Household Earnings and Consumption: A Nonlinear Framework

Manuel Arellano* Richard Blundell† Stéphane Bonhomme‡

10 March 2014

PRELIMINARY AND INCOMPLETE

Abstract

We develop a flexible framework to study the nonlinear relationship between shocks to household earnings and consumption over the life cycle. Log-earnings are the sum of a general Markovian persistent component and a transitory innovation. Consumption is modelled as an age-dependent nonlinear function of assets and the two earnings components. We establish the nonparametric identification of the earnings process and the consumption rule, and apply quantile-based methods for estimation. Exploiting the enhanced consumption and asset data in the PSID for 1999-2009, we find the impact of earnings shocks varies substantially across households' earnings histories, and that this nonlinearity drives heterogeneous consumption responses. We also find that the insurability of earnings shocks is higher late in the life-cycle, and that assets play an important role.

JEL code: C23, D31, D91.
Keywords: Earnings dynamics, consumption, latent variables.

*CEMFI, Madrid.
†University College London and Institute for Fiscal Studies.
‡University of Chicago.

1 Introduction

Consumption decisions and earnings dynamics are intricately linked. Together with the net value of assets, the size and durability of any income shock dictates how much consumption will need to adjust to ensure a reasonable standard of living in future periods of the life-cycle. Understanding the persistence of earnings is therefore of great interest not only because it affects the permanent or transitory nature of inequality, but also because it drives much of the variation in consumption. Our aim in this study is to examine the interrelationship between income and consumption dynamics in a flexible manner. To preempt our results, we find the impact of earnings shocks varies substantially across households' earnings histories, and that this nonlinearity is a key driver of heterogeneous consumption responses.

With some notable exceptions (see the discussion in Meghir and Pistaferri, 2011), the literature on earnings dynamics has focused on linear models. The random walk permanent/transitory model is a popular example (Abowd and Card, 1989). Linear models have the property that all shocks are associated with the *same* persistence, irrespective of the household's earnings history. Linearity is a convenient assumption, as it allows to study identification and estimation using standard covariance techniques. However, by definition linear models rule out nonlinear transmission of shocks and nonlinearities in income dynamics are likely to have a first-order impact on consumption choices. Thus the twin objectives of this paper. First, to develop a flexible earnings model that allows to capture interesting nonlinearities. Second, to assess the impact of nonlinear earnings shocks on household consumption.

The existing literature on earnings shocks and consumption follows two main approaches. One approach is to take a stand on the precise mechanisms that households use to smooth consumption, for example saving and borrowing or labor supply, and to calibrate a fully-specified life-cycle model to the data; see Gourinchas and Parker (2002) for example. Except in very special cases (as in Hall and Mishkin, 1982) the consumption function is generally a complex nonlinear function of earnings components.[1]

Another approach is to estimate the degree of "partial insurance" from the data without precisely specifying the insurance mechanisms; see Blundell, Pistaferri and Preston (2008) for example. Linear approximations to the optimization problem deliver tractable estimating equations. However, linear approximations may not be accurate (Kaplan and Violante,

[1]An interesting recent exception is Heathcote, Storesletten and Violante (2013).

2008). Moreover, some aspects of consumption smoothing such as precautionary savings or asset accumulation in the presence of borrowing constraints are nonlinear in nature, making a linear framework possibly problematic.

In this paper we develop a flexible framework to study the nonlinear relationship between shocks to household earnings and consumption over the life cycle. Log-earnings are the sum of a general Markovian persistent component and a transitory innovation. This modelling allows to capture the intuition that different shocks may be associated with different persistence. For example, our framework allows for "unusual" shocks to wipe out the memory of past shocks. This feature, which could empirically correspond to job losses, changes of career, or health shocks, is at odds with linear models commonly used in the literature.

We model consumption as an age-dependent nonlinear function of assets and the two earnings components. We motivate our specification using a standard life-cycle model of consumption and saving with incomplete markets, see Huggett (1993), for example. The consumption rule is nonlinear, thus allowing for interactions between asset holdings and the persistent or transitory earnings components. This flexible modelling allows to capture an array of response coefficients that provides a rich picture of the extent of consumption insurance in the data.

As the consumption model is unrestricted, our framework nests standard life-cycle models. In particular, there is no need for approximation arguments as we directly estimate the nonlinear consumption rule. In addition, we show how to extend our baseline model to allow for household heterogeneity in preferences or discounting, advance information on earnings shocks, and habits in consumption.

In contrast to linear models, our nonlinear model cannot be studied using standard techniques. As a result, a large part of the paper is devoted to the econometric analysis. We establish nonparametric identification, building on a recent literature on nonlinear models with latent variables. Identification of the earnings process builds on Wilhelm (2012). Identification of the consumption rule relies on related but different arguments. In the identification analysis we emphasize the link to instrumental-variables techniques commonly used in linear models.

We take the model to PSID data for 1999-2009. Unlike earlier waves of the PSID, these data contain enhanced information on asset holdings and consumption expenditures in addition to labor earnings, see Blundell, Pistaferri and Saporta-Eksten (2012), for example.

This is the first household panel to include detailed information on consumption and assets across the life-cycle for a representative sample of households. Without the panel information on earnings, consumption and assets, our approach would not be feasible.

We estimate the earnings process and consumption rule using a nonlinear specification (a sieve) that combines quantile modelling and linear expansions in bases of functions. This flexible approach builds on Wei and Carroll (2009) and Arellano and Bonhomme (2013). The sequential estimation algorithm consists in iterating between quantile regression estimation, and draws from the posterior distribution of the latent persistent components of earnings. We propose a parametric inference method, and build on Nielsen (2000a) to analyze its large-sample properties.

The preliminary results that we report in this version of the paper suggest that the impact of earnings shocks varies substantially across households, and that this nonlinearity drives heterogeneous consumption responses. We also find that the insurability of earnings shocks is higher late in the life-cycle, and that assets play an important role.

The outline of the paper is as follows: we present the models of earnings and consumption, establish identification as well as several extensions of the baseline model, describe our estimation strategy and the dataset, and discuss our preliminary results.

2 Model (I): Earnings process

We start by describing the model of earnings dynamics. In the next section we will present the consumption model.

2.1 The model

We consider a cohort of households, $i = 1, ..., N$, and denote as t the age of the household head. Let Y_{it} be the pre-tax labor earnings of household i at age t, and let y_{it} denote log-Y_{it}, net of a full set of age dummies. We decompose y_{it} as follows:[2]

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

The first component η_{it} is assumed to follow a general first-order Markov process. We denote the τth conditional quantile of η_{it} given $\eta_{i,t-1}$ as $Q_t(\eta_{i,t-1}, \tau)$, for each $\tau \in (0, 1)$. The

[2]Model (1) is additive in η and ε. Given our nonlinear approach, it is in principle possible to allow for interactions between the two earnings components, for example in $y_{it} = H_t(\eta_{it}, \varepsilon_{it})$ subject to suitable normalization. Identification could then be established along the lines of Hu and Shum (2012).

following representation is then without loss of generality given the Markov assumption[3]

$$\eta_{it} = Q_t(\eta_{i,t-1}, u_{it}), \quad \text{where } (u_{it}|\eta_{i,t-1}, \eta_{i,t-2}, ...) \sim \textit{Uniform}(0,1). \tag{2}$$

The second component ε_{it} is assumed to have zero mean, and to be independent over time. We denote its marginal cumulative distribution function as F_{ε_t}. Even though more general moving average representations are commonly used in the literature, the biennial nature of our data makes this assumption more plausible. Nevertheless, allowing for serial dependence in ε_{it} would be an important extension.

Both earnings components are mean independent of age t. However, the conditional quantile functions Q_t, and the marginal distributions F_{ε_t}, may all depend on t. For a given cohort of households, age and calendar time are perfectly collinear, so this dependence may capture age effects as well as aggregate shocks. The distribution of the initial condition η_{i1}, which we denote as F_{η_1}, is left unrestricted.

A special case of model (1)-(2) is obtained when $\eta_{it} = \eta_{i,t-1} + v_{it}$ is a random walk. When v_{it} is independent of $\eta_{i,t-1}$ and follows a cdf F_t, (2) becomes: $\eta_{it} = \eta_{i,t-1} + F_t^{-1}(u_{it})$. We will refer to the random walk plus independent shock as the *canonical model* of earnings dynamics.

Throughout the paper we will refer to η_{it} as the persistent component of earnings, and to ε_{it} as the transitory one. The earnings shocks u_{it} and $e_{it} = F_{\varepsilon_t}(\varepsilon_{it})$ are expressed in percentile terms. In our framework, the dependence of η_{it} is not restricted beyond the Markov assumption. The identification assumptions will only require that η_{it} be dependent over time, without specifying the degree of dependence.

2.2 Nonlinear dynamics

Model (1)-(2) allows for nonlinear dynamics of earnings. Here we focus on its ability to capture nonlinear persistence, and general forms of conditional heteroskedasticity.

Nonlinear persistence. Let us consider the following quantities

$$\rho_t(\eta_{i,t-1}, \tau) = \frac{\partial Q_t(\eta_{i,t-1}, \tau)}{\partial \eta}, \quad \text{and} \quad \rho_t(\tau) = \mathbb{E}\left[\frac{\partial Q_t(\eta_{i,t-1}, \tau)}{\partial \eta}\right], \tag{3}$$

[3] We assume that the probability distributions of η's and ε's are absolutely continuous. Note that, given that our earnings data are recorded every other year, this specification is consistent with both first or second-order Markov assumptions at the yearly frequency.

where $\partial Q_t/\partial\eta$ denotes the partial derivative of Q_t with respect to its first component, and where the expectation is taken with respect to the distribution of $\eta_{i,t-1}$.

The ρ's in (3) are measures of nonlinear persistence of the η component. $\rho_t(\eta_{i,t-1},\tau)$ measures the persistence of $\eta_{i,t-1}$ when it is hit by a shock u_{it} that has rank τ. This quantity depends on the state $\eta_{i,t-1}$, and on the percentile of the shock τ. Average persistence across η values is $\rho_t(\tau)$.

In the canonical model of earnings dynamics, $\rho_t(\eta_{i,t-1},\tau) = 1$ irrespective of $\eta_{i,t-1}$ and τ. In contrast, in model (2) the persistence of $\eta_{i,t-1}$ may depend on the magnitude and direction of the shock u_{it}. The interaction between the shock u_{it} and the lagged persistent component $\eta_{i,t-1}$ is a key feature of our nonlinear approach.

It is useful to consider the following specification of the quantile function

$$Q_t(\eta_{i,t-1},\tau) = \alpha_t(\tau) + \beta_t(\tau)'h(\eta_{i,t-1}), \tag{4}$$

where h is a multi-valued function. Examples are $h(\eta) = |\eta|$ or $h(\eta) = (\max(\eta,0), \min(\eta,0))$, which correspond to the CAViaR quantile regression models of Engle and Manganelli (2004). Our empirical specification will be based on (4), taking the components of h in a polynomial basis of functions capable of approximating any continuous function arbitrarily well as the number of polynomial terms tends to infinity. Persistence and average persistence in (4) are $\rho_t(\eta_{i,t-1},\tau) = \beta_t(\tau)'\frac{\partial h(\eta_{i,t-1})}{\partial\eta}$ and $\rho_t(\tau) = \beta_t(\tau)'\mathbb{E}\left[\frac{\partial h(\eta_{i,t-1})}{\partial\eta}\right]$, respectively, thus allowing shocks to affect the persistence of $\eta_{i,t-1}$ in rather general ways.

Empirically, this nonlinear persistence may capture a number of labor market events in a reduced-form fashion. For example, it can be that job losses, changes of career, or health shocks within the household have a large effect that leads the current state to be less persistent. Conversely, it could be that earnings shocks that hit households having stable employment histories are associated with a high persistence. In Section 7 we will see that empirical estimates of model (1)-(2) on PSID data are consistent with this interpretation.

Conditional heteroskedasticity. As (2) does not restrict the form of the conditional distribution of η_{it} given $\eta_{i,t-1}$, it allows for general forms of heteroskedasticity. In particular, a measure of period-t uncertainty generated by the presence of shocks to the persistent earnings component is

$$\sigma_t(\tau) = \mathbb{E}\left[Q_t(\eta_{i,t-1},\tau) - Q_t(\eta_{i,t-1},1-\tau)\right].$$

As a special case, in the canonical model with $v_{it} \sim \mathcal{N}(0, \sigma_{vt}^2)$, $\sigma_t(\tau) = 2\sigma_{v_t}\Phi^{-1}(\tau)$. An analogous measure of uncertainty generated by the transitory shocks is

$$\sigma_{\varepsilon_t}(\tau) = F_{\varepsilon_t}^{-1}(\tau) - F_{\varepsilon_t}^{-1}(1-\tau).$$

In addition, the model allows for conditional skewness and kurtosis in η_{it}. Along the lines of the measures proposed by Kim and White (2004), one can consider

$$sk_t(\eta_{i,t-1}, \tau) = \frac{Q_t(\eta_{i,t-1}, \tau) + Q_t(\eta_{i,t-1}, 1-\tau) - 2Q_t(\eta_{i,t-1}, .5)}{Q_t(\eta_{i,t-1}, \tau) - Q_t(\eta_{i,t-1}, 1-\tau)},$$

and

$$kur_t(\eta_{i,t-1}, \tau, \alpha) = \frac{Q_t(\eta_{i,t-1}, 1-\alpha) - Q_t(\eta_{i,t-1}, \alpha)}{Q_t(\eta_{i,t-1}, \tau) - Q_t(\eta_{i,t-1}, 1-\tau)},$$

for example taking $\tau = .75$ and $\alpha = .025$. The empirical estimates below suggest that conditional skewness is a feature of the earnings process.[4]

3 Model (II): Consumption rule

In order to motivate our empirical specification, we start by outlining a simple life-cycle model of consumption and savings and derive the form of the policy rule for household consumption. We then describe the empirical consumption model that we will take to the data.

3.1 A simple life-cycle model

Households live for a finite number of periods T^*. For simplicity we assume that they work up to period T^*. Accounting for a retirement period would not fundamentally affect the analysis. Households maximize expected life-time utility given by

$$\max_{C_{i1}, C_{i2}, \ldots, C_{iT^*}} \mathbb{E}_1 \left[\sum_{t=1}^{T^*} \beta^{t-1} u(C_{it}) \right], \tag{5}$$

where C_{it} denote consumption levels, β is a discount factor, and the expectation is conditional on the agent's information set in period 1. We assume that $u(\cdot)$ is twice differentiable, increasing and strictly concave.

[4]In a recent paper on US Social Security Data for 1978-2010, Guvenen, Ozcan and Song (2012) find that the left-skewness of earnings shocks is counter-cyclical. It would be interesting to apply our framework to study distributional dynamics over the business cycle.

