



Measuring National Income Growth Democratically: Methods and Estimates for the United Kingdom

Andrew Aitken and Martin Weale

ESCoE Discussion Paper 2022-17

June 2022

ISSN 2515-4664

DISCUSSION PAPER

Measuring National Income Growth Democratically: Methods and Estimates for the United Kingdom

Andrew Aitken and Martin Weale

ESCoE Discussion Paper No. 2022-17

June 2022

Abstract

Measures of aggregate real income and average real income per head are some of the best-known and frequently cited statistics around the world; but they are a poor measure of human progress, even in terms of income. By construction they ‘weight’ households by their share in total income, rather than their share in the population. We show how a democratic measure of real national disposable income growth can be produced. Our measure allocates the whole of national income, as shown in the national accounts, to a sample of individual households. We also use a democratic deflator rather than a plutocratic price index to derive our measure of real income. To make this possible, we present new methods to reconcile discrepancies between survey data and the national accounts. The data we produce for the UK show declining inequality over the period 2006 to 2017 and, consistent with that, we find that the growth rate of real democratic national income per household (after adjusting for household size) grew by 0.1% more than the real plutocratic income per household over the period in question.

Keywords: Real national income, Inequality, Measure of income growth, Data imputation

JEL classification: E01, I31, D31, C43

Andrew Aitken, National Institute of Economic and Social Research
andrew.aitken@oecd.org

Published by:
Economic Statistics Centre of Excellence
National Institute of Economic and Social Research
2 Dean Trench St
London SW1P 3HE
United Kingdom
www.escoe.ac.uk

ESCoE Discussion Papers describe research in progress by the author(s) and are published to elicit comments and to further debate. Any views expressed are solely those of the author(s) and so cannot be taken to represent those of the Economic Statistics Centre of Excellence (ESCoE), its partner institutions or the Office for National Statistics (ONS).

© Andrew Aitken and Martin Weale

Measuring National Income Growth Democratically: Methods and Estimates for the United Kingdom

Andrew Aitken* and Martin Weale†
Economic Statistics Centre of Excellence
National Institute of Economic and Social Research
2, Dean Trench Street,
London SW1P 3HE

June 2022

Abstract

Measures of aggregate real income and average real income per head are some of the best-known and frequently cited statistics around the world; but they are a poor measure of human progress, even in terms of income. By construction they ‘weight’ households by their share in total income, rather than their share in the population. We show how a democratic measure of real national disposable income growth can be produced. Our measure allocates the whole of national income, as shown in the national accounts, to a sample of individual households. We also use a democratic deflator rather than a plutocratic price index to derive our measure of real income. To make this possible, we present new methods to reconcile discrepancies between survey data and the national accounts. The data we produce for the UK show declining inequality over the period 2006 to 2017 and, consistent with that, we find that the growth rate of real democratic national income per household (after adjusting for household size) grew by 0.1% more than the real plutocratic income per household over the period in question.

Keywords: Real national income, Inequality, Measures of income growth, Data imputation

JEL Codes: E01, I31, D31, C43

Disclaimer: This work contains statistical data from ONS which are Crown Copyright. The use of the ONS statistical data in this work does not imply the endorsement of the ONS in relation to the interpretation or analysis of the statistical data. This work uses research datasets which may not exactly reproduce National Statistics aggregates.

*National Institute of Economic and Social Research and Economic Statistics Centre of Excellence.
andrew.aitken@oecd.org

†King’s College, London, Economic Statistics Centre of Excellence and Centre for Macroeconomics.
martin.weale@outlook.com

Contents

1	Introduction	3
2	Methodology to distribute UK national income to households	5
3	Survey and tax data	7
4	Reallocation of sectoral income accruing outside households	8
4.1	Non-financial corporations	10
4.2	Financial corporations and the treatment of occupational pensions	10
4.3	General government	12
4.4	Non-profit institutions serving households	13
4.5	Income reallocated	13
4.6	Depreciation	14
5	Imputation and scaling	14
5.1	Imputation methods	17
5.2	Imputation of pension wealth	18
5.3	Imputation of interest and dividend income	19
5.4	Imputation of housing rent	26
5.5	Estimation of correlations between stochastic imputations	26
5.6	Scaling labour income	28
5.7	Scaling mortgage interest payments	29
5.8	Scaling tax payments	30
5.9	Other discrepancies	32
6	Deflation	32
7	Estimates of measures of inequality and growth in democratic and plutocratic income	34
8	Conclusion	40
9	References	41
	Appendices	44
A	Models used to impute missing data	44
A.1	Interest and dividends	44
A.2	Pension rights	50
A.3	Housing rent	54
B	Deflators	57

1 Introduction

This paper describes the construction of a democratic measure of national income growth. Frequent headlines discuss the rise and fall of GDP *per capita*, and there is frequent public discussion about people who have not shared equally in this growth. This disconnect between movements in aggregate output or income, and the distribution of income to individuals has led to initiatives to explore measures of individual well-being that go beyond GDP *per capita*. Strong GDP growth is considered evidence of good fortune for all the people in a country, and yet it is entirely possible for GDP *per capita* or average household income to be rising year-on-year, reflecting growth in top or middle incomes, while incomes at the bottom of the distribution may in fact be falling. The fact that many people may be “left behind” makes these statistics misleading for policy makers and politicians who rely on them, making it difficult clearly to understand the effects of income inequality or fashion more equitable economic policies. At the same time GDP or GDP *per capita* has a long history and a dominant hold on the popular imagination; the simplicity of a single number is surely one of the virtues of GDP. Other indicators of wellbeing, such as those currently produced by the ONS, are often in the form of a dashboard of multiple metrics¹, or alternatively composite measures that aggregate different measures in often arbitrary ways. While other metrics are needed to measure dimensions of wellbeing other than income, the usual metrics measure even income growth poorly. In this paper we propose a democratic measure of national income growth that treats the income growth of each household equally, and has the advantage that it is a *single number*. The stages by which we compile our measure of democratic income growth also allow, for the first time for the UK, estimates of Gini coefficients when the whole of national income has been allocated to households.

The limitations of GDP are widely known. It is, in contrast to the original measure used in the national accounts (Meade and Stone, 1941) measured gross rather than net of depreciation and it does not take any account of property income due abroad or received from abroad. It ignores inequalities in income, destruction of the natural environment, differences in longevity, or the quality of social relations, for example. Nor does it include childcare or other work in the home. Thus the Stiglitz-Sen-Fitoussi Commission established by the French government in 2008 argued that the emphasis of our measurement system should shift “from measuring economic production to measuring people’s well-being” (Stiglitz *et al.*, 2009). They also recommended more focus on income rather than GDP, and it is natural to focus on measures *per capita* or per household. However treating GDP or income growth as a measure of progress is problematic even in terms of income. Macroeconomic aggregates like these are conventionally

¹See the ONS dashboard here: [Measures of National Well-being Dashboard](#).

constructed such that the growth experiences of the individual people/households are aggregated according to their shares in total income. Thus the experience of high-income people/households contributes more towards the aggregate growth rate than does the experience of low-income people/households. One may therefore describe growth in such conventional measures as plutocratic rather than democratic. A democratic measure of income growth would in contrast treat every individual or household (after appropriate adjustment for household size) equally.

Aitken and Weale (2020) set out the theory behind the democratic measure of income, showing that it can be interpreted in terms of a simple logarithmic welfare function. They produced estimates for household income in the UK. In this paper we extend that work to provide estimates for real national disposable income. In practice, our democratic measure of income growth is structured around the growth in the geometric mean of household income, deflated by a democratic price index. While Fleurbaey (2009) is sceptical about the use of income-based aggregates as indicators of well-being, Sefton and Weale (2006) establish a link between income growth and welfare growth, at least subject to the assumptions that consumption and savings decisions are the result of inter-temporal optimisation. Much micro-economic work has focused on household consumption as a guide both to welfare and to inequality. It is often argued that consumption may be a better guide to life-time income than is actual income at any point in time. Nevertheless, as Sefton and Weale (2006) show, income measures life-time welfare accruing and, for policy purposes that may also be of interest. Furthermore, of course, much of the public discussion is in terms of income rather than consumption and, on these grounds alone, it is important to produce indicators which provide metrics relevant to the public. One of these is a measure of the average real income growth experience of households.

This paper shows how it is possible to allocate the whole of UK national income to a sample of individual households, addressing both the discrepancies between the survey data and the national aggregates, and the allocation of income accruing outside the household sector. That makes it possible to estimate, using as broad a definition of income as possible, the geometric mean of household incomes after adjusting for household size.

In the next section we discuss the reallocation to households of those components of national income which accrue outside households. The subsequent section focuses on the imputations that we make, both as a part of the reallocation process and to address the issue of mis-recording in the Living Costs and Food Survey. These imputation methods are in part stochastic and, after discussing deflation, we therefore present stochastic estimates of inequality indicators and of growth in nominal and real democratic net disposable income. This is followed by our conclusions.

2 Methodology to distribute UK national income to households

In this section we describe the main concepts and methodology we use to distribute UK national income to a sample of UK households. Our work is closely related to Piketty *et al.* (2018) who introduce the concept of distributional national accounts (DINA). For the United States they produce a set of national accounts in which the whole of national income is allocated to households. Their work on DINA builds on a succession of studies which have constructed top income shares for a large number of countries (see Piketty and Saez (2003a), Piketty and Saez (2003b), and the two multi-country volumes on top incomes edited by Atkinson-Piketty 2007, 2010; see also Atkinson *et al.* (2011) and Alvaredo *et al.* (2013) for surveys of this literature).

In fact there is a long tradition of national accountants taking an interest in the distribution of income. The UK national accounts, first published in 1941, included tables of the distribution of taxable incomes from 1942 until 1982. Kuznets, one of the founders of the US national accounts, was also one of the first to combine national income series and income tax data in order to estimate the evolution of the share of total income going to different percentiles of the income distribution (Kuznets, 1953). In the UK, Atkinson and Harrison (1978) combined historical inheritance tax data with capital income data and wealth surveys to study the long-run evolution of wealth distribution for Britain over the period 1922-1972. The purpose of these studies is to provide estimates of the distribution of income (and often wealth) using concepts of income and wealth that are consistent with the macroeconomic national accounts. Our approach broadly follows this previous work in combining micro survey data, national accounts data and imputation techniques; at the same time of course it reflects the particular characteristics of UK data sources.

We make two important departures from the approach advocated by Piketty *et al.* (2018). First, our work focuses on the household and thus reflects the economies of scale enjoyed by multi-person households over individuals, while their work looks at individuals. Secondly, we allocate public consumption across households either on the basis of estimated use by households of the services in question, or, for those services which cannot be allocated in this way, with the assumption that consumption is proportional to effective household size. This contrasts with their approach which assumes that consumption of public services depends on income.

Jenkins (2017) and Burkhauser *et al.* (2018) show the evolution of top incomes in the UK and illustrate the problem in relying on a single number such as GDP or GDP per capita when income growth has

diverged between those at the top of the income distribution and the rest of society. These papers emphasise that household survey data do not typically capture income at the extreme top of the distribution very well, which means that survey-based estimates of overall income inequality are biased downwards. The ‘top incomes’ literature mentioned above, in contrast tends to use administrative record data on personal income tax returns, reporting estimates of top income shares - the share of total income received by the richest 1% or richest 10%, and so on. While it is important to gain a greater understanding of the evolution of the distribution of top incomes, by definition, the focus on the top decile is not particularly relevant to the experience of the majority of the population. These studies also focus solely on individual income (total taxable income from the labour market, plus taxable government transfers and investment income), and before or after the deduction of income tax. In contrast, our work focuses on national income; including sources of income that are not directly attributable to households, such as undistributed corporate income, pension income, and other benefits such as social transfers in kind. Furthermore, there are often large discrepancies between the survey totals and the equivalent variables shown in aggregate national accounts data as Tonkin (2015) and Brewer *et al.* (2017) show for the UK and Fesseau and Matteonetti (2013) demonstrate internationally; we resolve these differences.

In analysing trends in UK income inequality, Burkhauser *et al.* (2018) and Jenkins (2017) have combined estimates from tax return data and household survey data. We also follow this approach of combining administrative and survey data. The Department for Work and Pensions (DWP) carries out an adjustment of income data in the the HBAI (Households Below Average Income) by using HMRC tax data from the Survey of Personal Incomes (SPI) to correct for under-coverage of top incomes in the HBAI. The ‘HBAI-SPI’ variables are the source of official UK income distribution statistics. Burkhauser *et al.* (2018) provide an alternative to the DWP approach to improve the top income coverage of the HBAI data. They place density mass over a greater income range than the DWP adjustment and they also use population and income control totals from the World Wealth and Income Database (WID) data series. The latter is necessary as the SPI does not cover all individuals in the UK or all of their income.

In contrast to Burkhauser *et al.* (2018) and Jenkins (2017), the definition of income that we use is based on the idea of “national income”, i.e. gross domestic product minus consumption of fixed capital, plus net foreign income. According to the 2008 SNA national income can be defined using either a production or an income approach. According to the production approach, national income is defined as the sum of net domestic product (i.e. gross domestic product, minus consumption of fixed capital) and foreign income (net foreign inflow of capital and labour income). Net domestic product can then be broken down as the sum of the net value added of each institutional sector (household sector, financial

and non-financial corporate sector, government sector, non-profit sector) and of “taxes on products” (i.e. value-added type taxes and other product taxes, which according to SNA 2008 are not attributed to the value-added of any particular sector).

According to the income approach, national income is defined as the sum of primary incomes of each institutional sector plus taxes on production. Primary income of the household sector (including unincorporated businesses) is the largest component, and is equal to the addition of total employee compensation (including all employer social contributions), net mixed income (i.e. self-employment income), net operating surplus of the household sector (i.e. rental value of housing owned by property income received by households (“property income”, as defined by SNA 2008, can be further decomposed into interest, dividends, etc., and other financial income flows).

As noted above, the household is the unit of observation in our work, in contrast to many studies in the ‘top incomes’ or ‘distribution’ literature that focus on individuals. Most work (see for example Lewbel (1997), Schafer (1999) and Crossley and Pendakur (2010)) suggests, however, that it is better to work with households, adjusted for household size using an equivalence scale, as the fundamental unit for welfare analysis. Historically and, we suspect for some time into the future, the information on incomes that we require to construct our democratic measure of national income will be provided by household surveys rather than tax records despite the Digital Economy Act (2017).

3 Survey and tax data

We use a combination of administrative tax data and data from several household surveys. The HMRC income tax return data are from the public-release files of the Survey of Personal Incomes (SPI) for each year from 2005/06 to 2018/19, with the exception of 2008/9, for which no data have been released (HM Revenue and Customs, 2021). Each year’s SPI is a stratified sample of the universe of tax returns. There are 756,015 individuals in the 2018/19 data representing 31,357,371 million taxpayers (weighted). The SPI data underlie the UK top income share estimates in the WID.world database.

The main unit record survey data we use come from the ONS Living Costs and Food Survey (LCFS), and the associated subfiles of derived variables called Earnings, Taxes and Benefits (ETB) (Office for National Statistics and Dept for Environment and Affairs, 2021). The LCFS is a cross-sectional random survey of about 5,500 households (12,700 individuals) carried out each year. In addition to income, there are survey questions on employment, benefits and assets. Expenditure data are also collected, largely by means of an expenditure diary kept for two weeks. Information on age, occupation, education received, family composition and housing tenure is also obtained. The LCFS is the main source of data

for the ETB, which is the main data used by the ONS to examine trends in income inequality, for example.

The ETB also uses data from administrative sources to enhance the data, and particularly VAT data to improve estimates of indirect taxes and NHS and education system data for benefits in kind.²

We also use data from the Wealth and Assets Survey (WAS) (Office for National Statistics Social Survey Division, 2020). This is a longitudinal survey of approximately 20,000 households designed to collect data on wealth but also asking information on income. It began in July 2006 with interviews carried out over two years, and consists of an individual and household questionnaire of all those aged 16 years and over.

The HBAI data set is produced by the Department of Work and Pensions using the results of the Family Resources Survey (FRS) The FRS has a sample of approximately 19,000 households.

4 Reallocation of sectoral income accruing outside households

In producing democratic estimates of income growth it is necessary to allocate the whole of net national disposable income to households. Here we explain the basis on which this is done.

The Living Costs and Food Survey (LCFS) is often used to provide estimates of the distribution of household income in the United Kingdom (see Stone and Stone (1972) for an early reference to this work), again after adjusting for household size. These estimates show the distribution of income both before and after the redistributive activities of the government. The latter involve the collection of taxes used to finance the provision of benefits and publicly-provided consumption. Benefits and direct taxes are reported in the survey and indirect taxes are allocated on the basis of reported consumption patterns. The large individual items of public consumption are education and health services (social transfers in kind); these are allocated to the individual households on the basis of what is known about the use of the education system and the use of the health service.³ No allocation is made of collective

²Social benefits in kind that are allocated in the ETB include the National Health Service (NHS) (including Health and Social Care in Northern Ireland), state education, adult social care, school meals and Healthy Start vouchers (including nursery milk and school milk, universal infant free school meals in England, free school meals for children in primary 1 to 3 in Scotland and the provision of free breakfast to all pupils in maintained primary schools in Wales), housing subsidies, railway travel subsidies, and bus travel subsidies (including concessionary fares schemes).

