Institute for Fiscal Studies

Income changes and their determinants over the lifecycle

IFS Working Paper W15/33

Andrew Hood
Robert Joyce

Income changes and their determinants over the lifecycle

Andrew Hood and Robert Joyce[1]

December 2015

Abstract

What explains the variation in how income changes as people age? Using household panel data, we investigate the contribution of different time-varying factors in explaining variation in income changes over prime working-age life (between 35-44 and 50-59). We find that demographic changes, such as acquiring or losing a partner and the entry or exit of children to and from the household, account for a larger share of the variation in household income changes than shifts in employment status or occupation. This is particularly true for women, for whom demographic changes explain 82% of ex-post predictable variation in household income changes, compared to only 12% explained by employment status and occupation. We find a similar result when looking at the transition into retirement (between 50-59 and 66-75). These results illustrate an important limitation of the extensive literature examining consumption and savings behaviour over the lifecycle: focusing on earnings and income whilst ignoring changes in household composition excludes the largest source of ex-post predictable variation in income changes.

Acknowledgements:

The authors would like to thank Carl Emmerson and members of the IFS Retirement Savings Consortium for providing useful comments throughout the course of this work. This report was funded by the IFS Retirement Savings Consortium and the Economic and Social Research Council. The IFS Retirement Savings Consortium comprises Age UK, Department for Work and Pensions, Financial Conduct Authority, HM Treasury, Institute and Faculty of Actuaries, Investment Association, Just Retirement and Money Advice Service. Responsibility for interpretation of the data, as well as for any errors, is the authors' alone.

[1]Institute for Fiscal Studies, London. Correspondence to: robert_j@ifs.org.uk.

1. Introduction

The evolution of individuals' incomes across the course of their lives, and the extent to which that evolution is predictable from the perspective of the individual, are central questions in microeconomics. They have implications for the measurement of economic welfare, the modelling of consumption and savings behaviour, and assessments of whether individuals are prepared for retirement. However, the extensive existing literature addressing these questions provides, for the most part, a partial analysis. There are numerous papers seeking to better understand the process governing the evolution of (male) individual earnings (for example Meghir and Pistaferri (2004), Guvenen (2009)). By focusing on the earnings process rather than the income process, this literature explicitly restricts its analysis of the determinants of changes in the economic welfare of individuals to labour market factors. A related literature seeks instead to characterise the process governing household incomes (for example Blundell, Pistaferri and Preston (2008), Etheridge (2015)). While this might seem to allow a more comprehensive analysis of the determinants of income changes, in practice this literature restricts its attention to those households whose demographic characteristics remain constant, as Burgess et. al. (2000) point out. This paper provides a broader analysis of the income process, quantifying the relative importance of changes in labour market status and occupation, demographic characteristics and other factors (such as health) in explaining the variation in income changes across different parts of the lifecycle.

Following a sample of UK households over nearly twenty years using the British Household Panel Survey (BHPS), we investigate the distribution of changes in earnings and incomes over two different periods of the lifecycle: prime working-age life (35-44 to 50-59) and the move into retirement (50-59 to 66-75). We estimate regressions of income on time-varying characteristics, exploiting the panel element of the data to control for fixed unobserved individual heterogeneity. With the resulting parameter estimates, we identify the contribution of different factors – such as changes in labour market status or family structure – to (ex-post) predictable variation in the evolution of incomes across the lifecycle.

The key finding is that changes in family structure and the characteristics of partners are more important than changes in an individual's employment status and occupation in explaining the variance in income changes between mid and late working-age life. There are also important differences by gender. Women see greater variability in individual earnings over this part of the lifecycle, and changes in family circumstances are a particularly important factor in determining changes in their household income – they explain 82% of ex-post predictable variation in income changes, compared to only 12% explained by changes in employment status and occupation. Looking at the transition from late working-age life to retirement, the gender difference remains: while

labour market changes explain more than demographic changes for men, the opposite is true for women.

This paper yields insights relevant to several literatures. As well as complementing the literature on earnings and income processes discussed above, the quantification of the relative importance of non-labour market factors in explaining income changes has implications for the assessment of preparedness for retirement and the 'optimality' of wealth holdings. For example, the conclusion drawn by Scholz, Seshadri and Khitatrakun (2006) and Crawford and O'Dea (2014) that most US and UK households have more wealth than 'optimal' is premised on a model of future income uncertainty that does not incorporate the possibility of separation from a partner. Hence, it is possible that the 'excess' savings are partly explained by the existence of a source of uncertainty not incorporated in such models. The important gender differences in the source of variability in income changes found by this paper suggest potential extensions to the literature on non-unitary models of household behaviour (surveyed in Chiappori and Donni (2009)). For example, Browning (2000) provides a model of the savings behaviour of a two-person household that accounts for mortality differences across couples, but does not account for the differential impact of separation on the future income of each member of the couple.

The paper to which ours is perhaps most closely related is Burgess et. al. (2000). They also draw attention to the importance of demographic change in determining income changes, in addition to labour market factors. An important difference is that, rather than focusing on what explains year-to-year volatility in incomes over a six-year period as they do, our focus is on the more persistent (and perhaps permanent) changes in income - changes that are likely to have more significant repercussions for economic welfare and optimal savings behaviour than transitory fluctuations. We are able to do this by using the full 18-year panel of the BHPS.

Some of the variation in the path of income across individuals will be the result of planned choices, while some will reflect the uncertainty that individuals face about their income when planning for the future. Ultimately, a challenge for research is to distinguish between these two kinds of variation. That task inevitably involves making some assumptions about what people know about their future and what they do not know (as Burgess et al 2000 do, in order to try to separate uncertainty from ex-ante predictable heterogeneity when looking at year-to-year volatility in incomes). In this paper we take a simpler first step, which is to document the observed factors which explain the variation in how incomes evolve over the lifecycle, without taking a definitive view about which of these are known in advance to the individuals concerned.

The rest of the paper proceeds as follows. Section 2 provides a brief description of the BHPS data and the sample selection criteria. Section 3 documents the variability of income changes over the two periods considered, and how that varies across different measures of income and different groups. Section 4 describes the methodology used in

our regression analysis, before the results are presented in Section 5. Section 6 concludes.

