Nonlinear Persistence and Partial Insurance
Income Dynamics and Consumption Insurance in the PSID

Manuel Arellano (Cemfi)
Richard Blundell (UCL & IFS)
Stephane Bonhomme (Chicago)

PSID 50^{th} Celebration Session

AEA Meetings, Philadelphia, 2018

January 5, 2018

Income Dynamics and Partial Insurance

The idea behind this research is to examine the transmission of income "shocks" to consumption.

Income Dynamics and Partial Insurance

The idea behind this research is to examine the transmission of income "shocks" to consumption.

> Much of the existing work on the PSID - Blundell and Preston (BP, 1998), Krueger and Perri (2005), Blundell, Pistaferri and Preston (BPP, 2008),..., Attanasio and Pistaferri (2016),....

> Recent work on the PSID incorporates family labor supply and nonseparabilities, e.g. Blundell, Pistaferri and Saporta (BPS, 2016).

> The focus in this talk is on *nonlinear persistence and partial insurance*, developed in Arellano, Blundell and Bonhomme (ABB, 2017).

Income Dynamics and Partial Insurance

The idea behind this research is to examine the transmission of income "shocks" to consumption.

> Much of the existing work on the PSID - Blundell and Preston (BP, 1998), Krueger and Perri (2005), Blundell, Pistaferri and Preston (BPP, 2008),..., Attanasio and Pistaferri (2016),....

> Recent work on the PSID incorporates family labor supply and nonseparabilities, e.g. Blundell, Pistaferri and Saporta (BPS, 2016).

> The focus in this talk is on *nonlinear persistence and partial insurance*, developed in Arellano, Blundell and Bonhomme (ABB, 2017).

In particular,

1. to consider *alternative ways of modelling persistence*, and
2. to explore *the nonlinear nature of income shocks and the implications for consumption dynamics and inequality*.

Income Dynamics and Partial Insurance

The idea behind this research is to examine the transmission of income "shocks" to consumption.

> Much of the existing work on the PSID - Blundell and Preston (BP, 1998), Krueger and Perri (2005), Blundell, Pistaferri and Preston (BPP, 2008),..., Attanasio and Pistaferri (2016),....

> Recent work on the PSID incorporates family labor supply and nonseparabilities, e.g. Blundell, Pistaferri and Saporta (BPS, 2016).

> The focus in this talk is on *nonlinear persistence and partial insurance*, developed in Arellano, Blundell and Bonhomme (ABB, 2017).

In particular,

1. to consider *alternative ways of modelling persistence*, and
2. to explore *the nonlinear nature of income shocks and the implications for consumption dynamics and inequality*.

⇒ key data: PSID, also Norwegian Population Registers ....

New data on consumption and family income sources

I. Newly designed panel data: PSID since 1999.

- Collection of consumption and assets had a major revision in 1999
  - ˜70% of consumption expenditures, [90+% since 2004.]
  - The sum of food at home, food away from home, gasoline, health, transportation, utilities, etc.
- Earnings and hours for all earners; Assets measured in each wave.
  - e.g. Blundell, Pistaferri and Saporta-Eksten (2016).

II. Administrative linked data: e.g. Norwegian population register.

- Linked registry databases with unique individual identifiers.
  - Containing records for *every Norwegian from 1967 to 2014.*
  - Detailed demographic and socioeconomic information (market income, cash transfers). Hours of work. Recent *links to real estate and assets*; => New consumption measurements.
- Family identifiers allow to match spouses and children.
  - e.g. Blundell, Graber and Mogstad (2015).

The prototypical panel data model

- A prototypical "canonical" of (log) family (earned) income y_{it} is:

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

where y_{it} is net of a *systematic component*, η_{it} is a *random walk* with innovation v_{it},

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

and ε_{it} is a *transitory shock*.

The prototypical panel data model

- A prototypical "canonical" of (log) family (earned) income y_{it} is:

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

where y_{it} is net of a *systematic component*, η_{it} is a *random walk* with innovation v_{it},

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

and ε_{it} is a *transitory shock*.

