Review of Economic Studies (2024) **00**, 1–27 doi:10.1093/restud/rdae042

Advance access publication 15 April 2024

Wage Risk and Government and Spousal Insurance

Mariacristina De Nardi

University of Minnesota, Federal Reserve Bank of Minneapolis, CEPR, and NBER, USA

Giulio Fella

Queen Mary University of London and University of Bologna, CEPR, CFM, and IFS, Italy

and

Gonzalo Paz-Pardo

European Central Bank, Germany

First version received December 2020*; Editorial decision November* 2023*; Accepted April* 2024 *(Eds.)*

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 labour 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—IW versus income floor. The canonical wage process underestimates wage persistence for women and thus implies that IW 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 U.K. This stresses that allowing for rich wage dynamics is important to properly evaluate policy.

JEL codes: D1, D12, D14, D15, H11, H2

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 labour 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 labour 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 labour supply of both partners to at least partly counteract individual wage fluctuations. In addition, couples benefit from economies of scale in consumption.

The editor in charge of this paper was Kurt Mitman.

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 modelling 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 U.K. survey data from the British Household Panel Study (BHPS), at the household level, and U.K. 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 U.K. 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 labour supply behaviour.[1]

We then develop a dynamic, structural, life-cycle model with an active female labour-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 (NL) 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 (IW) benefits, as well as the rate at which benefits should phase-out as a function of labour income. Our calibration matches key aspects of the data that include government policy and household labour market outcomes over the life cycle, during the time period preceding the 2016 Universal Credit (UC) benefit reform in the U.K.

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 labour 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 U.C. reform. It privileges the income floor with a very limited role for IW 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 IW benefits and effectively eliminate income support

1. Guvenen *et al*. (2021) document rich dynamics for pre-tax individual earnings in the U.S., Arellano *et al*. (2017) for household pre-tax earnings in the U.S. and Norway, De Nardi *et al*. (2020) for household disposable earnings in the U.S. Ozkan *et al*. (2017) and De Nardi *et al*. (2021) study the relative contribution of wages and hours to male earnings dynamics, respectively, in Norway, and in the Netherlands and the U.S.

(IS). 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 labour market by lowering their out-of-work benefits and increasing their IW benefits. In reality, women's wages are 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 labour 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 U.C. benefit reform that was subsequently introduced in the U.K. 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 U.K. We find that, irrespective of the wage process the move to U.C. implies overall welfare gains, but significantly reduces welfare for single men.

Because many women do not work and the relevant wage dynamics for our analysis are the potential ones, rather than those observed just for labour 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 *et al*. (2003), Arellano and Bonhomme (2017), and Chiappori *et al*. (2018). However, we also evaluate the robustness of our results to alternative imputation procedures and show that, while the NL 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 labour supply. Keane and Wolpin (2010) study the effect of the U.S. welfare system on women's welfare participation, labour supply, marriage, fertility, and schooling. Blundell *et al*. (2016) study how the U.K. tax and welfare system affects the career choices of women. Guner *et al*. (2012a) and Bick and Fuchs-Schündeln (2017) investigate the effect of taxation on household labour supply, while Guner *et al*. (2012b) evaluate gender-based taxes, Nishiyama (2019) and Groneck and Wallenius (2017) evaluate Social Security spousal provisions, and Borella *et al*. (2023) study the effects of marriage-related taxes and Social Security rules for different cohorts of women whose labour supply behaviour has been changing. The paper closer in spirit to ours is possibly Guner *et al.* (2020) that compares the implications of IW childcare credits to those of child benefits independent of the mother's labour 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 labour supply elasticities, we take men's labour supply as exogenous while we model women's labour supply. Thus, in our empirical analysis, we study men's earnings and women's wages.

Our main data source is the BHPS. The BHPS is a household survey of the U.K. 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 Panel Study of Income Dynamics (PSID) in the U.S., because instead of being asked about their total labour earnings in the last 12 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 labour 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 U.K. 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 36,042 usable observations (pairs of earnings in t and $t+1$) for men and 54,178 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 U.K. 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 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 labour market participants as a function

2. The typical threshold is around 5% 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 NESPD, an administrative data set with individual data from the U.K. Social Security. It shows that the results from NESPD are very similar to those from the BHPS.

