96 onwards (except for 2008-09).

us with precious information on the distribution" (2003 (1896), p.451).[4] Since then, the income tax data have provided the basis for many classic studies of income distribution in the UK, such as Bowley (1914), Stamp (1916), and Champernowne (1973 (1936)) and the resurgence of interest in top income shares (Atkinson, 2005 and 2007). The essential statistical ingredient is information on the distribution of total incomes: the number of taxpayers by ranges of gross income and their total gross incomes. Gross income is the total from all sources (earnings, investments and transfers) before deduction of tax. Until 1990, taxpayers are "tax units", defined as single persons or married couples; from 1990, taxpayers are individuals.

In the twentieth century, the collection of this information begins with the special investigation of total incomes carried out by the Inland Revenue for 1918/19, which was repeated for 1919/20 and 1937/38. After the Second World War, it was established as a regular Survey of Personal Incomes (SPI), conducted quinquennially in 1949/50, 1954/55, 1959/60, and from the 1960s becoming annual. These, together with the super-tax (later surtax) returns, have provided the basis for the estimates of the top income shares in the UK contained in the World Wealth and Income Database (WWID): http://www.wid.world/ . The sources of the SPI data are given in Appendix 1. The results presented here cover the SPI years up to 1964/65, then 5 yearly intervals up to 2004/05, and finally all available years after 2004/05 up to 2013/14. The percentages of the total population are calculated using control totals for total tax units up to 1989 and total individuals from 1990 (in both cases children aged under 15 are excluded). In the case of tax units, no account is taken of the differing needs of couples and single persons (incomes are not equivalised). The shares of total income are calculated using control totals derived from the national accounts and other information. In both cases, the sources are described in Atkinson (2007), and subsequent updates in WWID (Alvaredo et al, 2015).

The data relate to the upper part of the distribution, and cannot be used directly to measure overall income inequality. For this, they have to be supplemented by the household survey data that now provide the main source of evidence on income inequality across the population (Department for Work and Pensions, 2015), although the latter also make use of income tax data on the upper tail and the reconciliation of the two sources is an active area of research (Burkhauser et al, 2016, and Jenkins, 2016). Household surveys with national coverage are however a relatively recent innovation. In the UK, the existing series based on survey data start in 1961 (Jenkins, 2015), and it is a signal advantage of the income tax data that they allow a much longer historical story to be told.

What is meant by the "upper part" of the distribution? A major issue in measuring the Pareto coefficient is the choice of threshold above which the distribution is assumed to have the Pareto form. Throughout the paper, as in Figure 2, the estimates are based on

[4] The limitations of the income tax data are discussed in Atkinson (2007) and Atkinson, Piketty and Saez (2011, Section 3.2).

ranges of income that include the top 5 per cent, which in 2013/14 meant having an annual gross income of around £55,000 or more. This definition of the scope is arbitrary, although it has the merit of being readily explained and understood. The evident drawback is that the estimated value of α may depend on the choice of threshold. In the case of the US income data, Burkhauser et al (2012, Figure A1) report that α falls as the threshold rises. A natural alternative would be to estimate the threshold statistically, as in the study of the recent years of the UK income distribution by Jenkins (2016). One consequence of this alternative approach is that, where the shape of the distribution changes over time, the population covered would vary from year to year. The same difficulty does however lie beneath the surface of the approach adopted here, in that, not only the level of α, but also the estimated pattern of changes over time may depend on the selection of 5 per cent rather than another sub-group of the population, such as the top 1 per cent (above £120,000 in 2013/14). In view of this, it is important to consider the sensitivity to the choice of threshold, and for this reason results are also shown where α is estimated using data covering only the top 1 per cent.

The approach adopted here to determining the threshold is pragmatic, and the same applies to the method of estimation of α. Rather than employ a method of estimation with superior statistical properties, I retain the original Pareto approach of estimating relations such as (2a) by ordinary least squares (OLS), which provides estimates that are consistent but biased in small samples.[5] This choice of estimation method is based on two considerations. The first is that OLS lends itself to graphical representation – see Figures 3, 4 and 5 below - which (pace Cirillo, 2013) is instructive for those who believe "in looking at the data". The second is that I wish to concentrate on a different aspect of estimation that has been largely ignored and that provides a new insight into the extent to which the upper tail is Paretian in form.

Three approaches to measurement

Pareto estimated the alpha coefficient from equation (2a), and his approach has been largely followed in the succeeding literature, but there are two other approaches to estimating the Pareto parameter, as may be seen if one lists the three pieces of information that are typically available in tabulated data concerning the cumulative distribution of income:

- The range of income: from y upwards (e.g. above £50,000);
- The proportion of income units with incomes of y or higher, denoted by the survivor function S(y);

[5] A recent review of the estimation problems is provided by Bottazi, Pirino and Tamagni (2015), who emphasize the potential magnitude of the bias. Of the two methods for grouped (binned) data that they consider, that used here performs better, but the results presented here should be viewed in the light of this important qualification. For an earlier review of estimation methods for grouped data, see Aigner and Goldberger (1970).

- The total income received by these units, divided by the total population, denoted by $\Omega(y)$.

It should be noted that these use a control total for population (to express income units as a percentage of the total) but no control total for total income (if the mean income is known, then $\Omega(y)$ divided by the mean is the income share of those in the range from y upwards).[6]

The method employed by Pareto is based on the first two pieces of information, and its application is illustrated in Figure 3 to the top 5 per cent for the UK in 1969-70 (the reason for choosing this year is explained below). The top right hand quadrant of Figure 3 shows the "people curve", mapping $\log_e[S]$ against $\log_e[y]$ where S is measured in 000ths of 1 per cent. The value of α estimated in this way for the UK in 1969-70 is 2.45. However, the method ignores the third piece of information contained in the Inland Revenue tabulations: the amounts of income in each range of the tax data. This point was emphasised by Champernowne (1973 (1936)), who distinguished between the standard approach (a) where loge[S] declines with $\log_e[y]$ with slope α and the curve (b) based on the first and third pieces of information, where the equation estimated relates to the logarithm of total income received by those with y or more

$$\log_e[\Omega] = \log_e[\alpha A/(\alpha-1)] - (\alpha-1)\log_e[y] \qquad (2b)$$

which falls with the logarithm of income with slope $(\alpha-1)$. In Figure 3, the income curve is shown in the bottom right hand quadrant, where total income is measured downwards (normalised so that the total income for the top range is unity). The estimated value of α is very close, at 2.46, to that found by method (a).

In the recent studies of top incomes, a third approach has been adopted, making use (method (c)) of the second and third pieces of information: by eliminating y, the term $\log_e[\Omega]$ is expressed as a function of $\log_e[S]$:

$$\log_e[\Omega] = \log_e[\alpha/(\alpha-1)A^{\alpha}] + \{(\alpha-1)/\alpha\}\log_e[S] \qquad (2c)$$

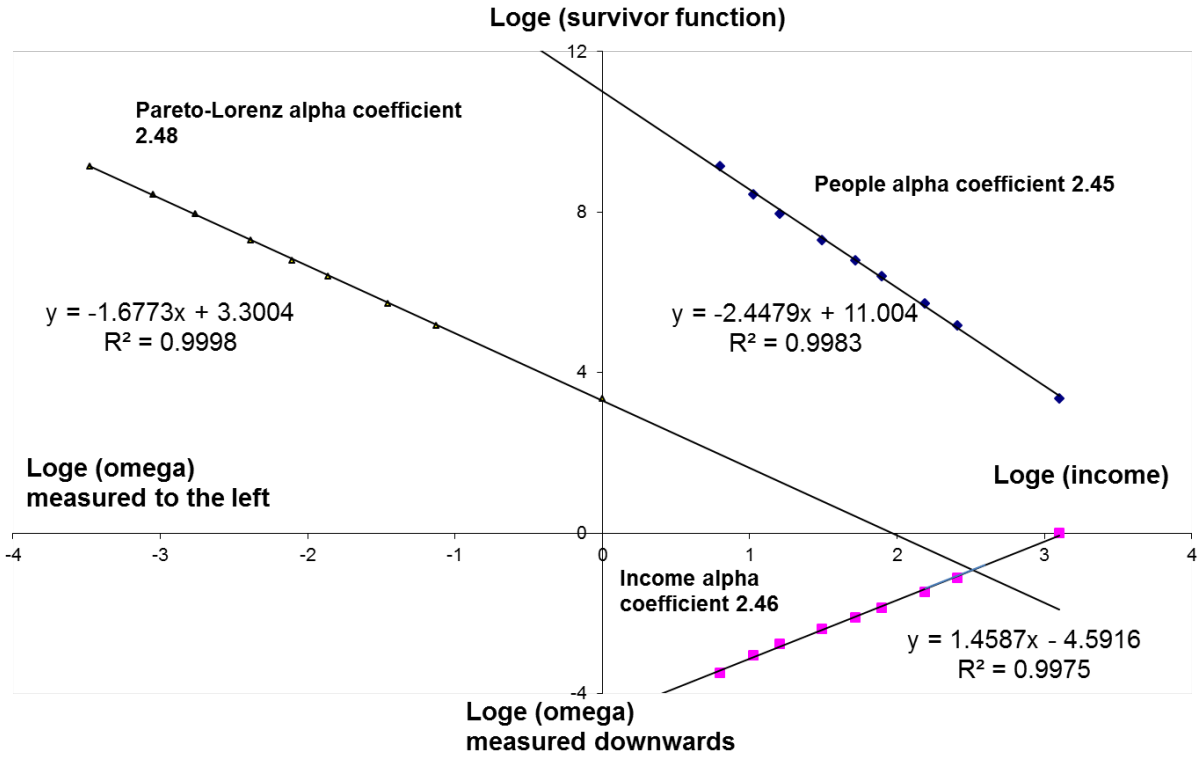
This defines the upper part of the Lorenz curve, and could be named after Macgregor who described such a "bridge between Pareto and the Lorenz ratios" (1936, p.86), or after the French mathematician Fréchet who proposed the approach used here in 1945 (see his equation at the top of p.25). Method (c) again ignores part of the information – the values of the ranges – since we are combining the two curves by eliminating $\log_e[y]$. This third approach is shown in the top left hand quadrant, in inverse form with $\log_e[S]$ plotted against $\log_e[\Omega]$; the slope is therefore equal to $\alpha/(\alpha-1)$.[7] This expression is the beta coefficient ($\beta = \alpha/(\alpha-1)$) preferred by Piketty (2001). From the slope β shown in Figure 3, it may be calculated that the estimated value of α is 2.48. This again is very close to that obtained by method (a), but Fréchet argued that the third approach provides results that were "more

6 Hence is related to the Generalized Lorenz curve, GL(F), by $\mu - \Omega(y(F)) = GL(F)$.

7 It should be noted that no constraint is imposed on the constant term.

regular and better aligned" (1945, p.26). (From Figure 3, it may be seen that the fit as measured by the R^2 is fractionally better with method (c).)

