Journal of Economic Inequality (2005) 3: 43–61
DOI: 10.1007/s10888-004-8309-1

Accounting for income distribution trends: A density function decomposition approach

STEPHEN P. JENKINS[1] and PHILIPPE VAN KERM[2,⋆]
[1] *Institute for Social and Economic Research, University of Essex, Colchester CO4 3SQ, United Kingdom, E-mail: stephenj@essex.ac.uk*
[2] *CEPS/INSTEAD, B.P. 48, L-4501 Differdange, G.-D. Luxembourg, E-mail: philippe.vankerm@ceps.lu*

(Received: 6 May 2004; accepted: 21 December 2004)

Abstract. This paper develops methods for decomposing changes in the income distribution using subgroup decompositions of the income density function. Overall changes are related to changes in subgroup shares and changes in subgroup densities, where the latter are broken down further using elementary transformations of individual incomes. These density decompositions are analogous to the widely-used decompositions of inequality indices by population subgroup, except that they summarize multiple features of the income distribution (using graphs), rather than focusing on a specific feature such as dispersion, and are not dependent on the choice of a specific summary index. Nonetheless, since inequality and poverty indices can be expressed as PDF functionals, our density-based methods can also be used to provide numerical decompositions of these. An application of the methods reveals the multi-faceted nature of UK income distribution trends during the 1980s.

Key words: density functions, income distribution, inequality, subgroup decompositions.

1. Introduction

Decompositions of inequality indices by population subgroup have been much used to account for trends in the income distribution. Given a partition of the population into subgroups defined by, for example, age, education, or employment status, inequality in a given year can be written as a function of subgroup population shares, subgroup mean incomes, and subgroup inequalities. The change in inequality between two years can then be related to changes in subgroup population shares, means and inequalities. Explanations of distributional trends are constructed by examining which of the three types of change accounted most for the aggregate change, and for which subgroups. For applications using these methods, see, *inter alia,* Mookherjee and Shorrocks [15], Atkinson [1], Jenkins [12], and Goodman et al. [9], who analyzed trends in the UK, and Tsakloglou [22] who analyzed trends in Greece. Trends in poverty and welfare have been analyzed similarly [17, 14]; so too have cross-national differences [11].

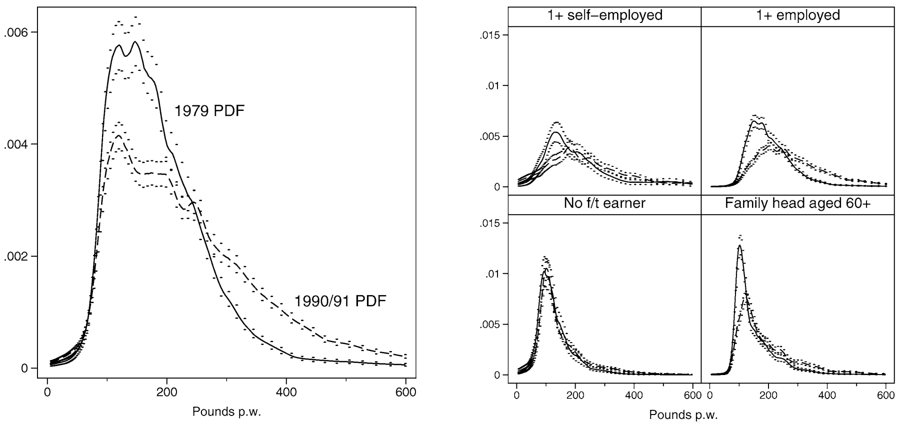
⋆ Corresponding author.

By their very nature, scalar indices of inequality (or poverty) focus on a particular feature of the income distribution and may not capture other aspects such as polarization. Moreover the estimated importance of the different factors in a decomposition may be sensitive to the choice of index used. It is therefore of interest to explore an approach to decomposition that summarizes multiple features of the income distribution, and yet is amenable to subgroup decomposition in the same way that inequality indices are. In this paper, we propose a decomposition method based on income probability density functions (PDFs) that has these capabilities.

The importance of considering multiple features of the income distribution is illustrated by the UK experience over the 1980s. Consider Figure 1(a) which plots estimates of the PDFs for the UK income distribution in 1979 and 1990/91.[1] Income trends over the 1980s were complex in nature. The peak of the income density changed from being near-bimodal to a more complex shape. There was a large shift in the concentration of incomes away from the 1979 peak and down to the right, combined with a small increase of concentration at the very lowest incomes. To characterize these changes as an increase in inequality (as an inequality decomposition analysis might) would omit much of the multi-faceted nature of the changes. A more general approach is required, and the aim of this paper is to provide one.

We show in Section 2 that a PDF can be decomposed by population subgroup in a manner analogous to the subgroup decomposition of inequality indices, with the crucial difference that our decomposition is applied to non-parametrically estimated income PDFs (as in Figure 1), rather than to scalar indices. The method takes account of (changes in) the complete distribution of income within each subgroup, rather than only the mean and spread. We propose a way to characterize within-group income changes, based on 'elementary income transformations' that characterize changes in location, changes in spread, and changes in higher moments of the subgroup distributions. Although our methods focus on PDFs, and results are summarized graphically, they provide a unified framework that may also be used to derive decompositions of changes in scalar inequality and poverty indices, since these indices are functionals of the PDF and so may be calculated from them. And, although our application focuses on trends over time, the methods themselves are more widely applicable, to differences in income distributions across countries, for example.

Our proposed methods bear a familial resemblance to those proposed by DiNardo et al. [8] to account for trends in US wage inequality. Both methods use the subgroup decomposability property of the aggregate PDF, but a key difference is that DiNardo et al. [8] treat each individual as a separate subgroup. Each counterfactual that they examine – for example a change in the minimum wage or in the percentage unionized – is cleverly characterized in terms of a change in each wage earner's sample weight, and the impact on the aggregate PDF is then examined by reweighting the base-year PDF. By contrast, our counterfactual transformations involve not only changes to subgroup shares (analogous to changing the weights



a. Aggregate PDFs b. Subgroup PDFs

Figure 1. Income PDFs for the UK, 1979 and 1990/91: adaptive kernel density estimates, with pointwise 2-SE variability bands.

for groups of individuals), but also changes to the subgroup income distributions themselves.