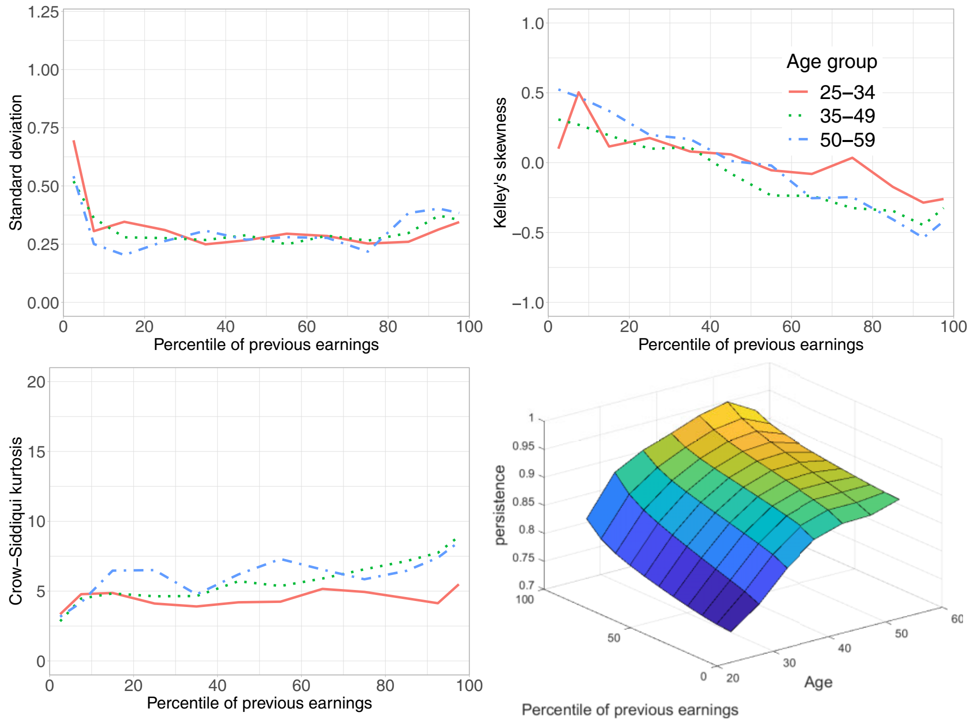

FIGURE 1

Moments of male earnings changes in BHPS data

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.

2.1. *Estimating the distribution of potential female wages*

There is an extensive empirical literature that finds that the elasticity of women's labour 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 labour 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

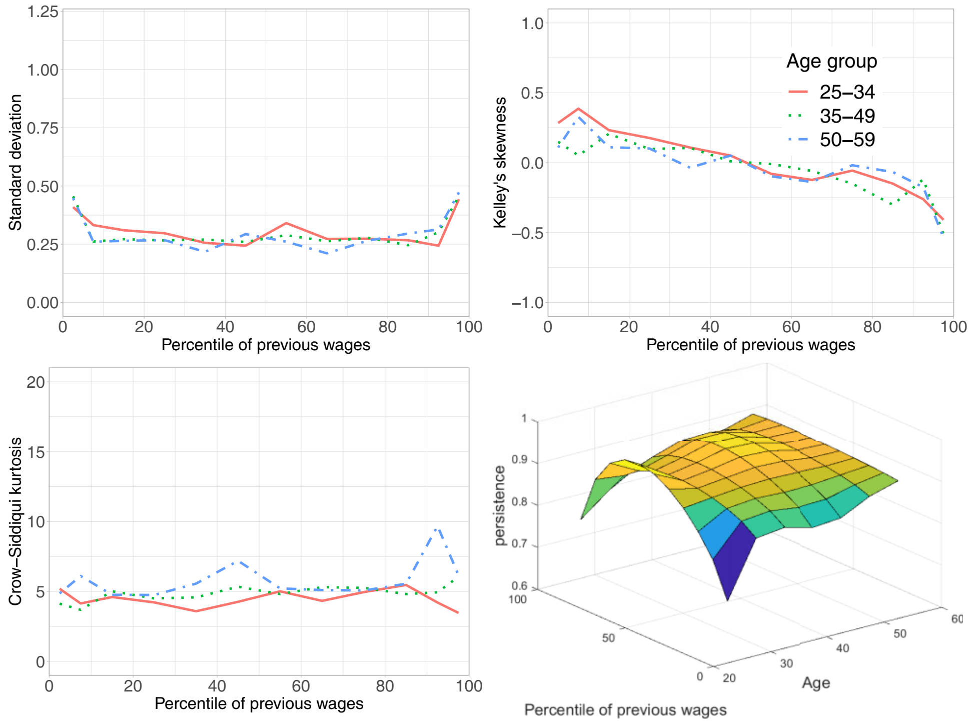

FIGURE 2

Moments of female wage changes in BHPS data

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 NL 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.

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).

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.

2.2. *Estimating the wage and earnings processes*

Our structural model of household behaviour 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 labour 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

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

Our flexible, or 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

TABLE 1
Estimates for the canonical processes

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

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

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

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 *et al.* (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 1 shows the estimated parameters for male earnings and female wages for the canonical process. To estimate the flexible NL process, we follow Arellano *et al.* (2017). Online Appendix B.3 shows how the persistent component preserves the non-normal and NL 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 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

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.

