NBER WORKING PAPER SERIES

WAGE RISK AND GOVERNMENT AND SPOUSAL INSURANCE

Mariacristina De Nardi
Giulio Fella
Gonzalo Paz-Pardo

Working Paper 28294
http://www.nber.org/papers/w28294

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2020, Revised March 2023

We thank Richard Blundell, Mike Daly, Robert Joyce, Peter Levell and Hamish Low for useful comments and suggestions, and Barra Roantree for generously sharing his knowledge of the NESPD dataset. The project has been funded by the Nuffield Foundation (visit www.nuffiedfoundation.org), but the views expressed are those of the authors and not necessarily of the Foundation, the NBER, the CEPR, any agency of the federal government, the Federal Reserve Bank of Minneapolis, the European Central Bank, or the IFS.

Wage Risk and Government and Spousal Insurance
Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz-Pardo
NBER Working Paper No. 28294
December 2020, Revised March 2023
JEL No. D1,D12,D14,D15,H11,H2

ABSTRACT

The extent to which households can self-insure depends on family structure and wage risk. We calibrate a model of couples and singles' savings and labor supply under two types of wage processes. The first wage process is the canonical—age independent, linear—one that is typically used to evaluate government insurance provision. The second wage process is a flexible one. We use our model to evaluate the optimal mix of the two most common types of means-tested benefits—in-work versus income floor. The canonical wage process underestimates wage persistence for women and thus implies that in-work benefits should account for most benefit income. In contrast, the richer wage process that matches the wage data well, implies that the income floor should be the main benefit source, similarly to the system in place in the UK. This stresses that allowing for rich wage dynamics is important to properly evaluate policy.

Mariacristina De Nardi
University of Minnesota
1925 S 4th St
Minneapolis, MN 55455
and Federal Reserve Bank of Minneapolis
and CEPR
and also NBER
denardim@nber.org

Giulio Fella
Queen Mary University of London
Mile End Road
London E1 4NS
United Kingdom
and Centre for Macroeconomics
and Institute for Fiscal Studies
g.fella@qmul.ac.uk

Gonzalo Paz-Pardo
European Central Bank
Sonnemannstrasse 20
60314
Frankfurt am Main
Europe
gonzalo.paz_pardo@ecb.europa.eu

1 Introduction

The necessity, efficacy, and cost-effectiveness of government welfare policies depends on the risks that households face and the actions that they can take to self-insure, for instance by adjusting their saving and labor supply. Wage risk is a key driver of household risk and being single rather than in a couple is an important factor affecting both a household's sources of risk and tools for self-insurance. This is because single people are solely exposed to their own wage risk and can only use their own savings and labor supply to smooth consumption and welfare fluctuations. In contrast, couples face the wage risk of both household members but can use their joint savings and the labor supply of both partners to at least partly counteract individual wage fluctuations. In addition, couples benefit from economies of scale in consumption.

Better understanding the dynamics of wage and earnings risk is key to study the ability of households to self-insure and to properly design an efficient benefit system. In addition, explicitly modeling couples and singles, as well as the dynamics of fertility and saving over the life-cycle, is crucial to understand how wage and earnings risks interact with self-insurance depending on family structure.

We begin our analysis by studying both UK survey data from the British Household Panel Study (BHPS), at the household level, and UK administrative data from the New Earnings Survey Panel Dataset (NESPD), at the individual level. We find that the individual-level earnings and wage dynamics that we observe in these data sets are remarkably similar, and that they display dynamics that are substantially richer than those implied by the canonical linear model (see MaCurdy, 1982; Abowd and Card, 1989)typically used for policy evaluation. Thus, we propose a much richer model for wage risk that, unlike the canonical model, allows for the distribution of wage shocks to be non-normal and for wage risk to vary by age and worker's rank in the wage distribution. This richer process can capture, for instance, that shocks are less persistent for younger and lower-income workers.

Our analysis shows that the canonical process, which imposes more restrictive assumptions that are at odds with the UK data, overestimates wage persistence for men, and underestimates it for women. Compared to the previous literature, our contribution in this part of our analysis is to estimate wage, rather than earnings, dynamics and to estimate both canonical and richer processes, for both men and women. Looking at wage dynamics is important because earnings are endogenous to the choice of hours worked. Allowing for heterogeneity in gender and family structure is important as single and married men and women have different labor supply behavior.[1]

We then develop a dynamic, structural, life-cycle model with an active female labor-supply decision at the extensive margin. The model features a rich menu of sources of heterogeneity. Individuals differ in gender, marital status, number of children, and wage realizations. We account for the presence of children across married and single households, the timing of their arrival, as well as marital transitions. We calibrate our model under the canonical and nonlinear wage processes described above and use it to evaluate the optimal provision of two important types of government transfers, an income floor and in-work benefits, as well as the rate at which benefits should phase-out as a function of labor income. Our calibration matches key aspects of the data that include government policy and household labor market outcomes over the life cycle, during the time period preceding the 2016 Universal Credit benefit reform in the UK.

We find that, while the model fits key aspects of the observed data under both wage processes, their optimal policy implications are starkly different. While in both cases the optimal reform involves halving—from 1.1 to 0.5—the rate at which benefits phase-out with labor income, the mix of the two benefits is very different under the two systems. The optimal benefit configuration under the richer wage process is similar to the one that was in place during the period preceding the Universal Credit reform. It privileges the income floor with a very limited role for in-work benefits. In contrast, if one were to assume a canonical wage process, one would conclude that optimal benefits during the same period should have been very different. In particular, that optimal policy would incorrectly prescribe a trebling of in-work benefits and effectively eliminate income support. The intuition for the difference is that the canonical wage process underestimates the persistence of shocks to women's wages relative to the richer process, and thus implies that it is less costly to induce women to participate in the labor market by lowering their out-of-work benefits and increasing their in-work benefits. In reality, women's wages are

[1] Guvenen, Karahan, Ozkan and Song (2021) document rich dynamics for pre-tax individual earnings in the US, Arellano, Blundell and Bonhomme (2017) for household pre-tax earnings in the US and Norway, De Nardi, Fella and Paz-Pardo (2020) for household disposable earnings in the US. Ozkan, Storesletten, Holter and Halvorsen (2017) and De Nardi, Fella, Knoef, Paz-Pardo and Van Ooijen (2021) study the relative contribution of wages and hours to male earnings dynamics respectively in Norway, and in the Netherlands and the US.

more persistent and thus such a reform would have negative impact on the welfare of a subset of persistently low-income women with high costs of labor market participation (which could be related, for example, to health issues), and would be pushed into low-paid work by the reform.

We also use the model to study the Universal Credit benefit reform that was subsequently introduced in the UK in 2016 and completed by the end of 2018. Our model with endogenous savings is particularly well suited to study this reform, which, in addition to introducing an earnings disregard for households with children, generalized asset means testing for benefit eligibility in the UK. We find that, irrespective of the wage process the move to Universal Credit implies overall welfare gains, but significantly reduces welfare for single men.

Because many women don't work and the relevant wage dynamics for our analysis are the potential ones, rather than those observed just for labor market participants, we infer the distribution of potential wages for all women, whether working or not, from the data. To recover potential wages, we impute them for non-working women by using a state-of-the-art Heckman selection procedure that uses a measure of potential out-of-work welfare income (potential benefit income for the household if the woman were not working, conditioning on family circumstances, geographic location, and yearly variation in policies) as an instrument. This approach has been previously adopted by Blundell, Reed and Stoker (2003), Arellano and Bonhomme (2017), and Chiappori, Costa-Dias and Meghir (2018). However, we also evaluate the robustness of our results to alternative imputation procedures and show that, while the nonlinear and canonical processes imply distinctly different optimal benefit systems, the differences generated by our alternative wage imputation procedures are minor.

Our work builds on the important, but still relatively small, literature that studies the effects of taxation and welfare policies taking into account household composition. A robust finding of this literature is the importance of accounting for the response of female labor supply. Keane and Wolpin (2010) study the effect of the US welfare system on women's welfare participation, labor supply, marriage, fertility, and schooling. Blundell, Costa Dias, Meghir and Shaw (2016) study how the UK tax and welfare system affects the career choices of women. Guner, Kaygusuz and Ventura (2012) and Bick and Fuchs-Schündeln (2017) investigate the effect of taxation on household labor supply, while

Guner, Kaygusuz and Ventura (2011) evaluate gender-based taxes, Nishiyama (2019) and Groneck and Wallenius (2017) evaluate Social Security spousal provisions, and Borella, De Nardi and Yang (2021) study the effects of marriage-related taxes and Social Security rules for different cohorts of women whose labor supply behavior has been changing. The paper closer in spirit to ours is possibly Guner, Kaygusuz and Ventura (2020) that compares the implications of in-work childcare credits to those of child benefits independent of the mother's labor market participation. Our focus is instead on benefits other than child-related ones and we allow for marital transitions. Furthermore, none of these papers allows for the richer wage dynamics that we observe in the data.

2 Earnings and wage risks

For tractability, and because most men work full time and display very small labor supply elasticities, we take men's labor supply as exogenous while we model women's labor supply. Thus, in our empirical analysis, we study men's earnings and women's wages.

Our main data source is the British Household Panel Survey (BHPS). The BHPS is a household survey of the UK population that started in 1991 by sampling 5,500 households and 10,300 individuals, and then followed them and their children over time. Its design suggests that its measurement error in self-reported earnings is likely to be lower than in other surveys, such as the PSID in the US, because instead of being asked about their total labor earnings in the last twelve months, respondents were asked to check their last pay slip and report about it. Furthermore, in a relevant proportion of the observations (around 30%), the interviewer saw the pay slip. An important advantage of the BHPS is that, in addition to income data, it includes rich information, including off-sample labor market histories. Furthermore, it collects information on all household members, and is thus suited for the study of family and government insurance. This is important because even though taxation in the UK is at the individual level, most subsidies and benefits are at the household level. Since 2008, the BHPS has been replaced by the wider Understanding Society survey, which kept most of its panel component. We provide more information about the BHPS in online Appendix A.1.

Our sample is composed of individuals between the ages of 25 and 60, which restricts attention to individuals who have completed full-time education and have not yet retired. The online Appendix A.2 details our requirements for sample construction, which are in line with most of the literature on earnings dynamics. The most important difference is that, rather than excluding individual observations below a minimum earnings threshold as typically done, we bottom-code men's earnings observations below the threshold.[2] This allows us to take into account the most negative outcomes that workers may face, such as staying out of work for a very long time, and for which government insurance might be particularly valuable. Our bottom coding is low enough (around £100 per month) to capture the high marginal utility of consumption of people in this situation.

Our earnings/wage measure is the residual obtained by regressing the logarithm of earnings on year and age dummies. Most of the moments that we present refer to changes in residual log-earnings/wages. This leaves us with 42,659 usable observations (pairs of earnings in t and $t+1$) for men and 43,198 for women.

We start by documenting the properties of male, pre-tax earnings in the BHPS by using a set of moments that has become rather standard in the literature.[3] The top left panel of Figure 1 plots the standard deviation of male earnings changes against the percentile of last period's earnings. The standard deviation follows a U-shaped pattern which is inconsistent with the assumption of linearity that underpins the canonical model.

The top right and bottom left panels plot the skewness and kurtosis of male earnings changes. The skewness is positive for low realizations of previous earnings and falls as one moves to the right in the distribution of previous earnings, becoming negative above the median. The kurtosis is somewhat higher than its value of 3 for the normal distribution, but overall UK male earnings display substantially smaller from normality than those found in the studies for other countries that we quote in footnote 1.

The bottom right panel plots the persistence of male earnings as a function of age and percentile of the previous earnings realization. As the moments discussed above, persistence is not independent of previous earnings levels (or age) which again is inconsistent with the linearity of the canonical model. More specifically, the picture shows that the

[2] The typical threshold is around 5 per cent of median earnings, which corresponds to £1,300 at 2015 constant prices in our dataset.

[3] To ease potential concerns about measurement error in the BHPS, online Appendix A.3 compares our findings with those from the New Earnings Survey Panel Dataset (NESPD), an administrative data set with individual data from the UK Social Security. It shows that the results from NESPD are very similar to those from the BHPS.

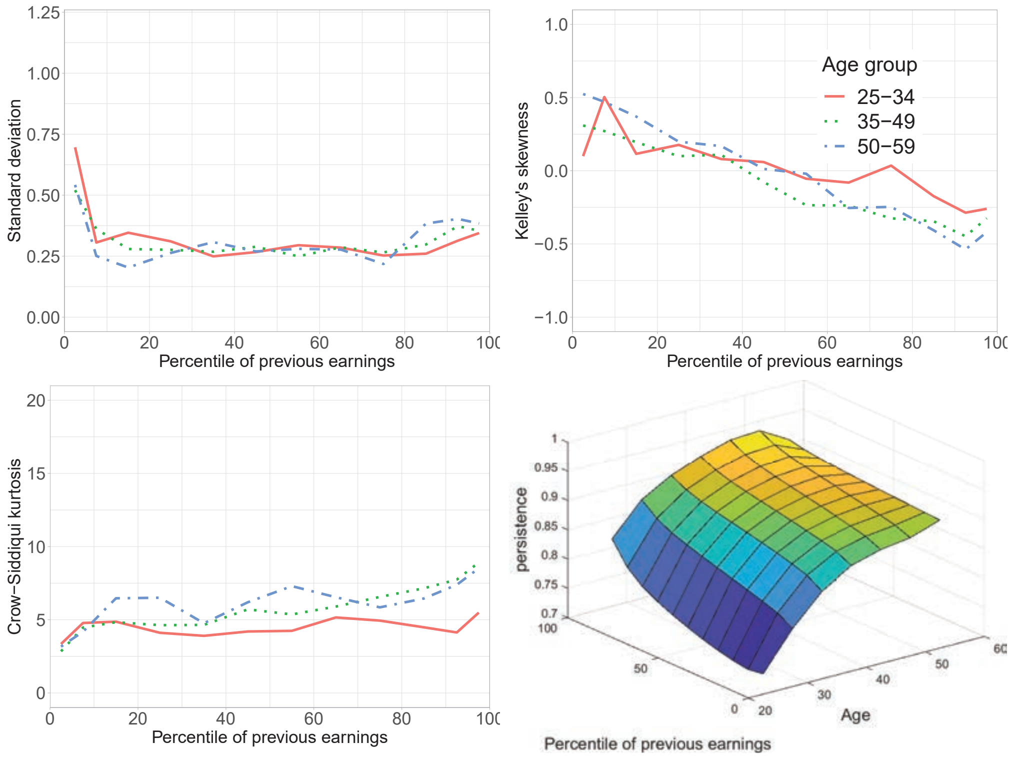


Figure 1: Moments of male earnings changes in BHPS data

persistence of male earnings is lowest at young ages and low earnings levels, is about 0.7.

Turning to observed female wages, the first three panels of Figure 2 plot the variance, skewness and kurtosis of observed female wage changes for labor market participants as a function of the rank of the previous period's realization. Their properties are remarkably similar to those of male earnings changes: the variance has a U-shaped pattern, skewness is positive below the median and declines with the rank of previous earnings and kurtosis is higher than for the normal distribution, but not too much so.

The bottom right panel of the same picture, instead, plots the persistence of female wages as a function of age and the percentile of the previous wage realization. Similarly to male earnings, the pattern of persistence is inconsistent with the standard, linear canonical model. Persistence is hump shaped as a function of the previous realization, though it displays much less variability with respect to age than in the case of male earnings.

These pictures make it apparent that both male earnings and female wages display strong deviations from the assumption of linearity underpinning the canonical model.

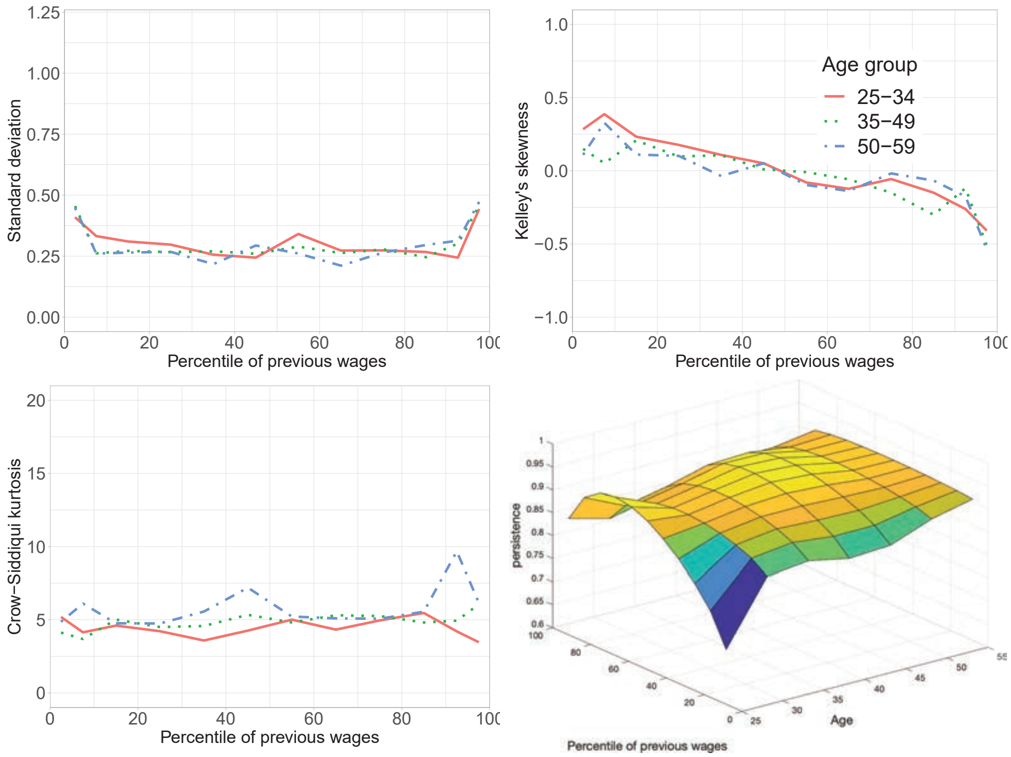


Figure 2: Moments of female wage changes in BHPS data

2.1 Estimating the distribution of potential female wages

There is an extensive empirical literature that finds that the elasticity of women's labor market participation is sizeable (see, for instance, Meghir and Phillips (2010)). This indicates that endogenous selection is likely important, and that the distribution of observed wages may differ from the distribution of potential wages. The latter is the relevant input for our structural model, in which women's labor supply decisions are endogenous. It is therefore crucial to capture it appropriately.

To recover the distribution of potential wages, we impute wages for women who are not currently working. In our preferred specification, we do so using a control function approach that allows for endogenous selection, as in Heckman (1979). More specifically, we use a reduced-form, binary choice probit model of employment as a function of a single index γZ_{it} and we estimate the following equation

$$\log w_{it} = u_i + \beta X_{it} + \lambda_P(\gamma Z_{it}) + \epsilon_{it} \tag{1}$$

where w_{it} is the wage of woman i working in year t, u_i an individual-specific fixed effect, X_{it} a set of covariates, and $\lambda_P(\gamma Z_{it})$ is the control function which we approximate with the inverse Mills ratio from the employment probit. The covariates X_{it} include the number of children, the number of children under 5, a college dummy, a marital status dummy, a birth decade dummy, and a year polynomial. The control variables Z_{it} in the employment probit including the same explanatory variables as in the wage imputation equation (1), except for the fixed effect, and a measure of potential out-of-work welfare income (potential benefit income for the household if the woman were not working, conditioning on family circumstances, geographic location, and yearly variation in policies). The exogenous variations in the benefit system through tax and benefit reforms, together with the nonlinear features of the tax system (which might strongly affect participation at the thresholds) and the large cross-sectional variation, make this a particularly suitable instrument for our purposes.[4] We use the fitted values from Equation (1) to impute (potential) wages for women who are not employed in a given year.

We evaluate the robustness of our findings under this baseline imputation procedure (**H-Ben**) with those for the following three alternatives. Two of them use alternative Heckman selection corrections based on a different set of excluded instruments entering Z. In particular, the first one, which we denote by **H**, uses homeownership status, a college dummy, marital status, a decade of birth dummy, and interactions of the decade of birth dummy and marital status as instruments. The second one, which we label **H-Child**, adds dummies for the years that have passed since the birth of the first child, interactions of those with marital status, a dummy for husband employment and a dummy for whether grandparents are present in the household to the set of excluded instruments in **H**. Finally, the third imputation, denoted by **FE**, uses a richer set of controls X, rather than a control function approach

Online Appendix A.5 reports the details of these imputation procedures and their results. It also shows that the implied average earnings profiles by age and the distributions of potential wages are very similar across all imputation procedures.

[4]Blundell et al. (2003) were the first to employ potential welfare income as an instrument which affects wages only through Z in the control function but not directly through X in equation (1). The same restriction is also used in Arellano and Bonhomme (2017) and Chiappori et al. (2018).

