BANCO DE ESPAÑA
Eurosistema

Income Uncertainty and Non-linear Dynamics: A Subjective Expectations Framework

Henrique S. Basso* Olympia Bover*
Jose M. Casado† Julio Galvez* Laura Hospido‡

* Banco de España
† AIReF
‡ Banco de España, CEMFI and IZA

EEA-ESEM, 29 August 2023

The views expressed here do not necessarily reflect the position of the Banco de España or the Eurosystem.

DIRECCIÓN GENERAL DE ECONOMÍA Y ESTADÍSTICA

Motivation

Although uncertainties and risks affect the way households make decisions, uncertainty is not directly observable.

Uncertainty about future income is a crucial factor that affects households' consumption and investment decisions.

As the income of a household fluctuates from one period to the next, the estimation of dynamic income processes is a key ingredient of consumption and savings models.

Large literature estimates idiosyncratic income risk considering statistical models that use realized income data.

Alternatively, surveys provide direct measures of uncertainty based on replies to probabilistic subjective expectations questions that also can be used for estimation.

Model-based measures of income uncertainty: linear

The literature has mainly focused on linear models that assume all shocks are associated with the same persistence, irrespective of the size of the shock and of the household's income history.

Formally, for a given household i and age t (relative to $t = 1$), log income y_{it} is assumed to follow an AR(1) process:

$$y_{it} = \rho y_{it-1} + \epsilon_{it}$$

These models typically assume:

1. Normality in shock distributions;
2. Linearities in persistence; and
3. Age-independence of moments of the income distribution.

Model-based measures of income uncertainty: non-linear

There has been an increasing interest on allowing for more flexible income processes and their consequences on consumption decisions:

- Existing evidence of asymmetries and non-linearities in the income process (Guvenen, Karahan, Ozkan and Song, 2021; Arellano, Blundell and Bonhomme, 2017; De Nardi, Fella and Paz-Pardo, 2020)
- ABB reported a downside risk for high-income households and an upside income risk for low-income households.
- These asymmetries induce heterogeneous consumption responses for households in different points of the distribution (high-income HHs would consume less than what a linear model would predict).

Model-based measures of income uncertainty: non-linear

Now, y_{it} is assumed to follow a general first-order Markov process.

We consider the τ conditional quantile of y_{it} given y_{it-1}:

$$y_{it} = Q_t(y_{it-1}, \tau) = Q_t(y_{it-1}, u_{it}), \ \ (u_{it} | y_{it-1}, y_{it-2}, \ldots) \sim U[0, 1]$$

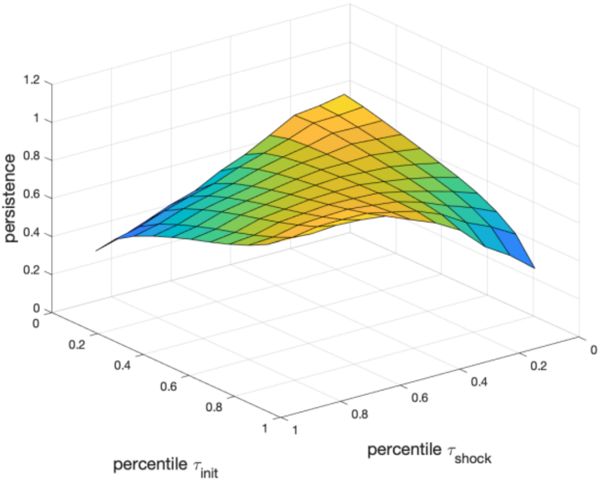
where

- u_{it} represents the rank of the income shock at age t (hence, it uniformly varies between 0 and 1)
- since u_{it} does not enter additively, the interaction between y_{it-1} and τ is allowed

The persistence of y_{it}, given y_{it-1}, when hit by a shock of rank τ:

$$\rho_t(y_{it-1}, \tau) = \frac{\partial Q_t(y_{it-1}, \tau)}{\partial y}$$

Non-linear persistence



Survey-based measures of income uncertainty: non-linear

Survey participants' probabilistic assessments of their future income provide useful information for the estimation of non-linear models.