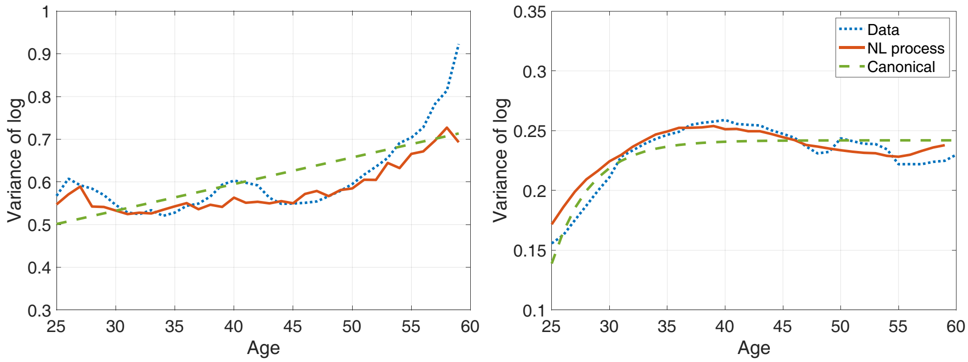

FIGURE 3

Variance of log earnings for men (left) and log wages for women (right). Dotted line, data; solid line, NL processes; dashed line, canonical process

(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 NLs, 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 labour 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 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 NL and canonical process imply distinctly different optimal benefit systems, the differences implied by our alternative imputation procedures are minor.

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 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 NL one, which we have described in the previous section.

Marital status evolves stochastically as in Cubeddu and Ros-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.

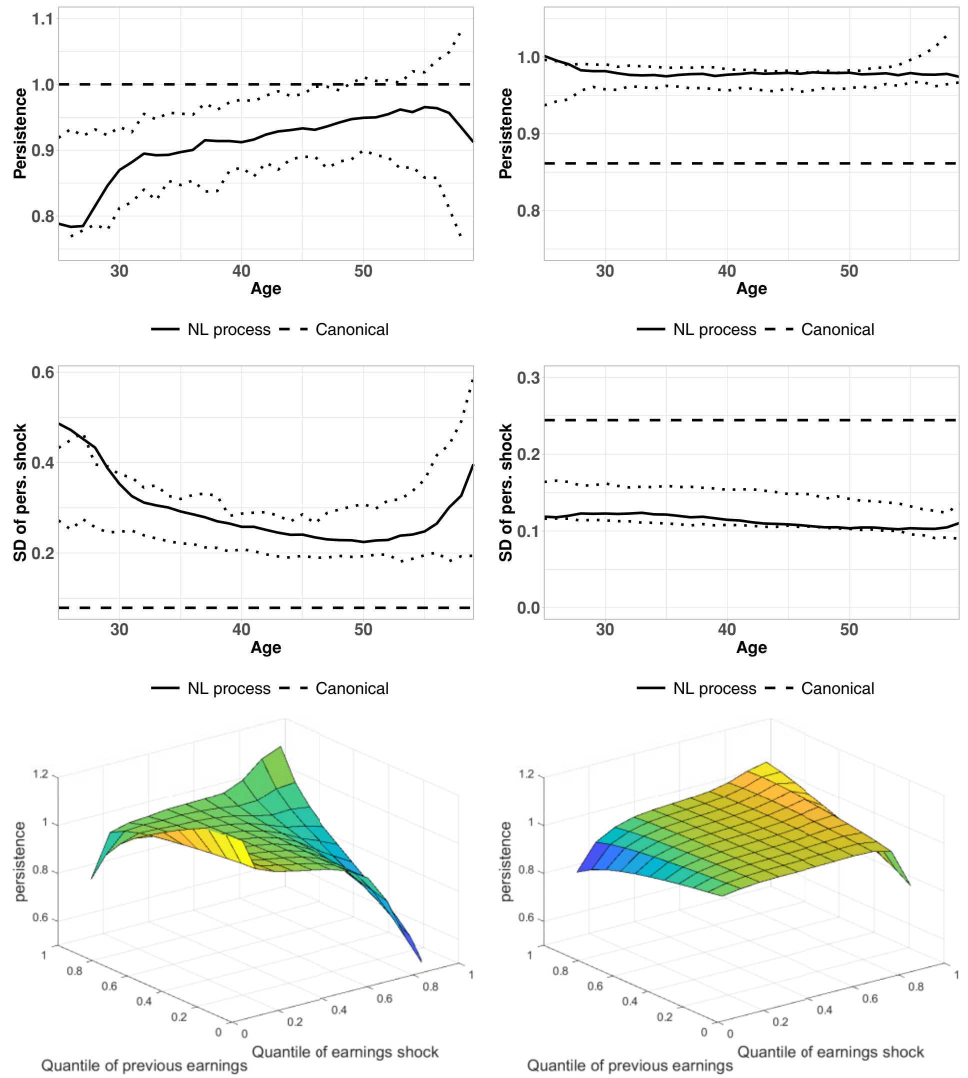

FIGURE 4

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

TABLE 2

Estimates for the canonical processes under alternative imputation procedures for potential wages

Group	σ_{ϵ}^2	$\sigma_{z_1}^2$	σ_{ν}^2	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

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 0–4, $n_{04}(p,t,n)$, and 5–11, $n_{511}(p,t,n)$, 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 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 labour earnings and τ, and s and ζ are parameters that we estimate.

