Wage Risk and Government and Spousal Insurance

Mariacristina De Nardi, Giulio Fella, and Gonzalo Paz-Pardo*

December 14, 2020

Abstract

The extent to which households can self-insure and the government can help them to do so depends on the wage risk that they face and their family structure. We study wage risk in the UK and show that the persistence and riskiness of wages depends on one's age and position in the wage distribution. We also calibrate a model of couples and singles with two alternative processes for wages: a canonical one and a flexible one that allows for the much richer dynamics that we document in the data. We use our model to show that allowing for rich wage dynamics is important to properly evaluate the effects of benefit reform: relative to the richer process, the canonical process underestimates wage persistence for women and generates a more important role for in-work benefits relative to income support. The optimal benefit configuration under the richer wage process, instead, is similar to that in place in the benchmark UK economy before the Universal Credit reform. The Universal Credit reform generates additional welfare gains by introducing an income disregard for families with children. While families with children are better off, households without children, and particularly single women, are worse off.

*Mariacristina De Nardi: University of Minnesota, Federal Reserve Bank of Minneapolis, CEPR, and NBER. Giulio Fella: Queen Mary University of London, CEPR, CFM, and IFS. Gonzalo Paz-Pardo: European Central Bank. 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 the Foundation, or the NBER, CEPR, any agency of the federal government, the Federal Reserve Bank of Minneapolis, the European Central Bank, or the IFS.

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 wage fluctuations. In addition, couples benefit economies of scale in the 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 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 display dynamics that are substantially richer than those implied by the canonical linear model previously 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 by the position of a worker 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 behaviors.[1]

We then develop a dynamic structural life-cycle model of married and single men and women under these two alternative wage processes and use our calibrated model to evaluate the optimal provision of two important types of government transfers, an income floor and in-work benefits. We allow these benefits to differ in their levels and phase-out rates as a function of income. We calibrate our model to match 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 both wage processes fit key aspects of the observed data, their optimal policy implications are starkly different. In particular, the optimal benefit configuration under the richer wage process is very similar to the one that was in place during the period preceding the Universal Credit reform. 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 a much faster (from 40 to 100 per cent) phase-out rate of benefits. The intuition for the difference in optimal benefits 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 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), who 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, generalised asset

[1]Guvenen, Karahan, Ozkan and Song (2016) 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 (2019) for household disposable earnings in the US and Ozkan, Storesletten, Holter and Halvorsen (2017) for household earnings in Norway both before and after taxes. De Nardi, Fella, Knoef, Paz-Pardo and Van Ooijen (2021) study the relative contributions of wages and hours to male earnings dynamics.

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. However, this average improvement masks very heterogeneous effects. The main beneficiaries are households with children, who constitute the majority of households and benefit from the earnings disregard. Singles without children, particularly women, lose out.

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 the response of female labor supply to understand these effects. 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, 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 (2016) study the effects of marriage-related taxes and Social Security rules for different cohorts of women whose labor supply behavior has been changing. Importantly, 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 (and household's savings). 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 and sampled initially 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, like 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 himself saw the pay slip. An important advantage of the BHPS is that, in addition to income data, it includes a wide variety of information (such as off-sample labor market histories). Furthermore, it collects information on all household members, and is thus suited forthe 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.

To ease potential concerns about measurement error in the BHPS, we compare our findings with the implications of the New Earnings Survey Panel Dataset (NESPD), an administrative data set with individual data from the UK Social Security. We provide more information about both data sources and their differences in Appendix A.1. Appendix A.2 details our requirements for sample construction, the most important of which is dropping observations below 5% median earnings (roughly £1,300 a year), which is usual practice in the literature. 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 57,659 usable observations (pairs of earnings in t and $t+1$) for men and 63,014 for women.

We document the properties of male, pre-tax earnings in the UK by using a set of moments that has become rather standard in the literature. Our first finding is that, for the case of the UK, the NESPD generates very similar implications in both quantitative and qualitative terms to the BHPS. Thus, we report the comparison in Appendix A.3 and in what follows we focus on the BHPS.

The top left panel of Figure 1 plots the standard deviation of earnings changes against the percentile of the last period's earnings. In both data sets, the standard deviation follows a U-shaped pattern which is inconsistent with the assumption of linearity which underpins the canonical model.

The top right and bottom left panels plot the skewness and kurtosis of earnings changes in the two datasets, respectively. Skewness is positive for low realization of previous earnings and falls as one moves to the right in the distribution of previous earnings. Kurtosis is somewhat higher than its value of 3 for the normal distribution, but, overall, UK male earnings display substantially smaller deviations 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 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, with a persistence of about 0.55.

Turning to female wages, the first three panels of Figure 2 plot the variance, skewness and kurtosis of female wage changes 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 a previous wage realization, though it displays much less variability with respect to age than in the case of male wages.

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 processes for earnings and wages

Our structural model of household behavior requires that we estimate the stochastic processes that households face for earnings (for men) and wages (for women). 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 their age. We assume that the logarithm of the potential wage $\tilde{w}_{it}^{g}$, net of time fixed-effects, can be decomposed into a deterministic (η_t^{gp}) and a stochastic (y_{it}^{g}) component

$$\log w_{it}^{gp} = \eta_t^{gp} + y_{it}^{g}. \tag{1}$$

For men, the potential wage in equation (1) above is actual measured earnings because we abstract from the labor supply margin and restrict attention to individuals with

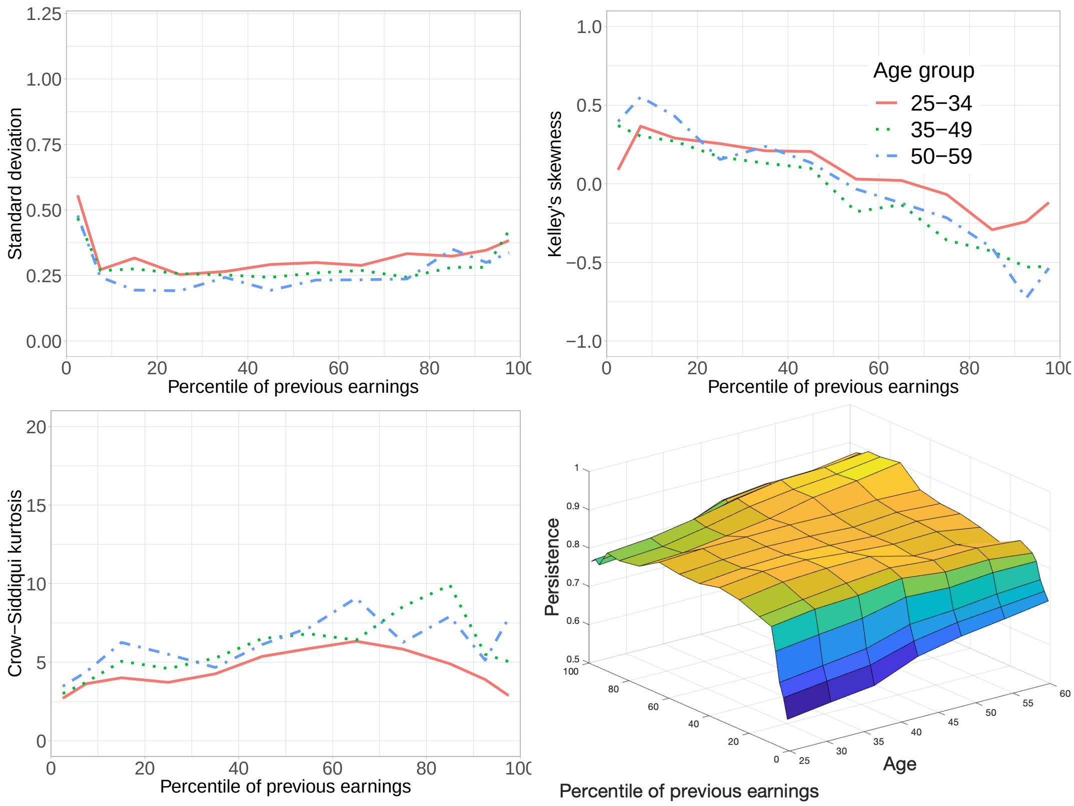

Figure 1: Moments of male earnings changes in BHPS data

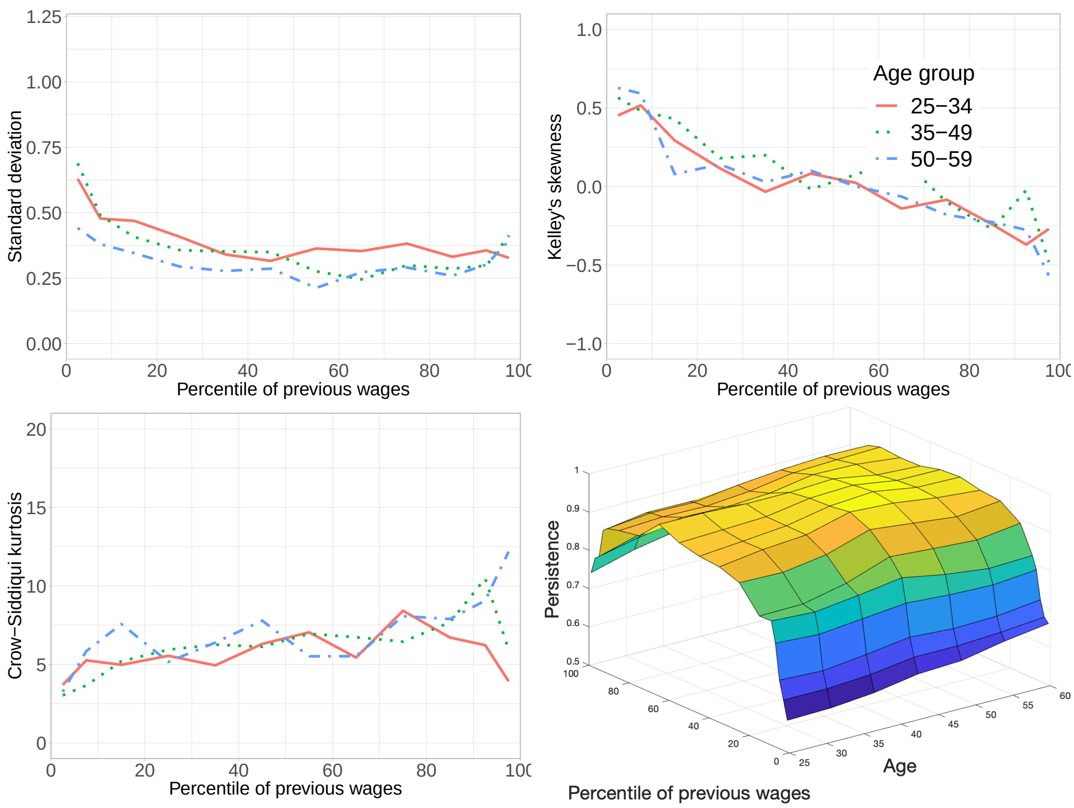

Figure 2: Moments of female wage changes in BHPS data

substantial attachment to the labor market (which make up for the vast majority of the data). For women we impute potential wages for the years when they are not working. See Appendix A.4 for details.

We estimate two alternative processes for the stochastic wage component[2] y_{it} from equation (1). Both assume that it can be decomposed as a *persistent* shock that follows a first-order Markov process, z_{it}, and a *transitory* shock that is independently distributed over time, ϵ_{it}

$$y_{it} = z_{it} + \epsilon_{it.} \tag{2}$$

The *canonical* (linear) model assumes that the persistent component of the shock evolves as

$$\begin{aligned} &z_{i,t} = \rho z_{i,t-1} + \nu_{it}, \\ &z_{i1} \overset{id}{\sim} N(0, \sigma_{z_1}), \quad \nu_{it} \overset{iid}{\sim} N(0, \sigma_\nu), \quad \epsilon_{it} \overset{iid}{\sim} N(0, \sigma_\epsilon). \end{aligned} \tag{3}$$

Our flexible, or nonlinear (NL), process comes from Arellano et al. (2017), which we will denote as ABB from now on: let $Q_z(q|\cdot)$, the conditional quantile function for the variable $z = z, \epsilon$, denote the qth conditional quantile of z. The process generalizes (3) to

$$\begin{aligned} &z_{i,t} = Q_z(v_{it}|z_{i,t-1}, t) \\ &z_{i1} = Q_{z_1}(u_{it}), \quad \epsilon_{it} = Q_\epsilon(e_{it}). \end{aligned} \tag{4}$$

Comparing equations (3) and (4) makes clear that the canonical process imposes constant persistence (linearity) and age-independent conditional distribution. The standard assumption of normality implies a further restriction on the quantile function Q_z for the innovation in equation (3). As we have discussed in Section 2 all these additional assumptions are inconsistent with earnings and wage data in the BHPS and NESPD. Instead, the methodology proposed by ABB does not impose any strong functional form assumption on the quantile functions $Q_z(q|\cdot)$. In particular, the first line of (4) does not restrict the dependence of z_{it} on $z_{i,t-1}$ to be linear.

We take out time effects before estimating our processes. We estimate the canonical earnings process following the procedure described in Storesletten, Telmer and Yaron

[2]We omit the gender superscript in what follows to streamline notation.