³The allocation of health services is based on average cohort use rather than individual use. It is not assumed that someone who has been provided with expensive health treatment has been a large consumer. Rather it is as though insurance were provided and the insurance is imputed.

consumption, such as public administration and defence. Additionally no allocation is made of the undistributed income of companies or pension fund accruals.

While this ONS work on the LCFS provides the core of an allocation of national income to households there are a number of further steps which are needed. First, in common with many household surveys, the aggregates implied are well below those shown in the national accounts. Second, even when that has been addressed, some basis is needed for allocating the undistributed income of companies and pension funds. It is, of course, possible to allocate undistributed corporate income on the basis of dividend income but given the issues we identify with the dividend data we do not attempt that at this stage. We also explain how we handle pension funds in a manner consistent with the current System of National Accounts and this points to a straightforward treatment of the pension fund transactions which do not directly impinge on households.

Table 1: The Sectoral Allocation of Net Disposable Income, £m 2017 Financial Year

	Sector code	Net disposable income
UK total economy	S.1	1,727,159
Non-financial corporations	S.11	26,207
Financial corporations	S.12	24,104
General government	S.13	365,607
Non-profit institutions serving households (NPISH)	S.15	38,186
Households	S.14	1,273,056

Source: UK Economic Accounts 2020Q2. Rounding of the depreciation estimates means that the components do not sum exactly to the total.

We work on a financial year (FY) basis because our basic survey sources, the LCFS/ETB are on a financial year basis. We structure our discussion around the data by sector for net disposable income in FY 2017 shown in Table 1. These are calculated gross of depreciation from the sectoral tables in the 2020Q2 vintage of the UK Economic Accounts.⁴

The current European System of National Accounts sets out, step by step the transactions which take us from Gross Domestic Product to net disposable income. Thus table 1.7.1 of the UK’s national accounts shows production with Gross Domestic Product as the “residual” use of the resources created in the production process. Table 1.7.2 shows the generation of income - the income accruing to capital and labour and also, because transactions are valued at market prices, the value of indirect taxes net of subsidies accruing to the government. In table 1.7.3 primary income is redistributed by means of

⁴We have used 75% of the FY 2017 figure and 25% of the 2018 figure to give an estimate of sectoral depreciation in financial year FY 2017. See section 4.6

dividend and interest payments. A secondary redistribution of income, shown in table 1.7.4 occurs as a result of transfer payments. This second category includes not only government direct taxes and benefits but also pension contributions and pensions received by households. Disposable income gross of depreciation (B.6g) is the amount left within each sector after these primary and secondary receipts have been credited and payments have been debited, while net disposable income (B.6n) is calculated after deduction of depreciation.

The data in Table 1 are shown at market prices. Thus the revenue from indirect taxes, net of subsidies is, in effect, treated on the same basis as factor income. We need to allocate the income accruing to corporations (non-financial and financial) and general government and non-profit institutions serving households to the household sector. We discuss subsequently the treatment of indirect taxes.

We now discuss sector by sector the approach we use to reallocating disposable income.

4.1 Non-financial corporations

Non-financial corporations include manufacturers, utilities, airlines, retailers and some farms for example; firms that produce goods and services, but do not, as a primary activity deal in financial assets and liabilities. They exclude small businesses which are not incorporated; the income from these is self-employment income. We allocate the income accruing to non-financial corporations to households in proportion to their dividend receipts.⁵ The allocation is done after we have imputed a figure for dividend receipts to each household and scaled the imputed data to be consistent with the figure for household dividend receipts shown in the national accounts (see section 5.3).

4.2 Financial corporations and the treatment of occupational pensions

In the European System of Accounts transactions between households and pension schemes are treated like taxes and benefits. Contributions depress household disposable income while receipts of pensions increase it. The investment income earned by pension schemes is described as a social contribution supplement. It is treated as a component of households' primary income but then deducted, along with pension contributions in order to arrive at household disposable income. Households' transactions with pension funds in FY 2017 can be summarised in table 2 taken from tables 1.7.4 and 1.7.5 of the national accounts.

Since we are concerned about the income of each household in the LCFS/ETB survey, it is not adequate

⁵Piketty *et al.* (2018) suggest using dividends plus capital gains. In the UK there are no data available on capital gains.

Table 2: Household Transactions with Pension Funds, 2017 Financial Year

	Code	£m
Employers' Actual Contributions	D.611	59,533
Employers' Imputed Contributions	D.612	12,895
Households' Actual Contributions	D.613	10,794
Social Contribution Supplements	D.614	62,896
Less Service Charges	D.6SC	-22,576
Total Resources	D.61	123,542
Benefits paid	D6.22	87,756
Pension adjustment	D.8	35,786
Total Uses		123,542

Source: *National Accounts*, Table 1.7.4 and 1.7.5

simply to allocate the pension adjustment shown in table 2. The households on whose behalf the employers' contributions are made are different from the households receiving benefits. Each of the three types of contribution, the contribution supplement net of the service charge and the benefits paid need to be allocated to households in a manner consistent with the data in table 2. We then treat contributions and contribution supplements as additions to observed household disposable income, while benefits paid become a form of dis-saving and are deducted from the national accounts figure which includes them as household income.

This treatment then moves private pension schemes from being a system of taxes and benefits to being a form of saving. Even if aggregate benefits were exactly paid out of contributions so that there was no net contribution supplement and no pension adjustment, those households on whose behalf the contributions were made would be seen as saving; their income would increase but their consumption would be unaffected. Those households receiving benefits would, however, see their income written down. With consumption unaffected that would appear as a reduction in their saving.

It should be noted that this is, nevertheless, less than ideal. The definition of pensions in the national accounts includes only those provided through schemes sponsored by employers, and does not cover arrangements such as self-invested pension plans (SIPPs) or pensions managed by insurance companies. One might think therefore that this adjustment is an understatement of the amount which should be allocated to households in this way. A further adjustment would need to be done with reference to global figures for pension liabilities which distinguished those identified in the national accounts from those which are not. We are not aware of any such data.

After dealing with pensions, we then have to reallocate the residual income of the financial corporations.

We deduct the D.8 adjustment of table 2 from the total shown in table 3 below. This adjustment is in fact larger than the disposable income of the financial corporations. We allocate the balance, -£11682m, in proportion to receipts of dividends as with non-financial corporations.

Taking financial and non-financial corporations together, corporate net disposable income is, after deducting the pension adjustment, £14,525m.⁶

4.3 General government

Like the household sector, the government uses its disposable income for consumption or saving. Government consumption is done on behalf of households and should be allocated as far as possible to those households which benefit from the services that it provides. Total government consumption in FY 2017 was £388,973m. This exceeds the total of government disposable income in FY 2017 resulting in savings of -£23,366bn.

For social benefits in kind (non-monetary transfers) we rely on the ETB. The ETB also allocates education and health expenditures to individual households on the basis of household structure. The benefits of education attributed to each household are done on the basis of the composition of the household. Data from a range of sources are used to provide the costs per full-time equivalent student in maintained special schools, nursery, primary and secondary schools, universities, and other further education establishments.

Health expenditure is allocated on an insurance principle, i.e. the expenditure on households with particular demographic characteristics is averaged across the households with those characteristics whether they actually obtain treatment or not. This is entirely consistent with our broad aim of producing a welfare measure; while a general increase in the availability of health treatment would be assumed to increase welfare, a high level of consumption by any particular household is not assumed to increase the welfare of that household. Rather it is implicitly assumed that the household is protected from a loss of welfare.

The allocation of education expenditures in ETB results in totals which are lower than those shown in the national accounts; scaling is used to correct this. Data consistent with the national accounts show consumption of £145,042m on health and £68,714m on education, leaving £175,217m to allocate

⁶Of course it is possible that the European System of Accounts does not define income appropriately for the financial sector. In particular the boundary between income and capital gains will always be relatively arbitrary. At the same time it would, of course, be wrong to infer that the negative accrual of disposable income means that income must be understated.

in addition to these amounts in FY 2017.⁷

There are broadly two possible approaches to allocating this sum, which includes provision of public administration, defence, social protection, justice and fire services among others. One, adopted by Piketty *et al.* (2018) is to allocate it in proportion to household income while the alternative is to allocate on a uniform basis, after adjusting for household size. A case can be made for each of these; we adopt the second, allocating this spending to each household in proportion to its size as measured by the household equivalence scale. The overall reallocation we describe as the government consumption adjustment.

Because government revenue does not normally add up to total government expenditure, it is necessary to impute the primary government deficit or surplus to households. Like Piketty *et al.* (2018) we allocate this government saving adjustment equally into two halves. Since the overall amount is negative, the first is allocated across households in proportion to their tax payments. The second is allocated in proportion to allocated government consumption, and is deducted from these. Thus the assumption is that a deficit has to be covered by means of reduced spending or increased taxation.⁸

4.4 Non-profit institutions serving households

Non-profit institutions serving households (NPISH) are institutions that a) provide goods and services, either free or below the market prices, b) mainly derive their income from grants and donations, and c) are not controlled by government. In the UK examples of NPISH organisations include: most charities, trade unions, religious organisations, political parties, and the majority of universities. As with those categories of government consumption which are not attributable to individual households, we allocate the income of non-profit institutions serving households on the assumption that it accrues to households on the basis of their size as measured by the equivalence scale.

4.5 Income reallocated

Table 3 illustrates the reallocation of income discussed in the previous sections, showing the amounts reallocated by each of the four types of adjustment. We can see that much the biggest item is government consumption, comprising 85% of the total re-allocated to households.

⁷General government final consumption expenditure (GGFCE) sub-sectors quarterly data Quarter 3 (July to Sept), Month 3 2020.

⁸There is a case for allocating government net capital formation in the way that we have allocated residual government consumption. This will result in a larger deficit to be allocated across households.

Table 3: The Reallocation of Sectoral Net Disposable Income, 2017 Financial Year (£m)

	Initial Income	Pensions	Dividends	Govt. cons.	Govt. saving	NPISH	Reallocated Income
UK total economy	1,727,159						1,727,159
Non-fin. corps	26,207		-26,207				0
Fin. corps	24,104	-35,786	11,682				0
Gen. gov.	365,607			-388,973	23,366		0
NPISH	38,186					-38,186	0
Households	1,273,056	35,786	14,525	388,973	-23,366	38,186	1,727,159

Source: Authors' calculations from the UK Economic Accounts 2020Q2. Note that rounding errors associated with the depreciation estimates mean that the elements of the table do not always sum exactly to the totals.

4.6 Depreciation

Data on depreciation are available only on a calendar year basis; quarterly income figures are gross. We deduce financial year estimates of depreciation, by taking three quarters of the value for the calendar year in which the financial year starts and one quarter for the year in which it finishes. Since depreciation figures are not erratic, it is hard to see this as a material source of error.

5 Imputation and scaling

Table 4 shows the relevant national accounts data for FY 2017 and juxtaposes these with the relevant LCFS/ETB survey aggregates. The magnitude of the reconciliation task is clear. The table also indicates the imputation method we have used to reconcile the survey with the national accounts; by modelling (M), by scaling (S) or by allocating according to household size (H). We also identify those items for which there is no explicit entry in ETB and which we have imputed using an alternative data source

Much work on reconciliation is done simply by scaling (e.g. Fesseau and Matteonetti (2013) and Fixler and Johnson (2014)); this can be done at any desired level of disaggregation. Thus Jorgenson and Schreyer (2015) reconcile consumption figures by scaling expenditure on each of fifteen categories of consumption. A major concern with scaling is that any mis-reported zeros will remain as zeros. This is a very material issue with investment income in the LCFS since there are considerably more zero entries shown for tax payers than are suggested by the Survey of Personal Incomes (HM Revenue and Customs, 2021), compiled from tax returns. Reporting zero rather than actual receipts may be a material factor behind the mis-recording discussed above; if this is the case, use of upward scaling will tend to accentuate estimates of inequality. For this reason we avoid using scaling on its own if there is a practical alternative

Table 4: Reconciliation Table: 2017 Financial Year (£m)

	National accounts variable	National accounts value	LCFS value	LCFS/ETB source	Alignment method	Section reference
<i>Primary incomes</i>						
Wages and salaries	D.11	839,600	784,149	wagern+impinc	S	5.6
Net operating surplus	B.2n+D.423	122,182	13,591	b360	M	5.4
Self-employment income	B.3n+ D.422	158,846	111,066	p037+p047	S	5.6
Employers' contributions		175,069	Nets out			
Interest receipts	D.41	20,902	8,312	p081+p082	M	5.3
Dividend receipts	D.421+D.443 +D.45	73,103	11,121	p083	M	5.3
Attributable to policy-holders	D.441	20,902		Proportional to insurance spend	S	5.3
Payable on pension entitlements	D.442	62,896	Nets out			
Less Interest paid	D.4	-24,154	-29,854	b130+b150+b237	S	5.7
EQUALS Net primary income	B.5n	1,447,471				
<i>Secondary redistribution</i>						
Social security benefits in cash	D.621	101,685	87,488	p033	S	5.6
Other social insurance benefits	D.622	134,011	116,039	p049	S	5.6
Social assistance benefits in cash	D.623	120,721	82,388	casben-retpen	S	5.6
Miscellaneous current transfers	D.75	11,036	2,708	b265	S	5.6
Hholds' actual social contributions	D.613	-74,436	-71,805	p071+eesnic	S	5.6
Miscellaneous current transfers	D.75	-23,999	-42,668	b486+b168+b110 +b229+b188+b205 +b206+ck5221t +ck5222t+ck5223t +b334h	S	5.6
Taxes on employment	D.511pt	-151,978	-87,259	p077-p391-b370	M	5.8
Taxes on self-employment & other	D.511pt	-27,728	-6,288	b391	M	5.8
Tax residual	D.51-D.511pt	-7,634			M	5.8
Other current taxes	D.59	-40,704			M	5.8
Pensions supplements less service charge ²	D.614-D.61SC	-40,320	Nets out		M	5.2
Employers' contributions		-175,069	Nets out			
EQUALS Hhld net disposable income (A)	B.6n	1,273,056				
<i>MEMO</i>						
Depreciation	P.51c	67,536	Not recorded			
Gross disposable income	B.6g	1,340,592	Not recorded			
<i>Pensions</i>						
Employer Contributions	D.611+D.612	67,954	Scaled from wages	p008+p011	S	5.2
Household Contributions	D.613	10,275	22,263	p071	S	5.2
Supplement ²	D.614	62,896	Not recorded		M	5.2
Service charge	D.61SC	-22,576	Not recorded		M	5.2
Less Benefits paid to households	D.622	-87,756			S	5.2
EQUALS Pensions adjustment (B)	D.8	35,786	Not recorded	Derived		
<i>Retained elsewhere in private sector</i>						
Retained earnings of companies excl. pension funds (C)		-14,525	Adjust dividend receipts		S	4.1
Net Income of NPISH (D)		38,186	Allocate on basis of effective household size		H	4.4
<i>Government</i>						
Consumption of government		388,973	Allocate educ. and health from ETB.		S	4.3
			Other consump. on effective household size		H	4.3
Net saving of government		-23,366	Allocate half in prop. to taxes, half in prop. to govt. cons.			4.3
EQUALS Net income of government (E)		365,607				
Residual Income (F)		701				
Net National Disposable Income (A+B+C+D+E+F)		1,598,420				

Source: Authors' calculations from LCFS, ETB and UK national accounts.

¹Excludes imputed rent.

²National Accounts 2017 shows in table 6.2.4 a payment of £74,068m for contribution supplements by households, but receipts by financial companies in table 4.1.4 of only £73,692m. This means that the supplement net of the service charge is £54,308m in the household account and £54,239m in the pensions account. There is no discrepancy in earlier years

The main literature on imputation addresses the problem of missing observations in a single data source (e.g. Rubin (1987) and van Buuren (2007)), but the situation we face is one where we assume at least some of the observations are faulty but we do not know which to accept and which to reject. In such circumstances Piketty *et al.* (2018) and Saez and Zucman (2016) suggest drawing on other data sources such as other surveys, provided those surveys provide better data on the variable in question than are available in the core data source. If imputation is carried out on the basis of covariates which are observed in both the source data and in the secondary survey records of interest, then there must be a need for a stochastic term to represent the random variation between the relevant covariates and the data of interest. Neglect of this will not be an issue if only arithmetic averages are of interest. Since, however, one of the purposes of using survey data is to address distributional issues, the treatment of any stochastic terms is in fact likely to be of considerable importance. In addition to the estimation of a stochastic term for variables which we are imputing in the LCFS, we need to take into account the possible correlation between different sources of income. We therefore estimate covariances between the stochastic terms, reflecting these in our imputations. This is explained further in section 5.5.

