Nonlinear household earnings dynamics, self-insurance, and welfare

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

October 1, 2018

Abstract

Earnings dynamics are much richer than typically assumed in macro models with heterogeneous agents. This holds for individual-pre-tax and household-post-tax earnings and across administrative (Social Security Administration) and survey (Panel Study of Income Dynamics) data. We estimate two alternative processes for household after-tax earnings and study their implications using a standard life-cycle model. Both processes feature a persistent and a transitory component, but while the first one is the canonical linear process with stationary shocks, the second one has substantially richer earnings dynamics, allowing for age-dependence of moments, non-normality, and nonlinearity in previous earnings and age. Allowing for richer earnings dynamics implies a substantially better fit of the evolution of cross-sectional consumption inequality over the life cycle and of the individual-level degree of consumption insurance against persistent earnings shocks. The richer earnings process also implies lower welfare costs of earnings risk.

Keywords: Earnings risk, savings, consumption, inequality, life cycle.
JEL Classification: D14, D31, E21, J31.

*Mariacristina De Nardi, Federal Reserve Bank of Minneapolis, UCL, CEPR, and NBER, denardim@nber.org. Giulio Fella: Queen Mary University of London, CFM, and IFS, g.fella@qmul.ac.uk. Gonzalo Paz Pardo: UCL, gonzalo.pardo.13@ucl.ac.uk. De Nardi gratefully acknowledges support from the ERC, grant 614328 "Savings and Risks." De Nardi is grateful to the Federal Reserve Bank of Chicago, where she carried out part of this research. Fella is grateful to UCL for their generous hospitality while he was working on this paper. Paz-Pardo gratefully acknowledges ESRC financial support for his postgraduate studies. We thank our Editor, Dirk Krueger, three anonymous referees, Marco Bassetto, Richard Blundell, Stephane Bonhomme, Tony Braun, Jeremy Lise, Fabrizio Perri, Fabien Postel-Vinay, Victor Ríos-Rull, Ananth Seshadri and Gustavo Ventura for helpful comments and suggestions. We are grateful to Moritz Kuhn for providing us with additional statistics from the Survey on Consumer Finances data set. The views expressed herein are those of the authors and do not necessarily reflect the views of the CEPR, the National Bureau of Economic Research, or the Federal Reserve Bank of Minneapolis.

1 Introduction

Macroeconomic models with heterogeneous agents are ideal laboratory economies to quantitatively study a large set of issues that include household behavior under uncertainty, inequality, and the effects of taxes, transfers, and social insurance reforms.[1] Earnings risk is a crucial source of heterogeneity in these models and its stochastic properties determine how saving and consumption adjust to buffer the impact of earnings shocks on current and future consumption. Appropriately capturing earnings risk is therefore important to understand consumption and wealth inequality, the welfare implications of income fluctuations, and the potential role for social insurance.

With few notable exceptions, most quantitative macroeconomic models adopt earnings processes that imply that persistence and other second and higher conditional moments are independent of age and earnings histories, and that shocks are normally distributed. The *canonical* permanent/transitory process is a popular example.

A growing body of empirical work, though, provides evidence that households' earnings dynamics feature non-normality, age-dependence, and nonlinearities, and devises flexible statistical models that allow for these features. For instance, recent work takes advantage of large, administrative datasets (e.g., W2 confidential Social Security Administration earnings data in Guvenen, Karahan, Ozkan and Song, 2016) and new methodologies applied to survey data sets like the Panel Study of Income Dynamics (PSID) (Arellano, Blundell and Bonhomme, 2017) to show that changes to pre-tax, individual male earnings display substantial skewness and kurtosis and that the persistence of shocks depends both on age and current earnings.[2]

[1]For instance, Scholz, Seshadri and Khitatrakun (2006) study the adequacy of savings at retirement, Storesletten, Telmer and Yaron (2004*a*); Krueger and Perri (2006); Heathcote, Storesletten and Violante (2010) study the evolution of consumption and Castañeda, Díaz-Giménez and Ríos-Rull (2003); De Nardi (2004); Cagetti and De Nardi (2009) study the evolution of wealth inequality over the life cycle, while Conesa, Kitao and Krueger (2009) study the optimal taxation of capital income.

[2]These features are consistent with several factors that affect the working lives of individuals. For instance, younger people tend to change jobs more frequently and this implies that the persistence of their earnings is lower. In addition, for most workers, earnings vary little from year to year and shocks are infrequent but can be of large magnitude, such as job loss or a career change, when they happen. This is captured by the

We show that all of these rich dynamics are present not only in individual pre-tax earnings, both in the W2 tax data and the PSID, but also in household, after-tax earnings, which are the relevant source of labor income risk at the household level.[3]

Our main contribution is to analyze the effects of richer earnings dynamics on consumption, wealth, and welfare, both in the cross-section and over the life cycle. We use the econometric framework recently proposed by Arellano et al. (2017) to separately identify the distributions of the persistent and transitory components of earnings while allowing for non-normality of shocks, non-linear persistence, and, in general, a rich dependence of the two distributions on age and (in the case of the persistent component) previous earnings. We use PSID data on after-tax, household earnings to estimate two different earnings processes: a richer earnings process along the lines of Arellano et al. (2017) and a "canonical" linear earnings process with a persistent and transitory component and normal innovations, like the one used in Storesletten et al. (2004*a*). For each process, we use two Markov chains to approximate the conditional distributions, respectively, of persistent and transitory shocks and introduce them into a partial-equilibrium, life-cycle model with incomplete markets to compare their implications for consumption, wealth, and welfare.[4]

Our main findings are as follows. First, compared to the canonical earnings process, the richer earnings process better fits the observed evolution of consumption inequality over the life cycle. More specifically, under the canonical earnings process, the growth in the variance of consumption substantially overshoots its data counterpart at all ages, while our richer process generates a realistic profile up to ages 50-55, when early and partial retirement start being important. The improved fit is due to the rich features of the earnings data that we

high level of kurtosis displayed by earnings changes.

[3]One caveat is that, in line with much of the previous literature, we take earnings as a primitive while earnings fluctuations likely partly reflect endogenous labour supply choices.

[4]Although our paper is highly indebted to Arellano et al. (2017), it differs from it in important respects. First, Arellano et al. (2017) estimate their earnings process on *pre-tax* earnings of households headed by *participating and married males* while we use *after-tax* earnings for *all* households. Secondly, Arellano et al. (2017) estimate the consumption rule semi-parametrically while we obtain it using the full model structure and, for the same reason, we can study welfare implications. Finally, the emphasis of Arellano et al. (2017) is on the consumption response to earnings shocks while we consider a wider range of implications.

model and to the households' precautionary saving response to them. In particular, the age-dependent persistence and variance of earnings innovations account for the main share of the improvement of the fit between age 25 and 45, while non-normality and, in particular, nonlinearity (for instance, the fact that persistence varies with the level of previous earnings) drive the improvement between age 45 and 55.

An alternative, and possibly more intuitive, measure of self-insurance is the extent of consumption passthrough of shocks to disposable earnings onto consumption. Our second finding is that the richer earnings process implies a consumption passthrough of persistent earnings shocks broadly consistent with the data. Its value is 0.57 which is within one standard deviation of the point estimate of 0.64 by Blundell, Pistaferri and Preston (2008). This result too is driven by the age-dependent persistence of shocks and can be understood in light of Kaplan and Violante's (2010) finding that a persistent, but not permanent, earnings process can imply the "right" level of insurance against persistent shocks. The richer earnings process implies a significantly lower degree of persistence of the persistent earnings component, particularly at younger ages, compared to the canonical process. Consequently, persistent shocks, particularly at younger ages when assets are low, are more easily self-insured through borrowing and lending.

Our third finding is that our rich earnings process does not improve the fit of the right tail of the wealth distribution with respect to the canonical earnings process.[5] This is perhaps not so surprising given an established literature, surveyed in De Nardi and Fella (2017), pointing to the fact that accounting for the saving of the rich requires mechanisms—such as a non-homothetic bequest motive, medical-expense risk and entrepreneurship—that go beyond idiosyncratic earnings risk.

Finally, from a normative perspective we find that the welfare costs of earnings risk—as measured by the yearly consumption equivalent—are 2 percentage points lower under the

[5] In De Nardi, Fella and Paz-Pardo (2016) we show that this conclusion still holds if we estimate a similar richer process on synthetically generated W2 data from the parametric processes proposed in Guvenen et al. (2016). It is thus not related to issues of lack-of-oversampling and non-participation by higher income people that are usually associated with most survey data sets.

richer earnings process than under the canonical one. The main reason for this finding is again that, while under the canonical process earnings have a permanent, random-walk, component, the richer process implies a lower persistence, particularly in the first part of the working life. As a result, life-cycle risk can be more effectively self-insured under the richer earnings process. Interestingly, the reduction in welfare costs would be even higher—4 rather than 2 percentage points—in the absence of non-normality and nonlinearities that partly offsets the welfare gains due to age-dependent persistence and innovation variances.

An additional contribution of this paper is to propose a simple, simulation-based, method to discretize nonlinear and nonnormal stochastic processes to introduce them in a computational model. Standard discretization methods used in macroeconomics, such as Tauchen (1986) and Rouwenhorst (1995), require the continuous process to be approximated to be linear, typically an AR(1), and, in the case of Tauchen (1986), normal.[6] Our method applies to any, otherwise unrestricted, age-dependent, first-order Markov process. It relies on simulating a panel of individual earnings histories using the continuous process to be approximated and estimating an age-specific, first-order Markov chain on it. This is achieved by discretizing the simulated marginal distribution of earnings at each age—e.g. into percentiles—and by replacing the (heterogeneous) values of earnings in each rank percentile with their median. The associated, age-specific transition matrix is then obtained by computing the proportion of observations transiting from each percentile rank of the earnings distribution at age t to that at age $t+1$. The result is a non-parametric representation of the process that follows a Markov chain with an age-dependent transition matrix and a fixed number of age-dependent earnings states.[7]

Our paper is related to the econometric literature on earnings dynamics[8] as well as the

[6]Fella, Gallipoli and Pan (2017) show how Tauchen (1986) and Rouwenhorst (1995) can be extended to allow for age dependence. Their method still requires linearity though.

[7]It should be noted that our method can be generalized to allow for Markov processes of order higher than one.

[8]Besides Arellano et al. (2017) and Guvenen et al. (2016), discussed above, it includes Geweke and Keane (2000), Lillard and Willis (1978), Bonhomme and Robin (2009), Meghir and Pistaferri (2004), Blundell, Graber and Mogstad (2015), Browning, Ejrnaes and Álvarez (2010), and Altonji, Smith and Vidangos (2013). Recent developments are discussed in Meghir and Pistaferri (2011).

macroeconomic literature on the relationship between income and consumption and wealth inequality over the life cycle. Deaton and Paxson (1994) is the seminal empirical contribution. Storesletten et al. (2004*a*), Guvenen (2007), Primiceri and Van Rens (2009), Huggett, Ventura and Yaron (2011) and Guvenen and Smith (2014) analyze lifetime inequality from the perspective of the standard, incomplete markets model as we do here. Within this literature, many of the consequences of richer earnings processes on consumption, savings and welfare in structural models are still unexplored, with few exceptions. Castañeda et al. (2003) propose an "awesome or superstar" shock to earnings that is unlikely to be observed in the data but that can help to explain the emergence of super-rich people. Karahan and Ozkan (2013) study the implications of age-dependent persistence and variance of shocks. McKay (2017) finds that taking into account the procyclical nature of negative skewness in earnings growth rates substantially raises the volatility of aggregate consumption growth. Golosov, Troshkin and Tsyvinski (2016) show that higher order moments of earnings shocks are important determinants of optimal redistribution and insurance. Civale, Díez-Catalán and Fazilet (2016) study the implications of skewness and kurtosis for the aggregate capital stock in an economy à la Aiyagari (1994).

The rest of the paper is organized as follows. Section 2 describes the main features of the data on earnings dynamics for both individuals and households. Section 3 details the methods we use to estimate the canonical and nonlinear earnings processes and their implications. Section 4 explains the discretization procedure we propose to tractably introduce rich nonlinear earnings dynamics in a quantitative life-cycle model. Section 5 presents the model and its calibration. Section 6 discusses the consumption, wealth, and welfare implications of the two earnings processes that we consider, and decomposes the determinants of their differences. Section 7 concludes. The Appendix discusses key features of the PSID data, details of the estimation and reports a number of robustness checks.

2 Earnings data and their features

Recent empirical literature has called into question the established view that (log-)earnings dynamics are well approximated by a linear model of which the canonical random-walk permanent/transitory model (Abowd and Card, 1989) with normal innovations is a popular example. Linear models imply that persistence and other second and higher moments are independent of earnings histories. Instead, Guvenen et al. (2016) and Arellano et al. (2017) document that, contrary to the implications of the canonical model, individual pre-tax earnings display both substantial deviations from log-normality and non-linearity.

Guvenen et al. (2016) use confidential Social Security Administration (W2) tax data to establish these facts. The W2 data set has both advantages and disadvantages compared to the PSID data (and household survey data sets more generally). Regarding its advantages, the W2 data set has a large number of observations, is less likely to be contaminated by measurement error, and is not affected by top-coding and differential survey responses. Thus, it could provide better information on the top earners to the extent that they do not respond to surveys but do pay taxes on all of their earnings. An important disadvantage of the W2 data set is that it is collected at the individual level and lacks the information to identify households and thus to construct household earnings.

The latter is an important shortcoming. In the U.S., the majority of adults are married, 95% of married couples file their income taxes jointly, and taxation of couples and singles is different. Therefore, one needs to know the earnings of both people in a household in order to compute disposable earnings. In this respect household survey data sets that keep track of household structure, like the PSID, have a distinct advantage. This is particularly important if, as we do here, one wants to understand the implications of earnings risk for consumption insurance, which requires taking into account that households and taxes provide insurance against earnings shocks. For such a purpose, *disposable household earnings* is the relevant variable of interest.

The data used in this paper are from the Panel Study of Income Dynamics (PSID), 1968-

1992. Our sample consists of households who are in the representative core sample, whose head is between 25 and 60 years of age. Given the paper's focus on the implications of earnings risk for consumption insurance, our main variable of interest is *disposable household* labor earnings, although we also discuss the properties of individual pre-tax labor earnings for the purpose of comparison with some closely related work (e.g. Arellano et al., 2017; Guvenen et al., 2016).

Disposable, household labor earnings are defined as the sum of household labor income and transfers, such as welfare payments, net of taxes and Social Security contributions paid. Appendix A contains a more detailed description of the PSID data we use, our definition of household earnings and how we estimate taxes on labor following Guvenen and Smith (2014).

We adjust our earnings measure for demographic differences across households, since these have no counterpart in the model in Section 5. We do so by regressing log earnings on family size. We apply the same regression to the CEX consumption data we use to compute the moments reported in Section 6. The residuals from these regressions are interpreted as earnings and consumption per-adult equivalent in the analysis below. For both earnings and consumption, we extract business cycle effects by running a regression of their log levels on year dummies. Finally, we separate the predictable from the stochastic component of earnings by running a regression of our equivalized earnings measure on age dummies. We use the residuals to estimate the stochastic processes for earnings in what follows.

2.1 Individual pre-tax earnings in the PSID and the W2 data

We now turn to comparing the properties of *individual pre-tax* earnings data in the PSID with those in the W2 data reported by Guvenen et al. (2016).