(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 represents the estimated parameters for male earnings and women wages for the canonical process. To estimate the flexible, non-linear process we follow Arellano et al. (2017) and De Nardi et al. (2019). Appendix B.1 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.

Group	σ_ϵ^2	$\sigma_{z_1}^2$	σ_ν^2	ρ
Men earnings	0.11664	0.1971	0.0032	0.9999
Women's wages	0.01323	0.1354	0.05128	0.886

Table 1: Estimates for the canonical processes

Figure 3 shows that both processes fit the profile of variances of log earnings for men and log wages for women over the life cycle in the BHPS data. The canonical process does so by construction, while the NL process achieves this result by matching the conditional distribution of y_{t+1} given y_t at every age and thus more realistically capturing the dynamics of earnings. As a result, the two processes have economically meaningful differences that are driven by the restrictive assumptions of the canonical process and which can affect policy evaluation in a structural model. In the case of male earnings, the canonical process generates an increase in variance later in life by assuming that earnings have an (almost) unit root; the NL process, instead, captures this increase through the observed rise in the persistence and variance of earnings later in life (Figure 4, left panels). In the case of women's wages, the hump-shape in the variance of wages is driven by the combination of high persistence and low and decreasing variance of shocks, features of the data that the NL process captures (Figure 4, right panels). The 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 assuming that persistence is low (which helps to generate a relatively flat profile) and that the variance of shocks is high. 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 also, as a result of its restrictive assumptions, generates implications in terms of persistences and variances over the life cycle that are at odds with the data.

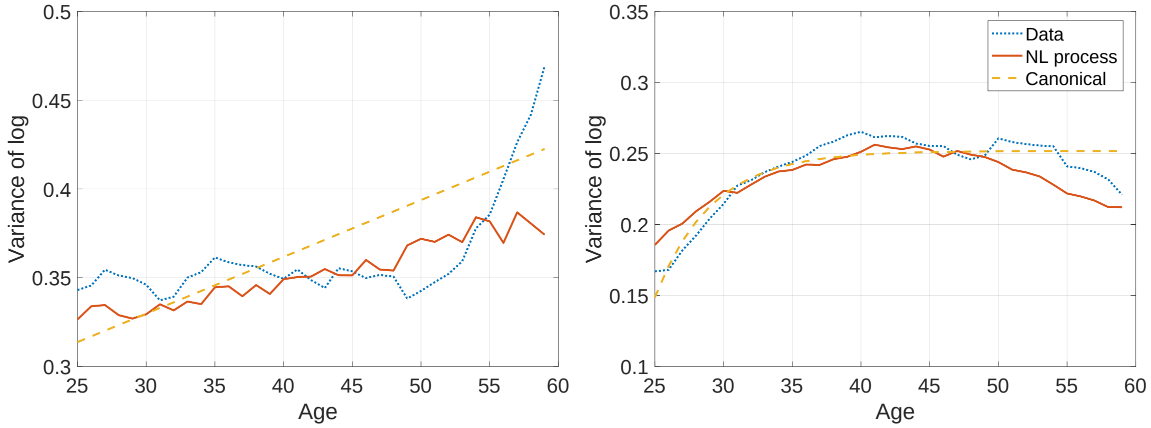

Figure 3: Variance of log earnings for men (left) and log wages for women (right).

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 insurances against income risk. It is this question that we address in the second part of the paper.

3 Our model

We develop a partial-equilibrium, life-cycle, dynamic, incomplete-markets model in the tradition of Bewley (1977). Time is discrete. Individuals start their economic life at age 25, with no assets and a given gender, marital status, initial number of children, and initial wage shock. Men face earnings shocks and women face wage shocks. There are two alternative processes describing the dynamics of earnings and wage shocks, which we have described in the previous section.

Children are born stochastically to households except in single-male households. The probability that children arrive and leave a household depends on their mother's age, marital status, and the number of children already in the household. Children increase

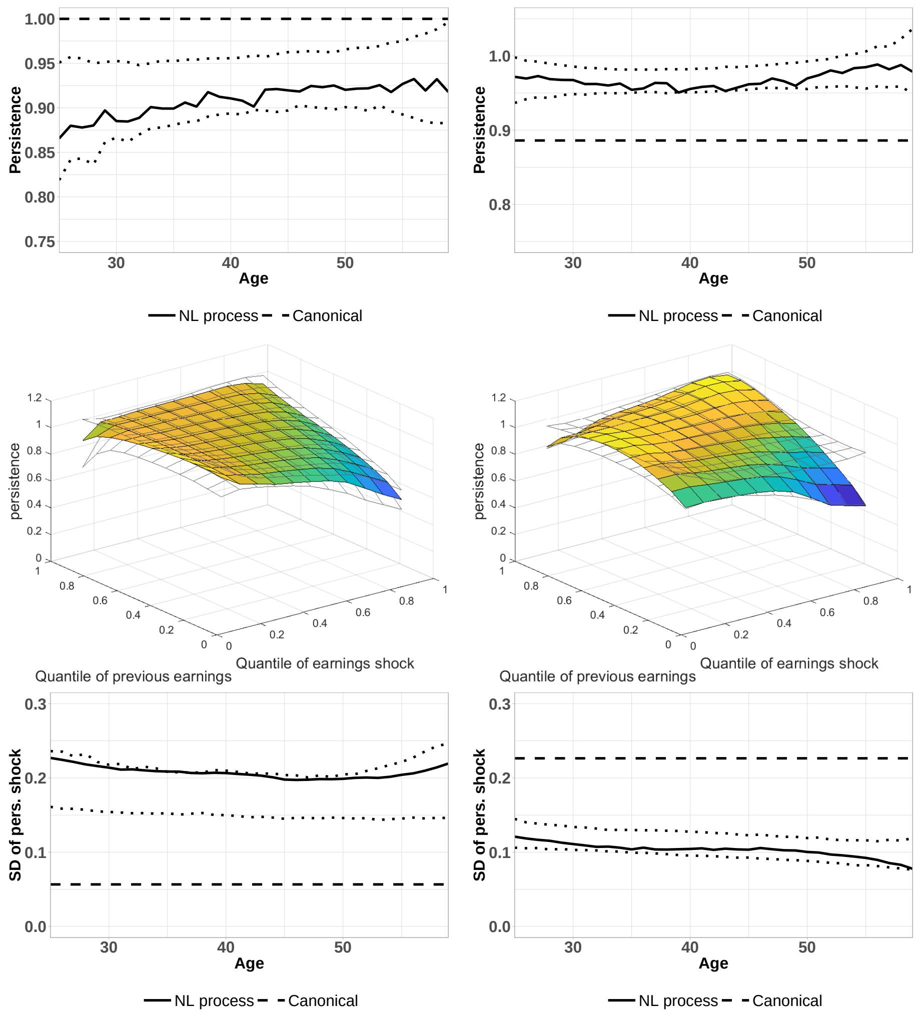

Figure 4: Persistence by age (top), by earnings and quantile of the shock (middle), and standard deviation of shocks (bottom), for male earnings (left) and women's wages (right), NL process vs canonical process, persistent component. Dotted lines and transparent surfaces represent 95% confidence intervals computed by bootstrapping

household consumption needs, entail child care costs if their mother works, and matter for benefit eligibility.

Retired people face mortality risk that depends on gender and age. Age 95 is the maximum possible age one can reach. For simplicity, we assume that people in a given couple have the same age, retire exogenously at age 60, and that marital status is fixed throughout one's lifetime, except for loss of a spouse due to death shocks.

There are no annuity markets to insure against mortality risk. The term t denotes age, $g = f, m$ denotes gender, and $p = s, c$ indicates marital status (single or couple).

Each period, households choose how much to consume and save in a risk-free asset subject to a borrowing limit. Available time 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 getting the children ready for daycare, commuting time, and time spent getting ready for work.[3]

3.1 Preferences and wages

Preferences are time-separable and β is the household's discount factor. The utility function for each person is given by

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

where c denotes total household consumption, l is leisure and μ denotes the equivalence scale, which depends on marital status and number of children. Couples maximize the sum of their utilities in a unitary fashion

$$U(c, l^f) = \frac{((c/\mu)\bar{l})^{1-\omega}}{1-\omega} + \frac{((c/\mu)l^f)^{1-\omega}}{1-\omega}.$$

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 age t, marital status, labor supply l as well as permanent, unobserved, individual heterogeneity $\theta = \{\theta_1, \theta_2\}$. $\mathbf{I}_{h>0}$ is a 0-1 indicator function equal to 1 when hours worked are positive:

$$\Psi^p(h_t, t; \theta) = \mathbf{I}_{h>0}\left\{\theta + \frac{\exp(\psi_0^{l,h} + \psi_1^{l,h}t + \psi_2^{l,h}t^2)}{1 + \exp(\psi_0^{l,h} + \psi_1^{l,h}t + \psi_2^{l,h}t^2)}\right\}. \quad (5)$$

Leisure for women is given by

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

The wage process for an individual of gender g and marital status p follows Equations 1 and 2, without the purely transitory shock ϵ_{it}, which we assume to also reflect measurement error. The first-order Markov stochastic component follows either the canonical

[3] As noted in French (2005), the fixed cost of working implies that, consistently with the observed data, most people will not choose to work just a few hours.

or the NL process. To reflect assortative mating, innovations to the stochastic wage components for people in couples are correlated at age 25.

3.2 Child care costs

The function $CC_t(p, h_t^f, n)$ captures child care costs for a mother of age t, marital status p, working h_t^f hours when n children are living in the household. To take into account the fact that children older than 5 are in school but require child care outside of school hours at least until age 11 and that children younger than age 5 are not yet in school, we specify the following child care cost function

$$CC_t(p, h_t^f, n) = [n_{04}(p, t, n)h_t^f + n_{511}(p, t, n) \max(h_t^f - sc_h, 0)] \times f \tag{6}$$

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 Gouveia and Strauss's (1994) tax schedule

$$\frac{T(y)}{y} = \tau - \tau(sy^\rho + 1)^{\frac{-1}{\rho}}, \tag{7}$$

where $y = wh$ is taxable individual labor earnings and τ, s and ρ are parameters.

The government provides benefits that depend on household labor income. In our benchmark economy, they are composed of an income floor or income support (IS), which is not conditional on working, and in-work benefits (IW). Both are unified in the Universal Credit (UC) system.

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:

$$\bar{Y}^X(p, n) = \phi_0^X + \phi_1^X \min\{n, km^X\} + \phi_2^X \mathbb{I}(p = c) \tag{8}$$

All three benefits are tapered away as labor income increases according to a proportionality factor ω^X. Before the Universal Credit reform, disposable income after taxes and benefits is given by

$$M(y^h) = \tilde{y}(y^h) + \max\{0, \bar{Y^{IS}}(p,n) - \omega^{IS} y^h\} + \max\{0, \bar{Y}^{IW}(p,n) - \omega^{IW} y^h\}\mathbb{I}(h_t > 0), \quad (9)$$

where y^h represents pre-tax household labor income and $\tilde{y}(y^h)$ represents post-tax household labor income. With the Universal Credit (UC) system, benefits are means-tested (households with more than $\bar{a}$ assets do not receive any benefits), there is an initial earnings disregard $y^{DR}(n)$ for families with children, and tapering is based on post-tax income. Thus, the flow of disposable income under this system M^{UC} is:

$$M^{UC}(y^h) = \tilde{y}(y^h) + \max\{0, \bar{Y}^{UC}(p,n) - \omega^{UC}(\max\{\tilde{y}(y^h) - y^{DR}(n), 0\}\}\mathbb{I}(a_t < \bar{a}) \quad (10)$$

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

3.4 Recursive representation

During the working period. Let $W_t^j(\cdot)$ denote the value function for a household of working age t, with $j = f, m$ for respectively single woman and man, and $j = c$ for couples. The state variables for a single woman during this stage 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, and her recursive problem is

$$W_t^f(a_t, z_t^f, n_t, \theta) = \max_{c_t, a_{t+1}, h_t} u(c_t, 1 - h_t - \Psi^1(h_t, t; \theta)) + \beta E_t W_{t+1}^f(a_{t+1}, z_{t+1}^f, n_{t+1}, \theta) \quad (11)$$
$$\text{s.t. } a_{t+1} = (1 + r)a_t + M(h_t w_t^f) - CC_t(f, h_t, n_t) - c_t, \quad a_{t+1} \geq 0,$$

where the expectation is taken with respect to the conditional distributions for wages and the number of children.