In the United Kingdom the focus of work on mis-reporting has been on household income gross of taxation. Jenkins (2021) provides a clear summary of the approach described by Webber *et al.* (2020). This involves using tax records to make adjustments to the top end of the income scale. However, as table 4 shows, there is under-recording of incomes such as social security benefits in cash in much the same proportion as the under-recording of wages and salaries, suggesting that the problem is not confined to the top of the distribution. Brewer *et al.* (2017) point to considerable under-recording of incomes at the bottom end of the income scale, and these can exert a substantial influence on a democratic measure of real income growth calculated from the mean of the log of income and its growth rate. Separately we are concerned about allocating components of income which do not accrue directly to households, such as income accruing on pension rights. This income does not appear in tax records.

The methods by which we address mis-recording or imputation depend on the variable in question. In some cases there are alternative data sources which provide better estimates of the variables we need; thus the SPI provides for taxpayers, information on investment income derived from tax returns. We impute income flows using these sources, and then scale the resulting totals to be consistent with the national accounting aggregates; we impute pension wealth from the WAS as a basis for allocating pension fund investment income net of service charges in the same way. In other cases either we do not have any information beyond what is shown in the LCFS/ETB, or alternative sources suggest a similar degree of mis-recording. In such cases there is little we can do beyond scaling the LCFS aggregates to

match the national accounts.

5.1 Imputation methods

We experimented with modelling the net operating surplus, interest and dividend income, taxes and pension income using two different methods of imputation. The first uses Heckman’s approach to model jointly selection and the variable of interest. This is adopted by Flower (2017) to model rental payments and thus imputed rent. We use this approach in order to replicate her procedure which is used to generate the estimates of imputed rent shown in the national accounts. Heckman’s method however, relies on an assumption of normality or, if the logarithm of the variable of interest is modelled, log-normality. A wide body of evidence suggests, however, that the distribution of variables such as income or wealth is more dispersed than log normality would imply and experimentation suggested that Heckman’s model did not allow us to model the distribution of interest and dividend income, or pension wealth (needed to allocate pension fund investment income net of service charges) very satisfactorily. However we did use Heckman’s method to model both imputed rent and rent accruing to households from buy-to-let properties.

To impute interest and dividend receipts and investment income accruing to pension funds we adopt a more flexible approach using an ordered probit model. We categorize our observations and then use an ordered probit model to represent the distribution. This is completely flexible in its functional form because, although the ordered probit model assumes that there is an underlying latent variable which is normally distributed, the cut points are selected to accommodate the actual distribution of the variable in question. There is a loss of information because we do not have any representation of the distribution of observations within any category. But this is more than offset by the flexibility that the approach offers in modelling the distribution as a whole. When we finally produce estimates of household income we need to provide figures conditional on an imputed observation being in any particular category. We do this by means of interpolation; for the upper-most category we assume that the observations are pareto-distributed beyond the upper-most threshold identified in the ordered probit model.

With either method we need to address the issue of covariances between the stochastic terms used to impute different aggregates. Neglect of these during imputation would result in a distribution of income more equal than would be found if, as is likely, the covariances are positive. This is covered in [section 5.5](#).

We use the ordered probit model as follows. We allocate each observation of the variable in question to a number of categories, using thirty-three for pension wealth and ninety for dividends and interest

income. With N categories we allocate y_{ij} to category $K_{n,j}$ if it lies above the upper threshold for category $K_{n-1,j}$ denoted by $k_{n-1,j}$, and but is less than or equal to the upper threshold for category $K_{n,j}$ denoted by $k_{n,j}$, in other words if $k_{n-1,j} < y_{ij} \leq k_{n,j}$. k_0 can take any negative number since negative holdings of pension wealth cannot be observed and the treatment of the upper boundary ensures that with $k_1 = 0$ zero holdings are distinguished from strictly positive holdings. The ordered probit model assumes that, underlying each observation there is a latent variable

$$z_{ij} = X_{ij}\delta_{ij} + \nu_{ij}; \quad \nu_{ij} \sim N(0, 1)$$

and a set of thresholds $k_{1,j}^*$ to $k_{N-1,j}^*$ such that, if

$$k_{n-1,j}^* < z_{ij} \leq k_{n,j}^* \text{ then } y_{ij} \in K_{n,j}$$

This model can be estimated by means of maximum likelihood using the *oprobit* command in Stata. When we come to simulate our model in order to impute data into the LCFS, we take the value of the latent variable generated by the sum of the deterministic and random components. The lower and upper bounds of the category in which we impute are defined by this. For those observations with a latent variable below the threshold for zero we impute a value of zero. For those whose value lies between two cut points we interpolate linearly between the cut points. As noted above, however, a common feature of both income and wealth distributions (see for example Jenkins (2017)) is that they show a large number of extreme values. Even when using the ordered probit model, it is necessary to take some view about the distribution of observations above the upper-most threshold.

We assume that the data classified to the upper-most category are distributed with a Pareto distribution of the first type. The Pareto type-1 distribution for observations $x_i > x_m$ has the form

$$1 - F(x) = (x_m/x)^\alpha \text{ with } \alpha > 0$$

and the expected value conditional on $x > x_m$ is $x_m\alpha/(\alpha - 1)$ with $\alpha > 1$ but infinite otherwise.

5.2 Imputation of pension wealth

In order to impute pension and insurance wealth, we break the WAS sample into four categories: i) households with heads aged under sixty-five who report employment income or self-employment income, ii) households with heads aged under sixty-five who do not report any employment or self-employment income, iii) households with heads aged sixty-five or more who report receiving private pension income and iv) households with heads aged sixty-five or more who do not report receiving any private pension

income. All of the households in category iii) report owning pension assets while those in the other three categories may not own any pension/insurance assets. Households in category iv) may own pension rights but not be drawing any pension from them. The Wealth and Assets Survey provides adequate data on the components of income required for our model only from Wave 3 (2010-2012) onwards; we do not therefore attempt to estimate models for waves 1 and 2. The parameters estimated for wave 3 are used for years before 2010, but the greater the distance from the start of the survey, the greater must be the uncertainty surrounding the results.

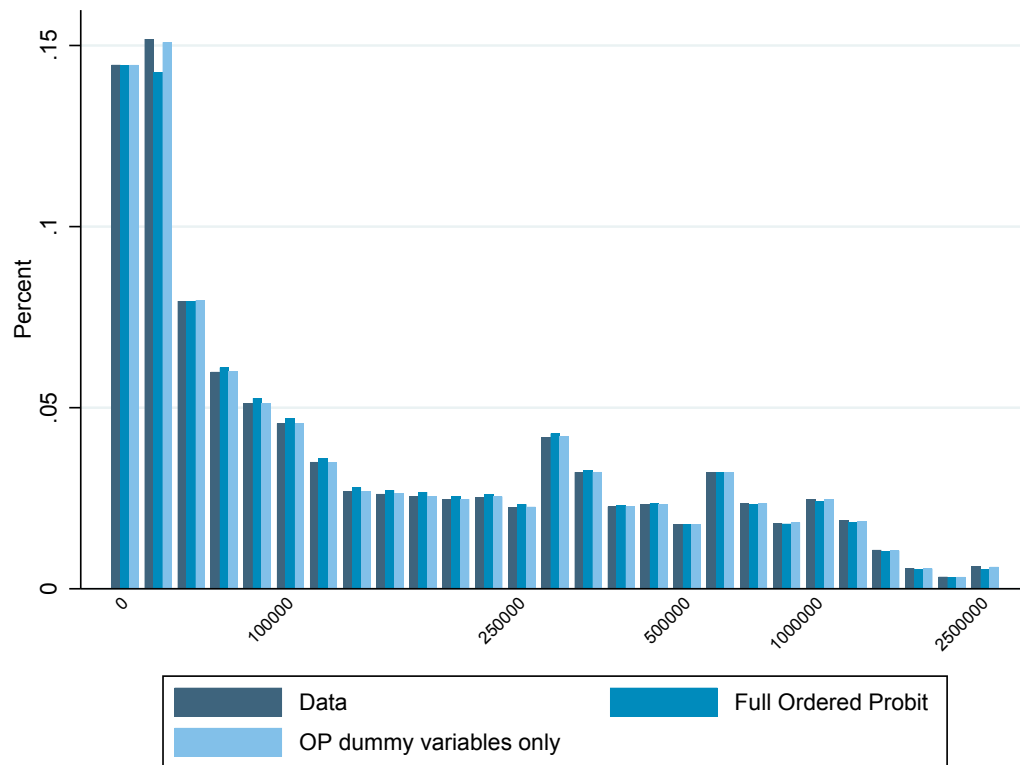
In figure 1 we show the distribution of the value of pension rights generated from the fourth wave of WAS. The ordered probit model mis-records the proportion with non-zero, but small pension pots but is otherwise close to the data. Our final model shows the ordered probit model estimated only with dummy variables for marital status and housing tenure. Here the fit of the distribution to the actual data is very good (as it is if the distribution is modelled without the use of any covariates). This suggests that the deviation of the expected distributional shape from the data arises from the use of continuous variables (age, age² and log employment/self employment income) as covariates. Parameter estimates for waves 3, 4, 5 and 6 are shown in tables 17 - 20, respectively in appendix A. As described in section 5.1, we fit a Pareto distribution to the observations above our categorical upper limit of £2m.

Scaling is also needed in handling pensions. Thus employer contributions are scaled from wages data; household contributions are scaled from the reported amounts and benefits received are similarly scaled from the reported amounts. One consequence of this treatment is that, while occupational pensions as reported in the LCFS include receipts from both funded and unfunded schemes, the pensions adjustment has to be calculated only on the basis of funded pension receipts. We scale all private pensions to this total taken from variable L8R3 in table 4.1.4 in the income/expenditure account of financial corporations.

5.3 Imputation of interest and dividend income

The two major categories of income from financial assets we observe are interest and dividends. We also need to address the investment income attributed to insurance policy holders, which we allocate in proportion to expenditure on insurance. The magnitude of the problem faced in developing satisfactory stochastic estimates of household interest and dividend income can be seen from table 5. This shows estimates of total investment income generated from four widely-used micro sources, the LCFS, the Wealth and Assets Survey (WAS, Office for National Statistics Social Survey Division (2020)), the data set Households Below Average Income (HBAI) and finally the SPI.

Figure 1: The Distribution of Pension Rights simulated for 2013 using an Ordered Probit Model applied to WAS Data



Source: Authors' calculations from WAS data.

The SPI is compiled as a random sample of tax records. Since this covers only people who pay income tax, it is not comprehensive but it provides a much better picture of high incomes than a survey source might be expected to. The tax threshold in 2013/4 for people aged under sixty-five was £9,440; the thresholds for older people were slightly higher. People who are only basic rate tax payers, those with incomes lower than £41,550 in 2012/3 are not required to complete tax returns. While they have tax records, tax on investment income is collected at source and is not linked to those records. For basic rate taxpayers therefore, the SPI is in turn enhanced by using survey data from the HBAI so as to provide estimates. While the SPI should be the most comprehensive source, it omits property income received through tax-free Individual Savings Accounts (ISAs) since interest and dividends received by holdings in these does not have to be declared.

It is clear that the LCFS under-records both dividend and interest income, while the coverage of WAS and HBAI is better but still well below the totals shown in the SPI and national accounts. In order to impute both dividend and interest income for households we use data from the Survey of Personal Incomes (SPI), splitting the sample into men and women. Given the origins of the SPI we look only at taxpayers. For people who do not pay tax we retain the figures shown in the LCFS.

Table 5 shows the estimates of interest and dividend receipts from the different sources, while table 6 shows the proportions of individuals reporting zero interest and dividend income in the LCFS and SPI. It is clear that, while many more people receive interest income than report it in the LCFS, the proportion reporting dividend income is similar to that found in tax records.

Table 5: Total Investment Income by Data Source (£m)

	LCFS			WAS		SPI			National Accounts		
	interest	dividends	total	total	interest	dividends	total	interest	dividends	total	
2005	9,390	9,493	18,883		18,629	33,293	51,922	42,282	34,538	76,820	
2006	12,433	8,412	20,846		20,232	37,189	57,420	47,569	38,450	86,019	
2007	12,309	10,299	22,608	21,952	26,863	40,436	67,300	58,919	42,744	101,663	
2008	12,026	7,693	19,719					54,443	41,762	96,205	
2009	9,264	6,768	16,032	17,957	15,385	41,968	57,353	18,503	47,618	66,101	
2010	7,579	6,343	13,922		12,761	35,753	48,514	18,530	33,933	51,923	
2011	6,905	5,897	12,802	23,867	13,226	41,003	54,229	21,140	39,783	60,923	
2012	9,453	9,002	18,455		13,217	44,410	57,627	20,632	42,812	63,444	
2013	8,461	12,396	20,858	31,974	13,326	53,302	66,629	18,353	52,148	70,501	
2014	7,226	8,334	15,560		12,071	58,937	71,008	24,214	57,872	82,086	
2015	6,799	7,994	14,793	28,002	11,147	78,705	89,852	23,959	80,659	104,618	
2016	7,680	10,046	17,726		11,471	58,034	69,505	22,825	62,617	85,442	
2017	8,312	11,121	19,433	24,820	11,316	63,538	74,854	18,397	69,558	87,955	
2018	7,173	18,352	25,526		8,146	61,038	69,184	25,266	78,952	104,218	

Source: Authors' calculations from the LCFS, WAS, and SPI.

Note: Fiscal year data except from WAS. Total investment income for all ages and all aged taxpayers only by data source. Identifying taxpayers in the WAS is not possible.

Table 6: Proportion receiving zero investment income in the LCFS and SPI

	LCFS			SPI		
	Total taxpayers	Report zero interest	Report zero dividends	Total taxpayers	Report zero interest	Report zero dividends
2006	39,247	0.71	0.89	31,566	0.17	0.74
2007	39,847	0.71	0.90	32,172	0.19	0.74
2008	38,869	0.71	0.91			
2009	37,863	0.74	0.92	30,276	0.20	0.85
2010	37,772	0.76	0.92	30,988	0.31	0.85
2011	37,149	0.75	0.92	30,462	0.26	0.85
2012	36,958	0.75	0.93	30,263	0.18	0.85
2013	37,860	0.75	0.92	30,108	0.21	0.85
2014	38,015	0.74	0.92	30,368	0.20	0.85
2015	38,075	0.74	0.92	30,788	0.21	0.84
2016	38,476	0.74	0.92	30,966	0.24	0.85
2017	34,687	0.75	0.90	30,871	0.25	0.85
2018	17,218	0.69	0.77	31,357	0.66	0.86

Source: Authors' calculations from the LCFS and SPI.

Note: Proportion of taxpayers aged 18 years and over who report receiving zero investment income in the LCFS and SPI. SPI data for 2008 are not currently available.

We use ordered probit models to allocate separately, interest income, and dividend income to categories as shown in figures 2 and 3. The explanatory variables we use are age bands (6), the log of labour income, and regional dummies.⁹ We estimate the model separately for men and women, and separately for each year before translating the coefficients into the LCFS for imputation. The estimated coefficients from the SPI are shown in Appendix A3.¹⁰ It is notable that the coefficients on log of income excluding investment income¹¹ (*linc*) are universally negative for women and at best only weakly positive for men. This suggests that investment income is a substitute for rather than a complement to labour income. One explanation may be that a high proportion of people who are in practice self-employed have chosen to incorporate and receive their income as dividends rather than as self-employment income appearing in *linc*. Old people are more likely to receive investment income, and receipts are also likely to be higher in regions such as London (region 7) and the South-East (region 8) relative to the North-East (omitted).

A small number of people have no reported income excluding their investment income. For these, again distinguishing men and women, we estimate models with only age and regional dummies present. These show similar age and regional patterns similar to the bulk of the population with reported non-

⁹The SPI is limited in having few characteristics of individuals in it.

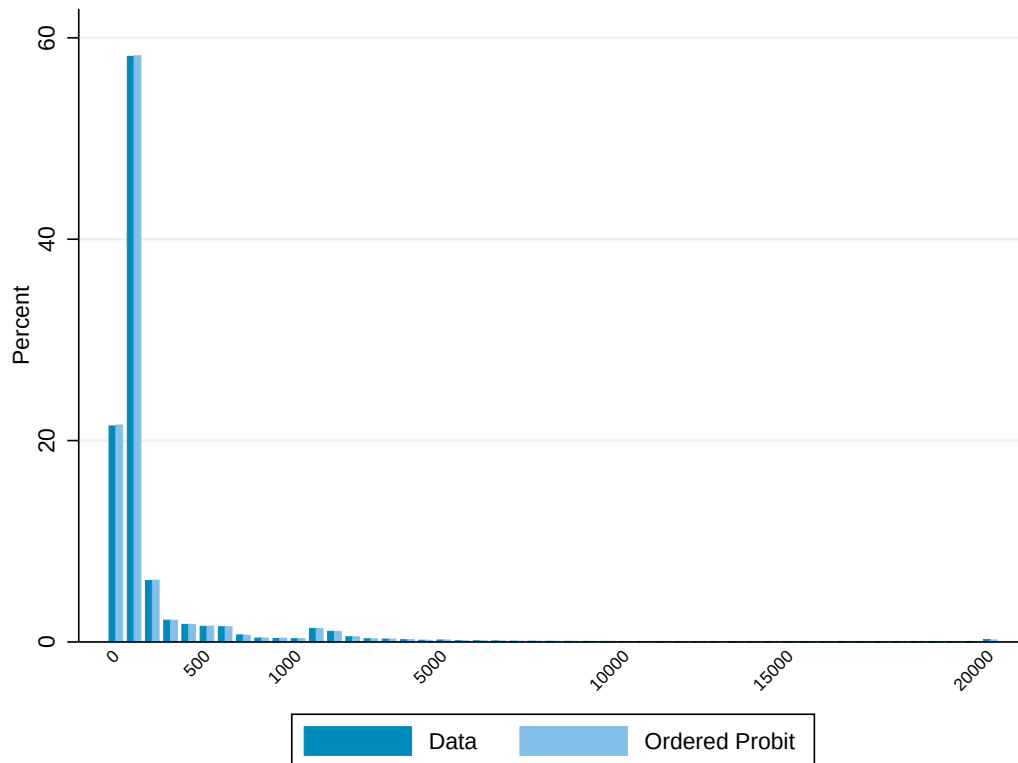
¹⁰The cut points are not shown, but are available on request.