The top two rows of Figure 1 compare the second to fourth moments of the W2 data and the PSID. The top row, derived from the moments reported by Guvenen et al. (2016), plots the conditional standard deviation, skewness and kurtosis (measured as the third and

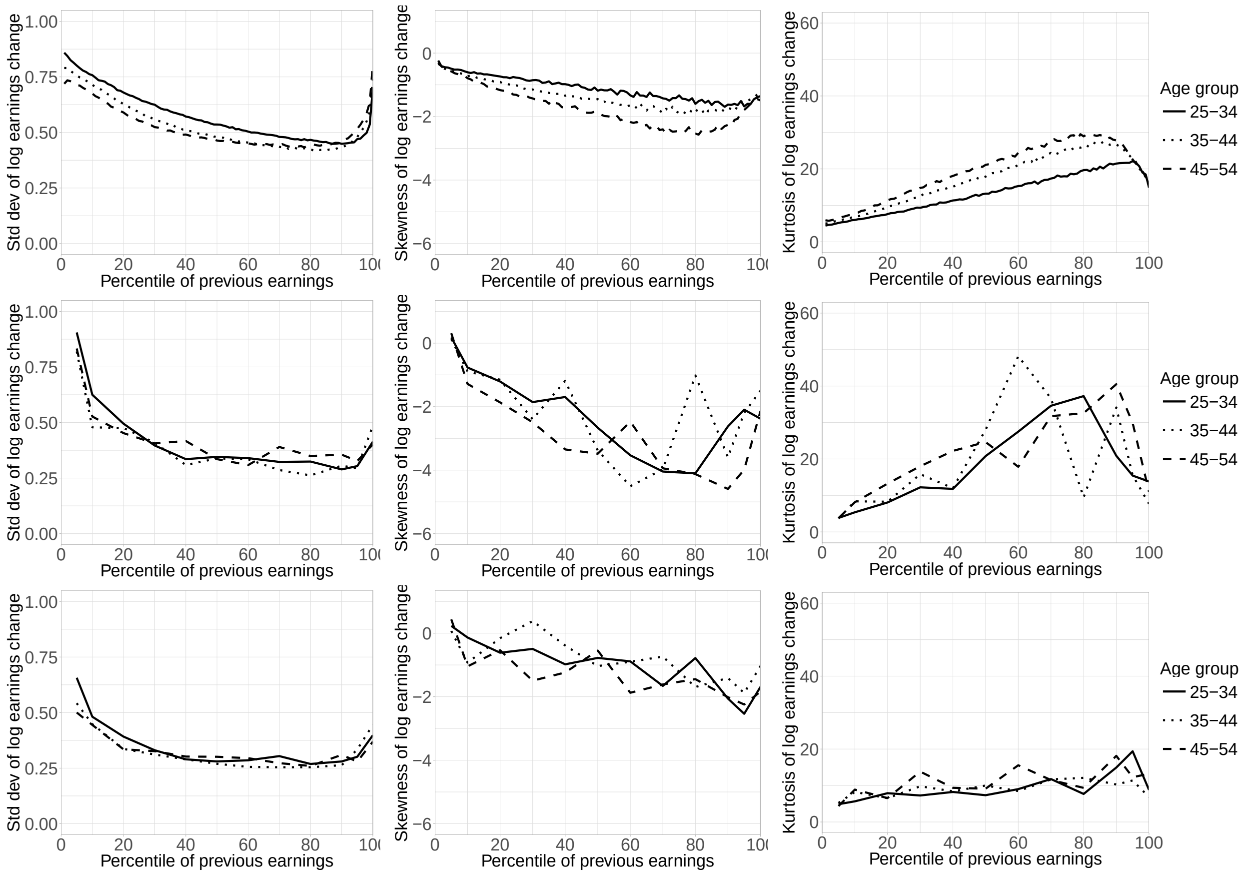

Figure 1: Standard deviation, skewness, and kurtosis of male pre-tax earnings growth in the W2 (top row) and PSID (middle row), and of household after-tax earnings growth in the PSID (bottom row)

fourth standardized moments) of *individual pre-tax* log earnings growth in the W2 data by age and percentile of previous earnings. The middle row of the same figure reports the same moments, by age and decile of previous earnings, for the PSID.[9]

Comparing these two sets of figures shows that, overall, the moments in the PSID data are both qualitatively and quantitatively close to those computed from the W2 data. More specifically, the conditional standard deviation of *individual pre-tax* log earnings growth is U-shaped across all age groups, declining until the 40th percentile and increasing again from the 90th percentile onwards. The increase is more pronounced for the top percentiles in the

[9]For comparability with Guvenen et al. (2016), we report moments for households whose head is a male. All moments are very close to those including female household heads. We show moments for 10-year age groups, which we obtain, in the W2 case, by averaging the moments reported by Guvenen et al. (2016) with equal weights for each 5-year age group.

W2, likely reflecting the coarser partition of the distribution in the PSID data. The most notable difference is the much higher variance at all percentiles above the 20th in the W2 data.

The figures also show that in both datasets *individual pre-tax* log earnings growth has strong negative skewness and very high kurtosis, and that these moments differ by both age and previous earnings. Skewness is more negative for individuals in higher earnings percentiles and for individuals in the 45-54 age group. This indicates that individuals face a larger downward risk in middle age.[10] The comparison of the implications of the two data sets also reveals that, if anything, there is more negative skewness in the PSID data than in the W2 data, except perhaps at the lowest earnings percentiles.

The kurtosis of *individual pre-tax* log earnings growth is hump-shaped by earnings percentile. Even for kurtosis, the maximum value is higher in the PSID, 40, against 30 in the W2 data (compared to 3 for a standard normal distribution).[11]

Taken together, these moments provide strong evidence against the standard assumption of a log-normal and linear earnings process for *individual pre-tax* log earnings growth for the PSID data, as well as for the W2 data.

2.2 Individual pre-tax versus household disposable earnings in the PSID

Now that we have shown that *individual pre-tax* earnings growth in both the W2 and PSID data displays remarkably similar deviations from normality and linearity, we turn to contrasting the properties of *individual pre-tax* and *disposable household* earnings in the PSID.

The bottom row in Figure 1 reports the same moments as the first two rows but for

[10] Graber and Lise (2015) generate this kind of earnings behavior in the context of a search and matching model with a job ladder.

[11] The levels and profiles of skewness and kurtosis of *individual pre-tax* log earnings growth are similar in the two datasets also when looking at report robust measures that exclude outliers (Kelly skewness and Crow-Siddiqui kurtosis). The main difference is a higher level of Crow-Siddiqui kurtosis in the W2 data than in the PSID (see Appendix C).

disposable household log earnings growth (bottom row) in the PSID. Comparing the bottom to the middle row reveals that, as one might have expected, disposable household earnings display lower variance, skewness, and kurtosis than pre-tax individual earnings. More specifically, the standard deviation is about 20% smaller at the lower end of the distribution of previous earnings, while skewness and kurtosis are about half as large. Thus, households and taxes perform an important insurance role in buffering individuals from pre-tax earnings changes (as shown by Blundell, Pistaferri and Saporta-Eksten (2016)). This has to be taken into account when considering the economic implications of earnings shocks.

To sum up, the above discussion has shown that, even after taking into account the insurance implied by pooling at the household level and the tax and welfare system, labor earnings display features that contrast with the *age-independence*, *normality,* and *linearity* (independence of variance, skewness and kurtosis of previous earnings realizations) implied by the canonical earnings process.

The same is true of another aspect on nonlinearity, nonlinear persistence, that has been documented by Arellano et al. (2017) using *pre-tax* earnings from the PSID. Figure 2 shows how this same feature is prominent also for *disposable* household earnings. It reports earnings persistence as a function of both the previous- and current-earnings rank in our PSID sample. In line with Arellano et al.'s (2017) findings, we also find that earnings persistence is lower (about 0.6) when previous earnings are highest and the current earning shock is lowest and when previous earnings are lowest and the current earning shock is highest (0.4).

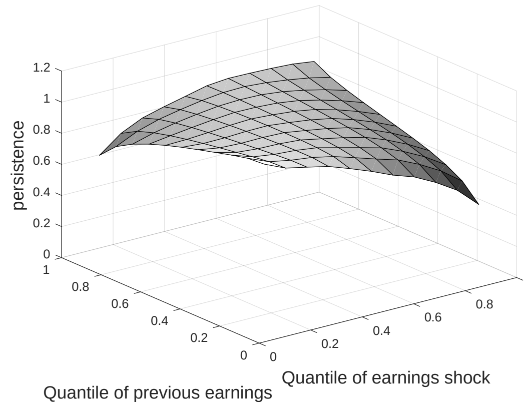

Figure 2: Persistence in log-earnings as a function of previous earnings rank and the rank of the shock received in the current period. PSID data.

3 Earnings processes and their estimation

3.1 Earnings processes

We start by introducing the canonical linear model of earnings dynamics used in macroeconomics before presenting its nonlinear generalization in Arellano et al. (2017).

Consider a cohort of households indexed by i and denote by $t = 1, \ldots, T$ the age of the household head. Let y_{it} denote the logarithm of (residual) disposable household earnings for household i at age t which can be decomposed as

$$y_{it} = \eta_{it} + \varepsilon_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T \tag{1}$$

where η and ε are assumed to have absolutely continuous distributions. The first component, η_{it}, is assumed to be *persistent* and to follow a first-order Markov process. The second component, ε_{it}, is assumed to be *transitory*, have zero mean, be independent over time and of η_{is} for all s.

The *canonical (linear) model* used in macroeconomics is described by

$$\eta_{i,t} = \rho\eta_{i,t-1} + \zeta_{it}, \tag{2}$$

$$\eta_{i1} \overset{id}{\sim} N(0, \sigma_{\eta_1}), \quad \zeta_{it} \overset{iid}{\sim} N(0, \sigma_\zeta), \quad \varepsilon_{it} \overset{iid}{\sim} N(0, \sigma_\varepsilon). \tag{3}$$

Thus, the persistent component η_{it} is an autoregressive process of order one with the innovation ζ_{it} independent of $\eta_{i,t-1}$, while the transitory component ε_{it} is white noise.

Equations (2)-(3) impose three types of restrictions

1. *Age-independence* (stationarity) of the autoregressive coefficient ρ and of the shock distributions, which imply age-independence of the second and higher moments of the conditional distributions of both the transitory and the persistent component. This is clearly at odds with the strong age-dependence in Figure 1.

2. *Normality* of the shock distributions, which is inconsistent with the negative skewness and high kurtosis discussed above.

3. *Linearity* of the process for the persistent component. Linearity implies: (a) the additive separability of the right hand side of equation (2) into the conditional expectation—the first addendum—and an innovation ζ_{it} independent of $\eta_{i,t-1}$, and (b) the linearity of the conditional expectation in $\eta_{i,t-1}$. Under separability, deviations of η_{it} from its conditional expectation are just a function of the innovation ζ_{it}. As a consequence, all conditional centered second and higher moments are independent of previous realizations of η. This is clearly inconsistent with the dependence of the moments reported in figures 1 and 2 on previous earnings realizations.

The evidence discussed in Section 2.1 is what motivates us to consider a more general process that relaxes the above three restrictions while maintaining the first-order Markov assumption for η. The question of how to easily introduce a richer and yet tractable earnings process in a structural model is non-trivial and part of what we propose in this paper.

We proceed in two steps. First, we use the quantile-based panel data method proposed by Arellano et al. (2017) to estimate a non-parametric model that allows for age-dependence, non-normality and nonlinearity, and that can be applied in datasets of moderate sample size like the PSID. This step gives us quantile functions for both the two (persistent and transitory) component of earnings. Second, we use the two quantile functions to simulate histories for the two earnings components and proceed to estimate, for each of them, a discrete Markov-chain approximation, which can then be easily introduced in a structural model (this latter step is discussed in detail in Section 4).

Let $Q_z(q|\cdot)$, the conditional quantile function for the variable z, denote the qth conditional quantile of z.[12] The process for η can be written in a very general form by replacing equation (2) with

$$\eta_{it} = Q_\eta(v_{it}|\eta_{i,t-1}, t), \quad v_{it} \overset{iid}{\sim} U(0,1), \; t > 1. \tag{4}$$

Intuitively, the quantile function maps random draws v_{it} from the uniform distribution over $(0,1)$ (cumulative probabilities) into corresponding random (quantile) draws for η. In the linear case in equation (2) the quantile function specializes to the linearly separable form $Q_\eta(v_{it}|\eta_{i,t-1}, t) = \rho\eta_{i,t-1} + \phi^{-1}(v_{it}; \sigma_\zeta)$, where $\phi^{-1}(v_{it}; \sigma_\zeta)$ is the inverse of the cumulative density function of a normal distribution with zero mean and standard deviations σ_ζ. So, age-independence, normality, and linearity can be seen as restrictions on the quantile function in equation (4).

In particular, one way to understand the role of nonlinearity is in terms of a generalized notion of persistence

$$\rho(q|\eta_{i,t-1}, t) = \frac{\partial Q_\eta(q|\eta_{i,t-1}, t)}{\partial \eta_{i,t-1}} \tag{5}$$

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

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

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

3.2 Estimating the nonlinear earnings process

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

$$Q_\varepsilon(q|age_{it}) = \sum_{k=0}^{K} a_k^\varepsilon(q)\psi_k(age_{it}) \tag{6}$$

$$Q_{\eta_1}(q|age_{i1}) = \sum_{k=0}^{K} a_k^{\eta_1}(q)\psi_k(age_{i1}) \tag{7}$$

$$Q_\eta(q|\eta_{i,t-1}, age_{it}) = \sum_{k=0}^{K} a_k^\eta(q)\psi_k(\eta_{i,t-1}, age_{it}) \tag{8}$$

where the coefficients $a_k^i(q)$, $i = \varepsilon, \eta_1, \eta$, are modelled as piecewise-linear splines in q on a grid $\{q_1 < \ldots < q_L\} \in (0, 1)$.[13] The intercept coefficients $a_0^i(q)$ for q in $(0, q_1]$ and $[q_L, 1)$ are specified as the quantiles of an exponential distribution with parameters λ_1^i and λ_L^i.

If the two earnings components ε_{it} and η_{it} were observable one could compute the polynomial coefficients simply by quantile regression for each point of the quantile grid q_j. To deal with the latent earnings components, the estimation algorithm starts from an initial guess for the coefficients and iterates sequentially between draws from the posterior distribution of the latent persistent components of earnings and quantile regression estimation until

[13] Following Arellano et al. (2017), we use tensor products of Hermite polynomials of degrees (3,2) in $\eta_{i,t-1}$, and age for $Q_\eta(q|\eta_{i,t-1}, age_{it})$ and second-order polynomials in age for $Q_\varepsilon(q|age_{it})$ and $Q_{\eta_1}(q|age_{i1})$.

σ^2_ε	$\sigma^2_{\eta_1}$	σ^2_ζ	ρ
0.0620	0.2332	0.0060	1
(0.0020)	(0.0061)	(0.0004)	*

Table 1: Estimates for the canonical earnings process (standard errors in parentheses)

convergence of the sequence of coefficient estimates.

Reported standard errors are computed by bootstrapping. In particular, we sample with replacement pairs of observations y_t, y_{t+1} from our PSID sample and then run the estimation algorithm for a large number of those samples.

3.3 Estimating the canonical linear earnings process

We estimate the canonical process for residual earnings in equations (1)-(3) using GMM, where the target moments are variance and autocovariance age profiles in the data.[14] The associated standard errors are obtained by bootstrapping. Table 1 reports our results. As common in the literature, we find that the persistent component has a unit root.[15] For this reason, though we have allowed for individual fixed effects at the estimation stage, their variance cannot be identified separately from the variance of the initial condition $\sigma^2_{\eta_1}$ and we have normalized it to zero.

3.4 Comparing the implications of the nonlinear and canonical earning processes

To understand the economic implications of the nonlinear and canonical earnings processes, it is useful to compare their implications in terms of (a) age-dependence of second moments; (b) non-normality; (c) nonlinearity.

[14] Appendix A.4.2 provides more information about our estimation method.

[15] The unrestricted GMM estimation returns an estimate of $\rho = 1.01$. Given that there has been little exploration of the explosive case in the literature, and that we want our canonical process to be standard, the above estimates are obtained under the restriction $\rho \leq 1$. The resulting estimate is at the bound.