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

$h_t = \bar{h}$, has no children and θ is normalized to zero:

$$W_t^m(a_t, z_t^m) = \max_{c_t, a_{t+1}} u(c_t, 1 - \bar{h}) + \beta E_t W_{t+1}^m(a_{t+1}, z_{t+1}^m), \tag{12}$$
$$\text{s.t. } a_{t+1} = (1+r)a_t + M(\bar{h}w_t^m) - c_t, \quad a_{t+1} \geq 0.$$

Working couples choose consumption and the labor supply of the wife to solve

$$W_t^c(a_t, z_t^f, z_t^m, n_t, \theta) = \max_{c_t, a_{t+1}, h_t^f} U(c_t, 1 - h_t^f - \Psi^2(h_t^f, t; \theta), \bar{l}) + \beta E_t W_{t+1}^p(a_{t+1}, z_{t+1}^f, z_{t+1}^m, n_{t+1}, \theta) \tag{13}$$
$$\text{s.t. } a_{t+1} = (1+r)a_t + M(\bar{h}w_t^m + h_t^f w_t^f) - CC_t(p, h_t^f, n_t) - c_t, \quad a_{t+1} \geq 0.$$

During retirement. People younger than age 95 die with positive probability that depends on both age and gender. We assume that children leave the household before their parents retire. Singles—$j = f, m$—solve the recursive problem

$$R_t^j(a_t) = \max_{c_t, a_{t+1}} u(c_t, 1) + \beta s_t^g R_{t+1}^j(a_{t+1}) \tag{14}$$
$$\text{s.t. } c_t + a_{t+1} = (1+r)a_t + Y_r - T(Y_r), \quad a_{t+1} \geq 0.$$

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 from that of the other one. Therefore the recursive problem of a retired couple can be written as

$$R_t^c(a_t) = \max_{c_t, a_{t+1}} U(c_t, 1, 1) + \beta \Big[s_t^f s_t^m R_{t+1}^c(a_{t+1}) + s_t^f(1 - s_t^m) R_{t+1}^f(a_{t+1}) + s_t^m(1 - s_t^f) R_{t+1}^m(a_{t+1}) \Big] \tag{15}$$
$$\text{s.t. } c_t + a_{t+1} = (1+r)a_t + 2Y_r - T(2Y_r), \quad a_{t+1} \geq 0.$$

4 Calibration

4.1 Externally calibrated parameters

Demographics. We use BHPS data to obtain the proportions of households by gender, marital status, and number of children, and the to estimate 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 equivalise consumption using an OECD-modified equivalence scale μ_t, where the first adult counts as 1, the second as 0.5 and all children as 0.3. Survival probabilities s_t^m, s_t^f are obtained from the UK life tables in the Human Mortality Database for the period 1980-2010. We use BHPS data to compute $n_{04}(p, t, n)$ and $n_{511}(p, t, n)$. We plot all of these variables in Appendix C.

Preferences and interest rate. We set the felicity function parameter to $\omega = 2.5$ and the after-tax interest rate $r = 0.02$.

Earnings and wages. We compute the deterministic profile for male earnings and female wages η_t^{gp} and the stochastic process for the persistent components of the canonical and NL process (z_t^f and z_t^m) using the BHPS and Understanding Society data (See Appendix C for details). For tractability, we discard the transitory components that we estimate, which also includes measurement error.[4] The estimated persistent component is discretized following the procedure in De Nardi et al. (2019).

To take into account that the labor markets of the two partners in a couple might be correlated, we use the correlation at age 25 between husband's earnings and wife's wages at age 25.

Taxes and government expenditure We estimate the tax function $T(y)$ in equation (7) by using BHPS data on pre-tax and net 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 $\rho = 5.38$

[4]In the case of women, the existence of an active labor supply margin would have required the transitory shock to enter the state space. Given that transitory shocks are typically well insured in the class of models considered omitting them should not affect our findings significantly.

Parameter	IW	IS	UC
ϕ_0	1960	4574	3856
ϕ_1	0	1366	2210
ϕ_2	2010	907	1733
ω	0.41	0.7	0.63
km	1	NA	2
y^{DR}	NA	NA	2304
$\bar{a}$	NA	NA	16000

Table 2: Parametrization of benefit functions for benchmark in-work benefits (IW), benchmark income support (IS), and Universal Credit (UC), 2016 pounds.

Benefit system. We use data from benefit programmes and benefit receipts to parametrize the benefit functions 8, 9, and 10. We show the resulting parameters in Table 2.

For in-work benefits in our benchmark benefit system, 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 programme is meant to replicate many benefits available to low-income households. These programs that have differential take-up rates and eligibility criteria which would be very complicated to explicitly model individually. Therefore, we use 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} based on the observed data.

More specifically, we look at average benefit receipts for households whose labor income in a given year is close to zero (below £2,000, although results are robust to changing the threshold to £1,000 or £3,000). This approach allows us to average across various types of benefits and weighting by the cross-sectional distribution of benefit receipts within this subset of the population. However, we cannot use the same approach to directly compute the tapering rate ω^{IS} from benefit data because most benefits have weekly or monthly eligibility criteria, while our data are annual, so the actual relationship between income and benefits received gets attenuated in the data due to time aggregation. For this reason, we estimate ω^{IS} as a weighted average of the statutory tapering rates of the relevant benefits taking into account cross-eligibility criteria and legal thresholds. For this programme, there is no limit on how many children the child component can be claimed for. Appendix D provides a more detailed description of the benefit programmes that we replicate and how we perform these computations.

Finally, we take the parameters for Universal Credit from their statutory values, given that we do not have sufficient years of benefit data to check actual benefit receipts. We scale all the fixed allowances $\phi_0^{UC}, \phi_1^{UC}, \phi_2^{UC}$ proportionally by a factor 0.86 so that the change to Universal Credit is revenue-neutral from the perspective of the government under our structural model and under the NL wage process. Table 2 reports the values after the scaling. The £2304 earnings disregard only applies to families with at least one child in the household.

Remaining government policy parameters. We replicate the UK (New) State Pension System. In 2016, all retired workers get a maximum amount of £156 per week, which amounts to about 28 percent of average male earnings, which is the numeraire in our model. The length of a school week sc_h is assumed to be 20 hours 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 model for male earnings and female wages, fits our target data as well as possible. Thus, the remaining seventeen parameters are calibrated inside the model, given each earnings and wage processes. These parameters 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 full-time/part-time employment status $h = \bar{h}/2, \bar{h}$, and thus a total of twelve parameters), the discount factor β, the hourly child care costs f, the disutility of work for the high-cost-of-work group θ_1, and the proportions of single and married women of each θ type. We normalize 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 145 moments. A wealth/income ratio of 2.9 (thus which corresponds 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/part-time status, for a total of 144 (36 × 4) targets.[5].

[5]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

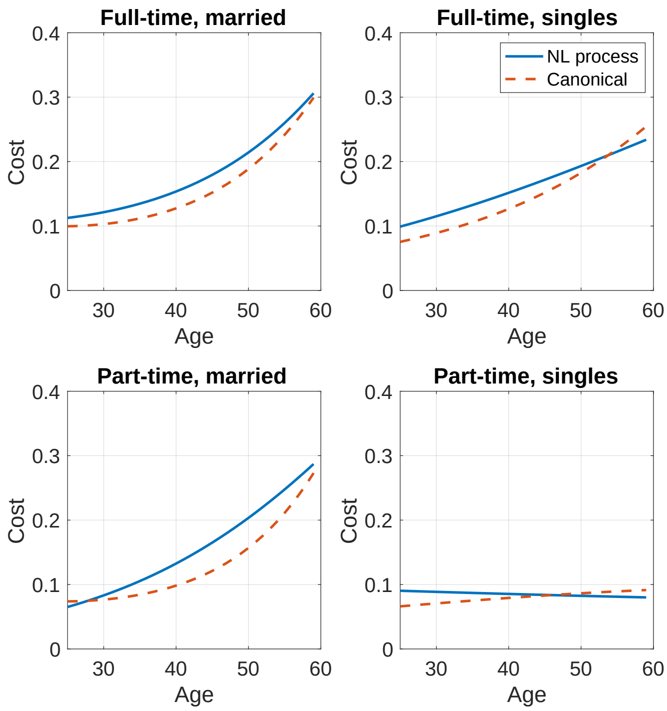

Figure 5: Calibrated fixed costs of working. The units are expressed as fractions of total day

5 Model calibrations and fits

Parameter	NL process	Canonical
Discount rate β	0.98	0.99
Cost of child care f	0.05	0.06
Disutility of work type θ_1	0.53	0.48
Share of θ_1, singles	0.25	0.23
Share of θ_1, couples	0.22	0.25

Table 3: Internally calibrated parameters

Table 3 reports the calibrated preference parameters and child care costs (as a share of the average male wage) for both processes. Figure 5 plots the calibrated fixed time costs (reported as fractions of a day) of part- and full-time work in red for the NL process and in black for the canonical process. Figure 6 compares the targeted participation rates in the data and in the model. In both calibrations the wealth-income ratio equals its target value.

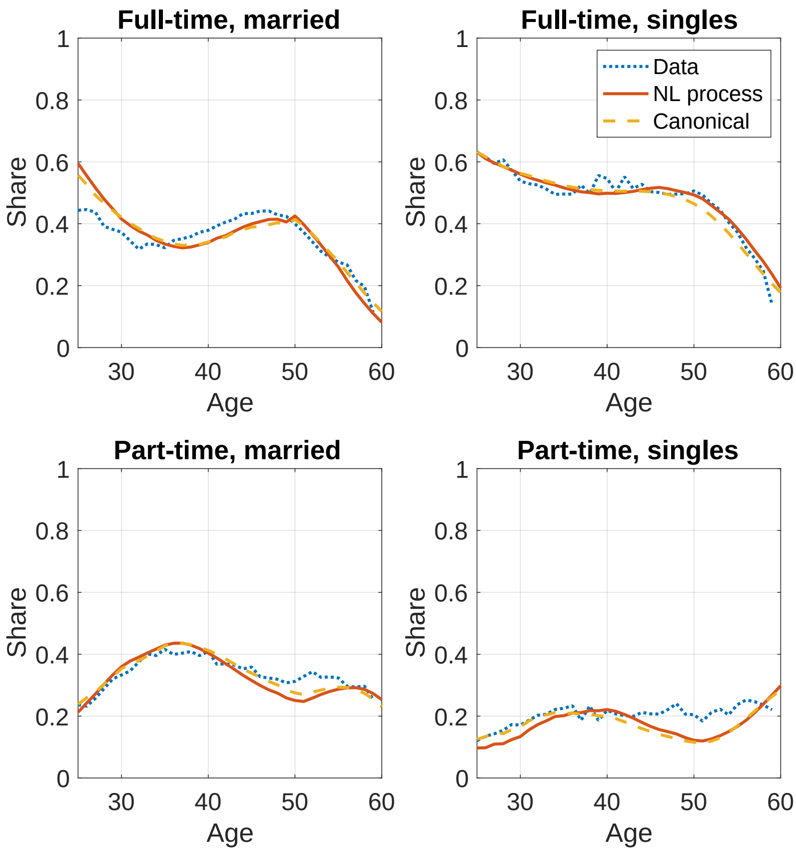

Figure 6: Fit of labor market participation, by marital status and working hours for the NL and canonical processes compared with the actual data

6 Policy evaluation

To evaluate whether benefit reforms have different implications under the canonical and nonlinear processes for earnings and wages, we look at both outcomes and welfare. Our welfare criterion is given by the utilitarian, un-weighted, average of the lifetime utilities of newborns. We report results both behind the full veil of ignorance and after the realization of gender, marital status and number of children.

6.1 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 Universal Credit. That is, we optimize over the intercepts ϕ_0^{IS}, ϕ_0^{IW} and slopes (tapering rates) ω^{IS}, ω^{IW} of functions (8) and (9) to find the system that

Parameters	Benchmark	Optimum (NL)	Optimum (Ca)
Income floor, level	4574	4504	3106
Income floor, tapering	0.70	0.62	0.56
In work, level	1960	2550	7500
In work, tapering	0.41	0.44	1.00

Table 4: Income floors and in-work benefits: benchmark vs optimum under NL and canonical processes.

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. Hence, this change is budget neutral for the government.

