The Panel Data Dynamics of Earnings and Consumption: A Nonlinear Framework

Richard Blundell (UCL and IFS)

Joint with Manuel Arellano (CEMFI) and Stéphane Bonhomme (Chicago)

21st PANEL DATA CONFERENCE

CEU, BUDAPEST, JUNE 29th 2015

The Panel Data Dynamics of Earnings and Consumption: A Nonlinear Framework

The idea behind this work is to examine the transmission of "shocks" from income to consumption using household panel data.

- To consider alternative ways of modelling persistence.

- To explore the nature of income persistence and consumption dynamics.

- To examine the link between consumption and income inequality.

⇒ Use a variety of US Household Panel data and Norwegian Population Register data.

Earnings and consumption dynamics

• A prototypical panel data model of (log) earned family income is the "canonical" model:

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

where y_{it} is net of a *systematic component*, η_{it} is a *random walk* with innovation v_{it}, and ε_{it} is an *independent shock*.

• Consumption is then related to income via the "partial insurance" model:

$$\triangle c_{it} = \phi_t(a_t) \cdot v_{it} + \psi_t(a_t) \cdot \varepsilon_{it} + \nu_{it}, \quad i = 1, ..., N, \quad t = 1, ..., T.$$

where c_{it} is log total consumption net of a systematic component, $\phi_t(a_t)$ is the *transmission* of persistence shocks v_{it}, and $\psi_t(a_t)$ the *transmission* of transitory shocks. The ν_{it} are taste shocks, typically assumed to be independent across periods.

$\Rightarrow$ The transmission or "*partial insurance*" parameters ϕ and ψ are known functions of age t and beginning of period (net) assets a_{it}, see Blundell, Low and Preston (QE, 2014).

Motivation

• This "*standard*" framework implies a set of covariance restrictions for panel data on consumption and income. Allowing parameters and variances to depend on age is key.

⇒ can show (over-)identification and efficient estimation via nonlinear GMM, see Blundell, Preston and Pistaferri (AER, 2008) and Blundell, Pistaferri and Saporta (NBER, 2014) - who also introduce family labor supply and taxes.

• Linearity of the income process simplifies identification and estimation. However, by construction it *rules out nonlinear transmission of shocks*.

• The aim here is to take a different tack and to develop a new approach to modeling persistence in which the impact of past shocks on current earnings can be altered by the size and sign of new shocks.

⇒ this new framework allows for "*unusual*" *shocks to wipe out the memory of past shocks*.
⇒ the future persistence of a current shock depends on the future shocks.

• We show the presence of "unusual" shocks matches the data and has a key impact consumption and saving over the life cycle.

Methodology and data

• Nonlinear dynamic model with latent variables (the unobserved earnings components).

– Nonparametric identification builds on Hu and Schennach (08) and Wilhelm (12).

– Flexible parametric estimation that combines quantile modeling and linear expansions in bases of functions.

• Panel data on household earned income, consumption ($\approx$ 70% of expenditures of nondurables and services) and assets holdings from the new waves of PSID (1999-2009). Recently (2004) further improved.

– Avoids need to use food consumption or imputed consumption data.

– Compare with population panel (register) data from Norway, see Blundell, Graber and Mogstad (2014) - not quite finished constructing consumption data.

Nonlinear Persistence

• Consider a cohort of households, $i = 1, ..., N$, and denote age as t. Let y_{it} denote log-labor income, net of age dummies.

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

$\triangleright$ η_{it} follows a general first-order Markov process (can be generalised).

• Denoting the τth conditional quantile of η_{it} given $\eta_{i,t-1}$ as $Q_t(\eta_{i,t-1}, \tau)$, we specify

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

$\triangleright$ ε_{it} has zero mean, independent over time (at a 2-year frequency in the PSID).

$\triangleright$ The conditional quantile functions $Q_t(\eta_{i,t-1}, u_{it})$ and the marginal distributions F_{ε_t} are age (t) specific.

A measure of persistence

- The model allows for nonlinear dynamics of income.

- To see this, consider the following measure of persistence

$$\rho_t(\eta_{i,t-1}, \tau) = \frac{\partial Q_t(\eta_{i,t-1}, \tau)}{\partial \eta}.$$

$\Rightarrow$ $\rho_t(\eta_{i,t-1}, \tau)$ measures the persistence of $\eta_{i,t-1}$ when it is hit by a shock u_{it} that has rank τ.

– Allows a general form of conditional heteroscedasticity, skewness and kurtosis.

- In the "canonical model" $\eta_{it} = \eta_{i,t-1} + v_{it}$, with v_{it} independent over time and independent of past $\eta's$,

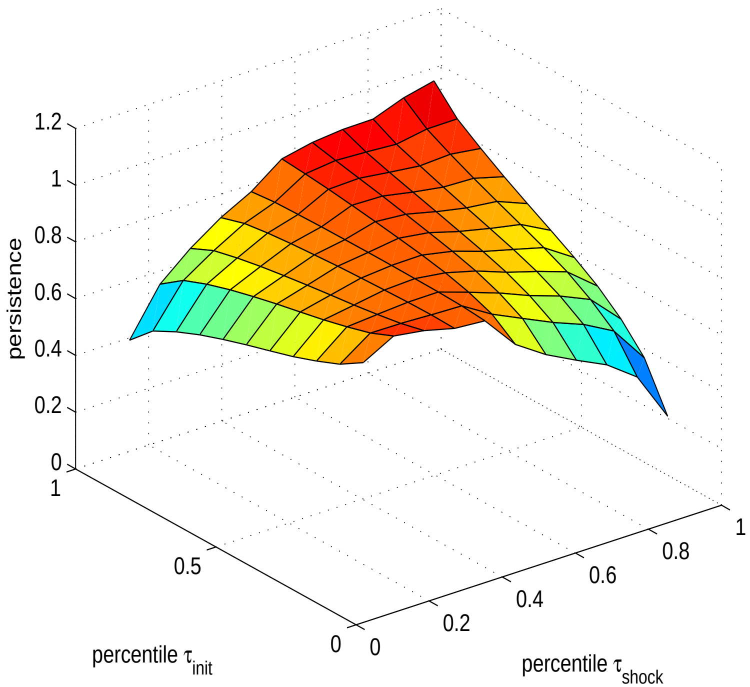
$$\eta_{it} = \eta_{i,t-1} + F_{v_t}^{-1}(u_{it}) \quad \Rightarrow \quad \rho_t(\eta_{i,t-1}, \tau) = 1 \text{ for all } (\eta_{i,t-1}, \tau).$$

– But what's the evidence for such nonlinearities in persistence?

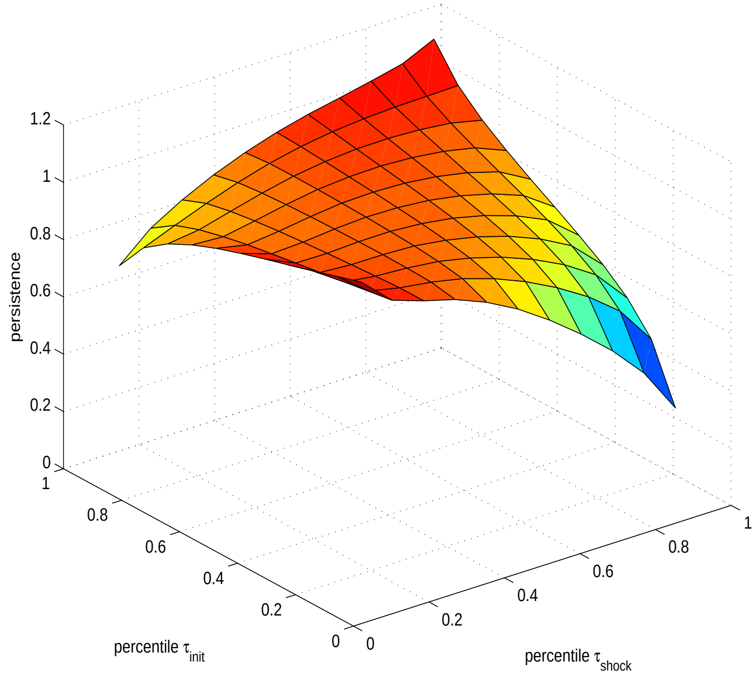
Some motivating evidence: Quantile autoregressions of log-earnings