Households receive earnings $Y_{it} = \exp(\mu_t + y_{it})$, where y_{it} are given by (1)-(2), and μ_t are deterministic functions of age t. They have access to a single risk-free, one-period bond whose constant return is $1 + r$. The evolution of beginning-of-period assets A_{it} is given by

$$A_{it} = (1 + r)A_{i,t-1} + Y_{i,t-1} - C_{i,t-1}, \quad t \in \{2, ..., T^*\}. \tag{6}$$

Initial assets A_{i1} are unrestricted. End-of-life assets are assumed to satisfy $A_{iT^*} \geq 0$.[5]

At each period t, households observe A_{it}, η_{it}, and ε_{it}, and all their past values since the start of life, that is

$$\mathcal{A}_t = \{A_{it}, \eta_{it}, \varepsilon_{it}, ..., A_{i1}, \eta_{i1}, \varepsilon_{i1}\}.$$

However, households do not have advance information about the future realizations of η (or alternatively future shocks u) and ε. All distributions are known to households, and there is no aggregate uncertainty.

The period-t value function is given by

$$V_t(A_{it}, \eta_{it}, \varepsilon_{it}) = \max_{C_{it}, C_{i,t+1}, ..., C_{iT^*}} \mathbb{E}\left[\sum_{s=t}^{T^*} \beta^{s-t} u(C_{is}) \,\middle|\, \mathcal{A}_t \right],$$

and the Bellman equation is

$$\begin{aligned} V_t(A_{it}, \eta_{it}, \varepsilon_{it}) &= \max_{C_{it}} \; u(C_{it}) + \beta \mathbb{E}\left[V_{t+1}\left((1 + r)A_{it} + Y_{it} - C_{it}, \eta_{i,t+1}, \varepsilon_{i,t+1} \right) \,\middle|\, \mathcal{A}_t \right] \\ &= \max_{C_{it}} \left\{ u(C_{it}) + ... \right. \\ &\left. \quad \beta \int\int V_{t+1}\left((1 + r)A_{it} + Y_{it} - C_{it}, \eta, \varepsilon \right) f_{\eta_{t+1}|\eta_t}(\eta|\eta_{it}) f_{\varepsilon_{t+1}}(\varepsilon) d\eta d\varepsilon \right\}, \end{aligned}$$

where we have used the fact that η_{it} is first-order Markov, and ε_{it} are independent over time, and where $f_{\eta_{t+1}|\eta_t}$ and $f_{\varepsilon_{t+1}}$ denote the conditional density of $\eta_{i,t+1}$ given η_{it} and the marginal density of $\varepsilon_{i,t+1}$, respectively, both of which are allowed to depend on t.

The consumption rule is then of the form

$$C_{it} = G_t\left(A_{it}, \eta_{it}, \varepsilon_{it} \right), \tag{7}$$

for some age-dependent function G_t.

[5]There could be additional borrowing constraints in each period. In that case, the nonparametric consumption rule below would no longer be differentiable.

When one is interested in documenting dynamic patterns of consumption and earnings, one strategy is to take a stand on the functional form of the utility function and the distributions of the shocks, and to calibrate or estimate the model's parameters by comparing the model's predictions with the data, as in Gourinchas and Parker (2002) for example. Another strategy is to linearize the consumption rule (7), with the help of the budget constraint. With a linear approximated problem at hand, standard covariance-based methods may be used for estimation, as in Blundell, Pistaferri and Preston (2008).

Our approach differs from the previous literature as we directly estimate the nonlinear consumption rule (7). Doing so, we avoid linearizing the model, and we estimate a flexible rule that is consistent with the life-cycle consumption model outlined above. In addition, this approach allows to document a rich set of derivative effects, thus shedding light on the actual amount of consumption insurance in the data.[6]

3.2 An empirical consumption rule

Let us consider a cohort of households. Let c_{it} denote log-consumption net of a full set of age dummies. Similarly, let a_{it} denote log-assets net of age dummies. Our empirical specification is based on

$$c_{it} = g_t\left(a_{it}, \eta_{it}, \varepsilon_{it}, \nu_{it}\right), \quad t = 1, ..., T, \tag{8}$$

where ν_{it} are independent across periods, and g_t is monotone in ν.

An economic interpretation for ν is as a taste shifter that increases marginal utility. Indeed, in the above single-asset life-cycle model monotonicity is implied by the Bellman equation, provided $\frac{\partial u(C,\nu')}{\partial C} > \frac{\partial u(C,\nu)}{\partial C}$ for all C if $\nu' > \nu$. Without loss of generality, we normalize the marginal distribution of ν_{it} to be standard uniform in each period.

Insurance coefficients. Average consumption, for given values of asset holdings and earnings components, is

$$\mathbb{E}\left[c_{it}|a_{it} = a, \eta_{it} = \eta, \varepsilon_{it} = \varepsilon\right] = \mathbb{E}\left[g_t\left(a, \eta, \varepsilon, \nu_{it}\right)\right].$$

[6]Estimating the consumption rule (7) could also be of interest in the perspective of the first, model-based approach. Indeed, one could take advantage of our estimates of the consumption rule and the distributions of the shocks in order to identify and estimate a fully-fledged life-cycle consumption model.

Our framework allows to document how average consumption varies with its arguments. In particular, the average derivative with respect to η is

$$\phi_t(a,\eta,\varepsilon)=\mathbb{E}\left[\frac{\partial g_t\left(a,\eta,\varepsilon,\nu_{it}\right)}{\partial\eta}\right], \tag{9}$$

while the average derivative with respect to ε is

$$\psi_t(a,\eta,\varepsilon)=\mathbb{E}\left[\frac{\partial g_t\left(a,\eta,\varepsilon,\nu_{it}\right)}{\partial\varepsilon}\right]. \tag{10}$$

The parameters $\phi_t(a,\eta,\varepsilon)$ and $\psi_t(a,\eta,\varepsilon)$ reflect the degree of insurability of shocks to the persistent and transitory earnings components, respectively. We will document how they vary along the life-cycle, and how they depend on households' asset holdings. Specifically, we will report estimates of

$$\overline{\phi}_t(a)=\mathbb{E}\left[\phi_t(a,\eta_{it},\varepsilon_{it})\right], \quad \text{and} \quad \overline{\psi}_t(a)=\mathbb{E}\left[\psi_t(a,\eta_{it},\varepsilon_{it})\right]. \tag{11}$$

Earnings shocks and consumption. The marginal effect on consumption of an earnings shock u_{it} to the persistent earnings component is, by the chain rule and equation (9),

$$\mathbb{E}\left[\left.\frac{\partial}{\partial u}\right|_{u=\tau} g_t\left(a,Q_t(\eta,u),\varepsilon,\nu_{it}\right)\right]=\phi_t\left(a,Q_t(\eta,\tau),\varepsilon\right)\frac{\partial Q_t(\eta,\tau)}{\partial u}.$$

This marginal effect depends on η through the insurance coefficient ϕ_t, but also through the quantity $\frac{\partial Q_t(\eta,\tau)}{\partial u}$ as the earnings model allows for general forms of conditional heteroskedasticity. In the empirical analysis we will report finite-difference counterparts to these derivative effects.

4 Identification

The earnings and consumption models take the form of nonlinear state-space models. A series of recent papers (notably Hu and Schennach, 2008, and Hu and Shum, 2012) has established conditions under which nonlinear models with latent variables are nonparametrically identified under conditional independence restrictions. Here we rely on techniques developed in this literature in order to establish identification of the models we consider.

4.1 Earnings process

Consider model (1)-(2), where η_{it} is a Markovian persistent component and ε_{it} are independent over time. We assume that the data contain T consecutive periods, which we denote as

$t = 1, ..., T$. So, for a given cohort of households, $t = 1$ corresponds to the age at which the household head enters the sample, and $t = T$ corresponds to the last period of observation. For that cohort, our aim is to identify the joint distributions of $(\eta_{i1}, ..., \eta_{iT})$ and $(\varepsilon_{i1}, ..., \varepsilon_{iT})$ given a random sample from $(y_{i1}, ..., y_{iT})$. In the following, all conditional and marginal densities are assumed to be bounded.

Operator injectivity. The identification arguments below rely on the concept of operator injectivity, which we now formally define. A linear operator $\mathcal{L}$ is a linear mapping from a functional space $\mathcal{H}_1$ to another functional space $\mathcal{H}_2$. $\mathcal{L}$ is *injective* if the only solution $h \in \mathcal{H}_1$ to the equation $\mathcal{L}h = 0$ is $h = 0$.

One special case of operator injectivity ("deconvolution") obtains when $Y_2 = Y_1 + \epsilon_1$, with Y_1 independent of ϵ_1, and $[\mathcal{L}h](y_2) = \int h(y_1) f_{\epsilon_1}(y_2 - y_1) dy_1$. $\mathcal{L}$ is then injective if the characteristic function of ϵ_1 has no zeros on the real line. The normal and many other standard distributions satisfy this property.[7] In particular, if the marginal distributions f_{Y_2} and f_{ϵ_1} are known, injectivity implies that $h = f_{Y_1}$ is the only solution to the functional equation $\int h(y_1) f_{\epsilon_1}(y_2 - y_1) dy_1 = f_{Y_2}(y_2)$. In other words, f_{Y_1} is identified from the knowledge of f_{Y_2} and f_{ϵ_1}. Deconvolution arguments of this type are commonly used in the literature.

Another special case of operator injectivity ("completeness") is obtained when $\mathcal{L}$ is the conditional expectation operator associated with the distribution of $(Y_1|Y_2)$, in which case $[\mathcal{L}h](y_2) = \mathbb{E}[h(Y_1) \,|\, Y_2 = y_2]$. $\mathcal{L}$ being injective is then equivalent to the distribution of $(Y_1|Y_2)$ being *complete*. Completeness is commonly assumed in nonparametric instrumental variables problems, see Newey and Powell (2003). Recent work provides primitive conditions for completeness in specific cases; see D'Haultfoeuille (2011) and Andrews (2011).

Wilhelm's result. To establish nonparametric identification of the earnings process, we rely on a result in Wilhelm (2012), which was originally derived in the context of a panel data model with measurement error. Wilhelm provides conditions under which the marginal distribution of ε_{i2} is identified, given three periods of observations (y_{i1}, y_{i2}, y_{i3}). We reproduce his identification argument in the appendix.

Wilhelm's result is derived under several high-level assumptions. In particular, it requires that the distributions of $(y_{i3}|\eta_{i2})$ and $(\eta_{i2}|y_{i1})$ be both complete. This requires that η_{i1} and

[7] In fact, injectivity also holds if the zeros of the characteristic function of ϵ_1 are isolated. See Carrasco and Florens (2011) and Evdokimov and White (2011).

η_{i2}, and also η_{i2} and η_{i3}, be statistically dependent, albeit without specifying the form of that dependence. An intuition for this is that if η's were independent over time there would be no way to distinguish them from the transitory ε's.

Identification of the earnings process. Returning to the earnings dynamics model (1)-(2), let now $T \geq 3$. Suppose that the conditions in Wilhelm (2012) are satisfied on each of the three-year subpanels $t \in \{1, 2, 3\}$ to $t \in \{T-2, T-1, T\}$. It follows from Wilhelm's result that the marginal distributions of ε_{it} are identified for all $t \in \{2, 3, ..., T-1\}$. By serial independence of the ε's, the joint distribution of $(\varepsilon_{i2}, \varepsilon_{i3}, ..., \varepsilon_{i,T-1})$ is thus also identified.

Hence, if the characteristic functions of ε_{it} do not vanish on the real line, then by a deconvolution argument the joint distribution of $(\eta_{i2}, \eta_{i3}, ..., \eta_{i,T-1})$ is identified. As a result, all Markov transitions $f_{\eta_t|\eta_{t-1}}$ are identified for $t = 2, ..., T-1$, and the marginal distribution of η_{i2} is identified as well. Moreover, it is easy to show that the conditional distributions of $\eta_{i2}|y_{i1}$ and $y_{iT}|\eta_{i,T-1}$ are identified.[8]

Note that, in the case where ε_{i1}, ..., ε_{iT} have the same marginal distribution, then the distributions of the initial and terminal components ε_{i1}, η_{i1}, and ε_{iT}, η_{iT} are also identified. However, the first and last-period distributions are generally not identified in the absence of stationarity assumptions.

4.2 Consumption rule

Let us now turn to the identification of the consumption rule (8). We make the following assumptions. First, we assume that η_{it} is independent of a_{i1} given $\eta_{i,t-1}$, and that ε_{it} is independent of a_{i1}. These assumptions require earnings shocks, which are independent of past components of earnings, to be independent of initial asset holdings as well. At the same time, we let η_{i1} and a_{i1} be arbitrarily dependent. This is important, because asset accumulation upon entry in the sample may be correlated with past earnings shocks.

Next, we assume that current assets are only determined by previous period assets, earnings, and consumption. Formally, we assume that a_{it} are independent of $(\eta_{i,t-1}, a_{i,t-2}, \eta_{i,t-2}, \varepsilon_{i,t-2}...)$ given $(a_{i,t-1}, c_{i,t-1}, y_{i,t-1})$. This assumption is consistent with a standard life-cycle model with one single risk-less asset, see equation (6). It is also consistent with a model

[8]Indeed we have $f_{y_2|y_1}(y_2|y_1) = \int f_{\varepsilon_2}(y_2 - \eta_2) f_{\eta_2|y_1}(\eta_2|y_1) d\eta_2$. Hence, as the characteristic function of ε_{i2} is non-vanishing, $f_{\eta_2|y_1}(\cdot|y_1)$ is identified for almost all y_1. A similar argument shows that $f_{y_T|\eta_{T-1}}(y_T|\cdot)$ is identified for almost all y_T.

where the interest rate r_t is time-varying and known to households. Note however that this assumption may be empirically strong if assets are risky and households choose their investment strategy in a similar way as they choose their consumption profile.

Lastly, we assume that all marginal and conditional distributions are bounded from above and below. With some abuse of notation we use $f(A|B)$ as a generic notation for the conditional distribution $f_{A|B}(\cdot|\cdot)$. We omit the i index for simplicity.

First period. Let us start by analyzing period $t = 1$. Letting $y = (y_1, ..., y_T)$, we have $f(c_1|a_1, \eta_1, y) = f(c_1|a_1, \eta_1, y_1)$ from the consumption rule. Hence

$$f(c_1|a_1, y) = \int f(c_1|a_1, \eta_1, y_1)f(\eta_1|a_1, y)d\eta_1. \tag{12}$$

Note that, if η_{i1} and a_{i1} were independent given y_i, an injectivity argument would identify the period-1 consumption rule.[9] Upon entry in the sample, however, this conditional independence restriction is unlikely to hold.