Table 4 shows the results of this optimization. Column 2 reports the parameter values for the two benefit functions in the benchmark economy, while columns 3 and 4 report the optimal values under, respectively, the NL and canonical wage process. Under the NL wage process, the optimal benefit system is close to the one in the benchmark economy. The main difference is that it features a 30 per cent higher level of in-work benefits and slightly higher tapering rate for in-work benefits and lower tapering for the income floor. The difference between the optimal and the benchmark benefit policies is possibly best appreciated with the help of Figure 7, which plots the relationship between benefit levels and after-tax labor income for single men, women and couples in the benchmark (blue lines) and under the optimal system under the NL (red lines) and canonical (yellow 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 the model). Under the NL wage process, benefits for working households are marginally higher than in the benchmark and are exhausted at a slightly higher level of disposable income due to the fall in the tapering rate for income-support. Single, working women with one child between the 10th and 30th percentile of labor income are the only group for which the switch to the optimal policy implies a significant increase in total benefits. Under the canonical wage process, instead, the optimal benefit system is substantially different from the one in the benchmark. In particular, the optimal system implies a 30 per cent reduction in benefit levels for non-working individuals, from 15 to 10 per cent of average earnings, accompanied by a more than three-fold increase in the level of in-work benefits. As a result, the net return to the first pound of labor income—the

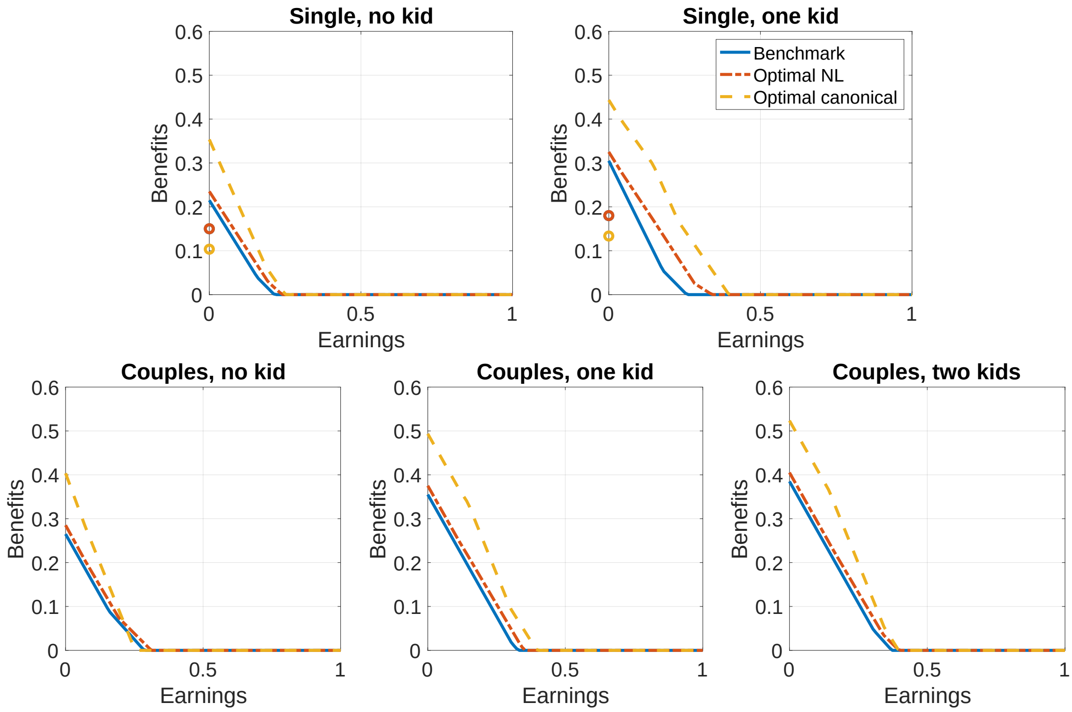

Figure 7: Implied total level of benefits, by income levels, marital status, and number of children. For singles, circles represent benefits for households where everyone is out of work, while lines represent benefits for households in which at least one member works. Earnings and benefits are expressed as the share of average male earnings

difference between the vertical intercept of the straight line and of the corresponding circle in Figure 7—is three times as large than in the benchmark and optimal systems under the NL process. Similarly, the reduction in the tapering rate for the income floor and its increase for in-work benefits are faster compared the NL case. In particular, the tapering rate for in-work benefits increases from 41 to 100 per cent. The net effect is a substantial increase in benefits for working individuals.

Increasing in-work benefits and reducing income support is welfare-improving because it increases incentives to participate in the labor market, which in turn increase tax revenues to be spent in the insurance system, but is welfare-decreasing because it reduces 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 for 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 low-wage women is lower under the canonical process, because it is a more transitory state, against which it is easier to self-insure. On the other hand, the NL process replicates the fact that

Group	Average	0 kids	1 kid	2 kids	3+ kids
		NL process			
Overall	0.15				
Single men	0.17				
Single women	0.42	0.34	0.59	0.55	0.47
Couples	0.06	0.05	0.06	0.07	0.07
		Canonical process			
Overall	0.19				
Single men	0.11				
Single women	0.52	-0.16	2.32	1.48	0.85
Couples	0.10	0.08	0.11	0.12	0.12

Table 5: Welfare change, by gender and marital status, for switch to optimal system.

low-wage status is a relatively persistent state for a group of women and reducing their income support to encourage them to work would drastically reduce their welfare, while a possible increase in-work benefits would not be enough to compensate the welfare cost of foregone leisure. As a result, the optimal welfare system under the more realistic NL process does not introduce these major changes and is much closer to the system that was in place before 2016.

Figures 8 and 9 show how, under either wage process, the optimal policy mix results in higher part-time and lower full-time labor market participation by single women, and a significant increase in participation overall. The rise in overall participation is driven by the increase in the relative return to work, as measured by the difference between the level of in-work benefits (the circles) and of total benefits (the corresponding lines) in Figure 7. This increase is particularly large under the canonical wage process and accounts for the larger rise in participation. The switch from full-time to part-time is due to the fact that the higher withdrawal rate for in-work benefits increases the effective tax rate and reduces the benefits of working full-time rather than part-time. The effect is, again, more pronounced under the canonical process, which features much larger increases in benefit phase-out rate compared to the NL process.

Table 5 reports the welfare change associated with the switch to the optimal benefit system. The welfare change is expressed as the percentage change in consumption (constant across ages and states) that would make a newborn agent in the benchmark economy indifferent to being born in the counterfactual economy. The first column reports the welfare change before the draw of the initial number of children. The "overall" measure in the first row is under the full veil of ignorance, including the realization of

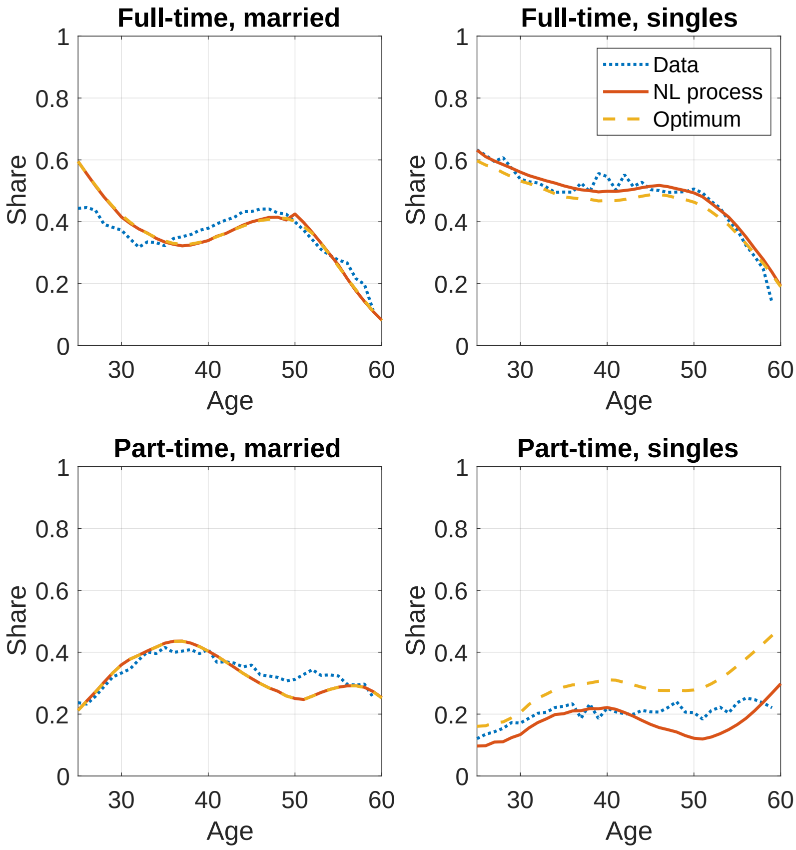

Figure 8: Labor force participation under NL process: optimal benefit system.

the gender and marital status draw.

Despite the substantial differences in the optimal benefit policies across the two wage processes, the overall welfare gains associated with moving from the benchmark to the optimal system are similar under the two wage processes. In both cases, the switch to the optimal benefit system implies an increase in welfare of approximately 0.2 percentage points. Under both wage processes, the main beneficiaries of the reform are single women, whose welfare gain is more that twice the overall one, while couples are hardly affected.

Among single women, though, the two wage processes imply a very different distribution of welfare gains and losses. Under the NL process, the gains are distributed rather evenly, while under the canonical process they accrue to single women with children, at the expense of single women without. The intuition is the following. Under the canonical process the benefit reform involves a substantial shift from income support to in-work benefits. In-work benefits entail an additional £2,010 for households with children, which accounts for the substantially bigger gap between the yellow and the blue line in the top-

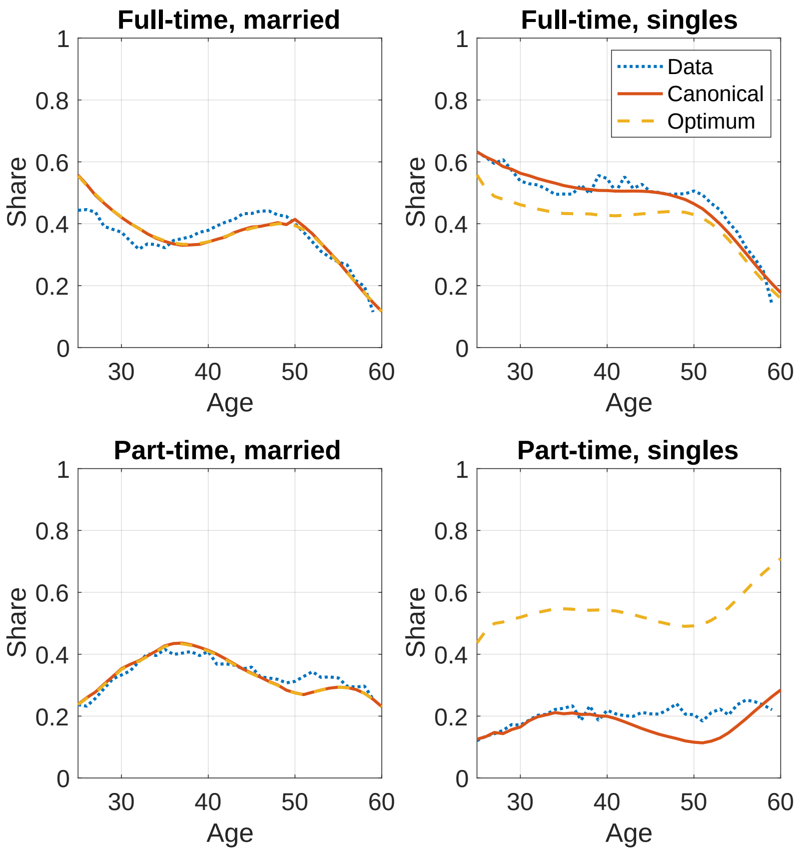

Figure 9: Labor force participation under canonical process: optimal benefit system.

right panel of Figure 7 compared to the top-left one. As women respond to the higher incentives to work, the child-related component of in-work benefits more than compensates, on average, single women with children switch to working under the reform. On the other hand, for single women without children the lower utility of leisure associated with post-reform labor market participation is not compensated by the child-related transfer.

6.2 Universal Credit

The aim of this section is to compare the allocations and welfare implied by the benefit system before and after Universal Credit.[6] 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 Chid Benefits)

[6]It is worth pointing out that, as we have shown in the previous section, the benchmark benefit system that was in place pre-Universal Credit is close to the optimal one that we compute by optimizing over the same set of policy instruments.

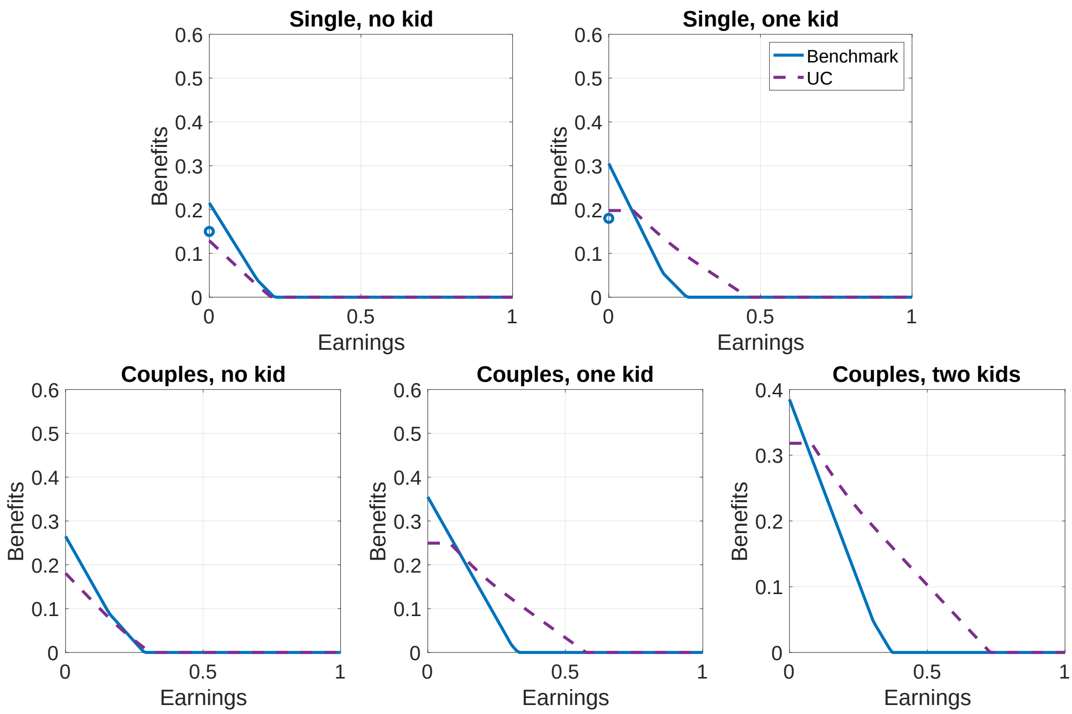

Figure 10: Implied total level of benefits, by income levels, comparing benchmark vs. Universal Credit. For singles, circles represent benefit entitlement for non-working individuals under our benchmark. Earnings and benefits are expressed as the share of average male earnings.