$$\frac{\partial Q_{y_t|y_{t-1}}(y_{i,t-1},\tau)}{\partial y}$$

PSID data

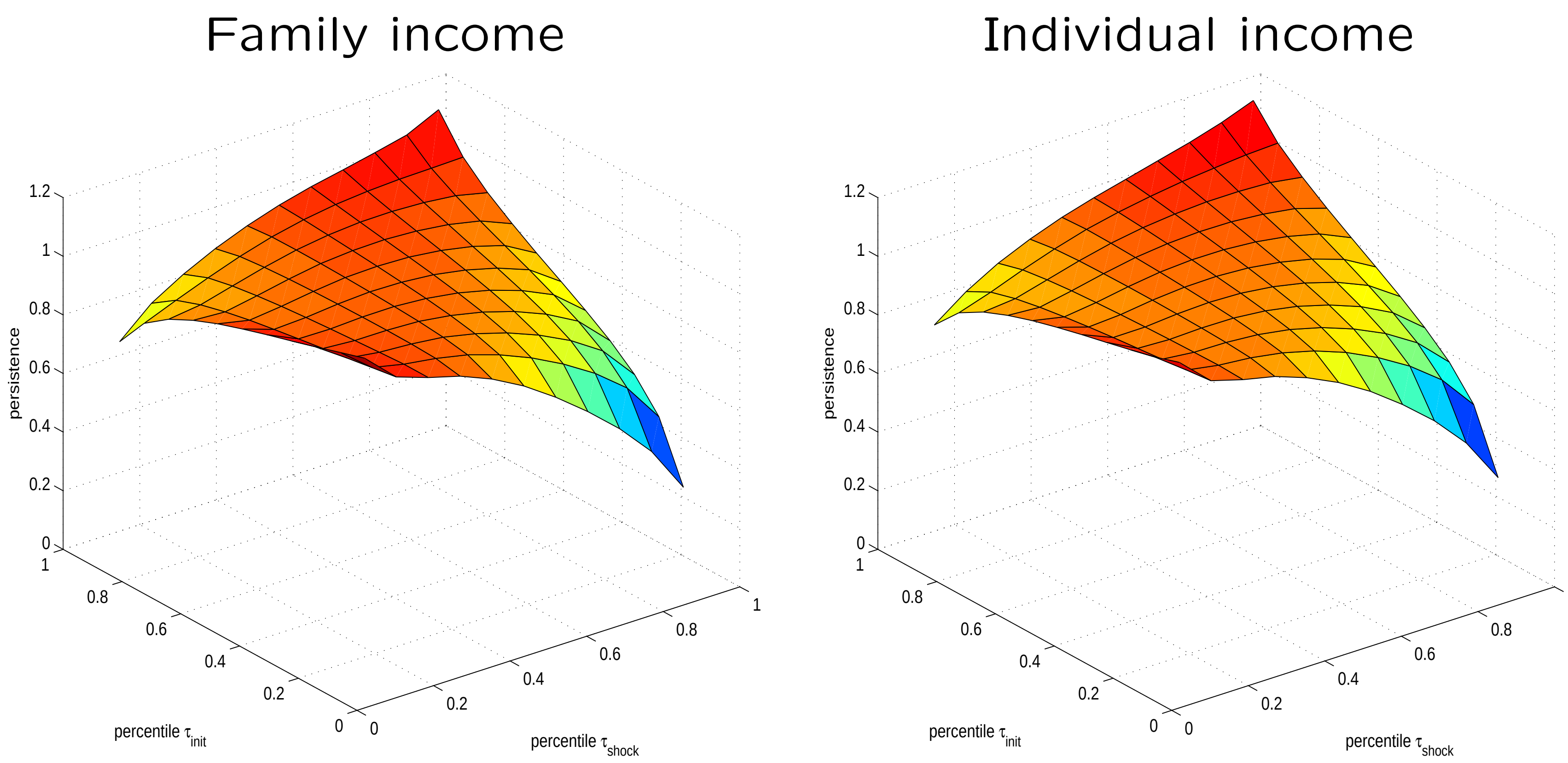


Norwegian administrative data



Note: Residuals of log pre-tax household labor earnings, Age 35-65 1999-2009 (US), Age 25-60 2005-2006 (Norway). Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, using a grid of 11-quantiles and a 3rd degree Hermite polynomial.

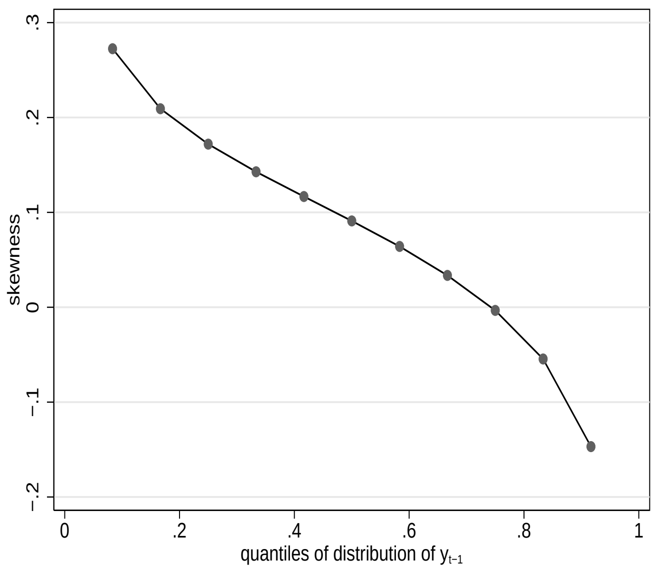
Nonlinear earnings persistence, Norwegian administrative data



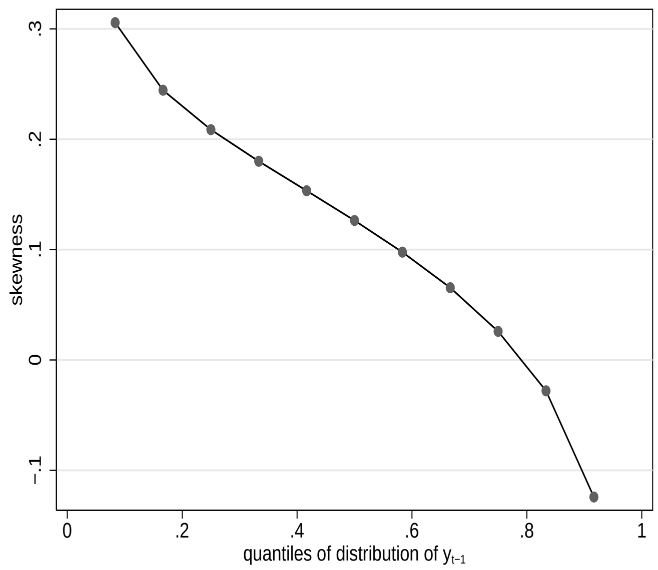
Note: Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the distribution of $y_{i,t-1}$, using a grid of 11-quantiles and a 3rd degree Hermite polynomial. Age 25-60, years 2005-2006.

Conditional skewness, Norwegian administrative data

Family income



Individual income



Note: Skewness measured as a nonparametric estimate of

$$\frac{Q_{y_t|y_{t-1}}(y_{i,t-1},.9)+Q_{y_t|y_{t-1}}(y_{i,t-1},.1)-2Q_{y_t|y_{t-1}}(y_{i,t-1},.5)}{Q_{y_t|y_{t-1}}(y_{i,t-1},.9)-Q_{y_t|y_{t-1}}(y_{i,t-1},.1)}.$$

Age 25-60, years 2005-2006.

Outline

- Consumption simulations and model specification
- Identification
- Data and estimation strategy
- Empirical results

Life-cycle model

- Setup and calibration based on Kaplan and Violante (10, KV).

- Households enter the labor market at age 25, work until 60, and die with certainty at age 95.

- They have access to a single risk-free, one-period bond whose constant return is $1 + r$ (where $r = .03$),

$$A_t = (1 + r)A_{t-1} + Y_{t-1} - C_{t-1}.$$

- Log-earnings are $\ln Y_t = \kappa_t + \eta_t + \varepsilon_t$, where κ_t is a deterministic age profile.

Life-cycle model (cont.)

• In period t agents know η_t, ϵ_t and their past values, but not η_{t+1} or ε_{t+1} (no advance information).

• Period-t optimization

$$V_t(A_t, \eta_t, \varepsilon_t) = \max_{C_t} \; u(C_t) + \beta \mathbb{E}_t \left[V_{t+1} \left(A_{t+1}, \eta_{t+1}, \varepsilon_{t+1} \right) \right],$$

where $u(\cdot)$ is CRRA ($\gamma = 2$), and $\beta = 1/(1+r) \approx .97$.

• We compare the results for the canonical earnings process used by KV, and for a parametric nonlinear process that roughly approximates the empirical autoregressions.

A simple nonlinear parametric model

• A parametric model for η_{it} is

$$\eta_{it} = \rho_t(\eta_{i,t-1}, v_{it})\eta_{i,t-1} + v_{it},$$

where $\rho_t(\eta, v) = 1-\delta$ if $(\eta > c_{t-1}, v < -b_t)$ or $(\eta < -c_{t-1}, v > b_t)$, and $\rho_t(\eta, v) = 1$ otherwise; and $v_{it} \sim \mathcal{N}(0, \sigma_t^2)$.

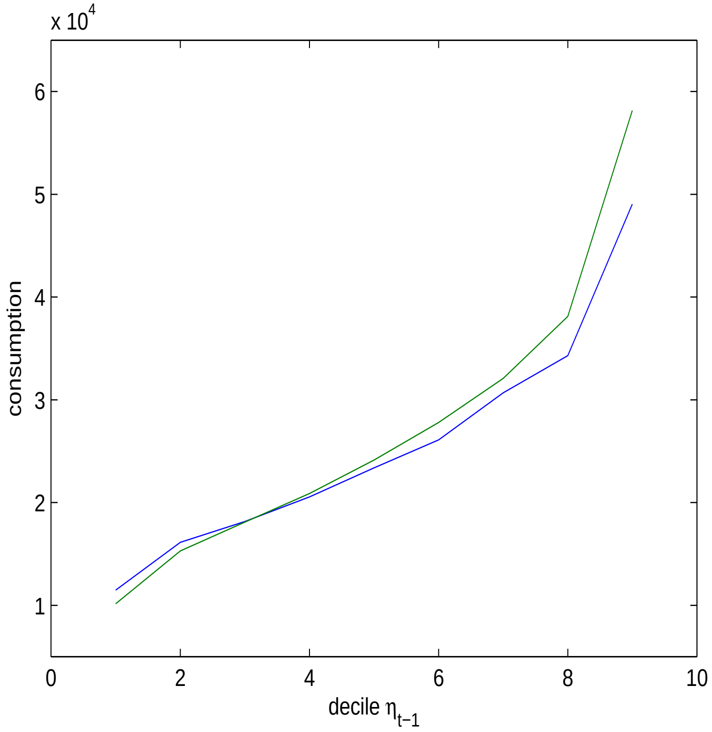
• Persistence is lower $(1-\delta < 1)$ when a bad shock hits a high earnings household ("individual disasters"), or a good shock hits a low earnings household.

• η_{it} features conditional skewness: positive for low $\eta_{i,t-1}$, negative for high $\eta_{i,t-1}$.

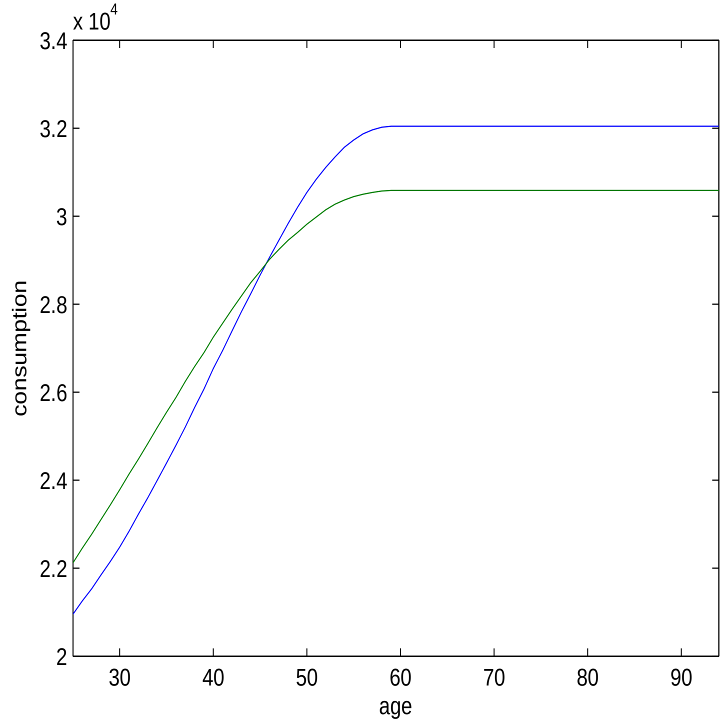
• In the simulation results we set $\delta = .2$ and the probability of a high or low "unusual shock" set to 15%.