2. Data

The British Household Panel Study is a longitudinal dataset on UK households from 1991 to 2008, similar in structure to the US Panel Study of Income Dynamics. In total, there are 252,592 observations of 35,002 separate individuals, with between 10,000 and 20,000 individuals appearing in each wave. Data on household incomes come from an auxiliary, derived dataset (see Levy and Jenkins (2008)), and is missing in 22% of cases (55,971 observations). The measure of household income used is total current net income: the household's labour income plus state benefits and private transfers, net of taxes. We also present results where household incomes are equivalised using the modified OECD equivalence scale.

When looking at income changes in prime working-age life we restrict our sample to those individuals we observe at least three times between the ages of 35 to 44 and at least three times between the ages of 50 and 59. This allows us to average out year-on-year volatility when comparing incomes from the two parts of the lifecycle. We also drop individuals who are observed less than ten times overall in order to ensure the fixed effects estimates of parameters are robust. Combining these restrictions with the requirement that the outcome variable (individual hourly wage, individual earnings, household earnings, household income) is neither missing nor zero (since we take the natural logarithm) leaves us with a sample of between 600 and 850 individuals, observed at least 10 times, for each outcome variable.

We implement similar restrictions when looking at income changes as individuals move into retirement. We restrict our sample to those individuals observed at least three times between the ages of 50 and 59 and at least once between the ages of 66 and 75. We do not require multiple observations in retirement as incomes are much less volatile in that part of the lifecycle. Alongside the restrictions that individuals are observed at least ten times overall and that household income is non-missing, this leaves us with a sample of 658 individuals.

3. Descriptives

In this section, we quantify the variability in the income changes seen by individuals and how that varies across different income measures and different groups within the population. We do so by describing the distribution of changes in log earnings and income over prime working-age life (35-44 to 50-59) and as individuals move into retirement (50-59 to 66-75).

Table 1 describes the distribution of changes in hourly wages, individual weekly earnings, household earnings, net household income and equivalised net household

income over prime working-age life. When looking at earnings, we restrict our sample to those with positive (individual or household) earnings, but we include those with zero earnings when looking at changes in income. Table A1 in Appendix A shows that our qualitative conclusions are unchanged if we instead focus on a common sample.

Looking first at the mean change, real individual hourly wages increase by an average of 10 log points, and real individual weekly earnings by an average of 8 log points. However, real household earnings are unchanged, reflecting a falling employment rate. The mean change in net household income is 12 log points (indicating stronger income growth in unearned income) and the mean increase in equivalised income is 27 log points, as children leave the home.

The variance of changes in log earnings and income provides a measure of how widely dispersed those changes were across individuals. Looking first at the variances for the whole sample, the variance in the change in individual hourly wages (as measured by the variance of logs) is 0.380, significantly less than the variance in individual and household weekly earnings (0.872 and 0.747 respectively). This reflects the inclusion of changes in hours as an additional source of variation. However, both earnings measures see significantly more variation than the change in household net income (0.269 unequivalised, 0.238 equivalised). This difference is the result of the insurance role performed by the tax and benefit system: falls in earnings are mitigated by lower tax payments and/or higher benefit receipt, and vice versa for rises in earnings.

In Section 4 we show that the determinants of income risk differ importantly by gender. The second and third panels of Table 1 provide context for that analysis, comparing the variability of income changes between men and women. Unsurprisingly, there is little difference across genders when looking at outcomes at the household level. However, while inequality in the change in hourly wages is slightly lower for women (0.321 compared to 0.433), the variance of weekly earnings is higher (0.911 compared to 0.798), reflecting more variation in hours. In particular, women are relatively likely to move from no or part-time work to full-time work over this part of the lifecycle as childcare responsibilities come to an end. In all cases, the distribution of changes is negatively skewed (as evident from the fact that the median is greater than the mean): there is a long 'bottom tail' of individuals who see sharp falls in their earnings or income.

Table 1. The distribution of changes in log earnings and income between the ages of 35-44 and 50-59

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised net household income
All					
Mean	0.099	0.082	0.001	0.123	0.268
Median	0.155	0.179	0.087	0.173	0.301
25th percentile	-0.055	-0.079	-0.266	-0.132	0.052
75th percentile	0.350	0.424	0.416	0.429	0.537
Variance	0.380	0.872	0.747	0.269	0.238
N	632	637	841	730	730
Male					
Mean	0.049	-0.055	0.009	0.143	0.261
Median	0.128	0.102	0.116	0.193	0.302
25th percentile	-0.076	-0.120	-0.235	-0.108	0.052
75th percentile	0.303	0.272	0.434	0.443	0.519
Variance	0.433	0.798	0.748	0.251	0.235
N	322	325	406	363	363
Female					
Mean	0.151	0.224	-0.006	0.104	0.274
Median	0.177	0.282	0.059	0.153	0.291
25th percentile	-0.004	0.008	-0.297	-0.156	0.051
75th percentile	0.413	0.588	0.385	0.419	0.554
Variance	0.321	0.911	0.748	0.288	0.241
N	310	312	435	367	367

Notes: Sample is individuals for whom we have 10 or more observations, with at least three in each age group.
Source: Authors' calculations.

Table 2 provides the same quantification of the distribution of changes in income, but for the stage in the lifecycle typically associated with a move into retirement (from the ages of 50 to 59 to the ages of 66 to 75). Since the majority of the population are not earning in the latter period, we restrict our focus to changes in household income (unequivalised and equivalised).

There are three key things to note about changes in income as individuals move into retirement. First, mean falls in unequivalised household income are relatively small – around 10 log points. After equivalisation, these falls are offset by the departure of children, meaning the mean change in equivalised household income is roughly zero.

Second, the variance of the change in log income over this period is slightly smaller than during prime working-age life: 0.201 compared to 0.269 on an unequivalised basis, 0.171 compared to 0.238 when incomes are equivalised. This indicates, perhaps surprisingly, that there is less variation in the income changes individuals see when moving from work to retirement than through the latter half of working life. Third, the variation in income changes over this part of the lifecycle is slightly larger for women than for men, even when income is measured at the household level, which may reflect greater uncertainty faced by women over this part of the lifecycle.