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 £2304 earnings disregard for families with children. Second, benefits are withdrawn as a function of *after-tax* income, rather than pre-tax income in the pre-reform system. Universal Credit was first piloted in 2013 in a few areas and then gradually rolled out to all of Great Britain from May 2016 to December 2018.

Figure 10 reports benefits levels as a fraction of pre-tax income in our benchmark economy and under Universal Credit. The 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 the rest of households with children.

Given that we find that the policy implications of the canonical and NL process can be different, and given that we show that the NL processes provide a much better representation of the dynamics of male earnings and female wages, we evaluate the effects of the introduction of the Universal Credit benefit reform under the NL process and show

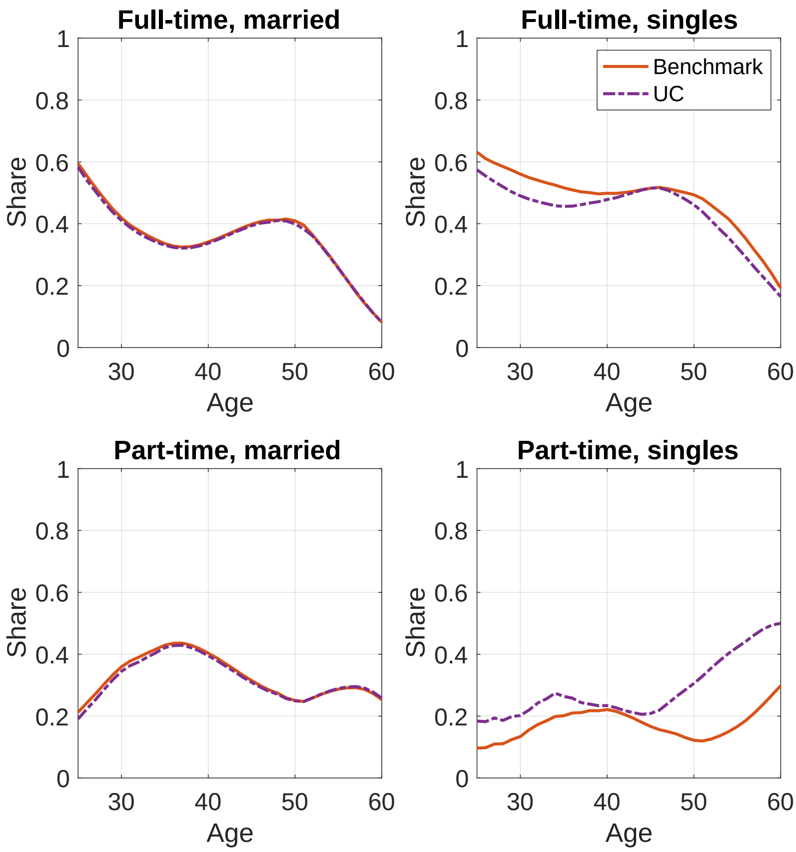

Figure 11: Labor force participation under NL process: Universal Credit vs optimal benefits.

the corresponding outcomes for the canonical process in Appendix E.

Figure 11 compares labor force participation under Universal Credit and in the benchmark, pre-UC, benefit system. Universal Credit substantially reduces benefits for single women with no children. This results in higher (part-time) labor force participation from middle age onward, when children have left the household. As a result, part-time labor force participation at older ages is higher than in the optimal formulation of the previous system. In contrast, participation is lower for single women between age 35 and 45, compared to the optimal system, because a significant number of them have children and receive higher benefits independently from labor force participation under Universal Credit than in our benchmark system.

Table 6 reports the steady-state changes in welfare associated with switching from the benchmark pre-UC benefit configuration to Universal Credit. The switch to Universal Credit entails an increase in average (overall) welfare of 0.45 percentage points. Looking into the distribution of the overall gains, though, reveals that the main beneficiaries are

Group	Average	0 kids	1 kid	2 kids	3+ kids
Overall	0.45				
Single men	-0.07				
Single women	-0.15	-1.65	2.44	3.98	3.44
Couples	0.69	0.44	0.75	0.97	1.02

Table 6: Welfare change for switch to Universal Credit. NL process.

households with children, who benefit from the disregard, the fact that only earnings above the first £2304 result in benefit withdrawal. On the other hand, singles without children are worse off. In particular, while the welfare loss for single men is relatively low, single women with children experience a very sizeable welfare loss of 1.65 percentage points. It may seem surprising that, unlike singles, couples with no children benefit from the reform. In fact, one has to realize that welfare in Table 6 is from the perspective of individuals at age 25. So the difference between couples and single women with no children is that the former have a much higher probability than the latter of having children, and therefore benefit from the earnings disregard, later in life. This drives their higher welfare, in ex ante, expected terms.

It is instructive to compare these welfare changes with those associated with a switch to the optimal pre-UC benefit system in Table 5. Relative to the benchmark, the optimal pre-UC system redistributes more towards singles, particularly single women, while Universal Credit mostly benefits couples. Since couples constitute the majority of the population, Universal Credit entails a two-to-three times larger welfare increase under the utilitarian (i.e. unweighted) social welfare criterion. Contrary to the optimal benefit reform, though, Universal Credit *reduces* overall benefits for singles without children at every level of income (top left panel in Figure 10) relative to the benchmark. This implies a negative income effect which induces higher labor force participation, and lower welfare, for single working women on low income.

The reason why Universal Credit can achieve higher average welfare than the optimal pre-UC benefit system is that possible because the reform in Section 6 is optimal within the class of linear in-work and income-floor benefit functions. The non-linear Universal Credit benefit function is therefore not nested in that class because it also adopts income disregards and asset testing.

7 Conclusion

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

This paper is the first to establish that the rich dynamics of labor income documented for other countries also applies to 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 show that ignoring such richer dynamics when estimating stochastic labor income processes implies biased estimates of important moments on the data. In particular, the canonical model underestimates the persistence of shocks to female wages and overestimates the persistence of shocks to male earnings relative to a richer, flexible earnings process which does not impose the constancy of variance and persistence.

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 dissaving (self-insurance). This is why we investigate how allowing for a richer labor income dynamics affects the evaluation of welfare policies compared to the standard, canonical income process. To do so, we build and estimate a structural life-cycle model with heterogeneity in family structure, which allows us to capture the following important channels. First, 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, 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 important margins of insurance at the household level.

We use our model to evaluate to evaluate alternative benefit reforms under 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 analyse a hypothetical reform that chooses the structure of the two main benefits – income support and in-work benefits – to maximize (utilitarian) welfare in the economy. This reform entails small welfare gains compared to the pre-2016, benchmark UK benefit configuration. Although the welfare results are robust to the way the earnings process is modelled, the optimal benefit configuration is very different under both canonical and flexible and nonlinear risks. Under the flexible earnings process, the optimal benefit configuration is very similar to the pre-reform one. 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 a much faster (from 40 to 100 per cent) withdrawal rate for benefits. 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 through borrowing, the optimal policy under the canonical process is skewed towards providing incentives to work rather than insurance against low labor income realisations.

We also consider a reform that mimics the switch to 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. We find that, thanks to the use of these additional instruments, the move to Universal Credit implies overall welfare gains which are larger than those under the optimal benefit system. However, this average improvement masks heterogeneous effects. The main beneficiaries are households with children, who benefit from the earnings disregard, while singles without children lose out.

References

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, 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.

Borella, Margherita, De Nardi, Mariacristina and Yang, Fang (2016), 'Marriage-related policies in an estimated life-cycle model of households' labor supply and savings for two cohorts', Working paper.

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

De Nardi, Mariacristina, Fella, Giulio and Paz-Pardo, Gonzalo (2019), 'Nonlinear household earnings dynamics, self-insurance, and welfare', *Journal of the European Economic Association* . Forthcoming.

French, Eric (2005), 'The effects of health, wealth, and wages on labour supply and retirement behaviour', *The Review of Economic Studies* **72**(2), 395–427.

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.

Guvenen, Fatih, Karahan, Fatih, Ozkan, Serdar and Song, Jae (2016), What do data on millions of U.S. workers reveal about life-cycle earnings risk? Working paper, University of Minnesota.

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

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.

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.

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 Data sources

A.1.1 NESPD

Individuals whose National Insurance Number (NIN) ends in a certain set of two digits are automatically selected for the sample. The NIN number is randomly issued to all UK residents at age 16, and kept constant throughout the lifetime of an individual. Data is currently available for all years between 1975 and 2015.

Every April, all employers whose employees qualify for the sample receive a form (currently online, although in the early years of the sample it was on paper) where they must provide payroll data about those employees.

This implies that, for individuals included in the survey, the New Earnings Survey Panel Dataset (NESPD) contains complete information on their working life from the first year they started working (or 1975 if later) until retirement age (or 2015 if earlier), for all years in 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 nonresponse 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 does not respond to the survey.

A.1.2 BHPS

Starting in 1991 and continuing until 2010, when it was discontinued to be included within the wider survey Understanding Society, the BHPS sampled initially 5,500 households and 10,300 individuals. All individuals that formed part of the initial sample were followed whenever contact was possible, thus 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. All of these features imply that this survey should remain representative of the UK population even as it evolved over the 1990s and 2000s.

As a household survey, 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, like 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 himself saw the pay slip.

A.2 Sample construction

The **earnings** sample is based on the "annual earnings" variable (NESPD) and on a reconstructed annual earnings measure in the BHPS. The latter is necessary because individuals are not asked about the total amount of earnings in the past year, but about their usual payments derived from all of the jobs they held. The derivation of annual earnings from these payments is straightforward - the only assumption we make is that we only consider jobs that have been held at least at some point in two calendar months.[7]

When reconstructing BHPS earnings, we take into account what the employee reports to be their "usual" payment in case the last payment they received was not the usual one. We drop individuals who say the last payment was not the usual one but did not report how much the usual amount is.

In both cases, we only consider people who have received at least 5% of median earnings (around £1,300 (2015)) in the year up to the moment when they are observed. Therefore, long-term unemployed and people permanently out of the labor force are not present in either of the samples. We drop self-employed people.

The population covered by the BHPS measure of earnings is wider than the population covered by the NESPD survey. 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. On the other hand, the BHPS, being a household survey, can capture those individuals who are non-employed but have worked at some point during the previous year.

The NESPD only considers the highest-paid job for each individual, so we limit our-

[7]Many individuals report not to remember the precise dates they started or finished an employment or unemployment spell.

selves to the main job in the BHPS too. In the NESPD, annual earnings go from 7th April to 6th April (tax year), while in the BHPS they go from 1st September to 31st August (by sample construction). We consider individuals between 25 and 60 years of age. To increase the sample size, we perform a rolling-sample transformation in the BHPS (similar to what De Nardi et al. (2019) do for the PSID case). We deflate earnings and wages with the CPI (2015=100).

In both samples, we reconstruct ages whenever the change of date in the interview may imply that the individual is reported to be the same age in two consecutive years. We only do so when there is enough information to safely assume so (namely, that reported age does not differ by more than 1 with the expected age). In the NESPD, we drop cases where there are two records with the same (identifier, year) pair. We also drop individuals for whom ages evolve unexpectedly, which can reflect, in the case of the NESPD, errors in recoding NINOs (as stated in the documentation for the data).

We also eliminate outliers that most likely reflect recoding errors and missing values. We drop individuals whose total working hours exceed 80 per week and individuals that display negative values in earnings or hourly wages. In the NESPD, we do not consider individuals whose hours worked or weekly pay are missing. In the BHPS, where wages are implied from earnings and hours of work, we eliminate extreme changes ($|\log w_t - \log w_{t-1}| > 2$) that probably represent errors in recording hours of work.

To compute all statistics related to wages, we also drop individuals whose wage is zero (i.e., that were not working in the week of reference). Hourly wages are directly available in the NESPD, while they can be inferred from weekly earnings and working hours in the BHPS.

To compute age-efficiency profiles η_a^{gp} by gender g and marital status p, we extract year fixed effects from our wage sample (earnings for men). To estimate them more precisely, we expand our sample to include the Understanding Society years (2010-2016).

$$\tilde{w}_{it} = \eta_a^{gp} + \xi_t + \epsilon_{it} \tag{16}$$

A.3 Comparing the BHPS and NESPD data

For the purposes of comparing moments from the two data sets, 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.

Figures 12 and 13 show that the implications of the BHPS and NESPD data are very similar in terms of all of the measures. The most salient difference is that average persistence is higher in the NESPD, which could reflect the presence of larger measurment error in the BHPS. However, the econometric procedure we describe in Section 2.1 allows us to separate the persistent and transitory components of earnings and wage changes; if present, most measurement error will be captured by the transitory component, which we do not include in our structural model.