Figure 3: Three different Pareto curves for UK 1969/70 SPI



Source: tabulated SPI data for 1969-70 (Appendix 1). The Pareto coefficients are estimated over the range of incomes that includes the top 5 per cent of tax units.
Reading note: the equation in the top right hand quadrant estimates an error-augmented version of equation (2a) in the text by ordinary least squares; the y variable is $\log_e[S]$ and the x variable is $\log_e[y]$; the coefficient of x is the estimate of the Pareto alpha; the bottom right hand quadrant refers to equation (2b) and the top left hand quadrant to a re-arrangement of equation (2c) such that $\log_e[S]$ appears on the left hand side (a regression with $\log_e[\Omega]$ on the left hand side yields an estimate of the alpha coefficient that is identical to 3 decimal places).

The results for 1969-70 are reassuringly coherent, and readers may wonder why I have emphasised the three different approaches. Should we not be focusing on the sensitivity to choice of threshold? In fact, using the data for 1969/70, the results for α when restricting attention to the top 1 per cent are close to those described above: 2.46 (in place of 2.45) using method (a), 2.49 (in place of 2.46) when using method (b) and 2.54 (in place of 2.48) using method (c). The fact that both sets of conclusion are robust – regarding method and threshold – provides grounds for believing that the Pareto distribution provides a close fit to the upper tail in this year.

Differences in the estimates

If we look at other years, however, the results are less reassuring. To begin with, it is not always the case that the three methods yield estimates that are in such close agreement. The first twentieth century income tax data covering the whole range of incomes relate to 1918/19. The results of the three methods for this year are shown in Figure 4. We now have three estimates of α that are distinctly different. The method (a) estimate is 1.46, whereas the results for the other two methods are 1.58 and 1.67, respectively. The salience of these differences may be seen from the fact that the move from method (a) to method (b) would take the 1918 position in Figure 2 from 1918/19 to that for 1937/38, and that the result from method (c) would take the value to that for 2009/10. (Lest it be thought that the findings for 1918/19 were unduly influenced by the ending of the First World War, it should be noted that the results for 1919/20 were very similar: 1.46, 1.57 and 1.66.)

The differences would not arise if the Pareto distribution provided a fully satisfactory representation of the data: if there were no evident departures from linearity. From Figure 4, it may be seen that, in all three quadrants, there is a distinct curvature. In the top right hand quadrant, the relation between $\log_e[S]$ and $\log_e[y]$is such that, in the middle of the range, the level of income associated with a particular value of y is greater than that predicted by the Pareto line, and in the upper range the level of income is less than predicted.[8] [9] Expressed in terms of the gradient between y and rank (measured by 1/S), the curve in 1918/19 appears to turn down as indicated by the left hand "baronial" version of Figure 1.

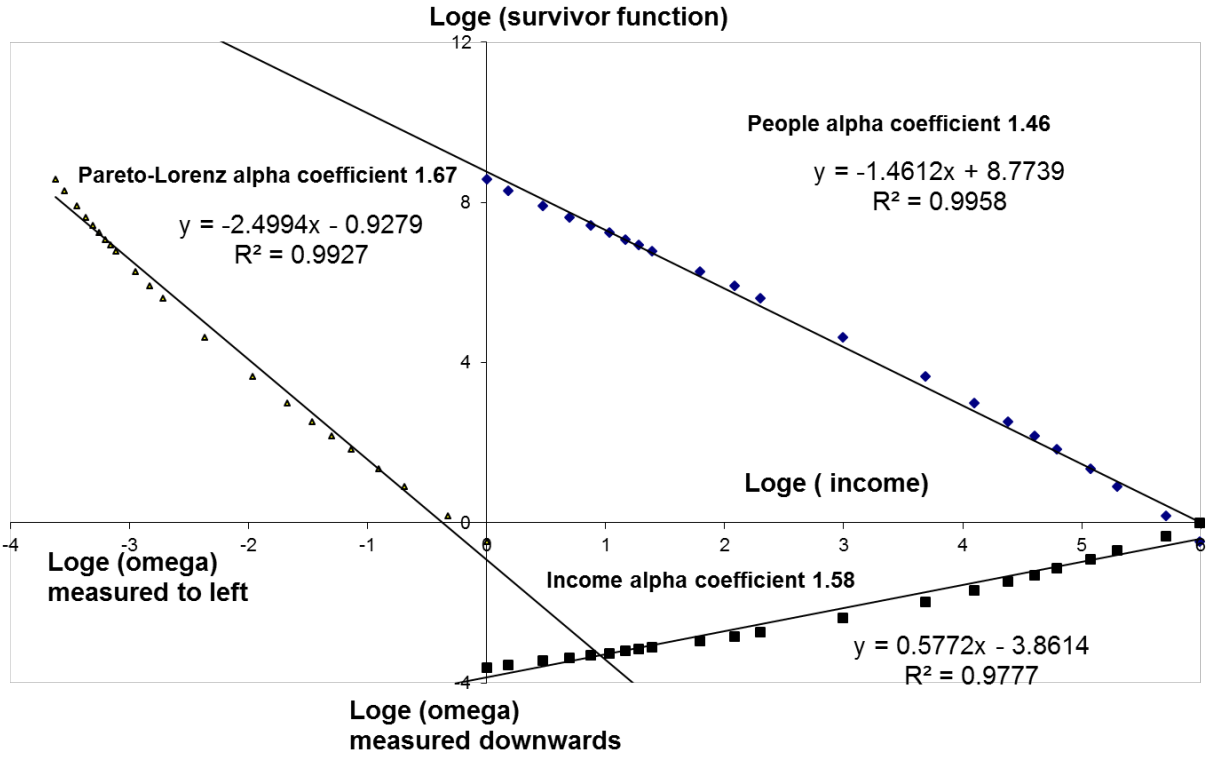
The differences between the results from the methods provide therefore a simple diagnostic device. The results for methods (a) and (c) covering the period from 1918/19 to 2012/13 are shown in Figure 5, and reveal an interesting pattern of change over the twentieth century. From Figure 5, it may be seen that the alpha coefficients cross-over around the 1970s, with the method (c) being initially higher and later lower. It was for this reason that the data for 1969-70 were used in the first example. The 1970s appear to have been a watershed. Figure 5 also shows the sensitivity of the results to the threshold: the dashed line indicates the effect, with method (a), of limiting the estimation to the top 1 per cent (in place of the top 5 per cent). In the early part of the period, the dashed line lies in general between those for methods (a) and (c), but all three series are close in the later years.

---

[8] Departure from the Pareto line in this form was noted by Shirras (1935, p.670) in his study of Indian income tax data for the period 1913-14 to 1929-30.

[9] Downward concavity was observed in the case of word frequency by Zipf (1949, p.82) who associated it with "informal colloquial speech" (Powers, 1998, p.153).

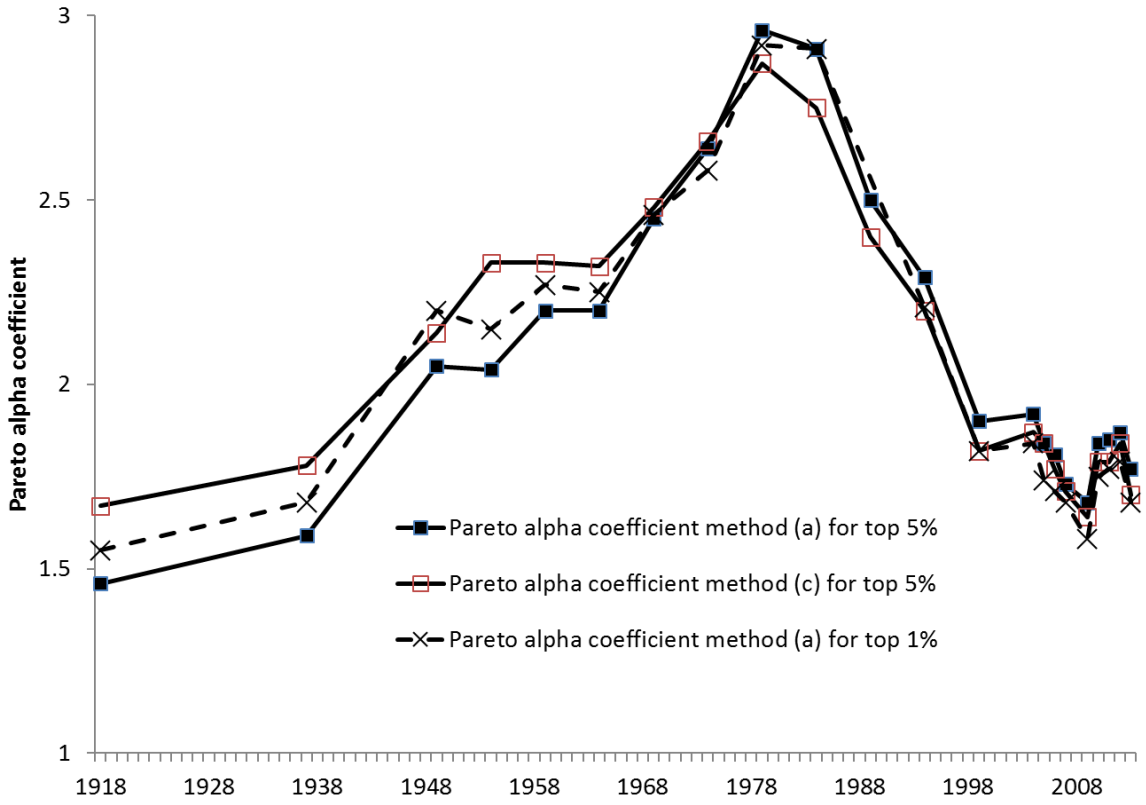
Figure 4: Three different Pareto curves for UK 1918/19



Source: IR tabulations for 1918-19 (Appendix 1). The Pareto coefficients are estimated over the range of incomes that includes the top 5 per cent of tax units. See the note to Figure 3.