Our approach provides a broad-brush documentation of the sources of distributional trends – just as inequality index decompositions do – rather than an examination of specific changes in economic institutions à la DiNardo et al. [8].[2] The two approaches are complementary, since broad-brush strategies are often useful when looking at trends in household income. The questions to be answered are often of a different nature to those relevant to analyzing trends in the distribution of wages, because there are multiple income sources and multiple persons per household. The questions include, for example, what is the impact on income of changes in the distribution of employment and unemployment? Are trends in the distribution of income among non-working households of as great a significance as changes in the distribution among working households? And what about the impact of the growth in self-employment and self-employment earnings? Put another way, changes in the distribution of wages are potentially only one part of a story about changes in the distribution of household income. Our methods help point to where subsequent, more detailed, analysis should focus – whether on explaining changes in the distribution of wages or in the prevalence of unemployment, for instance.

We consider these issues in detail in our empirical application (Section 3), but to get a flavour of the sorts of questions addressed, consider Figure 1(b). This shows the changes over the 1980s in the PDFs for four subgroups of the UK population, where individuals have been classified according to the labour market attachment of the family to which they belong (with families with head aged 60+ classified separately). Although the PDF for individuals in a family with no full-time earner hardly changed, substantial mass was shifted to right in the three other PDFs, with the effect most marked for individuals in families with at least one full-time employee. This suggests that changes in earnings from employment played a major role in accounting for the substantial rise in overall income inequality (increasing between-group inequality and inequality within that subgroup). This conclusion is premature, however. What, for example, was the impact of the large change in the distribution of income from self-employment? Moreover, Figure 1(b) does not show the changes in the subgroup population shares and, as we show below, there was a substantial rise in the proportion of persons in families with no full-time earner or with someone in self-employment, and a decline in the fraction of those in families with a full-time employee. What effects did they have? In addition, the figure shows that there were changes in modality, and in the prevalence of very low incomes, and we would also like to account for these, not only for changes in inequality.

2. Decomposing income density function changes

Our proposed decomposition approach has two elements, discussed in turn. First there is the decomposition of changes in a PDF into two terms, one summarizing the effects of changes in subgroup population shares and the other summarizing changes in subgroup distributions. Second, we break the latter term into components summarizing changes in subgroup income location, spread, and other distributional features.

2.1. CHANGES IN SUBGROUP POPULATION SHARES AND SUBGROUP DENSITIES

Our method relies on the additive decomposability of an income PDF. If a population of individuals is exhaustively partitioned into a set of mutually-exclusive subgroups $k = 1, \ldots, K$, the income density function $f(x)$, at each income x, can be expressed as the weighted sum of the PDFs for each subgroup:

$$f(x) = \sum_{k=1}^{K} v^k f^k(x), \tag{1}$$

where v^k is the population share of group k, and f^k is the PDF for group k. The change in the PDF between a base period 0 and a final period 1 can therefore be expressed as the sum of two components:[3]

$$\begin{aligned} \Delta f(x) &= \sum_{k=1}^{K} w^k \Delta f^k(x) + \sum_{k=1}^{K} z^k(x) \Delta v^k \\ &= C_D(x) + C_S(x). \end{aligned} \tag{2}$$

The first term, $C_D(x)$, is the contribution of the change in subgroup distributions to the total change in the density. The second term, $C_S(x)$, is the contribution of changes in subgroup population shares. The w^k and $z^k(x)$ terms are aggregation weights:

$$w^k = \pi v_0^k + (1-\pi)v_1^k, \tag{3}$$

$$z^k(x) = (1-\pi)f_0^k(x) + \pi f_1^k(x) \tag{4}$$

with $0 \leq \pi \leq 1$. Natural choices for the aggregation weights are either base period values v_0^k for w^k and final period values $f_1^k(x)$ for $z^k(x)$, or final period values v_1^k for w^k and base period values $f_0^k(x)$ for $z^k(x)$. These choices correspond to assuming $\pi = 1$ or $\pi = 0$, respectively. Mookherjee and Shorrocks [15, p. 896] pointed out that, for the decomposition of inequality indices, '(a)lthough this particular choice is unlikely to make much difference to the results, it is perhaps appropriate to adopt a compromise between the base and final period weights,' and they proceeded to use an average of base and final weights: $\pi = 0.5$.

The choice of π may also be cast as a sequence issue in marginal decomposition problems, an issue also faced by DiNardo et al. [8], for example. In the first polar case, $\pi = 0$, the effect of changing subgroup population shares is evaluated at the initial values for the subgroup densities whereas the effect of changing the latter is evaluated at the final values for the subgroup shares: changes in population shares are assumed to precede changes in subgroup densities. With the second polar choice, $\pi = 1$, changes in subgroup distributions are assumed to precede changes in subgroup shares. In both cases, the contribution of each factor is measured by its marginal impact on Δf. The Mookherjee and Shorrocks [15] choice, $\pi = 0.5$, corresponds to the contribution that would be assigned by averaging contributions from all possible sequences, i.e. the Shapley value algorithm in a marginal decomposition problem [19]. In our empirical application, we used $\pi = 0, 0.5$, and 1, but the principal conclusions were robust to the choice made. We therefore report results for the case $\pi = 0$ only (the other results are available on request).

2.2. DECOMPOSING SUBGROUP DENSITY CHANGES FURTHER: THE THREE 'S'S OF DISTRIBUTIONAL CHANGE

Accounts based on estimates of the 'changing subgroup shares' and 'changing subgroup densities' components tell only part of the story about income distribution change. It is also useful to be able to break the second component down further and to account for the changes in subgroup PDFs. The key features of each PDF that we focus on are its location and its spread (as in inequality decompositions by population subgroups), plus other features related to higher moments. We characterize changes in PDFs as arising in three different ways, which we label three 'S's of distributional change:

- a *sliding*: a *ceteris paribus* shift of the PDF along the income line;
- a *stretching*: a *ceteris paribus* increase in spread around a constant mean; and
- a *squashing*: a *ceteris paribus* disproportionate increase in density mass on one side of the mode.