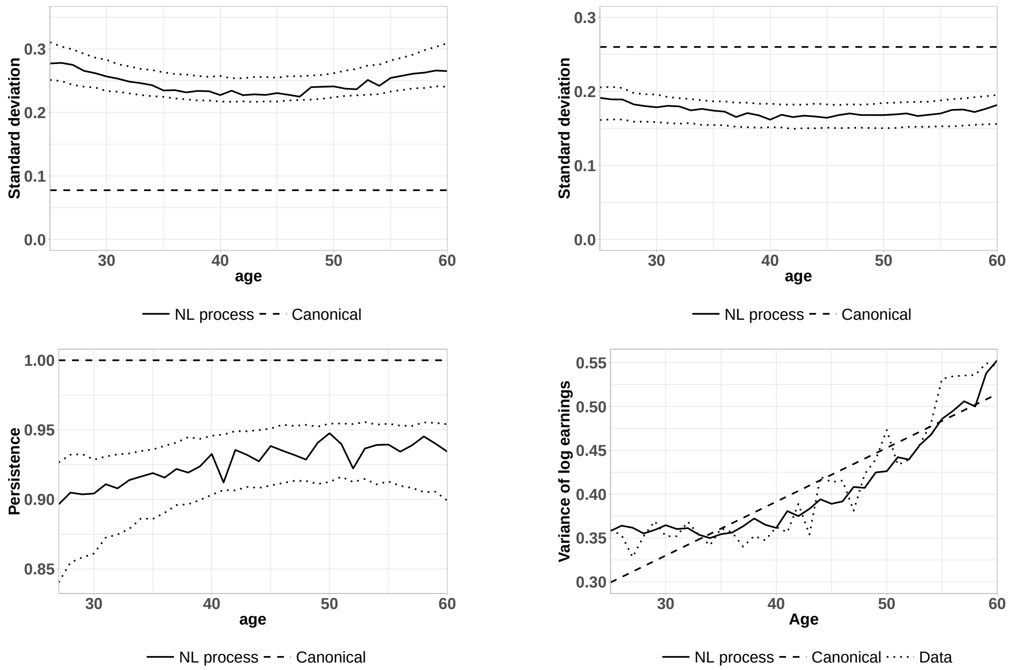

Figure 3: Age dependence of second moments: nonlinear vs canonical process. Top left, standard deviation of the innovation to the persistent component. Top right, standard deviation of the transitory shock. Bottom left, autocorrelation of the persistent component. Bottom right, cross-sectional variance of log earnings. The dotted bands represent bootstrapped 95% confidence intervals.

Starting from the **age-dependence** of second moments, the top row of Figure 3 plots the age profile of the standard deviations of the shocks to the persistent and transitory components of earnings. Both are age-independent by construction in the canonical process. The standard deviation of shocks to the persistent component is substantially higher for the nonlinear process and follows a U-shaped pattern by age. In contrast, the standard deviation of the transitory component of the nonlinear process displays little age variation and is lower in the nonlinear than in the canonical model. The bottom left panel of Figure 3 reports the age-profile of the first-order autocorrelation of the persistent earnings component for the two processes. In the nonlinear earnings process it is lower than in the canonical case for all ages, but it does increase between age 25 and 45. Given these differences, it is not surprising that the nonlinear process provides a substantially better fit of the age profile of the cross-sectional earnings dispersion, which we display in the bottom right panel of Figure 3[16]. More specifically, the canonical earnings process cannot capture the convex shape of the cross-sectional variance of earnings by age while the nonlinear process provides an extremely close fit, thanks to the combination of increasing persistence and declining variance of the persistent component over the ages 25 to 45. It is also apparent that the canonical model requires a low variance of the persistent shocks relative to the transitory ones to match the relatively low rate of growth of the cross-sectional variance of earnings over the life-cycle. Figure 4 displays more evidence on age-dependence, which also manifests itself in the skewness and kurtosis of the shocks.

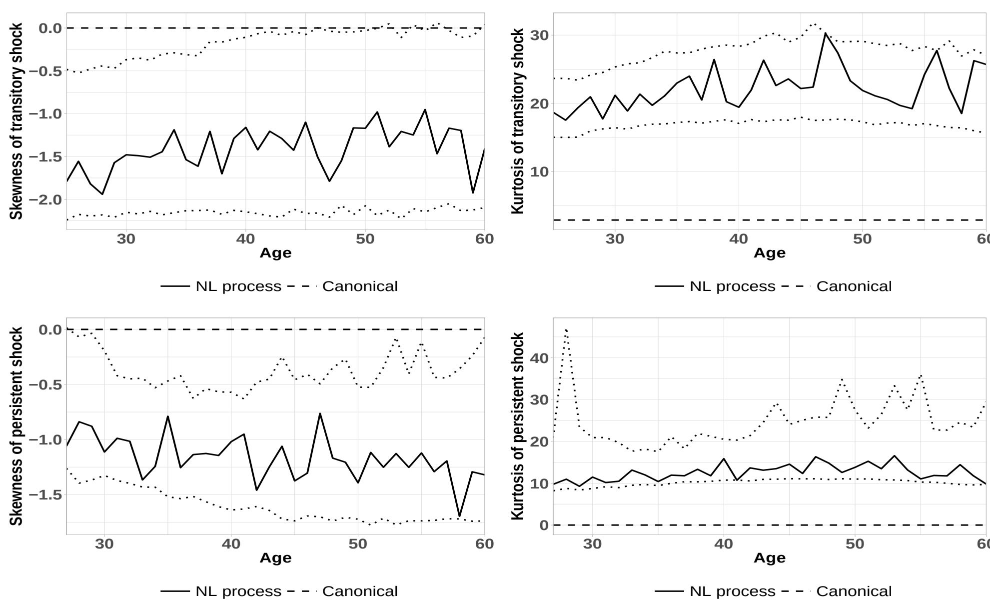

Figure 4: Skewness and kurtosis (by age) of the innovations to (a) the transitory component of earnings (top) and (b) the persistent component of earnings (bottom). The dotted bands represent bootstrapped 95% confidence intervals.

Turning to **non-normality**, Figure 4 reports skewness and kurtosis for the innovation to the transitory (top row) and persistent component of earnings (bottom row) by age and highlights that the earnings data display deviations from normality (the dashed line). They also highlight that non-normality, in particular kurtosis, is higher for transitory than persistent shocks.

[16]See Appendix B for details on the computation of this variance.

Turning to **nonlinearity**, Figure 5 plots the standard deviation of shocks to the innovation to the persistent component of earnings by previous earnings, while the right panel plots the persistence measure in equation (5)—namely the correlation between the percentile of η_{t-1} and of the innovation to it—averaged by age. In the right panel of this figure, we do not plot the persistence of the canonical model (which is constant at 1) for the sake of readability. These two panels clearly illustrate that the constant variance and persistence implied by the canonical process are strongly at odds with the highly nonlinear patterns in Figure 5 and the features of the observed data.

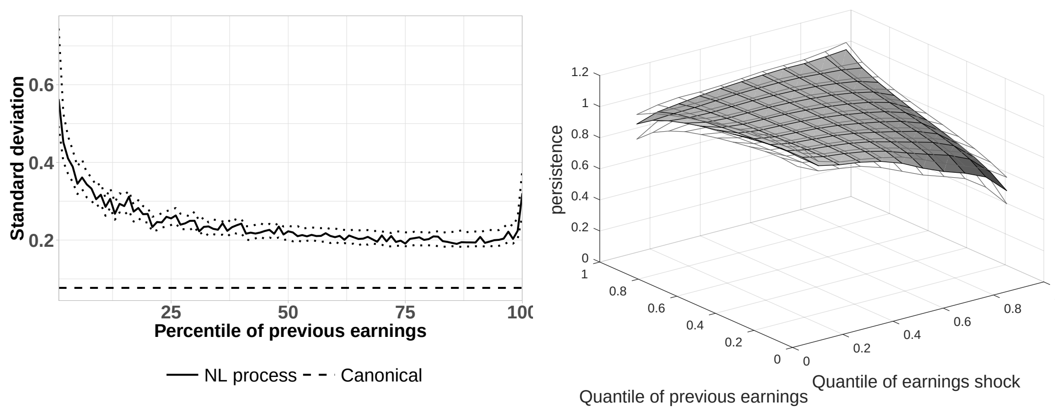

Figure 5: Standard deviations of persistent shocks by previous earnings (left) and nonlinear persistence of the persistent component, by quantile of previous earnings and quantile of shock received in the current period (right). The dotted bands and transparent mesh represent bootstrapped 95% confidence intervals.

4 The discretized nonlinear earnings process

To use the estimated process (1)-(3) in the life-cycle model, we discretize it using an age-dependent Markov chain.[17]

We start by simulating a large set of histories for the persistent and transitory component

[17]In theory, one could solve the model using a non-finite state space method, such as the Parameterized Expectations Algorithm (PAE) (e.g. den Haan and Marcet, 1990), simulating earnings histories using the estimated quantile functions directly. In practice, the computational costs of PAE are very large in life-cycle models with realistic lifetimes, given the non-stationarity of the policy functions. This likely explains why we are not aware of any paper using PAE to solve a life-cycle model.

of earnings. For each component in the simulated sample, we estimate a Markov chain of order one, with age-dependent state space $Z^t = \{\bar{z}_1, \ldots, \bar{z}_N\}, \; t = 1, \ldots, T$ and an age dependent transition matrices Π^t, of size $(N \times N)$. That is, we assume that the dimension N of the state space is constant across ages but we allow the set of states and the transition matrices to be age-dependent.

We determine the points of the state-space and the transition matrices at each age in the following way.

1. At each age, we order the realizations of each component by their size and we group them into N bins. Due to the limited sample size of the PSID, we want to strike a balance between a rich approximation of the earnings dynamics by earnings level (a large number of bins) and keeping the sample size in each bin sufficiently large. In our main specification we report the results for bins representing deciles, with the exception of the top and bottom deciles, that we split in 5. Thus, bins 1 to 5 and 14 to 18 include 2% of the agents at any given age, while bins $n = 6, \ldots, 13$ include 10% of the agents at any given age. This implies a total of 18 bins.

2. The points of the state space at each age t are chosen so that point z_t^n is the median in bin n at age t. Kennan (2006) proves that setting the gridpoint at the median of the bin (in the specific case of equally-sized bins) and attributing a weight of $1/N$ to each of the N bins constitutes the best discrete approximation of an arbitrary distribution.

3. The initial distribution at model age 0 is the empirical distribution at the first age we consider.

4. The elements π_{mn}^t of the transition matrix Π^t between age t and $t+1$ are the proportion of individuals in bin m at age t that are in bin n at age $t+1$.

Allowing for an age-dependent Markov chain allows to capture the non-constancy of moments of the earnings distribution over the life-cycle. The flexible form of the transition matrix allows to capture nonlinearities as a function of current earnings. The use of this kind of transition matrices is well established in the literature. Krueger and Perri (2003) use them to study the welfare consequences of an increase in earnings inequality. Studies of income mobility (e.g. Jäntti and Jenkins (2015)) and consumption mobility (e.g. Jappelli and Pistaferri (2006)) rely on them to analyze intra- or inter- generational mobility across relative rankings in the distributions. In this paper, instead, we are interested in capturing movements across earnings levels.

5 The model

The model is a partial-equilibrium, life-cycle, incomplete-markets model in the tradition of Bewley (1977). There is no aggregate uncertainty.

5.1 Demographics

Each year, a positive measure of agents is born. People start life as workers and work until retirement at age T^{ret}. The population grows at rate n.

An agent of age t faces a positive probability of dying $(1 - s_t)$ by the end of the period, where s_t denotes the one-period survival probability. Agents die with probability one by age T.

5.2 Preferences and endowments

Preferences are time separable, with a constant discount factor β. The intra-period utility is CRRA: $u(c_t) = c_t^{1-\sigma}/(1 - \sigma)$.

Agents are endowed with one indivisible unit of labor which they supply inelastically at zero disutility. Their earnings are subject to random shocks and follow the process described by equations (1)-(3).

5.3 Markets and the government

Asset markets are incomplete. Agents can borrow up to an exogenous borrowing limit $\underline{a}$ and can only invest in a risk-free asset at an exogenous rate of return r. There are no annuity markets to insure against mortality risk. As a result, there is a positive flow of accidental bequests in each period. We assume these are lost to the economy and thus are not received by any individual or the government.

Retired individuals receive an after-tax pension p from the government until they die. The pension is a function of the last earnings realization.

5.4 The household's problem

In any given period, a agent of age t chooses consumption c and risk-free asset holdings for the next period a', as a function of the relevant state vector. The optimal decision rules for consumption and savings solve the dynamic programming problems described below.

(i) Agents of working age $t < T_{ret}$ solve the recursive problem

$$V(t,z,\eta) = \max_{c,a'} \Big\{ u(c) + \beta s_t E_t V(t+1, z', \eta') \Big\} \tag{9}$$
$$\text{s.t. } a' = z - c, \quad a' \geq \underline{a},$$
$$z = (1+r)a + \eta + \varepsilon,$$

where z denotes total cash at hand.[18]

(ii) From the retirement age T^{ret} to the terminal age T agents no longer work and live off

[18] The choice of state vector avoids separately keeping track of the transitory component of earnings ε which is independently distributed over time.

their pension p and accumulated wealth. Their value function satisfies:

$$W(t,z,p) = \max_{c,a'} \Big\{ u(c) + s_t \beta W(t+1, z', p) \Big\} \tag{10}$$

$$\text{s.t. } a' = z - c, \quad a' \geq \underline{a},$$

$$z = (1+r)a + p.$$

The agent's pension p enters the state vector because it is a function of the agent's earnings pre-retirement. The terminal value function $W(T, a, p)$ is equal to zero (agents do not derive utility from bequests).

5.5 The model calibration

The model period is one year. Agents enter the labor market at age 25. The retirement and terminal ages are $T^{ret} = 60$ and $T = 85$. The population growth rate n is set to 1.2% per year. The survival probabilities s_t are from Bell, Wade and Goss (1992).

The coefficient of relative risk aversion is set to 2, a standard value. The risk-free rate is 4% and the discount factor β is calibrated to match a wealth to income ratio of 3.1. It equals 0.957 under the canonical earnings process and 0.939 under the nonlinear one.

We set the exogenous borrowing limit $\underline{a}$ to 12% of average disposable household earnings. This represents the average credit card limit in the SCF in 1989 and 1992, the two years within our sample period for which that information is available in the SCF.

As described in Section 2, our earnings processes are based on disposable earnings, hence we do not explicitly include taxation in the model.[19] In both cases, we impose the same average income profile, which we estimate from our PSID sample.

We discretize the two earnings processes as follows. In the case of the canonical earnings process, whose estimates we report in Table 1, we discretize the persistent component using the modified version of the Rouwenhorst method for non-stationary processes proposed by

[19] Appendix A provides more details about the earnings definition.

Fella et al. (2017). In the case of the nonlinear earnings process, we apply the procedure described in Section 4. For both cases, we use 18 gridpoints for the persistent component and 8 for the transitory component. In Appendix D we show that our results are robust to increasing the number of gridpoints and alternative discretization procedures.

The Social Security benefit p is a function of the last realization of disposable earnings $y_{ret} = \eta_{ret} + \varepsilon_{ret}$, which follows a fixed schedule of replacement rates. Namely, there is a 90% replacement rate for the part of y_{ret} below 18% of average earnings, of 32% for the fraction between 18% and 110% of average earnings, and 15 percent for the remainder. All benefits are then (very slightly) scaled up proportionately so that a worker that makes average earnings is entitled to a 45% replacement rate. The function is meant to mimic the US system and is based on Kaplan and Violante (2010).

6 Consumption, wealth, and welfare implications

This section studies the model's implications under the canonical and nonlinear earnings processes and compares them to U.S. consumption data. To do so, we first analyze the growth in consumption dispersion over the working life and then turn to measuring self-insurance insurance as proposed by Blundell et al. (2008). Finally, we compare the implications of these earnings processes for wealth inequality and welfare.