2.2 Estimating the wage and earnings processes

Our structural model of household behavior requires that we estimate the stochastic processes for male earnings and female wages. In this section, we describe our assumptions about these processes and how we estimate them.

Consider a cohort of individuals indexed by i and denote by g the individual's gender, p marital status, and t age. We assume that the logarithm of the potential wage $\tilde{w}_{it}^{gp}$, net of time fixed-effects, can be decomposed into a deterministic age component η_t^{gp} and a stochastic component y_{it}^{g} according to

$$\log \tilde{w}_{it}^{gp} = \eta_t^{gp} + y_{it}^{g}. \tag{2}$$

We assume that the stochastic component does not depend on marital status because, as we show in the online Appendix A.4, the features of male earnings changes and female wages are very similar for singles and married individuals.

For men, the potential wage in equation (2) is actual, measured earnings. This is because we abstract from the labor supply margin, which is much more important for women than for men. For women, the potential wage in year t is the actual, measured wage for women in employment and the wage imputed on the basis of equation (1) for the others.

In what follows, we omit the gender superscript to streamline notation. We estimate two alternative processes for the stochastic wage component y_{it} from equation (2). Both assume that it can be decomposed into a *persistent* shock that follows a first-order Markov process, z_{it}, and a *transitory* shock that is independently distributed over time, ϵ_{it}

$$y_{it} = z_{it} + \epsilon_{it.} \tag{3}$$

The *canonical* (linear) model assumes that the two components follow

$$\begin{gathered} z_{i,t} = \rho z_{i,t-1} + \nu_{it}, \qquad (4) \\ z_{i1} \overset{id}{\sim} N(0, \sigma_{z_1}), \;\; \nu_{it} \overset{iid}{\sim} N(0, \sigma_{\nu}), \;\; \epsilon_{it} \overset{iid}{\sim} N(0, \sigma_{\epsilon}). \end{gathered}$$

Our flexible, or nonlinear (NL), process is taken from Arellano et al. (2017) (ABB in what follows) and does not impose linearity or any parametric distributional assumption.

Formally, let $Q_\eta(q|\cdot)$, the conditional quantile function for the variable η, denote the qth conditional quantile of η.[5] The flexible, unrestricted, counterpart of the canonical process can then be written as

$$z_{i,t} = Q_z(v_{it}|z_{i,t-1}, t) \tag{5}$$
$$z_{i1} = Q_{z_1}(u_{it}), \;\; \epsilon_{it} = Q_\epsilon(e_{it}), \quad v_{it}, u_{it}, e_{it} \sim U(0,1).$$

The canonical process with normally-distributed shocks in equation (4) obtains when the quantile function for z specializes to the linear form $Q_{z_{i,t}}(v_{it}|z_{i,t-1}, t) = \rho z_{i,t-1} + \sigma_\nu \phi^{-1}(v_{it})$ and $Q_\epsilon(e_{it}) = \sigma_\epsilon \phi^{-1}(e_{it})$, where $\phi^{-1}(.)$ is the inverse of the cumulative density function of a standard, normal distribution.

Comparing equations (4) and (5) makes clear that the canonical process imposes constant persistence (linearity), age-independence, and normality. As we have discussed in Section 2, these assumptions are inconsistent with the earnings and wage data in the BHPS and NESPD. Instead, the methodology proposed by ABB is fully flexible along all these dimensions. We provide more details about the NL earnings process and its estimation in Online Appendix B.

We take out time and age effects before estimating our processes for residual earnings y_{it}. We estimate the canonical earnings process following the procedure described in Storesletten, Telmer and Yaron (2004), which implies fitting the parameters of interest (persistence of the persistent component ρ, variance of the persistent shocks σ_ν, variance of the initial realization σ_{z_1}, and variance of the transitory component σ_ϵ) to the profile of variances and autocovariances of log earnings over the life cycle. Table 2 shows the estimated parameters for male earnings and female wages for the canonical process. To estimate the flexible non-linear process, we follow Arellano et al. (2017). Online Appendix B.3 shows how the persistent component preserves the non-normal and non-linear features of interest of the earnings and wage data that we have described in Section 2.

Figure 3 reports the fit of the profile of variances of log earnings for men and log wages for women over the life cycle that are implied by both processes in the BHPS data. The canonical process aims at fitting these profiles by construction, while the NL process

[5] Intuitively, the conditional quantile function is the inverse of the conditional cumulative density function of the variable η mapping from the $(0, 1)$ interval into the support of η. Namely, $\eta_q = Q_z(q|\cdot)$ satisfies $P[\eta \leq \eta_q|\cdot] = q$, where $P[\cdot|\cdot]$ denotes the conditional probability.

Group	σ_ϵ^2	$\sigma_{z_1}^2$	σ_ν^2	ρ
Men earnings	0.1187	0.3827	0.0062	1.000
Women's wages	0.0106	0.1283	0.0597	0.861

Table 1: Estimates for the canonical processes.

achieves this result by matching the whole conditional distribution of y_{t+1} given y_t at every age. Figure 4 compares the estimated second moments for the two processes and shows that they have economically meaningful and statistically significant differences.

Comparing figures 3 and 4 reveals that, in the case of male earnings, the canonical process matches the increase in variance later in life through a unit root in the persistent component. The NL process, instead, captures this increase through a progressive rise in the persistence of the persistent component of earnings, coupled together with a large increase in variance of shocks and decrease in persistence at older ages (Figure 4, left panels). In the case of women's wages, the NL process captures the hump-shape in the variance of wages through a combination of relatively high persistence and low and decreasing variance of shocks to the persistent component (Figure 4, right panels). The, age-independent, canonical process cannot, by construction, generate a decreasing age profile in the variance of shocks. So it fits the profile as best as it can by matching the upward sloping part through a relatively low persistence and a high variance of shocks to the persistent component relative to the variance of the initial condition. Thus, the canonical process not only does not replicate the set of important facts about earnings risk that we have described, such as non-normalities or non-linearities, but as a result of its restrictive assumptions, it also generates implications for the profile of persistence and variance over the life cycle that are at odds with the data.

The differences in the estimated persistence of shocks implied by the two methods are potentially important, not only from a statistical, but also from an economic perspective. More persistent shocks are more difficult to self-insure through household borrowing and therefore imply a bigger role for complementary forms of insurance, such as public insurance. Our findings suggest that the canonical process overestimates labor income risk for men and underestimates it for women. This raises the question of the extent to which these differences are important for the evaluation of welfare policies aimed at providing insurances against income risk. It is this question that we address in the second

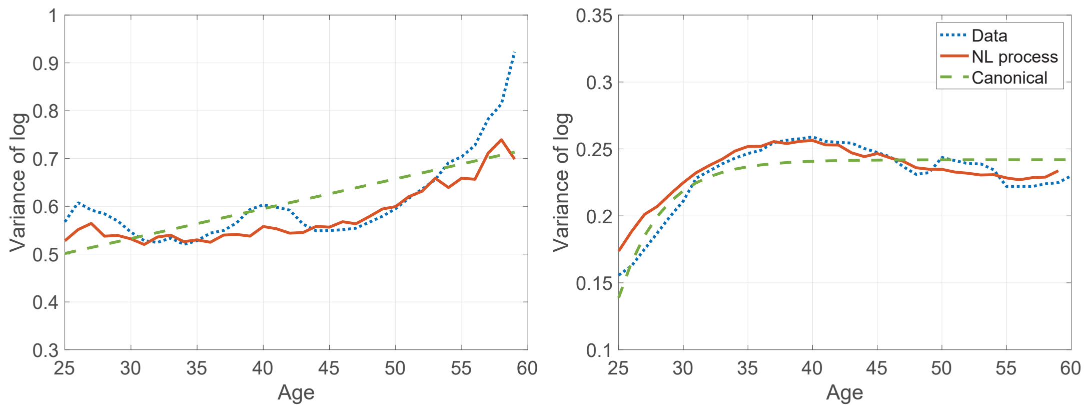


Figure 3: Variance of log earnings for men (left) and log wages for women (right). Dotted line: data. Solid line: nonlinear processes. Dashed line: canonical process.

part of the paper.

Robustness with respect to the distribution of potential wages. Table 2 shows that various imputation procedures yield very similar estimated parameters for the canonical process. While we do not report the coefficients associated with the Arellano et al. (2017) estimation for each of the alternative imputations, they are available in the online replication package. Importantly, we will show later that, while the nonlinear and canonical process imply distinctly different optimal benefit systems, the differences implied by our alternative imputation procedures are minor.

Group	σ^2_ϵ	$\sigma^2_{z_1}$	σ^2_ν	ρ
Women's wages, **H-Ben**	0.0106	0.1283	0.0597	0.861
Women's wages, **H**	0.0143	0.1285	0.0533	0.877
Women's wages, **H-Child**	0.0143	0.1288	0.0533	0.877
Women's wages, **FE**	0.0143	0.1289	0.0533	0.877

Table 2: Estimates for the canonical processes under alternative imputation procedures for potential wages.

3 Our model

We develop a discrete time, partial-equilibrium, life-cycle, incomplete-markets model in the tradition of Bewley (1977). Individuals start their economic life at age 25, which allows us to take education decisions as given. They do so with no wealth and with a given

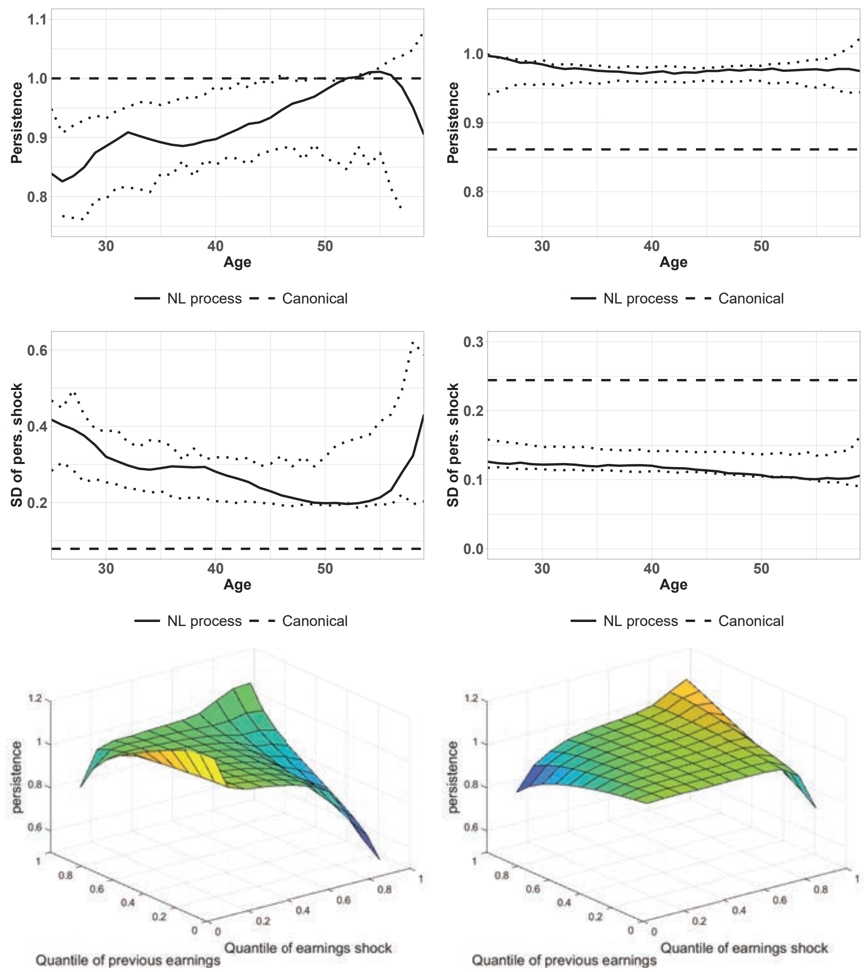


Figure 4: Features of the persistent component of male earnings (left) and women's wages (right) with the NL and canonical processes. Top: persistence by age; middle: standard deviation of persistent shocks; bottom: persistence by level of earnings and quantile of the shock. The dotted bands represent bootstrapped 95% confidence intervals.

gender, marital status, number of children, and wage shock. Men face earnings shocks and women wage shocks. There are two alternative processes describing the dynamics of earnings and wage shocks, the canonical and nonlinear one, which we have described in the previous section.

Marital status evolves stochastically as in Cubeddu and Ríos-Rull (2003). The probability of marriage and divorce depends on one's age and wage. Singles marry another single of the same age and opposite gender. Wealth is pooled upon marriage and divided equally upon divorce.

Children are born stochastically to single and married women. The probability that children are born into or leave a household depends on their mother's age, marital status, and the number of children already in the household. Children increase household consumption needs, entail child care costs if their mother works, and affect benefit eligibility.

For simplicity, we assume that people retire exogenously. Retired people face a mortality risk that depends on gender, age, and marital status. People die with probability one at age 95. There are no annuity markets to insure against mortality risk.

During each period, households choose how much to consume and save in a risk-free asset subject to a borrowing limit. Individuals have a total time endowment which is normalized to 1. Men of working age supply $\bar{h}$ hours of work inelastically, where this amount corresponds to full time work. Women, instead, optimally choose among three possible levels of working hours $\{0, \bar{h}/2, \bar{h}\}$ and bear a fixed time-cost of working which is meant to capture commuting, time spent getting ready for work, and the disutility of work.

In what follows, t denotes age, $g = f, m$ denotes gender, and $p = s, c$ indicates marital status (single or couple).

3.1 Preferences and wages

Preferences are time-separable and β is the household's discount factor. An individual's utility function is given by

$$u(c, l) = \frac{((c/\nu)l)^{1-\gamma}}{1-\gamma},$$

where c denotes total household consumption, l is leisure and ν is the equivalence scale (which depends on marital status and number of children). For men, leisure is exogenously given by $\bar{l} = 1 - \bar{h}$. For women, l^f is an endogenous leisure choice. Couples maximize the sum of their individual utilities in a unitary fashion.

The fixed time cost of working for men is normalized to zero. The fixed time-cost of working for women, $\Psi^p(h_t, t; \theta)$, depends on marital status $p = \{s, c\}$, whether working part-time or full-time h_t, age t, and one's permanent, unobserved, individual heterogeneity $\theta = \{\theta_1, \theta_2\}$. It is given by

$$\Psi^p(h_t, t, \theta) = \mathbf{I}_{h>0} \left\{ \theta + \frac{\exp(\psi_0^{p,h} + \psi_1^{p,h} t + \psi_2^{p,h} t^2)}{1 + \exp(\psi_0^{p,h} + \psi_1^{p,h} t + \psi_2^{p,h} t^2)} \right\}, \tag{6}$$

with $\mathbf{I}_{h>0}$ an indicator function equal to 1 when hours worked are positive and zero otherwise. The heterogeneity in θ is a simple way to capture differences in preferences for work and leisure across the population. It parsimoniously represents additional forces, such as poor health, that impact the marginal cost of work and which are persistent. Leisure for women is thus given by

$$l_t^f = 1 - h_t^f - \Psi^p(h_t^f, t, \theta).$$

The wage of an individual of gender g and marital status p follows the processes for the persistent component z in Equations 2 and 3. That is it follows either the canonical or the NL process. To capture assortative mating, the initial realization of wife's wages and husband's earnings are correlated, both for couples that start working life together, and for those marrying later. Additionally, we allow for the shocks to husband's earnings and wife's wages to be correlated (with correlation ρ_{HW}) for the duration of their marriage.

3.2 Child care costs

Child care costs depend on mother's age t, marital status p, working hours h_t^f, and number of children living in the household n_t. To take into account the fact that children younger than age 5 are not yet in school, and children older than 5 are, but require child care outside of school hours at least until age 11, we specify the following child care cost function

$$CC_t(p, h_t^f, n) = [n_{04}(p,t,n)h_t^f + n_{511}(p,t,n)\max(h_t^f - sc_h, 0)] \times f \tag{7}$$

where the numbers of children aged 0 to four, $n_{04}(p,t,n)$, and 5 to 11, $n_{511}(p,t,n)$, are a deterministic function of age, marital status, and the total number of children in the household, f is the hourly cost of child care and sc_h is the length of the school day.

3.3 The government

The government taxes individuals according to the Gouveia and Strauss's (1994) tax function

$$\frac{T(y)}{y} = \tau - \tau(sy^{\zeta} + 1)^{\frac{-1}{\zeta}}, \tag{8}$$

where $y = wh$ is taxable individual labor earnings and τ, and s and ζ are parameters that we estimate.

The government provides benefits that depend on household labor income. We consider two alternative structures for the benefit system. The first structure includes two types of benefits: an income floor or income support (IS) which is not conditional on working, and an in-work benefit (IW) which is conditional on a minimum working-hour requirement. This is the structure that was in place before the Universal Credit UK reform and over which we compute our optimal benefit reform.[6] The second benefit structure features no distinction between in-work and unconditional benefits, as is the case after the Universal Credit (UC) reform (introduced in the UK in 2016). We analyze the latter reform in Section 5.2.

Let $X \in \{IS, IW, UC\}$. We model the amount that a household with marital status p and children n gets for benefit X as the sum of a component that accrues to all households ϕ_0^X, a per-child component ϕ_1^X up to a child cap km^X, and a component that accrues only to couples ϕ_2^X and that we call marital benefit

$$\bar{Y}^X(p, n) = \phi_0^X + \phi_1^X \min\{n, km^X\} + \phi_2^X \mathbb{I}(p = c) \tag{9}$$

Benefit are tapered away at proportional rate ω as labor income increases.

In the pre-2016 benefit reform system, as well as in our benchmark calibration and optimal benefit structure, disposable income after taxes *and* benefits is given by

$$M(y^h) = \tilde{y}(y^h) + \max\{0, \bar{Y^{IS}}(p, n) + Y^{IW}(p, n)\mathbb{I}(h_t > 0) - \omega y^h\}, \tag{10}$$

where y^h is the household level pre-tax income, which is obtained by summing the labor income of head and spouse (if present), and $\tilde{y}(y^h)$ is household after-tax income.

The post-2016 Universal Credit (UC) system differs from the previous benefit structure in place and from our baseline calibration in a number of respects. First, there is no in-work benefit component. Second, there is an initial earnings disregard $y^{DR}(n)$ for families with children, and tapering is based on post-tax income. Finally, benefits are subject to an asset test, in addition to an income test. That is, households with assets in excess of $\bar{a}$ do not receive any benefits. As a result, the flow of disposable income under

[6]This dual structure is typical of many benefit systems, including the US one.

this system M^{UC} is:

$$M^{UC}(y^h) = \tilde{y}(y^h) + \max\{0, \bar{Y}^{UC}(p,n) - \omega^{UC}(\max\{\tilde{y}(y^h) - y^{DR}(n), 0\}\}\mathbb{I}(a_t < \bar{a}) \quad (11)$$

Finally the government provides old-age Social Security payments to retirees and wasteful government expenditure. When choosing the optimal policy or evaluating the introduction of Universal Credit, we impose that reforms are revenue-neutral for the government.

3.4 Assets

There is a single risk-free asset that yields a rate of return r.

3.5 Recursive representation

Working life. Let $W_t^j(\cdot)$ denote the value function for a single person of working age t, with $j = f, m$ for single women and men, respectively. The state variables for a single woman are age t, assets, a_t, the persistent wage shock z_t^g, the number of children n, and her disutility of work type θ. Her recursive problem is

$$\begin{aligned} W_t^f(a_t, z_t^f, n_t, \theta) = \max_{c_t, a_{t+1}, h_t} & \, u(c_t, 1 - h_t - \Psi^s(h_t, t, \theta)) + \\ & \beta(1 - \mu_t^f(z_t^f)) E_t W_{t+1}^f(a_{t+1}, z_{t+1}^f, n_{t+1}, \theta) + \\ & \beta \mu_t^f(z_t^f) E_t W_{t+1}^{fc}(a_{t+1} + a_{t+1}^h, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta) \\ \text{s.t. } & a_{t+1} = (1+r)a_t + M(h_t w_t^f) - CC_t(f, h_t, n_t) - c_t, \quad a_{t+1} \geq 0, \end{aligned} \quad (12)$$

where $\mu_t^f(z_t^f)$ represents the probability that a single woman of age t and wage z_t^f marries. The first expectation is taken with respect to the conditional distributions of own wages and number of children, while the second one is taken with respect to the conditional distributions of own wages, number of children, and the earnings and wealth of potential husbands.