These distributional changes are related to assessments of changes in welfare. If assessments are based on a social welfare function $W(x)$ that satisfies the property of monotonicity, then a sliding of the distribution to the right implies an increase in welfare [3, p. 99]. A stretching implies a decrease in welfare according to social welfare functions that are increasing and S-concave functions of incomes [5]. The process of decomposition can therefore also help identify situations in which welfare has unambiguously increased or decreased, for subgroups taken separately and for the population as whole or, more commonly, draw attention to potential conflicts in welfare assessments (for example if an increase in average income is combined with an increase in inequality). Welfare assessment criteria are not so well developed for the changes encapsulated in the squashing component: it reflects changes in higher-order moments of the distribution, polarization and other changes in modality. We discuss the interpretation of this component further below.

Given the definitions of sliding, stretching, and squashing, we decompose the 'changes in subgroup densities' term, $C_D(x)$ in Equation (2), into a sum of three components, to be added to the component reflecting changes in subgroup population shares. Thus the decomposition of the change in the aggregate density has four components:

$$\Delta f(x) = \underbrace{C_{D1}(x) + C_{D2}(x) + C_{D3}(x)}_{C_D(x)} + C_S(x), \tag{5}$$

where $C_{D1}(x)$, $C_{D2}(x)$ and $C_{D3}(x)$, measure the impacts of sliding, stretching, and squashing, respectively.

To estimate $C_{D1}(x)$, C_{D2} and $C_{D3}(x)$, we use an approach based upon elementary transformations of densities. Suppose that there is an income transformation function that describes the relationship between base and final period income for each individual within a given group. That is, for each subgroup k, we have $x_1 = g_k(x_0)$, for some arbitrary transformation g_k. This implies a relationship between subgroup k's income PDF in the two periods [21, p. 20]:

$$f_1^k(x) = \left| \frac{d(g_k^{-1}(x))}{dx} \right| f_0^k(g_k^{-1}(x)). \tag{6}$$

By choosing a particular g_k, we can construct counterfactual PDFs that reflect various characterizations of income changes. For example, controlling for the shifting and stretching of subgroup PDFs is straightforward. We assume that, within each subgroup, the relationship between income in year 1 and income in year 0 is linear: $x_1 = \alpha_k + \beta_k x_0$. The resulting PDF for group k is therefore

$$\zeta^k(x) = \left| \frac{1}{\beta_k} \right| f_0^k\left(\frac{x - \alpha_k}{\beta_k} \right) \tag{7}$$

and the counterfactual aggregate PDF is obtained by summing the subgroup densities. We use ζ to refer to counterfactual constructs based on linear income transformation functions; f refers to actual base or final period density functions.

Consider an income transformation consisting of an equal addition to all incomes: $\alpha_k = a$ and $\beta = 1$. The density function implied by Equation (7) is the initial PDF shifted along the income line. Mean income is increased by a but the variance is left unchanged. Hence to construct a counterfactual PDF that incorporates the change in subgroup means, we simply apply this income transformation and calibrate a so that the mean income in the counterfactual distribution, $E(\zeta_1^k)$, is equal to the mean income in the observed final distribution, $E(f_1^k)$, i.e.

$$\alpha_k = E(f_1^k) - E(f_0^k). \tag{8}$$

Denote the counterfactual PDF obtained after such a transformation $\zeta_1^k(x; \mu_1^k, \sigma_0^k)$, where μ_1^k and σ_0^k reflect the fact that mean income is at its final period value and the variance is at its base period value.

Now consider a second income transformation incorporating a Sandmo [18] increase in dispersion which stretches the PDF around a constant mean. Within each subgroup, each income in the second period is a fraction s of initial income and a fraction $(1 - s)$ of base-period subgroup mean income:

$$x_1 = sx_0 + (1 - s)E(f_0^k). \tag{9}$$

The parameters of the linear transformation in this case are $\alpha_k = (1 - s)E(f_0^k)$ and $\beta_k = s$. Mean income remains constant but the variance increases by a factor s^2. Hence we can construct a counterfactual PDF that incorporates a *ceteris paribus* increase in income dispersion by calibrating the transformation parameters so that $\mathrm{Var}(\zeta_1^k) = \mathrm{Var}(f_1^k)$ with

$$s = \sqrt{\frac{\mathrm{Var}(f_1^k)}{\mathrm{Var}(f_0^k)}}. \tag{10}$$

Denote the counterfactual PDF obtained after such a transformation $\zeta_1^k(x; \mu_0^k, \sigma_1^k)$.

The two preceding transformations can be combined to construct a counterfactual PDF that allows for changes in subgroup means and in variances with the following transformation parameters:

$$\alpha_k = E(f_1^k) - sE(f_0^k), \quad \text{and} \quad \beta_k = s. \tag{11}$$

These parameters imply a counterfactual density, $\zeta_1^k(x; \mu_1^k, \sigma_1^k)$, with the same mean and variance as the second period density: $E(\zeta_1^k) = E(f_1^k)$ and $\mathrm{Var}(\zeta_1^k) = \mathrm{Var}(f_1^k)$.

We combine the three counterfactual constructs just described to compute the elements of the decomposition set out in Equation (5). First each subgroup density change is decomposed into the estimated contributions of location, spread and squashing:

$$\begin{aligned}
&\Delta f^k(x) \\
&= \underbrace{\eta(\zeta_1^k(x;\mu_1^k,\sigma_0^k) - f_0^k(x)) + (1-\eta)(\zeta_1^k(x;\mu_1^k,\sigma_1^k) - \zeta_1^k(x;\mu_0^k,\sigma_1^k))}_{C_{D1}^k\text{: subgroup mean effect (sliding)}} + \\
&\quad + \underbrace{\eta(\zeta_1^k(x;\mu_1^k,\sigma_1^k) - \zeta_1^k(x;\mu_1^k,\sigma_0^k)) + (1-\eta)(\zeta_1^k(x;\mu_0^k,\sigma_1^k) - f_0^k(x))}_{C_{D2}^k\text{: subgroup variance effect (stretching)}} + \\
&\quad + \underbrace{f_1^k(x) - \zeta_1^k(x;\mu_1^k,\sigma_1^k)}_{C_{D3}^k\text{: subgroup residual effect (squashing)}} \quad . \qquad (12)
\end{aligned}$$