In order to identify the consumption rule in the case where a_{i1} and η_{i1} are dependent, we first note that, by Bayes' rule,

$$f(\eta_1|a_1, y) = \frac{f(y|\eta_1, a_1)f(\eta_1|a_1)}{f(y|a_1)} = \frac{f(y|\eta_1)f(\eta_1|a_1)}{f(y|a_1)} = f(\eta_1|y)\frac{f(y)f(\eta_1|a_1)}{f(y|a_1)f(\eta_1)},$$

where we have used that y_i is independent of a_{i1} given η_{i1}. Hence, using (12),

$$f(c_1|a_1, y)\frac{f(y|a_1)}{f(y)} = \int \frac{f(c_1|a_1, \eta_1, y_1)f(\eta_1|a_1)}{f(\eta_1)}f(\eta_1|y)d\eta_1, \tag{13}$$

where $f(\eta_1|y)$ is identified from the earnings process alone.

Let us now define the operator

$$[\mathcal{L}_1 h](c_1, a_1, y) = \int h(c_1, a_1, \eta_1, y_1)f(\eta_1|y)d\eta_1.$$

Let us assume that $\mathcal{L}_1$ is injective. Then there exists a unique function h that satisfies

$$[\mathcal{L}_1 h](c_1, a_1, y) = f(c_1|a_1, y)\frac{f(y|a_1)}{f(y)}.$$

This implies that $\frac{f(c_1|a_1,\eta_1,y_1)f(\eta_1|a_1)}{f(\eta_1)}$ is identified. As $f(\eta_1)$ is identified from the earnings process, it follows that $f(c_1|a_1, \eta_1, y_1)$ and $f(\eta_1|a_1)$ are both identified.[10]

[9] To see this, let us define $[\mathcal{L}h](c_1, a_1, y) = \int h(c_1, a_1, \eta_1, y_1)f(\eta_1|y)d\eta_1$. If $\mathcal{L}$ is injective, the solution to $[\mathcal{L}h](c_1, a_1, y) = f(c_1|a_1, y)$ is unique. Hence, as $f(\eta_1|y)$ is identified from the earnings process alone it follows that the density of period-1 consumption $f(c_1|a_1, \eta_1, y_1)$ is identified.

[10] This comes from $f(c_1|a_1, \eta_1, y_1)f(\eta_1|a_1)$ being identified, so $f(\eta_1|a_1)$ by integration with respect to c_1, and $f(c_1|a_1, \eta_1, y_1)$ upon dividing by $f(\eta_1|a_1)$, are identified as well.

An intuitive explanation for the identification argument is that $y_{i2}, ..., y_{iT}$ are used as "instruments" for η_{i1}. In particular, $T \geq 2$ is needed. Using leads of log-earnings for identifying consumption responses is a common strategy in linear models, see for example Hall and Mishkin (1982), and Blundell, Pistaferri and Preston (2008).

Injectivity of $\mathcal{L}_1$ only depends on the properties of the earnings process. Moreover, this condition is intuitive given the Markovian property of η_{it}. As an example, consider the case where $T = 2$, and $(\eta_{i1}, y_{i1}, y_{i2})$ follows a multivariate normal distribution with zero mean. Then $\eta_{i1} = \alpha y_{i1} + \beta y_{i2} + \zeta_i$, where ζ_i is normal $(0, \sigma^2)$, independent of (y_{i1}, y_{i2}). It can easily be shown that $\beta \neq 0$ if $\text{Cov}(\eta_{i1}, \eta_{i2}) \neq 0$, in which case $\mathcal{L}_1$ is injective.[11] As in the identification of the earnings process, identification of the consumption rule thus relies on η's being dependent over time.

Subsequent periods. Let us then consider the period-t problem, where $t \geq 2$. We denote $c^t = (c_1, ..., c_t)$, and use similar notation for a^t and η^t.

We proceed by induction. Let $t \in \{2, ..., T\}$, and assume that $f(c_s|\eta_s, a_s, y_s)$ is identified for all $1 \leq s \leq t-1$. We have

$$\begin{aligned} f(c^t, a_2, ..., a_t|a_1, y) &= \int \prod_{s=1}^{t} f(c_s|a_s, \eta_s, y_s) \prod_{s=2}^{t} f(a_s|a_{s-1}, y_{s-1}, c_{s-1}) f(\eta^t|a_1, y) d\eta^t \\ &= \prod_{s=2}^{t} f(a_s|a_{s-1}, y_{s-1}, c_{s-1}) \int \prod_{s=1}^{t} f(c_s|a_s, \eta_s, y_s) f(\eta^t|a_1, y) d\eta^t. \end{aligned} \tag{14}$$

Note that $f(a_s|a_{s-1}, y_{s-1}, c_{s-1})$ is identified for all s. Moreover, consider the following linear operator

$$[\mathcal{L}_t h](c^t, a^t, y) = \int h(c_t, a_t, \eta_t, y_t) \kappa_t \left(\eta_t, c^{t-1}, a^{t-1}, y\right) d\eta_t,$$

where

$$\kappa_t \left(\eta_t, c^{t-1}, a^{t-1}, y\right) = \int \prod_{s=1}^{t-1} f(c_s|a_s, \eta_s, y_s) f(\eta^t|a_1, y) d\eta^{t-1}.$$

[11] Formally, taking Fourier transforms (for a fixed y_1 value) in

$$\int_{-\infty}^{\infty} g(\eta_1, y_1) \frac{1}{\sigma} \phi \left(\frac{\eta_1 - \alpha y_1 - \beta y_2}{\sigma} \right) d\eta_1 = 0$$

yields $g(\cdot, y_1) = 0$ for almost all y_1, provided $\beta \neq 0$.

By Bayes' rule and the assumptions on initial assets and earnings, we have

$$\begin{aligned} f(\eta^t|a_1,y) &= \frac{f(y|\eta^t)\prod_{t=2}^T f(\eta_t|\eta_{t-1})f(\eta_1|a_1)}{f(y|a_1)} \\ &= f(\eta^t|y)\frac{f(y)f(\eta_1|a_1)}{f(y|a_1)f(\eta_1)}, \end{aligned}$$

which is identified from the earnings process and the period-1 consumption problem. As a result, given that the consumption rules up to period $t-1$ are assumed identified in the induction argument, the operator $\mathcal{L}_t$ is identified as well.

Let us assume that $\mathcal{L}_t$ is injective. Then there exists a unique function h that solves

$$[\mathcal{L}_t h](c^t,a^t,y) = \frac{f(c^t,a_2,...,a_t|a_1,y)}{\prod_{s=2}^t f(a_s|a_{s-1},y_{s-1},c_{s-1})}.$$

So by (14) $f(c_t|\eta_t,a_t,y_t)$ is identified. This shows that the consumption rules are all identified for $t=1,...,T$.

Unlike the first-period case, the operator $\mathcal{L}_t$ depends on lags and leads of earnings (at all periods) and lags of assets and consumption (up to period $t-1$). Intuitively, lagged consumption and assets, as well as lags and leads of earnings, are used as instruments for η_{it}. Injectivity depends on the relevance of these instruments (in a nonparametric sense).

5 Extensions

In this section we introduce several extensions of the baseline model, and we briefly discuss identification in each of them.

5.1 Advance information

If households have advance information about future earnings shocks, the consumption rule (8) takes future earnings components as additional arguments. For example, let us consider a model where households know the realization of the one-period-ahead persistent component, in which case

$$c_{it} = g_t\left(a_{it},\eta_{it},\eta_{i,t+1},\varepsilon_{it},\nu_{it}\right), \quad t=1,...,T-1. \tag{15}$$

Identification can be established using similar arguments as in the baseline model. In period 1 we have

$$\begin{aligned} f(c_1|a_1,y) &= \int\int f(c_1|a_1,\eta_1,\eta_2,y_1)f(\eta_1,\eta_2|a_1,y)d\eta_1 d\eta_2 \\ &= \frac{f(y)}{f(y|a_1)}\int\int \frac{f(c_1|a_1,\eta_1,\eta_2,y_1)f(\eta_1|a_1)}{f(\eta_1)}f(\eta_1,\eta_2|y)d\eta_1 d\eta_2, \end{aligned}$$

where we have used that η_{i2} is independent of a_{i1} given η_{i1}, and that y_i is independent of a_{i1} given (η_{i1}, η_{i2}).

Let us assume that the operator

$$[\mathcal{L}_1 h](c_1, a_1, y) = \int \int h(c_1, a_1, \eta_1, \eta_2, y_1) f(\eta_1, \eta_2 | y) d\eta_1 d\eta_2$$

is injective. Note in particular that we need $T \geq 3$. Then $f(c_1|a_1, \eta_1, \eta_2, y_1)$ and $f(\eta_1|a_1)$ are identified. An argument similar to the one we used in the baseline model then identifies the period-t conditional distribution $f(c_t|a_t, \eta_t, \eta_{t+1}, y_t)$ for $t = 1, ..., T-1$.

Lastly, similar arguments can be used to show identification in models where households have advance information about future transitory shocks ε, as well as in models where the consumption rule depends on lags of η or ε, for example in models where η_{it} follows a second-order Markov process.

5.2 Habits

In the presence of habits, the consumption rule takes the form

$$c_{it} = g_t\left(c_{i,t-1}, a_{it}, \eta_{it}, \varepsilon_{it}, \nu_{it}\right), \quad t = 2, ..., T. \tag{16}$$

To see under which conditions the model is identified, let us take $t = 2$. Let us denote $\widetilde{y}_i = (y_{i2}, ..., y_{iT})$. We have

$$f(c_2|c_1, a_2, \widetilde{y}) \quad = \quad \int f(c_2|c_1, a_2, \eta_2, y_2) f(\eta_2|c_1, a_2, \widetilde{y}) d\eta_2.$$

Assuming that $\widetilde{y}_i$ is independent of (c_{i1}, a_{i2}) given η_{i2} we obtain

$$f(c_2|c_1, a_2, \widetilde{y}) \quad = \quad \frac{f(\widetilde{y})}{f(\widetilde{y}|c_1, a_2)} \int \frac{f(c_2|c_1, a_2, \eta_2, y_2) f(\eta_2|c_1, a_2)}{f(\eta_2)} f(\eta_2|\widetilde{y}) d\eta_2.$$

Identification of $f(c_2|c_1, a_2, \eta_2, y_2)$ and $f(\eta_2|c_1, a_2)$ follows, provided that the operator

$$[\mathcal{L}_2 h](c_1, c_2, a_2, \widetilde{y}) = \int h(c_1, c_2, a_2, \eta_2, y_2) f(\eta_2|\widetilde{y}) d\eta_2$$

be injective. Here also, $T \geq 3$ is needed. Identification of $f(c_t|c_{t-1}, a_t, \eta_t, y_t)$ $(t = 3, ..., T)$ follows similarly.

5.3 Household heterogeneity

In the baseline model, households differ *ex-ante* in their earnings due to heterogeneous initial conditions η_{i1} and level of assets a_{i1}. In contrast, the consumption rule is fully homogeneous. As accounting for unobserved heterogeneity in preferences or discounting may be empirically important, we now develop an extension of the model that allows for a household-specific effect ξ_i. The consumption rule is then

$$c_{it} = g_t\left(a_{it}, \eta_{it}, \varepsilon_{it}, \xi_i, \nu_{it}\right), \quad t = 1, ..., T. \tag{17}$$

For simplicity we consider a scalar ξ_i. Depending on the number of available time periods, a vector of unobserved heterogeneity could be allowed for.

We assume that η_{it} is independent of (a_{i1}, ξ_i) given $\eta_{i,t-1}$ for all $t \geq 2$, and that ε_{it} is independent of (a_{i1}, ξ_i) for $t \geq 1$. At the same time, the distribution of $(\xi_i, \eta_{i1}, a_{i1})$ is left unrestricted. Simple calculations similar to the ones of Section 4 then yield

$$\begin{aligned} f(c^t, a_2, ..., a_t | a_1, y) \quad = \quad & \frac{f(y)}{f(y|a_1)} \prod_{s=2}^{t} f(a_s | a_{s-1}, y_{s-1}, c_{s-1}) \times ... \\ & \int \frac{1}{f(\eta_1)} \left(\int \prod_{s=1}^{t} f(c_s | a_s, \eta_s, y_s, \xi) f(\eta_1, \xi | a_1) d\xi \right) f(\eta^t | y) d\eta^t. \end{aligned} \tag{18}$$

Note that we have also assumed that a_{is} is independent of ξ_i (as well as η_{is}) given $a_{i,s-1}$, $y_{i,s-1}$, and $c_{i,s-1}$.

Let us now evaluate (18) at $t = 3$. We assume that the operator

$$[\mathcal{L}_3 h](c_1, c_2, c_3, a_1, a_2, a_3, y) = \int h(c_1, c_2, c_3, a_1, a_2, a_3, y_1, y_2, y_3, \eta_1, \eta_2, \eta_3) f(\eta_1, \eta_2, \eta_3 | y) d\eta^3$$

is injective. In particular, this requires that $T \geq 6$.[12] It then follows from (18) that

$$\int \prod_{t=1}^{3} f(c_t | a_t, \eta_t, y_t, \xi) f(\eta_1, \xi | a_1) d\xi = f(\eta_1 | a_1) \int \prod_{t=1}^{3} f(c_t | a_t, \eta_t, y_t, \xi) f(\xi | \eta_1, a_1) d\xi$$

is identified, for almost all (c^3, a^3, y^3, η^3). It thus follows that $f(\eta_1 | a_1)$ and

$$k(c^3; a^3, \eta^3, y^3) \equiv \int \prod_{t=1}^{3} f(c_t | a_t, \eta_t, y_t, \xi) f(\xi | \eta_1, a_1) d\xi \tag{19}$$

[12] It is not clear whether the requirement that $T \geq 6$ is necessary for establishing identification, even though we need it in the proof.

are identified for almost all (c^3, a^3, y^3, η^3).

For fixed values of (a^3, y^3, η^3), equation (19) has the structure of mixture model, where ξ_i is the unobserved latent component. The main result of Hu and Schennach (2008) thus applies to show that $f(\xi|\eta_1, a_1)$ and $f(c_t|a_t, \eta_t, y_t, \xi)$ for all $t \in \{1, 2, 3\}$ are nonparametrically identified under suitable assumptions. Hu and Schennach's assumptions include injectivity conditions analogous to the ones we have used in the baseline model, as well as a scaling condition. In particular, the latter is satisfied if the mean (or any quantile) of c_{i1} given $(a_{i1}, \eta_{i1}, y_{i1}, \xi_i)$ is increasing in ξ_i, in which case identification is to be understood up to an increasing transformation of ξ_i.