Table 2. The distribution of changes in log income between the ages of 50-59 and 66-75

	Household net income	Equivalised net household income
All		
Mean	-0.099	0.009
Median	-0.111	-0.011
25th percentile	-0.353	-0.229
75th percentile	0.145	0.241
Variance	0.201	0.171
N	658	658
Male		
Mean	-0.102	0.013
Median	-0.104	-0.008
25th percentile	-0.346	-0.237
75th percentile	0.137	0.241
Variance	0.183	0.161
N	315	315
Female		
Mean	-0.096	0.007
Median	-0.124	-0.023
25th percentile	-0.369	-0.223
75th percentile	0.153	0.244
Variance	0.219	0.181
N	343	343

Notes: Sample is individuals for whom we have 10 or more observations, with at least three between the ages of 50 and 59, and one between the ages of 66 and 75.
Source: Authors' calculations.

4. Methodology

We assume a model of earnings or income of the following form:

$$y_{it} = F_i + \sum_{k=1}^{K} \beta^k X_{it}^k + \varepsilon_{it}$$

where i indexes individuals and t indexes year (which runs from 1991 to 2008). y_{it} is one of the five (log) earnings or income variables from Table 1. F_i is an individual fixed effect, capturing the impacts on the outcome of all time-invariant characteristics of individual i (whether or not those characteristics are observed in the BHPS). $X_{it} = [X_{it}^1, \dots, X_{it}^K]$ is a vector of observed time-varying characteristics which may be correlated with F_i. ε_{it} is an error term, capturing the impacts of unobserved time-varying factors.

We estimate $\beta = [\beta^1, \dots, \beta^K]$ using fixed effects regression to obtain $\hat{\beta}$, which is a consistent estimator of the causal impacts of X_{it} under the assumption that $E(\varepsilon_{it} - \bar{\varepsilon}_i | X_{it} - \bar{X}_i) = 0$, where $\bar{\varepsilon}_i = \frac{1}{T}\sum_{t=1}^{T} \varepsilon_{it}$ and $\bar{X}_i$ is defined analogously.[2]

The samples which we use to estimate these models are as defined in Section 3. Essentially we use individuals in birth cohorts that we observe a sufficient number of times in the 1991-2008 BHPS between the ages of 35-44 and 50- 59 (for the analysis of changes over prime working-age life) and 50-59 and 66-75 (for the analysis of changes between late working-age life and retirement).

The explanatory variables we include in X_{it} are whether or not an individual is in paid employment, with those in work split according to five occupation groups (unskilled, partly skilled, skilled manual, skilled non-manual and professional or managerial); family type (single without children, single with children, couple without children, couple with children), with those in couples further split according to the employment status and education of their partner (out of work, in work with no formal qualifications, in work with less than degree-level qualifications, in work with a degree or higher); a health index constructed from a self-assessment of general health and a series of questions about specific health conditions[3]; a three-category variable about caring responsibilities (no caring responsibilities, less than 20 hours per week, at least 20 hours per week); English region or nation of the UK; and a full set of year dummies and a quadratic in age (included just as control variables).[4] For the analysis of income changes between late working-age life and retirement, we add more dummy variables which split those not in work into five sub-categories: aged under 55, between age 55 and state pension age[5] (to separate out those who are likely to have voluntarily taken early retirement), over state pension age and receiving no occupational pension, over

[2] When computing standard errors we allow for heteroscedasticity and clustering of the error term at the individual level using the Huber-White sandwich estimator. That is, we assume that ε_{it} is independently distributed across individuals, but it may be distributed differently across individuals and it may be serially correlated for a given individual.

[3] We broadly follow Disney, Emmerson and Wakefield (2006), based on a suggestion by Bound et al (1999), in running an ordered probit of a 5-category self-reported health variable on a number of specific health indicator variables, and taking the predicted underlying index value. The specific health indicators are problems with arms, legs or hands, sight, hearing, skin conditions/allergies, chest/breathing, heart/blood pressure, stomach or digestion, diabetes, anxiety or depression, alcohol or drugs, epilepsy, migraines or 'other'. This has the advantage of providing a single summary measure of health from a detailed set of information, and potentially purging the self-reported health variable of endogeneity which might arise if (for example) changes in people's perceptions of their health are related to changes in their income.

[4] We also tried interacting the quadratic in age with education, but the differences in age trends between education groups were very small and not statistically significant.

[5] Age 60 for women and 65 for men.

state pension age and receiving a public sector pension, and over state pension age and receiving a private sector pension.

Having estimated the model, our interest lies in what changes in circumstances explain the variance of (log) earnings/income changes between one period of life and another. Equation 1 implies that the change in earnings/income for individual i between any two years is given by:

$$\Delta y_i = \sum_{k=1}^{K} \beta^k \Delta X_i^k + \Delta \varepsilon_i$$

and so the predicted change, given X_i, is simply

$$\Delta \hat{y}_i = \sum_{k=1}^{K} \hat{\beta}^k \Delta X_i^k.$$

Without the contribution of variable X_j, the predicted change would instead be

$$\Delta \hat{y}_i^{-j} = \Delta \hat{y}_i - \hat{\beta}^j \Delta X_i^j.$$

The quantities of central interest in the next section are the proportions of the predictable variance in log-earnings/income changes that are due to differential changes in different variables, holding other variables constant. For example, for variable X_j, we report the fraction of the explained variance accounted for by X_j, which is:

$$1 - \frac{Var(\Delta \hat{y}_i^{-j})}{Var(\Delta \hat{y}_i)}$$

From this it is clear that a variable X_j will have no estimated impact on the variance of log-earnings/income changes if

$$Var(\hat{\beta}^j \Delta X_i^j) = \hat{\beta}^{j\,2} Var\left(\Delta X_i^j\right) = 0.$$

This will occur if the variable has no estimated impact on the outcome ($\hat{\beta}^j = 0$) or if it changes in the same way for everyone over time ($Var\left(\Delta X_i^j\right) = 0$). Conversely, all else equal a variable will be more important for these purposes if it has a larger impact on the outcome or if it changes in a more variable way across individuals.[6]

In practice, most explanatory variables in our model are categorical variables which we enter as a set of dummies. Hence in most cases we are considering the contribution of a set of dummy variables rather than a single regressor. In some cases it is informative to split further the contributions of a set of dummy variables into subsets. For example, for dummies underlying a 6-category variable for "out of work" or "in work" and in one of 5 different occupation categories, it is interesting to separate the contributions of changes

[6] We will also capture any impact of a variable that comes through its covariance with changes in other variables.

in work status from changes in occupation for those who are in work. We handle this with a simple extension, outlined in Appendix D.