Simulation results

Consumption (age 37) by decile of η_{t-1}



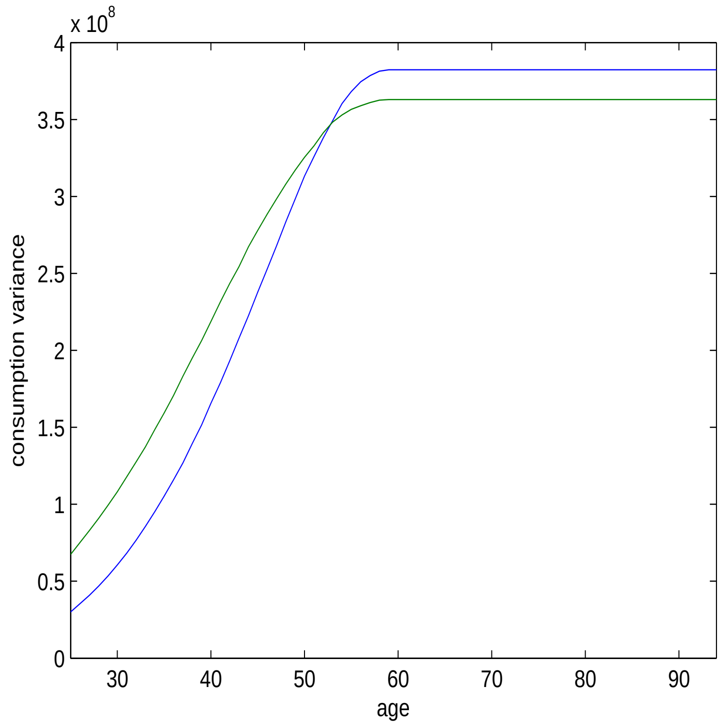
Average consumption over the life-cycle



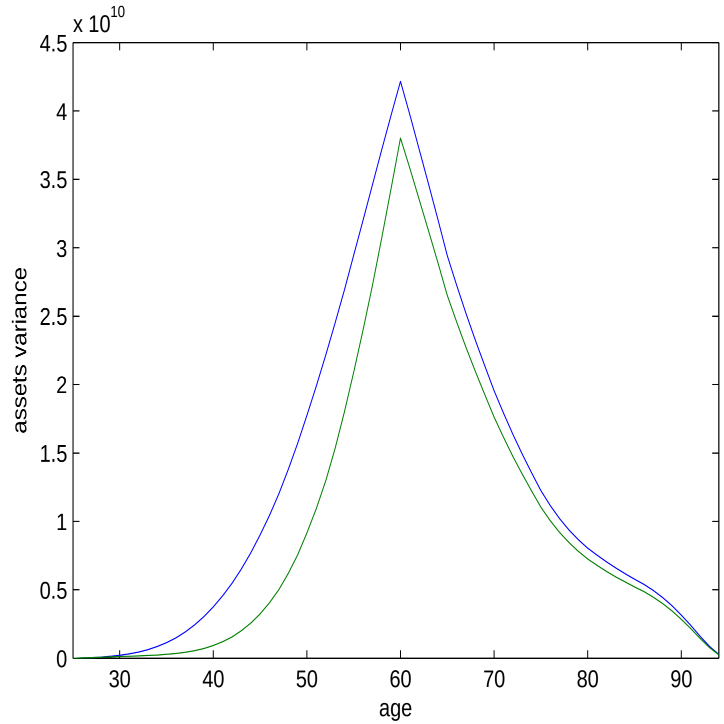
Note: Blue is nonlinear earnings process, Green is canonical earnings process.

Simulated Variance of Consumption and Assets

Consumption variance
over the life-cycle



Assets variance
over the life-cycle



Note: Blue is nonlinear earnings process, Green is canonical earnings process.

An Empirical Consumption Rule

• Let c_{it} and a_{it} denote log-consumption and log-assets (beginning of period) net of age dummies.

• Our empirical specification is based on

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

where ν_{it} are independent across periods, and g_t is a nonlinear, age-dependent function, monotone in ν_{it}.

– ν_{it} may be interpreted a taste shifter that increases marginal utility. We normalize its distribution to be standard uniform in each period.

• This consumption rule is consistent, in particular, with the standard life-cycle model of the previous slides. Can allow for individual unobserved heterogeneity and for advance information and habits.

Insurance coefficients

- With consumption specification given by

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

consumption responses to η and ε are

$$\phi_t(a, \eta, \varepsilon) = \mathbb{E}\left[\frac{\partial g_t\left(a, \eta, \varepsilon, \nu\right)}{\partial \eta}\right], \quad \psi_t(a, \eta, \varepsilon) = \mathbb{E}\left[\frac{\partial g_t\left(a, \eta, \varepsilon, \nu\right)}{\partial \varepsilon}\right].$$

– $\phi_t(a, \eta, \varepsilon)$ and $\psi_t(a, \eta, \varepsilon)$ reflect the transmission of shocks to the persistent and transitory earnings components, respectively. That is the lack of insurance to shocks.

- The marginal effect of an earnings shock u on consumption is

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

Earnings: identification

• For $T = 3$, Wilhelm (2012) gives conditions under which the distribution of ε_{i2} is identified.

– In particular completeness of the *pdf*s of $(y_{i2}|y_{i1})$ and $(\eta_{i2}|y_{i1})$. This requires η_{i1} and η_{i2} to be dependent.

• We build on this result to establish identification of the earnings model.

• Apply the result to each of the three-year subpanels $t \in \{1, 2, 3\}$ to $t \in \{T-2, T-1, T\}$

$\Rightarrow$ The marginal distribution of ε_{it} are identified for $t \in \{2, 3, ..., T-1\}$.

$\Rightarrow$ By independence the joint distribution of $(\varepsilon_{i2}, \varepsilon_{i3}, ..., \varepsilon_{i,T-1})$ is identified.

$\Rightarrow$ By deconvolution the distribution of $(\eta_{i2}, \eta_{i3}, ..., \eta_{i,T-1})$ is identified.

• The distribution of ε_{i1}, η_{i1}, and ε_{iT}, η_{iT} are not identified in general.

Consumption: assumptions

• u_{it} and ε_{it} are independent of a_{i1} for $t \geq 1$, where $\eta_{it} = Q_t(\eta_{i,t-1}, u_{it})$.

• 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 persistent shocks.

• Denoting $\eta_i^t = (\eta_{it}, \eta_{i,t-1}, ..., \eta_{i1})$, we assume (in this talk) that:
a_{it} is independent of $(\eta_i^{t-1}, a_i^{t-2}, \varepsilon_i^{t-2})$ given $(a_{i,t-1}, c_{i,t-1}, y_{i,t-1})$.

– Consistent with the accumulation rule in the standard life-cycle model with one single risk-less asset.

Consumption: initial assets

• Let $y = (y_1, ..., y_T)$. We have

$$f(a_1|y) = \int f(a_1|\eta_1, y) f(\eta_1|y) d\eta_1$$
$$= \int f(a_1|\eta_1) f(\eta_1|y) d\eta_1,$$

where we have used that u_{it} and ε_{it} are independent of a_{i1}.

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

• If $f(\eta_1|y)$ is complete, then $f(a_1|\eta_1)$ is identified.

– Structure is as in the NPIV problem where η_1 is the endogenous regressor and y is the instrument.

Consumption: first period

- We have

$$f(c_1, a_1|y) \equiv \int f(c_1, a_1|\eta_1, y) f(\eta_1|y) d\eta_1$$

and given our assumptions

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

– $f(a_1|\eta_1)$ can be treated as known.

– Provided we have completeness in $(y_2, ..., y_T)$ of $f(\eta_1|y_1, y_2, ..., y_T)$, then $f(c_1|a_1, \eta_1, y_1)$, is identified.

- Intuition: $y_{i2}, ..., y_{iT}$ are used as "instruments" for η_{i1}.

- Again requires dependence between $\eta_{i,2}$ and $\eta_{i,1}$.

Consumption: subsequent periods

- By the model's assumptions

$$f(c^t, a^t|y) = \prod_{s=2}^{t} f(a_s|a_{s-1}, y_{s-1}, c_{s-1}) \times \int \prod_{s=1}^{t} f(c_s|a_s, \eta_s, y_s) f(a_1|\eta_1) f(\eta^t|y) d\eta^t.$$

- Let

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

and consider

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

Identification of $f(c_t|a_t, \eta_t, y_t)$ follows by induction if $\mathcal{L}_t$ is injective.

- Intuition: lagged consumption and assets, as well as lags and leads of earnings, are used as instruments for η_{it}.

Identification: extensions

• Similar techniques can be used in the presence of *advance information*, e.g.

$$c_{it} = g_t\left(a_{it}, \eta_{it}, \eta_{i,t+1}, \varepsilon_{it}, \nu_{it}\right),$$

or consumption habits, e.g.

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

$\Rightarrow$ Can extend to habits and also cases where the consumption rule depends on lagged η, or when η follows a second-order Markov process. (See Section 5 in the paper).

• Households differ in their initial productivity η_1 and initial assets. Panel data provide opportunities to allow for additional, *unobserved heterogeneity* in earnings and consumption (the next slide deals with the latter).

– For example: heterogeneity ξ_i in discounting or preferences, or heterogeneity $\widetilde{\xi}_i$ in the Markovian transitions of η_{it}

Extensions (cont.)

• Consumption rule with *unobserved heterogeneity*:

$$c_{it} = g_t \left(a_{it}, \eta_{it}, \varepsilon_{it}, \xi_i, \nu_{it} \right).$$

• We assume that u_{it} and ε_{it}, for $t \geq 1$, are independent of (a_{i1}, ξ_i).

• The distribution of $(a_{i1}, \xi_i, \eta_{i1})$ is unrestricted.

• A combination of the above identification arguments and the main result of Hu and Schennach (08) identifies

– The period-t consumption distribution $f(c_t | a_t, \eta_t, y_t, \xi)$.

– The distribution of initial conditions $f(\eta_1, \xi, a_1)$.

Data and estimation strategy

New PSID

• PSID 1999-2009, 6 waves (every other year).

• y_{it} are residuals of log total pre-tax household labor earnings on a set of demographics.

– cohort and calendar time dummies, family size and composition, education, race, and state dummies.

• Assets holdings are the sum of financial assets, real estate value, pension funds, and car value, net of mortgages and other debt.

New PSID (cont.)

- Information on food expenditures, rents, health expenditures, utilities, car-related expenditures, education, and child care. Recreation, alcohol, tobacco and clothing are missing before 2004.