Given that $f(\xi|\eta_1, a_1)$ is identified, it is then easy to see that $f(c_t|a_t, \eta_t, y_t, \xi)$ is identified for all $t \geq 1$. To see this, let us suppose by induction that $f(c_s|a_s, \eta_s, y_s, \xi)$ is identified for $1 \leq s \leq t-1$. Let us assume that the operator

$$[\mathcal{L}_t h](c^t, a^t, y) = \int\int h(c_t, a_t, \eta_t, y_t, \xi)\kappa_t(\eta_t, \xi, c^{t-1}, a^{t-1}, y)d\eta_t d\xi$$

is injective, where

$$\kappa_t(\eta_t, \xi, c^{t-1}, a^{t-1}, y) = \int \frac{1}{f(\eta_1)} \prod_{s=1}^{t-1} f(c_s|a_s, \eta_s, y_s, \xi) f(\eta_1, \xi|a_1) f(\eta^t|y) d\eta^{t-1}.$$

Then the solution h to

$$[\mathcal{L}_t h](c^t, a^t, y) = \frac{f(y|a_1)}{f(y)} \frac{f(c^t, a_2, ..., a_t|a_1, y^t)}{\prod_{s=2}^{t} f(a_s|a_{s-1}, y_{s-1}, c_{s-1})}$$

is unique, so by (18) $f(c_t|a_t, \eta_t, y_t, \xi)$ is identified.

Lastly, it is possible to allow for unobserved heterogeneity in earnings as well, in addition to heterogeneity in η_{i1}. Specifically, one can let η_{it} be a first-order Markov process conditional on another latent component $\widetilde{\xi}_i$.[13] In that case, an extension of Wilhelm (2012)'s result allows to show identification of the marginal distribution of ε_{i3} given five periods of earnings data $(y_{i1}, ..., y_{i5})$. This extension, and the corresponding identification argument for the consumption rule, will be described in a future draft of the paper.

5.4 Measurement error

Survey data like the PSID are often contaminated by errors; see for example Bound *et al.* (2001). In the absence of information about measurement error, it is not possible to

[13] In particular, this extension nests linear earnings models with slope heterogeneity, see Guvenen (2007) and Guvenen and Smith (2010) for example.

disentangle the transitory innovation from the measurement error when true earnings y_{it}^* are not observed, and

$$\begin{aligned} y_{it} &= y_{it}^* + \zeta_{it} \\ &= \eta_{it} + \varepsilon_{it} + \zeta_{it}, \end{aligned}$$

is observed with classical measurement error ζ_{it}, independent of η_{it} and ε_{it}. Thus, an interpretation of our estimated distribution of ε_{it} is that it represents a mixture of transitory shocks and measurement error.

If additional information is available and the marginal distribution of ζ_{it} is known, then one can recover the distribution of ε_{it} using a deconvolution argument. A modification of the estimation algorithm described in the next section deals with this case.

Lastly, allowing for measurement error in asset holdings and consumption is not straightforward. Nevertheless note that, from an empirical perspective, the presence of the taste shifters ν_{it} in the consumption rule (8) – which we will assume to be additive in ν_{it} when taking the model to the PSID data – may partly capture measurement error in consumption expenditures.

6 Data and estimation strategy

6.1 Data

Since 1999 the PSID contains detailed data on consumption expenditures and asset holdings, in addition to household earnings and demographics. Data are available every other year. We use data for the 1999-2009 period (six waves).

Earnings Y_{it} are total pre-tax household labor earnings. We construct y_{it} as residuals of log household earnings on a set of demographics, which include cohort and calendar time dummies (thus implicitly including age dummies as well), family size and composition (including dummies for income recipients other than husband and wife, and for kids out of home), education, race, and state and big city dummies, the last four being interacted with time dummies. Controls for family size and composition are included so as to equivalize household earnings. We will proceed similarly for consumption and asset holdings. Education, race and geographic dummies are included in an attempt to capture individual heterogeneity beyond cohort effects and the initial persistent component of earnings η_{i1}.

Table 1: Descriptive statistics (means)

	1999	2001	2003	2005	2007	2009
Earnings	85,001.41	93,983.93	100,280.6	106,683.7	119,039.4	122,907.7
Consumption	30,182.19	35,846.02	39,843.4	47,635.89	52,174.8	50,582.89
Assets	266,957.8	315,865.7	376,484.9	399,901.4	501,590.3	460,262.2

Notes: Balanced subsample from PSID, $N = 749$, $T = 6$.

We use data on consumption C_{it} of nondurables and services. Since 1999, PSID data contains information on health expenditures, utilities, car-related expenditures and transportation, education, and child care. Recreation, alcohol, tobacco and clothing (the latter available from 2005) are the main missing items. Rent information is available for renters, but not for home owners. We follow Blundell, Pistaferri and Saporta-Eksten (2012, BPS hereafter) and impute rent expenditures for home owners.[14] In total, approximately 70% of consumption expenditures are covered. We construct c_{it} as residuals of log total consumption on the same set of demographics as for earnings.

Asset holdings A_{it} are constructed as the sum of financial assets (including cash, stocks and bonds), real estate value, pension funds, and car value, net of mortgages and other debt. We construct residuals a_{it} by regressing log-assets on the same set of demographics as for earnings and consumption, and use them as arguments of the consumption rule (8).

To select the sample we follow BPS and focus on a sample of participating and married male heads aged between 30 and 65. We drop all observations for which data on earnings, consumption, or assets, either in levels or log-residuals, are missing. Finally, we also drop observations with extreme "jumps" in household earnings between three consecutive waves (0.25% of observations). Moreover, in this version of the paper we focus on a balanced subsample of $N = 749$ households.

Table 1 shows mean total earnings, consumption and asset holdings, by year. Compared to BPS, households in our balanced sample have higher assets, and to a less extent higher earnings and consumption.

[14] Note that, as a result, consumption responds automatically to variations in house prices. An alternative would be to exclude rents and imputed rents from consumption expenditures.

6.2 Empirical specification

Earnings components. The earnings model depends on the Markovian transitions of the persistent component $Q_t(\cdot,\cdot)$, the marginal distributions of ε_{it}, and the marginal distribution of the initial persistent component η_{i1}. We now explain how we empirically specify these three components.

Let $\varphi_k(\cdot)$, for $k = 0, 1, ...$, denote a dictionary of functions defined on $\mathbb{R}^2$, with $\varphi_0 = 1$. Letting age_{it} denote the age of the head of household i in period t, we specify

$$\begin{aligned} Q_t(\eta_{t-1},\tau) &= Q(\eta_{t-1}, age_t, \tau) \\ &= \sum_{k=0}^{K} a_k^Q(\tau)\varphi_k(\eta_{t-1}, age_t), \end{aligned} \tag{20}$$

where the number of terms K is chosen by the researcher.

We specify the quantile functions of ε_{it} (for $t = 1, ..., T$) given age_{it}, and that of η_{i1} given age at the start of the period age_{i1}, in a similar way. Specifically, we set

$$\begin{aligned} Q_\varepsilon(age_t,\tau) &= \sum_{k=0}^{K} a_k^\varepsilon(\tau)\varphi_k(age_t), \\ Q_{\eta_1}(age_1,\tau) &= \sum_{k=0}^{K} a_k^{\eta_1}(\tau)\varphi_k(age_1), \end{aligned}$$

with outcome-specific choices for K and the dictionary φ_k.

The sieve quantile model (20) provides a flexible specification of the conditional distribution of η_{it} given $\eta_{i,t-1}$ and age. Similarly, our quantile specifications flexibly models how ε_{it} and η_{i1} depend on age, at every quantile. We include the age of the household head as an additional control, while ruling out dependence on calendar time. This choice is motivated by our desire to model life-cycle evolution, as well as by the relative stationarity of the earnings distributions during the 1999-2009 period that we consider.

Note that the identification argument of Section 4.1 allows to nonparametrically recover, for each cohort entering the sample at age j, the distributions of ε at ages $j+2$, $j+4$, $j+6$, and $j+8$ (based on biennial data). Now, in our dataset $30 \leq j \leq 60$. Pooling across cohorts, we obtain that the distributions of ε are nonparametrically identified at all ages between 32 and 63 years. In turn, the joint distribution of η's is nonparametrically identified as well in this age range. Identification at ages $30, 31$ and $64, 65$ intuitively comes from parametric extrapolation using the quantile models.

Consumption rule. Turning to consumption, we now explain how we specify the conditional distribution of consumption given current assets and earnings components, and the distribution of initial assets conditional on the initial persistent component.

We specify

$$\begin{aligned} g_t(a_t, \eta_t, \varepsilon_t, \tau) &= g(a_t, \eta_t, \varepsilon_t, age_t, \tau) \\ &= \sum_{k=1}^{K} b_k^g \widetilde{\varphi}_k(a_t, \eta_t, \varepsilon_t, age_t) + b_0^g(\tau), \end{aligned} \tag{21}$$

where $\widetilde{\varphi}_k$ is a dictionary of functions defined on $\mathbb{R}^4$.

In addition, we model the conditional quantile of a_{i1} given η_{i1} and age_{i1} as

$$Q^{(a)}(\eta_1, age_1, \tau) = \sum_{k=0}^{K} b_k^a(\tau)\widetilde{\varphi}_k(\eta_1, age_1), \tag{22}$$

for different choices of K and of the dictionary $\widetilde{\varphi}_k$.

Equation (21) is a nonlinear regression model. In contrast with (20) and (22), the consumption model is additive in τ. It would be conceptually straightforward to let all coefficients b_k^g depend on τ, although this would lead to a less parsimonious specification.

Implementation. The functions a_k^Q, a_k^ε and $a_k^{\eta_1}$ are indexed by a finite-dimensional parameter vector θ. Likewise, the functions b_0^g and b_k^a are indexed by a parameter vector μ that also contains b_1^g,...,b_K^g.

Following Wei and Carroll (2009), we model the functions a_k^Q as piecewise-polynomial interpolating splines on a grid $[\tau_1, \tau_2]$, $[\tau_2, \tau_3]$,... , $[\tau_{L-1}, \tau_L]$, contained in the unit interval. We extend the specification of the intercept coefficient a_0^Q on $(0, \tau_1]$ and $[\tau_L, 1)$ using a parametric model indexed by λ^Q. All a_k^Q for $k \geq 1$ are constant on $[0, \tau_1]$ and $[\tau_L, 1]$, respectively. Hence, denoting $a_{k\ell}^Q = a_k^Q(\tau_\ell)$, the functions a_k^Q depend on $\{a_{11}^Q, ..., a_{KL}^Q, \lambda^Q\}$.

In practice, we take $L = 11$ and $\tau_\ell = \ell/L + 1$. The functions a_k^Q are taken as piecewise-linear on $[\tau_1, \tau_L]$. An advantage of this specification is that the likelihood function is available in closed form. In addition, we specify a_0^Q as the quantile of an exponential distribution on $(0, \tau_1]$ (with parameter λ_-^Q) and $[\tau_L, 1)$ (with parameter λ_+^Q). As a result, we have

$$\begin{aligned} a_k^Q(\tau) &= \frac{1}{\lambda_-^Q}\log\left(\frac{\tau}{\tau_1}\right)\mathbf{1}\{0 < \tau < \tau_1\} + \sum_{\ell=1}^{L-1}\left(a_{k\ell}^Q + \frac{a_{k,\ell+1}^Q - a_{k\ell}^Q}{\tau_{\ell+1} - \tau_\ell}(\tau - \tau_\ell)\right)\mathbf{1}\{\tau_\ell \leq \tau < \tau_{\ell+1}\} \\ &\quad - \frac{1}{\lambda_+^Q}\log\left(\frac{1-\tau}{1-\tau_L}\right)\mathbf{1}\{\tau_L \leq \tau < 1\}. \end{aligned}$$

We proceed similarly to model a_k^{ε}, $a_k^{\eta_1}$, and b_k^a. Moreover, as our data show little evidence against consumption being log-normal, we set $b_0(\tau)$ to $\alpha + \sigma\Phi^{-1}(\tau)$, where (α, σ) are parameters to be estimated. Lastly, we use tensor products of Hermite polynomials for φ_k and $\widetilde{\varphi}_k$. Each component of the product takes as argument a standardized variable.[15]

6.3 Estimation algorithm

The algorithm is an adaptation of techniques developed in Arellano and Bonhomme (2013) in the context of quantile regression with time-invariant unobserved heterogeneity. The first estimation step recovers estimates of the earnings parameters θ. The second step recovers estimates of the consumption parameters μ, given a previous estimate of θ. Our choice of a sequential estimation strategy, rather than joint estimation of (θ, μ), is motivated by the fact that θ is identified from the earnings process alone.

Model's restrictions. Let $\rho_\tau(u) = u(\tau - \mathbf{1}\{u \leq 0\})$ denote the "check" function of quantile regression (Koenker and Bassett, 1978). Let also $\overline{\theta}$ denote the true value of θ, and let

$$f_i(\eta_i^T; \overline{\theta}) = f(\eta_i^T | y_i^T, age_i^T; \overline{\theta})$$

denote the posterior density of $(\eta_{i1}, ..., \eta_{iT})$ given the earnings data. As the earnings model is fully specified, f_i is a known function of $\overline{\theta}$.

We start by noting that, for all $t \geq 2$ and $\ell \in \{1, ..., L\}$,

$$\left(\overline{a}_{0\ell}^Q, ..., \overline{a}_{K\ell}^Q\right) = \underset{\left(a_{0\ell}^Q, ..., a_{K\ell}^Q\right)}{\text{argmin}} \ \mathbb{E}\left[\int \rho_{\tau_\ell}\left(\eta_{it} - \sum_{k=0}^{K} a_{k\ell}^Q \varphi_k(\eta_{i,t-1}, age_{it})\right) f_i(\eta_i^T; \overline{\theta}) d\eta_i^T\right], \qquad (23)$$

where $\overline{a}_{k\ell}^Q$ denotes the true value of $a_{k\ell}^Q = a_k^Q(\tau_\ell)$.

To see that (23) holds, note that the objective function is smooth (due to the presence of the integrals) and convex (because of the "check" function). Moreover, the first-order conditions are, for all $k \in \{0, ..., K\}$,

$$\mathbb{E}\left[\ \mathbb{E}\left[\varphi_k(\eta_{i,t-1}, age_{it})\mathbf{1}\left\{\eta_{it} \leq \sum_{k=0}^{K} \overline{a}_{k\ell}^Q \varphi_k(\eta_{i,t-1}, age_{it})\right\} \middle| y_i^T, age_i^T\right]\ \right]$$
$$= \mathbb{E}\left[\varphi_k(\eta_{i,t-1}, age_{it})\mathbf{1}\left\{\eta_{it} \leq \sum_{k=0}^{K} \overline{a}_{k\ell}^Q \varphi_k(\eta_{i,t-1}, age_{it})\right\}\right] = \tau_\ell,$$

[15] For example, $(a_t - mean(a))/std(a)$, $(\eta_t - mean(y))/std(y)$, $(\varepsilon_t - mean(y))/std(y)$, and $(age_t - mean(age))/std(age)$ are used as arguments of the consumption rule.

where we have used that, by (20),

$$\mathbb{E}\left[\mathbf{1}\left\{\eta_{it}\leq \sum_{k=0}^{K}\overline{a}_{k\ell}^{Q}\varphi_k(\eta_{i,t-1},age_{it})\right\}\middle|\eta_i^{t-1},age_i^T\right]=\tau_\ell.$$

Likewise, we have, for all $t\geq 1$ and all ℓ,

$$(\overline{a}_{0\ell}^{\varepsilon},...,\overline{a}_{K\ell}^{\varepsilon})=\underset{\left(a_{0\ell}^{\varepsilon},...,a_{K\ell}^{\varepsilon}\right)}{\text{argmin}}\ \mathbb{E}\left[\int\rho_{\tau_\ell}\left(y_{it}-\eta_{it}-\sum_{k=0}^{K}a_{k\ell}^{\varepsilon}\varphi_k(age_t)\right)f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right],\qquad(24)$$

and, for all ℓ,

$$(\overline{a}_{0\ell}^{\eta_1},...,\overline{a}_{K\ell}^{\eta_1})=\underset{\left(a_{0\ell}^{\eta_1},...,a_{K\ell}^{\eta_1}\right)}{\text{argmin}}\ \mathbb{E}\left[\int\rho_{\tau_\ell}\left(\eta_{i1}-\sum_{k=0}^{K}a_{k\ell}^{\eta_1}\varphi_k(age_{i1})\right)f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right].\qquad(25)$$

In addition to (23)-(24)-(25), the model implies other restrictions on the tail parameters λ, which are given in appendix. All the restrictions depend on the posterior density f_i. Given the use of piecewise-linear splines, the joint likelihood function of $(\eta_i^T,y_i^T|age_i^T;\overline{\theta})$ is available in closed form, and we provide an explicit expression in the appendix. In practice, this means that it is easy to simulate from f_i. We take advantage of this feature in our estimation algorithm below.