Each of the ‘S’ factors is evaluated in terms of its marginal impact on $\Delta f^k(x)$ in a sequential approach. Just as π controls the sequence in which changes in shares and changes in PDFs are introduced, η controls the order in which changes in means (sliding) and variances (stretching) are introduced and, again, either polar values (0 or 1) or compromise values (e.g., 0.5) can be adopted. We report results for $\eta = 1$.[4]

Observe that the squashing component is defined as a residual, and so in principle might reflect changes in the modality of the income distributions as well as changes in skewness and kurtosis of the underlying subgroup distributions (or other higher moments). Our view, however, is that, conditional on an appropriate definition of subgroups, the second type of changes is the relevant one. That is, in our experience, subgroup distributions are virtually always unimodal, and so changing modality in the aggregate distribution reflects changes in the mixture of unimodal subgroup distributions. This situation is illustrated in Figure 1(b).

In the second and final step, the decomposition of $\Delta f^k(x)$ is plugged into Equation (2) and the terms are arranged so that $C_{D1}(x)$ is the sum of each subgroup's sliding component ($C_{D1}^k(x)$) weighted by w^k as in (3), and similarly for the spread and squashing components, $C_{D2}(x)$ and $C_{D3}(x)$. This gives us estimates for each term in Equation (5).

2.3. ADDITIVE *versus* MULTIPLICATIVE CHANGES IN INCOMES

The methods developed above for constructing counterfactual distributions are based on additive transformations: we keep spread constant by assuming equal additions to all incomes, and use the variance as a measure of spread. The additive approach appears well-suited to a PDF decomposition since the visual impact of equal additions is a sliding of the PDF along the x-axis with no change in shape. However the Lorenz curves of the base year and counterfactual distributions are *not* held constant; relative inequality changes. In constrast, equi-proportionate income changes shift the mean but keep the Lorenz curve, and therefore relative inequality, unchanged. In this ‘multiplicative’ case, the variance changes and the rightward shift of the PDF is accompanied by a flattening (or shrinking) of the shape of the distribution.

The multiplicative approach is consistent with the most popular way of summarizing inequality – using relative measures.[5] Counterfactuals based on this approach are straightforwardly obtained by first taking a logarithmic transformation of incomes, and then applying the methods to the distribution of log income rather than income. Equal additions in the log-scale result in a change of location such that the geometric means of the counterfactual and final year distributions are equal, while keeping the income shares, and therefore the variance of log income, constant. Similarly, a Sandmo increase in the spread of log income changes the variance of log income while keeping the geometric mean constant. No new tools are required.

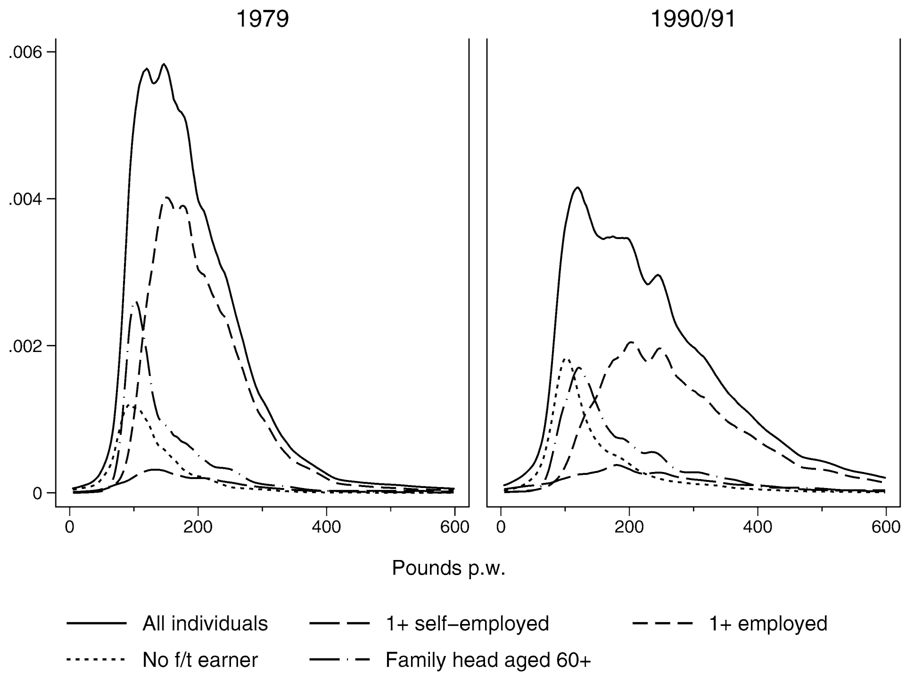
3. UK income distribution changes between 1979 and 1990/91

We illustrate the methods with analysis of the changes in the UK income distribution between 1979 and 1990/91 using data from the 'Households Below Average Income' (HBAI) subfiles of the Family Expenditure Survey (pooled data for 1990 and 1991).[6] The two years span a period of high inequality growth.[7] Most previous analysis has focused on a specific feature of the distribution – inequality, poverty, or mean income – rather than looking at changes in the distribution as a whole.

We concentrate on results for one partition of the population, characterized by the attachment of each individual's family to the labour market. This partition was chosen because previous analysis has shown that it provides the most insightful picture into UK income distribution trends over the 1980s, and because it led to subgroup PDFs that were clearly unimodal (see the earlier discussion). The four groups of individuals identified were: (i) individuals living in a family with one or more full-time self-employed persons (the '1+ self-employed' subgroup); (ii) individuals living in a family with one or more full-time employees (the '1+ f/t employed' subgroup); (iii) individuals living in a family with no full-time income earner (the 'no f/t earner' subgroup), and (iv) individuals living in a family with a household head aged 60 years or more (the '60+ family head' subgroup). The family heads in subgroups (i)–(iii) were aged less than 60. Full-time employment was taken to mean working at least 30 hours per week.[8]

The first step was estimation of income density functions for each subgroup and for the population as a whole. We used an adaptive kernel density estimator [23]. The advantage of this estimator is that it does not over-smooth the distribution in zones of high income concentration, while keeping the variability of the estimates low where data are scarce, for example in the highest income ranges (see, e.g., [20, 16]). Standard errors of all the estimates were obtained by bootstrap resampling. The whole procedure (PDF estimation and decomposition) was repeated for each of 500 bootstrap resamples, and the standard errors reported summarize the variability of the estimates from the 500 replications.