- Consumption growth is then related to income shocks:

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

where c_{it} is log total consumption net of a systematic component,
> ϕ_t is the *transmission* of persistence shocks v_{it}, and
> ψ_t the *transmission* of transitory shocks;
- the ν_{it} are taste shocks, assumed to be independent across periods.

Covariance Restrictions and Partial Insurance Parameters

Baseline panel data model specification:

$$\triangle c_{it} = \phi v_{it} + \psi \varepsilon_{it} + \nu_{it},$$

$$\triangle y_{it} = v_{it} + \triangle \varepsilon_{it},$$

Implying covariance restrictions:

$$var(\triangle c_{it}) = \phi^2 \sigma_v^2 + \psi^2 \sigma_\varepsilon^2 + \sigma_v^2$$

$$var(\triangle y_{it}) = \sigma_\eta^2 + 2\sigma_\varepsilon^2$$

$$cov(\triangle y_{it} \triangle y_{it-1}) = -\sigma_\varepsilon^2$$

$$cov(\triangle c_{it} \triangle y_{it}) = \phi \sigma_v^2 + \psi \sigma_\varepsilon^2$$

$$cov(\triangle c_{it-1} \triangle y_{it}) = -\psi \sigma_\varepsilon^2$$

> *For T>3, BPP* include time(age) variation in the σ_*^2 and insurance parameters,
> Allow for measurement error and extend to MA(1) transitory shocks,
> *BP* develop these covariance restrictions for repeated cross-sections.

Nonlinear Income Dynamics

This "*standard*" framework implies a set of covariance restrictions,

$\Rightarrow$ BPP show (over-)identification & efficient estimation via nonlinear GMM.

Nonlinear Income Dynamics

This "*standard*" framework implies a set of covariance restrictions,

$\Rightarrow$ BPP show (over-)identification & efficient estimation via nonlinear GMM.

Linearity of the income (or wage) process simplifies estimation,

$\triangleright$ however, it *rules out the nonlinear transmission of shocks and* general interactions between assets and shocks in the measures of partial insurance.

The aim in the new work is to step back and take a different tack:

$\triangleright$ develop *an alternative approach to modeling persistence* in which the impact of past shocks can be altered by the size and sign of new shocks.

- The idea is to have a framework that allows:

$\Rightarrow$ "*unusual*" shocks to wipe out the memory of past shocks, and

$\Rightarrow$ future persistence of a current shock to depend on the future shocks.

Nonlinear Income Dynamics

This "*standard*" framework implies a set of covariance restrictions,

⇒ BPP show (over-)identification & efficient estimation via nonlinear GMM.

Linearity of the income (or wage) process simplifies estimation,

▷ however, it *rules out the nonlinear transmission of shocks and* general interactions between assets and shocks in the measures of partial insurance.

The aim in the new work is to step back and take a different tack:

▷ develop *an alternative approach to modeling persistence* in which the impact of past shocks can be altered by the size and sign of new shocks.

- The idea is to have a framework that allows:

⇒ "*unusual*" shocks to wipe out the memory of past shocks, and

⇒ future persistence of a current shock to depend on the future shocks.

- Nonlinear persistence and the presence of "unusual" shocks is shown to match the data well, and has a key impact on consumption and saving over the life cycle.

Nonlinear Persistence

• Let y_{it} denote log-labor income, net of a systematic component, for i of age t

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

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

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

• Consider the following measure of persistence:

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

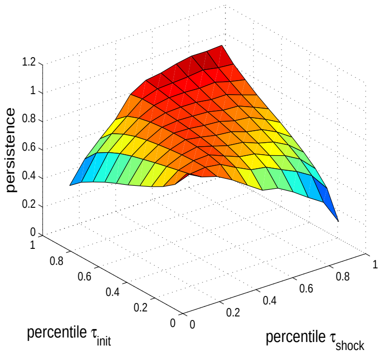
$\Rightarrow \rho_t(\eta_{i,t-1}, \tau)$ measures the persistence of $\eta_{i,t-1}$ when, at age t, it is hit by a shock u_{it} that has rank τ. Measures the *persistence of histories*.

$\triangleright$ Allows general conditional heteroscedasticity, skewness and kurtosis.