Turning to consumption we have, for all $t\geq 1$,

$$\left(\overline{\alpha},\overline{b}_1^g,...,\overline{b}_K^g\right)=\underset{\left(\alpha,b_1^g,...,b_K^g\right)}{\text{argmin}}\ \mathbb{E}\left[\int\left(c_{it}-\alpha-\sum_{k=1}^{K}b_k^g\widetilde{\varphi}_k(a_{it},\eta_{it},y_{it}-\eta_{it},age_{it})\right)^2g_i(\eta_i^T;\overline{\theta},\overline{\mu})d\eta_i^T\right],$$

where

$$g_i(\eta_i^T;\overline{\theta},\overline{\mu})=f(\eta_i^T|c_i^T,a_i^T,y_i^T,age_i^T;\overline{\theta},\overline{\mu})$$

denotes the posterior density of $(\eta_{i1},...,\eta_{iT})$ given the earnings, consumption, and asset data.

Moreover, the variance of taste shifters satisfies

$$\overline{\sigma}^2\ =\ \mathbb{E}\left[\int\left(c_{it}-\overline{\alpha}-\sum_{k=1}^{K}b_k^g\widetilde{\varphi}_k(a_{it},\eta_{it},y_{it}-\eta_{it},age_{it})\right)^2g_i(\eta_i^T;\overline{\theta},\overline{\mu})d\eta_i^T\right].$$

Lastly we have, for all ℓ,

$$\left(\overline{b}_{0\ell}^a,...,\overline{b}_{K\ell}^a\right)=\underset{\left(b_{0\ell}^a,...,b_{K\ell}^a\right)}{\text{argmin}}\ \mathbb{E}\left[\int\rho_{\tau_\ell}\left(a_{i1}-\sum_{k=0}^{K}b_{k\ell}^a\widetilde{\varphi}_k(\eta_{i1},age_{i1})\right)g_i(\eta_i^T;\overline{\theta},\overline{\mu})d\eta_i^T\right],$$

with additional restrictions characterizing tail parameters given in the appendix.

Overview of the algorithm. Here we describe the estimation of the earnings parameters θ. Estimation of the consumption parameters μ is similar. The estimation algorithms are described in detail in the appendix.

A compact notation for the restrictions implied by the earnings model is

$$\overline{\theta} = \underset{\theta}{\text{argmin}} \ \mathbb{E}\left[\int R(y_i, \eta; \theta) f_i(\eta; \overline{\theta}) d\eta\right],$$

where R is a known function and $\overline{\theta}$ denotes the true value of θ.

Our estimation algorithm is closely related to a "stochastic EM" algorithm of Celeux and Diebolt (1993). Stochastic EM is a simulated version of the classical EM algorithm of Dempster *et al.* (1977), where new draws from η are computed in every iteration of the algorithm.[16] One difference is that, unlike in EM, our problem is not likelihood-based. Instead, we exploit the computational convenience of quantile regression and replace likelihood maximization by a sequence of quantile regressions in each M-step of the algorithm. We explain below how this difference affects standard errors calculation.

Starting with a parameter vector $\widehat{\theta}^{(0)}$, we iterate the following two steps on $s = 0, 1, 2, ...$ until convergence of the $\widehat{\theta}^{(s)}$ process:

1. *Stochastic E-step:* Draw $\eta_i^{(m)} = (\eta_{i1}^{(m)}, ..., \eta_{iT}^{(m)})$ for $m = 1, ..., M$ from $f_i(\cdot; \widehat{\theta}^{(s)})$.

2. *M-step:* Compute
$$\widehat{\theta}^{(s+1)} = \underset{\theta}{\text{argmin}} \ \sum_{i=1}^{N} \sum_{m=1}^{M} R(y_i, \eta_i^{(m)}; \theta).$$

Note that, as the likelihood function is available in closed form, the E-step is straightforward. In practice we use a random-walk Metropolis-Hastings sampler for this purpose. The M-step consists of a number of ordinary regressions and quantile regressions. For example, the parameters $a_{k\ell}^Q$ are updated as

$$\min_{(a_{0\ell}^Q, ..., a_{K\ell}^Q)} \sum_{i=1}^{N} \sum_{t=2}^{T} \sum_{m=1}^{M} \rho_{\tau_\ell}\left(\eta_{it}^{(m)} - \sum_{k=0}^{K} a_{k\ell}^Q \varphi_k(\eta_{i,t-1}^{(m)}, age_{it})\right), \ \ell = 1, ..., L,$$

which is a set of standard quantile regressions, with convex objectives. We proceed in a similar way to update all other parameters. See the appendix for details.

[16] Nielsen (2000b) compares the "stochastic EM" algorithm with the "simulated EM" algorithm of McFadden and Ruud (1994), where in contrast the same underlying uniform draws are re-used in every iteration.

Properties. Nielsen (2000a) studies the statistical properties of the stochastic EM algorithm in a likelihood case. He provides conditions under which the Markov Chain $\widehat{\theta}^{(s)}$ is ergodic, for a fixed sample size. In addition, he also characterizes the asymptotic distribution of $\widehat{\theta}^{(s)}$ as N increases. Specifically, he shows that the Markov Chain $\sqrt{N}\left(\widehat{\theta}^{(s)}-\overline{\theta}\right)$ tends to an autoregressive Gaussian process as N tends to infinity, and provides analytical expressions for its asymptotic variance.

In the appendix we rely on Nielsen's work to characterize the asymptotic distribution of $\widehat{\theta}^{(s)}$ in our model, where the optimization step is not likelihood-based but based on different estimating equations. If s corresponds to a draw from the ergodic distribution of the Markov Chain, then

$$\sqrt{N}\left(\widehat{\theta}^{(s)}-\overline{\theta}\right) \xrightarrow{d} \mathcal{N}(0, V_{\theta}),$$

where the expression of V_{θ} is given in the appendix.

In practice, we stop the chain after a large number of iterations and we report an average across the last $\widetilde{S}$ values $\widehat{\theta}=\frac{1}{S-\widetilde{S}}\sum_{s=\widetilde{S}+1}^{S}\widehat{\theta}^{(s)}$. Then, given estimates of the earnings parameters $\widehat{\theta}$, we use a similar algorithm to estimate the consumption-related parameters $\widehat{\mu}$. The form of the asymptotic variance of $\widehat{\mu}$ accounts for the sequential nature of the estimation.

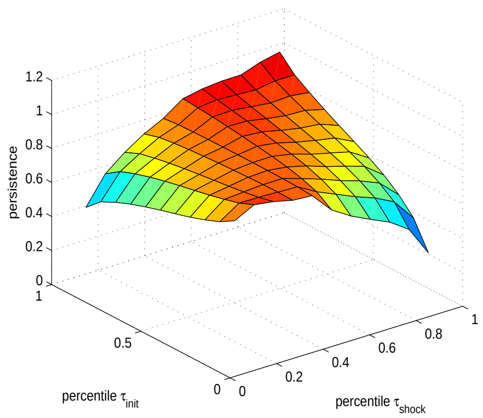
Finally, note that, in the above asymptotic analysis, we have let the sample size N increase for fixed values of K and L. This amounts to assuming that the fixed-K, L parametric model is well-specified. An alternative, nonparametric approach, would be to let K and L increase with N at an appropriate rate so as to let the approximation bias tend to zero. Arellano and Bonhomme (2013) show consistency in a case where K is fixed and $L = L_N$ tends to infinity with N. Belloni, Chernozhukov and Fernandez-Val (2012) consider a quantile regression model for fixed τ and $K = K_N$ tends to infinity, and derive inference results. Studying inference in our problem as (N, K, L) jointly tend to infinity is an interesting avenue for future work.

7 Preliminary empirical results

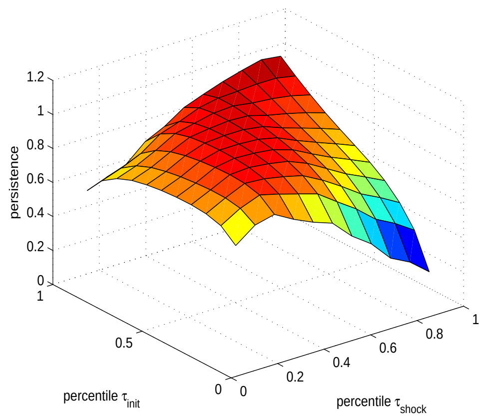
Earnings. We start by commenting on the empirical estimates of the earnings process. Figure 1 (a) shows estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$ in the PSID sample. To compute these numbers, we use a modelling analogous to (20), except that the dependent variable is y_{it} and the conditioning set only includes $y_{i,t-1}$. The specification is based on a third-order Hermite polynomial.

Figure 1: Nonlinear persistence

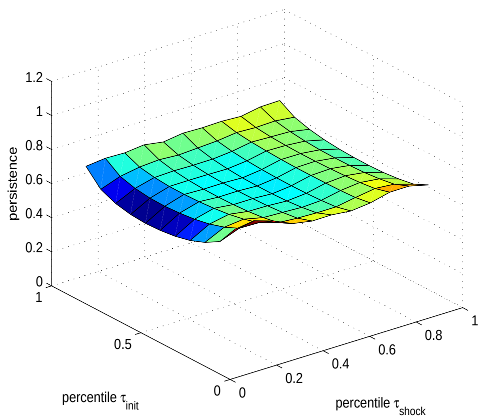
(a) Earnings, PSID data



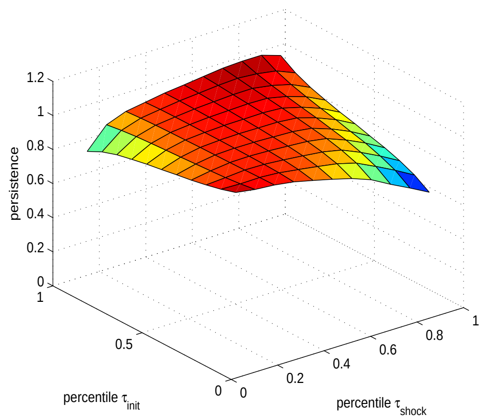
(b) Earnings, nonlinear model



(c) Earnings, canonical model



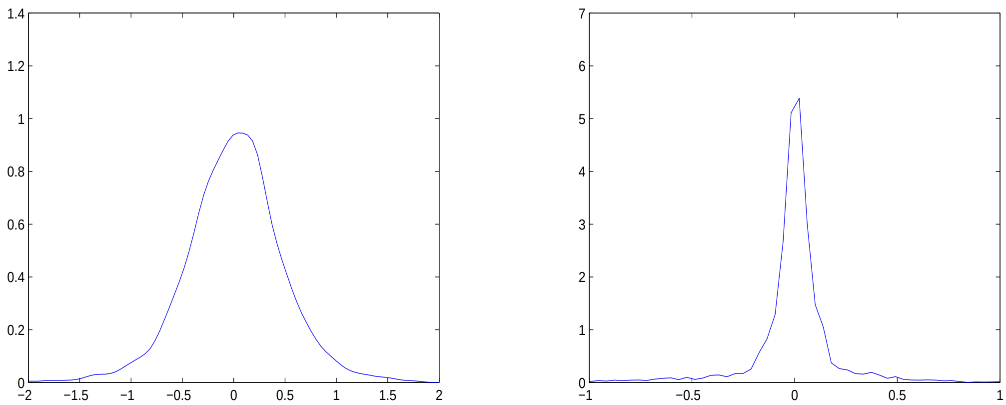
(d) Persistent component η_{it}, nonlinear model



Note: Graphs (a), (b), and (c) show estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the distribution of $y_{i,t-1}$. Graph (a) is based on the PSID data, graph (b) is based on data simulated according to our nonlinear earnings model with parameters set to their estimated values, and graph (c) is based on data simulated according to the canonical random walk earnings model. Graph (d) shows estimates of the average derivative of the conditional quantile function of η_{it} on $\eta_{i,t-1}$ with respect to $\eta_{i,t-1}$, based on estimates from the nonlinear earnings model.

Figure 2: Densities of persistent and transitory earnings components

(a) Persistent component η_{it} (b) Transitory component ε_{it}



Note: Nonparametric estimates of densities based on simulated data according to the nonlinear model, using a Gaussian kernel.

Figure 1 (a) shows clear evidence of nonlinear persistence. Persistence is highest when high earnings households (that is, high τ_{init}) are hit by a good shock (high τ_{shock}), and when low earnings households (that is, low τ_{init}) are hit by a bad shock (low τ_{shock}). In both cases, estimated persistence is close to $.9 - 1$. In contrast, bad shocks hitting high-earnings households, and good shocks hitting low-earnings ones, are associated with much lower persistence, as low as $.3 - .4$.

Figure 1 (b), which is based on simulated data, shows that our nonlinear model reproduces these patterns well. In contrast, standard models have difficulty fitting this empirical evidence. For example, we estimated a simple version of the canonical earnings dynamics model with a random walk component and independent transitory shocks.[17] Figure 1 (c) shows that the average derivative of the quantile function is approximately constant (up to simulation error) with respect to τ_{shock} and τ_{init}. This stands in sharp contrast with the data, and suggests that interaction effects between earnings shocks and past earnings components are key.[18]

Figure 1 (d) then shows the estimated persistence of the earnings component η_{it} in model (1)-(2). Specifically, the graph shows $\rho_t(\eta_{i,t-1}, \tau)$ from equation (3), evaluated at percentiles

[17] Estimation is based on equally-weighted minimum distance using the covariance structure predicted by the canonical model.