The government provides benefits that depend on household labour income. We consider two alternative structures for the benefit system. The first structure includes two types of benefits: an income floor or IS which is not conditional on working, and an IW benefit which is conditional on a minimum working-hour requirement. This is the structure that was in place before the UC U.K. reform and over which we compute our optimal benefit reform.[6] The second benefit structure features no distinction between IW and unconditional benefits, as is the case after the UC reform (introduced in the U.K. 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 labour 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 labour income of head and spouse (if present), and $\tilde{y}(y^h)$ is household after-tax income.

The post-2016 UC system differs from the previous benefit structure in place and from our baseline calibration in a number of respects. First, there is no IW 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

6. This dual structure is typical of many benefit systems, including the U.S. one.

is, households with assets in excess of $\bar{a}$ do not receive any benefits. As a result, the flow of disposable income under 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 UC, 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:

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

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

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

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

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

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

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

Retirement. Retirees do not 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

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

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

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

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 U.K. 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 labour 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 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 U.K. (even for married couples), we separately apply the tax schedule $T(y)$ to the earnings of husbands and wives. Our estimated 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.

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

TABLE 3

Parameterization of the benefit functions for benchmark IW benefits, benchmark IS, and UC, 2015 pounds

Description	Parameter	Benchmark		UC
		IW	IS	
Intercept	ϕ_0^X	1,960	4,574	4,035
Marriage benefit	ϕ_1^X	0	1,366	2,312
Per kid	ϕ_2^X	2,010	907	1,805
Max. kids	km^X	1	–	2
Tapering rate	ω	1.11		0.63
Earnings disregard	y^{DR}	–	–	2,304
Asset test	$\bar{a}$	–	–	16,000

For the IW 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 IS 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 labour 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 U.K. pre-2016 system. These were 0.7 for IW benefits and 0.41 for IS, 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 UC, 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 UC revenue-neutral from the perspective of the government under the NL wage process. Table 3 reports the values after the scaling. The £2,304 earnings disregard only applies to families with at least one child. All amounts expressed in pounds correspond to 2015 prices.

Retirement. We replicate the U.K. (New) State Pension System. All retired workers receive a maximum amount of £156 per week (in 2016), which corresponds to about 28% 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 U.K. 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, 2006).

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.

TABLE 4
Internally calibrated parameters

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

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 NL 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 ψ_0^{lh}, ψ_1^{ih}, ψ_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 *et al.* (2004) divided by average household income in the same BHPS wave) and the profiles of female labour 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. 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 NL model match participation rates equally well. However, as

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.

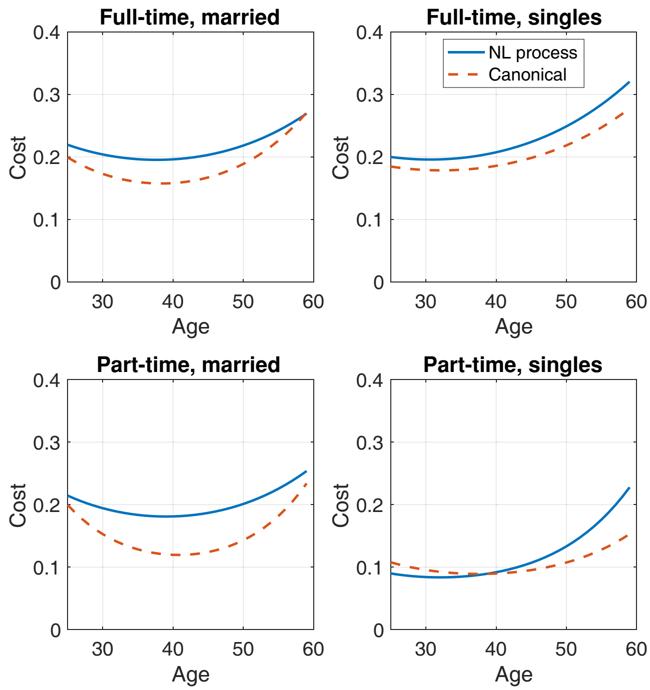

FIGURE 5

Calibrated fixed costs of working over the working period as a fraction of time endowment

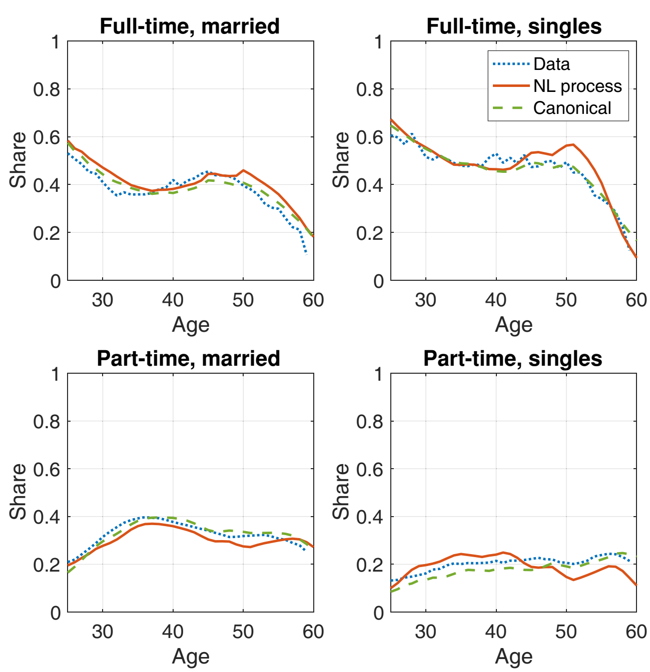

FIGURE 6

Fit of labour market participation for women by marital status and working hours for the NL and canonical processes compared with the data