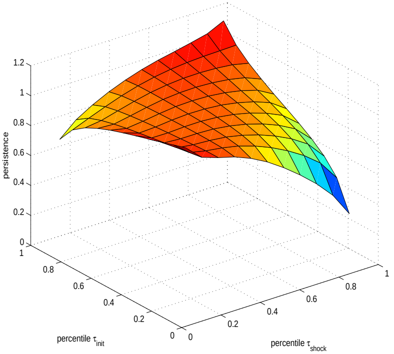
– But what is the evidence for such nonlinearities in persistence?

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

PSID data



Norwegian administrative data



Note: Household labor earnings, Age 30-59, 1999-2009 (US), 2005-2014 (Norway).

Estimates of the average derivative of the conditional quantile function of y_{it} *given* $y_{i,t-1}$

An empirical model for consumption and partial insurance

- Model consumption as a flexible/'nonparametric' nonlinear rule.

- Let c_{it} and a_{it} denote log-consumption and 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} taste shifter.

An empirical model for consumption and partial insurance

- Model consumption as a flexible/'nonparametric' nonlinear rule.
- Let c_{it} and a_{it} denote log-consumption and assets (beginning of period) net of age dummies. Our empirical specification is based on

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

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

> Consistent with the standard life-cycle model, e.g. Kaplan and Violante (2010).

> Consumption responses to η and ε are given by transmission/**partial insurance** coefficients

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

Extensions for habits and unobserved heterogeneity

• 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.

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

$\triangleright$ also cases where the consumption rule depends on lagged η, or when η follows a second-order Markov process. (See Section 3 in *ABB*, 2017).

Extensions for habits and unobserved heterogeneity

• 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.

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

▷ also cases where the consumption rule depends on lagged η, or when η follows a second-order Markov process. (See Section 3 in *ABB*, 2017).

• Households can also differ in their initial productivity η_1 and initial assets, the panel data provide opportunities to allow for additional, *unobserved heterogeneity* in earnings and consumption.

• E.g. consumption rule with *unobserved heterogeneity:*

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

Data: PSID

- (New) PSID 1999 - 2009, we use 6 waves (every other year), as in *BPS*.

- C_{it}: Information on food expenditures, rents, health expenditures, utilities, car-related expenditures, ...

- A_{it}: Assets holdings are the sum of financial assets, real estate value, pension funds, and car value, net of mortgages and other debt. (Net worth).

- y_{it} are residuals of log total pre-tax household labor earnings on a set of demographics. Note, c_{it} and a_{it} are residuals, using the same set of demographics as for earnings.

$\triangleright$ cohort and calendar time dummies, family size and composition, education, race, and state dummies.

- As in BPS, we select married male heads aged between 25 and 59.
- In this talk we focus on a balanced sub-sample of $N = 792$ households.

Specification and estimation: summary

- 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). Similarly for ε_{it} etc.

- Consumption (log) function:

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

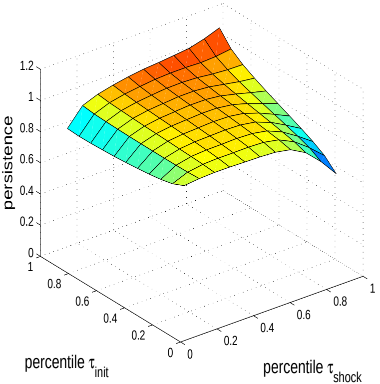
1. The first estimation step recovers estimates of the income parameters.
2. The second step recovers estimates of the consumption parameters, given an estimate of the income parameters.

- The estimation algorithm alternates between draws of latent variables from candidate posteriors and quantile regressions using those draws.