[18] Models with variance dynamics such as Meghir and Pistaferri (2004) do not seem able to reproduce the nonlinear asymmetric effects apparent in Figure 1 (a) either.

τ_{init} and τ_{shock} and at the mean age in the sample (47.5 years). Persistence in η_{it} is higher than persistence in y_{it}, consistently with the fact that Figure 1 (d) is net of transitory shocks. Persistence is close to 1 for high earnings households hit by good shocks, and for low earnings households hit by bad shocks. At the same time, persistence is lower, down to $.6 - .8$, when bad shocks hit high-earnings households or good shocks hit low-earnings ones.

Figure 2 shows estimates of the marginal distributions of the persistent and transitory earnings components. While the persistent component η_{it} shows small departures from Gaussianity, the density of ε_{it} is clearly non-normal and presents high kurtosis and fat tails. These results are qualitatively consistent with empirical estimates of non-Gaussian linear models in Horowitz and Markatou (1996) and Bonhomme and Robin (2010).

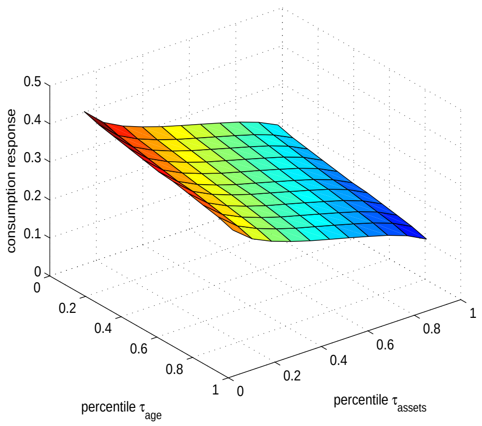
Consumption. We next turn to consumption-related parameters. Figure 3 (a) shows estimates of the average derivative of the conditional mean of c_{it} given y_{it}, a_{it} and age_{it} with respect to y_{it}. The function is evaluated at percentiles τ_{assets} and τ_{age}, and averaged over y_{it}. We use tensor products of Hermite polynomials with degrees $(2, 2, 1)$. The consumption coefficient is about .3 on average. Moreover, the results suggest that the consumption of older households, and of households with higher assets, responds less to variations in earnings. Figure 3 (b) shows the same response surface based on simulated data from our full nonlinear model of earnings and consumption. The fit of the model, though not perfect, seems reasonable.

Note that our model fits poorly the dynamics of consumption (not shown). While the covariances between earnings and consumption levels are well reproduced, the model underestimates the first and higher-order autocorrelations in consumption. Allowing for household unobserved heterogeneity in consumption, as outlined in Section 5, should help improve the fit.

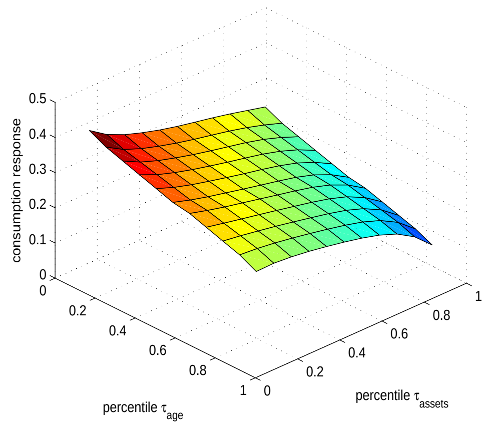
Figure 3 (c) shows estimates of the average consumption response $\overline{\phi}_t(a)$ to variations in the persistent component of earnings, see equation (11). This parameter measures the degree of consumption insurance to persistent earnings shocks, as a function of age and assets. On average this parameter is about .4, suggesting that more than half of earnings fluctuations is effectively insured. Moreover, variation in assets and age suggests the presence of an interaction effect. In particular, older households with high assets seem virtually fully insured against earnings fluctuations.

Figure 3: Consumption responses to earnings shocks, by assets and age

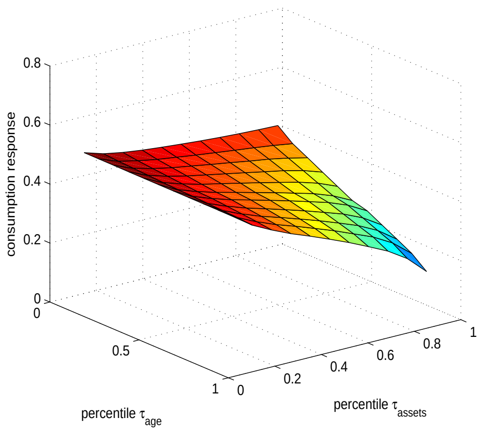
(a) Consumption response to earnings
PSID data



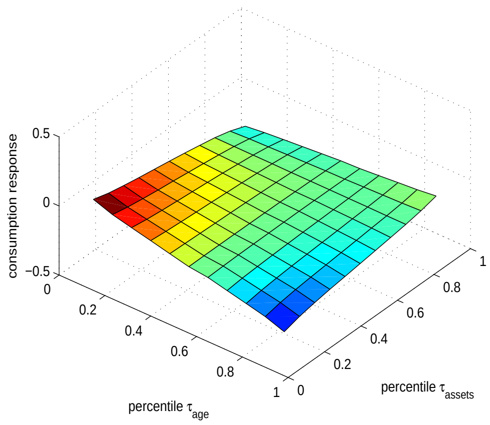
(b) Consumption response to earnings
Nonlinear model



(c) Consumption response to η_{it}
Nonlinear model



(d) Consumption response to ε_{it}
Nonlinear model



Note: Graphs (a) and (b) show estimates of the average derivative of the conditional mean of c_{it} given y_{it}, a_{it} and age_{it} with respect to y_{it}, evaluated at values of a_{it} and age_{it} that corresponds to their τ_{assets} and τ_{age} percentiles, and averaged over the values of y_{it}. Graph (a) is based on the PSID data, and graph (b) is based on data simulated according to our nonlinear model with parameters set to their estimated values. Graphs (c) and (d) show estimates of the average consumption responses $\overline{\phi}_t(a)$ and $\overline{\psi}_t(a)$ to variations in η_{it} and ε_{it}, respectively, evaluated at τ_{assets} and τ_{age}; see equation (11).

In Figure 3 (d) we report estimates of $\overline{\psi}_t(a)$ to variations in the transitory component of earnings. The coefficient is negative, and shows little variation with age and assets. While consistent with some of the results obtained by Blundell, Pistaferri and Saporta-Eksten (2012), a negative response coefficient seems puzzling.

Model's simulation. Next, we simulate life-cycle earnings and consumption according to our nonlinear model, and document the evolution of earnings and consumption following a persistent earnings shock. In Figure 4 we report the difference between the earnings paths of two types of households: households that are hit at age 37 by either a large negative shock to the persistent earnings component ($\tau_{shock} = .10$), or by a large positive shock ($\tau_{shock} = .90$), and households that are hit by a median shock $\tau = .50$ to the persistent component. We report age-specific averages across 100,000 simulations of the model. At the start of the simulation (age 35) all households have the same persistent component indicated by the percentile τ_{init}.

In order to simulate consumption paths, we need to specify the evolution of assets. We estimate the following rule from the PSID data

$$a_{it} = d_0^a + d_1^a a_{i,t-1} + d_2^a y_{i,t-1} + d_3^a c_{i,t-1} + \upsilon_{it}, \ t = 2, ..., T, \tag{26}$$

and simulate assets according to (26) while assuming that υ_{it} is iid and normally distributed. It will be important to assess the sensitivity of our results to the assets accumulation rule.

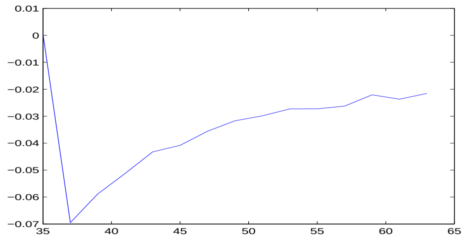
Earnings responses are consistent with the presence of strong interaction effects between the rank in the distribution of earnings component (τ_{init}) and the magnitude of the shock to the persistent component (τ_{shock}). While a large negative shock ($\tau_{shock} = .10$) is associated with a 7% drop in earnings for low and medium earnings households ($\tau_{init} = .10$ or $.5$), it is associated with a 17% drop for high-earnings households ($\tau_{init} = .90$). We also find strong interaction effects in the response to large positive shocks ($\tau_{shock} = .90$). Moreover, the persistence of these shocks over the life cycle also depends on the initial condition. For example, Figure 4 (e) shows a very slow recovery from a negative earnings shock when starting from a high-earnings position, while graph (a) shows a quicker recovery. Lastly, in graphs (g) and (h) we report results based on the canonical model, where interaction effects are absent by assumption. The implications of the nonlinear earnings model thus differ markedly from those of standard linear models.

Figure 4: Impulse responses, earnings

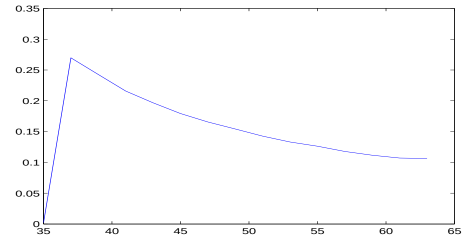
Nonlinear model

$\tau_{init} = .1$

(a) $\tau_{shock} = .1$

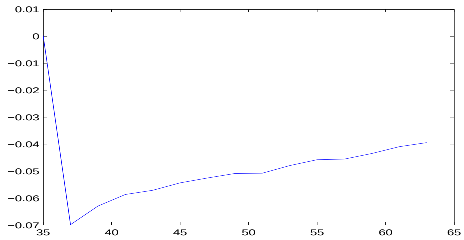


(b) $\tau_{shock} = .9$

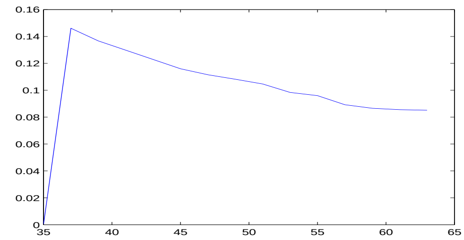


$\tau_{init} = .5$

(c) $\tau_{shock} = .1$

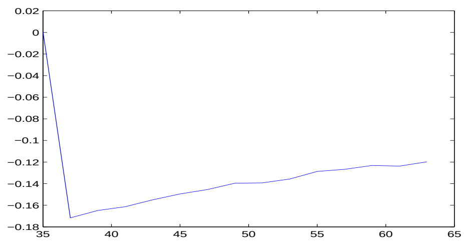


(d) $\tau_{shock} = .9$

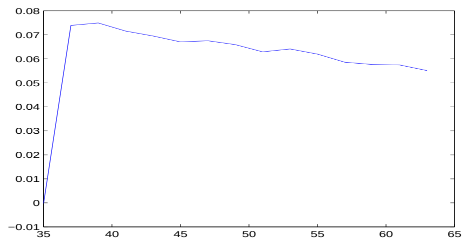


$\tau_{init} = .9$

(e) $\tau_{shock} = .1$

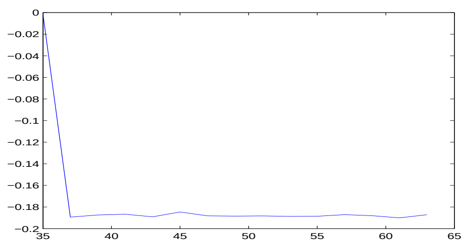


(f) $\tau_{shock} = .9$

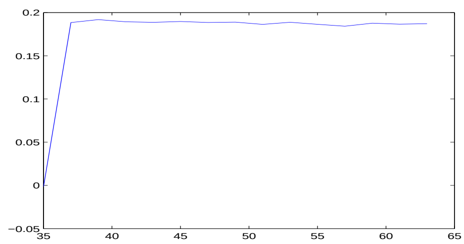


Canonical model

(g) $\tau_{shock} = .1$



(h) $\tau_{shock} = .9$



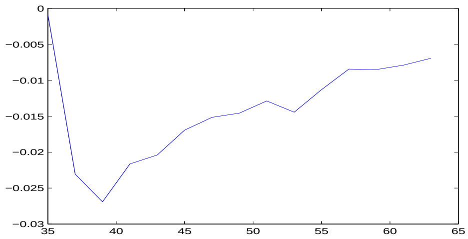
Note: Persistent component at percentile τ_{init} at age 35. The graphs show the difference between a household hit by a shock τ_{shock} at age 37, and a household hit by a .5 shock at the same age. Age-specific means across 100,000 simulations. Graphs (a) to (f) correspond to the nonlinear model. Graphs (g) and (h) correspond to the canonical model of earnings dynamics.

Figure 5: Impulse responses, consumption

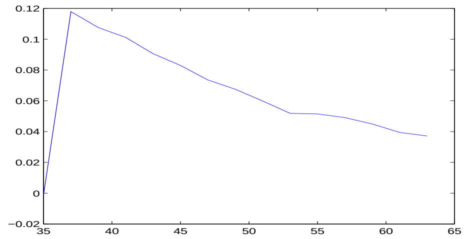
Nonlinear model

$\tau_{init} = .1$

(a) $\tau_{shock} = .1$

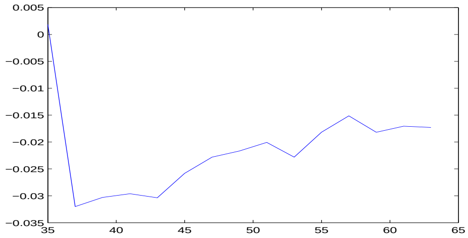


(b) $\tau_{shock} = .9$

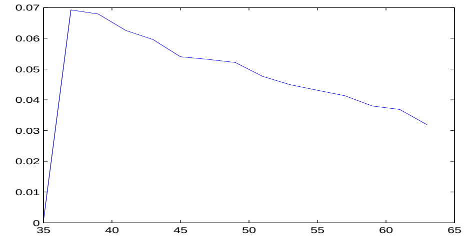


$\tau_{init} = .5$

(c) $\tau_{shock} = .1$

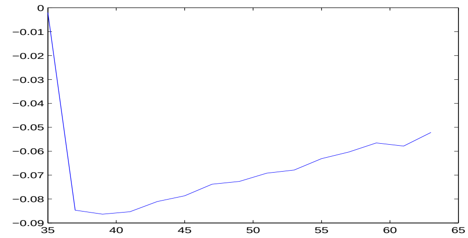


(d) $\tau_{shock} = .9$

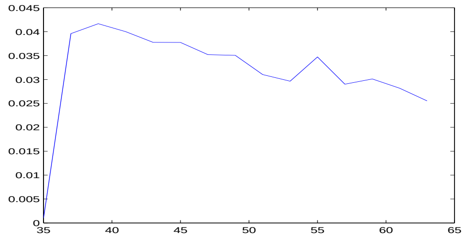


$\tau_{init} = .9$

(e) $\tau_{shock} = .1$

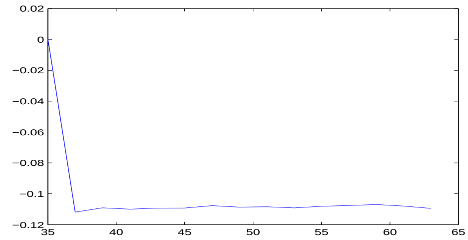


(f) $\tau_{shock} = .9$

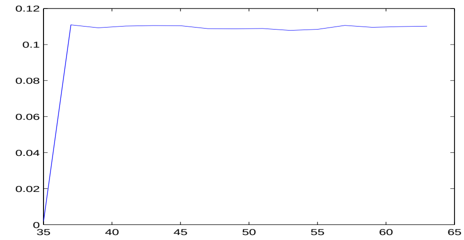


Canonical model

(g) $\tau_{shock} = .1$



(h) $\tau_{shock} = .9$



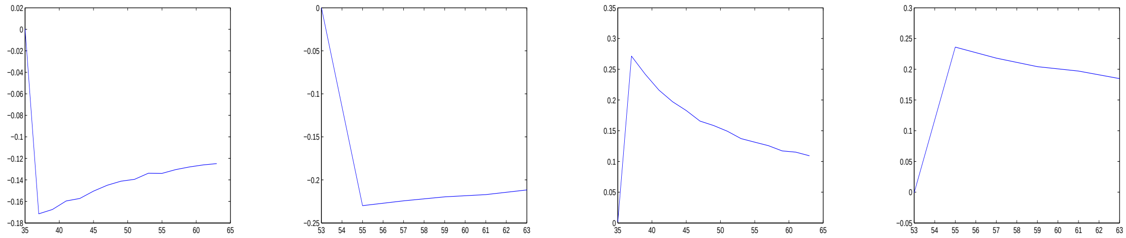
Note: Persistent component at percentile τ_{init} at age 35. The graphs show the difference between a household hit by a shock τ_{shock} at age 37, and a household hit by a .5 shock at the same age. Age-specific means across 100,000 simulations. Graphs (a) to (f) correspond to the nonlinear model. Graphs (g) and (h) correspond to the canonical model of earnings dynamics and a linear consumption rule.