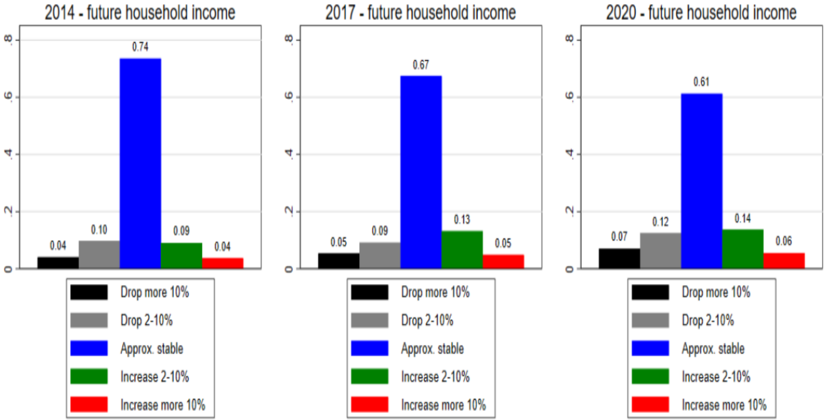
Since 2014, the Spanish Survey of Household Finances asks about:

- the expected change in household income for the next 12 months under 5 different scenarios (explicitly provided), and
- the probabilities assigned to each scenario

▸ Subjective Probabilistic Question ▸ More on the EFF

Replies to this question provide an alternative approach to estimate the conditional distribution of future household income.

Average probability distribution of future income



▸ Coherence

Subjective Expectations Framework: this paper

What we do:

- We estimate the τ conditional quantile of y_{it+1}, given y_{it}, $y_{it+1} = Q_{t+1}(y_{it}, \tau)$, and compute quantile-based measures of persistence, skewness and variance.
  - We can do this with just one cross-section.
- We compare estimates to the ones obtained from realized income.

If both y_{it} and y_{it+1} are observed, one can estimate the parameters of this process via ordinary quantile regressions, and obtain estimates of the income process and the corresponding moments (persistence, conditional skewness, and conditional variance).

Subjective Expectations Framework: this paper

However, y_{it+1} is unobserved. The survey asks households to distribute 10 points among those 5 different scenarios. What is observed are M realizations of the categorical variable j_{it}. This variable takes on five categories (hence, $J = 5$) and $M = 10$:

$$j_{it} = \begin{cases} 1, & \text{if } \Delta y_{it} \leq -0.10 \\ 2, & \text{if } -0.10 \leq \Delta y_{it} \leq -0.02 \\ 3, & \text{if } -0.02 \leq \Delta y_{it} \leq 0.02 \\ 4, & \text{if } 0.02 \leq \Delta y_{it} \leq 0.10 \\ 5, & \text{if } \Delta y_{it} \geq 0.10 \end{cases}$$

When households allocate the points, for a given y_{it}, they have a particular realization of y_{it+1} in mind that comes from a given distribution $f(y_{it+1}|y_{it};\theta)$.

We do not observe that distribution, but we can estimate it using the information coming from the j_{it}'s.

Subjective Expectations Framework: estimation

Stochastic EM algorithm: from an initial parameter vector $\widehat{\theta}^{(0)}$, we iterate on the following two steps until convergence of the $\widehat{\theta}^{(s)}$

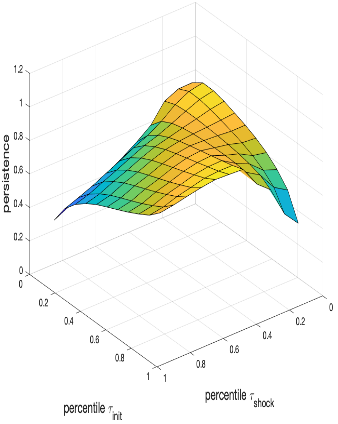
1. **E-step**: Draw realizations of $y_{it+1}^{(m,s)}$, $m = 1, \ldots, M$ from the density $f(y_{it+1}, j_{it}^{(M)} | y_{it}; \widehat{\theta}^{(s)})$ with $j_{it}^{M} = (j_{it}^{(1)}, j_{it}^{(2)}, \ldots, j_{it}^{(M)})$
   - This density is available in closed form, and from this likelihood we will generate the unobserved y_{it+1}

     ▸ Likelihood

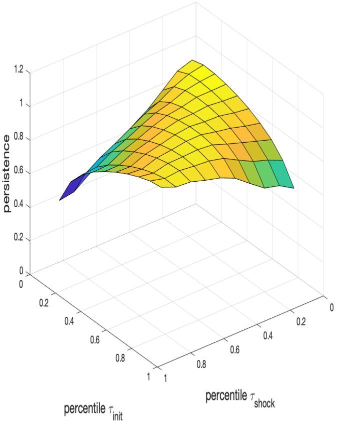
2. **M-step**: Estimate quantile regressions of $y_{it+1}^{(m,s)}$ on y_{it}, pooling across households, time periods, and realizations. This results in an updated estimate of θ, $\widehat{\theta}^{(s+1)}$

Non-linear persistence

Realized income (2014-2020)