6.1 Consumption inequality over the working life

We start by studying the evolution of cross-sectional consumption dispersion over the life cycle. Following Deaton and Paxson (1994) and Storesletten, Telmer and Yaron (2004*b*), it is common to interpret its growth, relative to the growth of cross-sectional earnings dispersion, as a measure of risk sharing. For reference, Figure 6 plots the cross-sectional dispersion of consumption and earnings over the life cycle. The (solid) earnings variance profile is from our PSID sample, while the dashed line plots the variance profile of nondurable consumption

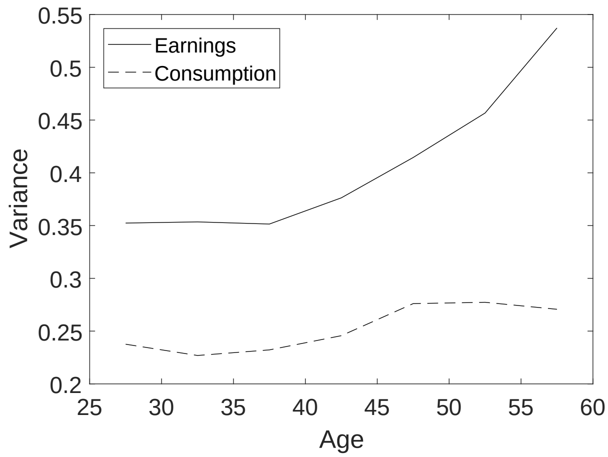

Figure 6: Cross-sectional variance of log earnings and log consumption by age. See Appendix B for details on their computation.

is from the CEX 1980-2007 as in Heathcote, Perri and Violante (2010). Given the relatively small sample size, we group observations in 5-year age groups. As it is well known, both earnings and consumption inequality increase over the working life, but earnings inequality increases substantially faster than consumption inequality from about age 40.

Because the increase in consumption inequality over the working life is informative about people's ability to insure against earnings risk, it provides a useful benchmark against which to assess the ability of the model to capture the degree of insurability of earning shocks in the data. Figure 7 reports the age profile of cross-sectional consumption dispersion for both the CEX data and the model under, respectively, the canonical and nonlinear earnings processes.[20]. The series are normalized so that each starts at zero at age 27, which is the midpoint of the first 5-year age group (25–29). The canonical earnings process fails to match both the overall growth and the shape of the profile of consumption dispersion. Its overall growth rate is more than double that in the data and its profile is monotonic and roughly linear. Conversely, in the data, the age profile of consumption is significantly convex between age 25 and 47. The nonlinear process, instead, matches well both the overall growth in consumption dispersion and its convexity in the first part of the life cycle. The one part

[20]We perform this comparison recalibrating beta so as to keep the wealth to income ratio constant across earnings processes. Appendix D.5 shows that the profiles are very similar if we keep the discount factor constant across processes instead.

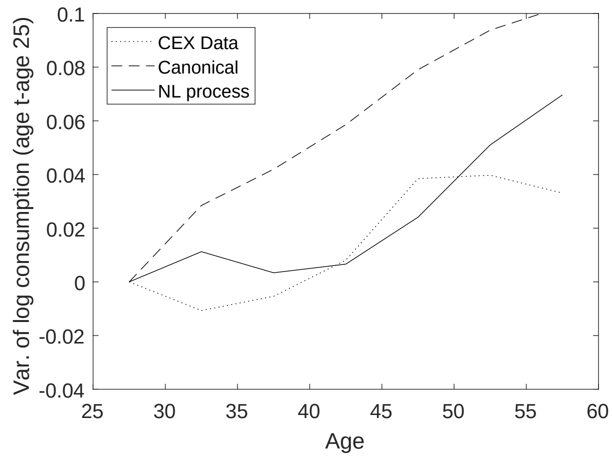

Figure 7: Growth in the cross-sectional variance of log consumption, data and implications of two earnings processes.

that it misses is the flattening out after age 47.

The finding that the estimated richer earnings processes implies a profile of consumption dispersion in line with the data is remarkable. Standard models with linear earnings processes (see Storesletten et al. (2004*a*)) generate a profile similar to the one implied by the canonical earnings process in Figure 7, and thus overstate the rate of growth of consumption dispersion, unless the process for earnings has an idiosyncratic deterministic time trend, or Heterogeneous Income Profile (Guvenen, 2007; Primiceri and Van Rens, 2009). Intuitively, heterogeneity in individual, life-cycle trend growth generates a substantially smaller rise in consumption dispersion because the individual-specific trend growth is known to consumers but not to the econometrician. Huggett et al. (2011) show that heterogeneity in earnings growth rates can be also generated by the endogenous response of human-capital investment over the life cycle to heterogeneity in initial human capital levels.

Our findings suggest a novel explanation: the age profile of cross-sectional consumption dispersion can be generated by the response of saving to the richer earnings dynamics that we consider, without resorting to heterogeneity in income profiles. It should also be noted that allowing for heterogeneity in income profiles cannot generate (cfr. Guvenen, 2007; Primiceri and Van Rens, 2009) the strong convexity that characterizes the consumption data (dotted line in Figure 7).

As we have discussed in Section 3.1, our rich earnings process deviates from the canonical linear process along three main dimensions: (1) age-dependence, (2) non-normality, and (3) nonlinearity. To understand the contribution of each of these factors to the growth of consumption dispersion over the life cycle, we conduct a series of counterfactual experiments, simulating the model under progressively richer stochastic processes for earnings.

We start by restricting the functional form of the earnings process to be the sum of an AR(1) plus a white noise component, as in the canonical process, but allowing for both age-dependent persistence and variance of shocks (as in Karahan and Ozkan (2013)), as well as non-normality of their distributions. Compared to the fully general nonlinear earnings process, this one imposes linearity in $\eta_{i,t-1}$; namely, that persistence and other second and higher conditional moments are independent of $\eta_{i,t-1}$. We estimate this process on our PSID data, following the procedure described in Section 3.1 for the nonlinear process, but restricting the quantile function for the persistent component in equation (4) to be linear in its past value.

To further disentangle the effect of the age dependence of persistence and variance from that of non-normality, we perform two set of simulations using the restricted estimates that we have just described. In the first one, we simulate earnings using the estimated persistence and variances but drawing shocks from a normal distribution. In the second experiment, we simulate earnings using the estimated distribution (i.e. quantile function), that also allows for non-normality. We discretize each of the resulting processes using the method in Section 4. The recalibrated value of the discount factor equals 0.939 in the economy with normal shocks and 0.940 in the other one.

Figure 8 plots the cross-sectional variance profiles reported in Figure 7, with the addition of the two profiles implied by (a) only age-dependence and (b) age-dependence together with non-normality.

The solid line marked with circles in Figure 8 corresponds to the case of an age-dependent

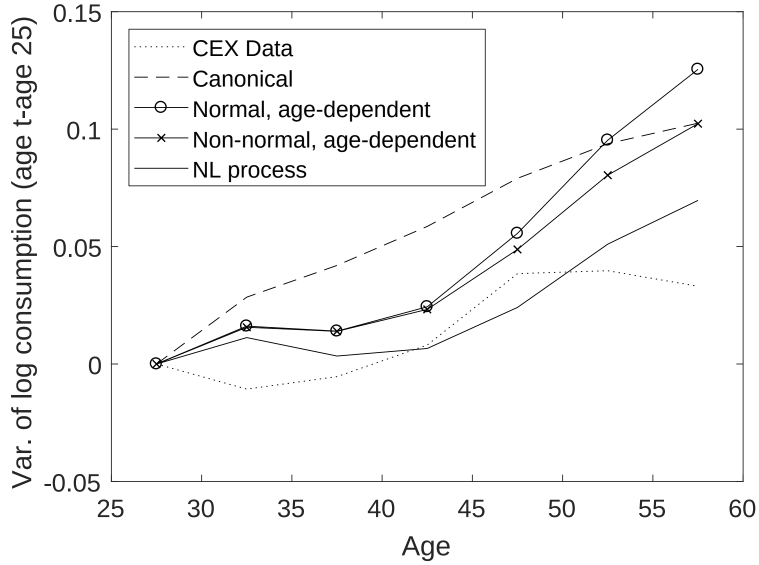

Figure 8: Growth in the cross-sectional variance of log consumption. Contribution of (1) age-dependence, (2) non-normality and (3) nonlinearity to the growth of the cross-sectional variance of log consumption.

linear process with normal innovations. Compared to the canonical case, allowing for age dependence substantially improves the fit of consumption dispersion in the first part of the life cycle, but counterfactually implies an even larger growth rate of consumption dispersion from age 43 onwards. The net effect for the age-dependent earnings process is an overall rate of growth in consumption dispersion between ages 25 and 60 that is three percentage points higher than in the canonical case.

The intuition behind the above finding is the following. Allowing for age-dependence implies that the estimated process for earnings matches the age-profile of the cross-sectional earnings variance in the bottom right panel of Figure 3; namely, relatively flat until age 43 but growing at a rate substantially above its working-life average afterwards. The forces underpinning this pattern are: (a) the U-shaped profile of the variance of the persistent component of earnings; and (b) a persistence below one that increases until age 45 but flattens out afterwards (see Figure 3). Compared to the canonical process with a unit root and constant shock variance, the interaction of these two forces implies that self-insurance through precautionary saving is more effective and, as a consequence, the growth in consumption dispersion is lower until middle age. In the second half of the working life, though, the increase in the variance of the persistent earnings shocks reduces the ability to self-insure and results in a substantial increase in consumption dispersion. This is confirmed by comparing the age profile of average wealth reported in the left panel of Figure 9 under the canonical (dashed curve) and age-dependent earnings process with normal shocks (solid curve with circles). Though the aggregate wealth-to-earnings ratio is the same in the two economies, average saving is higher before and lower after age 50 in the economy with age-dependent earnings process.

We now turn to the linear process with the same (age-dependent) first and second moments as above but with non-normal innovations. The solid line marked with crosses in Figure 8 plots the associated age profile of variance. Compared to the normal case, the rate of growth of the consumption variance is everywhere lower. The difference is particularly pronounced towards the end of the working life. To understand the mechanism at work, it is important to understand the impact of negative skewness and kurtosis on precautionary saving and the wealth distribution. Civale et al. (2016) study the issue in an Aiyagari economy and show that, everything else equal, negative skewness reduces both the cross-sectional mean and dispersion of wealth while kurtosis increases both.

The effect of higher kurtosis is in line with intuition. By increasing the probability of tail events higher kurtosis increases precautionary saving for all agents and therefore the mean and variance of the wealth distribution. The effect of negative skewness, though, is less intuitive. Basically, for a distribution to have higher negative skewness keeping the other moments constant, some probability mass has to move towards the top of the distribution. Wealthy agents are not sensitive to left skewness but, confronted with a higher probability of positive shocks, save less. Conversely, agents who are close to the borrowing constraint are more sensitive to skewness than to the higher probability of positive shocks and save more. In the aggregate, the response of wealth-rich agents dominates that of the wealth-poor and average wealth falls. More intuitively, so does the variance of wealth holdings. Comparing

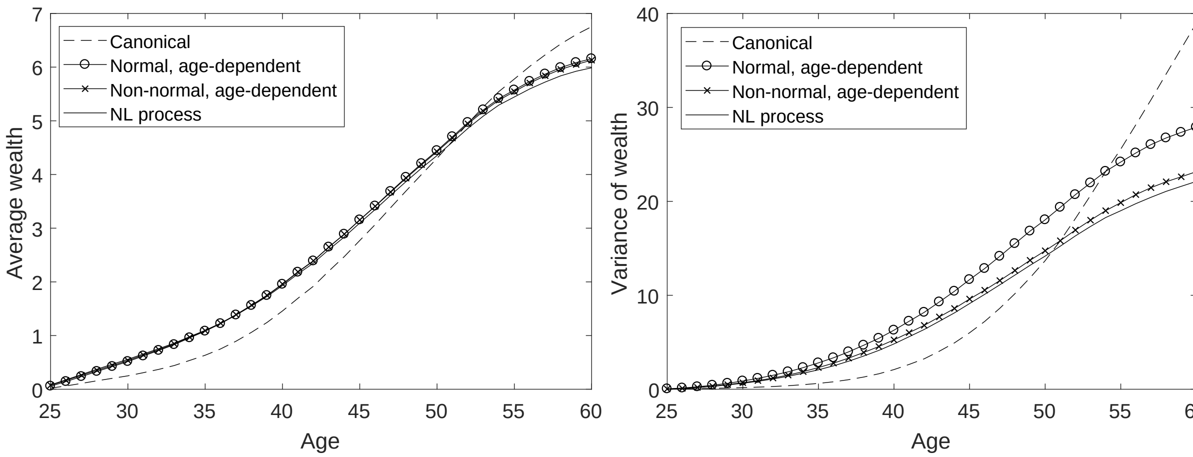

Figure 9: Cross-sectional average wealth (left) and variance of wealth holdings (right), by age and earnings process.

the lines marked with circles and crosses in Figure 9 reveals that, in our model, the net effect of negative skewness and kurtosis hardly affects the life-cycle profile of average wealth (left panel), but substantially reduces the rate of growth of the variance of wealth holdings (right panel) compared to the case with normal shocks. This fall in wealth dispersion accounts for the fall in consumption dispersion in Figure 8 when skewness and kurtosis of shocks are introduced.

Finally, comparing the line marked with crosses and the solid line in Figure 8 shows that allowing for nonlinearity brings the overall fit of life-cycle inequality substantially closer to the data, compared to all of the other earnings processes considered. Figure 10 provides some insight into the mechanism associated with the nonlinearity in earnings. It plots the persistence (averaged over age) of the persistent earnings component by previous earnings and current shock for both the age-dependent non-normal case (light surface) and the nonlinear one (dark surface). By assumption persistence is constant in the former case. For individuals with previous earnings realizations below the median, negative shocks (below the median) increase persistence relative to the linear case, while positive shocks reduce it. This implies that good shocks partially wipe out the memory of previous bad realizations while bad shocks reinforce it. The average persistence of a bad previous realization is hardly affected

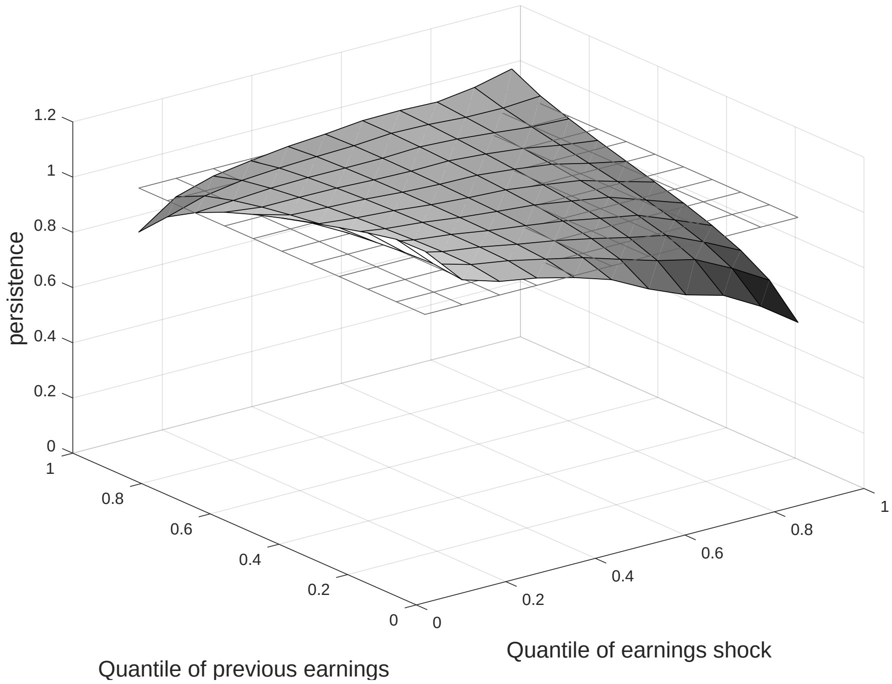

Figure 10: Persistence of the persistent component by quantile of previous earnings and current shock: NL (dark surface) vs non-normal age-dependent process (transparent surface).

but, since the precautionary motive implies that saving responds more to downside than upside risk, individuals with bad earnings realizations save more in the nonlinear case. The nonlinearity is also present, though much less pronounced, for individuals with previous earnings realizations in the top two deciles. For these, shocks below the second decile reduce persistence, while shocks above it increase it, relative to the linear case. The increase in average persistence, relative to the linear case, tends to reduce the saving response. On the other hand, the fact that bad shocks reduce earnings more than linearly (reduce persistence) increases precautionary saving. Overall, saving increases for agents with adverse realizations of previous persistent earnings, while it falls for individuals with good previous earnings realization. This is reflected in the flattening in the age-profile of the variance of wealth holdings in the right panel of Figure 9. This fall in wealth dispersion accounts for the further fall in the growth of consumption dispersion over the life cycle which brings it much closer to the data, particularly for ages up to 50 (see Figure 8).