The problem of a single man is similar, except that he works a fixed number of hours

$h_t = \bar{h}$ and has no children, and can thus be simplified as follows

$$W_t^m(a_t, z_t^m) = \max_{c_t, a_{t+1}} u(c_t, 1 - \bar{h}) + \beta(1 - \mu_t^m(z_t^m)) E_t W_{t+1}^m(a_{t+1}, z_{t+1}^m) + \tag{13}$$
$$\beta \mu_t^m(z_t^m) E_t W_{t+1}^{mc}(a_{t+1} + a_{t+1}^w, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta),$$
$$\text{s.t. } a_{t+1} = (1+r)a_t + M(\bar{h} w_t^m) - c_t, \quad a_{t+1} \geq 0.$$

where the second expectation is also taken with respect to the distribution of all the state variables of potential wives.

The value function for a married woman in a couple with *household* wealth a_t is

$$\begin{aligned} W_t^{fc}(a_t, z_t^f, z_t^m, n_t, \theta) =& u(\hat{c}_t, 1 - \hat{h}_t^f - \Psi^c(\hat{h}_t^f, t; \theta)) + \\ & \beta(1 - \delta_t(z_t^f, z_t^m)) E_t W_{t+1}^{fc}(a_{t+1}, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta) + \\ & \beta \delta_t(z_t^f, z_t^m) E_t W_{t+1}^f(a_{t+1}/2, z_{t+1}^f, n_{t+1}, \theta) \end{aligned} \tag{14}$$

and the value function for a married men in a couple with *household* wealth a_t is

$$\begin{aligned} W_t^{mc}(a_t, z_f^t, z_t^m, n_t, \theta) =& U(\hat{c}_t, 1 - \bar{l})) + \\ & \beta(1 - \delta_t(z_t^f, z_t^m)) E_t W_{t+1}^{mc}(\hat{a}_{t+1}, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta) + \\ & \beta \delta_t(z_t^f, z_t^m) E_t W_t^m(\hat{a}_{t+1}/2, z_{t+1}^m), \end{aligned} \tag{15}$$

where $\delta_t(z_t^f, z_t^m)$ denotes the divorce probability for a couple of age t, wife's wage z_t^f and husband's earnings z_t^m. The optimal policy functions $\{\hat{c}_t, \hat{a}_{t+1}, \hat{h}_t^f\}$ in the two value function above maximize the couple's joint problem

$$\begin{aligned} W^c(a_t, z_t^f, z_t^m, n_t, \theta) = & \max_{c_t, a_{t+1}, h_t^f} u(c_t, 1 - \bar{h}) + u(c_t, 1 - h_t^f - \Psi^c(h_t^f, t; \theta)) + \\ & \beta(1 - \delta_t(z_t^f, z_t^m)) E_t[W^c(a_{t+1}, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta)] + \\ & \beta \delta_t(z_t^f, z_t^m) E_t[W_t^f(a_{t+1}/2, \cdot) + W_{t+1}^m(a_{t+1}/2, \cdot)] \\ & \text{s.t. } a_{t+1} = (1+r)a_t + M(\bar{h} w_t^m + h_t^f w_t^f) - CC_t(p, h_t^f, n_t) - c_t, \quad a_{t+1} \geq 0. \end{aligned} \tag{16}$$

Retirement. Retirees don't marry or divorce and have no children living with them. If younger than 95 they die with positive probability s_t^j that depends on age, gender and marital status.

Singles retirees ($j = f, m$) solve the recursive problem

$$R_t^j(a_t) = \max_{c_t, a_{t+1}} u(c_t, 1) + \beta s_t^j R_{t+1}^j(a_{t+1}) \tag{17}$$
$$\text{s.t. } c_t + a_{t+1} = (1+r)a_t + Y_r - T(Y_r), \quad a_{t+1} \geq 0.$$

where Y_r is the old-age Social Security payment from the government.

For couples, we assume that the death of each spouse is independent of each other. Therefore the recursive problem of a retired couple can be written as

$$\begin{aligned} R_t^c(a_t) = \max_{c_t, a_{t+1}} U(c_t, 1, 1) + \beta \Big[& s_t^{cf} s_t^{cm} R_{t+1}^c(a_{t+1}) + s_t^{cf}(1 - s_t^{cm}) R_{t+1}^f(a_{t+1}) + \\ & s_t^m (1 - s_t^{cf}) R_{t+1}^m(a_{t+1}) \Big] \\ \text{s.t. } c_t + a_{t+1} = (1+r)a_t + 2Y_r - T(2Y_r), & \quad a_{t+1} \geq 0. \end{aligned} \tag{18}$$

4 Calibration

4.1 Externally calibrated parameters

Demographics. We use demographic information from the BHPS data. We estimate the proportions of households by gender, marital status, and number of children, and the first-order Markov chain governing the evolution of the number of children as a function of mother's age and marital status. The number of children, n, can take values {0, 1, 2, 3}, where 3 is associated with three or more children. We also estimate marriage and divorce probabilities by age and wage.

We compute the functions for the average number of children in the 0-4 and 5-11 age brackets ($n_{04}(p, t, n)$ and $n_{511}(p, t, n)$) as a function of maternal age, marital status, and total number of children in the household n. We plot these variables in the online Appendix C.

We use the survival probabilities s_t^j from the UK life tables in the Human Mortality Database for the period 1980-2010. Because they condition on gender but not marital status, we use the BHPS data to adjust them to be marital-status specific (see online Appendix C.3).

Preferences and interest rate. We set the curvature of the utility function, γ to 2.5, and the after-tax interest rate r to 2%. We equivalise consumption using an OECD-modified equivalence scale ν_t, according to which the first adult counts as 1, the second as 0.5 and each child as 0.3.

Earnings and wages. We compute the deterministic profile for male earnings and female wages η_t^{gp}, and the stochastic processes for the persistent components of the canonical and NL process (z_t^f and z_t^m) using the BHPS data and the Understanding Society (US) survey. Adding the latter dataset allows us to expand our sample to 2016 and better extract year and cohort effects (See online Appendix A.2 for details). For tractability, we discard the transitory components that we estimate, which also includes measurement error. Given that transitory shocks are typically well insured in these models, omitting them should not have an important effect on our findings. We discretize its estimated persistent component following the procedure in De Nardi et al. (2020).

Correlations across partners. Couples tend to be positively sorted by wages and wealth. To capture this, we have three parameters governing the correlation of husband's earnings and wife's wages. The first correlation pertains to couples who enter our model as married. We compute this one, which turns out to be 0.32, as the unconditional correlation of husband's earnings and wife's wages between age 25 and 30. The second correlation is the one that occurs when single people get married after they enter our model. It turns out to be 0.27, and is obtained by regressing husband's earnings[7] during the first year of marriage/cohabitation on wife's wages during the year before marriage. This avoids avoid potential selection issues due to changes in labor supply at marriage. The third correlation is the one between husbands' earnings shocks and wives' wage shocks after marriage. We estimate ρ_{HW} within our model by targeting the cross-sectional correlation between husband's earnings and wife's wages over the life cycle (as described in Section 4.2). We implement this correlation using a normal copula for both the NL and canonical process.

Turning to the correlation in wealth, we assume that the wealth of the partner that a person marries is a function of that person's wages (for women) or earnings (for men). We find that a 1% increase in a woman's wages translates into a husband's wealth that is on

[7] In our model, male earnings are given by wages times the exogenously fixed hours.

Description	Parameter	Benchmark		UC
		IW	IS	
Intercept	ϕ_0^X	1960	4574	4035
Marriage benefit	ϕ_1^X	0	1366	2312
Per kid	ϕ_2^X	2010	907	1805
Max. kids	km^X	1	–	2
Tapering rate	ω	1.11		0.63
Earnings disregard	y^{DR}	–	–	2304
Asset test	$\bar{a}$	–	–	16000

Table 3: Parameterization of the benefit functions for benchmark in-work benefits (IW), benchmark income support (IS), and Universal Credit (UC), 2015 pounds.

average 2.4% higher, and a 1% increase in a man's income translates into a wife's wealth that is on average 1.7% higher. For simplicity, we assume that individuals marry partners with the expected level of wealth conditional on their own characteristics and that we estimate from the data. We report details about all these model inputs in Appendix C.

Taxes and government expenditure We estimate the tax function $T(y)$ in equation (8) by using BHPS data on pre- and post-tax household income (we obtain the latter from the Derived Current and Annual Net Household Income Variables). Our measure of taxes includes income taxes, National Insurance, and (state) pension contributions of all household members (see Section 4.1). Because income taxation is at individual level in the UK (even for married couples), we separately apply the tax schedule $T(y)$ to the earnings of husbands and wives. Our estimates tax parameters are $\tau = 0.31$, $s = 0.00004$, and $\zeta = 5.38$

Benefit system. We use data from the benefit programs and benefit receipts to parameterize the benefit functions in Equations 9, 10, and 11. We display the resulting parameters in Table 3.

For the in-work benefits in our benchmark economy, we follow the statutory rules of the Working Tax Credit. The child component of WTC is independent of the number of children, which is equivalent to setting $km^{IW} = 1$.

Our income-support program is meant to replicate many benefits available to low-income households. These programs have differential take-up rates and eligibility criteria which would be complicated to model individually. Hence, we use the benefit data available in the BHPS and in the BHPS Derived Net Household Income Variables to estimate ϕ_0^{IS}, ϕ_1^{IS}, and ϕ_2^{IS}. More specifically, we look at average benefit receipts for households whose labor income in a given year is close to zero (below £2,000, but our results are robust to changing this threshold to £1,000 or £3,000). This allows us to average across various types of benefits and to weight by the cross-sectional distribution of benefit receipts within this subset of the population. For this program, there is no limit on how many children the child component can be claimed for.

The tapering rate ω for our benchmark economy is the one implied by the different tapering rates for the two types of benefits in the UK pre-2016 system. These were 0.7 for in-work benefits and 0.41 for income-support, respectively. The former is the statutory one, while the latter is estimated as a weighted average of the statutory tapering rates of the relevant benefits, taking into account cross-eligibility criteria and legal thresholds.[8] Online Appendix D provides a more detailed description of the relevant benefit programs and our computations.

Finally, we take statutory values of the parameters for Universal Credit, because we do not have sufficient years of benefit data to check actual benefit receipts, but we scale all the fixed allowances $\phi_0^{UC}, \phi_1^{UC}, \phi_2^{UC}$ proportionally by a factor 0.9. This makes the switch to Universal Credit revenue-neutral from the perspective of the government under the NL wage process. Table 3 reports the values after the scaling. The £2304 earnings disregard only applies to families with at least one child. All amounts expressed in pounds correspond to 2015 prices.

Retirement. We replicate the UK (New) State Pension System. All retired workers receive a maximum amount of £156 per week (in 2016), which corresponds to about 28 percent of average male earnings (the numeraire in our model). We assume that men retire at age 65 and women at age 60. Age 60 was the statutory retirement age for women in the UK before the Pensions Act 1995, which equalized the retirement age of women with that of men, and established that the transition would be phased in between 2010 and 2020. Given that our data spans 1991-2008, we keep it fixed at 60, which was also the median and mode retirement age for women during this period (Banks and Smith,

[8]We let the two types of benefits taper at their respective rates in the simulation of the benchmark economy, but report their sum to simplify notation and ease comparison. As can be seen from the implied relationship between earnings and benefits in Figure 9, benefits do taper at the sum of the two rates over most of the relevant income range.

2006).

Time use. All components that are measured in units of time (θ_1, θ_2, Ψ) are expressed as fractions of a full day. We assume that full time work is 8 hours a day ($\bar{h} = 0.3$) and that the length of a school week sc_h is 20 hours (4 hours per day), as in Blundell et al. (2016).

4.2 Internally calibrated parameters

We require that each version of our model, whether with the canonical or nonlinear processes, fits our target data as well as possible. To do so, we calibrate a total of eighteen parameters. They include the fixed cost of working for women (three parameters $\psi_0^{lh}, \psi_1^{ih}, \psi_2^{ih}$ for each marital status ($p = s, c$) and for full-time or part-time employment status, hence a total of twelve), the discount factor β, the hourly child care cost f, the correlation coefficient ρ_{HW} between husband's earnings and wife's wages, the disutility of work for the high-cost-of-work group θ_1, and the proportions of single and married women of the θ_1 type. We set θ_2, the value of the disutility of work for the low-cost-of-work women, to its value for men (zero).

These parameters are calibrated to target the following 146 moments. A wealth/income ratio of 2.9 (equal to the average wealth measure for the 1995 BHPS constructed by Banks, Blundell and Smith (2004) divided by average household income in the same BHPS wave) and the profiles of female labor market participation by age, marital status and full-time and part-time status, this amounts to 144 (36×4) targets.[9]. Finally, we target the average correlation between husband's earnings and observed wife's wages from our BHPS sample.

4.3 Model fit

Table 4 reports the calibrated preference parameters and child care costs for both processes. Child care costs are reported as shares of the average male earnings per unit of time. In the NL Parameterization, a family with a mother working full time and a young child pays $0.14 \times \bar{h} = 0.14 \times 0.3 = 4.2\%$ of average male earnings in childcare.

[9]We target the 1991-2008 BHPS profile, which is similar to that implied by the longer panel that also includes the Understanding Society data until 2016

Parameter	NL process	Canonical
Discount rate β	0.97	0.97
Cost of child care f	0.14	0.15
Disutility of work type θ_1	-0.35	-0.31
Share of θ_1, singles	0.27	0.28
Share of θ_1, couples	0.18	0.22
Shock correlation ρ_{HW}	0.11	0.09

Table 4: Internally calibrated parameters

These child care costs should be interpreted as net of other sources of child care which we do not directly model and which include subsidized childcare and help from relatives and friends. Figure 5 plots the calibrated fixed time costs over the working period (reported as fractions of a day) of part- and full-time work (as a solid line for the NL process and as a dashed line for the canonical process).

Figure 6 illustrates how our two calibrated models fit the targeted participation rates by age and marital status in the data. In both calibrations the wealth-income ratio equals its target value. Both the canonical and nonlinear model match participation rates equally well. However, as Figure 7 shows, the NL model matches participation rates by decile of potential wages better.[10] This is untargeted and shows that the NL model generates the right endogenous selection into labor market participation. The success of the model along this key margin further strengthens our confidence in the reliability of its implications, including in terms of optimal policy. In contrast, under the canonical earnings process the labor force participation of women appears to respond too strongly to potential wages. Additionally, our model generates reasonable patterns of participation by number of children (online Appendix E.2) and also matches the persistence of labor market participation well (online Appendix E.3). Turning to the correlations between husbands' earnings and wives' wages (Figure 8), the estimated values from the data do not vary much between the ages of 25 and 60. They stay approximately between 0.33 and 0.17, compared with a maximum potential range between -1.0 and +1.0. The correlations from the NL and canonical processes display a similar variation. Consistently with the data, both processes generate a relatively flat profile in the correlations between ages 35

[10] In the data, we impute potential wages as described in Section 2.2. Women whose wages we cannot impute because we never observe them working are still included in the computation of participation rates. We assume that the distribution of their potential wages is identical to that of the non-participants whose wages we can impute.

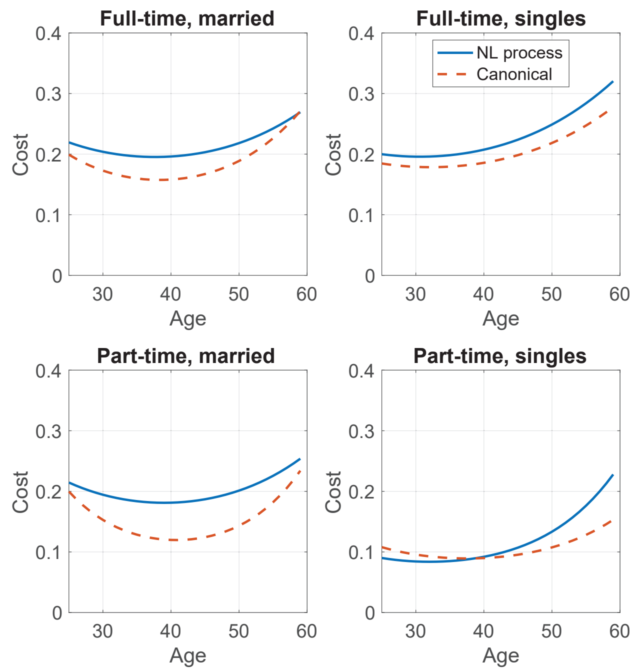


Figure 5: Calibrated fixed costs of working over the working period as a fraction of time endowment.

and 50. However, both imply an increase in the correlation after age 50, while the data display a mild decrease after that age. Overall, both models seem to fit the age profile of correlations reasonably well.

5 Policy evaluation

We now turn to evaluating the implications of the nonlinear and canonical processes in terms of the optimal composition of in-work and income support and their phase out rate. Our welfare criterion is given by the utilitarian, expected lifetime utility of newborns. We report results both behind the full veil of ignorance and after the realization of gender, marital status and number of children.

5.1 Optimal benefit system

We start by evaluating the provision of government insurance by optimizing over the parameters of the welfare system for the income floor and the in-work benefit that were in place before the introduction of Universal Credit. More specifically, we optimize over

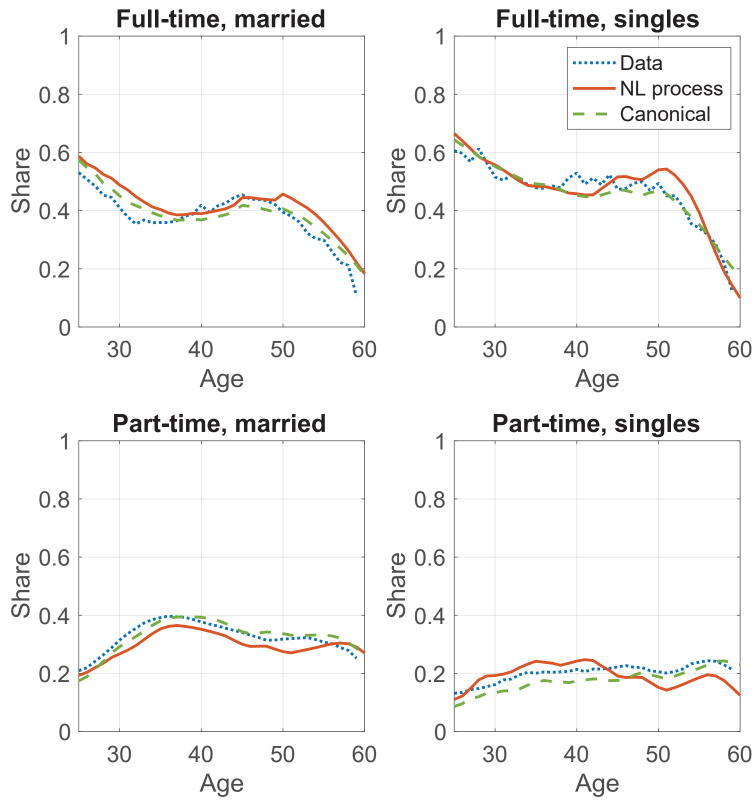


Figure 6: Fit of labor market participation for women by marital status and working hours for the NL and canonical processes compared with the data

the intercepts ϕ_0^{IS}, ϕ_0^{IW} and the slope (tapering rate) ω of the function (10) to find the system that maximizes ex-ante welfare (under the veil of ignorance) while maintaining the tax function unchanged and keeping total tax revenues minus total benefit outlays constant. As a result, this change is budget neutral for the government. Because the purpose of our experiment is to evaluate the relative role of out-of-work and in-work benefits and how they should be optimally related to income, rather than studying the distributive effects of the benefit system across family types, we keep the marital status and child-specific components of benefits ϕ_2^X, ϕ_3^X for $X = \{IS, IW\}$ constant across our experiments.

Table 5 shows the result of this optimization. Column 2 reports the parameter values for the two benefit functions in our benchmark economy, while columns 3 and 4 report their optimal values under, respectively, the NL and canonical wage process. Under the NL wage process, the optimal income floor level is close to the one in the benchmark

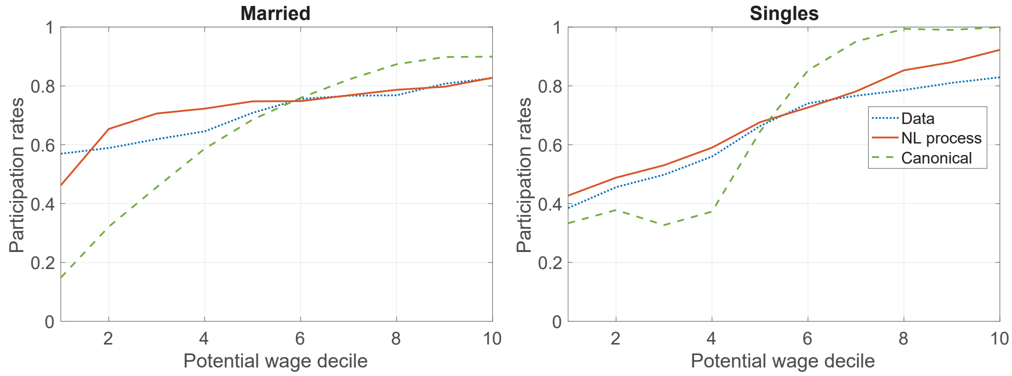


Figure 7: Women's labor market participation by marital status and decile of potential wages, compared with the data.