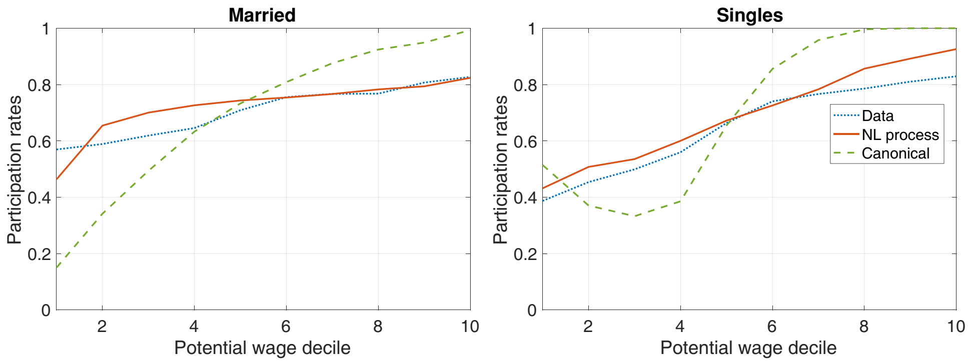

FIGURE 7

Women's labour market participation by marital status and decile of potential wages, compared with the data

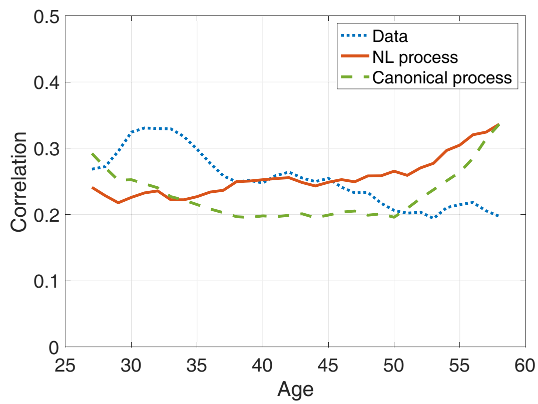

FIGURE 8

Correlation between observed men's earnings and observed wife's wages

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 labour 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 labour 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 labour 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 and 50. However, both imply

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.

TABLE 5

Income floors and IW benefits: benchmark versus optimum under NL and canonical processes. Pound values correspond to 2015

Parameters		Benchmark	Optimum (NL)	Optimum (Ca)
Income floor, level	ϕ_0^{IS}	0.15 (£4,574)	0.143	0.001
In work, level	ϕ_0^{IW}	0.07 (£1,960)	0.010	0.169
Tapering rate	ω	1.11	0.53	0.55

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 NL and canonical processes in terms of the optimal composition of IW and IS 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 IW benefit that were in place before the introduction of UC. More specifically, we optimize over 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 IW 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 economy while the IW 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 53%. 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 labour 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% of average male earnings income have lower benefits as a result while those above 15 and below 50% gain under the new policy. While the optimal benefit system under the NL process implies a reduction in the net return to working, the optimal benefit configuration