None of our earnings processes captures the flattening out in the variance of consumption that we measure after age 47 because the variance of earnings in the data keeps increasing. Our structural model misses two aspects of the data that could be important in this regard. The first one is early retirement. For retirees, income is mainly composed of Social Security payments and does not vary much. Thus, consumption is no longer exposed to earnings fluctuations and medical expense risk is not very high until well into retirement age, as shown by De Nardi, French and Jones (2010). The second one is the role of durables and housing, that become substantial by that age and might affect both measured consumption (we only look at nondurable consumption) and one's ability to self-insure.

6.2 Measuring self-insurance against earnings shocks

An alternative, and possibly more intuitive, measure of self-insurance is related to the extent of pass through from shocks to disposable earnings onto consumption. Blundell et al. (2008) propose estimating consumption insurance coefficients on persistent and transitory earning shocks by positing the following equation

$$\Delta c_{it} = (1-\psi^p)\nu_{it} + (1-\psi^{tr})\varepsilon_{it} + \xi_{it}, \tag{11}$$

where $\nu_{it} = \eta_{it} - E[\eta_{it}|t, \eta_{i,t-1}]$ denotes the innovation to the persistent component of earnings and ε_{it} the transitory component. The insurance coefficients with respect to persistent (ψ^p) and transitory (ψ^{tr}) shocks

$$\psi^p = 1 - \frac{\mathrm{cov}(\Delta c_{it}, \nu_{it})}{\mathrm{var}(\nu_{it})}, \quad \psi^{tr} = 1 - \frac{\mathrm{cov}(\Delta c_{it}, \varepsilon_{it})}{\mathrm{var}(\varepsilon_{it})} \tag{12}$$

capture the fraction of the variance of either type of shock that does not translate into movements in consumption. Similarly, one can compute age-specific insurance coefficients ψ_t^p, ψ_t^{tr} where moments are computed only over agents of age t.

To compute the insurance coefficients implied by our model, we simulate a panel of working lives under both the benchmark and nonlinear processes and compute the associated consumption c_{it} and insurance coefficients in equation (12) on the simulated data.

Computing the coefficients in equation (12) within the model is straightforward since the shocks are observable. In contrast, estimating them from the data requires identifying the two types of earning shocks at the individual level. Blundell et al. (2008) propose an identification strategy under the assumption that earnings follow the canonical linear process (1)-(3). The estimators for the insurance coefficients based on the BPP methodology are given by

$$\psi^{p}_{BPP} = 1 - \frac{\text{cov}(\Delta c_{it}, y_{i,t+1} - y_{i.t-2})}{\text{cov}(\Delta y_{it}, y_{i,t+1} - y_{i,t-2})}, \; \psi^{tr}_{BPP} = 1 - \frac{\text{cov}(\Delta c_{it}, \Delta y_{i,t+1})}{\text{cov}(\Delta y_{i,t}, \Delta y_{i,t+1})}. \tag{13}$$

As pointed out by Kaplan and Violante (2010), comparing the coefficients in equation (13) estimated within the model to the estimates in Blundell et al. (2008) conveys information on the degree of shock insurability in the model relative to the data.

The coefficients in equation (13), though, may provide biased estimates of the true coefficients in equation (12) to the extent that the identification assumption on which they are based is violated. The assumption can be violated for two reasons. First, if earnings do not follow the canonical linear process in equation (1)-(3). This is obviously true in the more flexible cases we consider. Second, as pointed out by Kaplan and Violante (2010), even if earnings follow a canonical linear process the ψ^{p}_{BPP} estimator may be biased whenever consumption does not equal permanent income, as is the case in the presence of a precautionary saving motive.[21] For this reason, we compute both types of coefficients. Table 2 reports their values under the alternative income processes.

Columns 1 and 2 in Table 2 report the coefficients in equation (13). As a reference, the first row reports the estimates by Blundell et al. (2008)—respectively 0.36 for permanent and 0.95 for transitory shocks—on the PSID using similar data to ours.[22] The corresponding

[21] Formally, the bias is present whenever present consumption responds to past persistent income changes, which implies that $\text{cov}(\Delta c_{it}, y_{i,t+1} - y_{i.t-2})$ is a biased estimator of $\text{cov}(\Delta c_{it}, \nu_{it})$. Kaplan and Violante (2010) show that this is indeed the case in a life-cycle model similar to ours with a canonical earnings process and occasionally-binding borrowing constraints.

[22] Blundell et al. (2008) conduct their analysis using disposable household earnings for continuously married coupled headed by a male head.

Process/Coefficients	ψ^{p}_{BPP}	ψ^{tr}_{BPP}	ψ^{p}	ψ^{tr}
	Data: BPP (2008)			
Canonical (S.E. in parenthesis)	0.36	0.95	–	–
	(0.09)	(0.04)		
	Model			
Canonical	0.13	0.89	0.31	0.92
Nonlinear process	0.43	0.82	0.46	0.91
Normal, age-dependent	0.42	0.83	0.46	0.88
Non-normal, age-dependent	0.42	0.83	0.46	0.88

Table 2: Insurance coefficients

values for the model, when earnings follow the canonical earnings process, are 0.13 and 0.89, which confirms the finding by Kaplan and Violante (2010) that the extent of self-insurance of permanent earnings shocks implied by the model is substantially lower than the degree of insurance in the data. On the other hand, the estimates for the model with a nonlinear earnings process imply an insurance coefficient for persistent shocks of 0.43 which is substantially more in line with, and even marginally larger than, the BPP estimate in the first row.

From a qualitative perspective, this result is very much in line with our findings in Section 6 that agents are more able to self-insure against income fluctuations when earnings follow the nonlinear process than in the canonical case. Interestingly, our finding that allowing for a richer earnings process implies a substantially different estimate of the insurance coefficient for persistent shocks is confined to *disposable* household earnings. Using the same earnings process we use here, Arellano et al. (2017) estimate an average insurance coefficient for persistent shocks to *pre-tax* household earnings between 0.6 and 0.7 which is in line with an estimate of 0.69 in Blundell et al. (2008) under the identifying assumption that earnings follow the canonical process. As discussed in Blundell et al. (2008), the nearly double magnitude of the insurance coefficients with respect to *pre-tax* rather than *disposable* earnings is due to the of insurance implied by the tax and transfer system.

Turning to the insurance coefficient for transitory shocks in column 2, it may seem surprising that it is higher under the canonical than under the nonlinear earnings process. As pointed out in Kaplan and Violante (2010), though, the intuition is that the increased insurability of persistent shocks induces households to shift the use of savings from the smoothing of transitory shocks to the smoothing of persistent shocks.

Columns 3 and 4 in Table 2 report the estimates of the *true* insurance coefficients in equation (12) within the model. Comparing them to the BPP estimates in columns 1 and 2 reveals that the downward bias of the insurance coefficient for persistent shocks implied by the BPP procedure is sizeable (0.13 against 0.31) in the case of the canonical income process but small (0.43 against 0.46) for the nonlinear process. The intuition is that, as pointed out by Kaplan and Violante (2010), the bias is exacerbated in an economy in which the borrowing constraint is occasionally binding. As discussed above, when earnings follow the nonlinear process shocks are more insurable, and precautionary saving larger. For this reason, the economy spends less time close to the borrowing constraint and the bias is lower.

Finally, the last two lines reports the same coefficient for the case with age dependence and normal shocks and the one that also allows for non-normality. Comparing the three set of estimates reveals that the feature that drives the better match of the insurance coefficient for persistent shock estimated by BPP is essentially the age dependence of the earnings process. This is consistent with the finding in Karahan and Ozkan (2013) that the (true) insurance coefficient for persistent shocks in a life-cycle economy with an age-dependent earnings process with normal shocks is 0.38.[23]

While Table 2 reports the average insurance coefficients, Figure 11 plots the true insurance coefficient for persistent shocks ψ_t^p at each age. The coefficients are increasing with age, as: (a) wealth is accumulated; and (b) the fall in the residual working life reduces the effective shock persistence. The degree of insurability at all ages but the last working age is substantially higher under the nonlinear earnings process than under the canonical one.

[23]The earnings process used by Karahan and Ozkan (2013) is similar to our age-dependent process with normal shocks. Their estimate of 0.38 for the true coefficient ψ^p is in the ballpark of our estimate of 0.46 in Table 2.

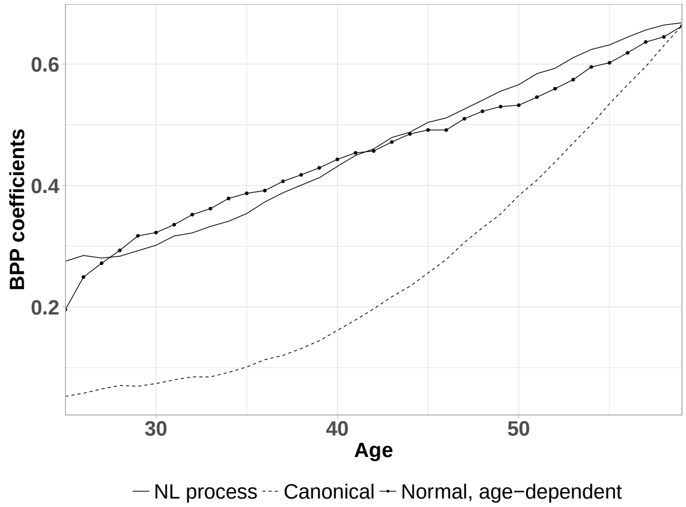

Figure 11: Partial insurance coefficients on persistent shocks, ψ_t^p, by age

	Wealth Gini	Percentage wealth in the top 1%	5%	20%	40%	60%	80%
Data (SCF 1989)	.79	30	54	81	94	99	100
Model: Canonical	.64	9	28	65	87	97	100.1
Model: Nonlinear process	.61	7	25	62	86	96	99.8

Table 3: Wealth distribution

For the same reason, the age profile of the coefficients is substantially flatter in the former case. In line with the discussion above, most of the difference is due to the age-dependence of earnings. It is only from age 40 onwards that the coefficients are marginally higher under the nonlinear process than under the age-dependent earnings process with normal shocks.

6.3 Wealth

Table 3 compares the implied wealth distribution of the canonical and nonlinear earnings processes with data from the U.S 1989 SCF (Kuhn and Ríos-Rull, 2015).

As known in the literature (see Quadrini and Ríos-Rull (2014), Cagetti and De Nardi (2008), and De Nardi and Fella (2017)), the model with a canonical earnings process is unable to generate the substantial level of wealth concentration that we observe in the data. For instance, the top 1% of agents holds about 30% of total wealth in the data, while the corresponding share is only holds 9% in the model. Comparing the second and third rows in the table reveals that allowing for nonlinear earnings does not improve the fit of the wealth the distribution. If anything it marginally reduces the degree of wealth concentration at the top.[24] One may be concerned that this may be due to the nature of the PSID data, which top-codes earnings and does not oversample the rich. However, De Nardi et al. (2016) conduct a similar exercise using synthetically-generated W2 Social Security Administration tax data, and find similar results for the concentration of wealth at the top. As pointed out by De Nardi and Fella (2017), non-homothetic preferences for bequests, entrepreneurship, and medical-expense risk are important for life-cycle models to be able to account for top wealth concentration.

6.4 Welfare

The differences in the evolution of the variance of log consumption and the pass-through of income shocks to consumption show that income risk affects households in a different way in the two economies. A natural question is to which extent these differences affect welfare.

To measure welfare, Table 4 displays the constant fraction of consumption that households are willing to give up to live in a world with no income uncertainty; i.e., a world where earnings are equal to the common and deterministic average earnings profile. We compute this measure under the veil of ignorance (before people enter the labor market and draw the first earnings realization) and, for comparability, we keep the discount factor the same for both processes fixing it at its calibrated value for the nonlinear process.

The nonlinear process features larger variance and lower persistence of persistent shocks that, as we have discussed in Section 6.1, improve shock insurability, but also negative skewness and high kurtosis, as well as nonlinear persistence. Vice versa in the canonical

[24] We target a wealth to income ratio of 3.1, but this has little effect on wealth concentration.

	Welfare cost
Canonical process	28.2%
Nonlinear process	26.1%
Normal, age-dependent	24.3%
Non-normal, age-dependent	25.4%

Table 4: Consumption measure of welfare costs

model, the lower variance of shocks at all ages after the first one is counteracted by their high persistence (unit root) and the higher variance of the initial condition. The net effect of all these forces is that overall risk is higher under the canonical process. In particular, households would be willing to give up 28.2 per cent of their consumption in every state to eliminate earnings risk under the canonical earnings process compared to 26.1 per cent under the nonlinear earnings process.

In order to understand the respective contribution of the various features of our rich earnings process we have also computed the welfare cost under the two *intermediate* earnings processes considered in sections 6.1 and 6.2. The results are reported the third and fourth row of Table 4. They show that the lower welfare costs of earnings risk relative to the canonical process are all due to the age-dependence of second moments. Allowing only for age-dependence reduces the welfare costs of earnings risk by 4 percentage points, from 28.2 to 24.3 per cent, relative to the canonical process. Introducing, non-normality lowers welfare by one percentage point, relative to the normal age-dependent case, while allowing also for non-linearity reduces it by an additional percentage point.

7 Conclusions

We estimate a richer stochastic process for *household disposable earnings* featuring a transitory and persistent component and allowing for age-dependence, non-normality and nonlinearity. We use a standard life-cycle model with incomplete markets to compare the implications of our richer process to those of canonical permanent/transitory linear process with age-independent, normally-distributed shocks. Our main findings are as following. Compared to the canonical process, the richer process implies a much better fit of the growth in cross-sectional consumption dispersion over the life cycle and a degree of self-insurance of persistent earnings shocks in line with the empirical estimates in Blundell et al. (2008). It also implies smaller welfare costs of earnings fluctuations. In terms of wealth inequality, we find that the two earnings processes have similar implications, including at the upper tail of the wealth distribution.

References

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

Aiyagari, S. R. (1994), 'Uninsured idiosyncratic risk, and aggregate saving', *Quarterly Journal of Economics* **109**(3), 659–684.

Altonji, Joseph G., Smith, Anthony A. and Vidangos, Ivan (2013), 'Modeling earnings dynamics', *Econometrica* **81**(4), 1395–1454.

Arellano, Manuel, Blundell, Richard and Bonhomme, Stéphane (2017), 'Earnings and consumption dynamics: A non-linear panel data framework'. Econometrica, forthcoming.

Bell, Felicitie C., Wade, Alice H. and Goss, Stephen C. (1992), 'Life tables for the United States Social Security Area: 1900-2080', (Social Security Administration, Office of the Actuary).

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

Blundell, Richard, Graber, Michael and Mogstad, Magne (2015), 'Labor income dynamics and the insurance from taxes, transfers and the family', *Journal of Public Economics* **127**, 58–73.

Blundell, Richard, Pistaferri, Luigi and Preston, Ian (2008), 'Consumption inequality and partial insurance', *The American Economic Review* pp. 1887–1921.

Blundell, Richard, Pistaferri, Luigi and Saporta-Eksten, Itay (2016), 'Consumption inequality and family labor supply', *The American Economic Review* **106**(2), 387–435.

Bonhomme, Stéphane and Robin, Jean-Marc (2009), 'Assessing the equalizing force of mobility using short panels: France, 1990–2000', *The Review of Economic Studies* **76**(1), 63–92.

Browning, Martin, Ejrnaes, Mette and Álvarez, Javier (2010), 'Modelling income processes with lots of heterogeneity', *Review of Economic Studies* **77**, 1353–1381.

Cagetti, Marco and De Nardi, Mariacristina (2008), 'Wealth inequality: Data and models', *Macroeconomic Dynamics* **12**(S2), 285–313.