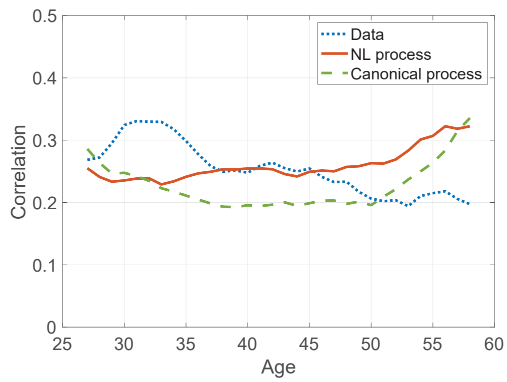


Figure 8: Correlation between observed men's earnings and observed wife's wages.

economy while the in-work benefit level is less than one third of its benchmark counterpart. The fall in the maximum total benefit for working individuals is compensated by the halving of the tapering rate from 110 to 54 per cent. The difference between the optimal and the benchmark benefit policies is possibly best appreciated with the help of Figure 9, which plots the relationship between benefit levels and after-tax labor income for single men, women and couples in the benchmark (solid lines) and under the optimal system under the NL (dash-dot lines) and canonical (dashed lines) wage processes. The continuous lines plot benefit levels for working individuals, while the circles in the top two panels denote benefits for non-working individuals (single women in our model). Under the NL wage process, benefits for working households are lower than in the benchmark but they are exhausted at a higher level of disposable income due to the fall in the tapering. Households earnings below 15 per cent of average male earnings income have lower benefits as a result while those above 15 and below 50 per cent gain under the new

Parameters		Benchmark	Optimum (NL)	Optimum (Ca)
Income floor, level	ϕ_0^{IS}	0.15 (£4574)	0.1396	0.0029
In work, level	ϕ_0^{IW}	0.07 (£1960)	0.0171	0.1655
Tapering rate	ω	1.11	0.5431	0.5448

Table 5: Income floors and in-work benefits: benchmark vs optimum under NL and canonical processes. Pound values correspond to 2015.

policy. While the optimal benefit system under the NL process implies a reduction in the net return to working, the optimal benefit configuration under the canonical process is dramatically skewed towards in-work benefits. In particular, the optimal system implies a 96 per cent reduction (from 0.15 to 0.01) in benefits for non-working individuals and a nearly three-fold increase in in-work benefits. As a result, at the optimum the net return to the first pound of labor income (the difference between the vertical intercept of the straight line and of the corresponding circle in Figure 9) is more than ten times as large under the canonical than in the NL process. The tapering rate, instead, is very similar for both processes.

Increasing in-work benefits and reducing income support have offsetting effects on welfare. On the one hand, welfare increases as a result of improved incentives to participate in the labor market, which in turn increase the tax revenues that can be spent to insure households. On the other hand, welfare falls as a consequence of the reduction in insurance provision for low-wage households and, in particular, single women. Under the canonical process, the benefits outweigh the costs, but the opposite is true under the NL process. The key reason for this difference is that the canonical process underestimates wage persistence for women. Thus, the cost of reducing insurance to women on low wages is lower under the canonical process because having a low wage is a more transitory state, against which it is easier to self-insure. In contrast, the NL process replicates the fact that having a low wage is a relatively persistent state. Hence, reducing income support to encourage labor market participation drastically reduces welfare for women on low wages and with a higher disutility from work for whom the increase in in-work benefits does not compensate the welfare costs of foregone leisure. As a result, the optimal welfare system under the more realistic NL process is much closer to the system in place before 2016.

Figures 10 shows that, under both wage processes, the optimal policy mix results in higher part-time and lower full-time labor market participation by single women, and

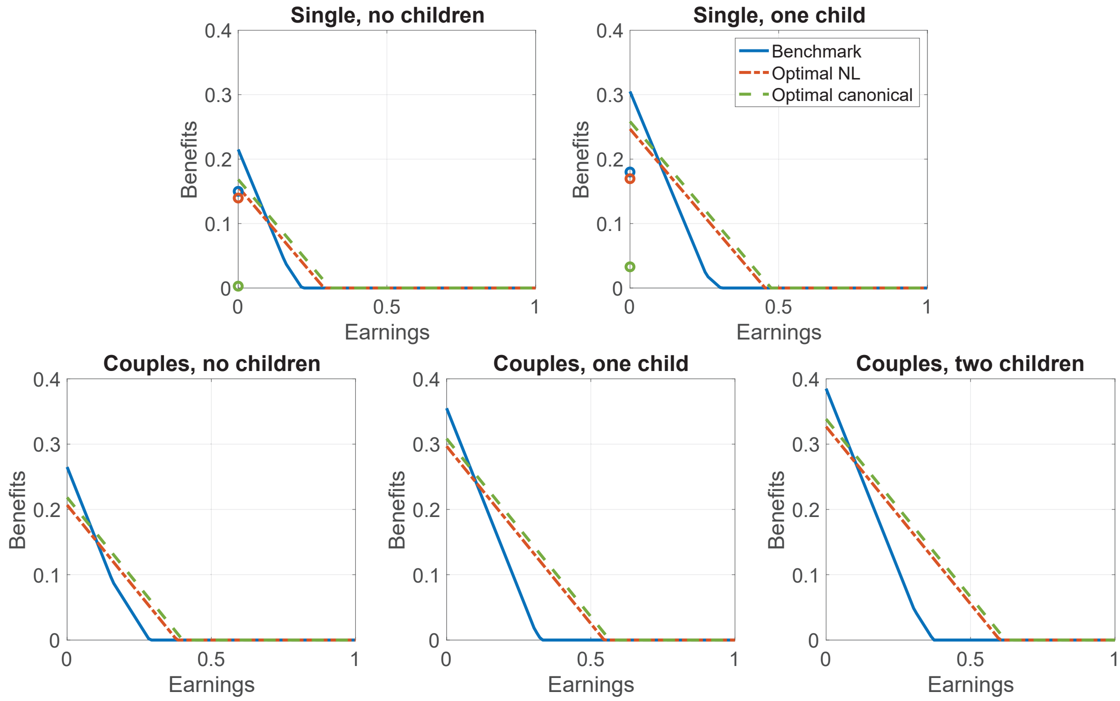


Figure 9: Implied total level of benefits, by income, marital status, and number of children. Circles represent benefits for households where everyone is out of work. Lines represent benefits for households in which at least one member works. Earnings and benefits are expressed as the share of average male earnings

implies a significant increase in participation overall. The rise in overall participation is driven by the large reduction in the effective tax rate for benefit claimants stemming from the halving of the tapering rate. This increases the return to work part-time relative to both full-time work and non-participation. The increase in participation, and hence the overall employment response, is much larger under the canonical process, due to the dramatic shift from income support to in-work benefits.

Table 6 reports the welfare change associated with the switch to the optimal benefit system. Welfare is expressed as the percentage change in consumption (constant across ages and states) that would make a 25 year old in the benchmark economy indifferent to being in the counterfactual economy. The "overall" measure in the first row is under the full veil of ignorance, including the realization of the gender and marital status draw. The other three rows report the welfare change from the perspective of age 25, conditional on the realization of gender and initial marital status, but before the draw of the initial number of children. Although marital status may change over the life cycle, the comparison across gender and marital status at age 25 provides some insight into the

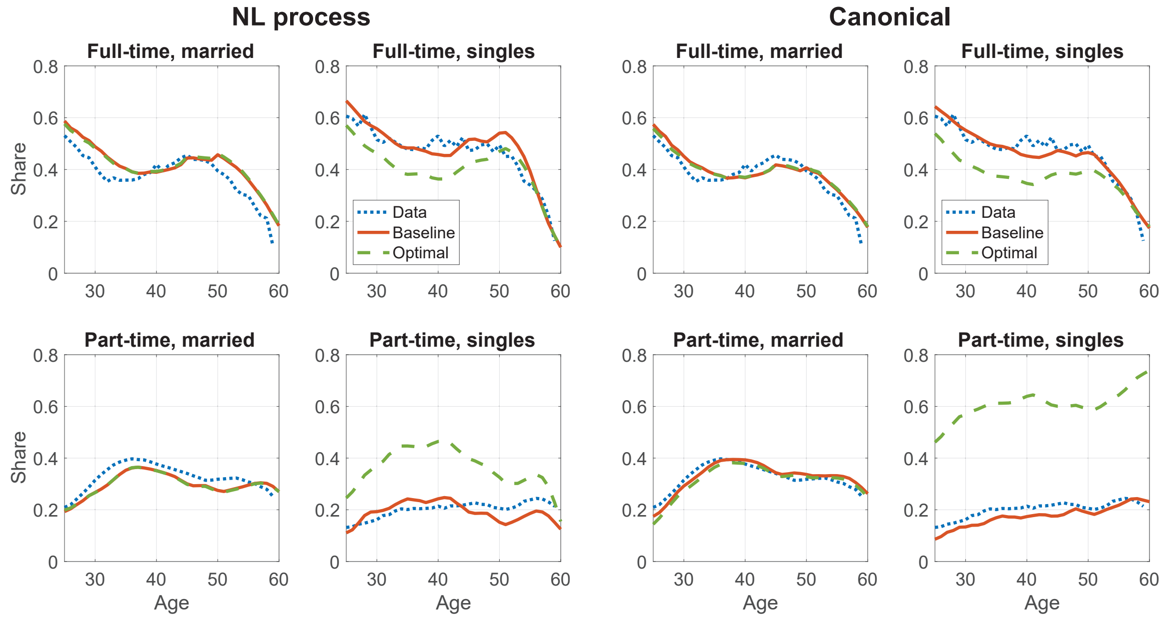


Figure 10: Labor force participation for women in the data, baseline model, and under its optimal benefit system.

Group	NL (opt)	Canonical (opt)
Overall	0.88	0.97
Single men	-0.09	1.20
Single women	1.04	-1.33
Couples	0.93	1.22

Table 6: Welfare change (measured as consumption equivalent compensations) implied by the switch to the optimal system, by gender and marital status, as of age 25.

distributional implication of the reform across these two dimensions.

The two processes have different implications for both the overall welfare gains and their distribution for various groups at age 25. The overall welfare gains from moving from the benchmark to the optimal system are respectively 0.89 and 0.98 percentage points under the NL and canonical process. Under the canonical process, the gains are mostly driven by households, single men and couples, who are unaffected by the reallocation from unconditional to in-work transfers and benefit from the lower tapering rate. Single women lose from the reform because many of them are being pushed into work as the result of the dramatic switch to in-work benefits. Conversely, the more generous unconditional transfers under the NL process imply that single women, who are more likely to choose to be non-participants, are the main beneficiaries of this policy reform.

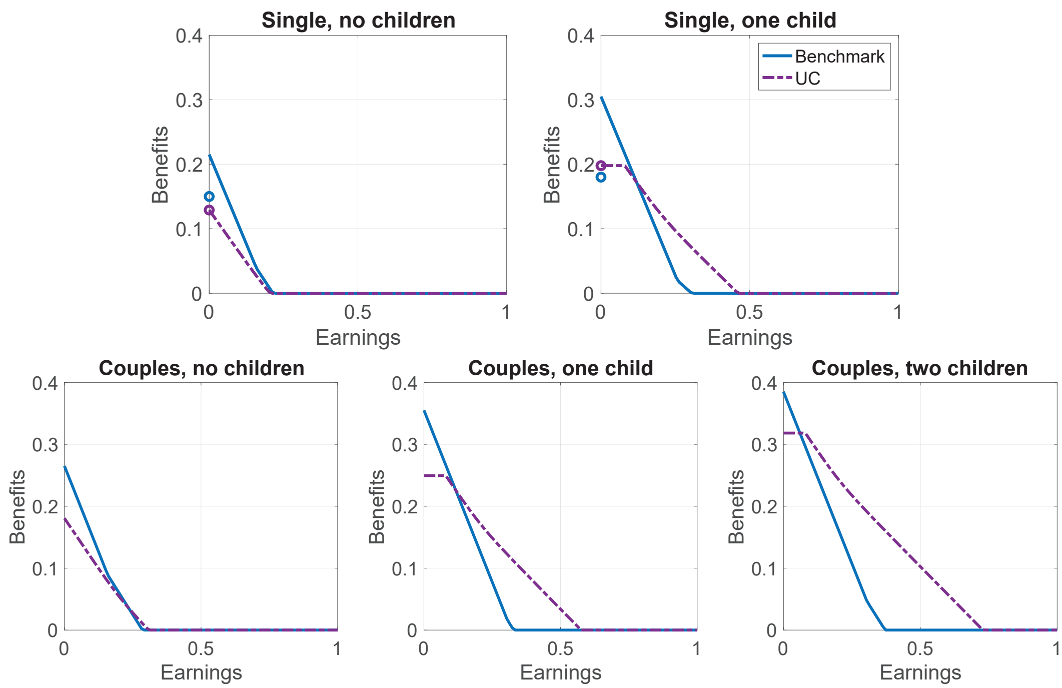


Figure 11: Implied total level of benefits, by income levels, comparing benchmark vs. Universal Credit. For singles, circles represent benefit entitlement for non-working individuals under our benchmark. Earnings and benefits are expressed as the share of average male earnings.

Robustness with respect to the distribution of potential wages. Table 7 reports the optimal benefit system for each of our female wage imputation procedures. It shows that the results are very similar to those implied by our benchmark imputation.

5.2 Universal Credit

The aim of this section is to compare the allocation and welfare implied by the benefit system before and after Universal Credit. Universal Credit replaced many key benefits (Income-Based JSA, Income-Related Employment and Support Allowance, Income Support, Working Tax Credit, Child Tax Credit and Housing Benefits, but not Child Benefits) that we have modeled in our benchmark economy (described in Section 4) with a unified benefit system. Two features of Universal credit are worth pointing out. First, it features a £2,304 earnings disregard for families with children. Second, benefits are withdrawn as a function of *after-tax* income, rather than pre-tax income (as in the pre-reform system). Universal Credit was piloted in 2013 in a few areas, and then gradually rolled out to all of Great Britain from May 2016 to December 2018.

Parameters		NL process, optimum			
		H-Ben	H	H-Child	FE
Income floor, level	ϕ_0^{IS}	0.1396	0.1405	0.1423	0.1405
In work, level	ϕ_0^{IW}	0.0171	0.02	0.01	0.02
Tapering rate	ω	0.5431	0.54	0.52	0.54
		Canonical process, optimum			
		H-Ben	H	H-Child	FE
Income floor, level	ϕ_0^{IS}	0.001	0.002	0.002	0.001
In work, level	ϕ_0^{IW}	0.166	0.166	0.166	0.166
Tapering rate	ω	0.55	0.55	0.55	0.55

Table 7: Income floors and in-work benefits: optimum under alternative potential wage imputations.

Figure 11 reports benefits levels as a fraction of pre-tax income in our benchmark economy and under Universal Credit. Its main takeaway is that, compared to our benchmark, Universal Credit entails lower benefits for households without children and for very low-income couples with children, and higher benefits for less poor households with children.

Given that we find that the policy implications of the canonical and NL process are different, we also evaluate the effects of the introduction of the Universal Credit benefit reform in both cases.

Figure 12 compares the labor force participation under Universal Credit and in the benchmark (pre-UC benefit system), under NL earnings and wages. It shows that the introduction of UC increases the part-time participation of single women and reduces their full-time participation, similarly to the welfare-improving reform that we study in the pre-UC benefit system. Due to space constraints, we report the corresponding outcomes for the canonical process in online Appendix E.4.

Group	NL process	Canonical
Overall	1.24	1.13
Single men	-2.02	-0.90
Single women	2.28	0.38
Couples	1.37	1.37

Table 8: Welfare change for switching to Universal Credit, measured as consumption equivalent compensations.

Table 8 reports the steady-state changes in welfare associated with switching from the benchmark pre-UC benefit configuration to Universal Credit for both the canonical

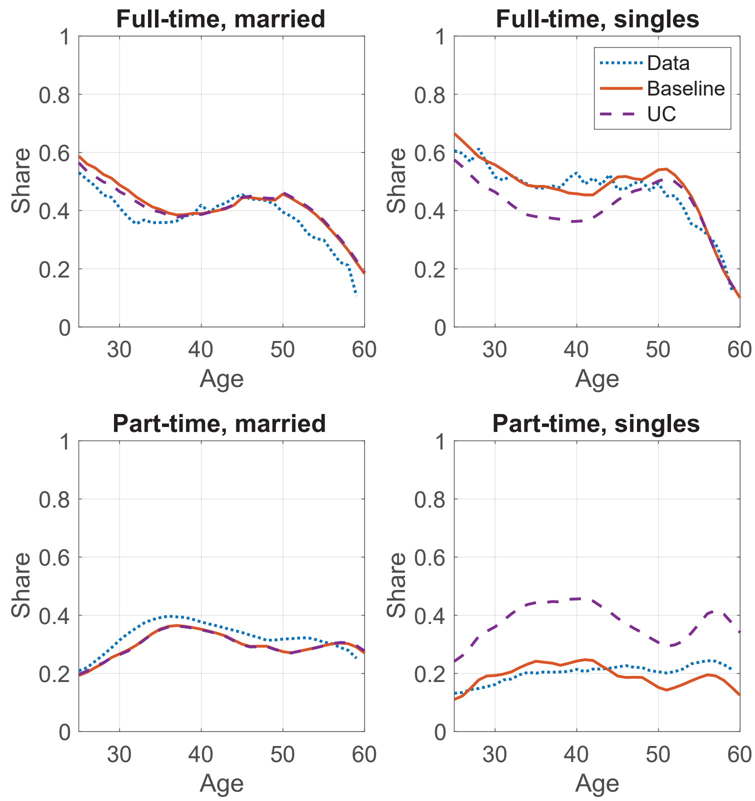


Figure 12: Labor force participation for women under NL process: Universal Credit vs baseline.

and NL earnings and wages. In both cases, the overall welfare gains are slightly higher than those of the optimal policy reform that we discuss in the previous subsection. While this may appear puzzling, it is not once one realizes that Universal Credit is not nested in the linear class of benefit functions in our previous experiment. In fact, it features additional policy tools (such as an earnings disregard) and these additional tools turn out to have positive welfare effects. On the other hand, single men are substantially worse off compared to the optimal reform of the previous system because Universal Credit implies a benefit reduction for households without children.

6 Conclusion and directions for future research

A growing body of empirical work takes advantage of large, administrative data sets and new statistical techniques to provide evidence that households' labor income dynamics are substantially richer than those implied by the *canonical* income processes—with constant variance and persistence—that are typically used in studies that evaluate welfare policies.

We establish that the rich dynamics of labor income documented for other countries also hold for the UK. Rather than being constant, the variance and persistence of labor earnings display substantial differences by age and labor income history. These rich dynamics are a feature not only of earnings, but also of wages. Hence, they reflect genuine labor income risk rather than being merely the byproduct of the adjustment of hours to wage shocks. We also find that these features of the data are present in both administrative and survey data sets.

We show that ignoring these richer dynamics when estimating stochastic labor income processes implies biased estimates of key moments. In particular, relative to a more flexible earnings process which does not impose constancy in variance and persistence, the canonical model underestimates the persistence of shocks to female wages and overestimates the persistence of shocks to male earnings.

Correctly estimating the persistence of labor income shocks is important to capture labor income risk because persistence crucially affects agents' ability to insulate consumption from income shocks through borrowing and lending (self-insurance). This is why we investigate how allowing for richer labor income dynamics affects the evaluation of welfare policies compared to the canonical income process. To do so, we build and estimate a structural life-cycle model with heterogeneity in family structure that captures the following important elements. First, that both the need for resources and the level of welfare benefits in the UK depend on the presence of a spouse and the number of dependent children. Second, that allowing for both single and married households is crucial because labor income pooling within families and the possibility of adjusting the labor supply of the secondary earner are potentially important margins of insurance at the household level.

We use our model to evaluate alternative benefit reforms under both richer and canonical labor income processes. Our findings confirm that correctly capturing the dynamics of labor income is important to evaluate the costs and benefits of welfare policies. In particular, we analyze a hypothetical reform that chooses the structure of two main benefits—income support and in-work benefits—to maximize (utilitarian) welfare in the economy. This reform entails relatively small welfare gains compared to the pre-2016, benchmark UK benefit configuration. More importantly, the optimal benefit configuration is very different under the canonical as opposed to the flexible, non-linear wage process. Under the flexible process, the optimal benefit configuration is similar to the pre-reform one and implies that income support, independent of labor force participation, should provide the main share of total benefit income. In contrast, if one were to ignore the rich wage dynamics that we estimate from the data and simply assume a canonical wage process, one would find an optimal policy which incorrectly prescribes a trebling of in-work benefits and basically no role for income support. The intuition is that the canonical wage process underestimates the average persistence of shocks to female wages, relative to the richer process. Since more transitory shocks are easier to self-insure, the optimal policy under the canonical process is skewed towards providing incentives to work, rather than insurance against low labor income realizations.