In one sense, the conclusion to be drawn from Figure 5 is that the broad story is robust to the method and threshold chosen. At the same time, the scale of the graph may be misleading with respect to the economic significance of the differences in α. If, for example, we compare 2013/14 with 1937/38, then the α estimated from the top 5 per cent shows that the top tail today is still less concentrated than before the Second World War (1.72 in 2013/14, compared with 1.59 in 1937/38), but, if the threshold is set at 1 per cent, then the estimated values of α are identical (1.68). One yardstick of the economic significance of differences in α is provided by the implied value of the Gini coefficient if the entire distribution were Pareto in form, which is $1/(2\alpha-1)$. The difference between methods (a) and (c) in 1918/19 is then that between an implied Gini coefficient of 42.7 and one of 52.1 per cent. The effect of limiting the estimation to the top 1 per cent is then, with method (a), to reduce the implied Gini coefficient from 52.1 per cent to 47.6 per cent. These magnitudes suggest that the issues cannot be ignored.

Figure 5: Pareto alpha coefficient: Different methods and sensitivity to threshold UK 1918/19 to 2013/14



Source: tabulated SPI data and earlier Inland Revenue tabulations (Appendix 1).

Baronial or regal?

The terms "baronial" and "regal" were employed earlier to distinguish two directions of departure from the Pareto straight line that links the logarithm of income to the logarithm of rank measured by 1/S. They are designed to recall the difference between the situation where a monarch was surrounded by powerful barons whose resources were not dissimilar in scale and a situation where the monarch had, for example by appropriating the income of the church or seizing mineral wealth, raced ahead. Closer to home, there is the pay situation in universities. In the 1960s, university heads were paid not dissimilar amounts to professors, and there was little differentiation within the professoriate. In more recent years, the structure has changed, with salaries rising rapidly at the top: the Vice-Chancellor (head) of one major UK university receives some 6 times the basic professorial pay.

The changing shape of the upper tail of gross incomes in the UK is illustrated by Figure 6, which shows the income corresponding to different points in the distribution, where income for each year is expressed relative to the mean for all tax units (individuals) in that year

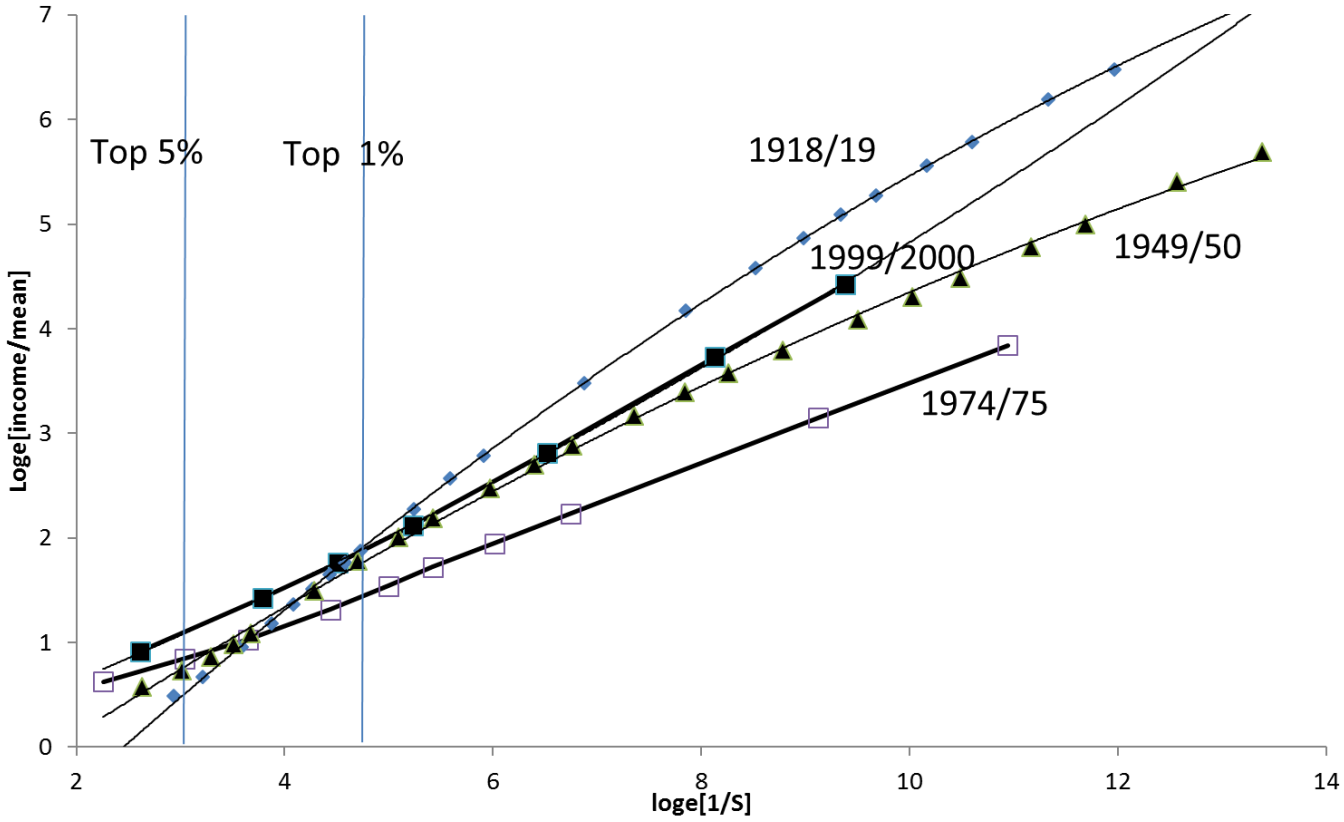
(based on the control totals covering the whole population). In each year, a quadratic trend is fitted by ordinary least squares. In 1918/19, there is a distinct departure from the Pareto linearity in the direction of concavity, as is illustrated by the fitted quadratic. Over the succeeding quarter century, the curve rotated clock-wise, so that within the top 1 per cent there is a lower level of income (relative to the mean) at any rank. At the same time, the degree of concavity fell, as is illustrated by 1949/50. The result of these changes is that by 1974/75 the curve had lost its concave shape and was close to linear. After the 1970s, the curve rose and rotated in the opposite direction, becoming mildly convex, as illustrated in Figure 6 by 1999/2000. Those at the top are leaving the rest behind, to an extent that becomes greater the further we penetrate the top 1 per cent. A further indicator of the changing shape is provided by the variation in the estimated α as the threshold is varied. In their study of the United States (US), Burkhauser et al found that over the period 1967 to 2006 the estimated value of α was lower using a threshold of 1 per cent in place of 5 per cent (2012, Figure A.1, which shows $\beta = \alpha/(\alpha-1)$), indicating greater concentration at the very top. As may be seen from Figure 5, the UK data examined here show the reverse pattern up to the 1970s, with the estimated α being lower when the threshold is set at 1 per cent, but after the 1970s the pattern resembles that in the US. This supports the view that the UK has moved from being baronial to being regal in the shape of its upper tail.

Richer functional forms

The change in shape of the upper tail suggests that we should explore a wider range of functional forms for the income distribution. A rich variety has indeed been proposed for the income distribution as a whole, such as those belonging to the five parameter Generalised Beta Distribution (McDonald and Xu, 1995). Pareto himself examined what has come to be known as the Pareto Type II distribution, where $S = A\,(y+b)^{-\alpha}$, with $b > 0$, which he describes as "probably the general form of the distribution curve" (Pareto, 2003 (1896), p.238). In terms of Figure 6, the introduction of a positive b generates a curve that approaches linearity from below (and hence does not allow for the very top racing away).[10] The improvement in fit from adopting the Pareto Type II has been investigated by Jenkins (2016), who concludes that, using UK data from 1996/97 to 2007/08, the evidence favours the extended model but that the advantage depends on the threshold, with the Type II fitting better at lower thresholds.

[10] Inverting the expression for S, $y = (A/S)^{1/\alpha} - b$, so that $\log_e[y] = (1/\alpha)\log_e[(A/S)]\,.\,g[S]$, where $g[S] = \log_e[(A/S)^{1/\alpha} - b]/\log_e[(A/S)^{1/\alpha}]$. The function g is less than 1 but converges to 1 as (A/S) tends to infinity.