- We follow Blundell, Pistaferri and Saporta (12, BPS) and impute rent expenditures for home owners.

- c_{it} and a_{it} are residuals, using the same set of demographics as for earnings.

- We follow BPS and select a sample of participating and married male heads aged between 30 and 65.

- In this talk I focus on a balanced subsample of $N = 749$ households.

Empirical specification: earnings

• The quantile function of η_{it} given $\eta_{i,t-1}$ is specified as

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

where φ_k, $k = 0, 1, ..., K$, are polynomials (Hermite).

• In addition, the quantile functions of ε_{it} and η_{i1} are

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

– Note that our data set has ages 30 - 65. The joint distribution of ε_{it} are nonparametrically identified in the age range for all ages between 32 and 63. in turn the joint distribution of and η_{it} is nonparametrically identified over the same age range.

Empirical specification: consumption

• We specify

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

– Additivity in the taste shifters, though not essential, is convenient given the sample size.

• In addition, the conditional quantiles of a_{i1} given η_{i1} and age_{i1} are

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

Implementation choices

• Following Wei and Carroll (09) we model $a_k^Q(\tau)$ as piecewise-linear interpolating splines on a grid $0 < \tau_1 < \tau_2 < ... < \tau_L < 1$.

– Convenient as the likelihood function is available in closed form.

• We extend the specification of the intercept coefficient $a_0^Q(\tau)$ on $(0, \tau_1]$ and $[\tau_L, 1)$ using a parametric model: exponential (λ).

• In practice, we take $L = 11$ and $\tau_\ell = \ell/L + 1$. φ_k and $\widetilde{\varphi}_k$ are low-dimensional tensor products of Hermite polynomials.

• We set $b_0(\tau) = \alpha + \sigma\Phi^{-1}(\tau)$, where (α, σ) are to be estimated.

Estimation algorithm

- The algorithm is an adaptation of techniques developed in Arellano & Bonhomme (2013) in the context of quantile models with time-invariant unobserved heterogeneity.

- The first estimation step recovers estimates of the income 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 income process alone.

Model's restrictions: income

- Let θ be the income-related parameters, and μ be the consumption-related ones, with true values $\overline{\theta}$ and $\overline{\mu}$. Denote the posterior density of $(\eta_{i1}, ..., \eta_{iT})$ given the income data as

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

- Let $\rho_\tau(u) = u(\tau - \mathbf{1}\{u \leq 0\})$ denote the "check" function of quantile regression, and let $\overline{a}_{k\ell}^Q$ denote the value of $a_{k\ell}^Q = a_k^Q(\tau_\ell)$ evaluated at the true $\overline{\theta}$.

- For all $t \geq 2$ and $\ell \in \{1, ..., L\}$ the model implies

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

with additional restrictions involving the other parameters in θ.

- The joint likelihood of $\left(\eta_i^T, y_i^T \mid age_i^T; \overline{\theta}\right)$ is available in closed form, so that it is easy to simulate from f_i.

Model's restrictions (cont.)

• Letting μ (true value $\overline{\mu}$) be the consumption-related parameters, the model implies

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

and

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

with additional restrictions involving the other parameters in μ.

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

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

Overview of estimation

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

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

• We use a "stochastic EM" algorithm (in a non-likelihood setup). Starting with $\widehat{\theta}^{(0)}$ we iterate on s=0,1,... the following two steps until convergence of the Markov Chain:

1. Stochastic E-step: draw $\eta_i^{(m)} = (\eta_{i1}^{(m)}, ..., \eta_{iT}^{(m)})$ for $m = 1, ..., M$ from $f_i(\cdot; \widehat{\theta}^{(s)})$. We use a random-walk Metropolis-Hastings sampler.

2. M-step: update

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

Overview of estimation (cont.)

• As the likelihood function is available in closed form, the E-step is straightforward.

• The M-step consists of a number of ordinary regressions and quantile regressions, such as

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

• We compute $\widehat{\theta}$ as an average of $\widehat{\theta}^{(s)}$ across S iterations.

• We estimate $\widehat{\theta}$ and $\widehat{\mu}$ sequentially.

Statistical properties

• Nielsen (00) studies the properties of this algorithm in a likelihood case. He provides conditions for the Markov Chain $\widehat{\theta}^{(s)}$ to be ergodic (for a fixed sample size).

• He also shows that $\sqrt{N}\left(\widehat{\theta}^{(s)} - \overline{\theta}\right)$ converges to a Gaussian autoregressive process as N tends to infinity.

– In the paper we adapt Nielsen's arguments to derive the form of the asymptotic variance in our case.

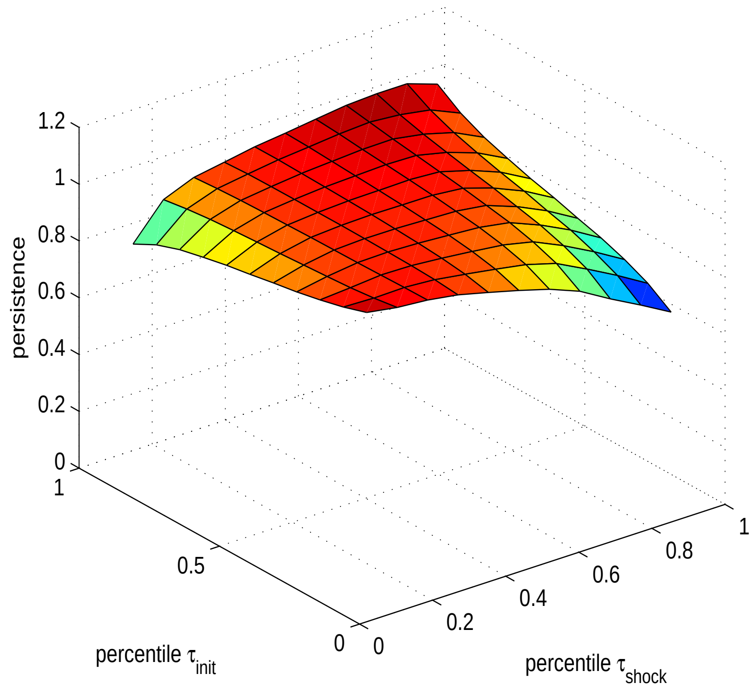
• Not done: Asymptotics as K (number of polynomial terms) and L (number of knots) tend to infinity with N.

– Special cases are treated in Belloni *et al.* (12) and Arellano and Bonhomme (13).

Empirical results

Nonlinear persistence of η_{it}

$$\rho_t(\eta_{i,t-1},\tau) = \frac{\partial Q_{\eta_t|\eta_{t-1}}(\eta_{i,t-1},\tau)}{\partial \eta}, \text{ nonlinear model}$$

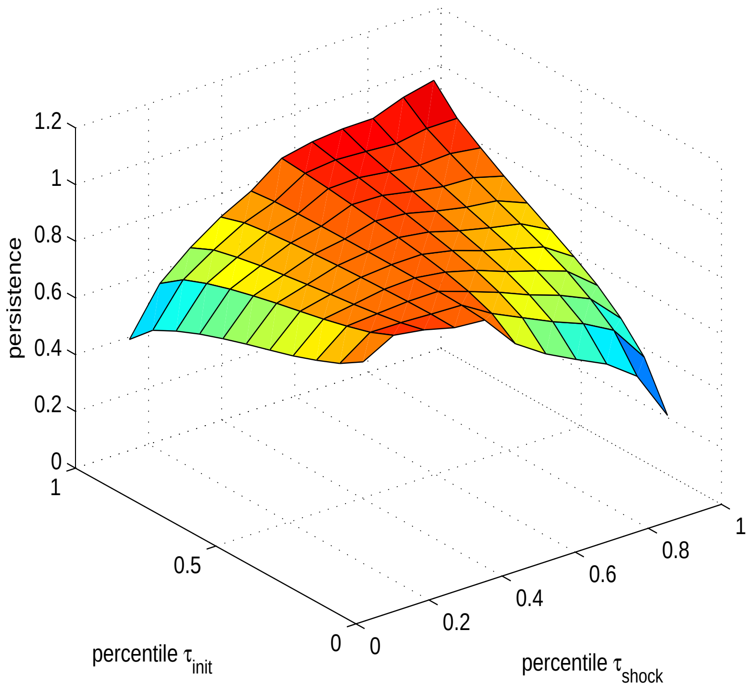


Note: Estimates of the average derivative of the conditional quantile function of η_{it} on $\eta_{i,t-1}$ with respect to $\eta_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $\eta_{i,t-1}$ that corresponds to the τ_{init} percentile of the distribution of $\eta_{i,t-1}$. Evaluated at mean age in the sample (47.5 years).