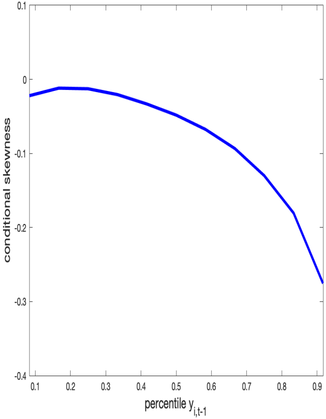


Subjective expectations (2014-2020)

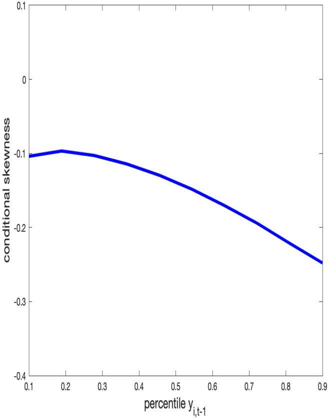


Skewness

Realized income (2014-2020)

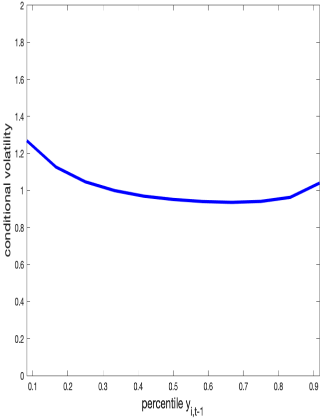


Subjective expectations (2014-2020)

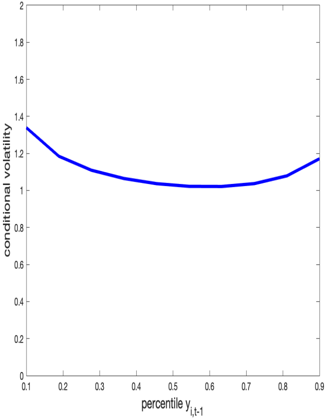


Variance

Realized income (2014-2020)



Subjective expectations (2014-2020)



Taking Stock

- This paper estimate households income uncertainty using predictive income distributions based on both realized income and subjective expectations data.
- When using realized income for estimation, income risk displays non-linear persistence:
  - (very) positive shocks for low income households and (very) negative shocks for high income households are associated with lower persistence than other shocks
  - *unusual* shocks have a higher propensity to wipe out the history of past shocks
- Asymmetries also present in the estimation from expectations.
- The existence of non-linearities in the income process matters because they induce heterogeneous consumption choices across households.

BANCO DE ESPAÑA
Eurosistema

THANK YOU!

DIRECCIÓN GENERAL DE ECONOMÍA Y ESTADÍSTICA

Data

EFF (Spanish Survey of Household Finances):

- contains measures of income, consumption and wealth for representative samples of households in 2002, 2005, 2008, 2011, 2014, 2017, and 2020
- rotating panel component

Sample: ▸ Descriptive Statistics

- we pool households present in 2002-2020 (14,818 observations)
- with household head aged 25-65 and at least one active adult per wave

Main variables:

- Total household income (before taxes)
- Total household consumption (durable & non-durable expenses)
- Household net wealth (real & financial assets − total debt)

▸ Back

Descriptive statistics

	Mean	Standard deviation (sd)
Age (years)	46.82	9.62
Women (=1)	0.40	0.49
Primary education (=1)	0.41	0.49
Secondary education (=1)	0.31	0.46
College education (=1)	0.28	0.45
Family size (number)	3.38	1.29
Second adult earner (=1)	0.52	0.50
Third (or more) adult earner (=1)	0.13	0.34

	Consumption		Earnings		Wealth	
Year	Mean	sd	Mean	sd	Mean	sd
2002	19.62	15.16	45.75	46.57	858.96	11300
2005	19.44	15.54	50.78	62.60	1030.87	3764
2008	19.63	15.96	53.51	73.89	1173.64	4947.7
2011	19.21	16.07	51.42	66.35	1018.82	3441.67
2014	19.33	19.36	49.85	92.54	983.2	3908.21
2017	19.71	18.35	57.44	133.18	1153.66	5755.2
2020	19.20	10.85	61.89	200.77	1211.17	8027.40

Notes: Monetary quantities in 2017 thousands €.

Back

Subjective Probabilistic Question

We are interested in knowing how you think the total annual income of your household will change in the next 12 months.

Share out 10 points among the five options given below, assigning more points to the options you think are more likely and 0 points to options you think are impossible:

- ☐ *Drop of more than 10%*
- ☐ *Drop between 2% and 10%*
- ☐ *Approximately steady (falls or rises of no more than 2%)*
- ☐ *Increase between 2% and 10%*
- ☐ *Increase of more than 10%*
- ☐ *DK/DA*

Back

Likelihood for a given i and j

To facilitate the discussion, we write the likelihood function using a two-period formulation: we would like to write the distribution of y_{i2}, j_{i1}, y_{i1} for a given individual: $f(y_{i2}, j_{i1}, y_{i1})$.