Of course, $Var(\Delta y_i)$ will be larger than $Var(\Delta \hat{y}_i)$, because of changes in unobserved time-varying factors affecting earnings/income, ε_{it}. For example, in the simplest case of iid errors, the variance of the total change would have an additional $2\sigma_\varepsilon^2$ term, where σ_ε^2 is the variance of ε_{it}. In fact, typically only a small fraction of $Var(\Delta y_i)$ is due to changes in observed factors. We separately report this fraction, i.e. $\frac{Var(\Delta \hat{y}_i)}{Var(\Delta y_i)}$.

The earnings/income changes that we are interested in are those occurring between ages 35-44 and 50-59, or between 50-59 and 66-75. We aim to focus on persistent (or permanent) income changes, rather than year-to-year volatility. Hence the difference that we actually focus on is that between average income over the age range 35-44 and average income over the age range 50-59 (or between 50-59 and 66-75).

5. Results

Coefficient estimates

Focusing first on our prime working-age sample, Tables 3 to 5 present the main coefficient estimates of interest from estimation of the model in equation 1, for the same 5 earnings/income outcomes as in Tables 1 and 2, estimated on the full sample and for men and women separately.

Changes in occupation are unsurprisingly strongly associated with changes in workers' gross individual earnings, especially for men. For example, being in skilled manual work rather than unskilled work is associated with a 62 log-point increase in earnings for men and a 34 log-point increase for women (controlling for individual fixed effects, as all these estimates do). But their proportional impacts on net household income, whilst still present, are much more muted: the numbers above fall to 24 log-points and 16 log-points respectively. This reflects the insurance effect of the tax and benefit system, and for women it also reflects the fact that they are normally secondary earners and therefore their earnings account for only a minority of their household's income.

Conversely, changes in family circumstances become particularly important once one looks at changes in household incomes – though some transitions, such as that into lone parenthood, are associated with earnings changes too – and this is particularly true for women. For example, the estimates show that moving from being single without children to being in a couple with a working high-educated partner is, on average, associated with a much larger increase in household income for women (even after equivalisation) than a switch from non-employment into a professional/managerial occupation (73 log-points versus 20 log-points without equivalisation, or 61 versus 16 with equivalisation).

The estimates also suggest that acquiring substantial caring responsibilities (at least 20 hours per week) has a negative association with changes in net equivalised household income for women, with some evidence that this comes through impacts on hours of paid work (comparing coefficients on hourly wages and weekly earnings). We do not find significant associations between changes in our index measure of health and changes in earnings or income.

Table 3. Coefficients from fixed effects regressions of log income between ages of 35-44 and 50-59: all

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised household net income
Age – 35	0.085**	0.108***	0.061**	0.029	0.003
Age – 35, squared	-0.001*	-0.002***	-0.001**	-0.001***	-0.000
Employment and occupation					
Unskilled	-	-	0.347***	0.154***	0.118***
Partly skilled	0.251	0.404**	0.452***	0.165***	0.132***
Skilled manual	0.310*	0.514***	0.508***	0.184***	0.154***
Skilled non-manual	0.324**	0.545***	0.509***	0.154***	0.128***
Professional	0.455***	0.734***	0.640***	0.220***	0.194***
Demographic status					
Couple without children, partner out of work	-0.014	0.017	-0.089	0.209***	0.013
Couple without children, working low-educated partner	0.035	0.094	0.508***	0.343***	0.130**
Couple without children, working mid-educated partner	0.049	0.069	0.549***	0.425***	0.221***
Couple without children, working degree-educated partner	-0.065	-0.053	0.574***	0.426***	0.219**
Lone parent	-0.150*	-0.205**	-0.564***	-0.215***	-0.325***
Couple with children, partner out of work	-0.073	-0.055	-0.187*	0.129*	-0.205***
Couple with children, working low-educated partner	-0.021	-0.008	0.305***	0.276***	-0.055
Couple with children, working mid-educated partner	-0.029	-0.050	0.332***	0.304***	-0.041
Couple with children, working degree-educated partner	-0.128	-0.242	0.383**	0.311***	-0.063
Health index	0.029	0.025	-0.003	0.002	0.004
Care status					
Caring less than 20 hours	0.033	0.012	-0.043	-0.015	-0.007
Caring at least 20 hours	0.107*	-0.019	-0.036	-0.056	-0.065*

Notes: Omitted categories are not employed, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Table 4. Coefficients from fixed effects regressions of log income between ages of 35-44 and 50-59: male