Figure 6: Impulse responses, by age and initial assets

Earnings

$\tau_{init} = .9$, $\tau_{shock} = .1$ — (a) Young, (b) Old

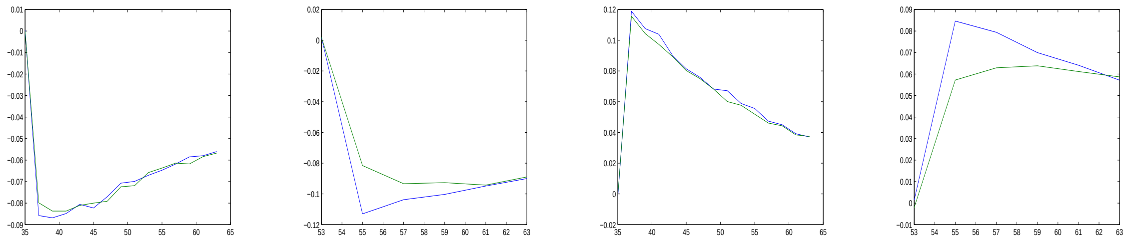
$\tau_{init} = .1$, $\tau_{shock} = .9$ — (c) Young, (d) Old



Consumption

$\tau_{init} = .9$, $\tau_{shock} = .1$ — (e) Young, (f) Old

$\tau_{init} = .1$, $\tau_{shock} = .9$ — (g) Young, (h) Old



Note: See notes to Figures 4 and 5. Initial assets at age 35 (for "young" households) or 53 (for "old" households) are at percentile .10 *(blue curves) and* .90 *(green curves).*

In Figure 5 we report the results of a similar exercise to Figure 4, but we now focus on consumption responses. We see that the nonlinearities observed in the earnings response matter for consumption too. For example, while a large negative shock ($\tau_{shock} = .10$) is associated with a 3% drop in consumption for low or medium earnings households, it is associated with a 9% drop for high-earnings households. We also observe differences in persistence across the different scenarios. In addition, on graphs (g) and (h) we report results based on the canonical model with a linear consumption rule.[19] Here also, interaction effects are absent by construction.

Lastly, in Figure 6 we perform similar exercises, while varying the timing of shocks and the asset holdings that individuals possess. Graphs (a) to (d) suggest that a negative shock ($\tau_{shock} = .10$) for high-earnings households has a higher impact on earnings at later ages: the earnings drop is 23% when the shock hits at age 55, compared to 17% when is hits at age 37. The impact of positive shocks for low earnings individuals does not seem to vary with age.

Graphs (e) to (h) show the consumption responses. The results suggest that the consumption of older households responds less than the one of younger households. Moreover, while the presence of asset holdings does not seem to affect the insurability of earnings shocks for younger households, it does seem to attenuate the consumption response for households who are hit later in the life-cycle. These results are consistent with the estimates of $\overline{\phi}_t(a)$ reported in Figure 3.

8 Conclusion

Our nonlinear model sheds new light on the nonlinear transmission of earnings shocks and the nature of consumption insurance. It also provides a framework to assess the suitability of existing life-cycle models of consumption and savings, and potentially help guide the development of new models. The next step on our agenda is to assess the consumption responses to our nonlinear earnings process in the context of a standard life-cycle model, and to document to which extent the model's predictions are consistent with the empirical facts that we have uncovered.

[19]Specifically, c_{it} is modelled as a linear function of η_{it}, ε_{it}, and an independent additive error term i.i.d. over time. The model is estimated by equally-weighted minimum distance.

References

[1] TO BE COMPLETED

[2] Abowd, John, and David Card (1989): "On the Covariance Structure of Earnings and Hours Changes." *Econometrica*, 57(2): 411-45.

[3] Arellano, M. and S. Bonhomme (2013): "Random-Effects Quantile Regression," unpublished manuscript.

[4] Attanasio, Orazio, and Steven Davis (1996): "Relative Wage Movements and the Distribution of Consumption." *Journal of Political Economy*, 104(6): 1227-62.

[5] Blundell, Richard, Luigi Pistaferri and Ian Preston (2008): "Consumption Inequality and Partial Insurance," *American Economic Review*, 98(5): 1887-1921.

[6] Blundell, Richard, Luigi Pistaferri and Itay Saporta-Eksten (2012): "Consumption Smoothing and Family Labor Supply," NBER Discussion Paper, November.

[7] Blundell, Richard, and Ian Preston. (1998): "Consumption Inequality and Income Uncertainty." *Quarterly Journal of Economics*, 113(2): 603-640.

[8] Celeux, G., and J. Diebolt (1993): "Asymptotic Properties of a Stochastic EM Algorithm for Estimating Mixing Proportions," *Comm. Statist. Stochastic Models*, 9, 599-613.

[9] Gourinchas, P.O. and J. A. Parker (2002): "Consumption over the Life Cycle." *Econometrica*, 70, 47-91.

[10] Guvenen, F. (2007): "Learning your Earning: Are Labor Income Shocks Really Very Persistent?" *American Economic Review*, 97(3), 687-712.

[11] Guvenen, F., S. Ozcan, and J. Song (2012): "The Nature of Countercyclical Income Risk," to appear in the *Journal of Political Economy.*

[12] Guvenen, F., and A. Smith (2010): "Inferring Labor Income Risk from Economic Choices: An Indirect Inference Approach," Staff Report 450, Federal Reserve Bank of Minneapolis.

[13] Hall, Robert, and Frederic Mishkin (1982): "The sensitivity of Consumption to Transitory Income: Estimates from Panel Data of Households." *Econometrica*, 50(2): 261-81.

[14] Heathcote, J., K. Storesletten, and G. Violante (2013): "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," to appear in the *American Economic Review.*

[15] Hu, Y. and S. M. Schennach (2008): "Instrumental Variable Treatment of Nonclassical Measurement Error Models," *Econometrica*, 76, 195–216.

[16] Hu, Y. and M. Shum (2012): "Nonparametric Identification of Dynamic Models with Unobserved State Variables," *Journal of Econometrics*, 171, 32–44.

[17] Huggett, M. (1993), "The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies". *Journal of Economic Dynamics and Control*, 17, 953-969.

[18] Kaplan, Greg and Giovanni Violante (2010), "How Much Consumption Insurance Beyond Self-Insurance." *American Economic Journal*, 2(4), 53-87.

[19] Koenker, R. and G. J. Bassett (1978): "Regression quantiles," *Econometrica*, 46, 33-50.

[20] McLachlan, G. J., and T. Krishnan (2008): *The EM Algorithm and Extensions*, New York, Wiley, Second Edition.

[21] Meghir, Costas, and Luigi Pistaferri, (2004), "Income Variance Dynamics and Heterogeneity." *Econometrica*, 72(1): 1-32.

[22] Meghir, Costas and Luigi Pistaferri, (2011), 'Earnings, Consumption and Life Cycle Choices,' Handbook of Labor Economics, Elsevier.

[23] Nielsen, S. F. (2000a): "The Stochastic EM Algorithm: Estimation and Asymptotic Results," *Bernouilli*, 6(3): 457-489.

[24] Nielsen, S. F. (2000b): "On Simulated EM Algorithms," *Journal of Econometrics*, 96: 267-292.

[25] Wei, Y. and R. J. Carroll (2009): "Quantile Regression with Measurement Error," *Journal of the American Statistical Association*, 104, 1129–1143.

[26] Wilhelm, D. (2012): "Identification and Estimation of Nonparametric Panel Data Regressions with Measurement Error," unpublished manuscript.

APPENDIX

A Sketch of the argument in Wilhelm (2012)

We consider model (1)-(2) with $T = 3$. Let $L^2(f)$ denote the set of squared-integrable functions with respect to a weight function f. We define $\mathcal{L}_{y_2|y_1}$ as the linear operator such that $\mathcal{L}_{y_2|y_1}h(a) = \mathbb{E}[h(y_2)|y_1 = a] \in L^2(f_{y_1})$ for every function $h \in L^2(f_{y_2})$. Similarly, let $\mathcal{L}_{\eta_2|y_1}$ be such that $\mathcal{L}_{\eta_2|y_1}h(a) = \mathbb{E}[h(\eta_2)|y_1 = a] \in L^2(f_{y_1})$ for every function $h \in L^2(f_{\eta_2})$. We denote as $\mathcal{R}\left(\mathcal{L}_{y_2|y_1}\right)$ the range of $\mathcal{L}_{y_2|y_1}$, that is

$$\mathcal{R}\left(\mathcal{L}_{y_2|y_1}\right) = \{k \in L^2(f_{y_1}), \text{ s.t. } k = \mathcal{L}_{y_2|y_1}h \text{ for some } h \in L^2(f_{y_2})\}.$$

Assumption 1
(i) $\mathcal{L}_{y_2|y_1}$ and $\mathcal{L}_{\eta_2|y_1}$ are injective.
(ii) There exists a function $h \in L^2(f_{y_3})$ such that

$$\mathbb{E}[h(y_3)|y_1 = a] \in \mathcal{R}\left(\mathcal{L}_{y_2|y_1}\right), \text{ and} \tag{A1}$$
$$\mathbb{E}[y_2 h(y_3)|y_1 = a] \in \mathcal{R}\left(\mathcal{L}_{y_2|y_1}\right). \tag{A2}$$

Thus, there exist s_1 and s_2 in $L^2(f_{y_2})$ such that

$$\mathbb{E}[h(y_3)|y_1 = \cdot] = \mathcal{L}_{y_2|y_1}s_1, \text{ and } \mathbb{E}[y_2 h(y_3)|y_1 = \cdot] = \mathcal{L}_{y_2|y_1}s_2.$$

(iii) Let $\widetilde{s}_1(y) = ys_1(y)$. The Fourier transforms $\mathcal{F}(s_1)$, $\mathcal{F}(\widetilde{s}_1)$, and $\mathcal{F}(s_2)$ (where $\mathcal{F}(h)(u) = \int h(x)e^{iux}dx$) are ordinary functions. Moreover, $\mathcal{F}(s_1)(u) \neq 0$ for all $u \in \mathbb{R}$.

Part (i) is an injectivity/completeness condition. Part (ii) is not standard. It is related to the existence problem in nonparametric instrumental variables. Horowitz (2009) proposes a test for (A1) in the case where $\mathcal{L}_{y_2|y_1}$ is a compact operator. Part (iii) is a high-level assumption; see Wilhelm (2012) for more primitive conditions.

By Assumption 1-(ii) we have, almost surely in y_1,

$$\mathbb{E}[h(y_3)|y_1] = \mathbb{E}[s_1(y_2)|y_1],$$
$$\mathbb{E}[y_2 h(y_3)|y_1] = \mathbb{E}[s_2(y_2)|y_1].$$

Moreover, s_1 and s_2 are the unique solutions to these equations by Assumption 1-(i).

Hence, given the model's assumptions

$$\mathbb{E}[\mathbb{E}(h(y_3)|\eta_2)|y_1] = \mathbb{E}[\mathbb{E}(s_1(y_2)|\eta_2)|y_1] \text{ a.s.}$$

It thus follows from the injectivity of $\mathcal{L}_{\eta_2|y_1}$ in Assumption 1-(i) that, almost surely in η_2,

$$\mathbb{E}[h(y_3)|\eta_2] = \mathbb{E}[s_1(y_2)|\eta_2]. \tag{A3}$$

Likewise, $\mathbb{E}[y_2 h(y_3)|\eta_2] = \mathbb{E}[s_2(y_2)|\eta_2]$. Hence

$$\eta_2 \mathbb{E}[h(y_3)|\eta_2] = \mathbb{E}[s_2(y_2)|\eta_2] \text{ a.s.} \tag{A4}$$

Combining (A3) and (A4), we obtain

$$\eta_2 \mathbb{E}[s_1(y_2)|\eta_2] = \mathbb{E}[s_2(y_2)|\eta_2] \text{ a.s.}$$

That is, almost surely in η_2,

$$\eta_2 \int s_1(y) f_{\varepsilon_2}(y-\eta_2)dy = \int s_2(y) f_{\varepsilon_2}(y-\eta_2)dy. \tag{A5}$$

The functional equation (A5) depends on s_1 and s_2, which are both uniquely determined given the data generating process, and on the unknown f_{ε_2}. By Assumption 1-(*iii*) we can take Fourier transforms and obtain

$$\boldsymbol{i}\mathcal{F}(s_1)(u)\frac{d\psi_{\varepsilon_2}(-u)}{du} + \mathcal{F}(\widetilde{s}_1)(u)\psi_{\varepsilon_2}(-u) = \mathcal{F}(s_2)(u)\psi_{\varepsilon_2}(-u), \tag{A6}$$

where $\psi_{\varepsilon_2}(u) = \mathcal{F}(f_{\varepsilon_2})(u)$ is the characteristic function of ε_2.