¹¹This includes pensions and self-employment income.

investment income.

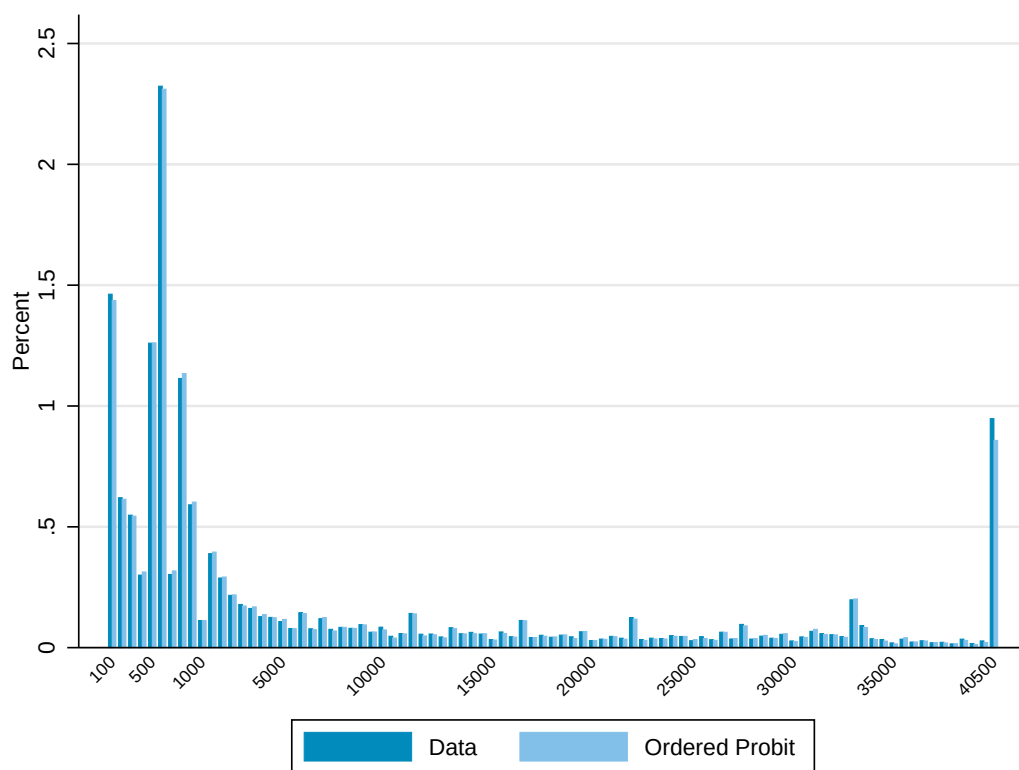
Given that between 30-50% of total investment income receipts accrue to people receiving more than £40,500 p.a., we use the Pareto distribution as discussed above to model the large number of extreme values of investment income. We estimate the Pareto parameter α from the SPI as a function of time and sex using *paretofit* in Stata (Jenkins and van Kearn, 2015), with the parameters shown in table 16 in the appendix. Figure 2 shows the distribution of the value of actual interest income for 2013 and the predictions using the ordered probit model in the SPI while figure 3 shows the analogous results for dividends. The figures show that the fits of the distribution to the actual data is very good.

Figure 2: The Distribution of Interest Income in the 2013 SPI and the Distribution Fitted by the Ordered Probit Models (Unweighted)



Source: Authors' calculations from SPI data.

Figure 3: The Distribution of Dividend Income in the 2013 SPI and the Distribution fitted by the Ordered Probit Models (Unweighted)



Source: Authors' calculations from SPI data. Note that a high proportion receive zero dividends, and this lowest category is omitted from the figure.

5.4 Imputation of housing rent

Owner-occupiers' imputed rental is an estimate of the housing services consumed by households who are not actually renting their house. Imputed rental should therefore represent the economic value to home owners of their dwellings, equivalent to if they were to rent out their properties. By definition, however, a homeowner does not receive payment on their property, and so the payment must be “imputed”. In the national accounts both imputed rent and rent accruing to households from buy-to-let properties are classified as operating surplus, and the aggregate against which we compare them is measured net of depreciation on housing. In order to estimate imputed rent at the household level we use an approach similar to that used by ONS for imputing rent for CPIH inflation (Flower, 2017). We use a Heckman selection model with the following covariates: household net income, housing type, number of rooms, socioeconomic group, council tax bands, region, and year dummies. The model is estimated from rent paid recorded in the LCFS, and is then used to impute rent on owner-occupied housing. There is no obvious exclusion restriction and the model is therefore identified by the assumption of normality. Table 7 shows the relevant aggregates computed from the LCFS. The estimated equations are presented in appendix A.3. The table shows that our model estimates for rent paid closely match the figures for rent paid in the LCFS, while our figures for imputed rent are slightly below the imputed rent produced by the ONS in the National Accounts.

5.5 Estimation of correlations between stochastic imputations

We have so far set out the stochastic models which we use to impute estimates of pension wealth, dividend and interest receipts and debt interest to the Living Costs and Food Survey. These models are estimated using different and unrelated sources, and the essence of stochastic imputation is that there is a random element to each imputed variable. This means that, while we cannot give any meaning to any particular record, we can hope to improve our understanding of the distribution of income. But the implications for household income of these imputations will depend very much on the extent to which the random components of the imputed variables are correlated. If we were to assume them not to be correlated while, in fact there are positive correlations on the receipts side, we would probably tend to understate the underlying inequality in incomes. Since, however, we have tried to use the most suitable source for each imputed item, we are unable to estimate our models jointly so as to provide estimates of the correlations simultaneously with the model parameters. Furthermore, in our categorical models we have divided the categories finely. This has merits in that it improves our representation of the underlying distributions. But, at the same time, the high number of parameters makes any joint

Table 7: Rent paid and rent received variables from the LCFS and the National Accounts (£m)

	LCFS		National Accounts	Model estimates	
	Rent paid	Rent received	Imputed rent	Imputed rent	Fitted rent paid
2006	37,444	3,904	137,863	119,558	36,290
2007	40,930	4,543	144,194	122,082	40,357
2008	43,641	3,916	148,083	125,520	42,469
2009	49,004	4,960	127,609	138,045	50,518
2010	53,173	5,039	126,922	142,262	52,634
2011	58,821	5,557	132,437	137,366	58,296
2012	62,198	7,666	139,500	153,382	61,277
2013	65,242	7,161	141,839	154,326	65,434
2014	64,319	8,182	156,041	157,423	63,189
2015	66,548	9,010	165,958	160,202	64,444
2016	66,311	9,967	175,387	170,856	63,992
2017	70,719	13,591	173,837	179,211	68,510
2018	71,743	11,342	183,264	185,261	69,149

Note: Rent paid and rent received are from the LCFS. Rent paid is $b010 + b020 + p220urmt$, where $b010$ (rent/rates - last net payment), $b020$ (net rent - service charge deducted), $p220urmt$ (rent, rate, council tax, rebate or allowance - renters only, this is based on advice from Ross Bowen at ONS; Rent received is $b360$ (Rent amount from property received). All figures have been converted to annual totals. Imputed rent is from the 2018 *Blue book*. The last two columns are the authors' calculations from imputation in the LCFS. Figures are weighted and annual. National accounts data are household sector gross operating surplus.

estimation difficult.

The only source which allows any form of joint estimation is the Wealth and Assets Survey. As well as the data on pension rights for which it is the only source, it provides information on individuals' holdings of shares and interest-bearing financial assets; for housing it has information on the current estimated value of home-owners' property. Since the financial asset data are provided on an individual basis they can be related to the individual data in the SPI. We model jointly, the distribution of pension rights by households with the value of financial assets (interest bearing assets, and the value of shares etc.) and the value of debt incurred by the household. Making the assumption that receipts of dividends and interest are proportional to financial assets, this then allows us to estimate the correlation matrix for the random components to the latent variables behind household pension rights and dividend and interest receipts of the household reference person and spouse.

The model we estimate specifies eight categories for pension assets, nine for both types of financial assets, and nine for housing assets. The equations we use correspond to those described above, with the exception that, since we can estimate the model only for households where there is a household reference person and spouse/partner, we do not include dummies representing marital status.

From a given wave of the WAS we use i to index the observations, and j to index the subcategory of WAS data (employment income and age of household reference person <65, no employment income and age of household reference person <65, and age ¹² of household reference person 65+). k indexes each of the six variables of interest (interest bearing assets of the reference person, interest bearing assets of the spouse, shares of the reference person, shares of the spouse, the pension rights of the household, and household debt including mortgage debt). We assume that there is a latent variable defined as

$$z_{ijk} = X_{ijk}\delta_{ijk} + \nu_{ijk}$$

with thresholds $k_{1,j,k}^*$ to $k_{N-1,j,k}^*$ and categories $K_{n,j,k}$ such that, if $k_{n-1,j}^* < z_{ij} \leq k_{n,j}^*$ then $y_{ijk} \in K_{n,j,k}$

Here

$$\begin{bmatrix} \nu_{ij1} \\ \nu_{ij2} \\ \nu_{ij3} \\ \nu_{ij4} \\ \nu_{ij5} \\ \nu_{ij6} \end{bmatrix} \sim N \left(0, \begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \rho_{14} & \rho_{15} & \rho_{16} \\ \rho_{12} & 1 & \rho_{23} & \rho_{24} & \rho_{25} & \rho_{26} \\ \rho_{13} & \rho_{23} & 1 & \rho_{34} & \rho_{35} & \rho_{36} \\ \rho_{14} & \rho_{24} & \rho_{34} & 1 & \rho_{45} & \rho_{46} \\ \rho_{15} & \rho_{25} & \rho_{35} & \rho_{45} & 1 & \rho_{56} \\ \rho_{16} & \rho_{26} & \rho_{36} & \rho_{46} & \rho_{56} & 1 \end{bmatrix} \right)$$

where the correlations shown by the $\rho_{m,n}$ are the focus of our interest. For example, ρ_{12} is the correlation between the latent variables driving interest income for each of the two adults; ρ_{13} is the correlation between interest income and dividend income of the first adult; ρ_{14} is the correlation between interest income of the first adult and dividend income of the second adult; ρ_{15} is the correlation between interest income of the first adult and pension rights; ρ_{16} is the correlation between interest income of the first adult and debt. For each of the six waves it is necessary to estimate a multivariate ordered probit model so as to extract the correlations.

Except for the correlation between assets held by the household reference person and their partner, the correlations are generally small. When simulating these variables we use the arithmetic averages of the three sets of correlations for each wave. These are provided in table 8. It is noteworthy that in wave 6 the correlations between pension wealth and debt (variables 5 and 6) and the income variables (variables 1 to 4) are generally lower than in the earlier waves.

5.6 Scaling labour income

In survey data we observe variables for wage income, self-employment income, and replacement income (i.e social insurance income, including pensions and unemployment benefits). The imputation of pension income is discussed in the following section. Both wages and salaries, and self-employment

¹²For this category we tried to distinguish recipients of pension income from others, but the relevant models did not solve in all waves.

Table 8: Estimates of the Correlation Matrix

Correlation	Wave					
	1	2	3	4	5	6
ρ_{12}	0.67	0.66	0.66	0.69	0.71	0.67
ρ_{13}	0.43	0.41	0.43	0.43	0.47	0.39
ρ_{14}	0.32	0.33	0.36	0.35	0.38	0.33
ρ_{15}	0.17	0.19	0.20	0.20	0.29	0.01
ρ_{16}	0.03	0.00	-0.04	0.00	0.01	0.06
ρ_{23}	0.38	0.37	0.37	0.36	0.43	0.35
ρ_{24}	0.44	0.43	0.43	0.43	0.44	0.42
ρ_{25}	0.21	0.21	0.25	0.22	0.32	0.01
ρ_{26}	0.02	-0.01	0.00	0.00	0.03	0.03
ρ_{34}	0.76	0.77	0.79	0.79	0.84	0.75
ρ_{35}	0.33	0.35	0.37	0.34	0.47	0.01
ρ_{36}	0.13	0.11	0.07	0.08	0.13	0.01
ρ_{45}	0.35	0.34	0.36	0.34	0.48	-0.02
ρ_{46}	0.13	0.04	0.04	0.07	0.12	0.02
ρ_{56}	0.09	0.07	0.07	0.03	0.17	0.11

Key

1. Interest income of reference person.
2. Interest income of partner.
3. Dividend income of reference person.
4. Dividend income of partner.
5. Value of household pension wealth.
6. Value of household debt.

Source: Authors' calculations from ONS data

income are scaled directly from the data collected in the LCFS. In the national accounts income of self-employed people (mixed income) is distinguished from the distributed incomes of quasi-corporations. The latter are unincorporated partnerships and the income from these is, like mixed income, taxed as self-employment income. Accordingly we aggregate these two categories of national accounting income and refer to the total as self-employment income. Table 9 shows the sum of self-employment income in both the LCFS and SPI. There is no obvious gain in trying to use SPI rather than LCFS, so we simply scale the data in the LCFS to be consistent with the national accounts. Social security benefits and social assistance benefits are also scaled using the variables in table 4. Secondary transfers paid are similarly scaled as shown in Table 4.

5.7 Scaling mortgage interest payments

We also have to adjust interest payments. Table 10 shows the sum of mortgage interest payments from both the LCFS and HBAI, together with all interest payments as shown in the national accounts. While

Table 9: Self-employed income in the LCFS and SPI (£m)

	LCFS	SPI	National Accounts
2005	63,255	71,673	96,235
2006	72,917	74,542	98,616
2007	74,172	75,083	99,849
2008	75,375		109,410
2009	73,718	70,474	111,101
2010	91,042	74,146	118,975
2011	62,502	77,237	119,819
2012	87,883	77,611	123,710
2013	83,649	83,401	133,193
2014	89,530	83,663	137,457
2015	88,847	87,953	146,399
2016	104,264	92,083	152,235
2017	111,066	96,358	158,846
2018	140,603	90,839	167,621

Source: Authors' calculations from ONS data. National accounts figures include income from quasi-corporations and are net of depreciation.

both surveys under-record payments in the early part of the period, they over-state payments in most years after the financial crisis. This over-recording is all the more pronounced because the national accounts figure covers all interest payments and not simply those associated with mortgages. There is no reason to think that HBAI should be preferred to LCFS, and we simply scale total interest payments shown in LCFS to the total in the national accounts.

5.8 Scaling tax payments

The LCFS records payments of current taxes on income and wealth and, as table 4 makes clear, these are materially under-recorded. In contrast employees' social contributions (National Insurance contributions collected on employment income) are, like employment income itself, only modestly under-recorded relative to the national accounts. If, as we suggest, taxable income is to be imputed to individuals, to replace reported data for investment income, then it is clearly necessary for tax liabilities to be recalculated on the basis of imputed gross investment income. It would not seem sensible to impute income receipts on the basis of another data source and then to reconcile tax paid simply by means of scaling.

Separately, although tax payments in aggregate are under-recorded, there are some households for which their tax payments grossly exceed their taxable income. This is obviously possible if people's tax payments refer to income of the previous year, while the income data themselves are current.

Table 10: Interest payments in the LCFS, HBAI and National Accounts (£m)

	LCFS	HBAI	National Accounts
2005	37,591	37,919	58,842
2006	44,395	41,722	66,648
2007	50,563	48,424	84,291
2008	47,951	45,146	78,586
2009	37,585	32,550	26,683
2010	36,661	31,471	22,071
2011	32,729	30,555	24,715
2012	31,941	30,479	24,081
2013	30,597	28,858	22,105
2014	29,632	29,043	28,345
2015	28,501	27,591	27,642
2016	31,006		27,510
2017	29,853		24,130
2018	31,824		31,134

Source: Authors' calculations from ONS data. *Note:* The LCFS and HBAI figures cover only mortgage and credit card interest. HBAI data for 2016-2018 are currently unavailable due to the absence of the relevant files in the SDS.

Nevertheless it raises difficulties which cannot be addressed by scaling; negative incomes have to be excluded from the calculation of a geometric mean.