The result that under the flexible earnings process, the constrained-optimal benefit configuration is very similar to the pre-reform one is an interesting finding. Although policy makers were not relying on a model with flexible earnings risk to find optimal policy, we understand this result as a product of the political process in which inputs from different parts of society are taken into account (existing academic and policy work, feedback from charities working with low-income families, etc.). As a result of balancing costs and benefits for different stakeholders, a solution was reached that was relatively close to the model constrained optimum. Although these are forces which are present in general in the policy process, it is difficult to know whether its resulting optimality in this case is likely to apply more broadly.

We also consider a reform that mimics the switch to the Universal Credit which was introduced in 2016 and completed in 2018. Universal Credit includes an earnings disregard for households with children, and thus does not belong to the class of linear benefit functions that we consider for optimality in the previous reforms. We find that the move to Universal Credit implies overall welfare gains which are similar to those under our optimal benefit system, but that this average improvement masks heterogeneous effects. The main beneficiaries of UC are households with children, while singles without children lose out.

For tractability and clarity, our model assumes that marriage, divorce, and children evolve as in the data, but exogenously. Endogenous marriage and fertility choices could affect our results to the extent that they generate additional insurance mechanisms for both singles and couples. For instance, couples could delay having children in response to a negative shock, and individuals could make decisions about marriage that depend on their own wage shock. As a result, marriage and divorce could imply less risk than they do in our model. However, for a single household, it is not clear that marriage as an insurance device is always available; for instance, the value of a single person in the marriage market might be lower after a negative earnings or wage shock. While these are very interesting questions, they are beyond the scope of the current paper and we leave them for future research.

References

Abowd, John and Card, David (1989), 'On the covariance structure of earnings and hours changes', *Econometrica* **57**(2), 411–45.

Adda, Jérôme, Dustmann, Christian and Stevens, Katrien (2017), 'The career costs of children', *Journal of Political Economy* **125**(2), 293–337.

Arellano, Manuel and Bonhomme, Stéphane (2017), 'Quantile selection models with an application to understanding changes in wage inequality', *Econometrica* **85**(1), 1–28.

Arellano, Manuel, Blundell, Richard and Bonhomme, Stéphane (2017), 'Earnings and consumption dynamics: A non-linear panel data framework', *Econometrica* **85**(3), 693–734.

Banks, James and Smith, Sarah (2006), 'Retirement in the UK', *Oxford Review of Economic Policy* **22**(1), 40–56.

Banks, James, Blundell, Richard and Smith, James (2004), Wealth portfolios in the United Kingdom and the United States, *in* 'Perspectives on the Economics of Aging', University of Chicago Press, pp. 205–246.

Bewley, Truman (1977), 'The permanent income hypothesis: A theoretical formulation', *Journal of Economic Theory* **16**(2), 252–292.

Bick, Alexander and Fuchs-Schündeln, Nicola (2017), 'Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis', *The Review of Economic Studies* **85**(3), 1543–1576.

Blundell, Richard, Costa Dias, Monica, Meghir, Costas and Shaw, Jonathan (2016), 'Female labor supply, human capital, and welfare reform', *Econometrica* **84**(5), 1705–1753.

Blundell, Richard, Reed, Howard and Stoker, Thomas M. (2003), 'Interpreting aggregate wage growth: The role of labor market participation', *American Economic Review* **93**(4), 1114–1131.

Borella, Margherita, De Nardi, Mariacristina and Yang, Fang (2021), 'Are marriage-related taxes and social security benefits holding back female labor supply?', *The Review of Economic Studies* . forthcoming.

Chiappori, Pierre-André, Costa-Dias, Monica and Meghir, Costas (2018), 'The marriage market, labor supply, and education choice', *Journal of Political Economy* **126**(S1), S26–S72.

Cubeddu, Luis and Ríos-Rull, José-Víctor (2003), 'Families as shocks', *Journal of the European Economic Association* **1**(2-3), 671–682.

De Nardi, Mariacristina, Fella, Giulio and Paz-Pardo, Gonzalo (2020), 'Nonlinear household earnings dynamics, self-insurance, and welfare', *Journal of the European Economic Association* **18**(2), 890–926.

De Nardi, Mariacristina, Fella, Giulio, Knoef, Marike, Paz-Pardo, Gonzalo and Van Ooijen, Raun (2021), 'Family and government insurance: Wage, earnings, and income risks in the netherlands and the us', *Journal of Public Economics* **193**, 104327.

Gouveia, Miguel and Strauss, Robert P (1994), 'Effective federal individual income tax functions: An exploratory empirical analysis', *National Tax Journal* pp. 317–339.

Groneck, Max and Wallenius, Johanna (2017), 'It sucks to be single! marital status and redistribution of social security', *SSE Working Papaer Series in Economics* **1**.

Guner, Nezih, Kaygusuz, Remzi and Ventura, Gustavo (2011), 'Taxation and household labour supply', *The Review of Economic Studies* **79**(3), 1113–1149.

Guner, Nezih, Kaygusuz, Remzi and Ventura, Gustavo (2012), 'Taxation and Household Labour Supply', *The Review of Economic Studies* **79**(3), 1113–1149.

Guner, Nezih, Kaygusuz, Remzi and Ventura, Gustavo (2020), 'Child-related transfers, household labour supply, and welfare', *The Review of Economic Studies* **87**, 2290–2331.

Guvenen, Fatih, Karahan, Fatih, Ozkan, Serdar and Song, Jae (2021), 'What do data on millions of U.S. workers reveal about lifecycle earnings dynamics?', *Econometrica* **89**(5), 2303–2339.

Heckman, James J (1979), 'Sample selection bias as a specification error', *Econometrica: Journal of the econometric society* pp. 153–161.

Hood, Andrew and Norris Keiller, Agnes (2016), 'A survey of the UK benefit system', *London: Institute for Fiscal Studies* .

Kaplan, Greg (2012), 'Inequality and the life cycle', *Quantitative Economics* **3**(3), 471–525.

Keane, Michael P. and Wolpin, Kenneth I. (2010), 'The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women', *International Economic Review* **51**(3), 851–892.

MaCurdy, Thomas E. (1982), 'The use of time series processes to model the error structure of earnings in a longitudinal data analysis', *Journal of Econometrics* **18**, 83 – 114.

Meghir, Costas and Phillips, David (2010), Labour supply and taxes, *in* J. A. Mirrlees et al., ed., 'Dimensions of Tax Design: The Mirrlees Review', Oxford University Press, chapter 3, pp. 202–274.

Nishiyama, Shinichi (2019), 'The joint labor supply decision of married couples and the us social security pension system', *Review of Economic Dynamics* **31**, 277–304.

Ozkan, Serdar, Storesletten, Kjetil, Holter, Hans and Halvorsen, Elin (2017), Dissecting Idiosyncratic Income Risk, 2017 meeting papers, Society for Economic Dynamics.

Shaw, Jonathan (2011), 'FORTAX: UK tax and benefit system documentation', IFS Working Paper 08/11.

Shephard, Andrew (2009), 'FORTAX: Reference manual', Unpublished Manuscript.

Storesletten, Kjetil, Telmer, Christopher I. and Yaron, Amir (2004), 'Consumption and risk sharing over the life cycle', *Journal of Monetary Economics* **51**(3), 609–633.

A Data and features of earnings and wages

A.1 The British Household Panel Survey (BHPS)

Starting in 1991 (and continuing until 2010, when it was discontinued to be included within the wider Understanding Society survey), the BHPS sampled 5,500 households and 10,300 individuals, which were then followed over time, hence generating a long panel. If an individual in the initial sample separated from his/her original household, all members of his/her new household were also interviewed. Children were interviewed once they reached the age of 16. These features imply that this survey should remain representative of the UK population.

As most household surveys, it has the limitation that all answers are self-reported and thus potentially subject to measurement error. However, the design of the survey suggests that measurement error in earnings is likely to be lower than in other surveys, such as the PSID in the US, because instead of just being asked about their total labor earnings in the last twelve months, respondents were asked to check their last pay slip and report about it. Furthermore, in a relevant proportion of the observations (around 30%), the interviewer saw the pay slip.

A.2 Sample and variable construction

We include individuals between 25 and 60 years of age. Given that our focus is on labour market income dynamics, we drop individuals from the sample if they have income from self-employment. We deflate earnings and wages with the CPI (2015=100).

We reconstruct age whenever the change of date in the interview implies that the individual is reported to be the same age in two consecutive years, but only when reported age does not differ by more than one year from one's expected age.

We use this broader sample to compute all of our measures of labour market participation. For the estimation of the earnings processes, the correlations of wages between members of a couple, etc. we impose further sample selection criteria related to the availability of earnings and wage data as follows.

Men's earnings. For men, our main variable of interest is total annual earnings. We construct this measure by adding up earnings for all jobs held over the past year (1st

September to 31st August) for each worker. We do so by using the information available on the start and end months of all employment spells, together with the "usual" payment per unit of time. We exclude jobs that were held for a period shorter than a calendar month because many individuals do not report the exact day when the employment or unemployment spell began. We also drop observations in which the respondent does not report their usual payment per unit of time and self-employed men, retirees, full-time students, and the long-term disabled.

After excluding these and using the complete BHPS sample period (1991-2009), we have 42,659 person-year observations for men's earnings. Typically, the literature on earnings dynamics (e.g. Kaplan, 2012; Guvenen et al., 2021) further excludes observations below some minimum threshold, that is those below 5% of yearly median earnings, or £1,300 a year in our data. There are 2,259 (5.2%) male-earnings observation below such threshold in our dataset, of which 2071 (4.8%) display earnings which are exactly zero. The vast majority of individuals in the latter group report being unemployed. Rather than excluding these observations, we bottom-code them to £1,300 and check for the robustness of our results to changing this threshold.

We make this choice because for our question it is important to include the most unfavorable earnings outcomes, such as staying out of work for a long time, for which government insurance is likely particularly valuable. However, the Arellano et al. (2017) procedure that separates persistent and transitory earnings and estimates their rich dynamics, requires taking logs of earnings. Bottom coding allows for the inclusion of all observations. Although the choice of a lower bound is somewhat arbitrary, our bottom coding is low enough (around £100 per month) to capture the really high marginal utility of consumption in this situation, and yet reflects other sources of insurance which are likely under-reported, such as help from family and friends, private charities, informal work, and so on.

Women's wages. For women, our main variable of interest is their hourly wage. For simplicity, we focus on the current job being held by the individual rather than an annual average. As we describe in Section 2.1, we focus on potential rather than observed wages. Therefore, we keep in our sample both the women who are currently working and those who are not, as long as we can impute a wage to the latter, for which we require that we observe them working at least once in the sample. We eliminate outliers that most likely reflect recording errors and missing values and drop individuals whose total working hours exceed 80 hours per week and those who have negative earnings or wages. Because wages are computed by dividing earnings by hours worked, we eliminate extreme changes ($|\log w_t - \log w_{t-1}| > 2$) that likely are errors in recording hours of work. After excluding all of these, we have 58,116 observations for women, of which 43,198 correspond to women for which we observe positive hours worked, and thus wages, and the remainder correspond to women for whom we can impute a wage.

While we estimate our wage process on imputed wages, the statistics that we report for female wages in the BHPS (for example, those in Section 2 or Appendix A.4) refer to female wages for labor market participants.

For both men and women, when we apply the Arellano et al. (2017) we increase our sample size by performing a rolling-sample transformation similar to that used by De Nardi et al. (2020) for the PSID data.

We decompose potential wages into a deterministic age-efficiency profiles η_t^{gp} which varies by gender g and marital status p and a stochastic residual component. To estimate the profiles η_t^{gp} more precisely, we expand our sample to include the Understanding Society survey (2010-2016). We report the resulting profiles in Appendix C.4.

A.3 Comparing the BHPS and NESPD data

A.3.1 The New Earnings Survey Panel Dataset (NESPD)

A National Insurance Number (NIN) number is randomly issued to all UK residents at age 16 and kept constant throughout one's lifetime. Individuals whose NIN ends in a certain set of digits are automatically selected for the NESPD sample. Its data is available for the years between 1975 and 2015. Every April, all employers whose employees qualify for the sample receive a form (currently online, although it was on paper in the early years of the sample) where they must provide payroll data about those employees.

This implies that, for individuals included in the survey, the NESPD contains complete information on their working life from the first year they started working (or 1975) until retirement age (or 2015), for all years during which the individual was working with the last recorded employer in April *and* the employer returned the questionnaire.

The most important limitation of the NESPD is that it is has a 25-30% employer non-response rate, implying that it only gathers 0.7% of all UK workers rather than 1%. Moreover, valid responses fell from 75% in the 1980s to 60% in 2012 (Adam, Phillips, Roantree 2016). This generates two main problems. First, endogenous non-responses might affect the randomness of the sample. Second, we cannot distinguish individuals who are not working from individuals whose employers do not respond to the survey.

As a result, the population covered by the BHPS' earnings measure is more comprehensive than that covered by NESPD. This is due to the fact that the latter is filled by the employer, so individuals who happen to be unemployed or out of the labor force in the week of reference will not appear in the sample. In contrast, the BHPS, being a household survey, can capture people who are non-employed but have worked at some point during the previous year.

A.3.2 Sample selection in the NESPD

We drop cases for which there are two records with the same identifier (year pair), as well as individuals whose hours worked or weekly pay are missing, or for whom age evolves unexpectedly, which can reflect, in the case of the NESPD, errors in recording NINs (as stated in the documentation for the data).

We apply the same transformation and sample selection criteria to the NESPD data as those for our main BHPS sample that we have described in Appendix A.2, with three main differences, which are motivated by the characteristics of the NESPD. For the purposes of this comparison, we apply the same screens to the BHPS data.

First, the NESPD only considers the highest-paid job for each individual, for which a direct measure of "annual earnings" is reported (earnings go from April 7th to April 6th, consistent with the tax year in the UK). Thus, for our comparisons with the NESPD, we also only keep the highest-paid job in the BHPS.

Second, we only consider men who have received at least 5% of median earnings (around £1,300 (2015)) in the year up to the moment when they are observed. This choice is motivated by the fact that the NESPD does not capture workers who spend all year out of the labor force.

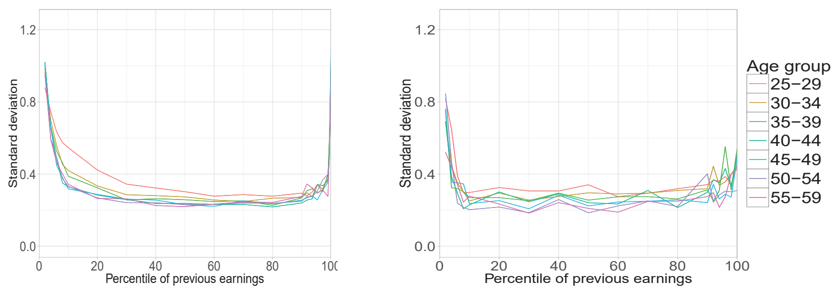
Third, we use data from 1996 to 2006 because of three considerations. First, annual earnings only start being available in the NESPD after 1996. Second, up to the mid-

90s there were many changes in the UK labor market (e.g. de-unionization) that could confound the analysis. Third, in the years 2007 and 2008 the New Earnings Survey suffered a budget cut that implied non-random attrition of part of the sample (those in smaller businesses which were still filling paper-based forms), and this was immediately followed by the financial crisis, whose specific effects are not the object of our study.

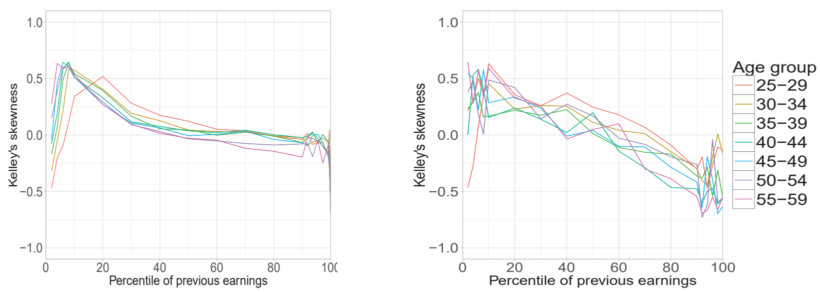
A.3.3 Comparison

Figures 13 and 14 show that the key implications of the BHPS and NESPD data are very similar. The most salient difference is that average persistence is higher in the NESPD, which could reflect the presence of larger measurement error in the BHPS. Luckily, the econometric procedure we use and describe in Section 2.2 separately identifies the persistent and transitory components of earnings and wage changes. Hence, whenever present, measurement error is mostly captured by the transitory component, which we do not include in our structural model.

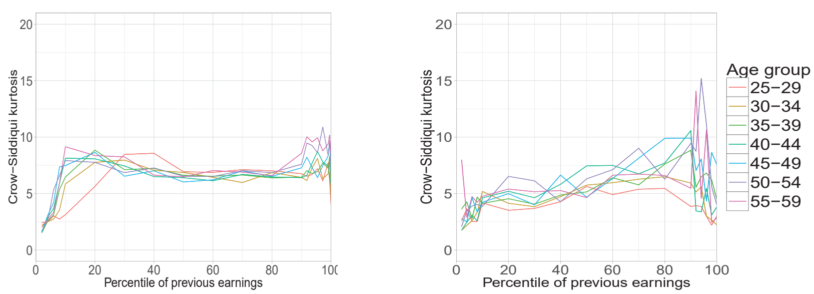
NESPD BHPS



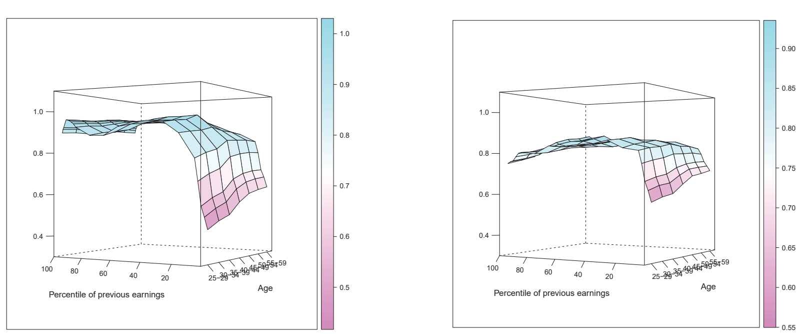
Standard deviation



Kelly’s skewness



Crow-Siddiqui kurtosis



Persistence

Figure 13: Moments of male earnings changes in the BHPS and NESPD. Top three panels: by previous earnings. Bottom panel, by previous earnings and age.