Figure 6: Income (relative to mean) in relation to rank measured by (1/S) UK 1918/19, 1949/50, 1974/75 and 1999/2000



Source: tabulated SPI Data (Appendix 1), with quadratic fitted by OLS.
Note: the top 1 per cent are those to the right of the top percentile shown by the right hand vertical line; the top 5 per cent are those to the right of the 95th percentile.

In discussing the upper tail in terms of Figure 6, we are approaching the problem via the inverse distribution function, regarding y as a function of S (=1-F), rather than the more usual practice of treating (1-F) as a function of y. As has been noted by Cowell, the inverse distribution, popularised as Pen's parade of incomes, has been "only rarely used" (1977, p.169). This approach via the inverse does however warrant more attention (see, for example, the powerful case made by Jasso, 1983). It provides a potential source of new functional forms that could capture the changing shape of the upper tail. In particular, many functional forms that have been proposed have the property of tending to a Pareto upper tail (as with the Pareto Type II). The racing away at the top that we have observed in the UK income distribution suggests, however, that we may want to allow for a slower rate of decay: a "super-heavy" tail. As is noted by Falk, Hüsler and Reiss, "the designation of

super-heavy concerns right tails decreasing to zero at a slower rate, as logarithmic, for instance" (2011, p.76). This, as they say, takes us "out of the 'power-law-world'".[11]

Conclusions

In this section, we have looked behind the picture of dramatic change over the twentieth century in the upper tail of the UK income distribution with which the paper began and obtained a more subtle view of the evolution over time. It is not just that the degree of concentration fell considerably sharply and then reverted by rising sharply after 1979. The distribution has changed shape. Using the different methods of estimating the Pareto alpha as a diagnostic device, we have seen that the distribution in the first part of the century (from 1918/19) departed from the Pareto pyramidic shape by being flatter, but that there was a shift over time, with a turning point in the 1970s, such that in recent years the (logarithmic) gradient increases with income. This means that those at the very top have raced away even faster. In terms of Pareto's interest in the shape of elites, the UK was transformed from being baronial to being regal. To capture this, we need to move on beyond assuming a Pareto form for the upper tail. The Pareto alpha is, at best, a convenient first summary of the extent of income concentration.

II The Mis-Understood Nineteenth Century

It may be unfair to question Pareto's ability to explain the twentieth century. It was nineteenth century data that he was studying, and it is to this century that I now turn. A natural motivation is the comparison with the twentieth century, but the nineteenth century is of independent interest as the locus for the application of the Kuznets curve to the British industrial revolution. In 1955, Kuznets described how income inequality could be expected to first increase and then fall as an economy industrialised. He cautiously suggested that "I would place the early phase in which income inequality might have been widening, from about 1780 to 1850 in England … I would put the phase of narrowing income inequality … in the last quarter of the 19th century" (1955, p.19). In his classic detailed study of the UK, Williamson adopted a similar periodization, with "inequality rising sharply up to somewhere in the middle of the nineteenth century and falling modestly thereafter" (1985, p.3). His conclusion is that:

> "British capitalism did breed inequality. … The French Wars interrupted the process, but the rise in inequality picked up following Waterloo [1815]. British inequality seems to have reached a peak somewhere around the 1860s or shortly thereafter. While not spectacular, the egalitarian leveling up to World War I was universal: the income shares at the top fell" (1985, p.200).

Can the same kind of analysis be reproduced for the nineteenth century? Can we fill the evident blanks shown in Figure 2? After all, the modern income tax in the UK was first

[11] For example, the log-Pareto model could be fitted (see Cormann and Reiss, 2009).

levied from 1799 to 1802 by the government of William Pitt the Younger as a means of financing the Napoleonic Wars; it was temporarily abolished during the Peace of Amiens; then re-introduced by Pitt's successor Addington in 1803 in a different form, with income being assessed under different "schedules" A to E, and with collection at source. Abolished again in 1816, the income tax remained in abeyance until 1842, when it was re-introduced by Peel and since then it has been in continuous operation.

Pitt's income tax

The first of these taxes – Pitt's income tax – was the subject of statistical investigation and the Inland Revenue published a detailed tabulation for Great Britain[12] of income taxpayers by ranges of income assessed in the year ending April 1801 and referring to incomes accruing in the year 1799/1800 ending April 1800 (reproduced in Stamp, 1916, Appendix IV). The figures are described here according to the year of accrual 1799/1800, and referred to as 1799. These statistics have to be regarded with considerable caution, since there is likely to have been a considerable shortfall in declared incomes in the early years of the operation of the tax. Deane and Cole draw attention to the increase in gross income assessed between 1801 and 1803, which they attribute "largely to the more effective coverage of the 1803 Act with its collection-at-the-source procedure" (1964, p.325).[13] The Inland Revenue in its history of the income tax stated that the introduction of taxation at source in 1803 "had a great effect on the productiveness of the Tax, the produce at Five per cent, having been almost equal to that in the year 1799 when the rate was Ten per cent" (Inland Revenue 43rd Annual Report for the year ended 31st March 1900, p.110). Top incomes are likely therefore to be more seriously under-stated in the 1799 data than in the twentieth century tabulations.

The 1799 distribution is, nonetheless, worth examination. Figure 7 shows the three versions of the Pareto diagram, estimated for broadly the top 5 per cent of tax units.[14] In each case, the estimated alpha coefficient is less in 1799 than that obtained using the corresponding method in 1918 (see Table 1).

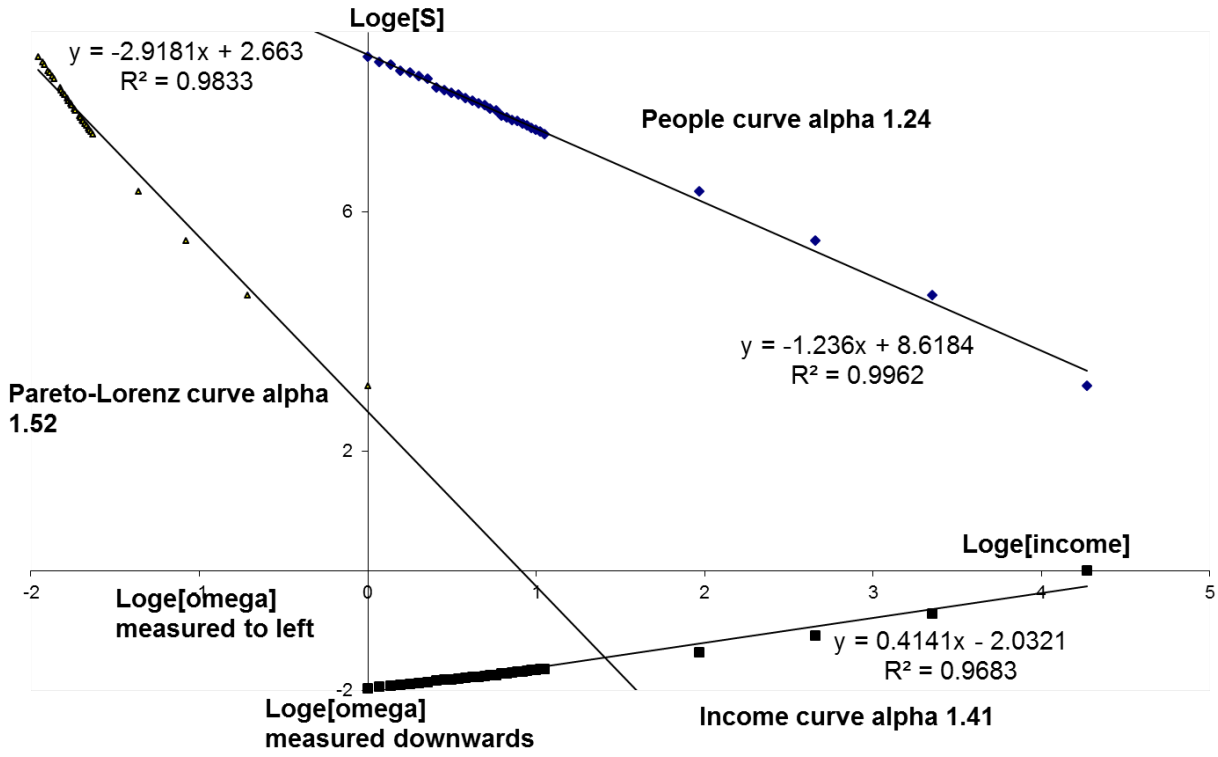
---

12 The figures therefore exclude Ireland.

13 The problems in relying on declarations of income are illustrated by the exchange between John Horne Tooke and the Clerk to the Income Tax Commissioners in 1799. The Clerk had written to say that the Commissioners had "reason to apprehend your income exceeds sixty pounds a year", to which Mr Tooke replied that "I have much more reason than the Commissioners can have to be dissatisfied with the smallness of my income" (quoted in Sabine, 1966, p.30).

14 Total tax units (total aged 15 and over minus married women) for Great Britain in 1801 have been estimated using the demographic information provided by Mitchell (1988), cited below as M. The total population is from M, p.9; the proportion aged 15 and over is based on the proportion in 1821 Census (M, p.15); the proportion of those aged 15 and over who were married women is based on the proportions married in 1851 (M, pp.20 and 24) and the number of women aged 15 and over (M, pp.16 and 17).

Figure 7: Three different Pareto curves for UK 1799/1800



Source: Inland Revenue tabulations, reproduced in Stamp (1916, Appendix IV).