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised household net income
Age – 35	0.112*	0.119*	0.090	-0.006	-0.028
Age – 35, squared	-0.001	-0.002*	-0.002**	-0.001*	-0.000
Employment and occupation					
Unskilled	-	-	0.928***	0.193***	0.140*
Partly skilled	0.546	0.518	1.053***	0.250***	0.227***
Skilled manual	0.622*	0.619*	1.092***	0.241***	0.228***
Skilled non-manual	0.684*	0.614*	1.149***	0.242***	0.224***
Professional	0.721**	0.718**	1.210***	0.272***	0.263***
Demographic status					
Couple without children, partner out of work	-0.132	-0.014	-0.138	0.022	-0.270***
Couple without children, working low-educated partner	-0.014	0.104	0.524***	0.257***	-0.103
Couple without children, working mid-educated partner	-0.032	0.038	0.379***	0.261***	-0.057
Couple without children, working degree-educated partner	-0.264**	-0.197	0.245*	0.121	-0.206**
Lone parent	-0.336	-0.359	-0.524*	-0.407*	-0.579**
Couple with children, partner out of work	-0.200	-0.118	-0.205*	-0.013	-0.483***
Couple with children, working low-educated partner	-0.082	-0.010	0.223**	0.152**	-0.326***
Couple with children, working mid-educated partner	-0.058	0.013	0.158	0.152**	-0.323***
Couple with children, working degree-educated partner	-0.321**	-0.245	-0.060	0.037	-0.451***
Health index	0.043	0.043	-0.011	0.005	0.010
Care status					
Caring less than 20 hours	0.066	0.031	-0.036	-0.009	0.004
Caring at least 20 hours	0.054	0.039	-0.001	-0.044	-0.030

Notes: Omitted categories are not employed, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Table 5. Coefficients from fixed effects regressions of log income between ages of 35-44 and 50-59: female

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised household net income
Age – 35	0.056	0.090	0.019	0.049**	0.024
Age – 35, squared	-0.001	-0.003***	-0.001**	-0.001***	-0.000
Employment and occupation					
Unskilled	-	-	0.147*	0.132**	0.096*
Partly skilled	0.056	0.312**	0.194***	0.116***	0.076**
Skilled manual	0.033	0.338**	0.218***	0.160***	0.104***
Skilled non-manual	0.108	0.516***	0.277***	0.125***	0.088***
Professional	0.301***	0.779***	0.398***	0.202***	0.155***
Demographic status					
Couple without children, partner out of work	0.083	0.075	-0.208	0.325***	0.198*
Couple without children, working low-educated partner	0.008	0.068	0.415***	0.360***	0.275***
Couple without children, working mid-educated partner	0.082	0.088	0.612***	0.515***	0.394***
Couple without children, working degree-educated partner	0.382	0.408	1.038***	0.733***	0.612***
Lone parent	-0.072	-0.135	-0.520***	-0.146**	-0.193***
Couple with children, partner out of work	0.131	0.160	-0.591***	0.196*	-0.020
Couple with children, working low-educated partner	-0.037	-0.039	0.336***	0.350***	0.157*
Couple with children, working mid-educated partner	-0.040	-0.093	0.429***	0.380***	0.143*
Couple with children, working degree-educated partner	0.268	0.052	0.866***	0.543***	0.265***
Health index	0.022	0.014	-0.014	-0.004	-0.000
Care status					
Caring less than 20 hours	-0.002	-0.004	-0.037	-0.021	-0.021
Caring at least 20 hours	0.140*	-0.065	-0.069	-0.055	-0.089**

Notes: Omitted categories are not employed, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Distribution of changes in time-varying characteristics

The importance of different factors in explaining variation in the path of income over the lifecycle also depends on how much their evolution varies across individuals. Something might have an important effect on income, but if it hardly ever changes, or it

changes in the same way for everyone over the lifecycle, it will not be important for our purposes.

Table 6. The distribution of changes in time-varying characteristics between the ages of 35-44 and 50-59

	Male	Female	All
Employment			
Moved into employment	5.1%	13.9%	9.6%
Moved out of employment	16.8%	19.78%	18.3%
Remained in employment	73.6%	54.0%	63.7%
Remained out of employment	4.6%	12.3%	8.4%
Occupation			
Moved up	20.9%	25.3%	22.8%
Moved down	16.3%	18.6%	17.3%
Unchanged	62.8%	56.2%	60.0%
Partner status			
Moved from single to couple	7.4%	5.2%	6.3%
Moved from couple to single	10.7%	12.3%	11.5%
Remained part of a couple	69.4%	69.2%	69.3%
Remained single	12.4%	13.4%	12.9%
Children			
Moved from no children to children	6.9%	1.9%	4.3%
Moved from children to no children	55.1%	67.3%	61.2%
Always children	10.7%	7.4%	9.0%
Always no children	27.3%	23.4%	25.3%
Health			
Improved	14.5%	21.0%	17.7%
Worsened	54.2%	49.7%	51.9%
Unchanged	31.3%	29.3%	30.3%
Hours caring			
Increased	18.9%	27.8%	23.6%
Reduced	7.5%	9.5%	8.6%
Unchanged	73.6%	62.7%	67.9%
Region			
Changed	13.5%	8.7%	11.1%
Unchanged	86.5%	91.3%	88.9%

Notes: Calculated using the first and last observations of each individual within the relevant age category. Rows may not sum due to rounding

Source: Authors' calculations.

Table 6 documents the distribution of changes in the variables above, over the relevant part of the lifecycle, comparing the first and last observations for the prime working-age sample. This shows, for example, that 74% of men simply remained in employment,

implying that all else equal changes in work status will explain relatively little variance in men's income changes; whereas for women transitions into and out of work are more common, with about one third changing their work status between the two periods. Partnership status is one of the more stable variables over this age range (though, as we saw above, when it does change the associated income change tends to be large). Relatively common changes in circumstance over this age range include dependent children moving out of the family home and worsening health.

What explains variation in income changes?

Putting this together, for the results shown in Table 7 we use the methods outlined in Section 4 to estimate how much of the (ex post) predictable variance in log-earnings/income changes is explained by specific life changes.

The first thing to note is the bottom row of each panel in the table. A large majority of the variation is not explained by any of the observed factors we control for. This is particularly true for the evolution of individual wages/earnings, which is unsurprising given the array of shocks or changes to productivity or hours of work choices which will be difficult to explain using observed factors, as opposed to some of the more mechanical and easily-observed determinants of household income changes like partnership status. However, in none of the regressions can we explain more than about one sixth of the variance in changes in the outcome variable.

Turning to the portion of the variances that we can explain, the major factor for earnings changes among those in work is, unsurprisingly, changes in occupation. This accounts for about half of the predictable variance in earnings changes overall (and about one third for men and one half for women). Changes in whether or not dependent children are in the household are the next most important factor. This is driven by the women in the sample, for whom it explains about 14% of the ex-post predictable variance in their earnings changes – and the fact that this drives differences in the income paths *between* men and women, which affects the overall results in the top panel.