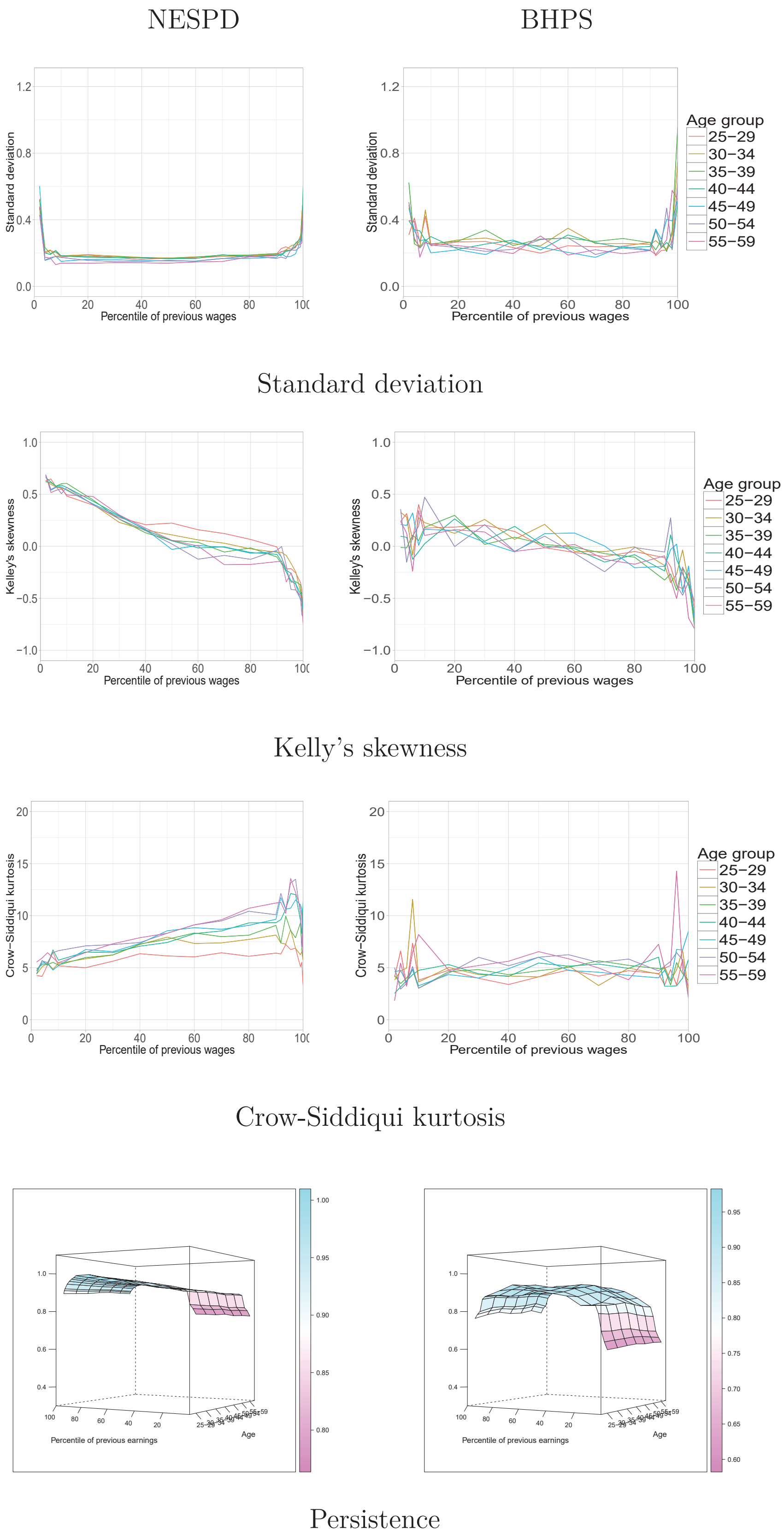


Figure 14: Moments of female wage changes in the BHPS and NESPD. Top three panels: by previous earnings. Bottom panel, by previous earnings and age.

A.4 Comparing the earnings dynamics of singles and married

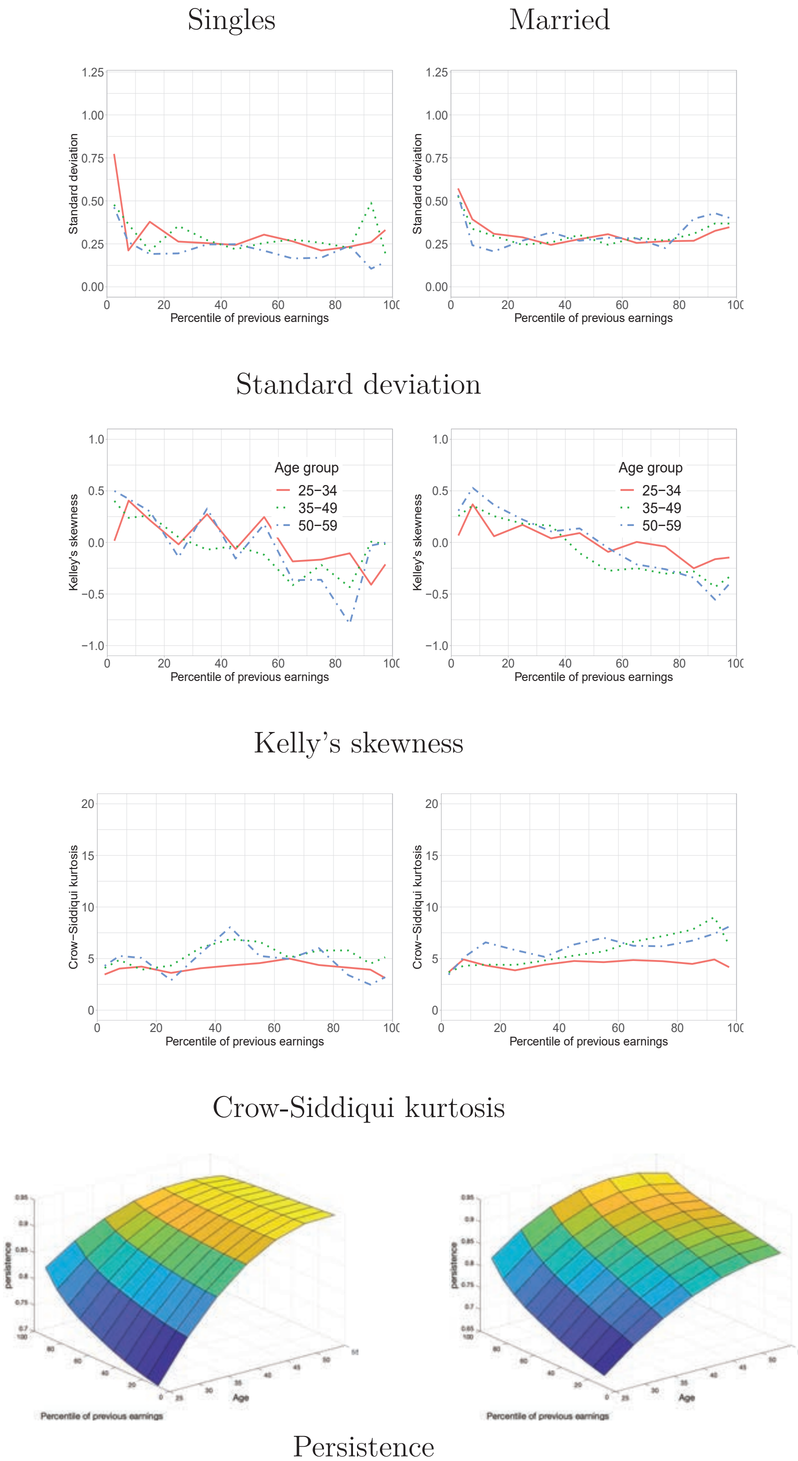
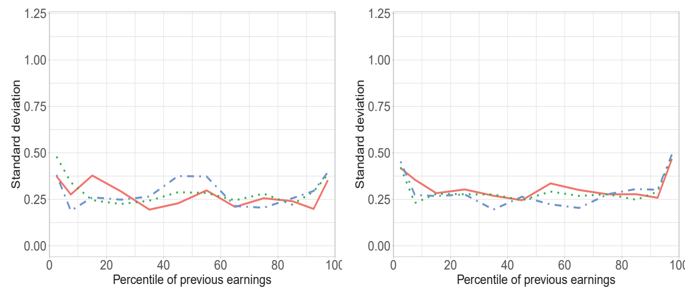


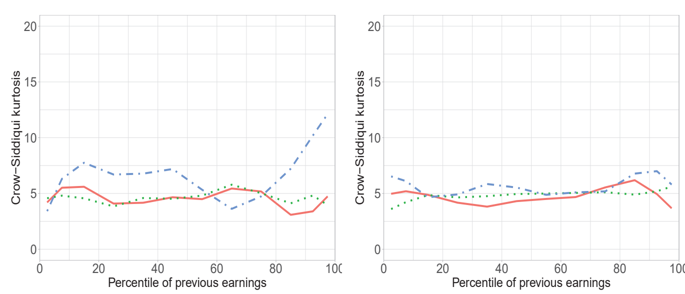
Figure 15: Moments of male earnings changes in the BHPS by marital status. Singles/married are defined as those observed single/cohabiting in both t and $t+1$. Top three panels: by previous earnings. Bottom panel, by previous earnings and age.

Singles Married

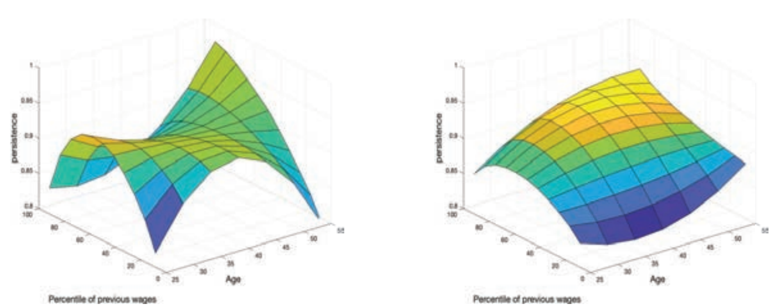


Standard deviation

Kelly's skewness



Crow-Siddiqui kurtosis



Persistence

Figure 16: Moments of female wage changes in the BHPS by marital status. Singles/married are defined as those observed single/cohabiting in both t and $t+1$. Top three panels: by previous earnings. Bottom panel, by previous earnings and age.

A.5 Estimating the distribution of potential female wages

As described in Section 2.1, our preferred imputation for the potential wage of women not working in a given period uses out-of-work benefit income as an instrument for selection into employment.

In this appendix, we first provide more details about the excluded instrument in the selection equation for our preferred imputation procedure. We then contrast its implications with those of the alternative imputation procedures that we discuss in Section 2.1 in the main body of the paper. Finally, we report the results for all our imputation regressions.

Computing out-of-work benefit income To compute potential out-of-work welfare income, we use the UK tax-benefit simulator FORTAX, developed by Shephard (2009) and Shaw (2011). More precisely, we utilize the code by Blundell et al. (2016). We rely on many observables from the BHPS and Understanding Society, including marital status, earnings and hours worked by the partner, age of both partners, number of children in the household and their ages, housing tenure, region, rents paid, childcare expenses, etc. We assume that all homeowners are in council tax band D. FORTAX captures the variations in the tax and benefit system over our sample period.

Implications for potential wages Figures 17 and 18 compare the implied profiles of average earnings over the life cycle and implied distributions of potential wages of our alternative imputation procedures. They show that they are remarkably similar.

Imputation regressions Tables 9-15 report below the participation and imputation regressions for all cases. For **H** and **H-Child** we effectively estimate one participation equation for every year. Here, in the interest of space, we report one equation for all years together and omit some of the lengthier interaction coefficients.

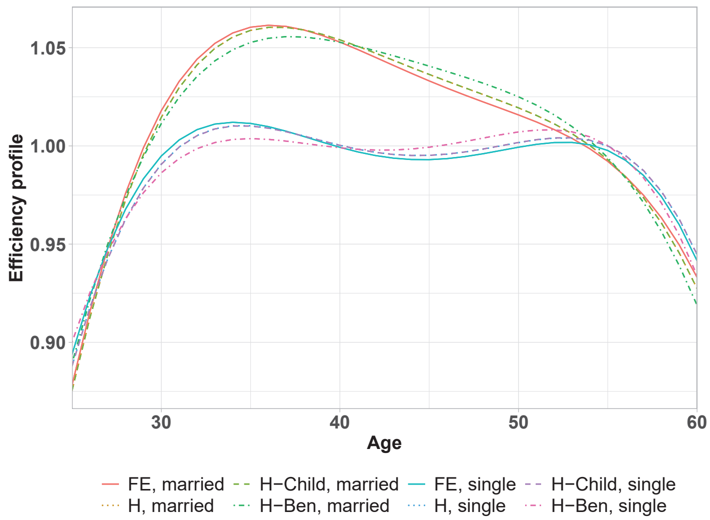


Figure 17: Age-efficiency profile for single and married women (estimated on the Understanding Society survey). **H-Child** and **H** mostly overlap.

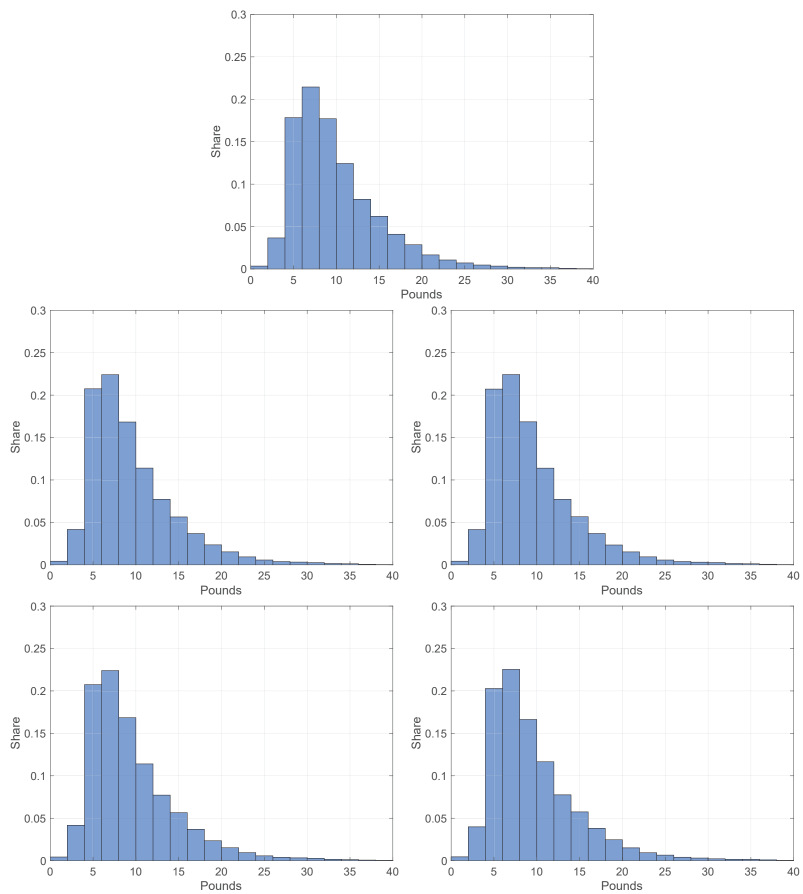


Figure 18: Distribution of women's wages. Top left: observed wages in the data. Middle left: **FE**. Middle right: **H**. Bottom left: **H-Child**. Bottom right: **H-Ben**

(Intercept)	6.7750×10^{8}** (2.1679×10^{8})
year	-1.3536×10^{6}** (4.3365×10^{5})
year^2	1.0142×10^{3}** (3.2530×10^{2})
year^3	−0.3377** (0.1085)
year^4	0.0000** (0.0000)
Born 40s	0.1993*** (0.0366)
Born 50s	0.5359*** (0.0371)
Born 60s	0.6492*** (0.0380)
Born 70s	0.8170*** (0.0417)
Born 80s	1.2436*** (0.1043)
College	0.2928*** (0.0181)
Married	0.0710+ (0.0408)
No. children under 4	−0.4451*** (0.0147)
No. children	−0.1756*** (0.0076)
Potential benefit	−0.0014*** (0.0002)
married:pot.benefit	−0.0002 (0.0002)
Num.Obs.	58 124
RMSE	0.37

Table 9: Participation equation - **H-Ben**

year	-2.4594×10^{3}*** (2.2703×10^{2})
year^2	1.2296*** (0.1135)
year^3	−0.0002*** (0.0000)
College	0.0780*** (0.0222)
Married	0.0118 (0.0098)
No. children under 4	0.0687*** (0.0131)
No. children	−0.0078 (0.0062)
Inv. Mills Ratio	−0.1662** (0.0593)
Num.Obs.	43 198
R2	0.754
R2 Within	0.159
RMSE	0.25
FE: individual	X

Table 10: Wage imputation - **H-Ben**

(Intercept)	−0.514*** (0.026)
Homeowner	0.572*** (0.012)
Age youngest child	0.009*** (0.001)
No. children	−0.258*** (0.007)
No. children under 4	−0.371*** (0.013)
College	0.391*** (0.016)
Married	0.159*** (0.013)
Born 40s	0.332*** (0.026)
Born 50s	0.783*** (0.026)
Born 60s	1.040*** (0.027)
Born 70s	1.132*** (0.029)
Born 80s	1.213*** (0.067)
Num.Obs.	70 165
Log.Lik.	−38 634.686
F	769.066
RMSE	0.43

Table 11: Participation equation - **H**

Age	0.2019* (0.0913)
Age^2	−0.0051 (0.0034)
Age^3	0.0001 (0.0001)
Age^4	0.0000 (0.0000)
Experience	0.0250*** (0.0042)
Experience^2	0.0005** (0.0002)
Experience^3	0.0000*** (0.0000)
No. children	−0.0358*** (0.0061)
Age youngest child	0.0001 (0.0004)
Married	0.0131 (0.0096)
Inv. Mills Ratio	0.0141 (0.0184)
Num.Obs.	43 250
R2	0.755
R2 Within	0.165
RMSE	0.25
FE: individual	X

Table 12: Wage imputation - **H**

(Intercept)	−0.362*** (0.030)
Years since child:1	−0.108 (0.132)
Years since child:2	−0.054 (0.108)
Years since child:3	−0.238* (0.101)
Years since child:4	−0.271** (0.091)
Years since child:5	−0.176* (0.087)
Years since child:6	−0.245** (0.084)
Years since child:7	−0.134 (0.082)
Years since child:8	−0.046 (0.081)
Years since child:9	−0.002 (0.080)
Years since child:10	−0.036 (0.080)
Years since child:11	−0.043 (0.077)
Years since child:12	−0.023 (0.078)
Years since child:13	0.016 (0.076)
Years since child:14	0.129+ (0.075)
Years since child:15	0.009 (0.074)
Years since child:16	−0.005 (0.074)
Years since child:17	−0.027 (0.076)
Years since child:18	0.031 (0.050)
Children in hh	0.019 (0.040)
Grandparents in hh	−0.097*** (0.027)
Husband has job	0.426*** (0.014)
Homeowner	0.487*** (0.012)
Age youngest child	0.005*** (0.001)
No. children	−0.303*** (0.010)
No. children under 4	−0.319*** (0.018)
College	0.376*** (0.016)
Married	−0.192*** (0.021)
Born 40s	0.284*** (0.026)
Born 50s	0.687*** (0.026)
Born 60s	0.944*** (0.027)
Born 70s	1.051*** (0.029)
Born 80s	1.152*** (0.068)
Y since child:1:married	0.106 (0.135)
Y since child:2:married	0.132 (0.112)
Y since child:3:married	0.162 (0.103)
Y since child:4:married	0.223* (0.093)
Y since child:5:married	0.122 (0.088)
Y since child:6:married	0.237** (0.084)
Y since child:7:married	0.135+ (0.081)
Y since child:8:married	0.153+ (0.080)
Y since child:9:married	0.172* (0.080)
Y since child:10:married	0.232** (0.079)
Y since child:11:married	0.258*** (0.076)
Y since child:12:married	0.307*** (0.076)
Y since child:13:married	0.326*** (0.074)
Y since child:14:married	0.233** (0.072)
Y since child:15:married	0.214** (0.073)
Y since child:16:married	0.185* (0.073)
Y since child:17:married	0.164* (0.075)
Y since child:18:married	−0.031 (0.036)
Num.Obs.	70 165
Log.Lik.	−37 998.485
F	188.537
RMSE	0.43

Table 13: Participation equation - **H-Child**

Age	0.2070* (0.0908)
Age^2	−0.0053 (0.0034)
Age^3	0.0001 (0.0001)
Age^4	0.0000 (0.0000)
Experience	0.0251*** (0.0042)
Experience^2	0.0005** (0.0002)
Experience^3	0.0000*** (0.0000)
No. children	−0.0340*** (0.0058)
Age youngest child	0.0001 (0.0004)
Married	0.0122 (0.0095)
Inv. Mills Ratio	0.0035 (0.0144)
Num.Obs.	43 250
R2	0.755
R2 Within	0.165
RMSE	0.25
FE: individual	X

Table 14: Wage imputation - **H-Child**

Age	0.2109* (0.0908)
Age^2	−0.0054 (0.0034)
Age^3	0.0001 (0.0001)
Age^4	0.0000 (0.0000)
Experience	0.0249*** (0.0042)
Experience^2	0.0005** (0.0002)
Experience^3	0.0000*** (0.0000)
No. children	−0.0336*** (0.0052)
Age youngest child	0.0001 (0.0004)
Married	0.0188+ (0.0104)
Husband has job	−0.0101 (0.0069)
Num.Obs.	43 250
R2	0.755
R2 Within	0.165
RMSE	0.25
FE: individual	X

Table 15: Wage imputation - **FE**

B Estimation and features of the earnings processes

B.1 Comparing the non-linear and canonical processes

As described in De Nardi et al. (2020), the canonical process, described in Equation 4, can be specified as a restricted version of the NL process in Equation 5, where:

$$Q_{z_{i,t}}(v_{it}|z_{i,t-1},t) = \rho z_{i,t-1} + \sigma_\nu \phi^{-1}(v_{it}) \tag{19}$$

$$Q_\epsilon(e_{it}) = \sigma_\epsilon \phi^{-1}(e_{it}), \tag{20}$$

where $\phi^{-1}(.)$ is the inverse of the cumulative density function of a standard, normal distribution. This specification allows to clearly see the restrictions the canonical process imposes on the earnings process:

1. *Age-independence* (stationarity) of the autoregressive coefficient ρ and of the shock distributions (both normal with constant standard deviations σ_ν and σ_ϵ), which imply age-independence of the second and higher moments of the conditional distributions of both the transitory and the persistent component.