Cagetti, Marco and De Nardi, Mariacristina (2009), 'Estate taxation, entrepreneurship, and wealth', *American Economic Review* **99**(1), 85–111.

Castañeda, Ana, Díaz-Giménez, Javier and Ríos-Rull, José-Víctor (2003), 'Accounting for the U.S. Earnings and Wealth Inequality', *Journal of Political Economy* **111**(4), pp. 818–857.

Civale, Simone, Díez-Catalán, Luis and Fazilet, Fatih (2016), 'Discretizing a process with non-zero skewness and high kurtosis', Mimeo, University of Minnesota.

Conesa, Juan Carlos, Kitao, Sagiri and Krueger, Dirk (2009), 'Taxing capital? not a bad idea after all!', *American Economic Review* **99**, 25–48.

Deaton, Angus and Paxson, Christina (1994), 'Intertemporal choice and inequality', *Journal of Political Economy* **102**(3), 437–67.

den Haan, Wouter J and Marcet, Albert (1990), 'Solving the Stochastic Growth Model by Parameterizing Expectations', *Journal of Business and Economic Statistics* **8**(1), 31–34.

De Nardi, Mariacristina (2004), 'Wealth inequality and intergenerational links', *Review of Economic Studies* **71**, 743–768.

De Nardi, Mariacristina and Fella, Giulio (2017), 'Saving and wealth inequality', *Review of Economic Dynamics* **26**, 280 – 300.

De Nardi, Mariacristina, Fella, Giulio and Paz-Pardo, Gonzalo (2016), 'The implications of richer earnings dynamics for consumption and wealth', NBER Working paper no. 21917. National Bureau of Economic Research.

De Nardi, Mariacristina, French, Eric and Jones, John B. (2010), 'Why do the elderly save? The role of medical expenses', *Journal of Political Economy* **118**(1), 39–75.

Fella, Giulio, Gallipoli, Giovanni and Pan, Jutong (2017), 'Markov-chain approximations for life-cycle models'.

Geweke, John and Keane, Michael (2000), 'An empirical analysis of earnings dynamics among men in the psid: 1968-1989', *Journal of Econometrics* **96**, 293–356.

Golosov, Mikhail, Troshkin, Maxim and Tsyvinski, Aleh (2016), 'Redistribution and social insurance', *American Economic Review* **106**(2), 359–86.

Graber, Michael and Lise, Jeremy (2015), 'Labor market frictions, human capital accumulation and consumption inequality', mimeo, University College London.

Guvenen, Fatih (2007), 'Learning your earning: Are labor income shocks really very persistent?', *The American Economic Review* **97**(3), 3.

Guvenen, Fatih and Smith, Anthony A (2014), 'Inferring labor income risk and partial insurance from economic choices', *Econometrica* **82**(6), 2085–2129.

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.

Haider, S. J. (2001), 'Earnings instability and earnings inequality of males in the united states: 1967-1991', *Journal of Labor Economics* **19**(4), 799–836.

Heathcote, Jonathan, Perri, Fabrizio and Violante, Giovanni L (2010), 'Unequal we stand: An empirical analysis of economic inequality in the united states, 1967-2006', *Review of Economic dynamics* **13**(1), 15–51.

Heathcote, Jonathan, Storesletten, Kjetil and Violante, Giovanni L. (2010), 'The macroeconomic implications of rising wage inequality in the united states', *Journal of Political Economy* **118**(4), 681–722.

Huggett, Mark, Ventura, Gustavo and Yaron, Amir (2011), 'Sources of lifetime inequality', *American Economic Review* **101**(7), 2923–54.

Jäntti, Markus and Jenkins, Stephen P. (2015), Income mobility, *in* Anthony B.Atkinson and FrançoisBourguignon, eds, 'Handbook of Income Distribution', Vol. 2 of *Handbook of Income Distribution*, Elsevier, pp. 807–935.

Jappelli, Tullio and Pistaferri, Luigi (2006), 'Intertemporal choice and consumption mobility', *Journal of the European Economic Association* **4**(1), 75–115.

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

Kaplan, Greg and Violante, Giovanni L. (2010), 'How much consumption insurance beyond self-insurance?', *American Economic Journal: Macroeconomics* **2**(4), 53–87.

Karahan, Faith and Ozkan, Serdar (2013), 'On the persistence of income shocks over the life cycle: Evidence, theory and implications', *Review of Economic Dynamics* **16**(3), 452–476.

Kennan, John (2006), 'A note on discrete approximations of continuous distributions', *University of Wisconsin Madison* pp. 211–251.

Krueger, Dirk and Perri, Fabrizio (2003), 'On the welfare consequences of the increase in inequality in the united states', *NBER macroeconomics annual* **18**, 83–121.

Krueger, Dirk and Perri, Fabrizio (2006), 'Does income inequality lead to consumption inequality? evidence and theory', *The Review of Economic Studies* **73**(1), 163–193.

Krueger, Dirk, Mitman, Kurt and Perri, Fabrizio (2016), 'Macroeconomics and household heterogeneity', *Handbook of Macroeconomics* **2**, 843–921.

Kuhn, Moritz and Ríos-Rull, José-Víctor (2015), '2013 update on the us earnings, income and wealth distributional facts: a view from macroeconomics', Working paper.

Lillard, Lee A. and Willis, Robert J. (1978), 'Dynamic aspects of earning mobility', *Econometrica* **46**(5), 985–1012.

McKay, Alisdair (2017), 'Time-varying idiosyncratic risk and aggregate consumption dynamics', *Journal of Monetary Economics* **88**, 1 – 14.

Meghir, Costas and Pistaferri, Luigi (2004), 'Income variance dynamics and heterogeneity', *Econometrica* **72**(1), 1–32.

Meghir, Costas and Pistaferri, Luigi (2011), 'Earnings, consumption and life cycle choices', *Handbook of Labor Economics* **4**, 773–854.

Primiceri, Giorgio E and Van Rens, Thijs (2009), 'Heterogeneous life-cycle profiles, income risk and consumption inequality', *Journal of Monetary Economics* **56**(1), 20–39.

Quadrini, Vincenzo and Ríos-Rull, José-Víctor (2014), Inequality in macroeconomics, *in* A.B.Atkinson and F.Bourguignon, eds, 'Handbook of Income Distribution, vol. 2B', Elsevier, chapter 15.

Rouwenhorst, Geert (1995), Asset pricing implications of equilibrium business cycle models, *in* Thomas F.Cooley, ed., 'Frontiers of business cycle research', Princeton University Press, chapter 10.

Scholz, John Karl, Seshadri, Ananth and Khitatrakun, Surachai (2006), 'Are americans saving optimally for retirement?', *Journal of Political Economy* **114**, 607–643.

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

Storesletten, Kjietil, Telmer, Chris C. and Yaron, Amir (2004*b*), 'Cyclical dynamics in idiosyncratic labor-market risk', *Journal of Political Economy* **112**(3), 695–717.

Tauchen, George (1986), 'Finite state markov-chain approximations to univariate and vector autoregressions', *Economic Letters* **20**(2), 177–181.

A Appendix: PSID data

We use PSID data to estimate both the canonical and the nonlinear earnings processes. In this Appendix we briefly describe the PSID, our sample selection criterion, the precise variable definition we use and the details of the estimation of the canonical earnings process.

A.1 The PSID

The Panel Study of Income Dynamics (PSID) follows a large number of U.S. households over time and reports information about their demographic characteristics and sources of income. The PSID was initially composed of two major subsamples. The first of them, the SRC (Survey Research Center) or core subsample, was designed to be representative of the U.S. population and is a random sample itself, including over 18,000 individuals living in 5,000 households. The second, the SEO (Survey of Economic Opportunity) subsample, was created to study the characteristics of the most deprived households. Later, Immigrant and Latino subsamples were also added to the PSID.

From 1968 to 1997, the survey was yearly. After 1997, it started having a biennial structure. We only consider the SRC or core subsample because the SEO oversamples the poor. After dropping the SEO and Latino samples we are left with a random sample, which makes computations simpler since weights are not needed (Haider, 2001).[25]

A.2 Sample selection

Since the model period is one year, we restrict ourselves to the yearly part of the survey, and focus on the years 1968-1992. We have dropped the 1993-1997 period because there was a major redesign of the survey in 1993. This affected the method through which information was collected, with the introduction of computer-based surveys, and the definitions of some variables we use (for instance, asset income of other family members was no longer available,

[25]It must be taken into account that the weighting of our final dataset can be affected by attrition and by the fact that we are neglecting observations of yearly income under $ 1500 (expressed in 2015 dollars).

and wife labor income was redefined). We have verified that results are not sensitive to including these five years.

Following standard practice in the literature, we only consider individuals between ages 25 and 60. We consider all households, whether or not male-headed. This differs from many other papers, but follows e.g. Krueger, Mitman and Perri (2016).

We deflate values to 2013 dollars, and only keep observations with earnings (as defined below) above \$1500. This is also in accordance with standard practice in the literature, where observations below a minimum earnings threshold are dropped (De Nardi (2004) or Guvenen et al. (2016), for instance).

A.3 Income definition

For the estimation of our earnings process, we use after-tax equivalized household earnings. We first construct nonfinancial pre-tax household earnings using PSID data. Then, we estimate a tax function to obtain after-tax earnings. Finally, we regress earnings on the number of family members for the purposes of equivalization.

We now describe each of these three steps in detail.

A.3.1 Nonfinancial pre-tax income

We construct nonfinancial pre-tax income closely following Guvenen and Smith (2014). The procedure is based on subtracting all asset income from total family income.

Before 1976, asset income was not directly available in the survey. Therefore, we take total family income, subtract head and wife taxable income (which includes labor and asset income) and then add back labor earnings for head and wife independently.

From 1976, we consider all the available measures of asset income. These include farm income, business income, rent and interests, with the addition of gardening and roomers income (from 1978), and asset income of family members other than head and wife (from 1984).

We keep top-coded observations, but drop the very small number (8) of households who, probably due to measurement error, would have nonfinancial income below zero.

A.3.2 Tax function

We obtain **disposable** labor income by subtracting an estimated measure of taxes on labor income from nonfinancial pre-tax income.

We first compute the total amount of income taxes paid by households by adding up the federal income tax variable (which is available in the PSID until 1990) with a constructed measure of payroll (FICA and Medicare) taxes, which is based on applying the historical rates and caps to labor earnings of husband and wife independently.

We then separate taxes on labor and asset income by running a regression of this total tax measure on nonfinancial income and its square, and asset income and its square. This also follows Guvenen and Smith (2014). The estimated coefficients allow us to predict taxes on labor income[26], which we subtract from nonfinancial pre-tax income to get after-tax labor income.

A.3.3 Equivalization

We then equivalize after-tax nonfinancial disposable income by running a regression of earnings on the number of family members and keeping the residuals. We also extract year fixed effects.

A.4 Estimating the canonical and nonlinear earnings processes

A.4.1 Estimating the nonlinear earnings process

To finally implement the Arellano et al. (2017) procedure, we create a sample with all sets of subsequent three-year observations (without replacement: once an observation in the PSID

[26] We use the coefficients estimated in the sample 1968-1990 to predict taxes on labor income for the period 1968-1992, given that the PSID federal income tax variable is not available for the last two years of our sample.

sample is in a 3-year set in our sample we drop it). This implies that we are also dropping all of those households that do not have three consecutive valid income observations in the PSID.

We then follow the procedure described in Section 3.2 and the discretization explained in Section 4.

A.4.2 Estimating the canonical earnings process

In Storesletten et al. (2004*a*) (and in many other papers in the literature, e.g. Krueger et al. (2016)) the earnings process is estimated by fitting a parametric process to the variance of earnings profile that we observe in the data. The standard way is to compute the variance of earnings by age-cohort-year cells, and then get the coefficients of a regression of those on either age and year or age and cohort. For consistence with our approach and with the consumption data we rely on, we use the one that controls for year effects (see discussion below).

We follow a GMM procedure in which we minimize the distance of the estimated process to the profile of variances and first-order autocovariances of earnings over the life cycle[27]. The weighting matrix is the identity matrix.

The canonical earnings process in equations (1)-(3) implies (for $t > 1$)

$$y_{it} = \rho^{t-1}\eta_{i1} + \sum_{j=2}^{t} \rho^{t-j}\zeta_{ij} + \varepsilon_{it} \tag{A.1}$$

from which

$$var(y_{it}) = \rho^{2(t-1)}\sigma^2_{\eta_1} + \sum_{j=2}^{t} \rho^{2(t-j)}\sigma^2_{\zeta} + \sigma^2_{\varepsilon}. \tag{A.2}$$

[27]We describe in Appendix B how we compute these variances.

and

$$\text{(A.3)} \qquad cov(y_{it}, y_{i,t+1}) = \rho^{2t-1}\sigma^2_{\eta_1} + \sum_{j=2}^{t} \rho^{1+2(t-j)}\sigma^2_{\zeta}$$

follow, allowing to identify moments.

B Appendix: Computation of the variances of log earnings and log consumption.

We estimate the canonical earnings process described in Section 3.3 by matching the variance and first autocovariance of log earnings.

To compute the variance of log earnings, which we report in Figure 3, we use the procedure described in Kaplan (2012) (Appendix C.3), controlling for year effects, with our PSID data. More specifically, we take log disposable and equivalized labor income $\tilde{y}_{it}$, where i indexes the household and t is the age of its head, and run the regression

$$\text{(B.1)} \qquad \tilde{y}_{it} = \beta_t' \mathfrak{D}_t + \beta_d' \mathfrak{D}_d + y_{it},$$

where $\mathfrak{D}_t$ and $\mathfrak{D}_d$ are matrices with columns corresponding to a full set of age and year (date) dummies, respectively. The vectors β_t and β_d are the corresponding coefficients and y_{it} the earnings residuals.[28]

We compute the variance of y_{it} by age group as

$$\text{(B.2)} \qquad Var_t(y) = \frac{1}{D} \sum_{d=1}^{D} \left(\sum_{i=1}^{N_{d,t}} \frac{y_{it}^2}{N_{d,t}} \right),$$

where D is the number of years in the dataset, and $N_{d,t}$ is the numerosity of each age-year cell. This implies that the variance of earnings at age t weighs equally the corresponding

[28] As described in Appendix A, we use the earnings residuals from equations (B.1) to estimate our earnings processes.

conditional variances of earnings in each year.

We also compute the variance of y_{it} by age group controlling for cohort instead of year effect, s using the cohort counterpart of equation (B.2)

$$Var_t(y) = \frac{1}{K(t)} \sum_{k=1}^{K(t)} \left(\sum_{i=1}^{N_{k,t}} \frac{y_{it}^2}{N_{k,t}} \right), \tag{B.3}$$

where $K(t)$ is the number of cohorts containing individuals of age t and $N_{k,t}$ is the numerosity of each cohort-age cell.[29] This approach weighs the conditional variances from each cohort equally.

Under both approaches, we obtain very similar age profiles (Figure B.1) and parameter estimates for the canonical process (Table B.1).

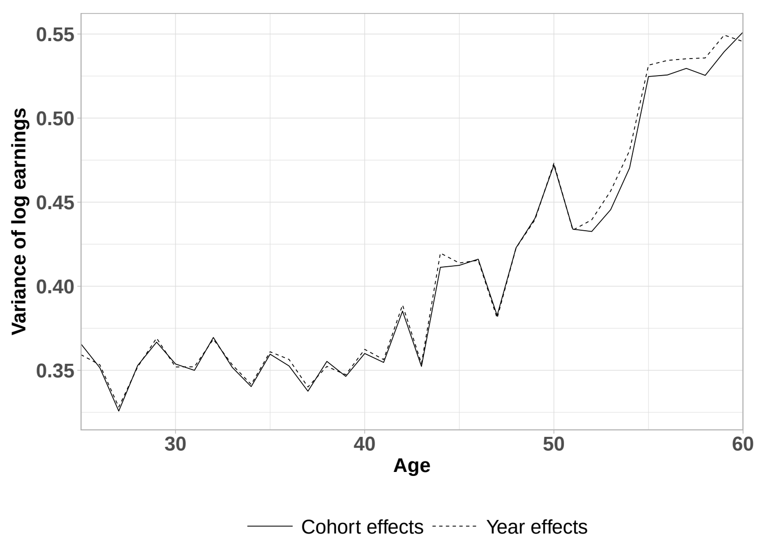

Figure B.1: Cross-sectional variance of log earnings over the life cycle, cohort effects vs year effects

Turning to consumption, we compute the variance of log consumption using data from the CEX for the period 1980-2007. Nondurable consumption includes food, clothing, gasoline, household operation, transportation, medical care, recreation, tobacco, and education.