Table 1 Estimated Pareto coefficients 1799 and 1918

	Method (a)	Method (b)	Method (c)
Estimated on top 5 per cent			
1799	1.24	1.41	1.50
1918	1.46	1.58	1.67
Estimated on top 1 per cent			
1799	1.30	1.50	1.65
1918	1.55	1.66	1.70

The alpha coefficients are lower in 1799 whether they are estimated on the top 5 per cent, as in Figure 7, or the top 1 per cent (not shown). On this basis, the degree of concentration in the upper tail was greater in 1799 than in 1918, and this conclusion would be reinforced if a greater degree of under-declaration in 1799 caused the alpha to be over-stated. At the same time, the Pareto fit in Figure 7 is not good, as is suggested by the large differences

between the values of alpha with the different methods. The Pareto line (plotting $\log_e[S]$ against $\log_e[y]$ has distinct downward curvature beyond an income of £200 a year.

After Pitt

Unfortunately, the changes made to the structure of the income tax – the adoption of a schedular system in 1803 – means that no further tabulations of taxpayers according to total income were available in the nineteenth century. This is a matter about which there have been frequent misunderstandings by scholars – including by Pareto himself – but the absence of the relevant tabulations means that only indirect, and incomplete, evidence can be brought to bear on the nineteenth century development of the upper tail.

The fact that, from 1803 onwards, the UK income tax was levied on a schedular basis had the consequence that the resulting administrative data could not be used to construct estimates of the distribution of income. It was indeed the express purpose of adopting a schedular system that the total income of a taxpayer should not be calculated. Income was assessed under different schedules: Schedule A on profits from the ownership of land, houses, etc., Schedule B on profits from the occupation of land, Schedule C on the income from British and other government securities, Schedule D on the profits from businesses, concerns, professions and employments, and Schedule E on the salaries of Government, Corporation and Public Company officials. So a taxpayer could be assessed under all these schedules. Even within a schedule a taxpayer could be assessed several times. Moreover, an assessment could cover more than one tax unit. The first Annual Report of the Inland Revenue Commissioners was quite explicit: "the system leaves unrevealed, to all those connected with the assessment of the Tax, the total Income of any Person, except those who claim entire exemption from it, or who seek to bring themselves under a lower rate of duty" (p.31).

Many students of income distribution have fallen foul of this administrative feature of the UK income tax. As noted at the beginning of the paper, Pareto employed data for England for 1843 and 1879/80. However, if we go back to the source (Giffen, 1904, pages 412 and 413), we see that the data do not pertain to individual total incomes. The data cover assessments under part of Schedule D of the income tax of the income from trades and professions. The data exclude public companies but, as explained by Giffen, partnerships make only one return. As a result, "there is no reason to believe that the number of separate assessments corresponds in any way to the number of individual incomes" (1904, p.412). Moreover, any individual taxpayer may appear several times in the statistics. The official Inland Revenue tables on Schedule D and E assessments carried a warning in bold that "The amounts do not represent 'Total Incomes from all sources`" (Annual Report for the Year ended 31 March 1915, Table 128). The Inland Revenue gave a hypothetical example of a person with total income of £5,000 a year who would have appeared six times under Schedule D (although only twice as a person) and once under Schedule E, whereas "the income of £5,000 as a whole would not appear in the tables at all" (*Annual Report for the Year ended 31 March 1915*, p.121).

Giffen, who tabulated the Schedule D figures used by Pareto, gave as a justification that "in comparing distant periods, it seems not unfair to assume that the increase or decrease of assessments would correspond to the increase or decrease of individual incomes" (1904, p.412). But this seems to be like whistling in the dark to keep up one's spirits. There is no reason to suppose that the difference between assessments and individual incomes is a fixed effect. A much more substantial argument is made by Williamson (1979) who makes use of individual returns for Edinburgh for 1800-01 and 1803-04. He concludes that "that the inequality trends in taxable Schedule D income … are good proxies for inequality trends in total taxable income, although the former exaggerates movements in the latter" (1979, p.37). The reassuring conclusion does however depend on a number of assumptions, including the absence of drift in the covariance of different types of income, and it is not evident that the underlying model allows adequately for people who appear under several assessments (it adds incomes but not assessments) nor for combined assessments as with partnerships (see Feinstein, 1988, p.718).

The problems with the Schedule D figures led contemporary writers to seek alternatives. In particular, there were efforts by those familiar with the tax statistics to combine them with other evidence to arrive at "mixed estimates", which have been used by Williamson (1985), including the work of Sayer (1833)[15] and Porter (1851). On the basis of Sayer's original data for income recipients and amounts of income by ranges for 1813/14 (Sayer, Appendix p.45), combined with a control total for tax units as described above, the fitted Pareto coefficient using method (a) is identical to that for 1799 at 1.24 - see the "mixed estimates" shown in Figure 8. If, as noted above, there was significant under-statement of income in the earlier year, then this would be consistent with some decline in concentration. Such a rise in α would be in the reverse direction from the fall in α (rise in concentration) given by Williamson (1985, Table 4.4) based on the same source.

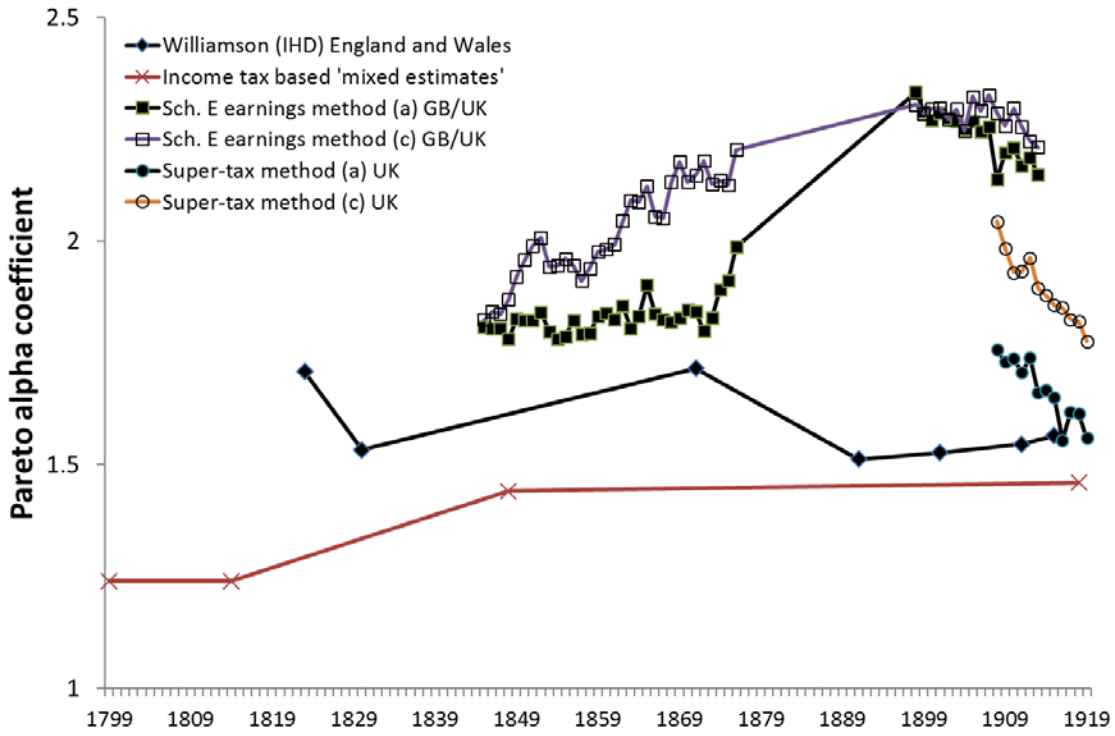
The Williamson estimate of the Pareto coefficient as 1.121 in 1813/14 would however have been extremely low. Of the 152 coefficients assembled by Clark (1951, pages 533-537), only two are below 1.2: 1.13 estimated by Pareto for the city of Augsburg in 1526 and 1.13 for one year in the series estimated by Shirras (1935) for India. My belief that 1.125 is too low for Great Britain in 1814 is re-inforced by the fact that method (b), based on Sayer's probably more reliable income totals (the numbers are derived using assumed mean incomes in each interval), yields an estimate of 1.45, close to the 1.41 obtained using the method (b) for 1799-1800. A second "mixed estimate" by Porter (1851, p.197)[16] for the UK in 1848 of the distribution of incomes by numbers in different ranges above £150 a

15 Sayer was arguing for the re-introduction of the income tax during the period of its abeyance (1816 to 1842). On the title page of his book, he described the income tax as "the most equitable, the least injurious, and (under the modified procedure suggested therein) the least obnoxious mode of taxation", and – with resonance today – "the most fair, advantageous, and effectual plans of reducing the national debt".

16 It should be noted that the reference is to Porter's journal article, not to his book (Porter, 1851a).

year leads to an estimate by method (a) of the Pareto coefficient of 1.441, as given by Williamson (1985, Table 4.3). Such an increase compared with 1799/1800 indicates a reduction in concentration at the top in the first half of the nineteenth century, and this is what has been shown in Figures 2 and 8. On the other hand, doubts about the quality of the data at both ends of the comparison suggest caution in drawing any firm conclusion. It should also be emphasised that we are concerned here with the upper tail. The degree of concentration at the top may have moved differently over time from the overall degree of income inequality, which was the main focus of Williamson (1985). What happened to top incomes may not throw light on the "standards of living debate" as to real wages during the Industrial Revolution.

Figure 8: Pareto alpha coefficients 1799 to 1919



Sources: (a) Williamson (IHD) from Williamson, 1985, Table 4.2; (b) Income tax based "mixed estimates" – see text; (c) Schedule E earnings calculated from tabulated data (Appendix 2); (d) calculated from super-tax data (Atkinson, 2007, Table 4A.1).