Figure 2 plots the 1979 and 1990/91 income PDFs for the population as a whole (solid line), together with the subgroup PDFs (dotted and broken lines). The density for each subgroup at each income level has been multiplied by the subgroup's population share, so that the weighted sum of the subgroup densities adds up to



Note: Each subgroup PDF is weighted by the relevant subgroup population share.

Figure 2. PDFs for 1979 and 1990/91.

the population density, in accordance with Equation (1). The subgroup population shares are reported in Table I. Observe the large rise in the proportion of individuals in families with no full-time earner (from 12 percent to 17 percent) and in the proportion in self-employed families (from 6 percent to 10 percent), and the corresponding fall in the proportion of individuals in families with at least one member in full-time employment (from 62 percent to 51 percent). Table I also summarizes well-known facts: mean income, inequality and relative poverty increased substantially for the overall population as well as in all subgroups taken separately. The statistics reported do not capture the other changes that occurred, however.

In both 1979 and 1990/91, the lower mode of the aggregate PDF corresponded with the modes of the PDFs for the 'no f/t earner' and '60+ family head' subgroups (people with little or no employment income), and the upper mode corresponded to the mode of the '1+ f/t employed' subgroup PDF. It is clear from Figure 2 that the movement of the '1+ f/t employed' PDF made a large contribution to the spreading of the aggregate PDF. But the effect of changes in subgroup shares is hard to identify in such a plot, and it is also difficult to disentangle the role of differential increases in mean income between groups and the role of the general increase in spread within all subgroups.[9] The pictorial representations of Figures 1(b) and 2 and the summary indices presented in Table I suggest important candidate

Table I. UK income distribution, 1979 and 1990/91: summary statistics

Subgroup	Population share		Mean		Coefficient of variation		Std. dev. of Log		Relative Gini		Absolute Gini		FGT(0) ×100		FGT(1) ×100	
	1979	1990/91	1979	1990/91	1979	1990/91	1979	1990/91	1979	1990/91	1979	1990/91	1979	1990/91	1979	1990/91
All individuals	1.00	1.00	188	245	0.49	0.61	0.46	0.59	0.25	0.31	47	75	8	18	1.4	4.1
			(1.3)	(1.6)	(0.009)	(0.007)	(0.006)	(0.005)	(0.003)	(0.002)	(0.8)	(0.9)	(0.4)	(0.4)	(0.09)	(0.12)
1+ self-employed	0.06	0.10	215	267	0.62	0.70	0.61	0.75	0.31	0.33	67	89	16	19	4.2	7.1
	(0.003)	(0.003)	(9.2)	(7.0)	(0.027)	(0.024)	(0.027)	(0.024)	(0.013)	(0.009)	(4.7)	(4.0)	(1.9)	(1.2)	(0.62)	(0.55)
1+ employed	0.62	0.51	209	291	0.41	0.49	0.37	0.46	0.21	0.25	44	73	3	9	0.4	1.7
	(0.006)	(0.004)	(1.5)	(2.1)	(0.011)	(0.007)	(0.005)	(0.006)	(0.004)	(0.003)	(1.0)	(1.2)	(0.3)	(0.4)	(0.05)	(0.10)
No f/t earner	0.12	0.17	130	155	0.54	0.67	0.51	0.55	0.25	0.30	32	46	7	10	2.2	2.4
	(0.004)	(0.004)	(2.6)	(2.4)	(0.048)	(0.024)	(0.022)	(0.012)	(0.009)	(0.007)	(1.7)	(1.5)	(0.8)	(0.8)	(0.36)	(0.24)
Family head	0.20	0.22	154	202	0.53	0.63	0.42	0.52	0.25	0.30	38	60	3	11	0.3	1.7
aged 60+	(0.004)	(0.003)	(2.1)	(2.0)	(0.025)	(0.012)	(0.009)	(0.006)	(0.006)	(0.004)	(1.4)	(1.2)	(0.4)	(0.6)	(0.05)	(0.12)

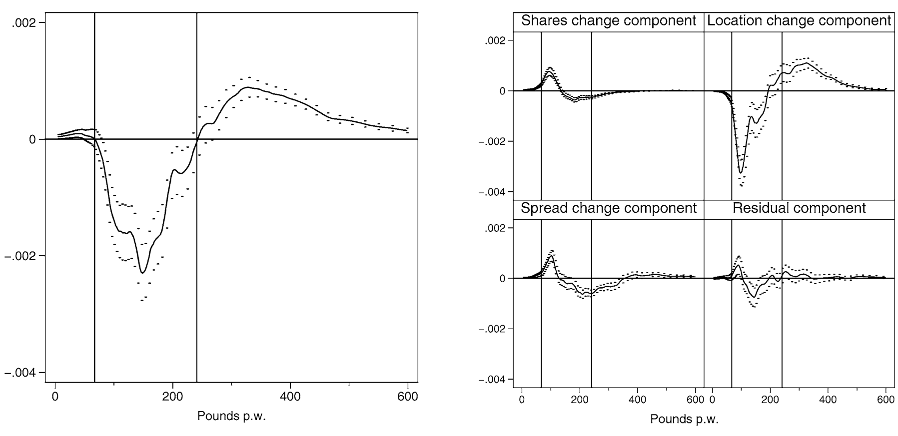
Notes: Incomes are expressed in constant 1993 pounds per week. All statistics (except population shares) estimated by numerical integration of the relevant density function estimates. Standard error estimates based on 500 bootstrap replications reported in parentheses. FGT(0) and FGT(1) are the headcount ratio and average normalized gap poverty indices. The poverty line is 50 percent of contemporary median income.

explanations for the distributional change but, at the same time, questions remain about which was the most important of them. Our density function decomposition methodology provides a means of isolating the contributions of the various factors and quantifying their impact.