RESULTS

Nonlinear persistence of η_{it} (PSID):

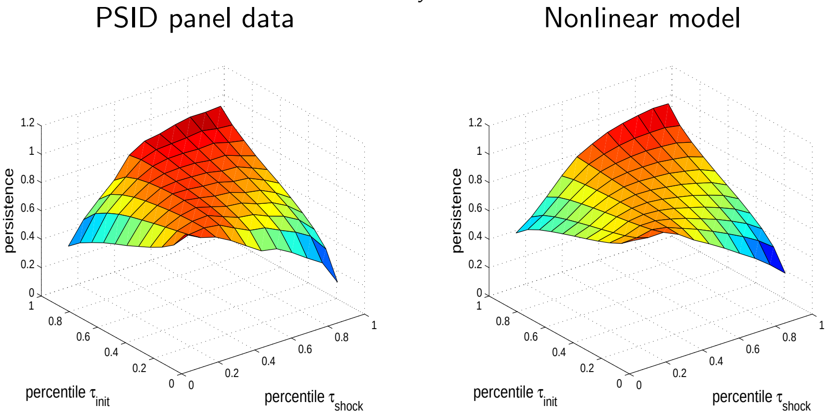
$$\rho_t(\eta_{i,t-1}, \tau) = \frac{\partial Q_{\eta_t|\eta_{t-1}}(\eta_{i,t-1}, \tau)}{\partial \eta}$$



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.

Nonlinear persistence of y_{it}

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

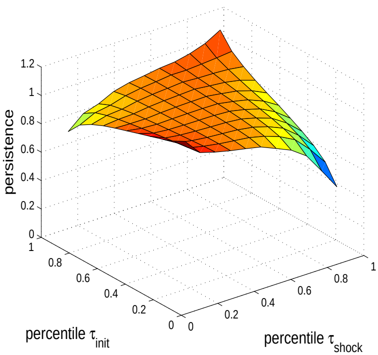


Note: Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the dist. of $y_{i,t-1}$.

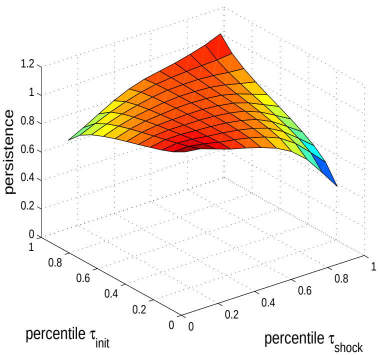
Nonlinear persistence of y_{it}

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

Norwegian register data



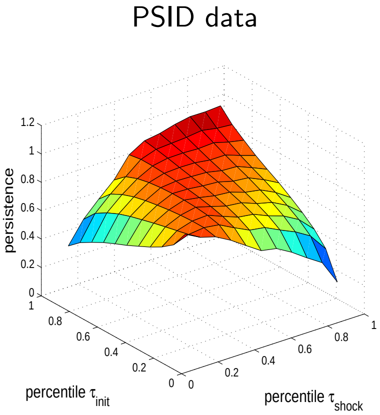
Nonlinear model

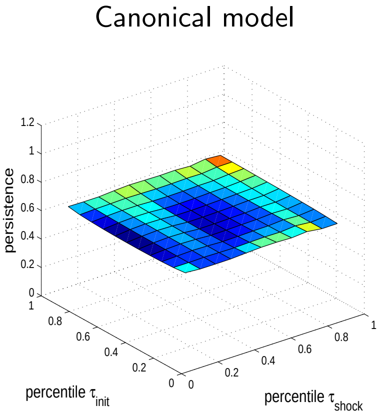


Note: Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the dist. of $y_{i,t-1}$.

Nonlinear persistence of y_{it} (cont.)

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

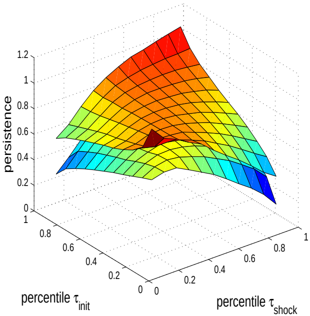




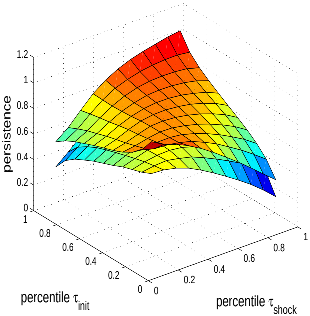
Note: Estimates of the average derivative of the conditional quantile function of y_{it} given $y_{i,t-1}$ with respect to $y_{i,t-1}$, evaluated at percentile τ_{shock} and at a value of $y_{i,t-1}$ that corresponds to the τ_{init} percentile of the dist. of $y_{i,t-1}$.