2. *Normality* of the shock distributions ($\phi^{-1}(\cdot)$).

3. *Linearity* of the process for the persistent component, which can be seen in the additive separability of equation 19 into the conditional expectation—the first addendum—and an innovation independent of $z_{i,t-1}$, and (b) the linearity of the conditional expectation in $z_{i,t-1}$. Under separability, deviations of z_{it} from its conditional expectation are just a function of the innovation ν_{it}. As a consequence, all conditional centered second and higher moments are independent of previous realizations of z.

One further way to understand the role of nonlinearity is in terms of a generalized notion of persistence

$$\rho(q|z_{i,t-1},t) = \frac{\partial Q_z(q|z_{i,t-1},t)}{\partial z_{i,t-1}} \tag{21}$$

which measures the persistence of $z_{i,t-1}$ when it is hit by a shock that has rank q. In the canonical model, $\rho(q|z_{i,t-1},t) = \rho$, independently of both the past realization of $z_{i,t-1}$ and of the shock rank q. Instead, the general model allows persistence to depend both on the past realization $z_{i,t-1}$, but also on the sign and magnitude of the shock realization. Basically, in the nonlinear model shocks are allowed to wipe out the memory of past shocks or, equivalently, the future persistence of a current shock may depend on future shocks.

Of course, a similar unrestricted representation can be used for the transitory component ϵ_{it} and the initial condition η_1, with the only difference that they are not persistent.

We proceed in two steps. First, we use the quantile-based panel data method proposed by Arellano et al. (2017) to estimate a non-parametric model that allows for age-dependence, non-normality and nonlinearity, and that can be applied in datasets of moderate sample size like the PSID. This step gives us quantile functions for both the two (persistent and transitory) component of earnings (see the next section, Appendix B.2 for details on the estimation). Second, we use the two quantile functions to simulate histories for the two earnings components and proceed to estimate, for the persistent component, a discrete Markov-chain approximation, which can then be easily introduced in a structural model.

B.2 Estimation

Following Arellano et al. (2017), we parameterize the quantile functions for the three variables as low order Hermite polynomials

$$Q_\epsilon(q|age_{it}) = \sum_{k=0}^{K} a_k^\epsilon(q)\psi_k(age_{it}) \tag{22}$$

$$Q_{z_1}(q|age_{i1}) = \sum_{k=0}^{K} a_k^{z_1}(q)\psi_k(age_{i1}) \tag{23}$$

$$Q_z(q|z_{i,t-1}, age_{it}) = \sum_{k=0}^{K} a_k^z(q)\psi_k(z_{i,t-1}, age_{it}) \tag{24}$$

where the coefficients $a_k^i(q)$, $i = \epsilon, z_1, z$, are modeled as piecewise-linear splines in q on a grid $\{q_1 < \ldots < q_L\} \in (0, 1)$.[11] The intercept coefficients $a_0^i(q)$ for q in $(0, q_1]$ and $[q_L, 1)$ are specified as the quantiles of an exponential distribution with parameters λ_1^i and λ_L^i.

If the two earnings components ϵ_{it} and z_{it} were observable one could compute the

[11]Following Arellano et al. (2017), we use tensor products of Hermite polynomials of degrees (3,2) in $z_{i,t-1}$, and age for $Q_z(q|z_{i,t-1}, age_{it})$ and second-order polynomials in age for $Q_\epsilon(q|age_{it})$ and $Q_{z_1}(q|age_{i1})$.

polynomial coefficients simply by quantile regression for each point of the quantile grid q_j. To deal with the latent earnings components, the estimation algorithm starts from an initial guess for the coefficients and iterates sequentially between draws from the posterior distribution of the latent persistent components of earnings and quantile regression estimation until convergence of the sequence of coefficient estimates.

B.3 Persistent and transitory earnings

In this section, we compare the non-linear and non-normal features of the BHPS data and the persistent and transitory components that result from the Arellano et al. (2017) decomposition.

Starting with male earnings, persistence is lowest for the young and for the lowest earners both for the BHPS data and the persistent component (Figure 19). As expected, the persistent component displays a larger overall persistence than the data, but shows the same patterns by age and over the earnings distribution.

Figure 20 shows the standard deviation, skewness, and kurtosis of earnings changes for the BHPS data and their persistent component. Their persistent component preserves most of the features of non-normality that are present in the data and the dependence on previous earnings realizations. The main difference lies in the Crow-Siddiqui kurtosis, which is significantly larger for the persistent component than in the raw data.

Transitory shocks, that we consider to be measurement error, are very leptokurtic, in particular for male earnings, and display negative skewness (see Figure 21).

Women's wages display similar patterns (see Figures 22, 23, and 24). The most noticeable difference is that the persistence of the persistent component is relatively high and close to 1, but still replicates the inverted U-shape by previous wages that we observe in the data.

Finally, in Figures 25 and 26 we show that most of the differences in dynamics between men's earnings and women's wages are also present if we compare male and female earnings. For example, the profile of persistence over the earnings distribution is much flatter for women than for men.

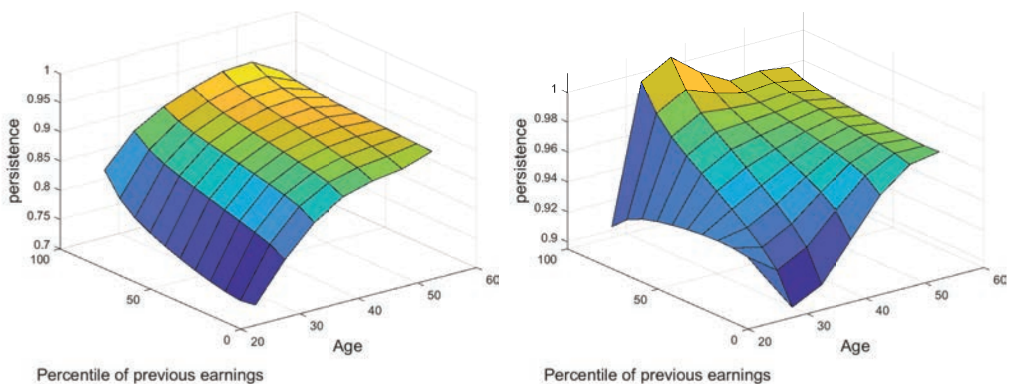


Figure 19: Non-linear persistence of male earnings by age and previous earnings in the BHPS. Left, data; right, persistent component

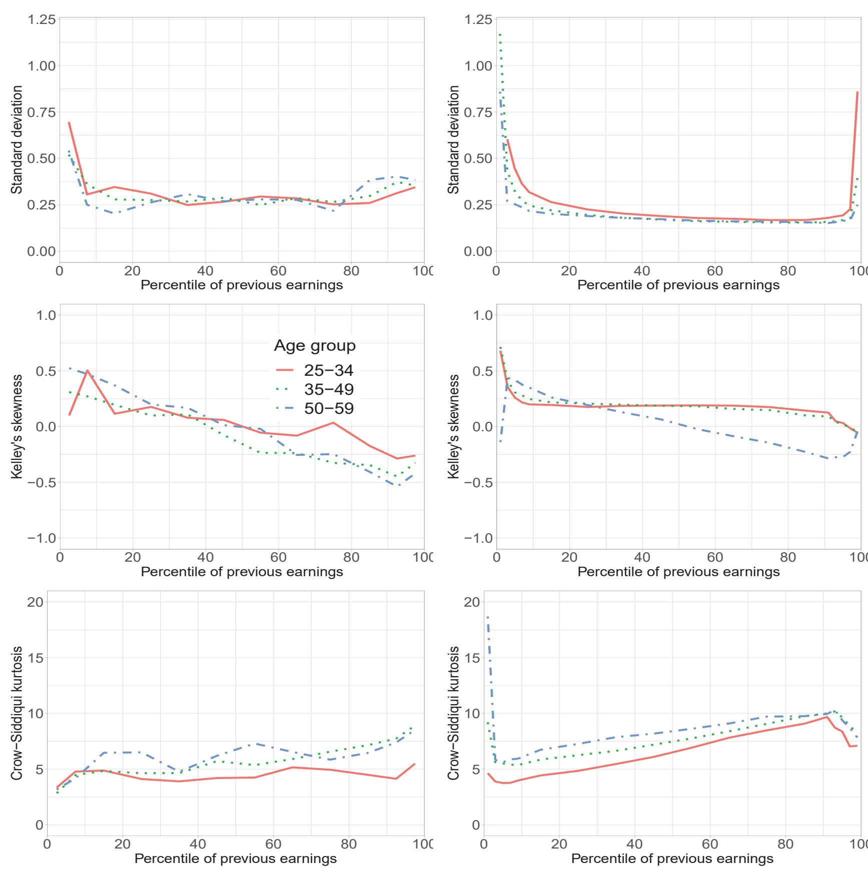


Figure 20: Standard deviation (top), skewness (middle) and kurtosis (bottom) of male earnings changes in the BHPS. Left, data; right, persistent component

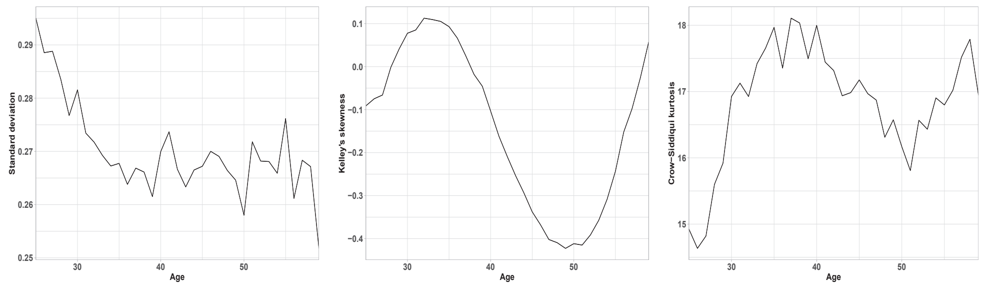


Figure 21: Transitory shock to male earnings: standard deviation, skewness and kurtosis by age

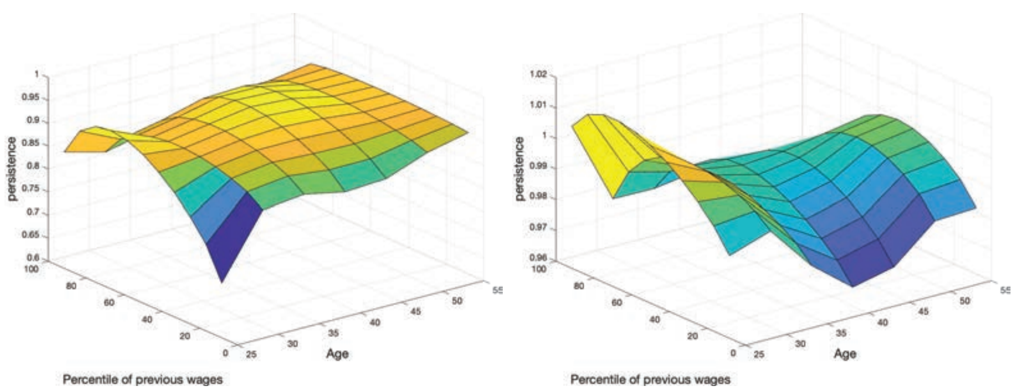


Figure 22: Non-linear persistence of female wages by age and previous wages in the BHPS. Left, data; right, persistent component

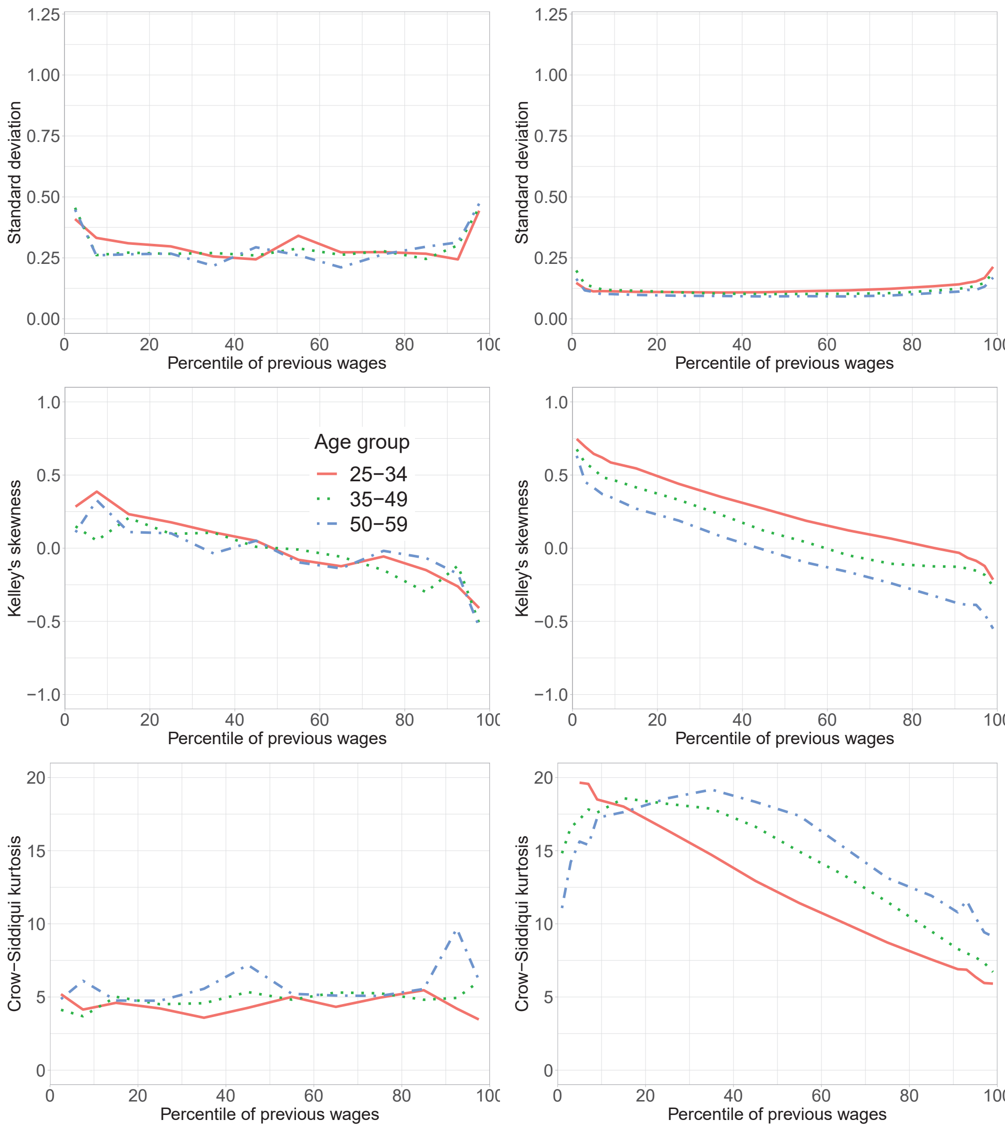


Figure 23: Standard deviation (top), skewness (middle) and kurtosis (bottom) of female wage changes in the BHPS. Left, data; right, persistent component

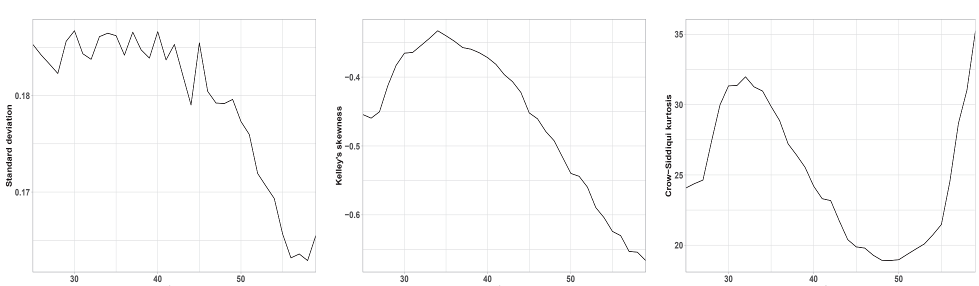


Figure 24: Transitory shock to female wages: standard deviation, skewness and kurtosis by age

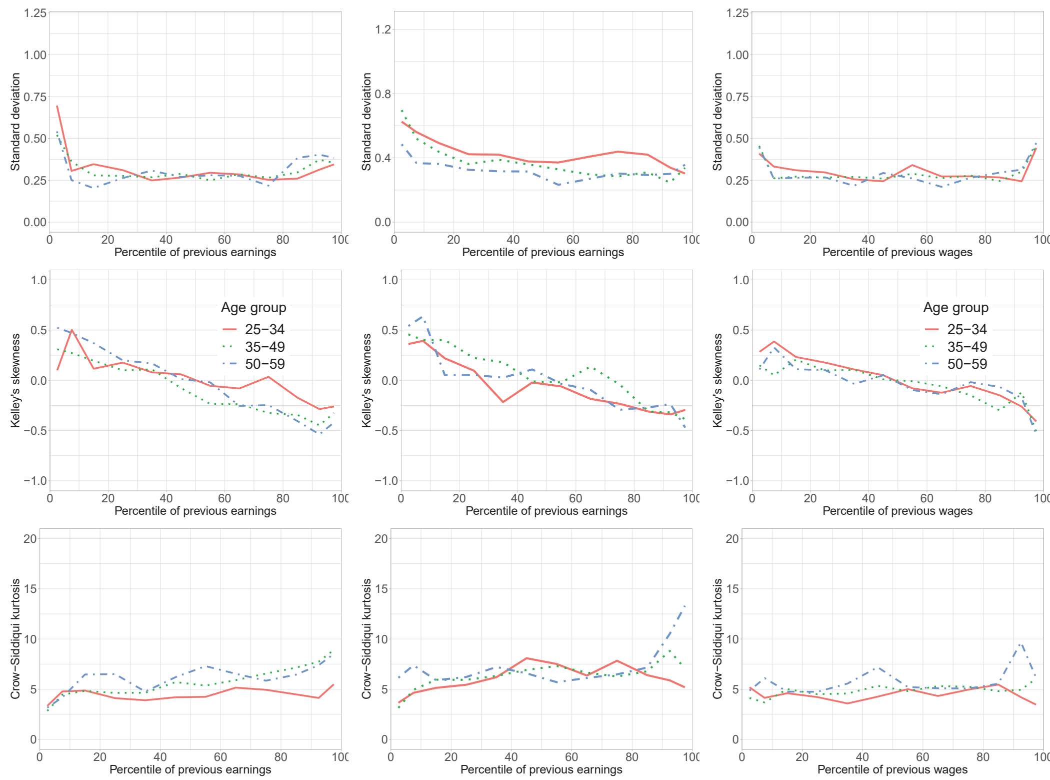


Figure 25: Standard deviation (top), skewness (middle) and kurtosis (bottom). Left: male earnings; middle: female earnings; right: female wages

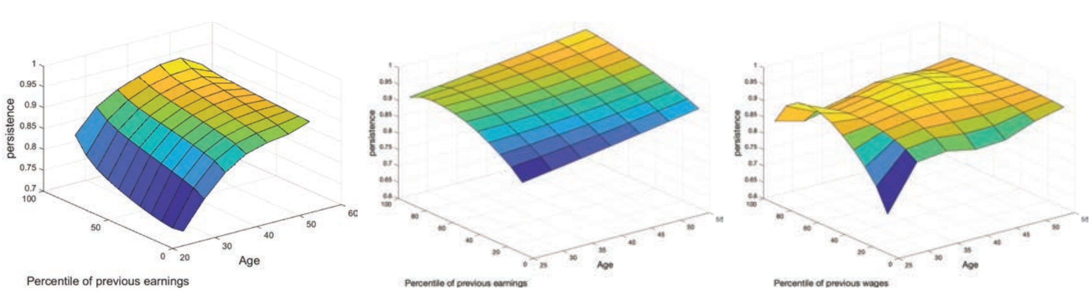


Figure 26: Non-linear persistence of male earnings (left), female earnings (middle) and female wages (right), by age and percentile of previous wages, BHPS data

C Other model inputs

C.1 Marriage and divorce

	Dependent variable:	
	marriage	divorce
	(1)	(2)
Age	-0.032^{***}	-0.018^{***}
	(0.002)	(0.003)
(log) Wife's imputed wage	0.041	-0.017
	(0.035)	(0.046)
(log) Husband's income		-0.007^{***}
		(0.002)
Constant	-0.161	-1.156^{***}
	(0.103)	(0.136)
Observations	11,350	22,014
Log Likelihood	-3,358.432	-1,898.706
Akaike Inf. Crit.	6,722.864	3,805.412
Note:	*p<0.1; **p<0.05; ***p<0.01	