3.1. DECOMPOSITION OF THE CHANGE IN THE PDF

Results from the decomposition exercise are presented in Figure 3. A multiplicative model provided a better fit of the changes than did an absolute model (apparent from a visual inspection of plots not shown), as well as more stable estimates of the contribution of the different factors when parameters π and η were varied. The results refer to the case $\pi = 0$ and $\eta = 1$. That is, we assessed the impact of changing subgroup population shares by comparing the 1979 PDF with a *ceteris paribus* change in shares, and we assessed the effect of the three 'S's of change with 1990/91 population shares, and allowed first for the change in means, then in the change in spread.[10]

Figure 3(a) shows the difference between the 1990/91 and 1979 PDFs, together with pointwise two-standard-error variability bands. The mass at very low incomes increased slightly, but there was a decrease in the density between £70 and £240. This was accompanied by a substantial increase in density over the range £240 to £500. The four components of the decomposition of the change are presented in Figure 3(b). The further away from zero the line is at any income level, the more of the change in the density that is accounted for that component (the contributions may be in the same direction or the opposite direction as the aggregate change); if a component had no impact, the relevant graph would be a horizontal line at zero.



a. Observed PDF change b. Explanatory components

Figure 3. PDF changes between 1979 and 1990/91, with pointwise two-SE variability bands: observed, and by explanatory component.

The vertical lines in each of the plots mark the incomes at which the change in the PDF was zero.

Changes in PDF location stand out as having made the largest contribution to changes in the aggregate PDF. The increase in mean income in every subgroup shifted density mass to the right, with a steep fall concentrated at about £100 and increases at all incomes greater than £200. The next most important components were the contributions from changes in spread and in population shares, which were of similar shape and magnitude. They accounted for the increase in mass at very low incomes. At most income levels (incomes below £400), they tended to offset the effect of higher mean incomes but, overall, their effects were dominated by it. The residual component is the least important. Its contribution was similar to that for the shares and spread components, but there was also much greater sampling variability compared to the other factors.

A quantitative summary of the relative importance of the different components to aggregate PDF change is provided by the areas under the curves in Figures 3(a) and 3(b). The integral of $|C_{D1}(x)|$ over x, 0.410, a measure of the total mass 'displaced' by the changes in means component, is much greater than the corresponding value for the population shares component (0.083), for the spread component effect (0.153) and for the residual component factor (0.081). For comparison, the displacement in observed mass was 0.405.

In sum, the change in PDF between 1979 and 1990/91 – the decrease in density mass in the middle-income range and the increase at higher incomes – is mostly accounted for by changes in the location of the subgroup PDFs (a sliding effect). The increase in mass at very low incomes is also identified well.

Our calculations also enabled us to assess whether each of the explanatory components was due to changes for a particular subgroup or by changes experienced more universally. Additional estimates (not shown) indicate that the large location change component arose mostly from rightward sliding in incomes for the '60+ family head' subgroup and, most especially, for the '1+ f/t employed' subgroup. (There were similar, but much smaller, changes for the other two subgroups.) Changes in spread among the same two groups accounted for virtually all of the overall stretching effect. Changes in the distribution of income for the '1+ self-employed' subgroup were relatively large, but they accounted for little of the change in the aggregate PDF, simply because the subgroup's population share was relatively small. (There was one exception to this: the increase in spread for this subgroup accounts for the increase in density mass at very low incomes.) In contrast, although the number of individuals in families with no full-time earner was relatively large (and almost doubled over the period), changes to the subgroup's income distribution were small by comparison with those for the other subgroups. The only explanatory component to which changes for this subgroup made much of a contribution was the residual one (and this component was itself small relative to the other three).

Table II. Estimates of marginal contributions to changes in summary statistics

				Marginal contribution of			
				subgroup	subgroup	subgroup	subgroup
Index	1979	1990/91	Change	shares	mean	spread	residual
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
A. *Quantiles:*							
P10	98	102	4	–5	20	–8	–4
	(0.8)	(0.8)	(1.2)	(0.4)	(1.4)	(0.8)	(1.2)
			[100]	[–125]	[548]	[–219]	[–104]
P25	126	140	14	–6	29	–8	–0
	(1.0)	(1.2)	(1.6)	(0.6)	(1.5)	(0.7)	(1.1)
			[100]	[–45]	[202]	[–56]	[–1]
P50	170	212	42	–7	47	–4	6
	(1.3)	(1.7)	(2.0)	(0.6)	(1.8)	(0.8)	(1.5)
			[100]	[–16]	[112]	[–10]	[14]
P75	230	311	81	–5	67	14	5
	(1.8)	(2.2)	(2.9)	(0.7)	(2.4)	(1.3)	(2.3)
			[100]	[–6]	[82]	[17]	[6]
P90	296	433	137	–3	89	44	7
	(2.9)	(4.3)	(5.4)	(1.1)	(3.3)	(3.4)	(4.0)
			[100]	[–2]	[65]	[32]	[5]
B. *Inequality measures:*							
P90/P10 ratio	3.01	4.24	1.23	0.12	0.23	0.67	0.22
	(0.037)	(0.052)	(0.065)	(0.014)	(0.040)	(0.054)	(0.059)
			[100]	[10]	[19]	[54]	[18]
P75/P25 ratio	1.83	2.23	0.39	0.05	0.09	0.21	0.04
	(0.016)	(0.021)	(0.027)	(0.007)	(0.015)	(0.017)	(0.026)
			[100]	[14]	[22]	[54]	[10]
P50/P10 ratio	1.73	2.08	0.35	0.01	0.10	0.10	0.13
	(0.015)	(0.019)	(0.024)	(0.006)	(0.018)	(0.018)	(0.028)
			[100]	[4]	[30]	[29]	[37]
P90/P50 ratio	1.74	2.04	0.30	0.06	0.02	0.25	–0.02
	(0.018)	(0.021)	(0.027)	(0.007)	(0.007)	(0.020)	(0.027)
			[100]	[18]	[7]	[82]	[–8]
Relative Gini	0.25	0.31	0.06	0.01	0.01	0.04	0.00
	(0.003)	(0.002)	(0.004)	(0.001)	(0.002)	(0.004)	(0.003)
			[100]	[17]	[14]	[62]	[6]
Absolute Gini	47	75	29	1	15	11	2
	(0.8)	(0.9)	(1.3)	(0.3)	(0.7)	(1.1)	(0.9)
			[100]	[2]	[53]	[39]	[6]