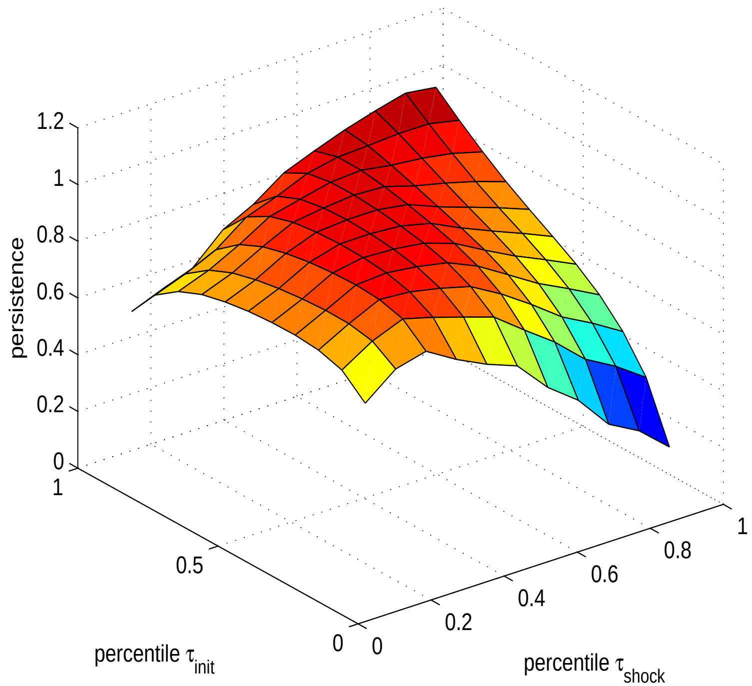
Nonlinear persistence of y_{it}

$$\frac{\partial Q_{y_t|y_{t-1}}(y_{i,t-1},\tau)}{\partial y}$$

PSID data



Nonlinear model

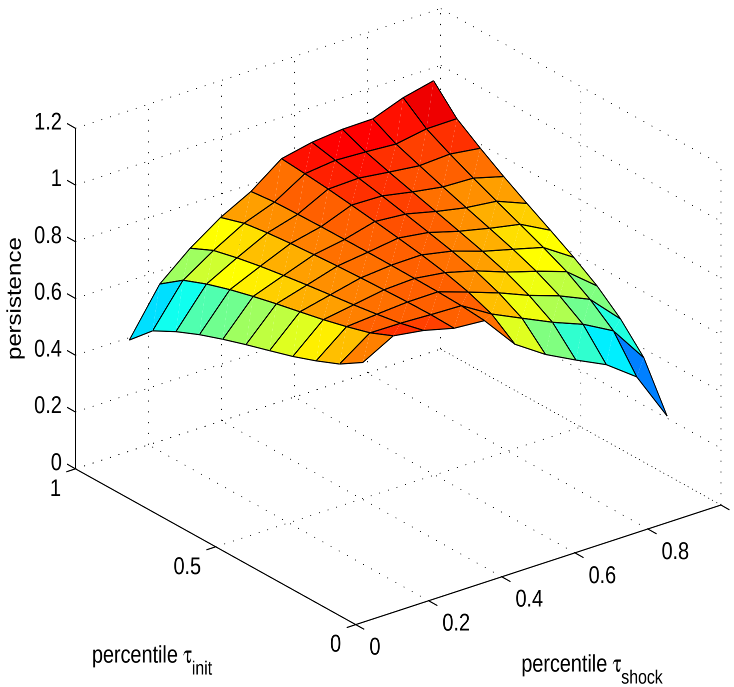


Note: Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the distribution of $y_{i,t-1}$.

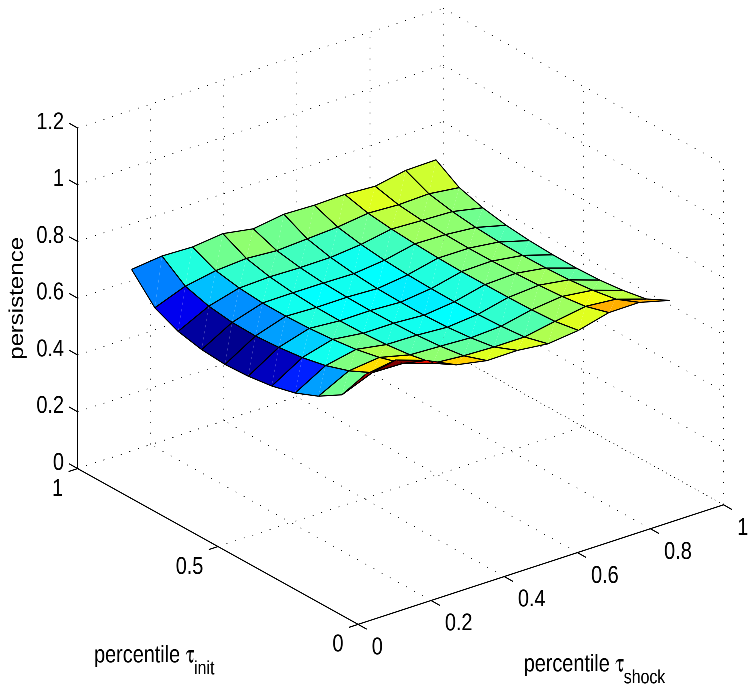
Nonlinear persistence of y_{it} (cont.)

$$\frac{\partial Q_{y_t|y_{t-1}}(y_{i,t-1},\tau)}{\partial y}$$

PSID data

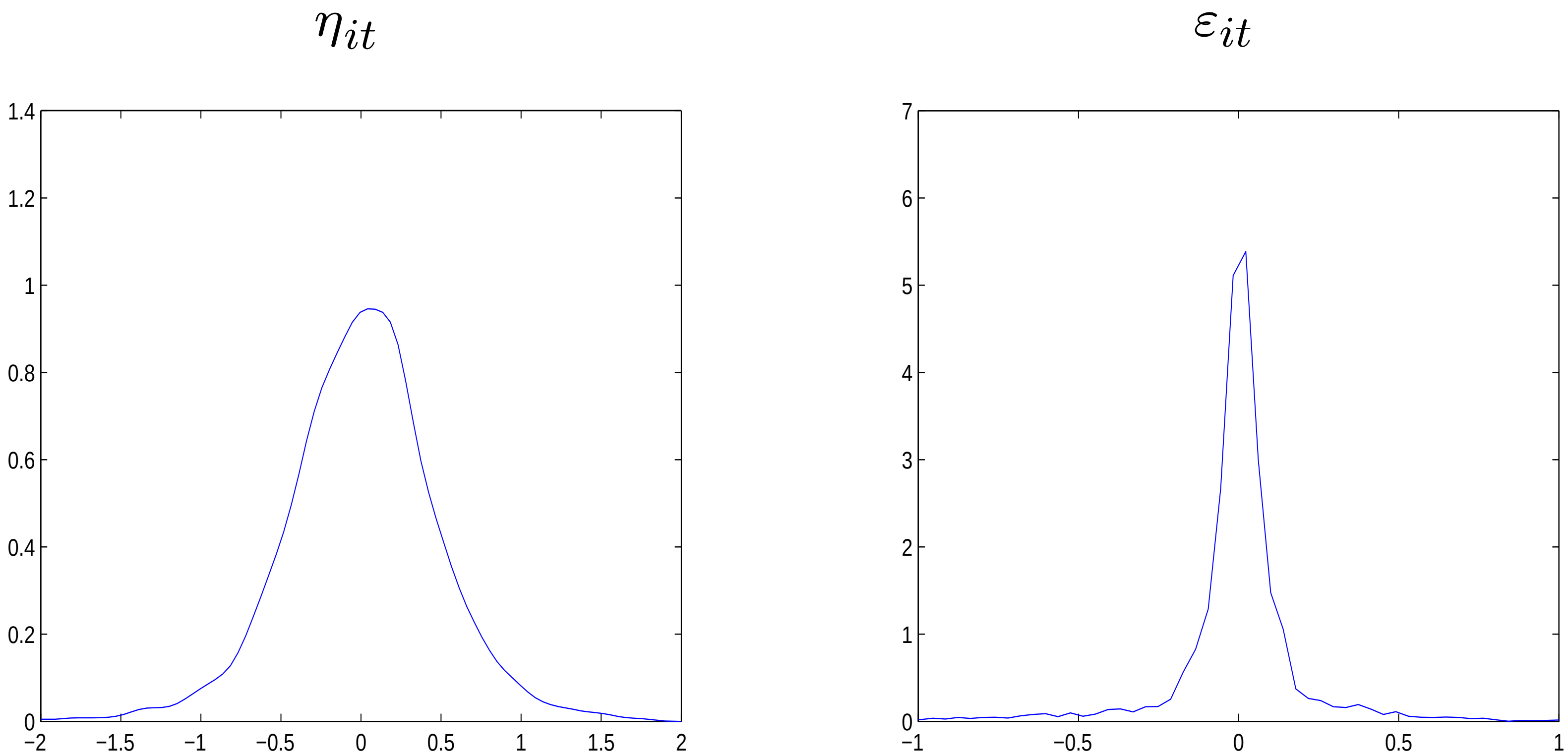


Canonical model



Note: 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 $y_{i,t-1}$.

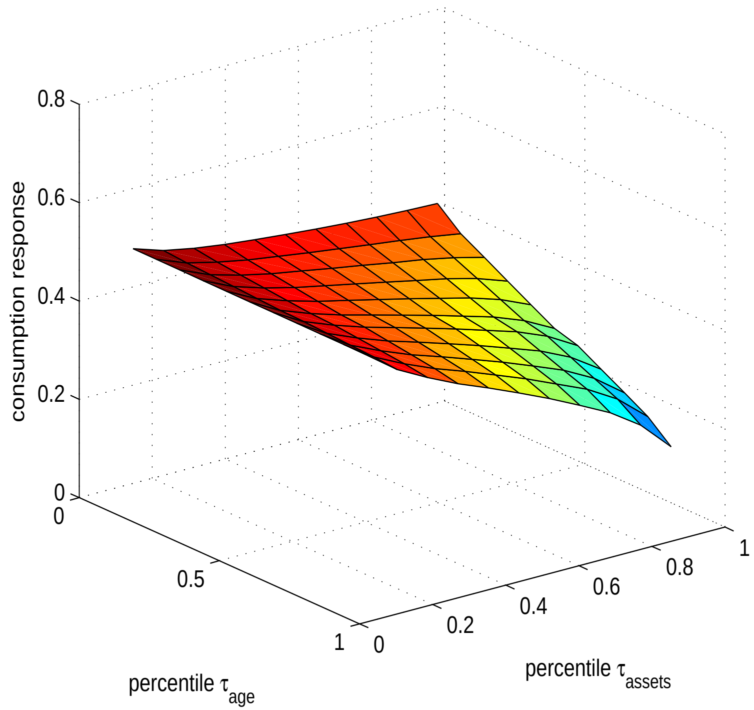
Densities of persistent and transitory earnings components



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

Consumption response to η_{it}, by assets and age

$$\overline{\phi}_t(a) = \mathbb{E}\left[\frac{\partial g_t(a,\eta_{it},\varepsilon_{it},\nu_{it})}{\partial \eta}\right], \text{ nonlinear model}$$

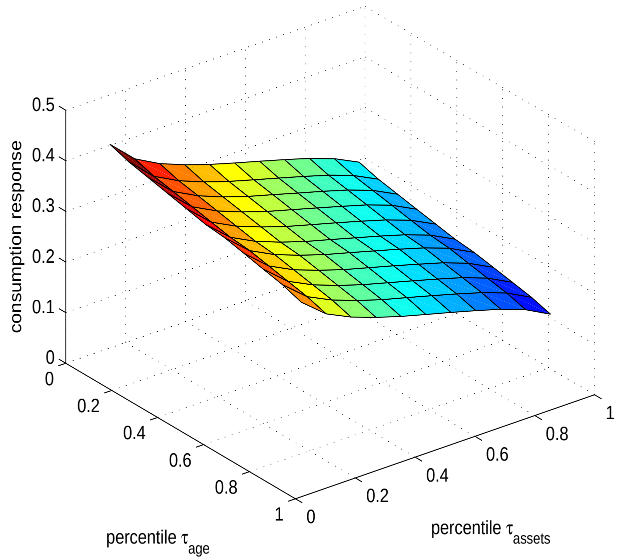


Note: Estimates of the average consumption response $\overline{\phi}_t(a)$ to variations in η_{it}, evaluated at τ_{assets} and τ_{age}.