Indirect sources

The long gap between 1799/1800 and 1918/19 is an irresistible challenge, and a number of indirect sources have been tapped in order to provide a picture of the evolution of income inequality in the UK over the nineteenth century. In reaching the conclusion cited earlier – that income concentration increased over the first part of the nineteenth century –

Williamson refers to the social tables of Gregory King and followers (revised by Lindert and Williamson, 1983), and makes new estimates based on the statistics on Inhabited House Duty (IHD).[17] The resulting IHD estimates of the Pareto coefficient are shown here in Figure 8. These have been described by Feinstein as "one of the most valuable contributions" of the book (1988, p.714), but Feinstein went on to argue that there are major shortcomings in the application of the IHD data. The criticisms of Feinstein are well summarised by Brandolini: "the partial utilisation of original sources, the incorrect deflation of rental values, and the improper treatment of the series as being homogeneous over time. Once that these errors are amended 'the peak is appreciably flattened and the valleys raised'" (2002, p.9). This led Feinstein to conclude that "the nineteenth century exhibited no marked fluctuations in inequality. Instead, the general picture is one of broad stability" (1988, p.728). In this context, we may note that the modest inverse-U shape in Figure 8 with the IHD data is the *reverse* of that predicted by the Kuznets curve. A rise in the Pareto alpha means less, not more, concentration of top incomes.

A partial and imperfect picture: The Schedules D and E distributions of earnings

Since the aim here is not to be totally negative, I now explore another indirect and, admittedly, partial and imperfect source of evidence about the changes in top incomes over the nineteenth century: the distribution of earned incomes by employees taxed under Schedule D (reported for years since 1898/99) and Schedule E (reported from 1845/46). These are a partial source, since they relate only to earned incomes. They are an imperfect source in that there remains the problem of multiple employments. Stamp gives the example of "a country solicitor, who is clerk to magistrates, clerk to rural district councils, clerk to income tax commissioners, to guardians, and to various institutional bodies and charities, may have twelve or fifteen separate assessments" (1916, pages 268-269). There is no way in which these can be aggregated in the statistics.

There is the further problem that earnings are reported in two different ways during this period: Schedule E (Sch.E) covered the salaries of those in the service of the Government, of Public Bodies, and of public companies, whereas Schedule D (Sch. D) covered those employed by private firms and private persons. As explained by Stamp:

> "the distinction between assessment under Sch. D and Sch. E rests not so much in the character of the duties performed as in the constitutional character of the employer. For example, a clerk performing exactly the same duties at exactly the same salary may one year be under Sch. D and the next under Sch. E merely because the employing firm has become registered as a limited company" (1916, p.264).

[17] Inhabited House Duty was a tax imposed on the annual value of houses wholly or partly occupied as dwellings, first imposed in 1696, and applied for much of the period (it was repealed in 1834 but re-introduced in 1851). It was finally repealed by the Finance Act 1924.

One consequence is that there was a constant shift from Sch. D to Sch. E: "the conversion of private concerns into public companies is a factor constantly tending to increase the assessments [under Sch. E] and to diminish the assessments on employees under Sch. D" (56th AR, p.117). Stamp comments that "the amount of this drain is important, but there is no way of determining it exactly" (1916, p.214). The number of Sch. E assessments certainly increased markedly over the period covered by the tabulations: in the first year (1845/46) there were 49,437 (for Great Britain). With the lowering of the threshold to £100 a year (from £150) in 1853/54, and the extension of coverage to the UK as a whole (adding Ireland), the number under Sch. E increased to 73,715; by 1898/99 it had reached 296,962, which was some 2 per cent of total employees. (The sources of the control totals for total employees and total earnings are given in Appendix 2.)

The existence of the two schedules would not be a matter for concern if they could be combined; this is however only possible from 1898/99 (when the separate Sch. D tabulations were first published). If we compare the two distributions (Sch. E and Sch. D and E combined), we find that the estimated Pareto coefficient (method (a)) is 2.33 in the former case and 2.37 in the latter case. These are reassuringly close, but Sch. E accounted for some two-thirds of the total observations, and the results might be different in earlier years when Sch. D was proportionately larger.

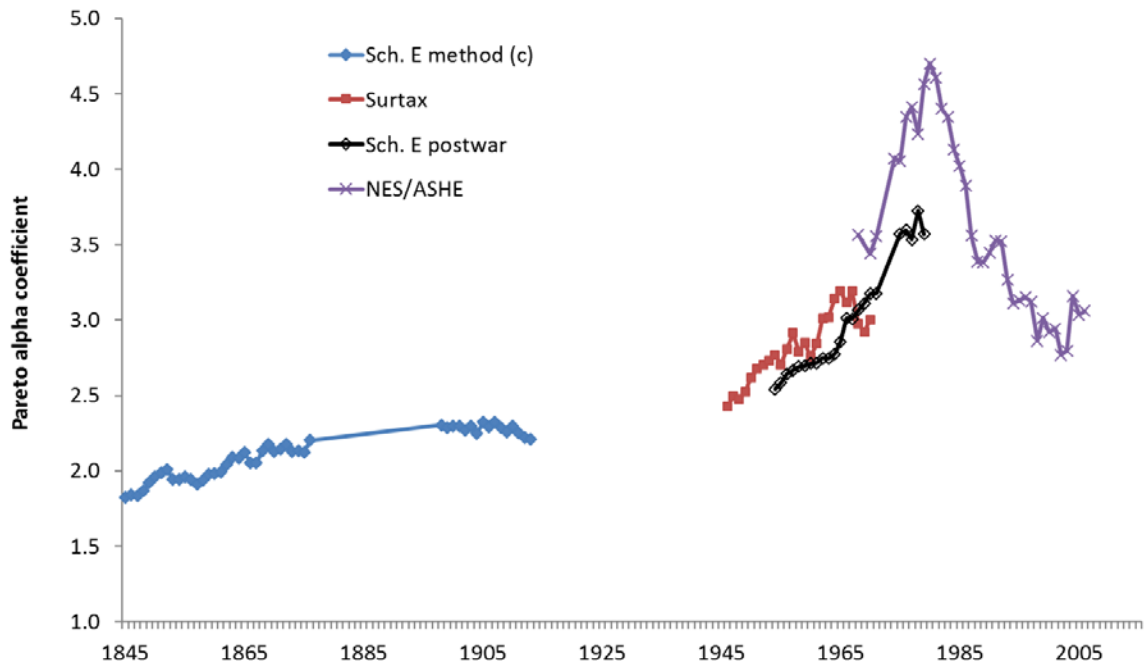
The results shown in Figure 8 for the Pareto coefficient of the upper tail of the earnings distribution for the period 1845 to 1913 should be viewed in the light of the above qualifications. The alpha coefficients are calculated on two bases: method (a) and method (c). The results are for Sch. E throughout. There is a gap between 1877-78 and 1897-98 when the statistics were not published. Nonetheless, there are data for a total of 48 years, and they tell an interesting story. They again appear to support the reverse of the Kuznets curve: in the early part of the period shown, from 1845 to 1876, the degree of concentration at the top is falling, as the coefficient rises; in the later part of the period, 1898 to 1913, concentration is rising, as the coefficient falls. The finding of a reverse-Kuznets curve should not be over-stated. The graph shows clearly that, while the two methods (a) and (c) give similar estimates for 1845, method (a) exhibits a much less marked subsequent increase and by 1876 the difference from method (c) is a distinctly salient 0.38. In the second part of the series, the two methods give results that move more closely together.

Coupling the two centuries

The paper has adopted a long-term perspective, but such a perspective also turns the spotlight on particular episodes of distributional change that may otherwise fall between the cracks. One such episode is revealed by Figure 8: the period from 1898 (following Queen Victoria's Diamond Jubilee) to 1914 (outbreak of the First World War). This is a period of considerable intrinsic interest. The economy was beginning to recover from the Great Depression of British Agriculture; and the landed wealthy were increasingly being displaced by those whose money came from industry and trade. Moreover, in contrast to much of the preceding century, there is annual evidence about the top of the earnings

distribution, as already discussed, and about the top of the overall distribution of income. The introduction of super-tax in 1909 meant that information became available about the total incomes of those liable to the new graduated income tax (Bowley, 1914 was quick to make use of these to estimate the Pareto coefficient). Figure 8 shows the full run of super-tax estimates from income year 1908 to income year 1919 on both methods (a) and (c).[18] Both sources indicate that the pre-First World War period was characterized by a falling Pareto alpha and hence greater concentration. The economic position of the wealthy was under attack from the new Estate Duty, introduced in 1894, and from the super-tax of 1909, but the period was one in which economic privilege was being reinforced, rather than the reverse.

Figure 9: Pareto alpha coefficients for EARNINGS 1845 to 2006



Sources: (a) Schedule E earnings calculated from tabulated data (Appendix 2); (b) other series from Atkinson and Voitchovsky (2010, Tables A1-A4).

If the pre-First World War period was the "Indian summer" for those at the top, the rest of the twentieth century brought a very different story. Figure 9 brings together the Sch. E estimates for the period 1845 to 1913 with more recent estimates of the Pareto coefficient

18 The sources and methods are described in Atkinson (2007).

for the distribution of individual earnings, bearing in mind that the coverage is now much more complete. Unfortunately, the Sch. E series ceased to appear after the First World War and the first income tax tabulations of earnings are those from surtax dating from 1946. The Sch. E series itself re-appeared in 1954. Estimates of the Pareto coefficient, based on the share of the top 0.5 per cent within the top 5 per cent (see Atkinson and Voitchovsky, 2010, p.439). Both of these are shown in Figure 9. There followed in 1968 the introduction of the employer survey, the New Earnings Survey (NES), now the Annual Survey of Hours and Earnings (ASHE). There is – as for total income - a striking inverse-U. All three different elements of the series show an increase of at least 0.5, and the NES/ASHE series depicts a fall from more than 4.5 to around 3.