Table II. (Continued)

				Marginal contribution of			
				subgroup	subgroup	subgroup	subgroup
Index	1979	1990/91	Change	shares	mean	spread	residual
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
C. *Poverty measures:*							
FGT(0) (×100)	7.77	18.27	10.50	0.81	2.86	4.99	1.84
	(0.410)	(0.409)	(0.604)	(0.267)	(0.483)	(0.554)	(0.576)
			[100]	[8]	[27]	[48]	[18]
FGT(1) (×100)	1.44	4.13	2.70	0.27	0.53	1.25	0.64
	(0.086)	(0.123)	(0.150)	(0.039)	(0.102)	(0.148)	(0.132)
			[100]	[10]	[20]	[46]	[24]

Notes: Incomes are expressed in constant 1993 pounds per week. All statistics estimated by numerical integration. Standard error estimates based on 500 bootstrap replications reported in parentheses. Numbers in brackets show the percentage change and marginal contributions expressed as a percentage of the total change. Any difference between the sum of marginal contributions and the actual change is due to rounding. FGT(0) and FGT(1) are the headcount ratio and average normalized gap poverty indices. The poverty line is 50 percent of contemporary median income.

3.2. DECOMPOSITION OF CHANGES IN SUMMARY INDICES

We have emphasized the relevance of looking at the income distribution as a whole but, of course, there is also interest in particular features such as inequality and poverty. To draw conclusions about more specific aspects of distributional change, we can derive counterfactual indices of poverty, inequality, and other other summary statistics, from the counterfactual distributions since the statistics are functionals of the PDFs. Table II reports changes in selected quantiles, five relative inequality indices, one absolute inequality index, and two relative poverty indices, together with the estimated contributions to the change of the four explanatory components.

The estimates for the quantiles are consistent with the results obtained from inspection of Figure 3. All quantiles increased, reflecting the sliding effect, but this was offset by the impact of changes in population shares (there were more people in the worst-off groups) and in spread (at the lower quantiles). At the tenth percentile the offsetting effect was large (and the residual component was offsetting too), so that the actual increase was only four percent. At higher quantiles, the effect of changes in location dominated. The median increased by 24 percent, and the ninetieth percentile increased by 46 percent. Note that, in the latter case, the effect of changes in population shares change was negligible, and the spread effect contributed substantially to the increase too (though by less than the location effect did).

The effects of changes in spread were more important for inequality and relative poverty. It was mostly the increase in spread within subgroups that accounted for the increase in each relative inequality index. Changes in location, which reflect changes in income between groups for these indices, played a secondary role. One exception is the P50/P10 ratio, for which changes were accounted for by location, spread and residual components in similar proportions. Unsurprisingly, changes in the absolute Gini were mainly driven by the location component. The large increase in relative poverty was principally accounted for by the increase in income spread within subgroups and changes in higher moments (the residual component). Perhaps surprisingly, the population shares component accounted for less than 10 percent of the poverty increase.

4. Concluding remarks

This paper has developed a PDF decomposition methodology to account for income distribution trends, analogous to that based on decompositions of inequality indices by population subgroup. We have shown that a change in a density may be decomposed into terms accounting for the effects of changes in subgroup population shares and in subgroup densities. The second term may itself be decomposed into three terms representing the impacts of the three 'S's of distributional change: sliding (changes in location), stretching (changes in spread), and squashing (changes in higher moments). Although we focused our discussion and empirical application on changes over time, the methodology has wider application, for example to analysis of differences in income distributions between countries.

Our application of the PDF decomposition methodology to UK income distribution trends between 1979 and 1990/91 unravelled what was, at first glance, a complex change. Two forces acted in opposite directions. On the one hand, increases in income levels shifted density mass towards higher income levels, and these were also responsible for some flattening of the PDF since the largest gains were obtained by the most well-off group of people with access to employment income. On the other hand, there was an increase in the proportion of the population in subgroups that had relatively low average income, accompanied by an increase in income spread within each subgroup. Although these were offsetting factors, their effects were much smaller than the effects of changes in average incomes. The most marked change was in the distribution for individuals in families with at least one member in full-time employment. This suggests that the trend in the distribution of household income was likely to have been driven by changes in the distribution of wages.

Previous research on inequality trends in Britain over the 1980s has emphasized the contribution of within-group inequality changes. We have also found this. But, in addition, we have shown that when one is interested in explaining changes in the overall income distribution, and its multiple features, then it is increases in income levels rather than increases in inequality – sliding rather than stretching –

that played the dominant role. The welfare implications of growth in both income levels and spread depend on assumptions about the shape of the social welfare function – an issue explored further by Jenkins [14] using the same data. Interestingly, although there was a striking change in the modality of the aggregate distribution, virtually all of this could be accounted for by changes in subgroup location and spread: subgroup squashing effects were negligible. Our results underline the usefulness of general and flexible tools for analysis of the multiple dimensions of distributional change.

Acknowledgements

This paper forms part of the research programme of the TMR network Living Standards, Inequality and Taxation (Contract ERBFMRXCT 980248) of the European Communities whose financial support is gratefully acknowledged. This research is also supported under the Transnational Access to major Research Infrastructures programme (Contract HPRI-CT-2001-00128) hosted by IRISS-C/I at CEPS/INSTEAD Differdange (G.-D. Luxembourg) and by core funding for ISER from the Economic and Social Research Council and the University of Essex. We thank Lars Osberg for helpful comments.

Notes

[1] The data are described in detail in Section 3.

[2] See also Hyslop and Maré [10] or Daly and Valletta [4].