Nonlinear persistence, 95% confidence bands

(a) Earnings, PSID data



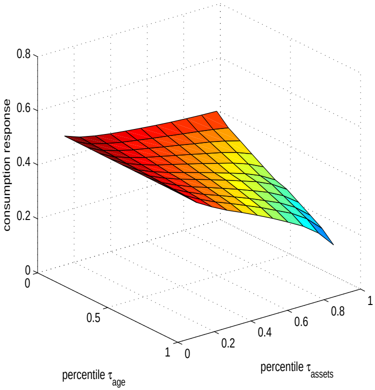
(b) Earnings, nonlinear model



Note: Pointwise 95% confidence bands. Parametric bootstrap, 500 replications.

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

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

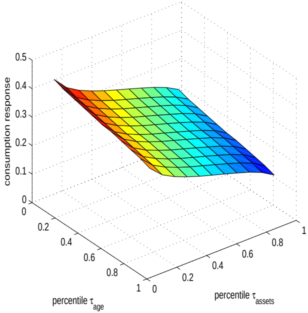


Note: Estimates of the average consumption response $\overline{\phi}_t(a)$ to variations in η_{it}, evaluated at τ_{assets} and τ_{age}.

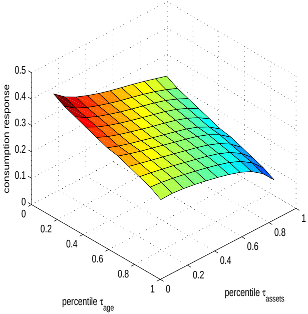
Consumption responses to y_{it}, by assets and age

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

PSID data



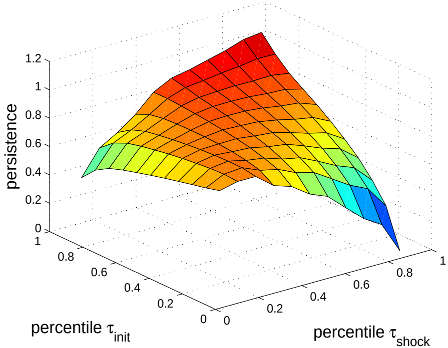
Nonlinear model



Note: Estimates of the average derivative of the conditional mean of c_{it} given y_{it}, a_{it} & age_{it} with respect to y_{it}, evaluated at values of a_{it} & age_{it} corresponding to their τ_{assets} & τ_{age} percentiles, and averaged over the values of y_{it}.

What about individual wages?

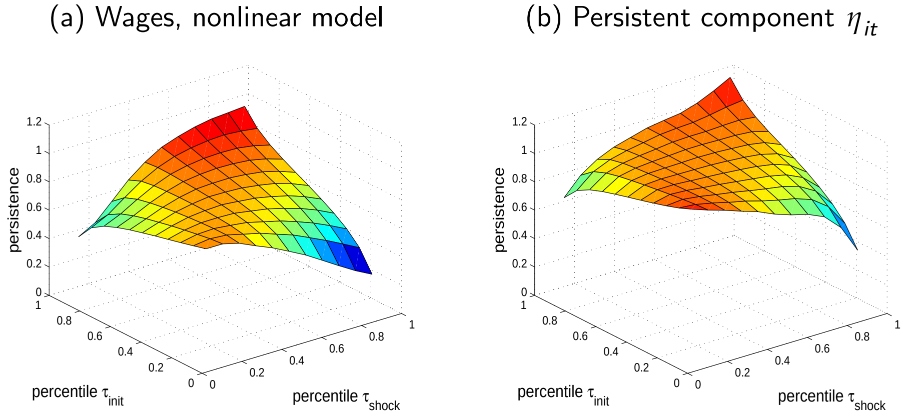
Measured Nonlinear Persistence in the PSID Male Wage Data:



Notes: Log male wages, Age 30-60 PSID 1999-2013 (US). Estimates of the average derivative of the conditional quantile function. Source: Authors calculations.

Results for hourly male wages from the nonlinear model

- find an important role for unusual shocks and nonlinear persistence.