The approach we therefore adopt is to recalculate tax liabilities on the basis of the final estimates of each individual's income, calculated after interest and dividend income has been imputed and after other income variables have been scaled to be consistent with the national accounts. In this calculation we include only those components of income which are taxable. The income data distinguish state retirement pensions from other cash benefits. We treat the former as taxable. The latter include some which are taxable and others which are tax-free. Since the large majority are tax-free, we treat all of these cash benefits as tax free. While we reflect tax relief on pension contributions, we have not modelled the withdrawal of this for people with very high incomes, or the effects of the life-time upper limit.

We model the attenuation of the personal allowance for people with high incomes, and also the different tax treatment of dividends from other forms of income. We are however unable to model a range of exemptions such as the tax-free nature of investment income for savings held in Individual Savings Accounts or the relief offered on charitable donations.

National income is now, by convention, measured gross of indirect taxes with relatively little attention paid to measures of national income at basic prices, a concept which corresponds more closely to income

at factor cost. The arguments for measuring at market prices are essentially that when people make spending decisions they do so on the basis of market prices; at the margin relative prices reflect relative values. In this paper we too, therefore work at market prices.

5.9 Other discrepancies

In the vintage of quarterly data that we have used, there are minor discrepancies in the financial year totals calculated from data which are not seasonally adjusted, while the corresponding calendar year totals do not show similar discrepancies. Thus the sum of sectoral gross disposable income is slightly different from the total shown for national gross disposable income. We address this by the introduction of a small category of residual income which we allocate in the same way as undistributed corporate income. We check our reconciliation by carrying it out for the calendar year, finding a discrepancy of no more than £5m between the components we identify in table 4 and the total for net national income computed by the ONS. We thus infer that our reconciliation correctly aggregates the components but that there are discrepancies in the quarterly accounts which cancel out when they data are aggregated to a calendar year.

6 Deflation

Aitken and Weale (2020) have shown that it is possible to design a coherent value/quantity/price system using democratically weighted Divisia deflators. The value counterpart is the geometric mean of income attributed to households, and deflation of the geometric mean by the democratic deflator gives a volume measure gives a measure of economic growth which is the average of the growth rates of real income attributed to each household. So far, however, the approach has been applied only to the measure of household income after housing costs available from the Households below Average Income data set. Ideally the approach should be applied with the whole of national income attributed to households. This means that the price index needs to be extended to cover all consumption and not just directly identified household consumption. We explain here how this can be done, and present the appropriate deflator.

The democratic price index presented here is a Törnqvist approximation to a Divisia index. It is computed from a range of price indices which are either publicly available, or have been derived from publicly available data. The weights used to produce the plutocratic consumption price index are similarly derived from publicly available data sources, while the weights used to produce the democratic price index draw also on the Earnings, Taxes and Benefits (ETB) data set which the ONS has kindly made available to us through its Secure Research Service. We have combined this with a version of

the LCFS data set in which the ONS have aligned each of the eighty-seven components of household consumption to be consistent with the national accounts. We have produced our index on a financial year basis, corresponding to current survey sources.

ONS has already computed a democratic price index for household consumption.¹³ To produce the aggregate consumption deflator it is necessary to combine this with the government deflators described below, and also to make some allowance for consumption by NPISH.

ONS data on general government final consumption expenditure provide quarterly figures for public consumption, distinguishing health, education and five other categories.¹⁴ Health and education take up just over half of the total. Data are provided in current prices and as chain-value measures, and we use the ratios of these to provide deflators for each of these seven categories of consumption.

As explained earlier, the ETB data set allocates education, health and some other components of public consumption to households. We scale the ETB figures for education to match the aggregate totals shown in the ONS tables and allocate health spending on a uniform *per capita* basis. thus we allocate health spending like the remainder of public spending (unallocated spending Five categories of unallocated spending are identified in aggregate: social protection, military defence, justice and fire services, other central government and other local government) between households on the assumption that it is spread between households in proportion to their size as measured by the equivalence scale. This allocation of public consumption allows us to calculate expenditure shares for each household in ETB, distinguishing private consumption, scaled to the national accounts totals, and the three components of public consumption. For each household we then produce a consumption aggregate equal to private consumption (which is already scaled to match the aggregate data) plus education, health and other public consumption. The weighted arithmetic averages of the shares of each of these across all households provide us with weights needed to combine the growth in the prices of household consumption and each of the three measures of public consumption. For education we use the deflators derived directly, as described above, For other public consumption we weight the growth in each of the six deflators (for health and the five other categories) of the component parts by the share of each of these six components in total unallocated spending in the previous financial year. It is striking that the weight of private consumption is about seventy-seven per cent in the plutocratic index, but only seventy-two per cent in the democratic index.

¹³Available from [Investigating the impact of different weighting methods on CPIH](#).

¹⁴[General government final consumption expenditure \(GGFCE\) sub-sectors quarterly data 2019 Quarter 3](#).

There remains the question of consumption by NPISH. Table 1.1.2 of the national accounts provides series on NPISH consumption in both current prices and volume terms, allowing us to derive the deflator. We have no information on the way in which the consumption of NPISH might be allocated across households. As with the allocation of NPISH income, we treat it in the same way as “other” public consumption, assuming that each household consumes in proportion to its effective size.

With these data we are able to produce a chain-linked index for aggregate consumption. We calculate, for each of the eight categories (private, NPISH and six public categories) of expenditure, shares in total expenditure for each financial year, and also the growth in the financial year average of each of the nine deflators compared with the previous year. The weighted sum of these gives the growth in the aggregate deflator from one financial year to the next. The democratic and plutocratic inflation rates, thus measured, are as shown in figure 4; the data themselves are in table 22 of appendix B. Over the eleven year period, 2006-2017, the average democratic inflation rate has been 0.09% p.a. higher than the plutocratic inflation rate.

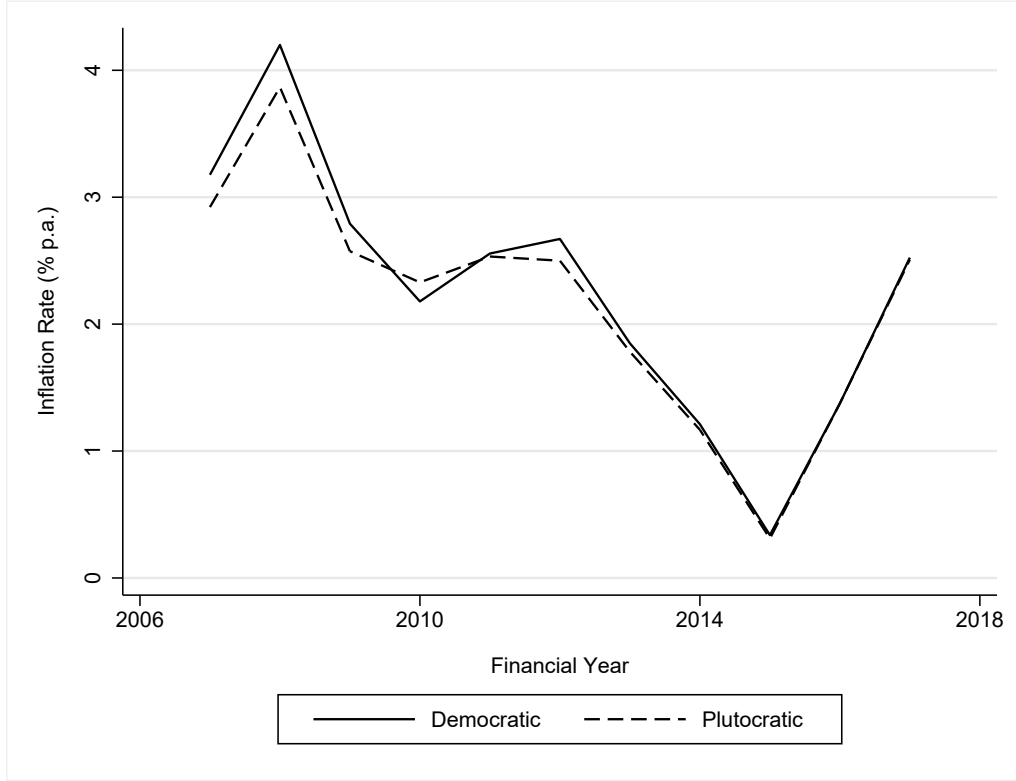
7 Estimates of measures of inequality and growth in democratic and plutocratic income

We are now in a position to bring the various strands together to produce our estimates of growth in democratic and plutocratic income per household, after adjusting for household size. We begin this section, however, by showing Gini coefficients for different measures of income. These are of interest in their own right, but also allow us to compare the estimates generated by our approach with official figures.¹⁵

We have modelled receipts of imputed rent, dividends, interest income and the investment earnings of pension funds on a stochastic basis. Our aim has been, as well as possible to represent the distribution shown in the alternative data sources which we have judged to be more robust than the LCFS. This means, however, that once we come to produce estimates of the income of each household in the LCFS, we can do this only on a stochastic basis. Different stochastic simulations will generate different receipts for each household and thus different values for the geometric mean of any definition of income. Since our measures of arithmetic income are adjusted for household size, even the arithmetic mean will vary materially from one simulation to the next. Thus, rather than produce a single figure for the levels and growth rates of plutocratic and democratic nominal income per household, we use the mean value of each variable of interest across 100 simulations as providing our single “best estimate” for each variable.

¹⁵ONS (2018). [The effects of taxes and benefits on household income, disposable income estimate.](#)

Figure 4: Democratic and Plutocratic Measures of the Growth in the Consumption Deflator

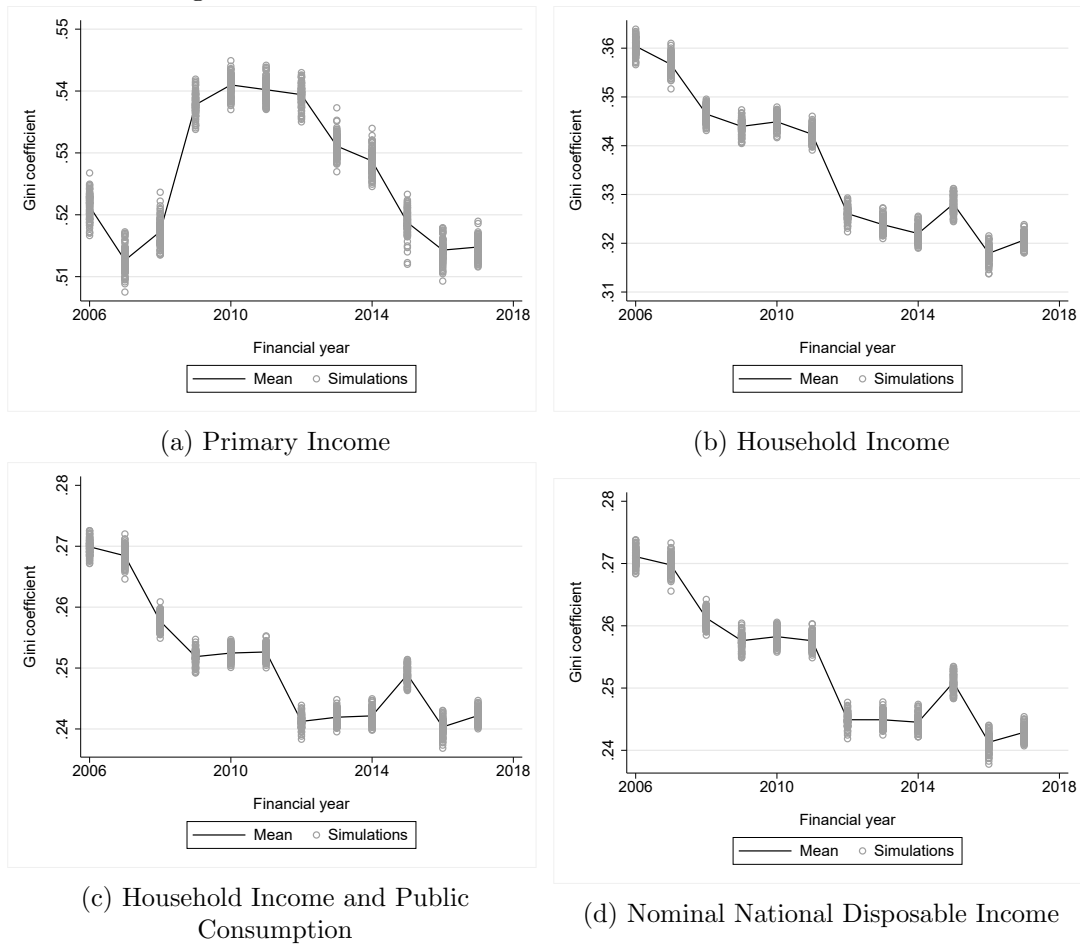


Source: Authors' calculations from ONS CPI data.

In figure 5 we show the Gini coefficients for different definitions of income. Figure 5a shows the Gini coefficient for our measure of primary income. This corresponds approximately to the ONS concept of original income in their analysis of inequality. We, however, include imputed rent on owner-occupied housing which the ONS excludes, and we also include our attribution of undistributed corporate income. Figure 5b shows the results after the redistribution arising from the tax and benefit system. Here it is important to remember that we treat occupational pension contributions as saving out of income and occupational pension receipts as dis-saving. The ONS treats the former as a deduction from income and the latter as a source of income. But the ONS measure also treats the indirect taxes collected on household consumption as deductions from income while, because we are working at market prices, these are seen on the same basis as factor income accruing to the government.

In figure 5c we show the Gini coefficient after crediting to households the consumption benefits they receive as a result of public consumption and consumption by non-profit institutions serving households. The allocation of this consumption, with health spending treated on a insurance basis rather than on the basis of treatment actually received and with all other spending except education allocated equally, means that this is strongly equalising. As we discussed, however, the government was dis-

Figure 5: Gini Coefficients for Various Definitions of Income

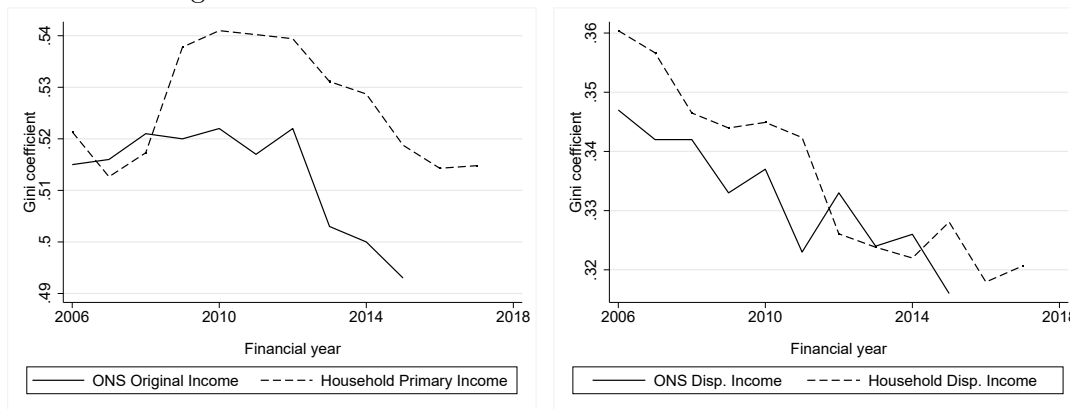


Source: Authors' calculations from ONS data.

saving throughout the period in question, and we allocate this dis-saving equally between government consumption and taxes. Making this adjustment leads to our allocation of national disposable income and, as is shown in figure 5d, the Gini coefficient is higher than we found in figure 5c. The difference is most acute when government dis-saving is high, as it was at the start of the current decade.

The most striking feature of these graphs is that they all show a pattern of declining inequality over the nine years in question. In figure 6a we compare our average Gini coefficient for primary income with that published by the ONS for original income, and in figure 6b we compare the Gini coefficient for household income with the ONS figure for disposable income. We can see that the pattern of declining inequality is also shown in the ONS data, giving reassurance that it is not the outcome of the various steps we have taken to arrive at our results.

Figure 6: Gini Coefficients for Various Definitions of Income



(a) Primary Income

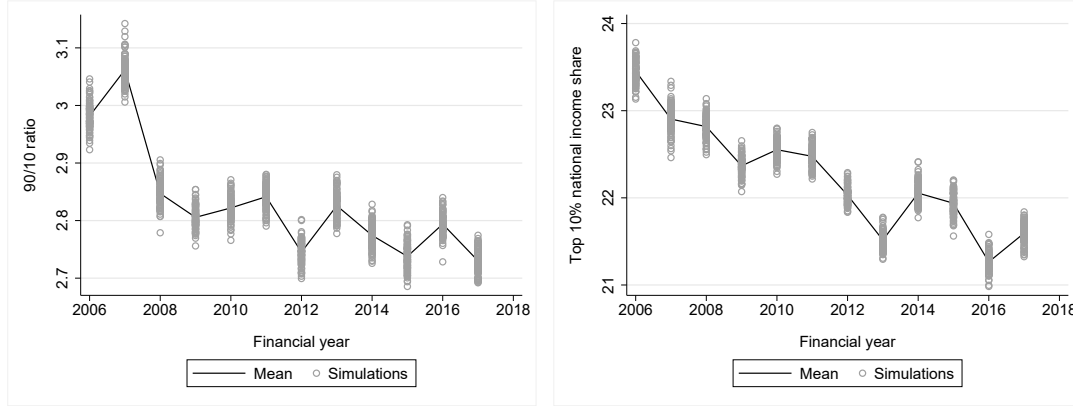
(b) Household Income

Source: Authors' calculations from ONS data. The bold lines show actual ONS data. The dotted lines are the authors' calculations.