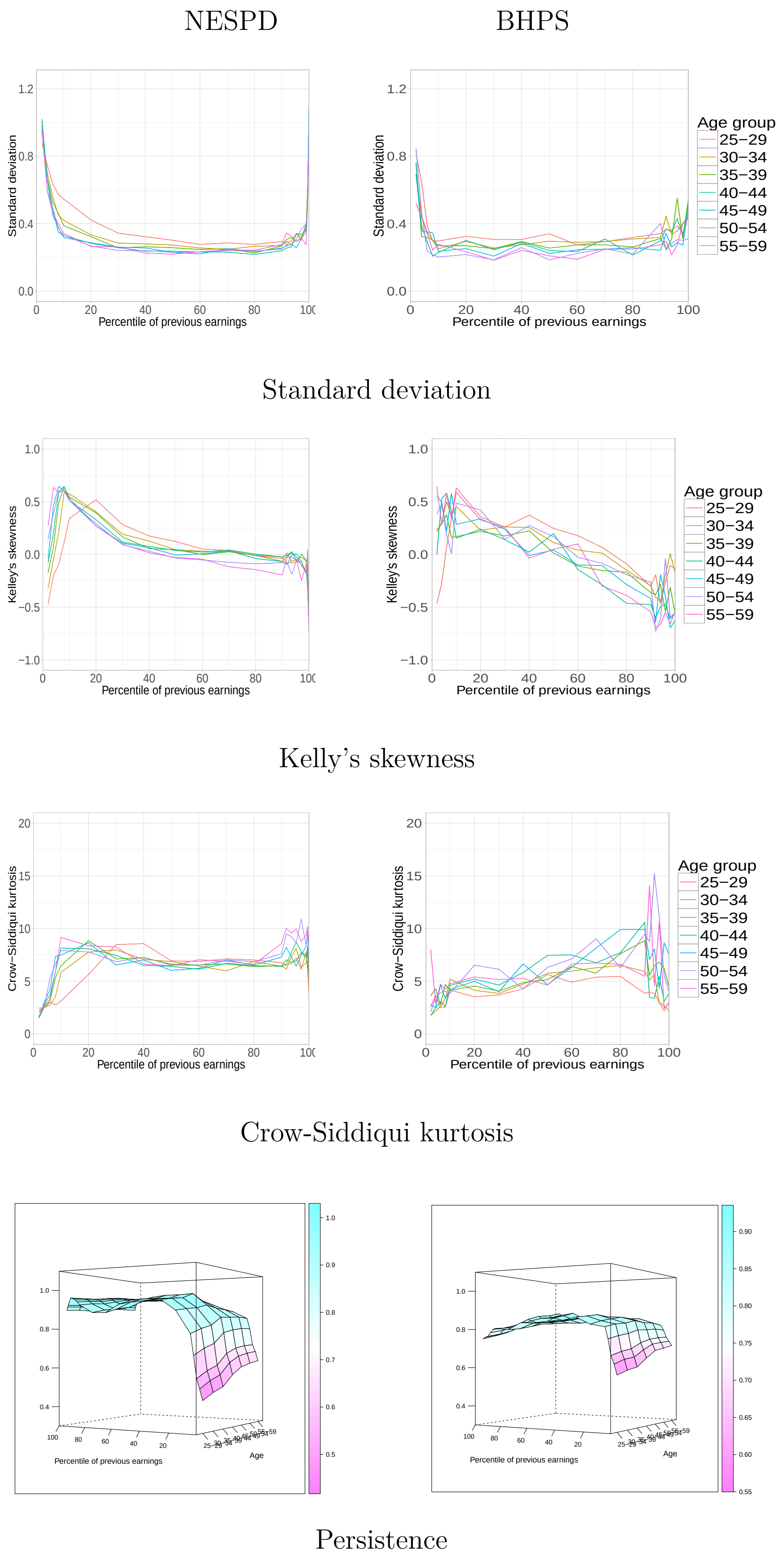

Figure 12: Moments of male earnings changes in the BHPS and NESPD. Top three panels: by previous earnings. Bottom panel, by previous earnings and age

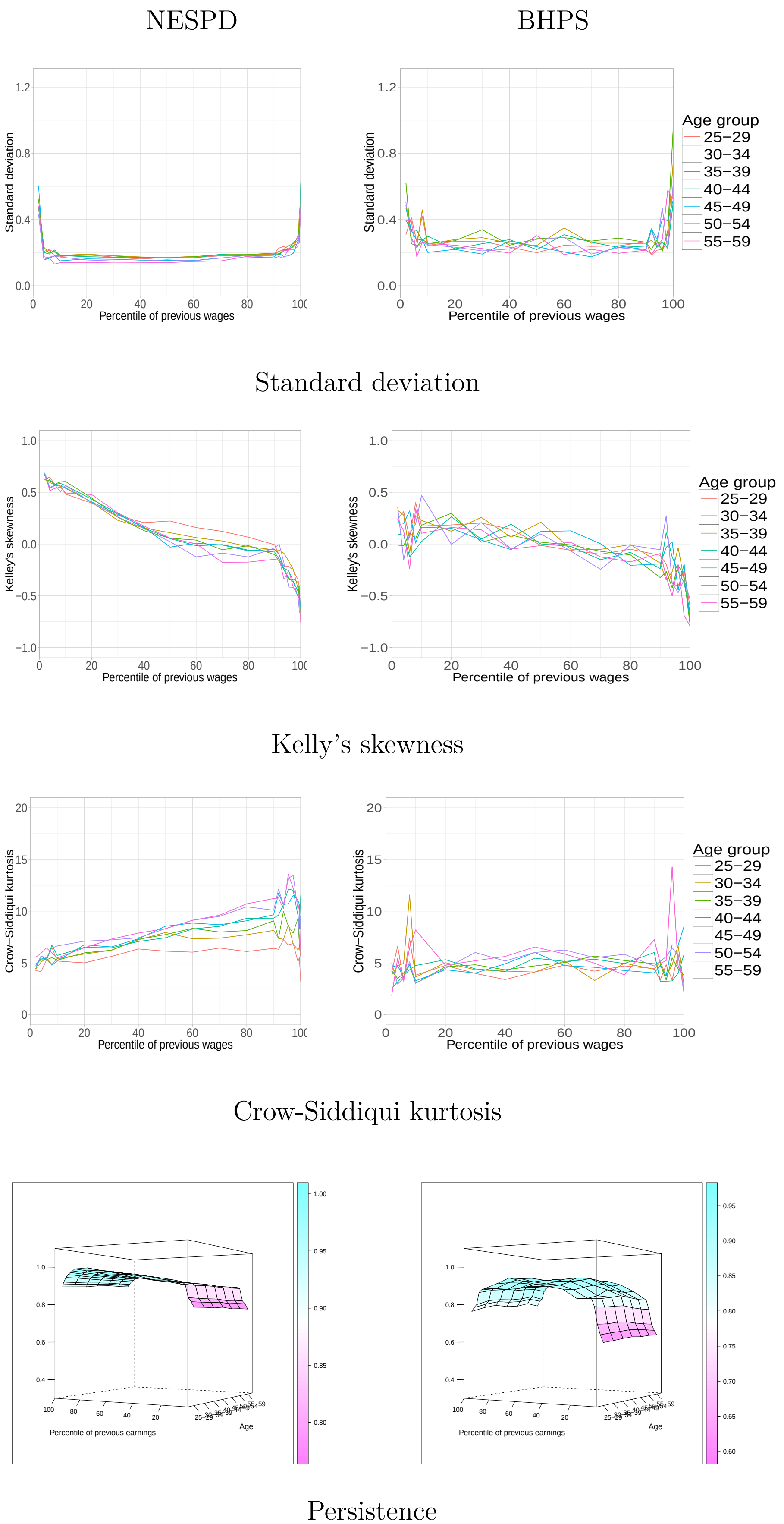

Figure 13: Moments of female wage changes in the BHPS and NESPD. Top three panels: by previous earnings. Bottom panel, by previous earnings and age

A.4 Women's wages imputation

We construct potential wages for women who do not work by using, whenever the age is missing, the predicted value from the regression

$$\log w_{it}^{f} = \alpha_i + \beta X_{it} + u_{it}, \tag{17}$$

which we estimated on all observations with a positive wage, where α_i is an individual fixed effect and X_{it} is a vector of demographic variables, including the first four powers of age, the first three powers of the number of years of work experience, the number of children, the age of the youngest child, marital status and whether the partner is currently working.

B Estimation and features of the earnings processes

B.1 Features of the persistent component

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 15). However, as expected, the persistent component displays a larger overall persistence than the data.

Figure 16 shows the standard deviation, skewness and kurtosis of earnings changes for BHPS data and the persistent component. The latter preserves most of the features regarding non-normality that are present in the data, and also the dependence on previous earnings realizations. The main exception is Crow-Siddiqui kurtosis, which is significantly larger for the persistent component than it is in the data.

Transitory shocks, that we consider to be measurement error, are very leptokurtic, in particular for male earnings, and display negative skewness (see Figure 17).

Women wages display similar patterns (see Figures 19, 21, 20, and 22).

Finally, in Figures 23 and 24 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. However, some salient features, in particular non-linear persistence, are significantly different between the earnings and the wage data for women, which suggests that labor supply decisions play an important role in explaining the non-linear features of earnings persistence.

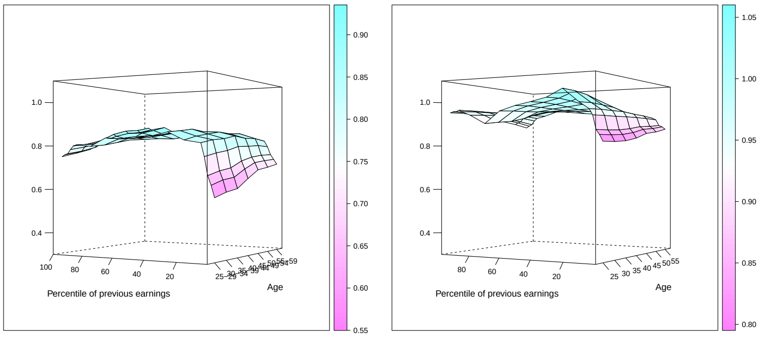

Figure 14: Non-linear persistence of male earnings by age and previous earnings in the BHPS. Left, data; right, persistent component

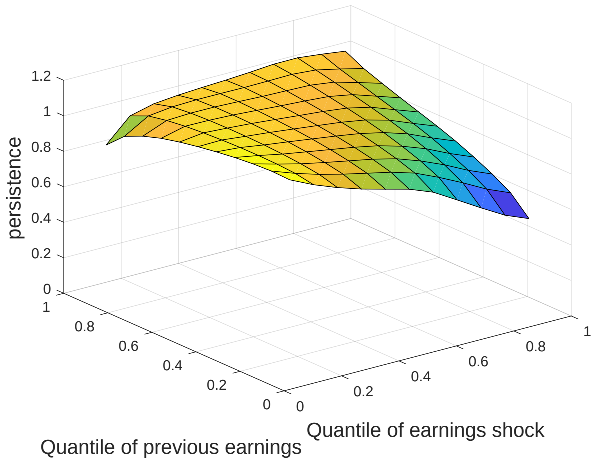

Figure 15: Non-linear persistence of male earnings by earnings and quantile of the shock: persistent component

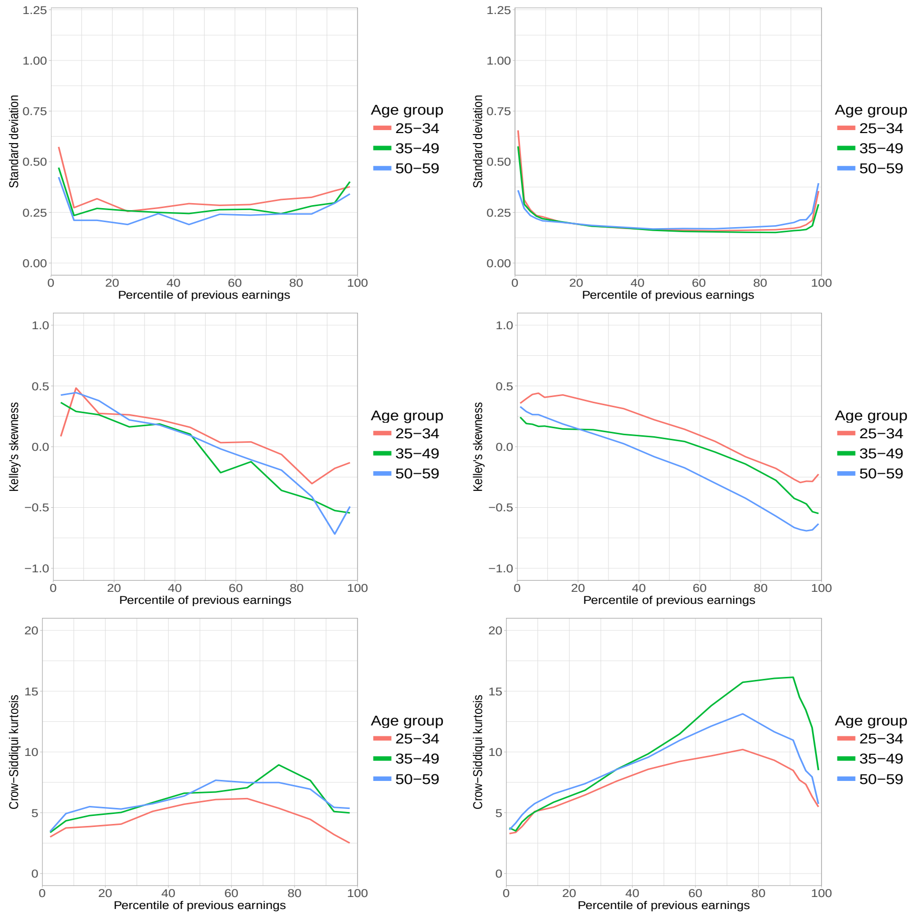

Figure 16: Standard deviation (top), skewness (middle) and kurtosis (bottom) of male earnings changes in the BHPS. Left, data; right, persistent component