Notes: Log male wages, Age 30-60 PSID 1999-2013 (US). Simulation of the average derivative of the estimated conditional quantile function from the η and ε.

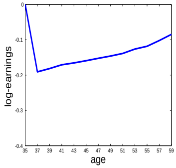
SIMULATIONS

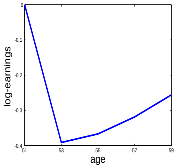
Impulse responses, by age and initial assets

Earnings

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

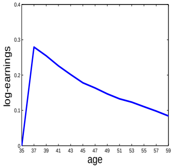
Young | Old

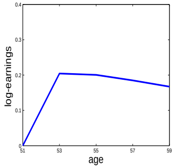




$\tau_{init} = .1,\ \tau_{shock} = .9$

Young | Old

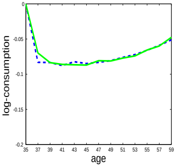


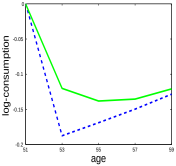


Consumption

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

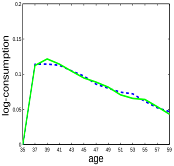
Young | Old

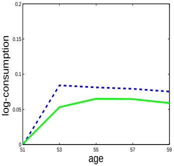




$\tau_{init} = .1,\ \tau_{shock} = .9$

Young | Old





Summary

- New framework to shed new light on income persistence and the nonlinear transmission of income shocks to consumption.

- Show the complementarities between ‘big’ administrative data, like the Norwegian registers, and purpose designed panel surveys, like the PSID.

- Exploit the important new measurements for consumption and assets in the PSID.

Summary

• New framework to shed new light on income persistence and the nonlinear transmission of income shocks to consumption.

• Show the complementarities between 'big' administrative data, like the Norwegian registers, and purpose designed panel surveys, like the PSID.

• Exploit the important new measurements for consumption and assets in the PSID.

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

Summary

• New framework to shed new light on income persistence and the nonlinear transmission of income shocks to consumption.

• Show the complementarities between ‘big’ administrative data, like the Norwegian registers, and purpose designed panel surveys, like the PSID.

• Exploit the important new measurements for consumption and assets in the PSID.

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

• This framework for nonlinear persistence in income dynamics leads to new empirical measures of the degree of partial insurance and the link between income and consumption inequality.

• Nonlinearities are also found to hold in individual male wages.

Next steps

1. Explore the impact of nonlinear persistence in wages on family labour supply and consumption smoothing, building on Blundell, Saporta-Eksten and Pistaferri (2010).
2. Include firm to firm transitions and lay-offs.
3. Separate housing equity and allow a role for local labour markets.
4. Health and other types of (partially insured) shocks.
5. and more...... [*would be great to have links with soc sec earnings etc*]

Congratulations PSID at 50!

Thanks for everything, for all those many micro-data innovations and looking forward to the next 50 years!

Additional slides

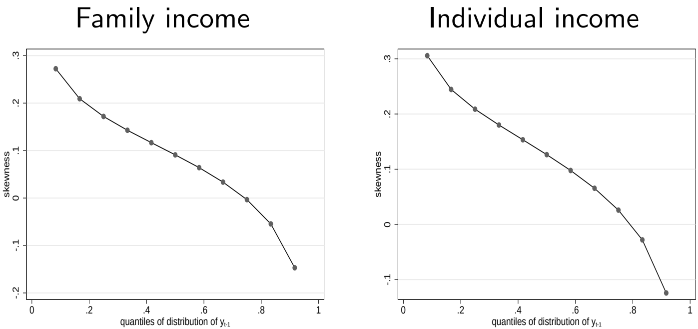
Descriptive statistics PSID (means)

	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.