Once we move to changes in (unequivalised) household earnings and income (for all individuals, rather than just those in work), occupation changes are proportionately far less important. Overall they explain less than one quarter of the ex-post predictable variance in changes, rather than one half when looking at changes in individual earnings. Changes in own work status become important, but to a far greater extent for men, for whom it explains about one half of the ex-post predictable variance. This is essentially because men earn more when in work than women, on average, so changes in work status make a larger difference to their household income (reflected in the coefficient estimates in Tables 4 and 5). This dominates an offsetting factor: namely that men's work statuses are substantially more stable than women's, as shown in Table 6.

Changes in partnership status and/or partners' characteristics also become particularly important once we look at household-level measures of earnings or income. This is true

Table 7. Proportion of explained variance in (log)-earnings/income changes between ages of 35-44 and 50-59 due to differential changes in specific characteristics

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised net household income
All					
Work status	0%	0%	18%	21%	15%
+ Occupation	46%	49%	21%	23%	18%
Partnership status	2%	-1%	27%	52%	15%
+ partner's work status and education	4%	3%	56%	72%	32%
+ Children	21%	18%	63%	77%	78%
Health	0%	0%	0%	0%	0%
Caring responsibilities	1%	1%	0%	1%	2%
Region	3%	1%	6%	4%	9%
% of variance explained	*2%*	*3%*	*15%*	*12%*	*12%*
Male					
Work status	0%	0%	59%	42%	21%
+ Occupation	32%	34%	60%	43%	21%
Partnership status	7%	6%	8%	24%	7%
+ partner's work status and education	28%	23%	17%	52%	14%
+ Children	34%	28%	27%	61%	75%
Health	2%	1%	0%	-1%	0%
Caring responsibilities	9%	2%	0%	1%	1%
Region	4%	7%	8%	9%	12%
% of variance explained	*3%*	*2%*	*17%*	*10%*	*15%*
Female					
Work status	0%	0%	1%	10%	6%
+ Occupation	43%	56%	4%	12%	9%
Partnership status	9%	0%	41%	62%	40%
+ partner's work status and education	39%	11%	85%	79%	63%
+ Children	49%	25%	87%	82%	84%
Health	2%	0%	0%	0%	0%
Caring responsibilities	1%	1%	0%	1%	2%
Region	-1%	2%	4%	5%	6%
% of variance explained	*7%*	*7%*	*21%*	*17%*	*16%*

Source: Authors' calculations.

to a far greater extent for women, presumably largely reflecting the fact that men's individual incomes tend to be higher and therefore they have larger impacts on their partners' household incomes than vice versa. Even for men, however, the results are striking. About 50% of the predictable variance in the path of net household income across men is due to differences in the evolution of partnership status and of their partners' characteristics (work status and education). For women the figure is about 80%. Hence the relative importance of labour market risk for women looks particularly low (at least when viewed from a household perspective under the assumption of full income sharing).

Once we move to an equivalised measure of net household income, the main effect is to make the presence (or otherwise) of dependent children a much larger explanatory factor relative to others. Qualitatively this is of course a mechanical effect of the equivalence scale. Nevertheless - given that this is the most typical measure of living standards - it is striking how much of the variance in its evolution with age (46%) is due to the presence or absence of dependent children.

The transition into retirement

We also conduct the same analysis for an older cohort observed over a later stage of the lifecycle, between late working-age life (50-59) and the age where they would typically be retired (66-75). We focus here just on the household income measures, as transitions out of work limit the sample size available to look at earnings measures over this age range. Results are shown in Table 8.

Again, a clear majority of the variance in income changes is unexplained, though if anything we tend to explain slightly more than for the younger age group.

Work status changes account for more of the variation in income changes for this older age group – for both men and women – reflecting the greater number of transitions out of work over this age range than over prime working age life. In fact for men, work status changes are more important over this age range than any of the changes in family structure or partner characteristics. For women, changes in family structure or partner characteristics do remain the most important factor, explaining almost two thirds of the predictable variance in household income changes.

Table 8. Proportion of explained variance in (log)-income changes between ages of 50-59 and 66-75 due to differential changes in specific characteristics

	Household net income	Equivalised net household income
All		
Work status	48%	48%
+ Occupation (in work)	49%	50%
+ Private pension provision (out of work)	52%	52%
Partnership status	5%	-1%
+ partner's work status and education	50%	50%
+ Children	50%	53%
Health	0%	0%
Caring responsibilities	1%	0%
Region	7%	4%
% of variance explained	*17%*	*19%*
Male		
Work status	51%	50%
+ Occupation (in work)	52%	51%
+ Private pension provision (out of work)	56%	54%
Partnership status	0%	1%
+ partner's work status and education	36%	36%
+ Children	36%	39%
Health	0%	0%
Caring responsibilities	2%	2%
Region	9%	6%
% of variance explained	*17%*	*19%*
Female		
Work status	39%	40%
+ Occupation (in work)	40%	41%
+ Private pension provision (out of work)	43%	44%
Partnership status	9%	0%
+ partner's work status and education	61%	60%
+ Children	64%	65%
Health	0%	-1%
Caring responsibilities	1%	0%
Region	5%	2%
% of variance explained	*19%*	*20%*

Notes: The
Source: Authors' calculations.

Variation in occupational pension provision for those not in work (whether any such pension is being received, and if so whether it is a public or private sector pension) does relatively little to explain variance in the path of income between late working-age life and retirement. This essentially reflects small estimated coefficients on occupational pension status, which might seem surprising, and it is worth being clear about what the fixed effects regressions underlying the analysis are identifying. A natural interpretation is that people who stop working and receive an occupational pension actually see a similar proportional change in income to people who stop working and do not receive one. This is likely to reflect, at least in part, the fact that higher-paying jobs also tend to offer better pension provision.[7] Hence, even conditional on occupation, the people who get occupational pensions in retirement are more likely to have had higher earnings pre-retirement, meaning that their replacement rate is not so different on average from those without an occupational pension who also had lower earnings. This does not imply, of course, that those who do have occupational pensions would be little affected if they were taken away.