Conclusions

The broad picture shown at the outset in Figure 2 is of a "dramatic" fall in the concentration of top incomes in the UK from 1918 to 1979 and then an almost equally "dramatic" rise in concentration in the next three decades. By these standards, the changes in income concentration in the nineteenth century were "modest". Indeed, placing the two centuries alongside each other has served to underscore the differences between them. One difference is the paucity of comparable information: we know less about the nineteenth century than is commonly believed. The aim of Section II of the paper has been to establish just what can and cannot be said. Moreover, the limited evidence that exists suggests that the widespread view that nineteenth century Britain exemplified the Kuznets curve has – as far as the top of the distribution is concerned – little validity. The "mixed estimates" indicate, if anything, a fall in concentration in the first half of the nineteenth century. The new set of estimates covering only earned incomes, and that imperfectly, suggest an inverse of the Kuznets curve, with a fall and then a rise in concentration, with signs that there was a rise in concentration in the years before the First World War. The latter evidence, coupled with that from the surtax returns, suggests that this period warrants closer examination.

What about Pareto? On the one hand, I believe that the Pareto distribution provides a valuable point of departure, and the Pareto coefficient alpha is a useful summary statistic. On the other hand, the upper tail of the UK income (and earnings) distribution departs from the Pareto in significant ways. The departures manifest themselves in the fact that the three approaches to estimating alpha can lead to different conclusions, and this provides a valuable diagnostic device. There has been a distinct change in the shape of the upper tail since 1918. At the outset, the income rank curve took the form of a concave relationship, but over the first half of the century, the curve comes to rise less steeply and becomes less concave. In the thirty years after 1949, the curve continued to rotate clock-wise, so that within the top 5 per cent there was a lower level of income (relative to the mean) at any rank, and by 1969-70 had become close to Pareto in form. After 1979-80, the curves rotated in the opposite direction and a degree of convexity emerged. In terms of the shape of the elite, the upper tail changed from a concave "baronial" shape to a convex "regal" shape

where the differences become more accentuated as one rises up the income scale. The conclusion is that one should indeed begin with Pareto, but not stop there: we need a richer representation of the upper tail of the income distribution.

Appendices

Appendix 1: Sources of Personal Income data

Table A1: Sources of Inland Revenue and HMRC data on distribution of total incomes

Income in tax year	Nature of survey	Source
1918-19	Special exercise	AR 1919-20, p.70
1919-20	Special exercise	Colwyn Committee, 1927, Appendix XIV
1937-38	Special exercise	AR 1939-40, p.30
1949-50	Quinquennial survey	AR 1950-51, p.97
1954-55	Quinquennial survey	AR 1955-56, p.67
1959-60	Quinquennial survey	AR 1961-62, p.93
1964-65	Quinquennial survey	AR 1965-66, p.120
1969-70	Quinquennial survey	SPI 1969-70, p.11
1974-75	Annual survey	IRS 1977, p.43
1979-80	Annual survey	SPI 1979-80, p.20
1984-85	Annual survey	SPI 1984-85, p.10
1989-90	Annual survey	IRS 1992, p.29
1994-95	Annual survey	IRS 1996, p.35
1999-2000	Annual survey	IR website, table 3.3
2004-05	Annual survey	HMRC website, table 3.5
2005-06	Annual survey	HMRC website, table 3.5
2006-07	Annual survey	HMRC website, table 3.5
2007-08	Annual survey	HMRC website, table 3.5
2009-10	Annual survey	HMRC website, table 3.3
2010-11	Annual survey	HMRC website, table 3.3
2011-12	Annual survey	HMRC website, table 3.3
2012-13	Annual survey	HMRC website, table 3.3
2013-14	Annual survey	HMRC website, table 3.3

Note: AR denotes Annual Report of the Inland Revenue, IR denotes the Inland Revenue, HMRC denotes Her Majesty's Revenue and Customs, SPI denotes Survey of Personal Incomes, and IRS denotes Inland Revenue Statistics.

Appendix 2: Sources of Schedules D and E earnings data and control totals for total employees and total earnings

Table A2: Sources of data on earnings by detailed ranges in the Inland Revenue Publications (UK except where indicated Great Britain (GB))

Income in Tax Year	***Data from Schedule E***	***Information in Annual reports of Inland Revenue***	***Periodical Return (PR) House of Commons Paper: Session and Number***
1842/43			PR 1844: 315
1843/44			PR 1846: 107
1844/45			Ditto
1845/46 GB	PR 1847: 747: first classification titled Return of Charge on Property and Income Tax, under Schedules D and E, 1845-46, p.3.		
1846/47 GB	PR 1849:317		
1847/48 GB	PR 1849:317		
1848/49 GB	PR 1852:480		PR 1851:27
1849/50 GB	PR 1852:480		
1850/51 GB	PR 1852:480		
1851/52 GB	PR 1853:616		
1852/53 GB	PR 1854:341		
1853/54	PR 1855:482 Ireland introduced		
1854/55	PR 1856:313		
1855/56	First AR for year ending 31 December 1856	First AR for year ending 31 December 1856	PR 1857: session 2:69
1856/57	PR 1858:465		
1857/58	PR 1860: 501	Second AR for year ending 31 March 1858	PR 1859 session 2:119
1858/59	PR 1861: 509		
1859/60	PR 1862: 466		
1860/61	PR 1863: 526		
1861/62	PR 1864: 565		
1862/63	PR 1865: 469		
1863/64	PR 1866: 488		
1864/65	PR 1867: 527		
1865/66	PR 1868: 460		

1866/67	PR 1868: 460		
1867/68	PR 1873: 397		PR 1873: 397
1868/69	PR 1873: 397	13th AR for year ending 31 March 1869	Supplement to 24th AR, pages 152-158
1869/70	PR 1873: 397	See below	Supplement to 24th AR, pages 152-158
1870/71	PR 1873: 397	14th AR for years ending 31 Mar 1870 and 1871	Supplement to 24th AR, pages 152-158
1871/72	PR 1873: 397		Supplement to 24th AR, pages 152-158
1872/73	PR 1879: 298, pages 3 and 7	Supplement to 24th AR, pages 152-158	
1873/74	PR 1879: 298, pages 3 and 7	Supplement to 24th AR, pages 152-158	
1874/75	PR 1879: 298, pages 3 and 7	Supplement to 24th AR, pages 152-158	
1875/76	PR 1879: 298, pages 3 and 7	Supplement to 24th AR, pages 152-158	
1876/77	PR 1879: 298, pages 3 and 7	Supplement to 24th AR, pages 152-158	
1877/78 to 1897/98	No detailed ranges		
1898/99	43rd AR, p.147		
1899/00	44th AR, p.137	44th AR for year ending 31 Mar 1901.	
1900/01	45th AR, p.205	45th AR for year ending 31 Mar 1902.	
1901/02	46th AR, p.209	46th AR for year ending 31 Mar 1903.	
1902/03	47th AR, p.222	47th AR for year ending 31 Mar 1904.	
1903/04	48th AR, p.228	48th AR for year ending 31 Mar 1905.	

1904/05	49th AR, p.229	49th AR for year ending 31 Mar 1906.	
1905/06	50th AR, p.225	50th AR for year ending 31 Mar 1907.	
1906/07	51st AR, p.191	51st AR for year ending 31 Mar 1908.	
1907/08	52nd AR, p.173	52nd AR for year ending 31 Mar 1909.	
1908/09	53rd AR, p.137	53rd AR for year ending 31 Mar 1910	
1909/10	54th AR, p.133	54th AR for year ending 31 Mar 1911	
1910/11	55th AR, p.131	55th AR for year ending 31 Mar 1912	
1911/12	56th AR, p.121	56th AR for year ending 31 Mar 1913	
1912/13	57th AR, p.125	57th AR for year ending 31 Mar 1914	
1913/14	58th AR, p.123	58th AR for year ending 31 Mar 1915	

Note: AR denotes Annual Report of the Inland Revenue and PR denotes Periodical Return.

The starting point for the total number of employees is the series of Feinstein (1972, Table 57) for the total in employment, which is given annually from 1855 to 1914. The figures cover the United Kingdom (Great Britain and Ireland). The total includes employees in employment (including members of the armed forces) and employers and self-employed persons. The series is extrapolated backwards from 1855 to 1842 using the estimates of Booth (1886) for 1841, 1851 and 1861, linearly interpolated. The estimates of Booth relate to the total "employed or independent" from which, following Feinstein (1972, p.224, note 1) are subtracted the categories "property owning" and "indefinite". The resulting figure for 1861 is 4.8% higher than the figure of Feinstein, and this adjustment is applied to the interpolated figures.

From the total in employment, we have to subtract employers and self-employed. This can only be done on the basis of strong assumptions. For 1911, Feinstein (1972, Table 11.10) gives an estimate of the total of employers and self-employed of 2.39 million, or 12.1% of the total in employment. However, the ratio of self-employed to employed may well have been changing over time. Here allowance is made for the higher rate of self-employment in agriculture: in 1911, the ratio of self-employed to wage and salary earners is given as 0.36 for agriculture but 0.11 for other sectors (Matthews et al, 1982, Table 6.4). These ratios are applied to the total working population in agriculture and non-agriculture (Feinstein, 1972, Table 60) for 1861, 1871, 1881, 1891, 1901, and 1911, and to estimates for 1841 and 1851 derived from Booth (1886, pages 352, 373, 394 and 426). The resulting adjustment factors are interpolated linearly between these years, and applied to the total employment figures to give the estimates of total wage and salary earners in Table A3.

The total of wages and salaries is based on the series of Feinstein (1972, Table 21) for total personal sector wages and salaries (including Forces' pay). This is available from 1855. The series is extrapolated backwards to 1841 using the series for total taxable income given by Stamp (1916, p.318). (This uses his "true comparative series".)