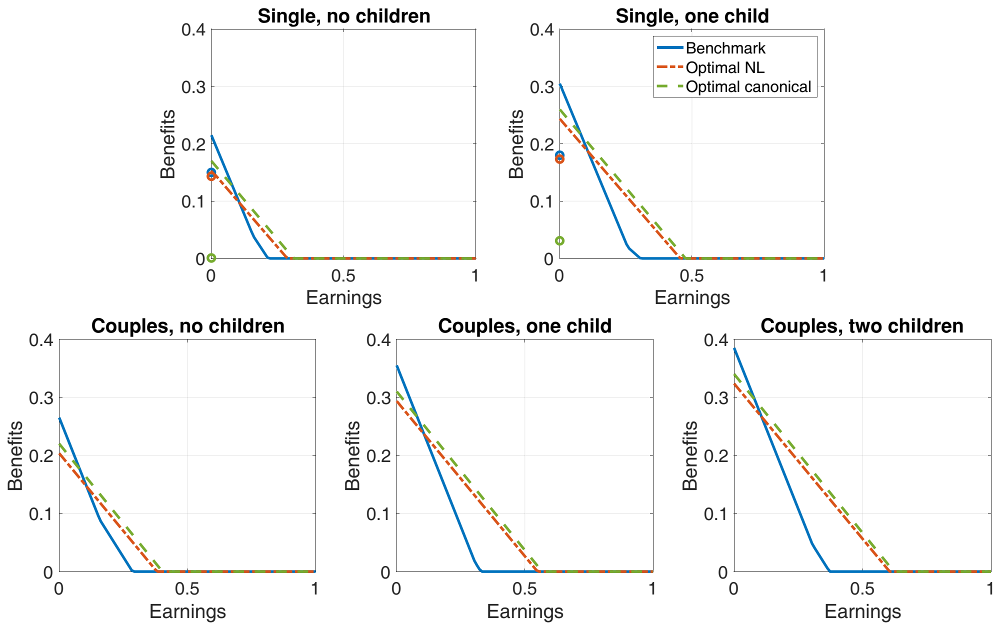

FIGURE 9

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

under the canonical process is dramatically skewed towards IW benefits. In particular, the optimal system implies the practical elimination of benefits for non-working individuals and a more than 2-fold increase in IW benefits. As a result, at the optimum the net return to the first pound of labour income (the difference between the vertical intercept of the straight line and of the corresponding circle in Figure 9) is more than 10 times as large under the canonical than in the NL process. The tapering rate, instead, is very similar for both processes.

Increasing IW benefits and reducing IS have offsetting effects on welfare. On the one hand, welfare increases as a result of improved incentives to participate in the labour 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 IS to encourage labour market participation drastically reduces welfare for women on low wages and with a higher disutility from work for whom the increase in IW 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 labour market participation by single women, and 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

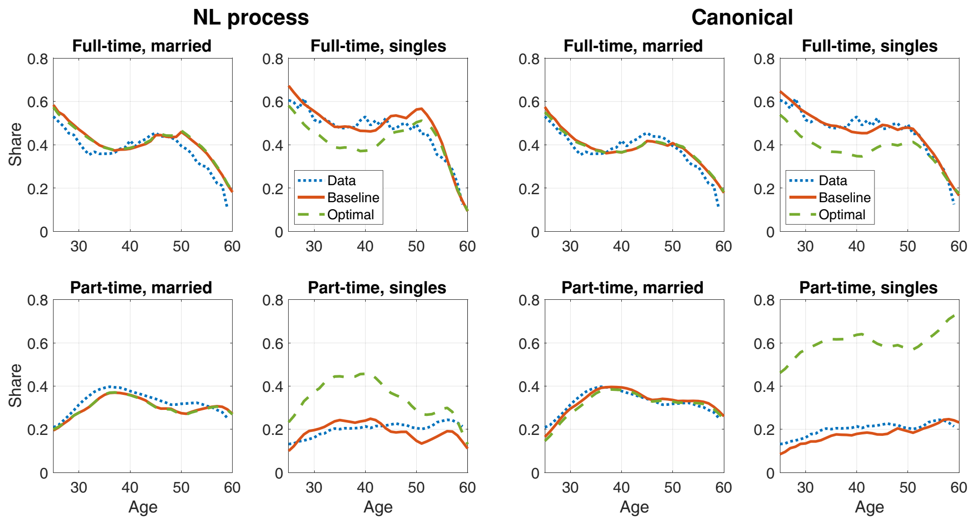

FIGURE 10

Labour force participation for women in the data, baseline model, and under its optimal benefit system

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

Group	NL (opt)	Canonical (opt)
Overall	0.92	1.00
Single men	−0.13	1.19
Single women	1.36	−1.29
Couples	0.95	1.25

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 IS to IW 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 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, 0.92 and 1 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 IW 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 IW 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.

TABLE 7
Income floors and IW benefits: optimum under alternative potential wage imputations

		H-Ben	H	H-Child	FE
Parameters			NL process, optimum		
Income floor, level	ϕ_0^{IS}	0.143	0.145	0.143	0.143
In work, level	ϕ_0^{IW}	0.010	0.007	0.007	0.014
Tapering rate	ω	0.53	0.54	0.52	0.55
			Canonical process, optimum		
Income floor, level	ϕ_0^{IS}	0.001	0.003	0.003	0.000
In work, level	ϕ_0^{IW}	0.169	0.159	0.159	0.166
Tapering rate	ω	0.55	0.52	0.52	0.54

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 UC. 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 modelled 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). UC was piloted in 2013 in a few areas, and then gradually rolled out to all of Great Britain from May 2016 to December 2018.

Figure 11 reports benefits levels as a fraction of pre-tax income in our benchmark economy and under UC. Its main takeaway is that, compared to our benchmark, UC 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 UC benefit reform in both cases.

Figure 12 compares the labour force participation under UC 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.

Table 8 reports the steady-state changes in welfare associated with switching from the benchmark pre-UC benefit configuration to UC for both the canonical 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 UC 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 UC implies a benefit reduction for households without children.