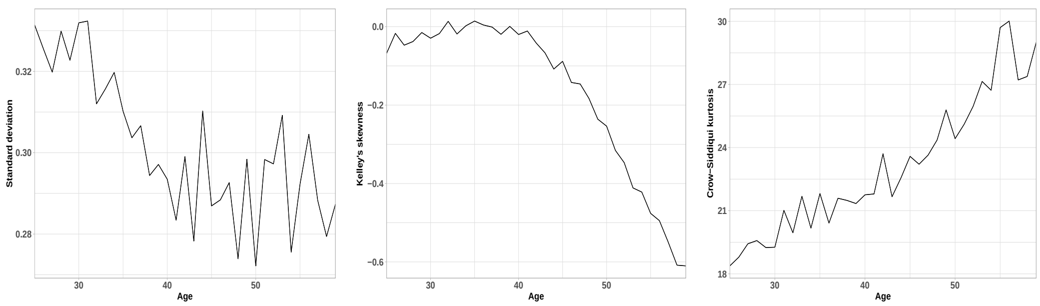

Figure 17: Transitory shock to male earnings: standard deviation, skewness and kurtosis by age

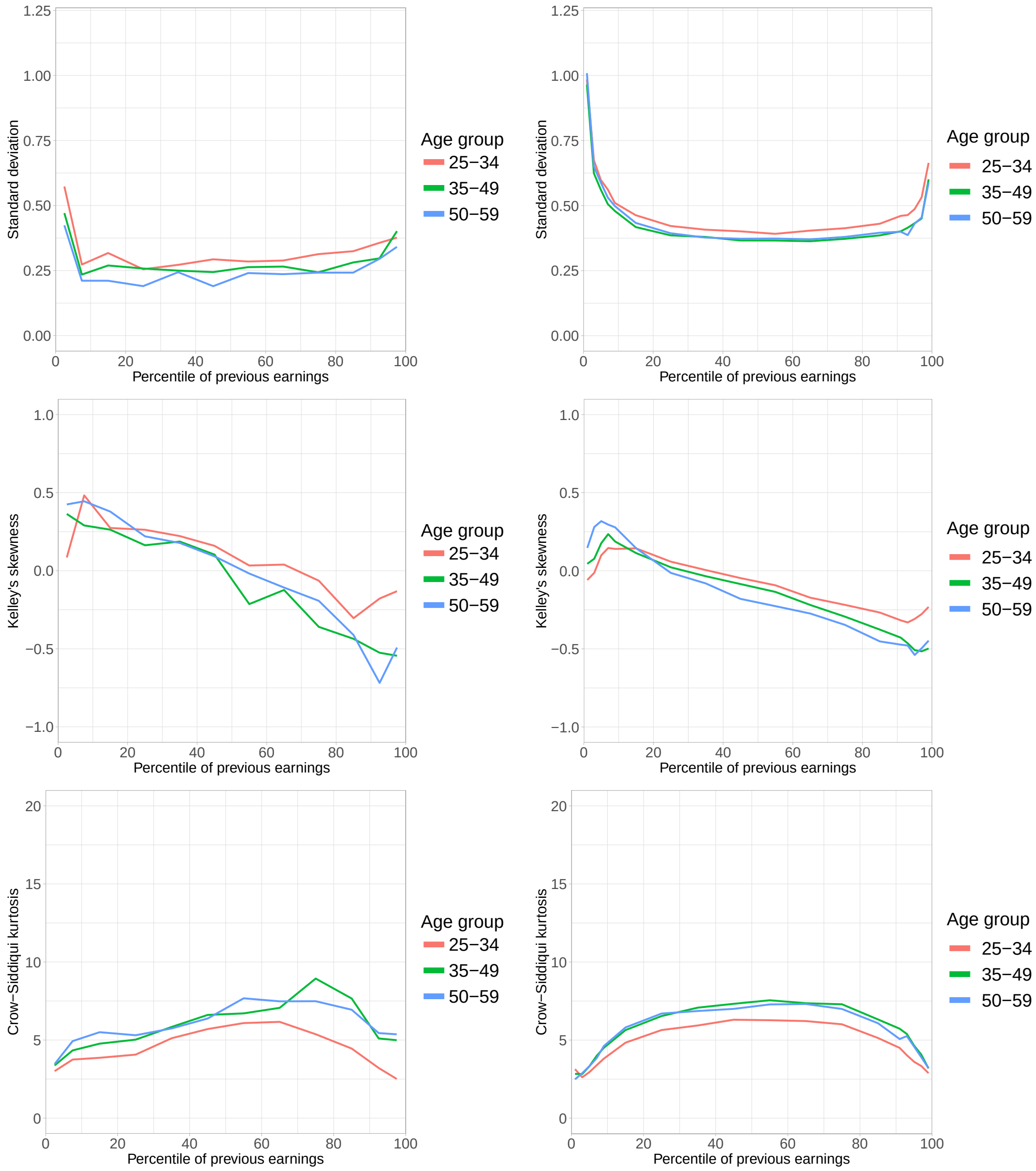

Figure 18: Standard deviation (top), skewness (middle) and kurtosis (bottom) of male earnings changes in the BHPS. Left, data; right, sum of persistent and transitory components

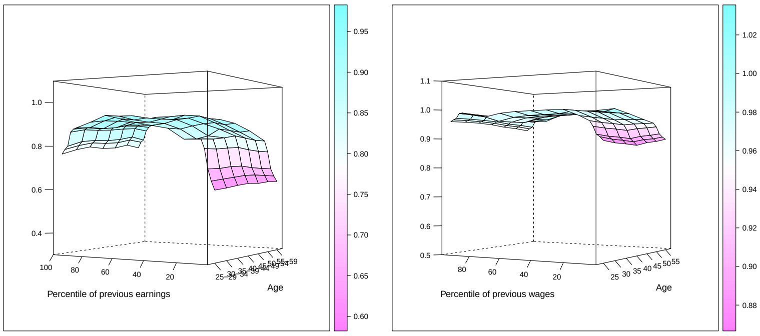

Figure 19: Non-linear persistence of female wages by age and previous earnings in the BHPS. Left, data; right, persistent component

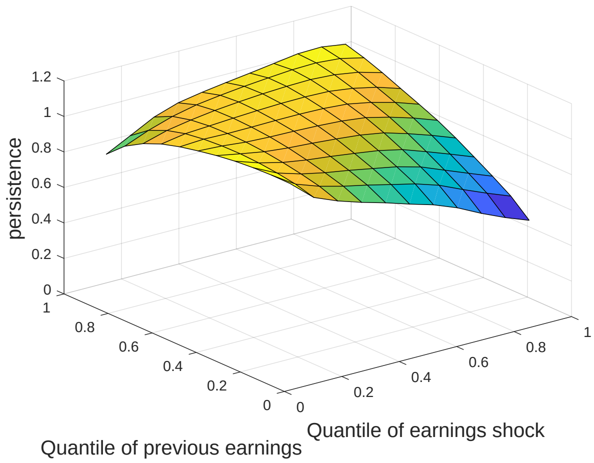

Figure 20: Non-linear persistence of female wages by earnings and quantile of the shock: persistent component

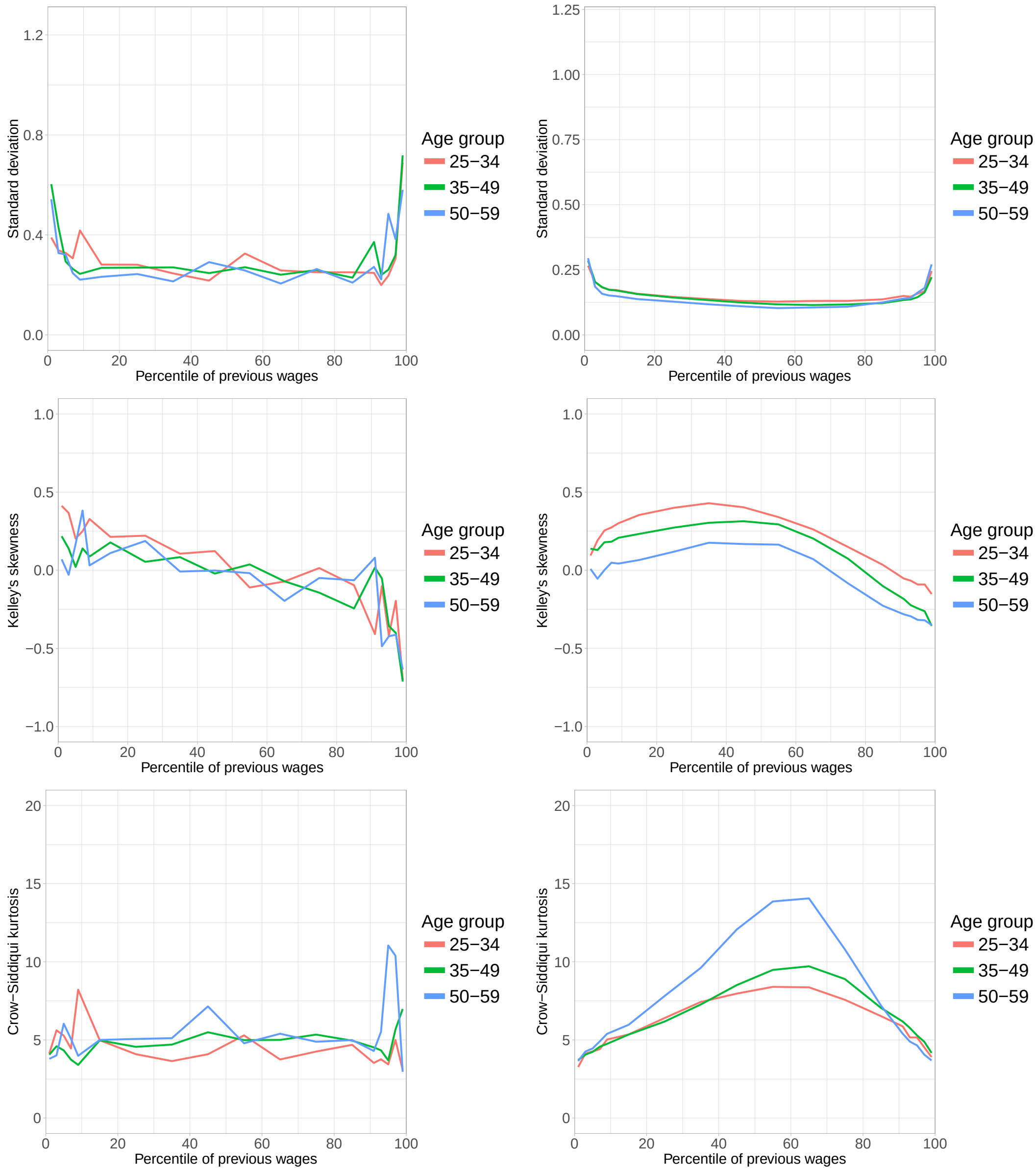

Figure 21: Standard deviation (top), skewness (middle) and kurtosis (bottom) of female wage changes in the BHPS. Left, data; right, persistent component

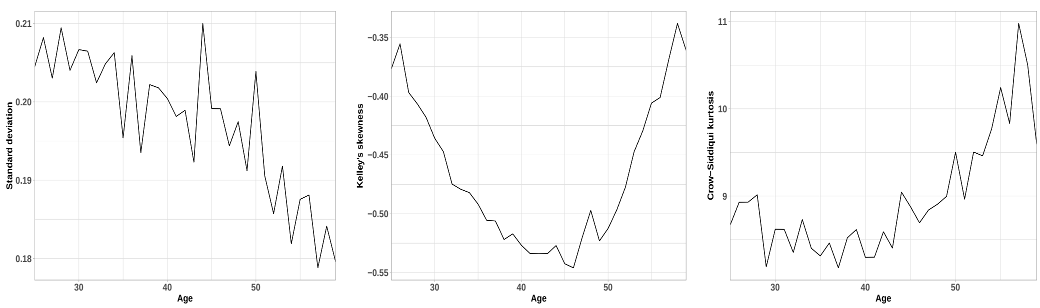

Figure 22: Transitory shock to female wages: standard deviation, skewness and kurtosis by age