Table A3: Control totals for total employees and total earned incomes

UK	Total Employees 000	Total wages and salaries (inc forces' pay) £m
1841		
1842		
1843		
1844		
1845	8,977	301
1846	9,138	304
1847	9,299	302
1848	9,461	303
1849	9,623	301
1850	9,785	303
1851	9,948	305
1852	10,031	309
1853	10,114	328
1854	10,197	326
1855	10,280	328
1856	10,419	336
1857	10,395	321
1858	10,158	308
1859	10,751	337
1860	10,925	350
1861	10,815	350
1862	10,645	352

1863	10,878	364
1864	11,275	376
1865	11,380	398
1866	11,373	409
1867	11,047	409
1868	11,083	400
1869	11,266	414
1870	11,613	431
1871	11,961	457
1872	12,133	512
1873	12,209	559
1874	12,234	547
1875	12,259	544
1876	12,197	542
1877		
1878		
1879		
1880		
1881		
1882		
1883		
1884		
1885		
1886		
1887		
1888		
1889		
1890		
1891		
1892		
1893		
1894		
1895		
1896		
1897		
1898	15,406	815
1899	15,717	848
1900	15,826	899
1901	15,910	898
1902	15,940	889
1903	15,970	897
1904	15,894	882
1905	16,206	902
1906	16,588	940
1907	16,724	996

1908	16,163	963
1909	16,334	974
1910	16,999	1,023
1911	17,453	1,051
1912	17,549	1,095
1913	17,928	1,136
1914	17,884	1,236

Source: see text.

References

Aigner, D.J. and Goldberger, A.S. (1970). 'Estimation of Pareto's Law from grouped observations'. *Journal of the American Statistical Association*, 65, 712-723.

Alvaredo, F., Atkinson, A.B., Piketty, T. and Saez, E. (2015). 'The World Top Incomes Database', http://topincomes.g-mond.parisschoolofeconomics.eu/#Database:

Atkinson, A.B. (2005). 'Top incomes in the United Kingdom over the twentieth century'. *Journal of the Royal Statistical Society*, 168, 325-343.

Atkinson, A.B. (2007). 'The distribution of top incomes in the United Kingdom 1908-2000'. In A.B. Atkinson and T. Piketty, editors, *Top incomes over the 20th century*, Oxford: Oxford University Press.

Atkinson, A.B., Piketty, T. and Saez, E. (2011). 'Top incomes in the long run of history'. *Journal of Economic Literature*, 49, 3-71.

Atkinson, A.B. and Voitchovsky, S. (2010). 'The distribution of top earnings in the UK since the Second World War'. *Economica*, 78, 440-459.

Benhabib, J. and Bisin, A. (2016). 'Skewed wealth distributions: Theory and empirics'. NBER Working Paper 21924, National Bureau of Economic Research, Cambridge, Mass.

Booth, C. (1886). 'Occupations of the people of the United Kingdom, 1801-1881'. *Journal of the Royal Statistical Society*, 49, 314-444.

Bottazi, G., Pirino, D. and Tamagni, F. (2015). 'Zipf Law and the firm size distribution: A critical discussion of popular estimators'. *Journal of Evolutionary Economics*, 25, 585-610.

Bowley, A.L. (1914). 'The British super-tax and the distribution of income'. *Quarterly Journal of Economics*, 28, 255-268.

Brandolini, A. (2002). 'A bird's eye view of long-run changes in income inequality'. Bank of Italy Research Department, Rome.

Bresciani-Turroni, C. (1939). 'Annual survey of statistical data: Pareto's Law and the index of inequality of incomes'. *Econometrica*, 7, 107-133.

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, 371-388.

Burkhauser, R.V., Hérault, N., Jenkins, S.P. and Wilkins, R. (2016). 'What has been happening to UK income inequality since the mid-1990s? Answers from reconciled and combined household survey and tax return data'. ISER Working Paper Series 2016-03, University of Essex.

Champernowne, D.G. (1973 (1936)). *The distribution of income between persons*, Prize Fellowship Dissertation 1936, King's College, Cambridge, published in 1973 by Cambridge University Press, Cambridge.

Champernowne, D.G. and Cowell, F.A. (1998). *Inequality and income distribution*, Cambridge: Cambridge University Press.

Chipman, J.S. (1974). 'The welfare ranking of Pareto distributions'. *Journal of Economic Theory*, 9, 275-282.

Cirillo, P. (2013). 'Are your data really Pareto distributed?' *Physica A*, 392, 5947-5962.

Clark, C. (1951). *The conditions of economic progress*, second edition. London: Macmillan.

Colwyn Committee (1927). *Report of the Committee on national debt and taxation*, Cmd 2800. London: HMSO.

Cormann, U. and Reiss, R-D. (2009). 'Generalizing the Pareto to the log-Pareto model and statistical inference'. *Extremes*, 12, 93-105.

Cowell, F.A. (1977). *Measuring inequality*. Oxford: Philip Allan, first edition. (2011), Oxford: Oxford University Press, third edition.

Deane, P. and Cole, W.A. (1964). *British economic growth 1688-1959*. Cambridge: Cambridge University Press.

Department for Work and Pensions (2015). *Households Below Average Income 1994/5–2013/14*. London: Department for Work and Pensions.

Falk, M., Hüsler, J. and Reiss, R-D. (2011). *Laws of small numbers: Extremes and rare events*. Basel: Springer.

Feinstein, C.H. (1972). *National Income, Expenditure and Output of the U.K. 1855-1965*. Cambridge: Cambridge University Press.

Feinstein, C.H. (1988). 'The rise and the fall of the Williamson curve'. *Journal of Economic History*, 48, 699-729.

Fréchet, M. (1945). 'Nouveaux essais d'explication de la répartition des revenus'. *Revue de l'Institut International de Statistique/Review of the International Statistical Institute*, 13, 16-32.

Gabaix, X. (2009). 'Power laws in economics and finance'. *Annual Review of Economics*, 1, 255–293.

Giffen, R. (1904). *Economic inquiries and studies*, volume 1. London: G Bell and Sons.

HMRC (2012). *The Exchequer effect of the 50 per cent additional rate of income tax*, London: Her Majesty's Revenue and Customs.

Jasso, G. (1983). 'Using the inverse distribution function to compare income distributions and their inequality'. *Research in Social Stratification and Mobility*, 2, 271-306.

Jenkins, S.P. (2015). 'The income distribution in the UK: A pattern of advantage and disadvantage'. CASEpaper 186, Centre for Analysis of Social Exclusion, STICERD, London School of Economics.

Jenkins, S.P. (2016). 'Pareto models, top incomes, and recent trends in UK income inequality'. ISER Working Paper 2016-07, University of Essex.

Kuznets, S. (1955). 'Economic growth and income inequality'. *American Economic Review*, 45, 1-28.

Lindert, P.H. and Williamson, J.G. (1983). 'Reinterpreting Britain's social tables, 1688-1913'. *Explorations in Economic History*, 20, 94-109.

Macgregor, D.H. (1936). 'Pareto's law'. *Economic Journal*, 46, 80-87.

Matthews, R.C.O., Feinstein, C.H. and Odling-Smee, J.C. (1982). *British Economic Growth 1856-1973*. Oxford: Clarendon Press.

McDonald, J.B. and Xu, Y.J. (1995). 'A generalization of the Beta Distribution with applications'. *Journal of Econometrics*, 66, 133–152, and (1995a). 'Errata'. *Journal of Econometrics*, 69, 427–428.

Mitchell, B.R. (1988). *British historical statistics*. Cambridge: Cambridge University Press.

Pareto, V. (1896). 'La courbe de la repartition de la richesse'. *Recueil publié par la Faculté de Droit à l'occasion de l'Exposition Nationale Suisse*. Lausanne: Ch. Viret-Genton.

Pareto, V. (2003 (1896)). Translation of Pareto, V. (1896). 'On the distribution of wealth and income'. In F.A. Cowell, editor, *The economics of poverty and inequality*, Volume II, Elgar Reference Collection, Cheltenham, reproduced from M. Baldassarri and P. Ciocca, editors, *Roots of the Italian school of economics and finance: From Ferrara (1857) to Einaudi (1944)*, Volume 2. Houndmills: Palgrave.

Piketty, T. (2001). *Les hauts revenus en France au 20ème siècle*. Paris: Grasset.

Porter, G.R. (1851). 'On the accumulation of capital by the different classes of society'. *Journal of the Statistical Society of London*, 14, 193-199.

Porter, G.R. (1851a). *The Progress of the Nation*. London: John Murray.

Powers, D.M.W. (1998). 'Applications and explanations of Zipf's Law'. In D.M.W. Powers, editor, *New Methods in Language Processing and Computational Natural Language Learning*, Association for Computational Linguistics, 151-160.

Sabine, B.E.V (1966). *A history of income tax*. London: Allen and Unwin.

Sayer, B. (1833). *An attempt to shew the justice and expediency of substituting an income or property tax for the present taxes*. London: Hatchard.

Schumpeter, J.A. (1949). ‘Vilfredo Pareto (1848-1923)’. *Quarterly Journal of Economics*, 63, 147-173.

Shirras, G.F. (1935). ‘The Pareto Law and the distribution of income’. *Economic Journal*, 45, 663-681.

Stamp, J.C. (1916). *British incomes and property*. London: P S King.

Williamson, J.G. (1979). ‘Two centuries of British income inequality: Reconstructing the past from tax assessment data, 1708-1919’, Graduate Program in Economic History, Discussion Paper Series, University of Wisconsin, Madison.

Williamson, J.G. (1985). *Did British capitalism breed inequality?* Boston: Allen and Unwin.

Zipf, G.K. (1932). *Selected studies of the principle of relative frequency in language*. Paris: Librairie du Recueil Sirey.

Zipf, G.K. (1949). *Human behavior and the principle of least effort*. Cambridge, Mass.: Addison-Wesley.