6. Conclusion

This paper provided new insight into the determinants of income changes over the lifecycle. Rather than focusing on models of the earnings or income process that assume constant household composition, we have considered a wide range of time-varying factors that could explain variation in the path of incomes, including demographic characteristics, health, caring responsibilities and occupational pension provision. We come to two key conclusions. The first is that changes in demographic characteristics (partner status and characteristics, the presence of children) account for more of the explained variation in household income changes than changes in employment status or occupation, both in prime working-age life and as individuals move into retirement. The second is that this is particularly true for women – as they are more likely to be the secondary earner in a household, changes in their labour market characteristics are relatively less important, and changes in the presence or characteristics of a partner relatively more important.

There are a number of ways in which one could seek to build on these findings. An important next step would be to attempt to distinguish between the variation explained by planned changes in demographic characteristics and that due to demographic shocks, in order to quantify the scale of demographic risk. Demographic risk could then be incorporated into a more realistic model of consumption and savings decisions, in which the model of household savings behaviour allows for different uncertainties across gender (along the lines of Browning (2000)).

[7] See Disney and Emmerson (2002)

References

Blundell, R., Pistaferri, L. and Preston, I. (2008) 'Consumption inequality and partial insurance', *The American Economic Review* 98(5), 1887-1921

Browning, M. (2000) 'The saving behaviour of a two-person household' *Scandinavian Journal of Economics* 102(2), 235-251

Burgess, S.M., Gardiner, K., Jenkins, S.P. and Propper, C. (2000) 'Measuring income risk', DIW Discussion Paper No. 213

Chiappori, P.A. and Donni, O. (2009) 'Non-unitary models of household behaviour: a survey of the literature', IZA Discussion Paper No. 4603

Crawford, R. and O'Dea C. (2014) 'Cash and Pensions: Have the elderly in England saved optimally for retirement?', IFS Working Paper W14/22

Disney, R. and Emmerson, C. (2002) 'Choice of pension scheme and job mobility in Britain', IFS Working Paper W02/09

Etheridge, B. (2015) 'A test of the household income process using consumption and wealth data', *European Economic Review* 78, 129-157

Guvenen, F. (2009) 'An empirical investigation of labor income processes', *Review of Economic Dynamics* 12, 58-79

Meghir, C. and Pistaferri, L. (2004) 'Income variance dynamics and heterogeneity', *Econometrica* 72(1), 1-32

Scholz, J.K., Seshadri, A. and Khitatrakun, S. (2006), 'Are Americans saving "optimally" for retirement?', *Journal of Political Economy* 114(4), 607-643

Appendix A: Additional results for the distribution of changes in earnings and incomes

Table A1. The distribution of changes in log earnings and income between the ages of 35-44 and 50-59 (common sample)

	Individual hourly wages	Individual earnings	Household earnings	Household net income	Equivalised net household income
All					
Mean	0.108	0.093	0.100	0.179	0.322
Median	0.164	0.175	0.167	0.224	0.345
25th percentile	-0.065	-0.153	-0.185	-0.072	0.103
75th percentile	0.415	0.506	0.557	0.505	0.608
Variance	0.436	1.085	0.767	0.281	0.266
Male					
Mean	0.026	-0.106	0.086	0.187	0.293
Median	0.110	0.083	0.176	0.228	0.330
25th percentile	-0.110	-0.241	-0.180	-0.073	0.067
75th percentile	0.342	0.311	0.560	0.510	0.583
Variance	0.541	1.120	0.731	0.298	0.288
Female					
Mean	0.207	0.334	0.117	0.170	0.358
Median	0.219	0.355	0.155	0.215	0.395
25th percentile	0.021	0.020	-0.198	-0.070	0.118
75th percentile	0.515	0.792	0.556	0.496	0.650
Variance	0.292	0.940	0.814	0.262	0.239

Appendix B: Additional detail on computation of estimates in Tables 6 and 7

Our regressions include a number of sets of dummy variables, created from a single categorical variable (so that exactly one dummy variable in the set is equal to one for each observation). We take as an example here the 6-category variable for "out of work" or "in work" and in one of 5 different occupation categories, which enters in the regression as 5 dummy variables (given the omitted category, which is "out of work"). When thinking about what explains variability in income changes across individuals, it is of interest to separate the contributions of changes in work status from changes in occupation for those who are in work.

Suppose we wanted to know how much less variability in income changes there would be in the counterfactual scenario where the occupation that workers are in makes no difference to their income (but the average level of income for both workers and non-workers, and hence the mean difference between them, is preserved). In reality, we postulate that

$$\Delta y_i = \sum_{o=1}^{5} \alpha^o \Delta Occ_i^o + \delta \Delta Z_i + \Delta \varepsilon_i$$

where $\alpha = [\alpha^1, \dots, \alpha^5]$ represents the coefficients on the 5 occupation categories $[Occ^1, \dots, Occ^5]$, and $X_i = [Occ_i{}^1, \dots, Occ_i^5, Z_i]$. But under the counterfactual scenario,

$$\Delta y_i^C = \bar{\alpha} \sum_{o=1}^{5} \Delta Occ_i^o + \delta \Delta Z_i + \Delta \varepsilon_i$$

where $\bar{\alpha}$ is the weighted average of $[\alpha^1, \dots, \alpha^5]$, with weights equal to the proportion of observations of people in work who are in each occupation.

This implies that, for individual i, the predicted change in log-earnings/income in the counterfactual scenario can be computed as:

$$\Delta \hat{y}_i^C = \Delta \hat{y}_i - \sum_{o=1}^{5} (\alpha^o - \bar{\alpha}) \Delta Occ_i^o .$$

Intuitively, we purge $\hat{y}_i$ of the component explained by the fact that specific occupation groups earn different (covariate-adjusted) amounts from the average – hence the $(\alpha^o - \bar{\alpha})$ term; but we retain the average difference between those in work and those not in work. The variance of $\Delta \hat{y}_i^C$ can then be compared with the variance of $\Delta \hat{y}_i$ to assess how much of the predictable variability in log-earnings/income changes is accounted for by differences between occupations, holding work status constant.