We can also examine other measures of inequality such as the 90/10 income inequality ratio, or the amount of income accounted for by the top 10%. Figure 7 shows our estimates of both of these measures calculated national disposable equivalised income. The 90/10 ratio shows how many times larger income at the 90th percentile is compared with income at the 10th percentile. Like the Gini coefficients, these figures show that inequality in terms of national income has generally decline over the period since 2006. Since 2008 the 90/10 ratio has been fairly flat, and at around 2.8, is substantially below the value of around 4 shown in ONS statistics from the Household Finances Survey.

We now turn to the main focus of this work, the resulting estimates of income growth when measured on a plutocratic and a democratic basis, and these are shown in Figure 9. As with the Gini coefficients we show the growth rates generated by each of the simulations together with the average of the 100 growth rates. Both are of course measured per household on an equivalised basis and are calculated

Figure 7: Alternative Measures of Income Inequality



(a) 90/10 ratio

(b) Top 10% income share

Source: Authors' calculations from ONS data.

before deflation.

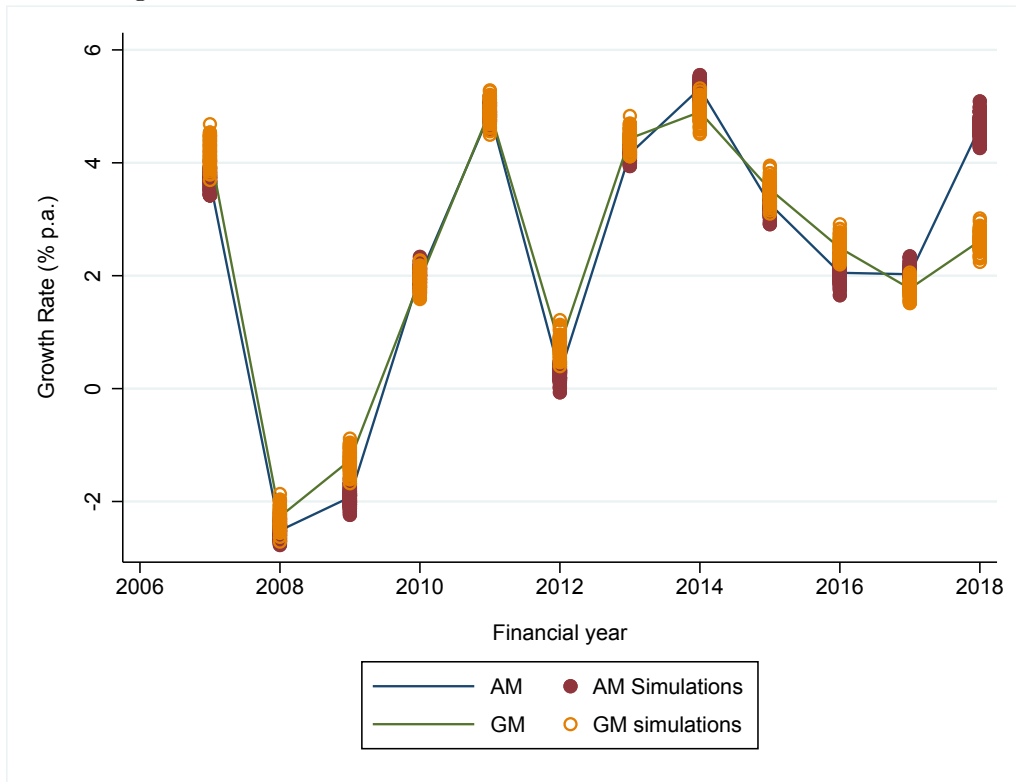
Given declining inequality we might expect democratic income to have grown faster than plutocratic income. Table 11, in which we present the average growth rates for both measures over the nine years confirms this. We also show the standard deviation of the average growth rates and the difference between them across each of the 100 simulations. We then use the deflators shown in table 22 of appendix B to convert these figures to real terms. There we recorded that the democratic price index rose by 0.09% p.a faster than the plutocratic price index. Working to one decimal place, however, the difference is exaggerated, with the former price index rising by 2.2% p.a and the latter by 2.3% p.a. Application of the deflators to the nominal figures confirms that over the period 2006 to 2017, while democratic national income per household grew by 0.02% p.a., plutocratic national income declined by 0.09% p.a. Thus half of the faster growth in nominal democratic income was absorbed by adverse price movements over the period examined. The year on year growth rates of real democratic and plutocratic national income per household are shown in figure 9.

Table 11: Average Growth Rates per Household: 2006-2017 (% p.a.)

	Plutocratic National Income	Democratic National Income	Difference
Avg. Nominal Growth 2006-2017	2.08%	2.28%	0.20%
Growth in Deflator	2.17%	2.26%	0.09%
Growth in Real Income	-0.09%	0.02%	0.11%

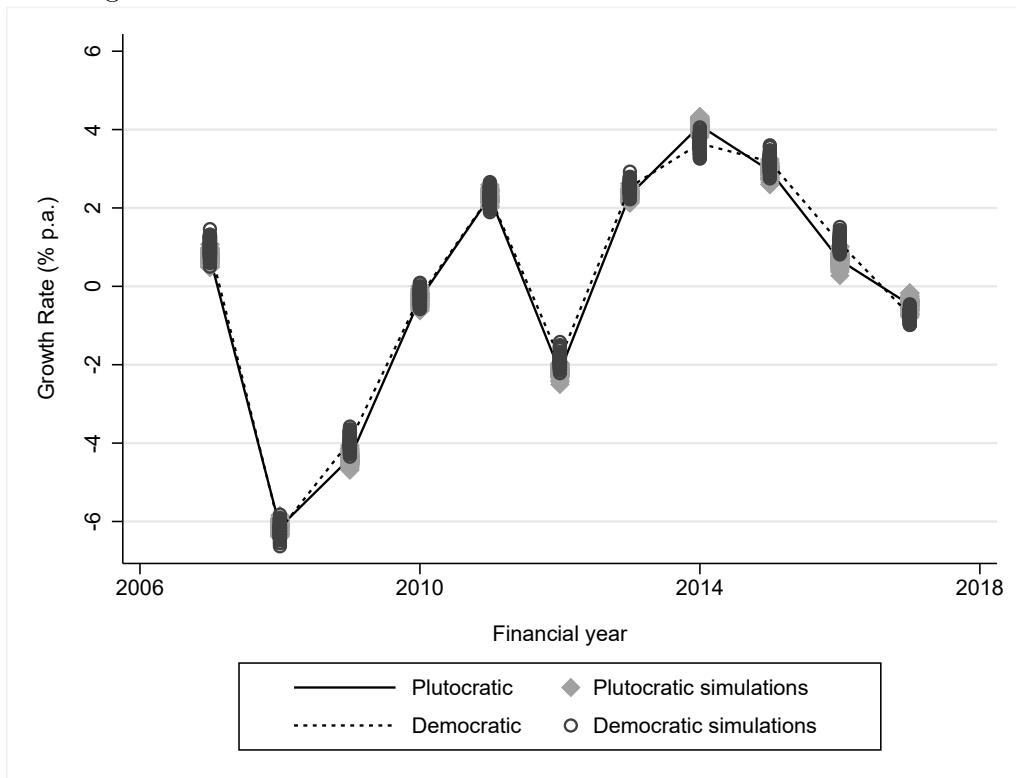
Source: Authors' calculations.

Figure 8: Plutocratic and Democratic Nominal National Income



Source: Authors' calculations from ONS data.

Figure 9: Plutocratic and Democratic Real National Income Growth



Source: Authors' calculations from ONS data.

8 Conclusion

We have demonstrated here the practicality of allocating the whole of national income to households. Taking as our starting point the Living Costs and Food Survey and the related Earnings, Taxes and Benefits data set, we use a combination of stochastic imputation and scaling to produce estimates of each household's receipts, either in money terms or imputed, of national disposable income. The use of stochastic methods implies that multiple simulations are needed; here we have used 100 but it would clearly be desirable to use more.

Over the period 2006-2017 we identify a materially faster growth rate in the geometric mean of nominal household income, after adjusting for household size, than is observed with the arithmetic mean. This is entirely consistent with the pattern of declining inequality (as measured by the Gini coefficient) shown in official data, and also in our income aggregates. However, the geometric mean needs to be deflated by a democratic price deflator, in contrast to the plutocratic deflator used for the arithmetic mean of household income. The democratic deflator shows slightly faster growth over the period, but, taking the two together we find that democratic real national income grew 0.1 per cent per annum faster than plutocratic real national income per household.

Our work sets out the practicality for statistics offices to provide measures of income growth which are robust to changes in inequality. In contrast to measures of well-being structured round a welfare function, they can be easily explained to the public. At the same time, as Aitken and Weale (2020) show, they can be given an interpretation in terms of a simple welfare function. Thus democratic measures of real income growth should form an important component of the evolving set of economic statistics.

9 References

- AITKEN, A. AND WEALE, M. (2020). ‘A Democratic Measure of Household Income Growth: Theory and Application to the United Kingdom’, *Economica*, vol. 87(347), pp. 589–610.
- ALVAREDO, F., ATKINSON, A.B., PIKETTY, T. AND SAEZ, E. (2013). ‘The Top 1 Percent in International and Historical Perspective’, *Journal of Economic Perspectives*, vol. 27(3), pp. 3–20.
- ATKINSON, A.B. AND HARRISON, A. (1978). *Distribution of Personal Wealth in Britain*, Cambridge: Cambridge University Press.
- ATKINSON, A.B., PIKETTY, T. AND SAEZ, E. (2011). ‘Top Incomes in the Long Run of History’, *Journal of Economic Literature*, vol. 49(1), pp. 3–71.
- BREWER, M., ETHERIDGE, B. AND O’DEA, C. (2017). ‘Why are Households that Report the Lowest Incomes So Well-off?’, *The Economic Journal*, vol. 127(605), pp. F1–F631.
- BURKHAUSER, R.V., HÉRALT, N., JENKINS, S.P. AND WILKINS, R. (2018). ‘Top incomes and inequality in the UK: reconciling estimates from household survey and tax return data’, *Oxford Economic Papers*, vol. 70(2), pp. 301–326.
- CROSSLEY, T.F. AND PENDAKUR, K. (2010). ‘The Common-Scaling Social Cost-of-Living Index’, *Journal of Business and Economic Statistics*, vol. 28(4), pp. 523–538.
- Digital Economy Act (2017). ‘Digital Economy Act. 2017’, .
- FESSEAU, M. AND MATTEONETTI, M. (2013). ‘Distributional Measures across Household Groups in a National Accounts Framework’, .
- Fixler, D. and Johnson, D. (2014). ‘Accounting for the Distribution of Income in the U.S. National Accounts’, in (D. Jorgenson, J. Landefeld and P. Schreyer, eds.), *Measuring Economic Sustainability and Progress*, pp. 213–244, Chicago: Chicago University Press.
- FLEURBAEY, M. (2009). ‘Beyond GDP: the Quest for Measures of Social Welfare’, *Journal of Economic Literature*, vol. 47(4), pp. 1029–75.
- FLOWER, T. (2017). ‘Methodology to calculate CPIH-consistent inflation rates for UK household groups’, .
- HM Revenue and Customs (2021). ‘Survey of Personal Incomes 2000-01 to 2018-19. Public Use Tape’, .

- JENKINS, S.P. (2017). ‘Pareto Models, Top Incomes and Recent Trends in UK Income Inequality’, *Economica*, vol. 84(334), pp. 261–289.
- JENKINS, S.P. (2021). ‘Top income adjustments and official statistics on income distribution: the case of the uk’, IZA Discussion Paper No. 14951 <https://ideas.repec.org/p/iza/izadps/dp14951.html>.
- Jenkins, S.P. and van Kearn, P. (2015). ‘Pareto t: Stata Module to t a Type 1 Pareto Distribution’, .
- JORGENSEN, D. AND SCHREYER, P. (2015). ‘Measuring Individual Economic Well-Being and Social Welfare within the Framework of the System of National Accounts.’, .
- KUZNETS, S. (1953). *Shares of Upper Income Groups in Income and Savings*, New York: National Bureau of Economic Research.
- Lewbel, A. (1997). ‘Consumer Demand and Household Equivalence’, in (H. Pesaran and P. Schmidt, eds.), *Handbook of Applied Economics*, Blackwell Publishers.
- MEADE, J.E. AND STONE, J.R.N. (1941). ‘The Construction of Tables of National Income, Expenditure, Savings and Investment’, *Economic Journal*, vol. 56(202/203), pp. 216–233.
- Office for National Statistics and Dept for Environment, F. and Affairs, R. (2021). ‘Living Costs and Food Survey, 2006-2020: Secure Access. [data collection]’, .
- OFFICE FOR NATIONAL STATISTICS SOCIAL SURVEY DIVISION (2020). ‘Wealth and Assets Survey, Waves 1-5 and Rounds 5-6, 2006-2018’, .
- PIKETTY, T. AND SAEZ, E. (2003a). ‘Income Inequality in the United States, 1913-1998’, *The Quarterly Journal of Economics*, vol. 118(1), pp. 1–39.
- PIKETTY, T. AND SAEZ, E. (2003b). ‘Income Inequality in the U.S., 1913-1998’, *Quarterly Journal of Economics*, vol. 118(1), pp. 1–41.
- PIKETTY, T., SAEZ, E. AND ZUCMAN, G. (2018). ‘Distributional National Accounts: Methods and Estimates for the United States’, *Quarterly Journal of Economics*, pp. 1–57.
- RUBIN, D.B. (1987). *Multiple Imputation for Non-response in Surveys*, New York: Wiley.
- SAEZ, E. AND ZUCMAN, G. (2016). ‘Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data’, *Quarterly Journal of Economics*, vol. 131(2), pp. 519–578.
- SCHAFER, J.L. (1999). ‘Multiple Imputation: a Primer’, *Statistical Methods in Medical Research*, vol. 8(1), pp. 3–15.

- SEFTON, J. AND WEALE, M.R. (2006). ‘The Concept of Income in a General Equilibrium’, *Review of Economic Studies*, vol. 73(1), pp. 219–249.
- STIGLITZ, J.E., SEN, A. AND FITOUSSI, J.P. (2009). ‘Report by the Commission on the Measurement of Economic Performance and Social Progress.’, INSEE, Paris.
- STONE, R. AND STONE, G. (1972). *National Income and Expenditure*, London: Bowes and Bowes.
- TONKIN, R. (2015). ‘The Distribution of Household Income, Consumption and Savings: an OECD Study’, .
- VAN BUUREN, S. (2007). ‘Multiple Imputation of Discrete and Continuous Data by Fully Conditional Specification’, *Statistical Methods in Medical Research*, vol. 16(3), pp. 219–242.
- Webber, D., Tonkin, R. and Shine, M. (2020). ‘Using tax data to better capture top incomes in official uk income inequality statistics’, in (*Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth* Chicago: NBER.