[29]The residuals used in equations (B.2) and (B.3) are the same. Given that year, age and cohort are linearly dependent, the residuals from equation (B.1) are the same that would obtain from projecting onto age and cohort dummies.

	σ_{ε}^2	$\sigma_{\eta_1}^2$	σ_{ζ}^2	ρ
Year effects	0.0620	0.2332	0.0060	1
Cohort effects	0.0669	0.2379	0.0057	1

Table B.1: Estimates for the canonical earnings process: cohort vs. year effects

We compute the variance of log consumption following the same procedures that we use for the variance of log earnings. Namely, we deal with year effects using the method proposed by Kaplan (2012) and equivalize consumption with a regression on the number of family members. We have also applied this procedure to OECD-equivalized consumption data and verified that it yields very similar results to Heathcote, Perri and Violante (2010) when they control for year effects.

C Appendix: Robust measures of skewness and kurtosis

Figure C.1 represents a robust measure of skewness (Kelley's skewness) and a robust measure of kurtosis (Crow-Siddiqui kurtosis) for male pre-tax earnings in the W2, and for both male pre-tax and household after-tax earnings in the PSID.

Equations C.1 and C.2 show that these measures are computed taking into account specific percentiles of the distribution of earnings changes and, as such, are robust to the effect of outliers. The inspection of C.1 shows that all of the main features highlighted in Section 2 are still present in these more robust measures.

$$KS = \frac{(P_{90} - P_{50}) - (P_{50} - P_{10})}{P_{90} - P_{10}} \tag{C.1}$$

$$CS = \frac{P_{97.5} - P_{2.5}}{P_{75} - P_{25}} \tag{C.2}$$

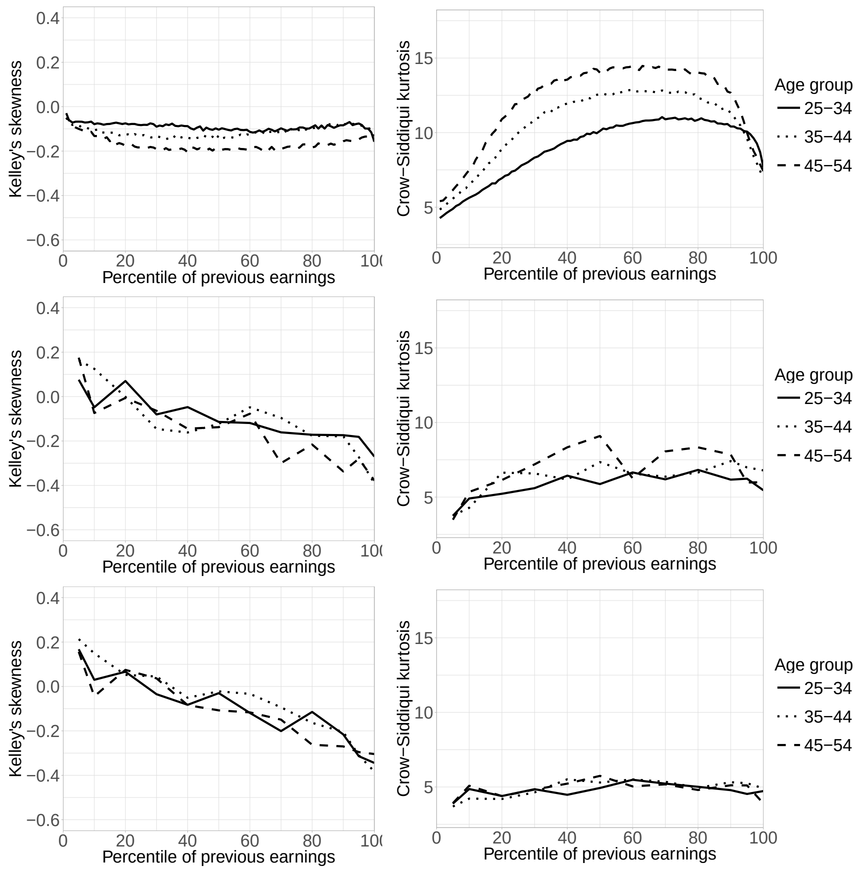

Figure C.1: Kelly skewness and Crow-Siddiqui kurtosis of male pre-tax earnings growth in the W2 (top panel) and PSID (central panel), and of household after-tax earnings growth in the PSID (bottom panel)

D Appendix: Fit of the earnings process to the data and robustness of the discretization

To sum up, our procedure for earnings requires two steps:

1. Apply the Arellano et al. (2017) decomposition to *PSID data* to obtain the *persistent*

and *transitory* components of earnings. This is described in Sections 3.1 and 3.2.

2. Discretize the simulated persistent component and transitory component using an age-dependent Markov chain to obtain the *discretized persistent* and *discretized transitory* components. This is described in Section 4.

In this appendix we show that (1) the central features of the data that we are interested in replicating are preserved by the discretized process that we use in the structural model and (2) that our main findings are robust to increasing the number of grid points for the two components of earnings. We also report the results for consumption inequality and self-insurance in the case in which the discount factor β is kept constant across earnings processes.

D.1 Fit of the earnings process to the data

D.1.1 Conditional moments of earnings changes

Figure D.1 plots the second, third, and fourth standardized moments of earnings changes. The left panel refers to our PSID sample (and thus replicates the bottom panel of Figure 1. The central panel displays the same measures computed on earnings (i.e. the sum of the *persistent* and *transitory* components), simulated using the estimated nonlinear process associated with the quantile functions in (6)-(8). The right panel reports the corresponding moments computed on earnings simulated using our discrete approximation of such continuous processes.

The polynomial quantile functions and their discretization smooths the skewness and kurtosis graphs (which are very noisy and affected by outliers). Yet, the main patterns of the data (negative skewness, large kurtosis and variation over previous earnings) are preserved.

Figure D.2 zooms in on just the estimated *persistent component* of earnings and compares its features of the *persistent component* (top) with those its discretized counterpart. It suggests that our flexible discretization, despite only having 18 bins per age group, captures

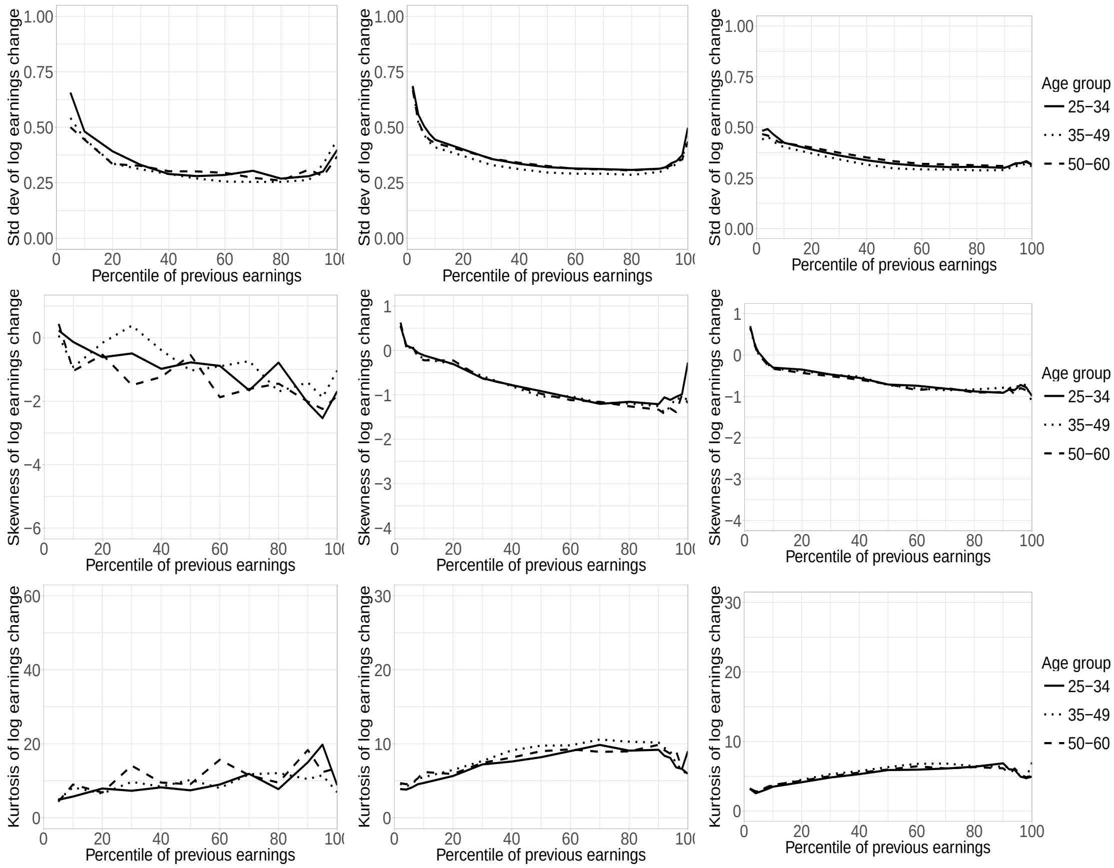

Figure D.1: Conditional moments of earnings changes. From top to bottom: standard deviation, skewness, kurtosis. Left: PSID data; center: simulated earnings using estimated nonlinear process; right: simulated earnings using discretized nonlinear process.

these feature of the data.

Figure D.3 conducts the same comparison for the estimated transitory component and shows that our discretization reproduces the observed age-dependence, high negative skewness, and large kurtosis. The 8-gridpoints discretization (top panel) generates much larger kurtosis than that of a normal distribution but falls a bit short of that in the data because it cannot accurately capture the effect of outliers. A finer, 16-gridpoints, discretization (bottom panel), which has additional bins on the tails, does match the kurtosis in the data. In Appendix D.2, we describe this alternative discretization more in detail and we show it does not make a difference for our results from our structural model.

Finally, Figure D.4 reports persistence by previous-earnings (percentile τ_{init}) and current-

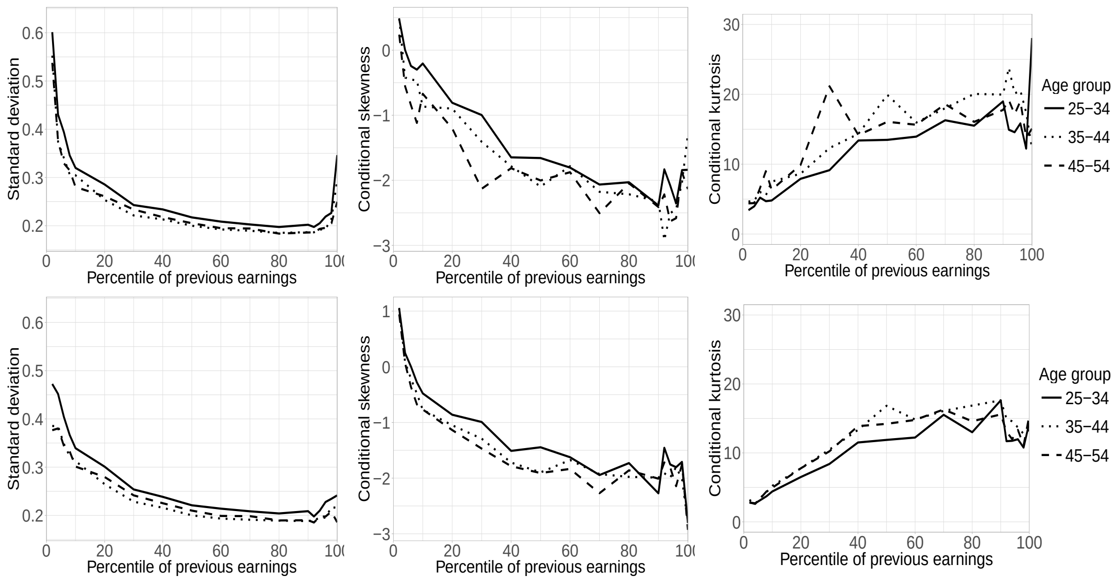

Figure D.2: Conditional moments of earnings changes (persistent component). From left to right: standard deviation, skewness, kurtosis. Top: persistent component; bottom: discretized persistent component.

shock (percentile τ_{shock}) rank. The top left panel refers to the PSID data, the top right panel to earnings simulated using the discretized process (persistent plus transitory component), the bottom left panel to the estimated *persistent component*, and the bottom right panel to its discretized counterpart. The discretization makes the graph for nonlinear persistence less smooth, but it preserves most of its important features. Namely, earnings are less persistent for high earners who receive a very bad shock and and low earners who receive a very good shock, while they are most persistent for high earners who receive a good shock and low earners who receive a bad shock.

D.1.2 Unconditional moments of persistent earnings

Figure D.5 plots the *unconditional* moments of the persistent earnings distribution, as opposed to the conditional moments of earnings changes in the previous sections. Our discretization captures very well their levels and variations by age, but similarly to Figure D.3 we need a finer discretization to match the very high levels of kurtosis in the data. We describe this finer discretization in Section D.2, where we also show that it generates very

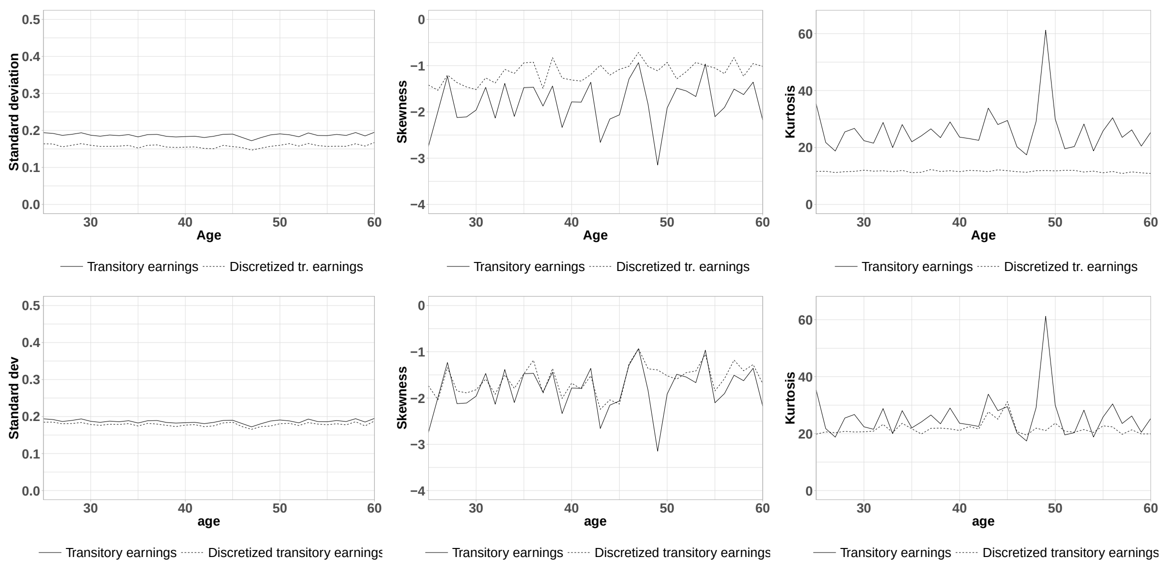

Figure D.3: Moments for the transitory shock. From left to right, standard deviation, skewness and kurtosis (top: main 8-gridpoint discretization; bottom: 16-gridpoint discretization).

similar results to our main discretization.