Table 16: Probability of marriage and divorce (probit regressions) between t-1 and t, conditional on income at t-1

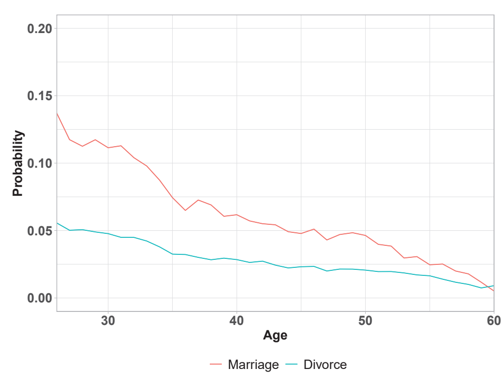


Figure 27: Marriage probabilities for single women, and divorce probabilities for married women, by age (BHPS data)

	Dependent variable:
	(log) Earnings of husband in t
(log) Woman's wage in t	0.325***
	(0.019)
Constant	9.359***
	(0.041)
Observations	3,728
R^2	0.076
Adjusted R^2	0.075
Residual Std. Error	0.506 (df = 3726)
F Statistic	304.354*** (df = 1; 3726)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 17: Correlation of husband's earnings and wife's wages before 30

	Dependent variable:
	(log) Earnings of husband in t
(log) Woman's wage in t-1	0.272***
	(0.055)
Constant	9.480***
	(0.123)
Observations	386
R^2	0.059
Adjusted R^2	0.056
Residual Std. Error	0.524 (df = 384)
F Statistic	23.987*** (df = 1; 384)
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 18: Correlation of husband's earnings and wife's wages at marriage

	Dependent variable:	
	log wealth of partner	
	(1)	(2)
Age	−0.003	−0.033
	(0.058)	(0.062)
(log) Woman's wage	2.362**	
	(1.074)	
(log) Men's income		1.688**
		(0.743)
Constant	3.701	−9.011
	(2.929)	(7.458)
Observations	86	117
R^2	0.055	0.044
Adjusted R^2	0.032	0.027
Residual Std. Error	4.257 (df = 83)	4.705 (df = 114)
F Statistic	2.424* (df = 2; 83)	2.625* (df = 2; 114)
Note:		*p<0.1; **p<0.05; ***p<0.01

Table 19: Correlation between partner's wealth before marriage and income of reference person at marriage year

C.2 Children

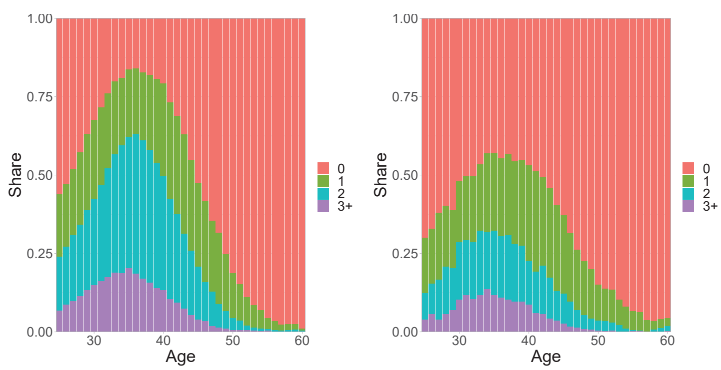


Figure 28: Distribution of number of children in the household, by age of the mother. Left: married mothers; right: single mothers

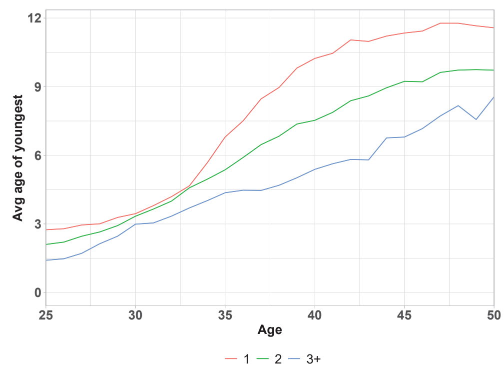


Figure 29: Average age of youngest cohabiting child by age of mother and number of children

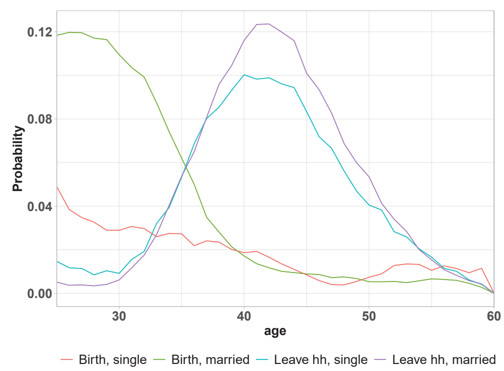


Figure 30: Probability that a child arrives (birth) and leaves (leave hh) a household, by age and marital status (of the woman), unconditionally on today's child status (BHPS data)

C.3 Mortality risk

Figure 31 shows the mortality risk by age and marital status in the model. Although there is information about mortality in the BHPS data, its sample size is too small to obtain reliable estimates for death probabilities that are age, gender, and marital-status specific. Thus, we turn to the life tables data from the Human Mortality Database (1980-2010), which are reported separately by gender and age. Then, to incorporate the increased mortality risk for singles, we estimate the average gap in mortality probabilities between singles and married people in the BHPS during the retirement period. We assume that this gap is constant during adult life, and compute the death probabilities for single and married people that are consistent with this gap and the observed mortality probabilities in the life tables. We report them in Figure 31.

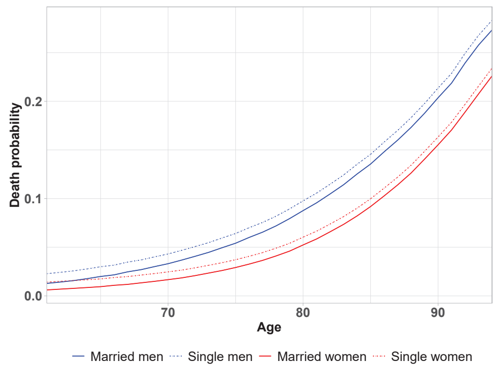


Figure 31: Annual death probabilities by age, gender, and marital status. Source: Human Mortality Database and BHPS data.

C.4 Average male earnings and female wages

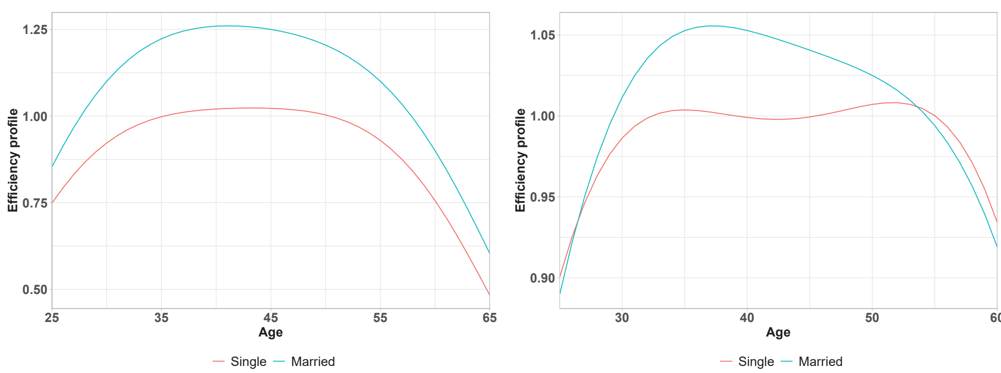


Figure 32: Age-efficiency profiles, left: men's earnings; right: women's wages. For this representation, both are individually normalized so that their average is 1

C.5 Population shares, data vs. model

In our sample, after excluding the retirees, long-term disabled, and full-time students, 9% of single men and 4% of married men display zero earnings in a given calendar year. The corresponding shares are 20% and 19.9% for single and married women, respectively. Within the male working population, 4.3% of singles and 3.2% of married people work part-time. The corresponding shares for women are 25% and 41%, respectively. As a result, we have chosen to model the labor supply decision of women explicitly and to assume that men always work full time.

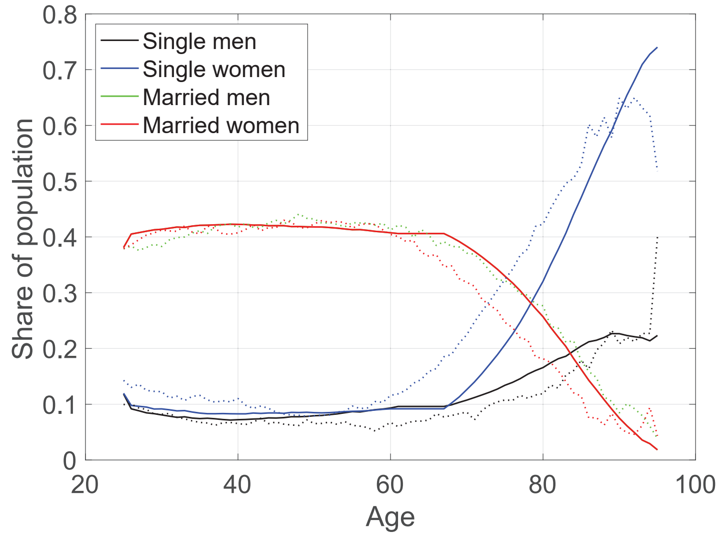


Figure 33: Share of people by age, gender, and marital status. Solid lines: model outcomes, dotted lines: data. In the model, the number of married men and married women at each age is identical by construction. Data: BHPS, whole sample

D UK Benefit system, details

Table 20 provides a brief overview of the main benefits for the working age population in the United Kingdom before the introduction of Universal Credit in 2016.[12]

In our model, in-work benefits are meant to capture the Working Tax Credit, while income support replicates a variety of benefits that low-income people receive, including Income-based Jobseeker's Allowance, Income Support, Housing Benefits, Child Benefits, and Child Tax Credits.

The tapering rate for in-work benefits corresponds to the statutory tapering rate for the Working Tax Credit (0.41). For income support, we compute an average tapering rate ω of the different benefits it summarizes, considering their respective sizes, tapering rates, and eligibility criteria, including how access to one of the benefits impacts the entitlement to the others. We do so in the following way. First, we calculate the benefit entitlement B_i^k by demographic group k (gender, marital status, and number of children) and household labor income y_i. We do so under the assumption that the household is eligible for all of the benefits that compose our income support, also taking into account that a household can only claim either Income-based Jobseeker's Allowance or Income Support, but not both at

[12]Given the gradual and too recent phase-in of Universal Credit, it would not have been appropriate to calibrate our steady-state benchmark economy to the post-2016 period.

the same time. We additionally assume that the household would be getting Working Tax Credit whenever eligible, which affects their eligibility criteria for other benefits (namely, Child Tax Credits and the Working Tax Credit are considered as income for purposes of computing eligibility for Income Support and Housing Benefits).

We then find the β_0^k and β_1^k that minimize:

$$\sum_i (B_i^k - max(\beta_0^k - \beta_1^k y_i^{hk}, 0))^2 \tag{25}$$

where the sum i is taking over all possible income levels between 0 and £100,000. We then obtain our estimate of ω by weighing the different β_1^k by the relative sizes in the population of each k group. The average tapering rate is then $-\beta_1$ is 0.70, which also corresponds to the tapering rate for couples with zero children.

E Additional model implications

E.1 Observed wages in the data and in the model

Figure 34 reports the distribution of potential wages that we use in our model, computed using our Heckman selection correction (left), the implied distribution of observed wages that the model delivers, under the assumption that we can only observe wages for women who choose to work (center), and the distribution of female wages in the data, which we can only observe for those who are actively participating (right).

Our model-implied distribution of observed wages is closer to the data than the distribution of potential wages, thus suggesting that the model replicates the patterns of selection in the data well. For instance, looking at the second bar of these histograms (which are computed in such a way that the binning is identical for all three), one can observe that it is taller in the potential wage distribution (a lot of women have low potential wages), but lower and closer to the data in the model-implied observed wage distribution (thus suggesting that many women select out of the labor force when they receive a low wage realization).

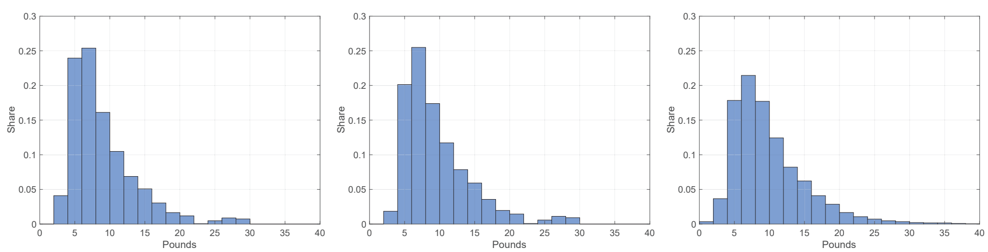


Figure 34: Distribution of women's wages. Left, potential wages in the model; middle: observed wages in the model; right: observed wages in the data.

Benefit	Time period	Eligibility (income)	Tapering	Wealth test	£M (2016)
Benefits for the unemployed					
Jobseeker's Allowance (Contributory)	1996-today	Work < 16h/week	100%	No	306
Jobseeker's Allowance (Income-based)	1996-today	Work < 16h/week	100%	Yes	2000
Benefits for low-income people					
Income Support		Work < 16h/week	100%	Yes	2700
Housing benefit		Tapering starts after JSA amount	65%	16k	24300
Council Tax Benefit	-2013	Being on IS, JSA, etc.	No	Yes	
Benefits for families					
Child benefit		Income < £50k	No	No	11300
Statutory Maternity Pay		None	No	No	2300
Maternity Allowance (Contributory)		Min £30 pw	No	No	443
Tax credits					
Child Tax Credit	2003-	Taper from £16,105 (2014)	41%	No	21700
Working Tax Credit	2003-	Working FT, taper from £6,420	41%	No	5900
Benefits for the sick and disabled					
ESA	2011-today	Work <16h/week	100%	No	14300
Personal Independence Payment	2013-	Work capability assessment	-	No	3000
Disability Living Allowance	-2013	Unable to work	-	No	13200
Carer's Allowance		No	No	No	2600
Industrial Injuries Benefits		Depends on disablement rate	No	No	869

Table 20: Main benefits for working age population in the UK (source: Hood and Norris Keiller (2016))

E.2 Labor market participation by number of children

Figure 35 shows the share of women that are working, conditional on the number of children they have, in the data and in the two versions of our model. We find that both are good at capturing the general patterns of participation by number of children, particularly for the largest groups (those with 0, 1 or 2 children). Our model is also successful at capturing the larger decrease in labor market participation for singles than for couples as the number of children increases. However, the model overestimates the share of women with 3 children or more who are working for both earnings process. This mismatch is related to the assumption that the maximum number of children is 3 in our model, but in the data the number of children might be larger than 3, which further discourages female labour supply. Furthermore, older cohorts in the data are both more likely to have more children and stay at home, and in our model we abstract from cohort heterogeneity.

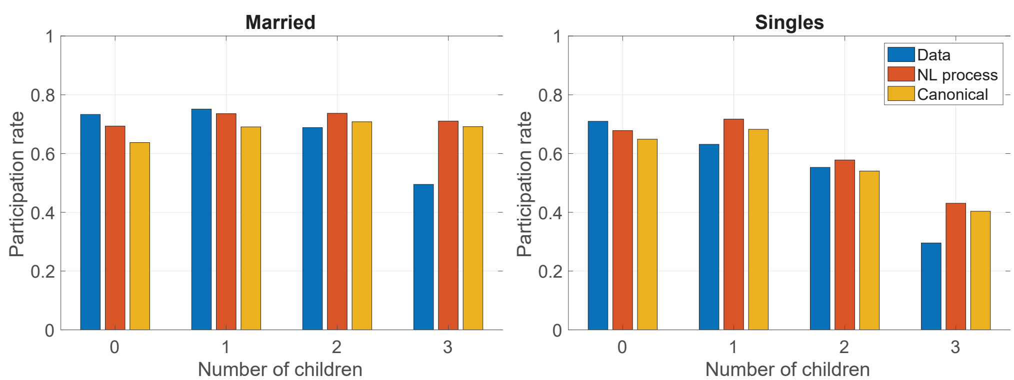


Figure 35: Women's labour market participation, by number of children.

E.3 Persistence of female labor force participation

In our model, we introduce heterogeneity in the disutility of work in line with previous literature on female labor supply, including Keane and Wolpin (2010), Blundell et al. (2016), or Adda, Dustmann and Stevens (2017), and as a parsimonious way to capture the large amount of heterogeneity in the data. For simplicity and transparency, we assume that there are two types of women, one with higher disutility from work than the other.

A way of evaluating whether the size of our fixed costs of work and their heterogeneity are quantitatively reasonable is to look at the dynamics of female labor force participation.

Aspects of the data that are pertinent for these purposes are the persistence of (a) being on benefits and (b) being unemployed or out of the labor force. In the data, the persistence of benefit receipt is 0.78; in the model it is also 0.78. In the data, the persistence of the unemployed/out-of-labor-force status for women is 0.80; in the model it is 0.88. Both are non-targeted moments by our estimation strategy.

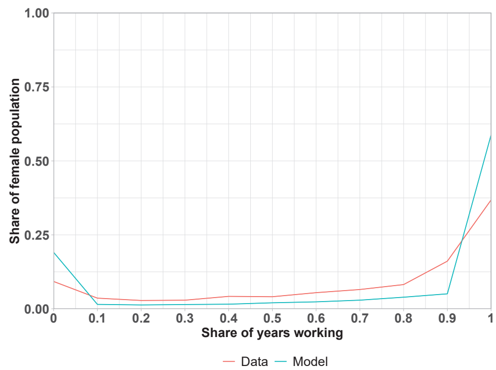


Figure 36: Distribution of the female population over the number of years worked in a 10-year period. Data: BHPS balanced sample of women who are observed for 10 years in a row.

In addition, Figure 36 shows, in a balanced 10-year sample, how many women work every year (1.0), don't work at all (0.0), or intermediate cases during that 10 year span. Both in the data and in the model there is large heterogeneity: some women work all the time, while others never work at all. The model does a reasonably good job of matching this untargeted distribution.

E.4 Universal Credit, canonical process

In this section, we report the welfare effects of the introduction of Universal Credit under the canonical wage process. As described in Section 4, in our main results with the NL process, we keep the change to Universal Credit budget neutral by multiplying all allowances with a proportional scaling factor of 0.9. For the purposes of this section, we keep budget neutrality under the canonical process, which implies that we scale these allowances by 0.82

Under the canonical wage process, the switch to Universal Credit generates a drop in full-time labor force participation and a large rise in part-time labor force participation, particularly at older ages (Figure 37).

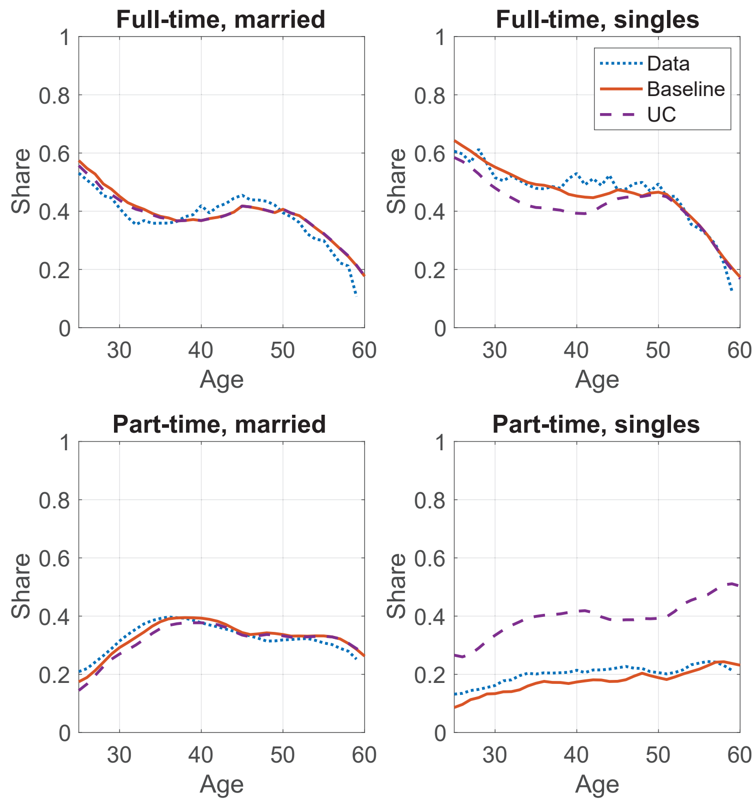


Figure 37: Labor force participation under canonical process: Universal Credit vs baseline, universal credit.