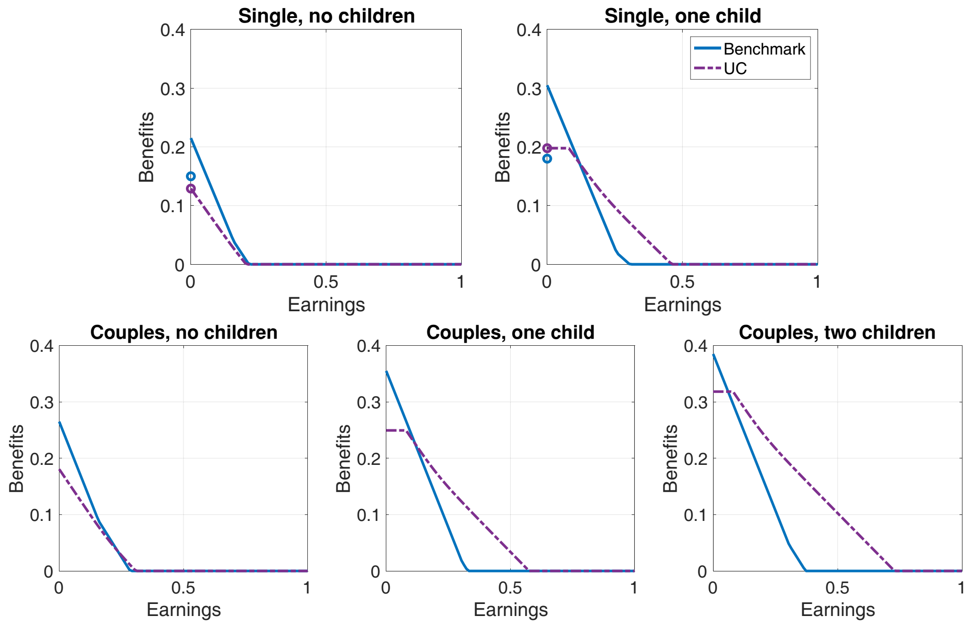

FIGURE 11

Implied total level of benefits, by income levels, comparing benchmark versus UC. 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

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' labour 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 labour income documented for other countries also hold for the U.K. Rather than being constant, the variance and persistence of labour earnings display substantial differences by age and labour income history. These rich dynamics are a feature not only of earnings, but also of wages. Hence, they reflect genuine labour 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 labour 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 labour income shocks is important to capture labour 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 labour 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 U.K. depend on the

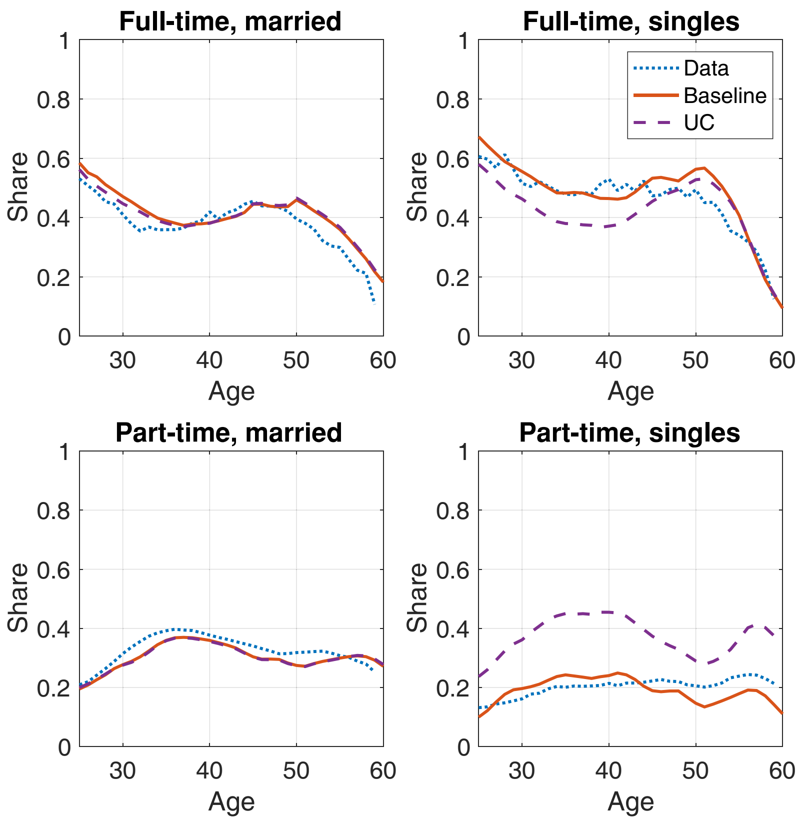

FIGURE 12

Labour force participation for women under NL process: UC versus baseline

TABLE 8

Welfare change for switching to UC, measured as consumption equivalent compensations

Group	NL process	Canonical
Overall	1.22	1.13
Single men	−2.00	−0.90
Single women	2.27	0.41
Couples	1.34	1.37

presence of a spouse and the number of dependent children. Second, that allowing for both single and married households is crucial because labour income pooling within families and the possibility of adjusting the labour 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 labour income processes. Our findings confirm that correctly capturing the dynamics of labour 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—IS and IW benefits—to maximize (utilitarian) welfare in the economy. This reform entails relatively small welfare gains compared to the pre-2016, benchmark U.K. benefit configuration. More importantly, the optimal benefit configuration is very different under the canonical as opposed to the flexible, NL wage process. Under the flexible process, the optimal benefit configuration is similar to the pre-reform one and implies that IS, independent of labour 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 IW benefits and basically no role for IS. 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 labour 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 UC which was introduced in 2016 and completed in 2018. UC 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 UC 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.