Appendices

A Models used to impute missing data

A.1 Interest and dividends

Table 12: probit coefficients from SPI (male)

	2005	2006	2007	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ointtot													
Age 25-34	0.021 (0.015)	-0.039* (0.012)	-0.029* (0.012)	-0.175*** (0.012)	-0.163*** (0.012)	-0.235*** (0.012)	-0.084*** (0.010)	0.025 (0.015)	0.072*** (0.014)	0.039** (0.014)	0.129*** (0.013)	0.132*** (0.015)	-0.149*** (0.016)
Age 35-44	0.824*** (0.014)	0.278*** (0.011)	0.567*** (0.011)	0.324*** (0.011)	0.500*** (0.012)	0.397*** (0.011)	-0.106*** (0.010)	0.364*** (0.015)	0.419*** (0.014)	0.488*** (0.014)	0.382*** (0.013)	0.340*** (0.015)	-0.134*** (0.016)
Age 45-54	0.899*** (0.014)	0.360*** (0.011)	0.663*** (0.011)	0.387*** (0.011)	0.578*** (0.012)	0.456*** (0.012)	-0.037*** (0.010)	0.425*** (0.015)	0.467*** (0.013)	0.530*** (0.014)	0.430*** (0.013)	0.385*** (0.015)	0.053*** (0.016)
Age 55-64	1.062*** (0.014)	0.581*** (0.011)	0.888*** (0.012)	0.594*** (0.011)	0.837*** (0.012)	0.710*** (0.012)	0.252*** (0.010)	0.694*** (0.015)	0.714*** (0.014)	0.762*** (0.014)	0.648*** (0.013)	0.626*** (0.015)	0.436*** (0.016)
Age 65-74	0.964*** (0.015)	0.590*** (0.012)	1.083*** (0.013)	0.834*** (0.012)	1.070*** (0.013)	0.866*** (0.013)	0.645*** (0.011)	1.114*** (0.015)	1.120*** (0.014)	1.170*** (0.014)	1.329*** (0.013)	1.322*** (0.015)	0.969*** (0.016)
Age 75+	1.148*** (0.017)	0.825*** (0.014)	1.284*** (0.015)	1.068*** (0.013)	1.185*** (0.015)	0.837*** (0.015)	0.620*** (0.014)	1.254*** (0.016)	1.216*** (0.015)	1.268*** (0.015)	1.466*** (0.014)	1.482*** (0.015)	0.998*** (0.017)
Ln Earned Income	0.230*** (0.003)	0.226*** (0.003)	0.205*** (0.003)	0.187*** (0.003)	0.132*** (0.004)	0.190*** (0.004)	0.247*** (0.004)	0.236*** (0.004)	0.238*** (0.004)	0.250*** (0.004)	0.327*** (0.004)	0.332*** (0.004)	0.351*** (0.004)
North-West	0.015 (0.015)	-0.023 (0.013)	-0.004 (0.013)	-0.055*** (0.013)	-0.023 (0.013)	-0.014 (0.013)	-0.018 (0.012)	-0.016 (0.013)	-0.005 (0.013)	-0.009 (0.013)	-0.012 (0.014)	-0.027 (0.015)	-0.031 (0.018)
Yorks/Humbs	0.021 (0.015)	-0.022 (0.014)	0.006 (0.014)	-0.044*** (0.013)	-0.020 (0.014)	-0.018 (0.013)	-0.020 (0.013)	-0.002 (0.014)	0.009 (0.014)	-0.013 (0.014)	0.007 (0.014)	0.005 (0.016)	0.069*** (0.018)
E. Midlands	0.056*** (0.016)	-0.005 (0.014)	0.037** (0.014)	-0.034* (0.013)	-0.000 (0.014)	-0.020 (0.013)	-0.012 (0.013)	-0.013 (0.014)	0.020 (0.014)	-0.009 (0.014)	0.001 (0.014)	-0.004 (0.016)	0.108*** (0.018)
W. Midlands	0.047** (0.015)	-0.010 (0.014)	0.035* (0.014)	-0.032* (0.013)	-0.004 (0.014)	0.019 (0.013)	0.002 (0.012)	0.008 (0.014)	0.004 (0.013)	-0.005 (0.014)	-0.006 (0.014)	-0.027 (0.015)	0.154*** (0.018)
E. England	0.029* (0.015)	-0.040** (0.013)	0.002 (0.013)	-0.090*** (0.013)	-0.043** (0.013)	-0.051*** (0.013)	-0.074*** (0.012)	-0.077*** (0.013)	-0.063*** (0.013)	-0.075*** (0.013)	-0.046*** (0.014)	-0.059*** (0.015)	0.133*** (0.017)
London	0.068*** (0.015)	-0.028* (0.013)	0.023 (0.013)	-0.072*** (0.013)	-0.009 (0.013)	-0.020 (0.013)	-0.103*** (0.012)	-0.071*** (0.013)	-0.064*** (0.013)	-0.111*** (0.013)	-0.085*** (0.013)	-0.110*** (0.014)	0.076*** (0.017)
South-East	0.050*** (0.014)	-0.027* (0.013)	0.040** (0.013)	-0.073*** (0.012)	-0.021 (0.013)	-0.036** (0.012)	-0.080*** (0.012)	-0.067*** (0.012)	-0.056*** (0.012)	-0.080*** (0.013)	-0.046*** (0.013)	-0.065*** (0.014)	0.186*** (0.017)
South-West	0.061*** (0.015)	0.015 (0.013)	0.062*** (0.013)	-0.027* (0.013)	0.011 (0.013)	-0.004 (0.013)	-0.036** (0.012)	-0.015 (0.013)	-0.018 (0.013)	-0.042** (0.013)	-0.018 (0.014)	-0.031* (0.015)	0.165*** (0.018)
Wales	0.025 (0.017)	-0.021 (0.015)	0.035* (0.015)	-0.056*** (0.015)	-0.023 (0.015)	-0.025 (0.015)	-0.033* (0.014)	-0.022 (0.015)	0.006 (0.015)	-0.024 (0.016)	-0.011 (0.016)	-0.037* (0.017)	0.080*** (0.020)
Scotland	0.037* (0.015)	0.000 (0.014)	0.015 (0.014)	-0.039** (0.013)	-0.009 (0.013)	0.001 (0.013)	-0.014 (0.012)	-0.010 (0.013)	0.026* (0.013)	-0.006 (0.014)	0.012 (0.014)	-0.015 (0.015)	0.098*** (0.018)
N. Ireland	-0.035 (0.021)	-0.087*** (0.019)	-0.022 (0.019)	-0.126*** (0.018)	-0.035 (0.018)	-0.050** (0.018)	-0.054** (0.017)	-0.072*** (0.019)	-0.051** (0.019)	-0.085*** (0.019)	-0.087*** (0.020)	-0.097*** (0.022)	-0.103*** (0.026)
N	292926	312710	324865	301874	348196	321424	326766	318559	338807	329150	324114	318203	343751
pseudo R^2	0.043	0.021	0.042	0.036	0.052	0.044	0.027	0.046	0.044	0.049	0.060	0.060	0.064

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: The Ordered Probit Models used to Fit Investment Income in the SPI: Women

	2005	2006	2007	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
ointtot													
Age 25-34	0.404*** (0.014)	0.454*** (0.011)	0.317*** (0.011)	0.351*** (0.010)	-0.078*** (0.011)	-0.352*** (0.010)	0.229*** (0.010)	0.223*** (0.010)	0.211*** (0.012)	0.232*** (0.011)	0.201*** (0.010)	0.234*** (0.010)	-0.099*** (0.010)
Age 35-44	1.295*** (0.012)	0.710*** (0.010)	0.898*** (0.010)	0.860*** (0.009)	0.405*** (0.011)	0.142*** (0.010)	0.493*** (0.011)	0.610*** (0.010)	0.542*** (0.012)	0.616*** (0.011)	0.559*** (0.010)	0.612*** (0.010)	-0.005 (0.010)
Age 45-54	1.388*** (0.012)	0.805*** (0.010)	0.999*** (0.010)	0.994*** (0.009)	0.510*** (0.011)	0.233*** (0.010)	0.614*** (0.010)	0.729*** (0.010)	0.650*** (0.012)	0.722*** (0.010)	0.667*** (0.010)	0.713*** (0.010)	0.225*** (0.010)
Age 55-64	1.457*** (0.013)	0.896*** (0.011)	1.117*** (0.011)	1.123*** (0.010)	0.719*** (0.011)	0.430*** (0.010)	0.793*** (0.011)	0.910*** (0.011)	0.819*** (0.013)	0.857*** (0.011)	0.802*** (0.011)	0.870*** (0.010)	0.563*** (0.010)
Age 65-74	0.998*** (0.015)	0.486*** (0.014)	1.050*** (0.013)	0.913*** (0.012)	0.826*** (0.013)	0.489*** (0.012)	0.917*** (0.012)	1.455*** (0.011)	1.363*** (0.012)	1.426*** (0.011)	1.458*** (0.011)	1.524*** (0.011)	0.991*** (0.011)
Age 75+	1.121*** (0.016)	0.599*** (0.014)	1.107*** (0.014)	0.843*** (0.012)	0.807*** (0.014)	0.542*** (0.012)	0.884*** (0.013)	1.533*** (0.011)	1.410*** (0.012)	1.464*** (0.012)	1.548*** (0.012)	1.606*** (0.011)	0.958*** (0.011)
North-West	0.022 (0.016)	0.038* (0.015)	0.041** (0.015)	0.036** (0.014)	0.024 (0.016)	-0.007 (0.015)	0.011 (0.015)	0.028* (0.014)	0.021 (0.015)	0.006 (0.014)	0.034* (0.014)	0.011 (0.013)	0.007 (0.015)
Yorks/Humbs	0.026 (0.017)	0.046** (0.015)	0.027 (0.016)	0.023 (0.015)	0.030 (0.017)	0.012 (0.015)	0.019 (0.016)	0.030* (0.015)	-0.005 (0.016)	-0.016 (0.015)	0.036* (0.015)	0.010 (0.014)	0.123*** (0.016)
E. Midlands	0.052*** (0.017)	0.041** (0.016)	0.022 (0.016)	0.040** (0.015)	0.029 (0.017)	0.008 (0.016)	0.021 (0.016)	0.018 (0.015)	0.004 (0.017)	-0.013 (0.015)	0.032* (0.015)	0.016 (0.014)	0.152*** (0.016)
W. Midlands	0.058*** (0.017)	0.064*** (0.015)	0.054*** (0.016)	0.045** (0.014)	0.040* (0.017)	-0.000 (0.015)	0.032* (0.016)	0.030* (0.014)	0.049** (0.016)	-0.001 (0.015)	0.024 (0.015)	0.016 (0.014)	0.193*** (0.015)
E. England	0.097*** (0.016)	0.108*** (0.015)	0.091*** (0.015)	0.079*** (0.014)	0.078*** (0.016)	0.022 (0.015)	0.081*** (0.015)	0.069*** (0.014)	0.048** (0.016)	0.027 (0.014)	0.047*** (0.014)	0.028* (0.014)	0.235*** (0.015)
London	0.131*** (0.016)	0.116*** (0.015)	0.088*** (0.015)	0.034* (0.014)	0.047** (0.016)	0.002 (0.014)	0.066*** (0.015)	0.071*** (0.014)	0.048** (0.015)	0.006 (0.014)	0.045** (0.014)	0.024 (0.013)	0.132*** (0.015)
South-East	0.132*** (0.015)	0.143*** (0.014)	0.146*** (0.015)	0.133*** (0.013)	0.104*** (0.016)	0.039** (0.014)	0.115*** (0.014)	0.087*** (0.013)	0.059*** (0.015)	0.043** (0.014)	0.075*** (0.014)	0.043*** (0.013)	0.287*** (0.014)
South-West	0.105*** (0.016)	0.112*** (0.015)	0.106*** (0.015)	0.103*** (0.014)	0.088*** (0.017)	0.036* (0.015)	0.074*** (0.016)	0.058*** (0.014)	0.042** (0.016)	0.017 (0.015)	0.017 (0.015)	0.006 (0.014)	0.237*** (0.015)
Wales	0.027 (0.019)	0.030 (0.017)	0.034 (0.017)	0.047** (0.016)	0.022 (0.019)	0.005 (0.017)	0.024 (0.017)	0.020 (0.016)	0.017 (0.018)	-0.033* (0.017)	0.015 (0.016)	-0.011 (0.016)	0.140*** (0.018)
Scotland	0.040* (0.016)	0.091*** (0.015)	0.073*** (0.015)	0.061*** (0.014)	0.032* (0.016)	0.006 (0.015)	0.047** (0.015)	0.045** (0.014)	0.050** (0.016)	0.025 (0.015)	0.047** (0.015)	0.020 (0.014)	0.171*** (0.015)
N. Ireland	-0.006 (0.022)	0.035 (0.021)	0.007 (0.021)	0.007 (0.019)	0.000 (0.022)	-0.056** (0.020)	-0.025 (0.020)	-0.013 (0.019)	-0.051* (0.021)	-0.053** (0.020)	-0.031 (0.019)	-0.023 (0.018)	-0.076*** (0.021)
N	195218	204448	214716	272786	247723	269759	276055	318821	297655	326449	321034	313742	318966
pseudo R^2	0.050	0.017	0.033	0.034	0.030	0.023	0.024	0.058	0.051	0.056	0.059	0.061	0.051

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 14: probit coefficients from SPI (male)

	2005	2006	2007	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
odividents													
Age 25-34	0.165*** (0.015)	0.085*** (0.014)	0.080*** (0.014)	0.206*** (0.017)	0.187*** (0.017)	0.282*** (0.018)	0.371*** (0.019)	0.369*** (0.023)	0.252*** (0.021)	0.275*** (0.021)	0.252*** (0.021)	0.298*** (0.025)	0.261*** (0.022)
Age 35-44	0.465*** (0.015)	0.287*** (0.013)	0.244*** (0.013)	0.490*** (0.016)	0.492*** (0.016)	0.587*** (0.017)	0.690*** (0.019)	0.707*** (0.023)	0.591*** (0.020)	0.590*** (0.020)	0.591*** (0.021)	0.638*** (0.024)	0.613*** (0.021)
Age 45-54	0.566*** (0.015)	0.377*** (0.013)	0.346*** (0.013)	0.624*** (0.016)	0.626*** (0.016)	0.720*** (0.017)	0.822*** (0.019)	0.842*** (0.023)	0.727*** (0.020)	0.725*** (0.020)	0.730*** (0.021)	0.795*** (0.024)	0.760*** (0.021)
Age 55-64	0.620*** (0.015)	0.429*** (0.013)	0.388*** (0.013)	0.650*** (0.016)	0.661*** (0.016)	0.758*** (0.017)	0.829*** (0.019)	0.851*** (0.023)	0.740*** (0.020)	0.750*** (0.020)	0.753*** (0.021)	0.800*** (0.024)	0.781*** (0.021)
Age 65-74	0.832*** (0.015)	0.470*** (0.014)	0.325*** (0.014)	0.571*** (0.016)	0.612*** (0.017)	0.718*** (0.017)	0.755*** (0.019)	0.741*** (0.023)	0.651*** (0.021)	0.637*** (0.021)	0.664*** (0.021)	0.721*** (0.024)	0.707*** (0.021)
Age 75+	0.974*** (0.017)	0.621*** (0.015)	0.452*** (0.015)	0.704*** (0.018)	0.632*** (0.019)	0.707*** (0.019)	0.712*** (0.020)	0.652*** (0.024)	0.531*** (0.022)	0.506*** (0.022)	0.526*** (0.022)	0.606*** (0.025)	0.596*** (0.022)
Ln Earned Income	0.006 (0.004)	-0.016*** (0.004)	-0.040*** (0.003)	-0.101*** (0.004)	-0.128*** (0.004)	-0.114*** (0.004)	-0.100*** (0.004)	-0.138*** (0.004)	-0.207*** (0.004)	-0.233*** (0.004)	-0.215*** (0.004)	-0.232*** (0.004)	-0.264*** (0.004)
North-West	0.095*** (0.017)	0.098*** (0.016)	0.076*** (0.016)	0.103*** (0.018)	0.099*** (0.017)	0.126*** (0.017)	0.153*** (0.017)	0.137*** (0.018)	0.110*** (0.018)	0.109*** (0.018)	0.112*** (0.018)	0.111*** (0.019)	0.140*** (0.018)
Yorks/Humbs	0.065*** (0.018)	0.073*** (0.017)	0.058*** (0.017)	0.055*** (0.019)	0.048*** (0.018)	0.087*** (0.018)	0.115*** (0.018)	0.112*** (0.019)	0.067*** (0.019)	0.074*** (0.019)	0.075*** (0.019)	0.087*** (0.020)	0.111*** (0.019)
E. Midlands	0.063*** (0.018)	0.104*** (0.017)	0.074*** (0.017)	0.112*** (0.019)	0.102*** (0.018)	0.133*** (0.018)	0.154*** (0.018)	0.159*** (0.019)	0.092*** (0.019)	0.123*** (0.019)	0.116*** (0.019)	0.135*** (0.020)	0.138*** (0.019)
W. Midlands	0.077*** (0.018)	0.094*** (0.017)	0.058*** (0.017)	0.098*** (0.018)	0.088*** (0.017)	0.136*** (0.018)	0.139*** (0.018)	0.139*** (0.018)	0.084*** (0.018)	0.099*** (0.018)	0.119*** (0.019)	0.121*** (0.019)	0.125*** (0.019)
E. England	0.141*** (0.017)	0.147*** (0.016)	0.123*** (0.016)	0.212*** (0.018)	0.199*** (0.017)	0.248*** (0.017)	0.277*** (0.017)	0.262*** (0.017)	0.225*** (0.017)	0.226*** (0.018)	0.246*** (0.018)	0.233*** (0.018)	0.270*** (0.018)
London	0.134*** (0.017)	0.161*** (0.016)	0.134*** (0.016)	0.244*** (0.017)	0.244*** (0.016)	0.279*** (0.017)	0.317*** (0.017)	0.312*** (0.017)	0.286*** (0.017)	0.288*** (0.017)	0.309*** (0.018)	0.317*** (0.018)	0.366*** (0.017)
South-East	0.167*** (0.016)	0.194*** (0.016)	0.170*** (0.016)	0.267*** (0.017)	0.263*** (0.016)	0.288*** (0.016)	0.330*** (0.016)	0.338*** (0.017)	0.312*** (0.017)	0.315*** (0.017)	0.320*** (0.017)	0.319*** (0.018)	0.356*** (0.017)
South-West	0.115*** (0.017)	0.149*** (0.017)	0.115*** (0.017)	0.194*** (0.018)	0.191*** (0.017)	0.223*** (0.017)	0.233*** (0.017)	0.225*** (0.018)	0.185*** (0.018)	0.209*** (0.018)	0.203*** (0.018)	0.202*** (0.019)	0.225*** (0.018)
Wales	0.034 (0.020)	0.016 (0.019)	0.011 (0.019)	0.031 (0.021)	0.020 (0.020)	0.033 (0.020)	0.066** (0.020)	0.046* (0.021)	0.023 (0.020)	0.009 (0.021)	-0.009 (0.022)	0.006 (0.022)	0.022 (0.021)
Scotland	0.068*** (0.018)	0.082*** (0.017)	0.066*** (0.017)	0.099*** (0.018)	0.074*** (0.017)	0.122*** (0.018)	0.154*** (0.018)	0.147*** (0.018)	0.093*** (0.018)	0.087*** (0.019)	0.087*** (0.019)	0.096*** (0.020)	0.124*** (0.019)
N. Ireland	-0.053* (0.024)	-0.007 (0.023)	-0.031 (0.023)	-0.017 (0.025)	-0.030 (0.023)	0.001 (0.024)	0.011 (0.024)	-0.022 (0.025)	-0.033 (0.025)	-0.047 (0.025)	-0.017 (0.026)	-0.015 (0.027)	-0.003 (0.026)
N	292926	312710	324805	301874	348196	321424	326766	318559	338807	329150	324114	318203	343751
pseudo R^2	0.016	0.008	0.005	0.011	0.012	0.012	0.012	0.013	0.014	0.015	0.015	0.016	0.018