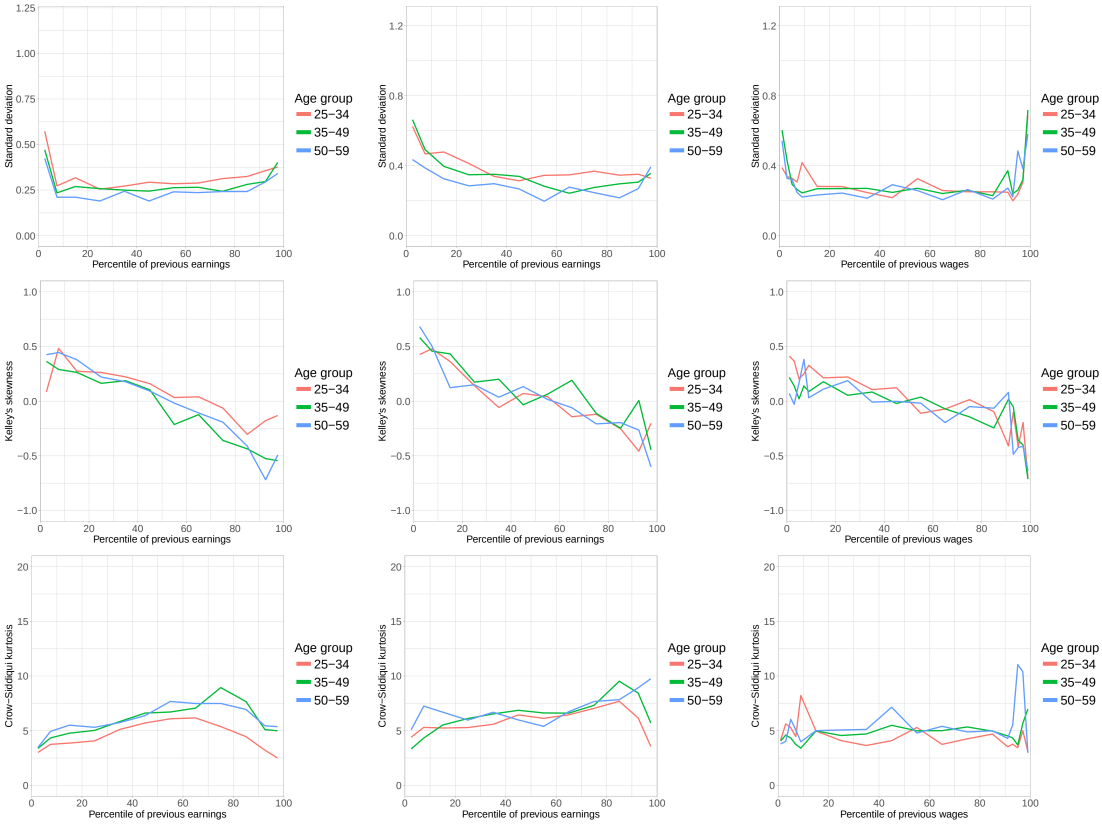

Figure 23: Standard deviation (top), skewness (middle) and kurtosis (bottom). Left: male earnings; middle: women earnings; right: women wages

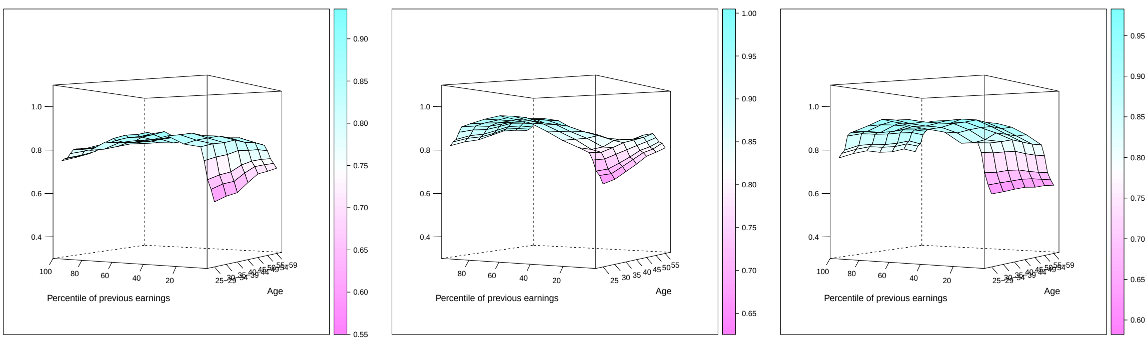

Figure 24: Non-linear persistence of male earnings (top left), female earnings (top right) and female wages (bottom) in the data

C Other model inputs

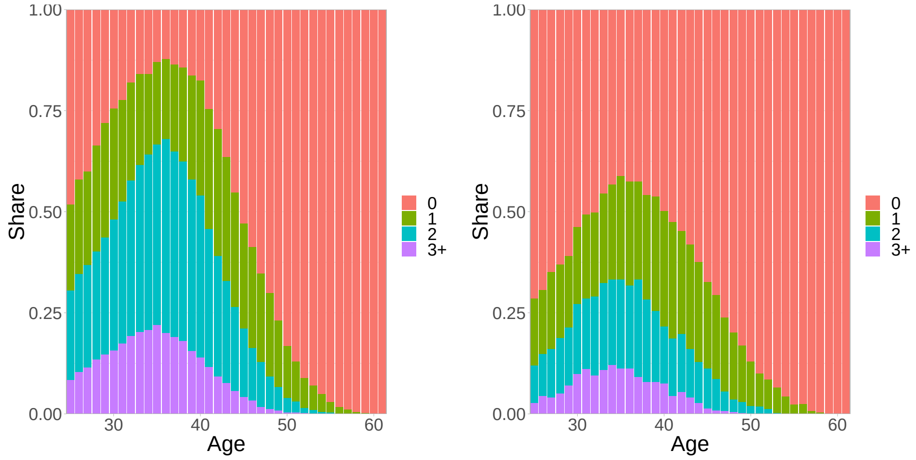

Figure 25: Distribution of number of kids in the household, by age of the mother. Left: married mothers; right: single mothers.

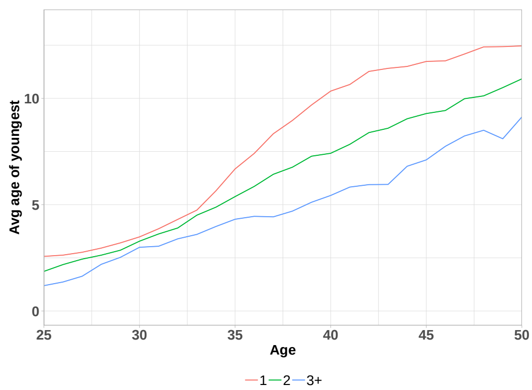

Figure 26: Average age of youngest cohabiting child by age of mother and number of kids.

Single men	Single women	Married couples
11%	19%	70 %

Table 7: Proportion of households by gender and marital status, BHPS data

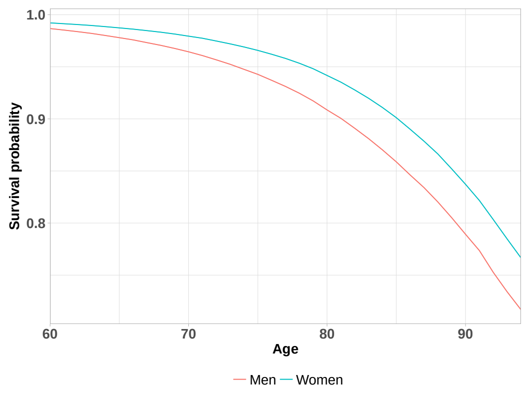

Figure 27: Survival probabilities by age and gender

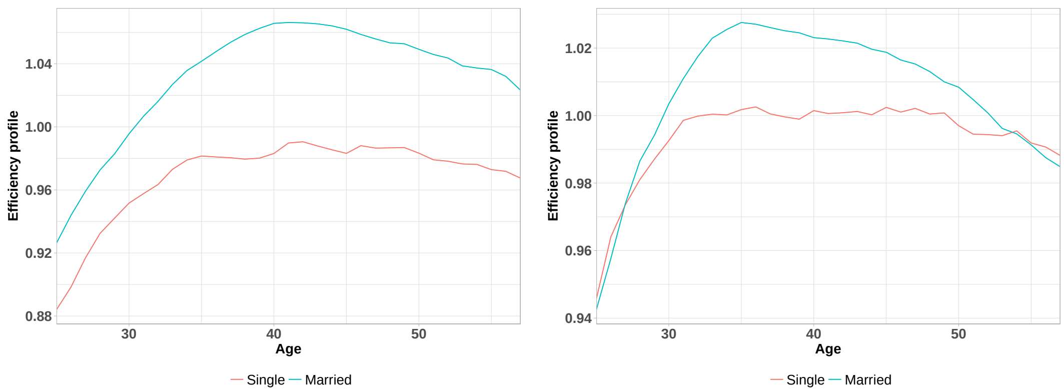

Figure 28: Age-efficiency profiles for wages, left: men; right: women. For this representation, both are individually normalized so that their average is 1

D UK Benefit system, details

Table 9 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.[8]

In the model, in-work benefits are intended to replicate the Working Tax Credit, while income support replicates a variety of benefits that low-income people receive under different circumstances, in particular Income-based Jobseeker's Allowance, Income Support, Housing Benefits, Child Benefits, and Child Tax Credits.

To obtain the tapering rate ω for the income support program in the model, we compute the benefit entitlement B_i^k of single men, single women and couples by number of children (k represents a group of gender/marital status and number of kids) and household labor income y_i^{hk}, jointly considering Income Support, Housing Benefits, Child Tax

[8]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.

Credits, and Child Benefits. We do not include disability benefits for this computation because they only accrue to a particular subset of workers, and we take into account that for our purposes Income-based Jobseeker's Allowance and Income Support are identical. We additionally assume that the household would be getting Working Tax Credit whenever eligible, which affects their eligiblity criteria for other benefits (namely, CTC and WTC are considered as income for purposes of computing eligibility for IS and HB).

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

$$\sum_i (B_i^k - max(\beta_0^k - \beta_1^k y_i^{hk}, 0))^2 \tag{18}$$

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 Universal Credit, canonical process

In this section, we compute the welfare effects of the introduction of Universal Credit under the canonical wage process. As decribed in Section 4, in our main results with the NL process we kept the change to Universal Credit budget neutral by multiplying all allowances with a proportional scaling factor of 0.86. For the purposes of this section, we keep that scaling factor constant, so the effective allowances of Universal Credit are also those reported in Table 2.[9]

Under the canonical wage process, the switch to Universal Credit also generates an increase in labor force participation amongst particularly older women (Figure 29). Yet, part-time labor force participation is lower at all ages under Universal Credit than under the optimal system. This is not surprising as we have seen that, under the canonical wage process, the optimal system provides very high incentive to labor force participation. The welfare implications are similar to those under the NL process (Table 8).

[9]Requiring budget neutrality under the canonical wage process would imply a slightly higher scaling factor (0.9).

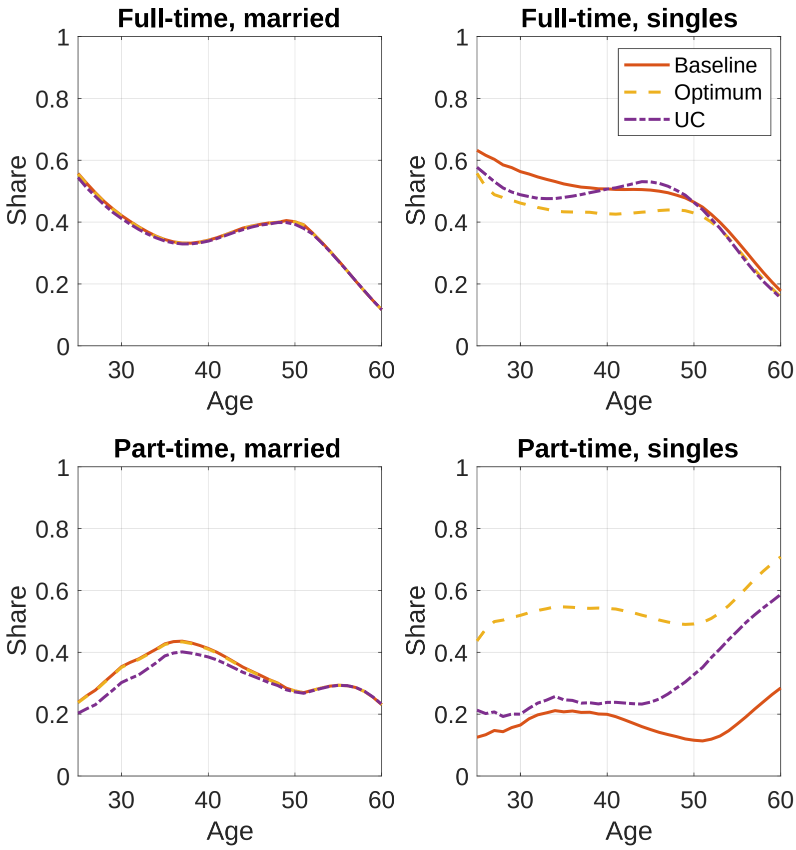

Figure 29: Labor force participation under canonical process: Universal Credit vs optimal benefits, universal credit.

	Canonical process				
Overall	0.38				
Single men	-0.01				
Single women	-0.41	-1.66	1.54	3.05	2.67
Couples	0.65	0.40	0.69	0.91	1.01

Table 8: Welfare change, by gender and marital status, for switch to Universal Credit. Canonical process.

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