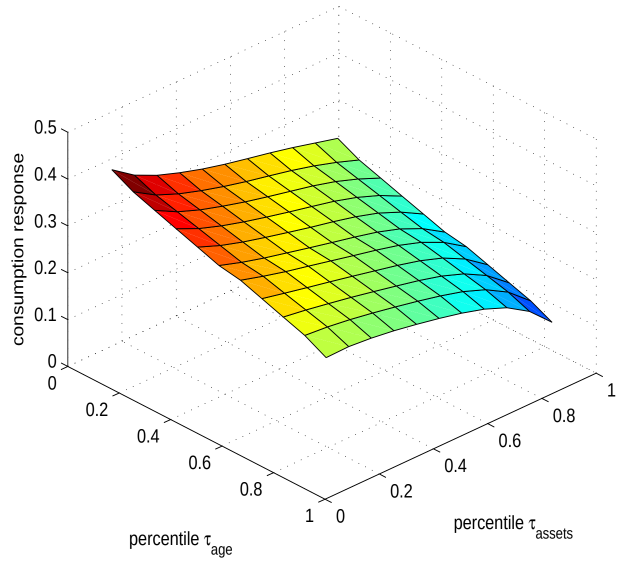
Consumption responses to y_{it}, by assets and age

$$\mathbb{E}\left[\left.\frac{\partial}{\partial y}\right|_{y_{it}} \mathbb{E}\left(c_{it}|a_{it}=a, y_{it}=y, age_{it}=age\right)\right]$$

PSID data



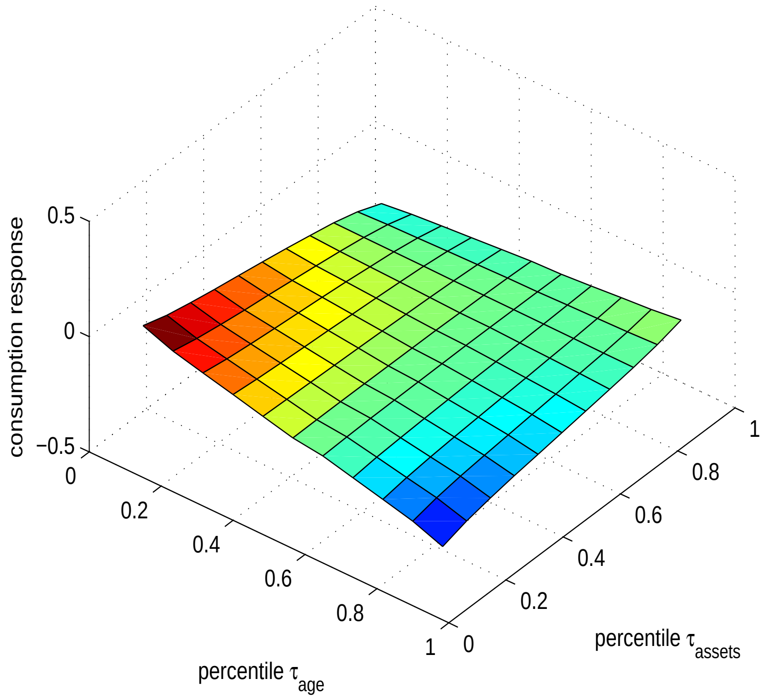
Nonlinear model



Note: 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}.

Consumption response to ε_{it}, by assets and age

$$\overline{\psi}_t(a) = \mathbb{E}\left[\frac{\partial g_t(a,\eta_{it},\varepsilon_{it},\nu_{it})}{\partial \varepsilon}\right], \text{ nonlinear model}$$



Note: Estimates of the average consumption response $\overline{\psi}_t(a)$ to variations in ε_{it}, evaluated at τ_{assets} and τ_{age}.

Model's simulation

• Simulate life-cycle earnings and consumption after a shock to the persistent earnings component (at age 37).

• We report the difference between:

– Households that are hit by a "bad" shock ($\tau_{shock} = .10$) or by a "good" shock ($\tau_{shock} = .90$).

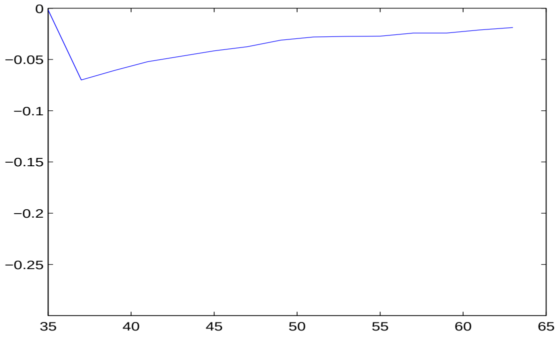
– Households that are hit by a median shock $\tau = .5$.

• Age-specific averages across 100,000 simulations. At age 35 all households have the same persistent component (percentile τ_{init}).