Conditional skewness, Norwegian administrative data

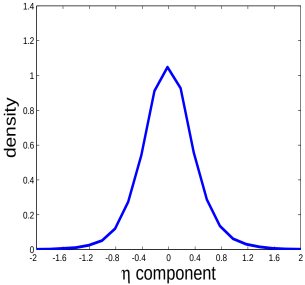


Note: Skewness measured as a nonparametric estimate of

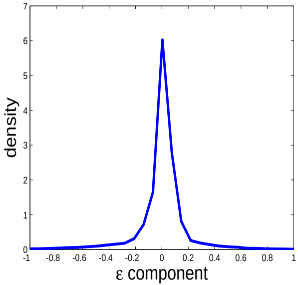
$$\frac{Q_{y_t|y_{t-1}}(y_{i,t-1},.9)+Q_{y_t|y_{t-1}}(y_{i,t-1},.1)-2Q_{y_t|y_{t-1}}(y_{i,t-1},.5)}{Q_{y_t|y_{t-1}}(y_{i,t-1},.9)-Q_{y_t|y_{t-1}}(y_{i,t-1},.1)}.$$

Figure: **Densities of persistent and transitory earnings components (PSID)**

(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.

Simulating the responses to income shocks

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

• We report the difference between:

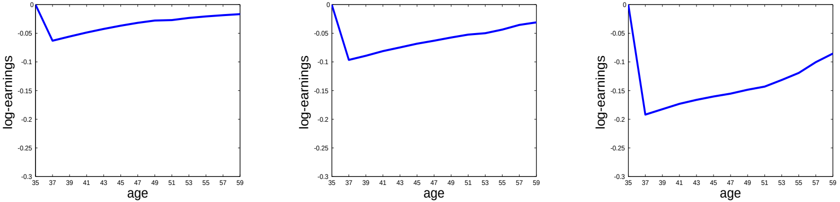
– Households that are hit by a "bad" shock ($\tau_{shock} = .10$) or by a "good" shock ($\tau_{shock} = .90$).
– Households that are hit by a median shock $\tau = .5$.

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

Impulse responses, earnings

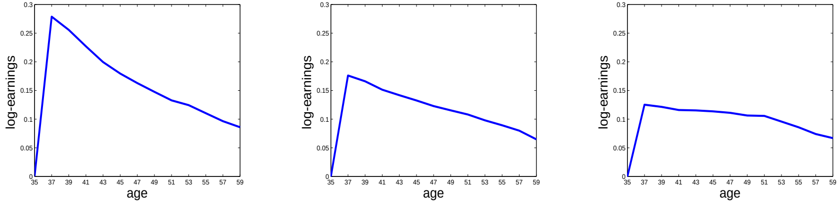
Bad shock: $\tau_{shock} = .1$

$\tau_{init} = .1$ | $\tau_{init} = .5$ | $\tau_{init} = .9$



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

$\tau_{init} = .1$ | $\tau_{init} = .5$ | $\tau_{init} = .9$



Linking Income and Consumption Dynamics

More specifically, to account for the impact of income shocks on the evolution of consumption inequality we introduce *transmission* or *partial insurance* parameters, writing consumption growth as:

$$\Delta \ln C_{it} \cong \gamma_{it} + \Delta Z_{it}'\varphi + \phi_t v_{it} + \psi_t \varepsilon_{it} + \xi_{it}$$

ϕ_t and ψ_t provide the link between the consumption and income distributions - v_{it} the permanent and ε_{it} the transitory shock to income.

Linking Income and Consumption Dynamics

More specifically, to account for the impact of income shocks on the evolution of consumption inequality we introduce *transmission* or *partial insurance* parameters, writing consumption growth as:

$$\Delta \ln C_{it} \cong \gamma_{it} + \Delta Z_{it}' \varphi + \phi_t v_{it} + \psi_t \varepsilon_{it} + \xi_{it}$$

ϕ_t and ψ_t provide the link between the consumption and income distributions - v_{it} the permanent and ε_{it} the transitory shock to income.

- For a simple benchmark intertemporal consumption model for consumer of age t, can show

$$\phi_t = (1 - \pi_{it}) \text{ and } \psi_t = (1 - \pi_{it})\gamma_{Lt}$$

where

$$\pi_{it} \approx \frac{\text{Assets}_{it}}{\text{Assets}_{it} + \text{Human Wealth}_{it}}$$

and γ_{Lt} is the annuity value of a temporary shock to income for an individual aged t retiring at age L.

[Easily extend to ARMA processes for income.]