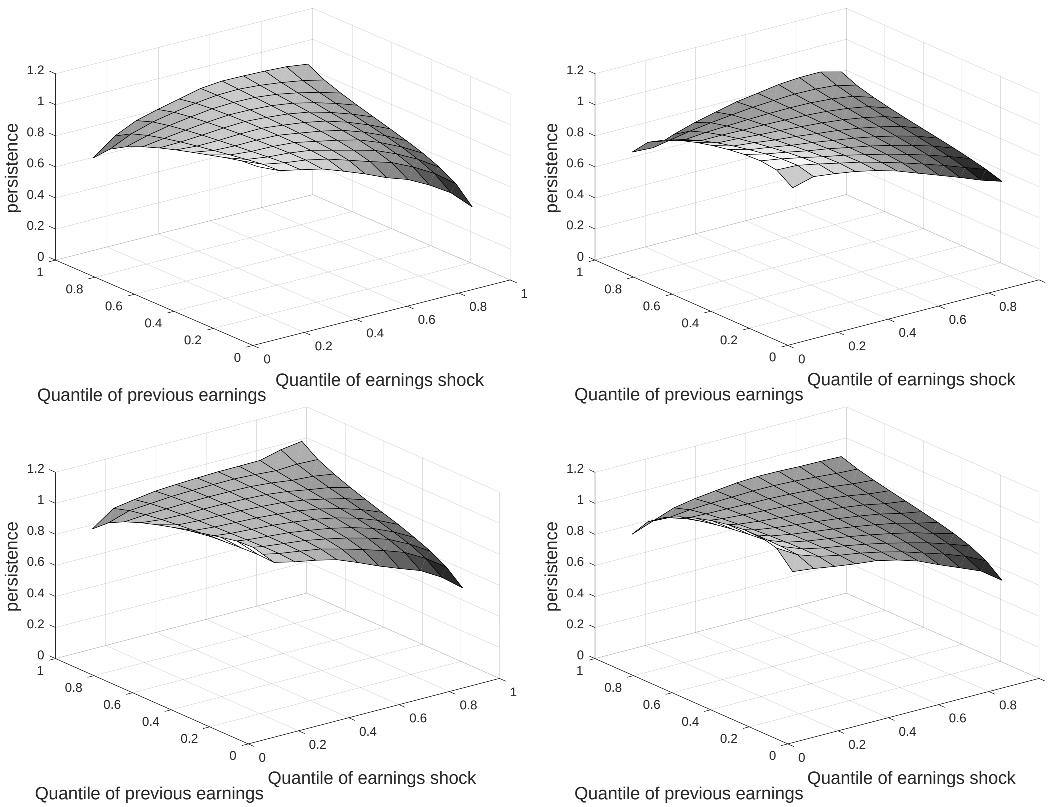

Figure D.4: Nonlinear persistence by quantile of previous earnings and quantile of the shock received in the current period (top left, PSID data; top right, persistent component; bottom left, discretized persistent + transitory component; bottom right, discretized persistent component).

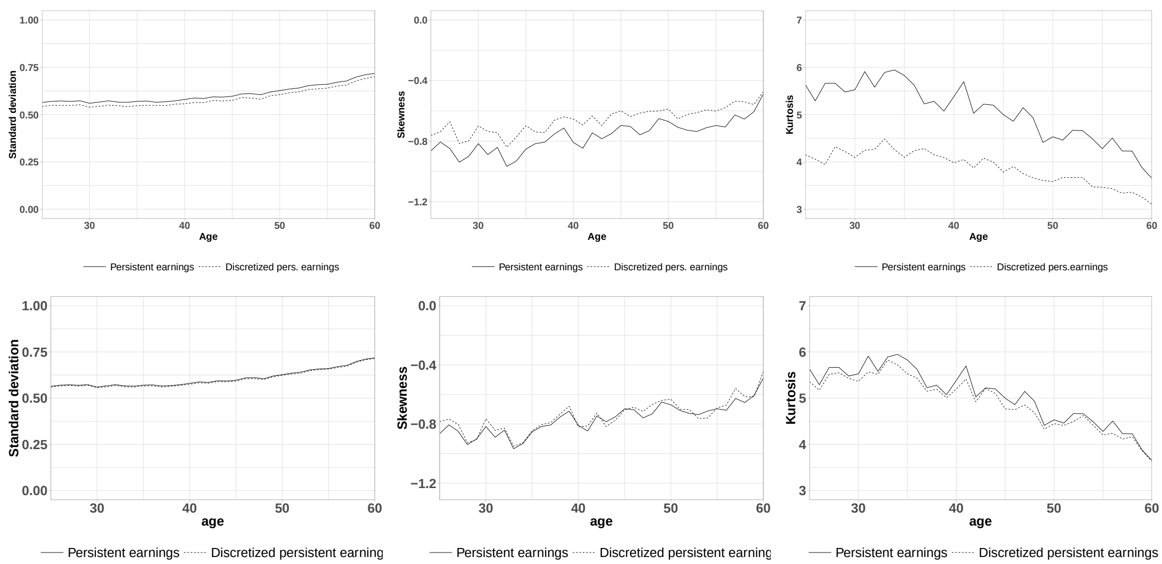

Figure D.5: Unconditional moments of the persistent earnings distribution: from left to right, standard deviation, skewness and kurtosis (top: main 18-gridpoint discretization; bottom: 36-gridpoint discretization).

D.2 Robustness to the number of earnings gridpoints

In our main results, at each age, we discretize the persistent component of earnings with 18 gridpoints and the transitory one with 8 gridpoints, as described in Section 4. These are the smallest grid sizes beyond which adding additional grid points makes little difference for the economic implications discussed in Section 6.

In this section, we report results on simulating our model with finer grids and show that our results are robust to these changes. The finer discretization has the advantage that it fits some moments of the data better (like the unconditional moments of the transitory and persistent earnings shocks, as seen in Figures D.3 and D.5).

Figure D.6 and Table D.1 show the growth in the variance of log consumption and the BPP coefficients under a finer discretization for the transitory component. More specifically, the transitory component is divided into, respectively, 8 gridpoints as in our main results (corresponding to the bottom 2.5%, next 2.5%, next 5%, next 40%, next 40%, next 5%, next 2.5% and top 2.5%) and 16 gridpoints (bottom 0.1%, next 0.4%, next 0.5%, next 2%, next 2%, next 5%, four quintiles, next 5%, next 2%, next 2%, next 0.5%, next 0.4% and finally top 0.1%). The differences between the 8- and 16- gridpoints specification are very small.

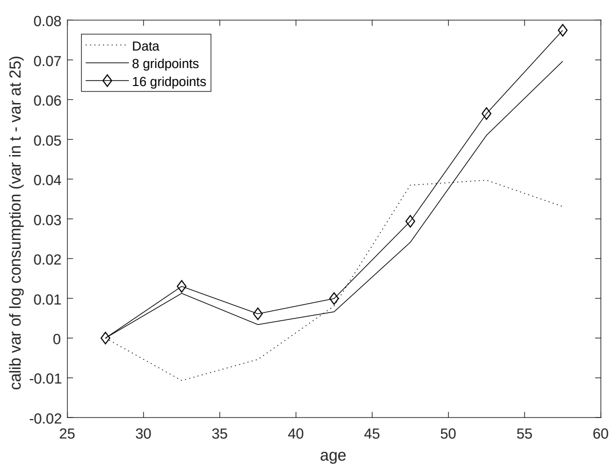

Figure D.6: Growth in the variance of log consumption, different discretizations for the transitory component.

Figure D.7 and Table D.2 show the results for alternative discretizations of the *persistent*

Coefficients	ψ^p_{BPP}	ψ^{tr}_{BPP}	ψ^p	ψ^{tr}
Data	0.36	0.95	–	–
NL process	0.42	0.82	0.45	0.89
NL process, 16 gridpoints	0.43	0.83	0.47	0.90

Table D.1: BPP coefficients, different discretizations for the transitory component.

component. Namely, it compares our main results with a 36-gridpoints discretization which adopts a very thin division of the bottom and top percentiles (for the top, there is a bin for the top 0.05%, following 0.05%, 0.2%, 0.2% and 0.5% and symmetrically for the bottom), percentiles for the rest of the top and bottom 5% and groups of 5% for the rest of the persistent earnings distribution. Differences are, again, minor.

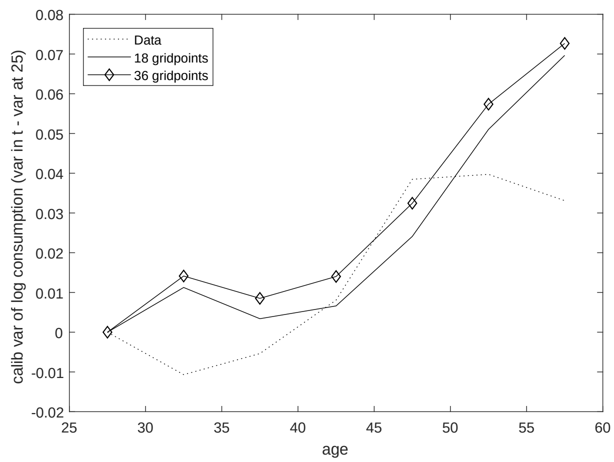

Figure D.7: Growth in the variance of log consumption, different discretizations for the persistent component.

Coefficients	ψ^p_{BPP}	ψ^{tr}_{BPP}	ψ^p	ψ^{tr}
Data	0.36	0.95	–	–
NL process	0.42	0.82	0.45	0.89
NL process, 36 gridpoints	0.41	0.83	0.45	0.92

Table D.2: BPP coefficients, different discretizations for the persistent component.

D.3 Alternative discretization of the canonical process

As described in Section 5.5, for our main results we discretize the canonical process using the modified version of the Rouwenhorst method for non-stationary processes proposed by Fella et al. (2017). Here we show that our findings that the nonlinear process provides a substantially better fit are not due to using to different discretizations for the canonical and the nonlinear processes. To this effect, we discretize the canonical process by taking the parametric estimates in Table 1, simulating a panel of earnings histories and applying the same the same age-varying Markov chain procedure as we followed for the NL process.

Figure D.8 and Table D.3 show the implied consumption profiles and BPP coefficients. Both discretizations give rise to qualitatively similar results. Under the alternative discretization, the canonical earnings process overshoots the growth in the variance of log consumption over the life-cycle by an even larger amount.

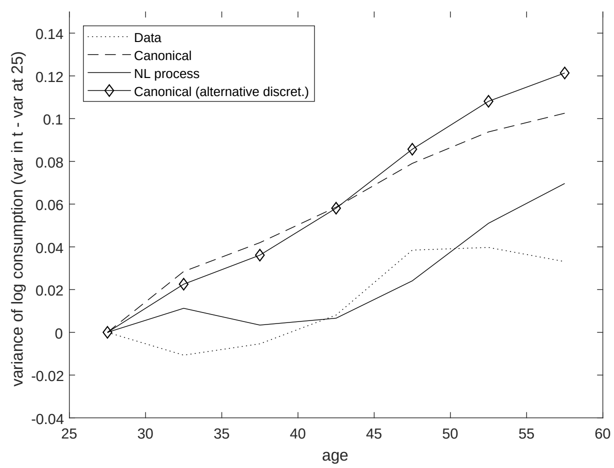

Figure D.8: Growth in the variance of log consumption, different discretizations for the canonical process.

Coefficients	ψ^p_{BPP}	ψ^{tr}_{BPP}	ψ^p	ψ^{tr}
Data	0.36	0.95	–	–
Canonical	0.09	0.89	0.29	0.91
Canonical, alternative discretization	0.25	0.89	0.36	0.92

Table D.3: BPP coefficients, different discretizations for the canonical process.

D.4 Results without persistent-transitory decomposition

In Section 3.1 we describe the flexible earnings process that we estimate, which is based in the persistent-transitory decomposition proposed by Arellano et al. (2017). However, an alternative, computationally less costly choice is to apply directly our Markov-chain flexible discretization method to the raw data. Figure D.9 and Table D.4 provide the results of applying that simpler method to our PSID sample.

Neglecting the persistent-transitory decomposition implies a substantial underestimation of the growth of the variance of log consumption over the life cycle and, consistently, an overestimation of the ability of households to self-insure against earnings shocks. This can partially reflect the existence of measurement error in the PSID, and provides further support to the procedure we follow in our main results. With administrative data, where the measurement error issue is smaller, we also the expect differences in implications when we take our the transitory shock to be smaller.

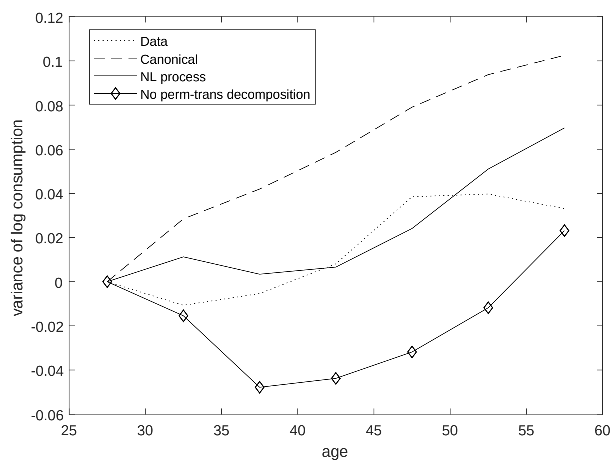

Figure D.9: Growth in the variance of log consumption, no permanent-transitory decomposition.

D.5 Results with a common discount factor

The profiles for the variance of log consumption and insurance coefficients that we report in Section 6 are computed by recalibrating the discount factor β to keep the wealth to income

Coefficients	ψ^p_{BPP}	ψ^{tr}_{BPP}	ψ^p	ψ^{tr}
Data	0.36	0.95	–	–
NL process	0.42	0.82	0.45	0.89
NL process, no pers-transitory	0.49	–	0.57	–

Table D.4: BPP coefficients, no permanent-transitory decomposition.

ratio identical in the economy under the canonical process and the economy under the NL process. On the other hand, to perform the welfare computations we keep β constant so that both economies can be comparable. Here we report the consumption and insurance results that correspond to this latter case. Namely, we fix β at its calibrated value for the NL process and then estimate the model for the canonical process. Given that the calibrated value for β is higher for the canonical process, there is less accumulation of wealth in this economy.

Figure D.10 reports the associated consumption inequality profile, which is very similar to that of the canonical process with a recalibrated discount factor, if not slightly worse. Table D.5 reports the insurance coefficients. The true coefficients ψ^p and ψ^{tr} are slightly smaller than in the case with a recalibrated β, which reflects that self-insurance is lower in an economy with less wealth accumulation. The BPP estimate of the insurance coefficients against permanent shocks ψ^p_{BPP} is, like we argued in Section 6.2, downward biased in an economy where the borrowing constraint is occasionally binding. With a lower discount factor, the borrowing constraint is more frequently binding, and the bias becomes so large that the coefficient is estimated to be negative.

Coefficients	ψ^p_{BPP}	ψ^{tr}_{BPP}	ψ^p	ψ^{tr}
Data	0.36	0.95	–	–
NL process	0.42	0.82	0.45	0.89
Canonical	0.09	0.89	0.29	0.91
Canonical, common β	-0.11	0.83	0.26	0.87

Table D.5: BPP coefficients, no permanent-transitory decomposition.

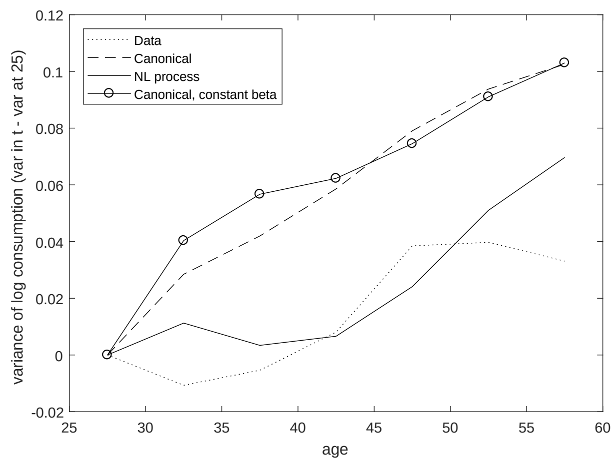

Figure D.10: Growth in the variance of log consumption, canonical process with constant β.