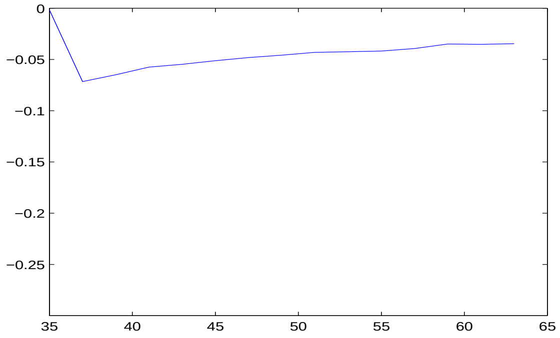
Impulse responses, earnings

Bad shock: $\tau_{shock} = .1$

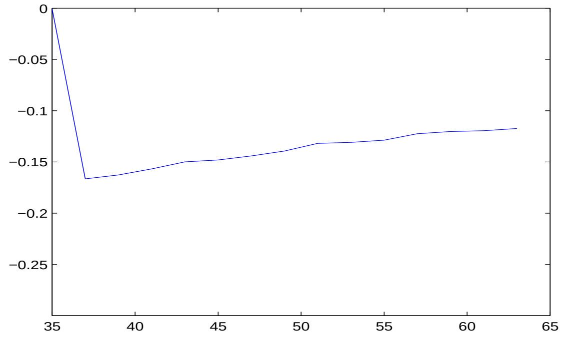
$\tau_{init} = .1$



$\tau_{init} = .5$

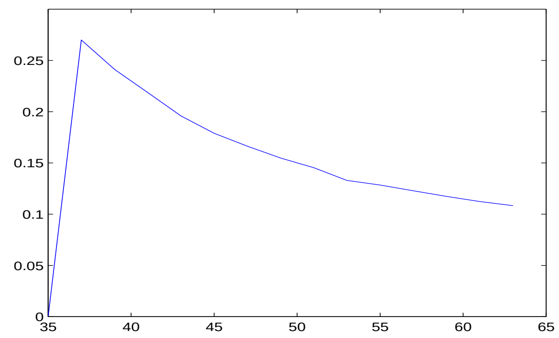


$\tau_{init} = .9$

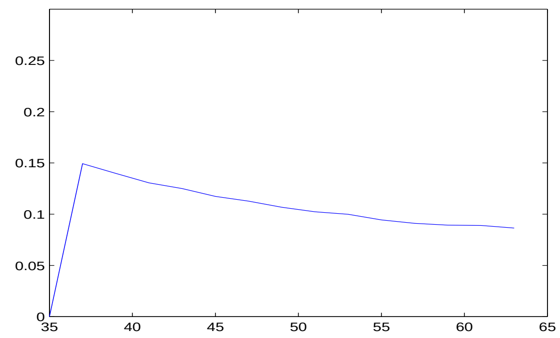


Good shock: $\tau_{shock} = .9$

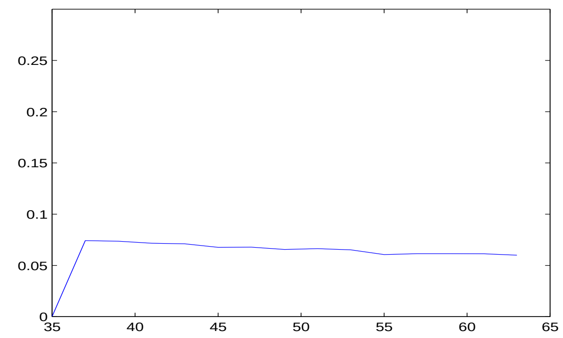
$\tau_{init} = .1$



$\tau_{init} = .5$



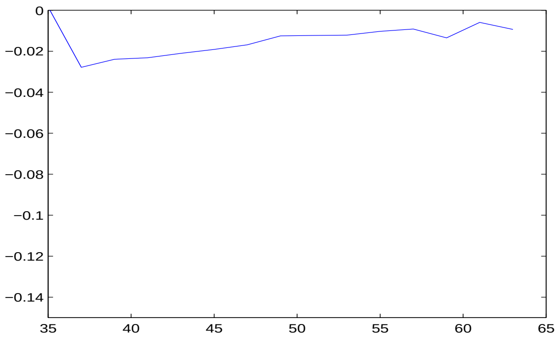
$\tau_{init} = .9$



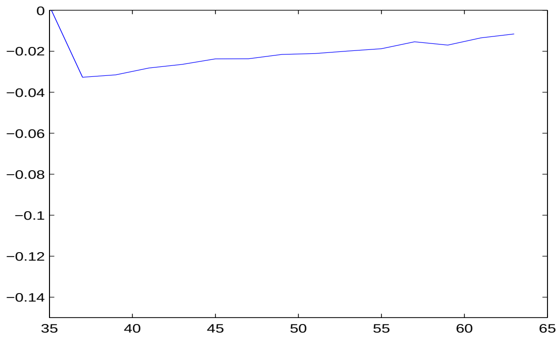
Impulse responses, consumption

Bad shock: $\tau_{shock} = .1$

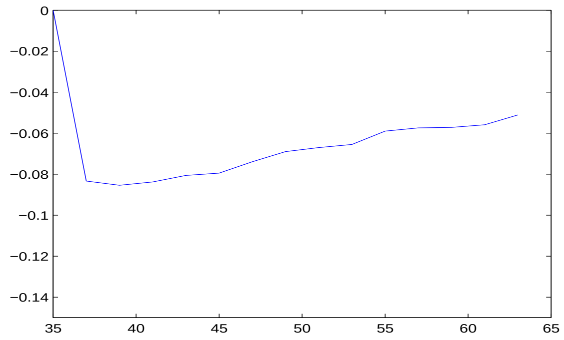
$\tau_{init} = .1$



$\tau_{init} = .5$

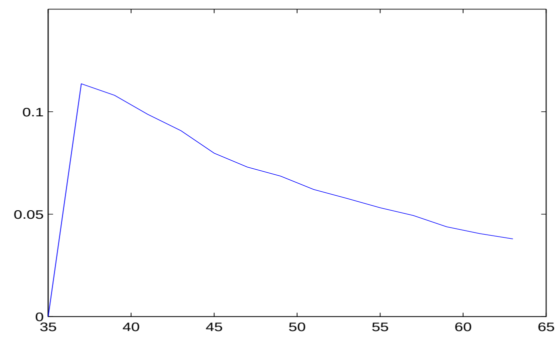


$\tau_{init} = .9$

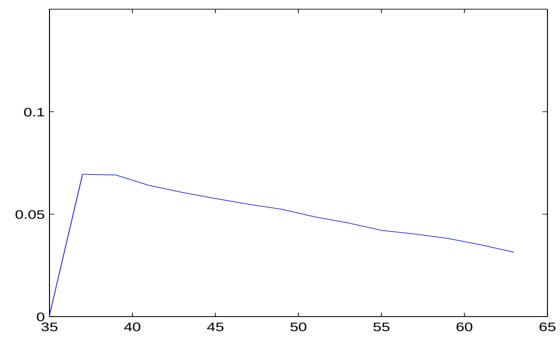


Good shock: $\tau_{shock} = .9$

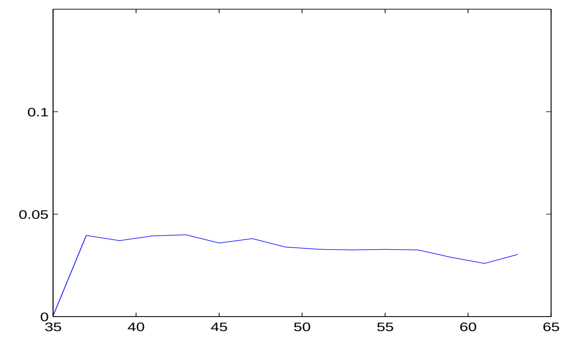
$\tau_{init} = .1$



$\tau_{init} = .5$



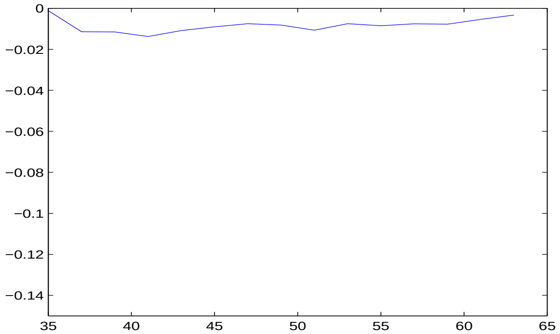
$\tau_{init} = .9$



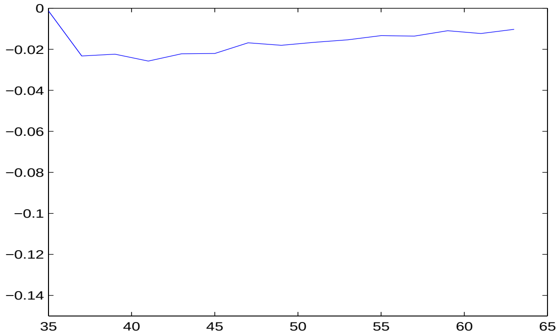
Impulse responses, consumption, household heterogeneity

Bad shock: $\tau_{shock} = .1$

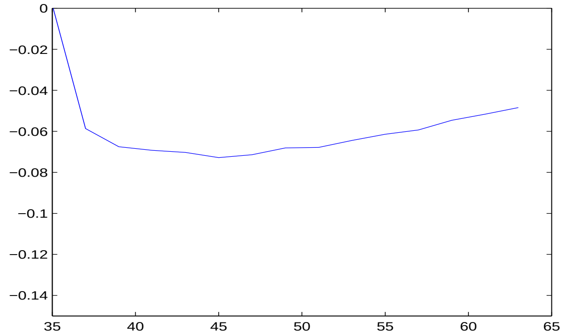
$\tau_{init} = .1$



$\tau_{init} = .5$

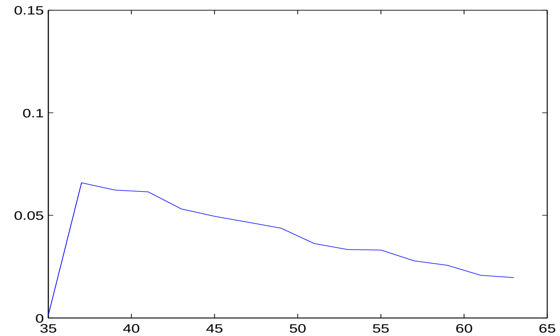


$\tau_{init} = .9$

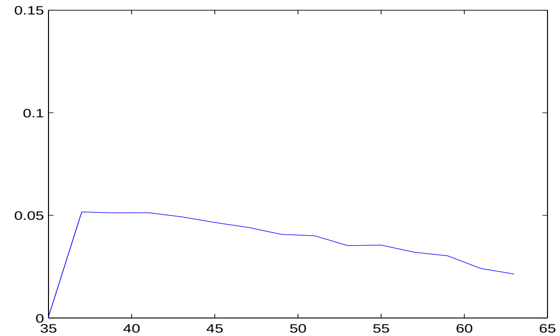


Good shock: $\tau_{shock} = .9$

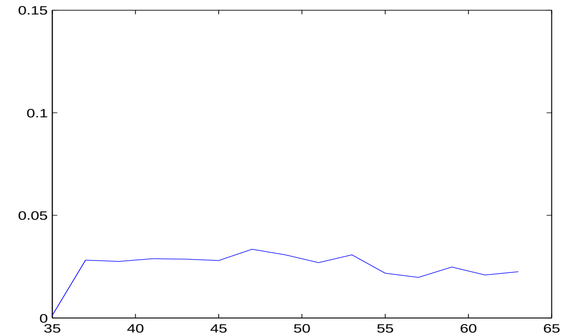
$\tau_{init} = .1$



$\tau_{init} = .5$



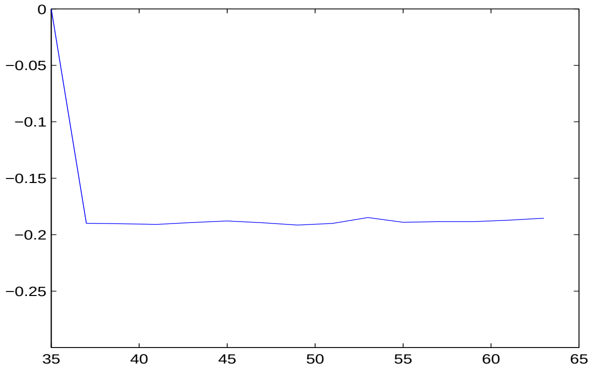
$\tau_{init} = .9$



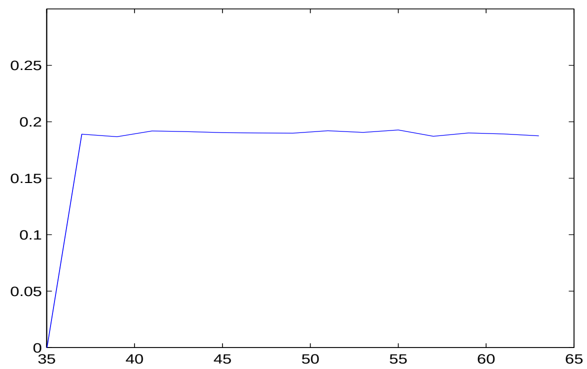
Impulse responses, canonical model

Earnings

$\tau_{shock} = .1$

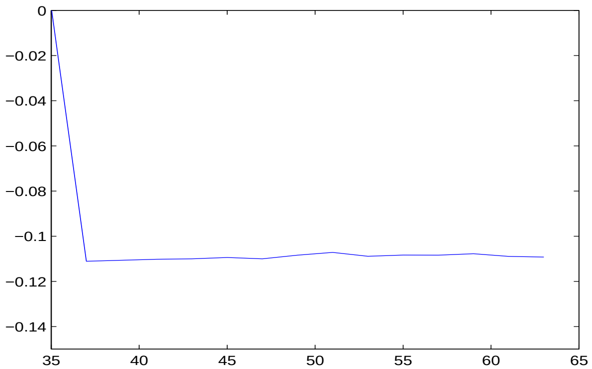


$\tau_{shock} = .9$

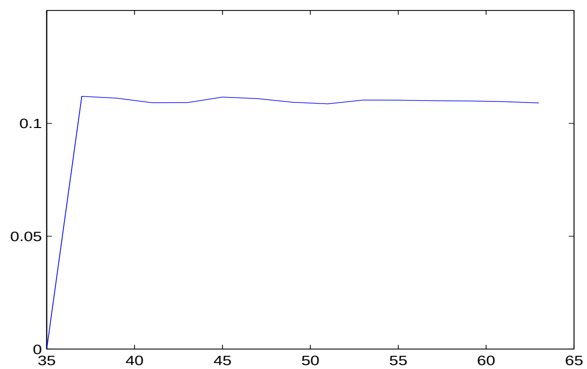


Consumption

$\tau_{shock} = .1$



$\tau_{shock} = .9$



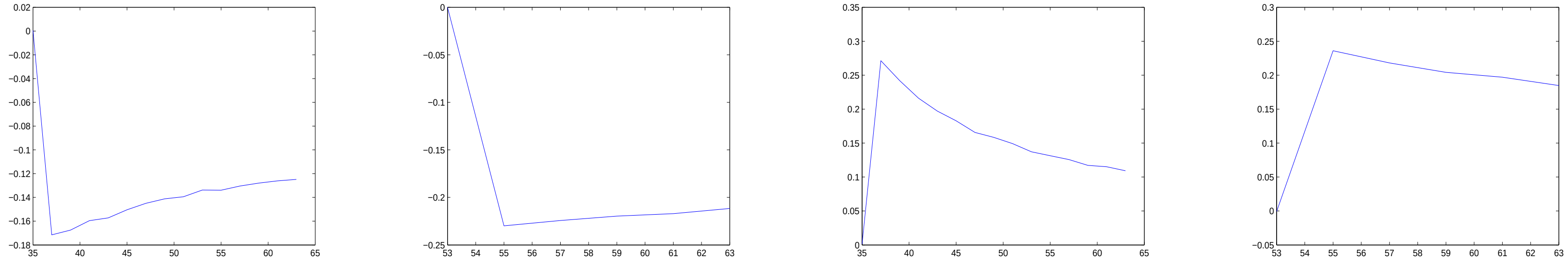
Note: Canonical earnings model and linear consumption rule.