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: probit coefficients from SPI (female)

	2005	2006	2007	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
odividents													
Age 25-34	0.096*** (0.017)	0.035* (0.015)	0.018 (0.015)	0.151*** (0.020)	0.045* (0.022)	0.058** (0.022)	0.072** (0.024)	0.078*** (0.022)	0.134*** (0.025)	0.110*** (0.022)	0.184*** (0.023)	0.128*** (0.024)	0.143*** (0.023)
Age 35-44	0.352*** (0.016)	0.170*** (0.014)	0.142*** (0.014)	0.349*** (0.020)	0.247*** (0.021)	0.311*** (0.021)	0.339*** (0.023)	0.368*** (0.022)	0.403*** (0.024)	0.431*** (0.022)	0.507*** (0.022)	0.490*** (0.023)	0.457*** (0.022)
Age 45-54	0.417*** (0.016)	0.239*** (0.014)	0.193*** (0.014)	0.447*** (0.019)	0.354*** (0.021)	0.420*** (0.021)	0.459*** (0.023)	0.506*** (0.021)	0.522*** (0.024)	0.538*** (0.021)	0.602*** (0.022)	0.602*** (0.023)	0.573*** (0.022)
Age 55-64	0.455*** (0.016)	0.309*** (0.015)	0.248*** (0.015)	0.518*** (0.019)	0.455*** (0.021)	0.538*** (0.021)	0.579*** (0.023)	0.607*** (0.021)	0.582*** (0.024)	0.589*** (0.022)	0.637*** (0.022)	0.633*** (0.023)	0.619*** (0.022)
Age 65-74	0.792*** (0.019)	0.510*** (0.017)	0.372*** (0.016)	0.712*** (0.020)	0.652*** (0.023)	0.703*** (0.022)	0.768*** (0.024)	0.739*** (0.023)	0.677*** (0.025)	0.732*** (0.023)	0.743*** (0.023)	0.707*** (0.023)	0.679*** (0.023)
Age 75+	0.894*** (0.019)	0.552*** (0.017)	0.403*** (0.017)	0.751*** (0.021)	0.652*** (0.024)	0.691*** (0.023)	0.759*** (0.025)	0.733*** (0.024)	0.600*** (0.026)	0.641*** (0.024)	0.630*** (0.024)	0.625*** (0.024)	0.598*** (0.023)
North-West	0.032 (0.021)	0.007 (0.020)	0.060** (0.020)	0.110*** (0.023)	0.133*** (0.025)	0.137*** (0.024)	0.091*** (0.025)	0.132*** (0.024)	0.084*** (0.024)	0.115*** (0.024)	0.119*** (0.023)	0.073** (0.023)	0.108*** (0.022)
Yorks/Humbs	0.017 (0.022)	0.008 (0.021)	0.062** (0.021)	0.091*** (0.024)	0.118*** (0.026)	0.116*** (0.025)	0.085*** (0.026)	0.133*** (0.025)	0.060* (0.025)	0.106*** (0.025)	0.139*** (0.024)	0.090*** (0.024)	0.133*** (0.023)
E. Midlands	0.045* (0.023)	0.020 (0.021)	0.095*** (0.021)	0.100*** (0.025)	0.146*** (0.026)	0.146*** (0.025)	0.119*** (0.026)	0.183*** (0.025)	0.122*** (0.025)	0.171*** (0.025)	0.160*** (0.024)	0.116*** (0.024)	0.118*** (0.024)
W. Midlands	0.035 (0.022)	0.034 (0.021)	0.080*** (0.021)	0.108*** (0.024)	0.151*** (0.026)	0.172*** (0.024)	0.125*** (0.025)	0.158*** (0.025)	0.129*** (0.025)	0.148*** (0.024)	0.152*** (0.024)	0.126*** (0.024)	0.140*** (0.023)
E. England	0.104*** (0.021)	0.089*** (0.020)	0.149*** (0.020)	0.222*** (0.023)	0.259*** (0.025)	0.242*** (0.023)	0.243*** (0.024)	0.287*** (0.024)	0.242*** (0.024)	0.253*** (0.023)	0.264*** (0.023)	0.229*** (0.023)	0.244*** (0.022)
London	0.116*** (0.021)	0.111*** (0.020)	0.156*** (0.019)	0.259*** (0.022)	0.301*** (0.024)	0.322*** (0.023)	0.290*** (0.024)	0.324*** (0.023)	0.291*** (0.023)	0.327*** (0.023)	0.329*** (0.022)	0.296*** (0.022)	0.317*** (0.021)
South-East	0.116*** (0.020)	0.117*** (0.019)	0.168*** (0.019)	0.273*** (0.022)	0.303*** (0.023)	0.331*** (0.022)	0.303*** (0.023)	0.342*** (0.023)	0.301*** (0.023)	0.322*** (0.023)	0.331*** (0.022)	0.305*** (0.022)	0.319*** (0.021)
South-West	0.067** (0.022)	0.078*** (0.020)	0.154*** (0.020)	0.215*** (0.023)	0.251*** (0.025)	0.270*** (0.023)	0.264*** (0.024)	0.298*** (0.024)	0.233*** (0.024)	0.284*** (0.024)	0.270*** (0.023)	0.232*** (0.023)	0.252*** (0.022)
Wales	0.013 (0.025)	-0.033 (0.023)	0.039 (0.023)	0.042 (0.027)	0.051 (0.028)	0.058* (0.027)	0.044 (0.028)	0.071* (0.028)	0.046 (0.028)	0.058* (0.028)	0.065* (0.027)	0.014 (0.027)	0.037 (0.026)
Scotland	0.046* (0.022)	0.009 (0.021)	0.086*** (0.020)	0.091*** (0.024)	0.130*** (0.025)	0.144*** (0.024)	0.087*** (0.025)	0.126*** (0.025)	0.070** (0.025)	0.079** (0.024)	0.130*** (0.024)	0.065** (0.024)	0.093*** (0.023)
N. Ireland	0.006 (0.029)	-0.006 (0.028)	0.033 (0.028)	0.021 (0.032)	0.073* (0.035)	0.043 (0.032)	-0.026 (0.034)	0.005 (0.033)	-0.002 (0.033)	0.012 (0.031)	-0.030 (0.032)	0.017 (0.031)	-0.004 (0.031)
N	154588	165206	174154	188020	188735	201202	201707	234788	225502	240950	238001	220985	222736
pseudo R^2	0.015	0.008	0.005	0.013	0.013	0.015	0.017	0.016	0.012	0.013	0.012	0.013	0.012

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 16: ML estimation of Pareto distribution for top of interest and dividend income

	Interest income	Dividend income
sex	0.162*** (0.013)	0.221*** (0.009)
year	-0.037*** (0.002)	-0.013*** (0.001)
Constant	75.104*** (3.115)	27.801*** (1.928)
N	52649	201105

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.2 Pension rights

Table 17: The Parameters of the Ordered Probit Model:
Wave 3 of the Wealth and Assets Survey

main	Under 65		65 and Older	
	Emp Inc	No Emp Inc	Pens Inc	No Pens Inc
Age	0.295*** (0.008)	0.175*** (0.017)	-0.814*** (0.033)	-0.203*** (0.040)
No. of adults	-0.134*** (0.018)	-0.209*** (0.062)	0.626*** (0.074)	0.488*** (0.091)
No. of children	-0.011 (0.012)	-0.099** (0.038)	-0.307** (0.116)	-0.300* (0.119)
House owned outright	0.836*** (0.044)	1.361*** (0.076)	0.150* (0.059)	0.507*** (0.096)
House mortgaged	0.614*** (0.033)	1.170*** (0.084)	0.135 (0.084)	0.901*** (0.186)
Married etc.	0.420*** (0.038)	0.608*** (0.089)	-0.549*** (0.084)	-0.231 (0.122)
Widowed	0.171*** (0.045)	0.065 (0.071)	-0.049 (0.073)	-0.154 (0.236)
Ln employment income	0.525*** (0.021)			
Ln pension income			2.186*** (0.095)	
N	11525	2483	5586	1549
Psuedo-R ²	0.11	0.15	0.37	0.11

Table 18: The Parameters of the Ordered Probit Model:
Wave 4 of the Wealth and Assets Survey

main	Under 65		65 and Older	
	Emp Inc	No Emp Inc	Pens Inc	No Pens Inc
Age	0.273*** (0.008)	0.183*** (0.016)	-0.802*** (0.025)	-0.239*** (0.041)
No. of adults	-0.116*** (0.018)	0.019 (0.050)	0.375*** (0.058)	0.564*** (0.081)
No. of children	-0.005 (0.012)	-0.130*** (0.034)	-0.311* (0.130)	-0.359** (0.137)
House owned outright	0.886*** (0.044)	1.326*** (0.070)	0.105* (0.050)	0.459*** (0.102)
House mortgaged	0.559*** (0.032)	0.902*** (0.067)	0.200* (0.096)	0.953*** (0.155)
Married etc.	0.289*** (0.042)	0.350*** (0.077)	-0.155* (0.073)	-0.149 (0.117)
Widowed	0.115* (0.050)	-0.079 (0.068)	-0.090 (0.066)	-0.126 (0.172)
Ln employment income	0.605*** (0.023)			
Ln pension income			2.127*** (0.081)	
N	9968	2841	5857	1574
Pseudo-R ²	0.11	0.14	0.35	0.13

Table 19: The Parameters of the Ordered Probit Model:
Wave 5 of the Wealth and Assets Survey

main	Under 65		65 and Older	
	Emp Inc	No Emp Inc	Pens Inc	No Pens Inc
Age	0.232*** (0.008)	0.215*** (0.022)	-0.894*** (0.035)	-0.333*** (0.056)
No. of adults	-0.123*** (0.020)	-0.171* (0.082)	0.356*** (0.067)	0.397*** (0.083)
No. of children	-0.049*** (0.014)	-0.062 (0.046)	-0.005 (0.178)	0.111 (0.131)
House owned outright	0.888*** (0.045)	1.565*** (0.087)	0.132* (0.063)	0.581*** (0.112)
House mortgaged	0.533*** (0.033)	1.355*** (0.126)	0.228* (0.093)	0.792*** (0.182)
Married etc.	0.265*** (0.043)	0.473*** (0.116)	-0.070 (0.086)	0.028 (0.140)
Widowed	0.029 (0.048)	-0.031 (0.087)	0.004 (0.083)	-0.151 (0.172)
Ln employment income	0.604*** (0.024)			
Ln pension income			2.272*** (0.111)	
N	9729	1880	5872	1505
Pseudo-R ²	0.10	0.17	0.37	0.13

Table 20: The Parameters of the Ordered Probit Model:
Wave 6 of the Wealth and Assets Survey

main	Under 65		65 and Older	
	Emp Inc	No Emp Inc	Pens Inc	No Pens Inc
Age	0.251*** (0.008)	0.177*** (0.024)	-0.785*** (0.039)	-0.287*** (0.045)
No. of adults	-0.141*** (0.020)	-0.135 (0.082)	0.264*** (0.069)	0.440*** (0.086)
No. of children	-0.021 (0.014)	-0.094 (0.052)	-0.103 (0.123)	-0.074 (0.144)
House owned outright	0.845*** (0.046)	1.545*** (0.091)	0.167** (0.062)	0.629*** (0.099)
House mortgaged	0.550*** (0.034)	1.054*** (0.132)	0.317*** (0.095)	0.544** (0.208)
Married etc.	0.393*** (0.048)	0.514*** (0.119)	-0.039 (0.085)	-0.214 (0.124)
Widowed	0.183*** (0.052)	-0.151 (0.087)	-0.044 (0.065)	-0.381* (0.170)
Ln employment income	0.563*** (0.025)			
Ln pension income			2.628*** (0.160)	
N	8935	1611	5923	1559
Pseudo-R ²	0.09	0.16	0.39	0.09

A.3 Housing rent

Table 21: Imputed rent - Heckman

	(1) log monthly rent	(2) selection equation
log hhold mthly income	0.075*** (0.008)	-0.262*** (0.014)
2.htype	-0.146*** (0.026)	0.436*** (0.027)
3.htype	-0.204*** (0.030)	0.904*** (0.034)
4.htype	-0.399*** (0.099)	0.004 (0.124)
nrooms2	0.097*** (0.007)	-0.264*** (0.009)
2.ctband	0.206*** (0.011)	-0.394*** (0.021)
3.ctband	0.351*** (0.015)	-0.620*** (0.024)
4.ctband	0.498*** (0.021)	-0.805*** (0.030)
5.ctband	0.723*** (0.035)	-0.895*** (0.042)
6.ctband	0.850*** (0.057)	-1.019*** (0.057)
7.ctband	1.241*** (0.084)	-1.001*** (0.076)
8.ctband	1.205*** (0.356)	-0.635** (0.200)
9.ctband	0.155 (0.086)	-0.115 (0.279)
2.socioeco	0.044 (0.043)	0.123* (0.057)
3.socioeco	-0.103* (0.043)	0.067 (0.053)
4.socioeco	-0.126** (0.044)	0.114* (0.057)
5.socioeco	-0.155*** (0.045)	0.249*** (0.058)
6.socioeco	-0.255*** (0.045)	0.281*** (0.057)
7.socioeco	-0.288*** (0.044)	0.532*** (0.056)
8.socioeco	-0.269*** (0.044)	0.491*** (0.057)
9.socioeco	-0.257*** (0.047)	1.228*** (0.079)
10.socioeco	-0.084 (0.054)	1.393*** (0.087)
11.socioeco	-0.200*** (0.048)	0.704*** (0.082)
12.socioeco	-0.199***	0.066

	(0.042)	(0.053)
2005.fyear	-0.259***	-0.357***
	(0.027)	(0.048)
2006.fyear	-0.228***	-0.265***
	(0.018)	(0.032)
2007.fyear	-0.199***	-0.227***
	(0.018)	(0.032)
2008.fyear	-0.152***	-0.242***
	(0.018)	(0.032)
2009.fyear	-0.112***	-0.169***
	(0.018)	(0.033)
2010.fyear	-0.098***	-0.101**
	(0.018)	(0.033)
2011.fyear	-0.058***	-0.081*
	(0.017)	(0.032)
2012.fyear	-0.004	-0.088**
	(0.017)	(0.032)
2013.fyear	-0.013	-0.010
	(0.017)	(0.033)
2014.fyear	-0.018	-0.021
	(0.016)	(0.032)
2.reg	0.067***	0.070*
	(0.017)	(0.035)
3.reg	0.004	0.071
	(0.018)	(0.036)
4.reg	0.026	0.095*
	(0.018)	(0.038)
5.reg	0.050**	0.126***
	(0.018)	(0.037)
6.reg	0.096***	0.261***
	(0.020)	(0.038)
7.reg	0.260***	0.598***
	(0.023)	(0.042)
8.reg	0.158***	0.329***
	(0.020)	(0.038)
9.reg	0.088***	0.200***
	(0.020)	(0.039)
10.reg	-0.147***	0.315***
	(0.024)	(0.044)
11.reg	-0.201***	0.026
	(0.019)	(0.039)
_cons	5.644***	2.505***
	(0.077)	(0.137)
	athrho	
	_cons	-0.958***
		(0.068)
	lnsigma	
	_cons	-0.725***
		(0.030)
N	53579	53579
Standard errors in parentheses		

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

B Deflators

Table 22: Democratic and Plutocratic Consumption Deflators: 2015=100

	Democratic	Plutocratic
2006	81.3	82.1
2007	83.9	84.5
2008	87.4	87.8
2009	89.9	90.0
2010	91.8	92.1
2011	94.2	94.5
2012	96.7	96.8
2013	98.5	98.5
2014	99.7	99.7
2015	100.0	100.0
2016	101.4	101.4
2017	103.9	103.9
2018		106.3

Source: Authors' calculations from ONS data.
See Aitken and Weale (2020).