Noting that $\psi_{\varepsilon_2}(0) = 1$, (A6) can be solved in closed form for $\psi_{\varepsilon_2}(\cdot)$, because $\mathcal{F}(s_1)(u) \neq 0$ for all u by Assumption 1-(*iii*). This shows that the characteristic function of ε_{i2}, and hence its distribution function, are identified.

B Inference for Stochastic EM

We rewrite the moment restrictions implied by the earnings problem in a compact notation

$$E[\Psi(y,\eta;\overline{\theta})] = 0,$$

where y is observed, η is latent, and θ (with true value $\overline{\theta}$) is a finite-dimensional parameter vector of same dimension as Ψ. The consumption problem may be written in a similar way.

Equivalently, we have

$$E\left[\int \Psi(y,\eta;\overline{\theta}) f(\eta|y;\overline{\theta})d\eta\right] = 0.$$

The stochastic EM algorithm for this problem works as follows, based on an i.i.d. sample $(y_1, ..., y_N)$. Iteratively, one draws $\widehat{\theta}^{(s+1)}$ given $\widehat{\theta}^{(s)}$ in two steps:[20]

1. Draw $(\eta_1^{(s)}, ..., \eta_N^{(s)})$ from the posterior distribution of $(\eta|y;\widehat{\theta}^{(s)})$.

2. Solve for $\widehat{\theta}^{(s+1)}$ in
$$\frac{1}{N}\sum_{i=1}^{N} \Psi(y_i, \eta_i^{(s)}; \widehat{\theta}^{(s+1)}) = 0.$$

This results in a Markov Chain $(\widehat{\theta}^{(0)}, \widehat{\theta}^{(1)}, ...)$, which is ergodic under suitable conditions. Moreover, under conditions given in Nielsen (2000a), asymptotically as N tends to infinity the process $\sqrt{N}(\widehat{\theta}^{(s)} - \overline{\theta})$ converges to a Gaussian autoregressive process. In the rest of this section we characterize its asymptotic mean and variance.[21]

Note that, using a conditional quantile representation,

$$\eta_i^{(s)} = Q_{\eta|y}(v_i^{(s)}|y_i; \widehat{\theta}^{(s)}),$$

where $v_i^{(s)}$ are standard uniform draws, independent of y_i.

[20]For simplicity we focus on the version of the algorithm with $M = 1$ draw per individual observation.

[21]Note that in our earnings and consumption model, some of the moment restrictions involve derivatives of "check" functions, which are not smooth. This is however not central to the discussion that follows, as it does not affect the form of the asymptotic variance.

We thus have

$$\frac{1}{N}\sum_{i=1}^{N}\Psi\left(y_i, Q_{\eta|y}(v_i^{(s)}|y_i;\widehat{\theta}^{(s)});\widehat{\theta}^{(s+1)}\right)=0.$$

Expanding around $\overline{\theta}$, we obtain

$$A\left(\widehat{\theta}^{(s+1)}-\overline{\theta}\right)+B\left(\widehat{\theta}^{(s)}-\overline{\theta}\right)+Z^{(s)} \quad = \quad o_p\left(N^{-\frac{1}{2}}\right),$$

where

$$\begin{aligned}
A &\equiv \mathbb{E}\left[\left.\frac{\partial}{\partial\theta'}\right|_{\overline{\theta}}\Psi\left(y, Q_{\eta|y}(v|y;\overline{\theta});\theta\right)\right] \\
&= \mathbb{E}\left[\frac{\partial\Psi\left(y,\eta;\overline{\theta}\right)}{\partial\theta'}\right], \\
B &\equiv \mathbb{E}\left[\left.\frac{\partial}{\partial\theta'}\right|_{\overline{\theta}}\Psi\left(y, Q_{\eta|y}(v|y;\theta);\overline{\theta}\right)\right] \\
&= \left.\frac{\partial}{\partial\theta'}\right|_{\overline{\theta}}\mathbb{E}\left[\int\Psi(y,\eta;\overline{\theta})f(\eta|y;\theta)d\eta\right] \\
&= \mathbb{E}\left[\Psi(y,\eta;\overline{\theta})\frac{\partial\ln f(\eta|y;\overline{\theta})}{\partial\theta'}\right], \text{ and} \\
Z^{(s)} &\equiv \frac{1}{N}\sum_{i=1}^{N}\Psi\left(y_i, Q_{\eta|y}(v_i^{(s)}|y_i;\overline{\theta});\overline{\theta}\right).
\end{aligned}$$

Note that,

$$A+B=\left.\frac{\partial}{\partial\theta'}\right|_{\overline{\theta}}\mathbb{E}\left[\int\Psi(y,\eta;\theta)f(\eta|y;\theta)d\eta\right]=\mathbb{E}\left[\left.\frac{\partial}{\partial\theta'}\right|_{\overline{\theta}}\int\Psi(y,\eta;\theta)f(\eta|y;\theta)d\eta\right].$$

The GMM identification condition thus requires $A+B<0$, so $(-A)^{-1}B<I$. This implies that the autoregressive process $\sqrt{N}\left(\widehat{\theta}^{(s)}-\overline{\theta}\right)$ is asymptotically stable.

We thus have

$$\sqrt{N}\left(\widehat{\theta}^{(s)}-\overline{\theta}\right)=\sum_{k=0}^{\infty}\left(-A^{-1}B\right)^k\left(-A^{-1}\right)\sqrt{N}Z^{(s-k)}.$$

Note that $\sqrt{N}Z^{(s)}$ are iid normal $(0,W)$, where

$$W=\mathbb{E}\left[\Psi\left(y,\eta;\overline{\theta}\right)\Psi\left(y,\eta;\overline{\theta}\right)'\right].$$

Hence

$$\sqrt{N}\left(\widehat{\theta}^{(s)}-\overline{\theta}\right)\xrightarrow{d}\mathcal{N}\left(0,V_\theta\right),$$

where $V_\theta=\sum_{k=0}^{\infty}\left(-A^{-1}B\right)^k\left(-A^{-1}\right)W\left(-A^{-1}\right)'\left(\left(-A^{-1}B\right)^k\right)'$.

Finally, V_θ can be recovered from the following matrix equation

$$A^{-1}BV_\theta B'(A^{-1})'=V_\theta-A^{-1}W(A^{-1})',$$

which can easily been solved in vector form.

C Estimation algorithm

Additional model restrictions. The tail parameters λ satisfy simple moment restrictions. For example, we have

$$\overline{\lambda}_{-}^{Q} = -\frac{\mathbb{E}\left[\int \mathbf{1}\left\{\eta_{it} \leq \sum_{k=0}^{K} \overline{a}_{k1}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right\} f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right]}{\mathbb{E}\left[\int \left(\eta_{it} - \sum_{k=0}^{K} \overline{a}_{k1}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right) \mathbf{1}\left\{\eta_{it} \leq \sum_{k=0}^{K} \overline{a}_{k1}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right\} f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right]}, \tag{C7}$$

and

$$\overline{\lambda}_{+}^{Q} = \frac{\mathbb{E}\left[\int \mathbf{1}\left\{\eta_{it} \geq \sum_{k=0}^{K} \overline{a}_{kL}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right\} f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right]}{\mathbb{E}\left[\int \left(\eta_{it} - \sum_{k=0}^{K} \overline{a}_{kL}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right) \mathbf{1}\left\{\eta_{it} \geq \sum_{k=0}^{K} \overline{a}_{kL}^{Q}\varphi_k(\eta_{i,t-1}, age_{it})\right\} f_i(\eta_i^T;\overline{\theta})d\eta_i^T\right]}, \tag{C8}$$

with similar equations for the other tail parameters.

Likelihood function. The likelihood function is (omitting the conditioning on age for conciseness)

$$\begin{aligned} f(y_i^T, c_i^T, a_i^T, \eta_i^T; \theta, \mu) &= \prod_{t=1}^{T} f(y_{it}|\eta_{it};\theta) \prod_{t=1}^{T} f(c_{it}|a_{it}, \eta_{it}, y_{it};\mu) \prod_{t=2}^{T} f(a_{it}|a_{i,t-1}, y_{i,t-1}, c_{i,t-1}) \\ &\quad \times \prod_{t=2}^{T} f(\eta_{it}|\eta_{i,t-1};\theta) f(a_{i1}|\eta_{i1};\mu) f(\eta_{i1};\theta). \end{aligned} \tag{C9}$$

Up to the term $\prod_{t=2}^{T} f(a_{it}|a_{i,t-1}, y_{i,t-1}, c_{i,t-1})$, which does not depend on η, the likelihood function is fully specified and available in closed form. For example, we have

$$\begin{aligned} f(y_{it}|\eta_{it};\theta) &= \mathbf{1}\left\{y_{it} - \eta_{it} < A_{it}^{\varepsilon}(1)\right\} \tau_1 \lambda_{-}^{\varepsilon} \exp\left[\lambda_{-}^{\varepsilon}\left(y_{it} - \eta_{it} - A_{it}^{\varepsilon}(1)\right)\right] \\ &\quad + \sum_{\ell=1}^{L-1} \mathbf{1}\left\{A_{it}^{\varepsilon}(\ell) \leq y_{it} - \eta_{it} < A_{it}^{\varepsilon}(\ell+1)\right\} \frac{\tau_{\ell+1} - \tau_\ell}{A_{it}^{\varepsilon}(\ell+1) - A_{it}^{\varepsilon}(\ell)} \\ &\quad + \mathbf{1}\left\{A_{it}^{\varepsilon}(L) \leq y_{it} - \eta_{it}\right\} (1 - \tau_L) \lambda_{+}^{\varepsilon} \exp\left[-\lambda_{+}^{\varepsilon}\left(y_{it} - \eta_{it} - A_{it}^{\varepsilon}(L)\right)\right], \end{aligned}$$

where $A_{it}^{\varepsilon}(\ell) \equiv \sum_{k=0}^{K} a_{k,\ell}^{\varepsilon}\varphi_k(age_{it})$ for all (i, t, ℓ). Note that the likelihood function is non-negative by construction. In particular, drawing from the posterior density of η automatically produces rearrangement of the various quantile curves (Chernozhukov, Galichon and Fernandez-Val, 2010).

Earnings algorithm. Start with $\widehat{\theta}^{(0)}$. Iterate on $s = 0, 1, 2, ...$ the two following steps.

Stochastic E-step: Draw M values $\eta_i^{(m)} = (\eta_{i1}^{(m)}, ..., \eta_{iT}^{(m)})$ from

$$f(\eta_i^T|y_i^T;\widehat{\theta}^{(s)}) \propto \prod_{t=1}^{T} f(y_{it}|\eta_{it};\widehat{\theta}^{(s)}) \prod_{t=2}^{T} f(\eta_{it}|\eta_{i,t-1};\widehat{\theta}^{(s)}) f(\eta_{i1};\widehat{\theta}^{(s)}),$$

where $a \propto b$ means that a and b are equal up to a proportionality factor independent of η.

M-step: Compute, for $\ell = 1, ..., L$,

$$\left(\widehat{a}_{0\ell}^{Q,(s+1)}, ..., \widehat{a}_{K\ell}^{Q,(s+1)}\right) = \underset{\left(a_{0\ell}^{Q},...,a_{K\ell}^{Q}\right)}{\text{argmin}} \sum_{i=1}^{N} \sum_{t=2}^{T} \sum_{m=1}^{M} \rho_{\tau_\ell} \left(\eta_{it}^{(m)} - \sum_{k=0}^{K} a_{k\ell}^{Q} \varphi_k(\eta_{i,t-1}^{(m)}, age_{it}) \right),$$

$$\left(\widehat{a}_{0\ell}^{\varepsilon,(s+1)}, ..., \widehat{a}_{K\ell}^{\varepsilon,(s+1)}\right) = \underset{\left(a_{0\ell}^{\varepsilon},...,a_{K\ell}^{\varepsilon}\right)}{\text{argmin}} \sum_{i=1}^{N} \sum_{t=1}^{T} \sum_{m=1}^{M} \rho_{\tau_\ell} \left(y_{it} - \eta_{it}^{(m)} - \sum_{k=0}^{K} a_{k\ell}^{\varepsilon} \varphi_k(age_{it}) \right),$$

$$\left(\widehat{a}_{0\ell}^{\eta_1,(s+1)}, ..., \widehat{a}_{K\ell}^{\eta_1,(s+1)}\right) = \underset{\left(a_{0\ell}^{\eta_1},...,a_{K\ell}^{\eta_1}\right)}{\text{argmin}} \sum_{i=1}^{N} \sum_{m=1}^{M} \rho_{\tau_\ell} \left(\eta_{i1}^{(m)} - \sum_{k=0}^{K} a_{k\ell}^{\eta_1} \varphi_k(age_{i1}) \right),$$

and compute

$$\widehat{\lambda}_{-}^{Q,(s+1)} = -\frac{\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{m=1}^{M} \mathbf{1}\left\{\eta_{it}^{(m)} \le \widehat{A}_{itm}^{Q,(s+1)}\right\}}{\sum_{i=1}^{N}\sum_{t=2}^{T}\sum_{m=1}^{M} \left(\eta_{it}^{(m)} - \widehat{A}_{itm}^{Q,(s+1)}\right) \mathbf{1}\left\{\eta_{it}^{(m)} \le \widehat{A}_{itm}^{Q,(s+1)}\right\}},$$

where

$$\widehat{A}_{itm}^{Q,(s+1)} \equiv \sum_{k=0}^{K} \widehat{a}_{k1}^{Q,(s+1)} \varphi_k(\eta_{i,t-1}^{(m)}, age_{it}),$$

with similar updating rules for $\widehat{\lambda}_{+}^{Q,(s+1)}$, $\widehat{\lambda}_{-}^{\varepsilon,(s+1)}$, $\widehat{\lambda}_{+}^{\varepsilon,(s+1)}$, $\widehat{\lambda}_{-}^{\eta_1,(s+1)}$, and $\widehat{\lambda}_{+}^{\eta_1,(s+1)}$.

In practice, we start the algorithm with different choices for $\widehat{\theta}^{(0)}$, and we select the parameter values that correspond to the highest average complete-data likelihood.

Consumption algorithm. Similar to the earnings algorithm. One difference is that in the stochastic E-step we draw $\eta_i^{(m)}$ from

$$f(\eta_i^T | y_i^T, c_i^T, a_i^T; \widehat{\theta}, \widehat{\mu}^{(s)}) \propto \prod_{t=1}^{T} f(y_{it}|\eta_{it}; \widehat{\theta}) \prod_{t=1}^{T} f(c_{it}|a_{it}, \eta_{it}, y_{it}; \widehat{\mu}^{(s)})$$
$$\times \prod_{t=2}^{T} f(\eta_{it}|\eta_{i,t-1}; \widehat{\theta}) f(a_{i1}|\eta_{i1}; \widehat{\mu}^{(s)}) f(\eta_{i1}; \widehat{\theta}).$$