[3] For comparisons of income distributions between countries, '0' and '1' would refer to a pair of countries.

[4] Conclusions were robust to the choice of η. Results are available from the authors on request.

[5] The multiplicative model has also been used in other contexts where analysts have needed to hold inequality constant when developing counterfactuals. See, e.g., Datt and Ravallion [6] and Van Kerm [24].

[6] The HBAI data are nationally representative, cover all income groups, and form the basis of the official income distribution statistics. See Department of Social Security [7] for further details. We focus on the distribution of income among individuals, attributing (in conventional fashion) each person with the income of the household to which they belong. We use the HBAI 'before housing costs' measure of household income, i.e. real net income, equivalized using the McClements equivalence scale. Net income includes cash income from all sources, minus income tax payments and employee National Insurance contributions. Sampling weights were used to account for differential non-response. All incomes were expressed in April 1993 prices.

[7] We should stress that our results refer to the period as a whole. The relative importance of different factors changed during the 1980s, since, for example, the rise in unemployment was sharpest at the very beginning of the decade. The episodic nature of UK distributional trends has been emphasized by Atkinson [2].

[8] Employment-related partitions were used by Atkinson [1], Jenkins [12], and Goodman et al. [9], to analyze inequality trends. The definitions employed here were necessary to ensure comparability over time. We also considered several other partitions: by 'receipt of Supplementary Benefit or Income Support (recipient vs. non-recipient)', 'gender of household head', 'age of household head',

and 'family type' (pensioner, childless couple, male-headed family with children, female-headed family with children). The results obtained with these partitions provided a less satisfactory account of the 1979 and 1990/91 changes than did the partition on which we focus here.

[9] Jenkins [13] also reported Figure 2, but did not undertake any formal decomposition analysis.

[10] Densities were estimated for log income and then these, and associated counterfactual densities, were back-transformed to the natural income scale (as in the figures).

References

1. Atkinson, A.B.: What is happening to the distribution of income in the UK?, In: A.B. Atkinson (ed.), *Incomes and the Welfare State: Essays on Britain and Europe*, Cambridge University Press, Cambridge, UK, 1995, pp. 15–40.
2. Atkinson, A.B.: Bringing income distribution in from the cold, *Economic Journal* **107**(441) (1997), 297–321.
3. Cowell, F.A.: Measurement of inequality, In: A.B. Atkinson and F. Bourguignon (eds.), *Handbook of Income Distribution*, North-Holland, Amsterdam, The Netherlands, 2000, pp. 87–166.
4. Daly, M.C. and Valletta, R.G.: Inequality and poverty in the United States: The effects of changing family behavior and rising wage dispersion, Working Papers in Applied Economic Theory 2000–06, Federal Reserve Bank of San Francisco, San Francisco, USA. Revised March 2004. Forthcoming in *Economica*.
5. Dasgupta, P.S., Sen, A.K. and Starrett, D.: Notes on the measurement of inequality, *Journal of Economic Theory* **6**(2) (1973), 180–187.
6. Datt, G. and Ravallion, M.: Growth and redistribution components of changes in poverty measures: A decomposition with applications to Brazil and India in the 1980s, *Journal of Development Economics* **38**(2) (1992), 275–295.
7. Department of Social Security: *Households Below Average Income. A Statistical Analysis. 1979-1990/91*, HMSO, London, UK, 1993.
8. DiNardo, J., Fortin, N.M. and Lemieux, T.: Labor market institutions and the distribution of wages, 1973–1992: A semiparametric approach, *Econometrica* **64**(5) (1996), 1001–1044.
9. Goodman, A., Johnson, P. and Webb, S.: *Inequality in the UK*, Oxford University Press, Oxford, UK, 1997.
10. Hyslop, D. and Maré, D.: Understanding changes in the distribution of household incomes in New Zealand between 1983–86 and 1995–98, Working Paper 01/21, New Zealand Treasury, Wellington, New Zealand, 2001.
11. Jäntti, M.: Inequality in five countries in the 1980s: The role of demographic shifts, markets and government policies, *Economica* **64**(255) (1997), 415–440.
12. Jenkins, S.P.: Accounting for inequality trends: Decomposition analyses for the UK, 1971–86, *Economica* **62**(245) (1995), 29–63.
13. Jenkins, S.P.: Did the middle class shrink during the 1980s? UK evidence from kernel density estimates, *Economics Letters* **49**(10) (1995), 407–413.
14. Jenkins, S.P.: Trends in real income in Britain: A microeconomic analysis, *Empirical Economics* **22**(4) (1997), 483–500.
15. Mookherjee, D. and Shorrocks, A.F.: A decomposition analysis of the trend in UK income inequality, *Economic Journal* **92**(368) (1982), 886–902.
16. Pagan, A. and Ullah, A.: *Nonparametric Econometrics*, Themes in Modern Econometrics, Cambridge University Press, New York, USA, 1999.
17. Piachaud, D.: Poverty in Britain 1899 to 1983, *Journal of Social Policy* **17**(3) (1988), 335–349.
18. Sandmo, A.: On the theory of the competitive firm under price uncertainty, *American Economic Review* **61**(1) (1971), 65–73.

19. Shorrocks, A.F.: Decomposition procedures for distributional analysis: A unified framework based on the Shapley value, Unpublished paper, University of Essex, Colchester, UK, 1999.
20. Silverman, B.W.: *Density Estimation for Statistics and Data Analysis*, Monographs on Statistics and Applied Probability, Chapman and Hall, London, UK, 1986.
21. Stuart, A. and Ord, J.K.: *Kendall's Advanced Theory of Statistics*, Vol. 1, 5th edn, Charles Griffen, London, UK, 1987.
22. Tsakloglou, P.: Aspects of inequality in Greece: Measurement, decompositions and intertemporal change: 1974, 1982, *Journal of Development Economics* **40**(1) (1993), 53–74.
23. Van Kerm, P.: Adaptive kernel density estimation, *Stata Journal* **3**(2) (2003), 148–156.
24. Van Kerm, P.: What lies behind income mobility? Reranking and distributional change in Belgium, Western Germany and the USA, *Economica* **71** (2004), 223–239.