Acknowledgments. We are grateful to the editor Kurt Mitman and to three anonymous referees for detailed comments. We also 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.nuffieldfoundation.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.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

Data availability

The data and code underlying this research is available on Zenodo at https://dx.doi.org/10.5281/zenodo.10402343

REFERENCES

ABOWD, J. and CARD, D. (1989), “On the Covariance Structure of Earnings and Hours Changes”, *Econometrica*, **57**, 411–445.

ARELLANO, M., BLUNDELL, R. and BONHOMME, S. (2017), "Earnings and Consumption Dynamics: A Non-Linear Panel Data Framework", *Econometrica*, **85**, 693–734.

ARELLANO, M. and BONHOMME, S. (2017), "Quantile Selection Models with an Application to Understanding Changes in Wage Inequality", *Econometrica*, **85**, 1–28.

BANKS, J., BLUNDELL, R. and SMITH, J. (2004), "Wealth Portfolios in the United Kingdom and the United States", in *Perspectives on the Economics of Aging* (University of Chicago Press) 205–246.

BANKS, J. and SMITH, S. (2006), "Retirement in the UK", *Oxford Review of Economic Policy*, **22**, 40–56.

BEWLEY, T. (1977), "The Permanent Income Hypothesis: A Theoretical Formulation", *Journal of Economic Theory*, **16**, 252–292.

BICK, A. and FUCHS-SCHÜNDELN, N. (2017), "Taxation and Labour Supply of Married Couples across Countries: A Macroeconomic Analysis", *The Review of Economic Studies*, **85**, 1543–1576.

BLUNDELL, R., COSTA DIAS, M., MEGHIR, C., *et al.* (2016), "Female Labor Supply, Human Capital, and Welfare Reform", *Econometrica*, **84**, 1705–1753.

BLUNDELL, R., REED, H. and STOKER, T. M. (2003), "Interpreting Aggregate Wage Growth: The Role of Labor Market Participation", *American Economic Review*, **93**, 1114–1131.

BORELLA, M., DE NARDI, M. and YANG, F. (2023), "Are Marriage-Related Taxes and Social Security Benefits Holding Back Female Labor Supply?", *The Review of Economic Studies*, **90**, 102–131.

CHIAPPORI, P.-A., COSTA-DIAS, M. and MEGHIR, C. (2018), "The Marriage Market, Labor Supply, and Education Choice", *Journal of Political Economy*, **126**, S26–S72.

CUBEDDU, L. and RÍOS-RULL, J.-V. (2003), "Families as Shocks", *Journal of the European Economic Association*, **1**, 671–682.

DE NARDI, M., FELLA, G., KNOEF, M., *et al.* (2021), "Family and Government Insurance: Wage, Earnings, and Income Risks in the Netherlands and the US", *Journal of Public Economics*, **193**, 104327.

DE NARDI, M., FELLA, G. and PAZ-PARDO, G. (2020), "Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare", *Journal of the European Economic Association*, **18**, 890–926.

GOUVEIA, M. and STRAUSS, R. P. (1994), "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis", *National Tax Journal*, **47**, 317–339.

GRONECK, M. and WALLENIUS, J. (2017), "It Sucks to be Single! Marital Status and Redistribution of Social Security", (Working Papaer Series in Economics, 1, SSE).

GUNER, N., KAYGUSUZ, R. and VENTURA, G. (2012a), "Taxation and Household Labour Supply", *The Review of Economic Studies*, **79**, 1113–1149.

GUNER, N., KAYGUSUZ, R. and VENTURA, G. (2012b), "Taxing Women: A Macroeconomic Analysis", *Journal of Monetary Economics*, **59**, 111–128.

GUNER, N., KAYGUSUZ, R. and VENTURA, G. (2020), "Child-Related Transfers, Household Labour Supply, and Welfare", *The Review of Economic Studies*, **87**, 2290–2321.

GUVENEN, F., KARAHAN, F., OZKAN, S., *et al.* (2021), "What Do Data on Millions of U.S. Workers Reveal about Lifecycle Earnings Dynamics?", *Econometrica*, **89**, 2303–2339.

HECKMAN, J. J. (1979), "Sample Selection Bias as a Specification Error", *Econometrica: Journal of the Econometric Society***47**, 153–161.

KEANE, M. P. and WOLPIN, K. 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**, 851–892.

MACURDY, T. 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, C. and PHILLIPS, D. (2010), "Labour Supply and Taxes", in Mirrlees, J. A. et al. (eds) *Dimensions of Tax Design: The Mirrlees Review* (Chapter 3) (Oxford University Press) 202–274.

NISHIYAMA, S. (2019), "The Joint Labor Supply Decision of Married Couples and the US Social Security Pension System", *Review of Economic Dynamics*, **31**, 277–304.

OZKAN, S., STORESLETTEN, K., HOLTER, H., *et al.* (2017), "Dissecting Idiosyncratic Income Risk", (2017 meeting papers, Society for Economic Dynamics).

STORESLETTEN, K., TELMER, C. I. and YARON, A. (2004), "Consumption and Risk Sharing Over the Life Cycle", *Journal of Monetary Economics*, **51**, 609–633.