Appendix C: Supplementary tables

Table C1. Coefficients from fixed effects regressions of log income between ages of 50-59 and 66-75: all

	Household net income	Equivalised household net income
Age – 35	-0.011	0.009
Age – 35, squared	0.000	-0.000
Employment, occupation and private pension		
No work, between 55 and SPA	0.067	0.070
No work, over SPA, no private pension	-0.028	-0.023
No work, over SPA, public sector pension	0.015	0.012
No work, over SPA, private sector pension	-0.077	-0.067
Unskilled	0.163***	0.155***
Partly skilled	0.244***	0.253***
Skilled manual	0.287***	0.283***
Skilled non-manual	0.241***	0.249***
Professional	0.268***	0.264***
Demographic status		
Couple without children, partner out of work	0.204**	-0.025
Couple without children, working low-educated partner	0.388***	0.165**
Couple without children, working mid-educated partner	0.443***	0.210***
Couple without children, working degree-educated partner	0.450***	0.201***
Lone parent	-0.279**	-0.312***
Couple with children, partner out of work	0.155	-0.127
Couple with children, working low-educated partner	0.273***	-0.074
Couple with children, working mid-educated partner	0.235**	-0.104
Couple with children, working degree-educated partner	0.374**	-0.091
Health index	0.009	0.009
Care status		
Caring less than 20 hours	-0.011	0.004
Caring at least 20 hours	0.036	0.035

Notes: Omitted categories are not employed and younger than 55, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Table C2. Coefficients from fixed effects regressions of log income between ages of 50-59 and 66-75: male

	Household net income	Equivalised household net income
Age – 35	-0.019	-0.009
Age – 35, squared	0.001*	0.000
Employment, occupation and private pension		
No work, between 55 and SPA	0.059	0.061
No work, over SPA, no private pension	-0.044	-0.050
No work, over SPA, public sector pension	-0.016	-0.000
No work, over SPA, private sector pension	-0.101	-0.091
Unskilled	0.148	0.136*
Partly skilled	0.221***	0.242***
Skilled manual	0.288***	0.294***
Skilled non-manual	0.262***	0.283***
Professional	0.258***	0.265***
Demographic status		
Couple without children, partner out of work	-0.000	-0.138*
Couple without children, working low-educated partner	0.174	0.037
Couple without children, working mid-educated partner	0.173	0.027
Couple without children, working degree-educated partner	0.304**	0.113
Lone parent	-0.168	-0.092
Couple with children, partner out of work	-0.055	-0.246**
Couple with children, working low-educated partner	0.056	-0.218**
Couple with children, working mid-educated partner	-0.051	-0.283***
Couple with children, working degree-educated partner	0.103	-0.154
Health index	0.011	0.006
Care status		
Caring less than 20 hours	0.002	0.008
Caring at least 20 hours	0.087**	0.073**

Notes: Omitted categories are not employed, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Table C3. Coefficients from fixed effects regressions of log income between ages of 50-59 and 66-75: female

	Household net income	Equivalised household net income
Age – 35	0.012	0.038
Age – 35, squared	-0.000	-0.001*
Employment, occupation and private pension		
No work, between 55 and SPA	0.048	0.044
No work, over SPA, no private pension	-0.043	-0.035
No work, over SPA, public sector pension	0.046	0.010
No work, over SPA, private sector pension	-0.022	0.005
Unskilled	0.153**	0.149**
Partly skilled	0.252***	0.252***
Skilled manual	0.233***	0.220***
Skilled non-manual	0.202***	0.197***
Professional	0.263***	0.249***
Demographic status		
Couple without children, partner out of work	0.316***	0.021
Couple without children, working low-educated partner	0.509***	0.223**
Couple without children, working mid-educated partner	0.625***	0.334***
Couple without children, working degree-educated partner	0.503***	0.231**
Lone parent	-0.319**	-0.390***
Couple with children, partner out of work	0.242	-0.101
Couple with children, working low-educated partner	0.469***	0.117
Couple with children, working mid-educated partner	0.550***	0.055
Couple with children, working degree-educated partner	0.775***	-0.018
Health index	0.009	0.013
Care status		
Caring less than 20 hours	-0.024	-0.004
Caring at least 20 hours	-0.001	0.003

Notes: Omitted categories are not employed, single without children, and not caring. Unskilled is omitted in the first two columns because all individuals in the sample are employed.
Source: Authors' calculations.

Table C4. The distribution of changes in time-varying characteristics between the ages of 50-59 and 66-75

	Male	Female	All
Employment			
Moved into employment	0.64%	0.59%	0.62%
Moved out of employment with no private pension	7.40%	23.01%	60.31%
Moved out of employment with private sector pension	40.84%	17.70%	28.77%
Moved out of employment with public sector pension	17.36%	14.75%	16.00%
Remained in employment	16.72%	7.96%	12.15%
Remained out of employment	17.04%	35.99%	26.92%
Partner status			
Moved from single to couple	1.59%	0.87%	1.22%
Moved from couple to single	8.57%	13.12%	10.94 %
Remained part of a couple	80.63%	64.72%	72.34%
Remained single	9.21%	21.28%	15.50%
Children			
Moved from no children to children	0%	0%	0%
Moved from children to no children	15.56%	5.54%	10.33%
Always children	0.32%	0%	0.15%
Always no children	84.13%	94.46%	89.51%
Health			
Improved	20.86%	17.47%	19.09%
Worsened	60.93%	65.66%	63.41%
Unchanged	18.21%	16.87%	17.51%
Hours caring			
Increased	20.66%	20.08%	20.37 %
Reduced	13.22%	22.18%	17.67%
Unchanged	66.12%	57.74 %	61.95%
Region			
Changed	10.16%	11.08%	10.64%
Unchanged	89.84%	88.92%	89.36%

Notes: Calculated using the first and last observations of each individual within the relevant age category
Source: Authors' calculations.