Impulse responses, by age and initial assets

Earnings

$\tau_{init} = .9,\ \tau_{shock} = .1$ Young | Old

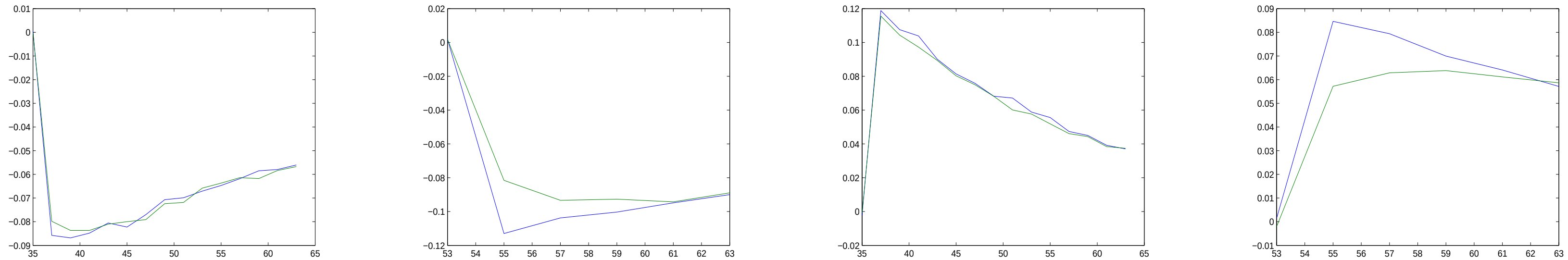
$\tau_{init} = .1,\ \tau_{shock} = .9$ Young | Old



Consumption

$\tau_{init} = .9,\ \tau_{shock} = .1$ Young | Old

$\tau_{init} = .1,\ \tau_{shock} = .9$ Young | Old



Note: Initial assets at age 35 (for "young" households) or 53 (for "old" households) are at percentile .10 (blue curves) and .90 (green curves).

Conclusion

• Developed a nonlinear framework for modeling persistence that sheds new light on the nonlinear transmission of income shocks and the nature of consumption insurance.

– A Markovian permanent-transitory model of household income, which reveals asymmetric persistence of unusual shocks in the PSID.
– An age-dependent nonlinear consumption rule that is a function of assets, permanent income and transitory income.

• We provide conditions under which the model is nonparametrically identified.

⇒ We explained how a simulation-based sequential QR method is feasible and can be used to estimate this model.

• This framework leads to new empirical measures of the degree of partial insurance.

⇒ Next step: generalize our nonlinear model to allow for other states or choices, such as evolution of household size and intensive/extensive margins of labor supply.

Additional slides

Identification when $T = 3$: Wilhelm (12)

- We work in L^2-spaces relative to suitable distributions.

- Let $g(y_2, y_3)$ such that there exists a $s(y_2)$ such that

$$\mathbb{E}\left[g(Y_2, Y_3)|Y_1\right] = \mathbb{E}\left[s(Y_2)|Y_1\right].$$

Under completeness of $Y_2|Y_1$, $s(\cdot)$ is unique.

- By conditional independence,

$$\mathbb{E}\left[\mathbb{E}\left(g(Y_2, Y_3)|\eta_2\right)|Y_1\right] = \mathbb{E}\left[\mathbb{E}\left(s(Y_2)|\eta_2\right)|Y_1\right].$$

- Under completeness of $\eta_2|Y_1$, it follows that

$$\mathbb{E}\left[g(Y_2, Y_3)|\eta_2\right] = \mathbb{E}\left[s(Y_2)|\eta_2\right].$$

The case $T = 3$ (cont.)

• Wilhelm (12) considers the functions $g_1(Y_3) = \mathbf{1}\{Y_3 \leq y_3\}$, and $g_2(Y_2, Y_3) = Y_2\mathbf{1}\{Y_3 \leq y_3\}$, for a given value y_3.

• This yields

$$
\begin{aligned}
\mathbb{E}\left[\mathbf{1}\{Y_3 \leq y_3\}|\eta_2\right] &\equiv G(\eta_2) = \mathbb{E}\left[s_1(Y_2)|\eta_2\right] \\
\mathbb{E}\left[Y_2\mathbf{1}\{Y_3 \leq y_3\}|\eta_2\right] &= \eta_2 G(\eta_2) = \mathbb{E}\left[s_2(Y_2)|\eta_2\right].
\end{aligned}
$$

• Hence, taking Fourier transforms (i.e., $\mathcal{F}(h)(u) = \int h(x)e^{\boldsymbol{i}ux}dx$),

$$
\begin{aligned}
\mathcal{F}(G)(u) &= \mathcal{F}(s_1)(u)\psi_{\varepsilon_2}(-u) \\
\boldsymbol{i}^{-1}d\mathcal{F}(G)(u)/du &= \mathcal{F}(s_2)(u)\psi_{\varepsilon_2}(-u),
\end{aligned}
$$

where $\psi_{\varepsilon_2}(u) = \mathcal{F}(f_{\varepsilon_2})(u)$ is the characteristic function of ε_2, and $\boldsymbol{i} = \sqrt{-1}$.

The case $T = 3$ (cont.)

• This yields the following first-order differential equation

$$\mathcal{F}(s_1)(-u)\frac{d\psi_{\varepsilon_2}(u)}{du} = \left[\frac{d\mathcal{F}(s_1)(-u)}{du} - \boldsymbol{i}\mathcal{F}(s_2)(-u)\right]\psi_{\varepsilon_2}(u).$$

• In addition, $\psi_{\varepsilon_2}(0) = 1$.

• This ODE can be solved in closed form for $\psi_{\varepsilon_2}(\cdot)$, provided that $\mathcal{F}(s_1)(u) \neq 0$ for all u (which is another injectivity condition).

• As a result, the distribution of ε_2, and the distribution of Y_3 given η_2, are both nonparametrically identified.

Descriptive statistics (means)

	1999	2001	2003	2005	2007	2009
Earnings	85,001	93,984	100,281	106,684	119,039	122,908
Consumption	30,182	35,846	39,843	47,636	52,175	50,583
Assets	266,958	315,866	376,485	399,901	501,590	460,262

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

- Compared to BPS (12), households in our balanced sample have higher assets, and to a less extent higher earnings and consumption.

Consumption response, two-period model

- CRRA utility. The Euler equation is (assuming $\beta(1+r)=1$)

$$C_1^{-\gamma} = \mathbb{E}_1\left[((1+r)A_2 + Y_2)^{-\gamma}\right],$$

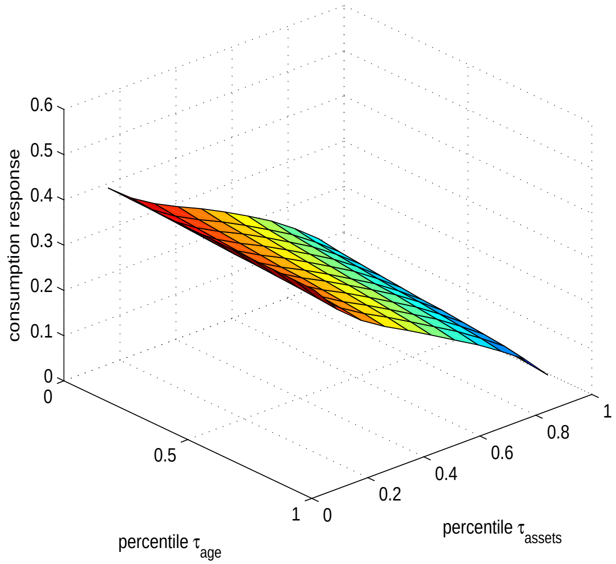
where $\gamma > 0$ is risk aversion and we have used the budget constraint $A_3 = (1+r)A_2 + Y_2 - C_2 = 0$.

- Let $X_1 = (1+r)A_1 + Y_1$, $R = (1+r)X_1 + \mathbb{E}_1(Y_2)$, and $Y_2 = \mathbb{E}_1(Y_2) + \sigma W$. Expanding as $\sigma \to 0$ we obtain

$$C_1 \approx \underbrace{\frac{(1+r)X_1 + \mathbb{E}_1(Y_2)}{2+r}}_{\text{certainty equivalent}} \underbrace{-\frac{\gamma+1}{2R}\mathbb{E}_1(W^2)}_{\text{precautionary-variance}} \underbrace{+\frac{(2+r)(\gamma+1)(\gamma+2)}{6R^2}\mathbb{E}_1(W^3)}_{\text{precautionary-skewness}}.$$

Consumption response to η_{it}, by assets and age, household heterogeneity

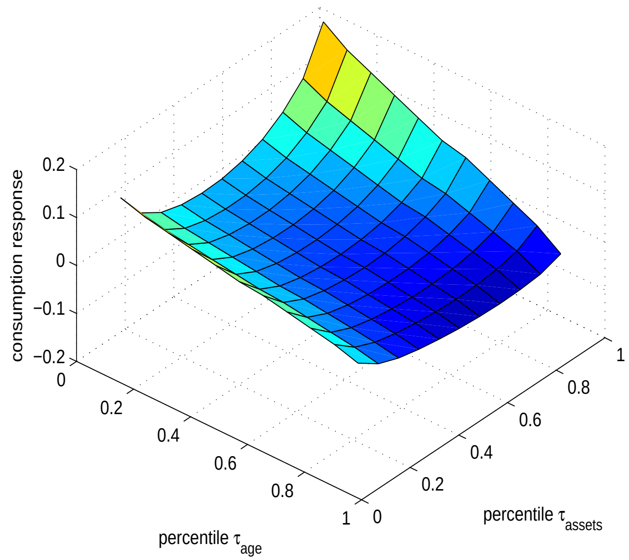
$$\overline{\phi}_t(a) = \mathbb{E}\left[\frac{\partial g_t(a,\eta_{it},\varepsilon_{it},\xi_i,\nu_{it})}{\partial \eta}\right], \text{ nonlinear model}$$



Note: Estimates of the average consumption response $\overline{\phi}_t(a)$ to variations in η_{it}, evaluated at τ_{assets} and τ_{age}.

Consumption response to ε_{it}, by assets and age, household heterogeneity

$$\overline{\psi}_t(a) = \mathbb{E}\left[\frac{\partial g_t(a,\eta_{it},\varepsilon_{it},\xi_i,\nu_{it})}{\partial \varepsilon}\right], \text{ nonlinear model}$$



Note: Estimates of the average consumption response $\overline{\psi}_t(a)$ to variations in ε_{it}, evaluated at τ_{assets} and τ_{age}.