Notice that $f(y_{i2}, j_{i1}, y_{i1}) = \Pr(j_{i1} = j|y_{i2}, y_{i1})f(y_{i2}|y_{i1})f(y_{i1})$. This is equivalent to:

$$f(y_{i2}, j_{i1}|y_{i1}) = \Pr(j_{i1} = j|y_{i2}, y_{i1})f(y^*_{i2}|y_{i1})$$

For each j, let d_{ij1} be a dummy variable equal to one when a household gives a point to a particular interval, and zero otherwise. We can re-write the density above as:

$$f(y_{i2}, j_{i1}|y_{i1}) = \prod_{j=1}^{J} \pi_{j1}^{d_{ij1}} f(y_{i2}|y_{i1}),$$

where $\pi_{j1} = \Pr(j_{i1} = j|y_{i2}, y_{i1})$.

Likelihood for all households

Because a household i distributes ten balls to the five categories, the entire likelihood contribution of a given household i is:

$$f(y_{i2}, j_{i1}|y_{i1}) = \prod_{m=1}^{M} \prod_{j=1}^{J} \pi_{j1}^{d_{ij1}^{(m)}} f(y_{i2}^{(m)}|y_{i1}).$$

For all households, we have the following likelihood function:

$$f(y_2, j_1|y_1) = \prod_{i=1}^{N} \prod_{m=1}^{M} \prod_{j=1}^{J} [\pi_{j1}^{(m)}]^{d_{ij1}^{(m)}} f(y_{i2}^{(m)}|y_{i1}),$$

where now $\pi_{j1}^{(m)} = \Pr(j_{i1}^{(m)} = j|y_{i2}^{(m)}, y_{i1})$. To simplify the problem, we assume that the probabilities are equal to each other: $\pi_{j1}^{(1)} = \pi_{j1}^{(2)} = \ldots = \pi_{j1}^{(10)} = \tilde{\pi}_{j1}$, leading to:

$$f(y_2, j_1|y_1) = \prod_{i=1}^{N} \prod_{m=1}^{M} \prod_{j=1}^{J} [\tilde{\pi}_{j1}]^{d_{ij1}^{(m)}} f(y_{i2}^{(m)}|y_{i1}).$$

Truncated distribution

Households' answers are based on the intervals that they are given. We can use following result for a truncated density function:

$$f(y_{it}|y_{it} \in (a,b)) = \frac{f(y_{it})}{F(b) - F(a)}.$$

Given this fact, we can write the following density function for each component density:

$$f_j(y_{i2}|y_{i1}) = \begin{cases} \frac{f(y_{i2}|y_{i1})}{F(y_{i1}+c_1)}, \text{for } j = 1 \\ \frac{f(y_{i2}|y_{i1})}{F(y_{i1}+c_j) - F(y_{i1}+c_{j-1})}, \text{for } j = 2,3,4 \\ \frac{f(y_{i2}|y_{i1})}{1 - F(y_{i1}+c_4)}, \text{for } j = 5 \end{cases}$$

Writing this density for all individuals, we have:

$$f(y_2, j_1|y_1) = \prod_{i=1}^{N} \prod_{m=1}^{M} \prod_{j=1}^{J} \left[\pi_{j1} f_j(y_{i2}^{(m)}|y_{i1})\right]^{d_{ij1,(m)}}.$$

Coherence

	Bunching in the middle	Negative change	Positive change	Uncertainty
EFF 2014	61.93	8.37	7.14	22.56
EFF 2017	54.54	8.71	10.78	25.96
EFF 2020	42.49	9.95	9.22	38.34

Back

Correlations with characteristics, HH heads $\geq$ 25 years

VARIABLES	(1) Bunching	(2) Drop	(3) Increase	(4) Uncertain
Reference person is a woman	0.053***	0.006	-0.014	-0.045***
	(0.017)	(0.009)	(0.010)	(0.016)
Reference person is bet 35 and 45	0.065	0.026*	-0.006	-0.085*
	(0.052)	(0.016)	(0.035)	(0.051)
Reference person is bet 45 and 55	0.046	0.052***	-0.030	-0.068
	(0.050)	(0.016)	(0.034)	(0.050)
Reference person is above 55	0.196***	0.050***	-0.068**	-0.178***
	(0.051)	(0.015)	(0.034)	(0.050)
Medium levels of education	-0.066***	-0.016	0.040***	0.042**
	(0.020)	(0.011)	(0.012)	(0.019)
High levels of education	-0.068***	-0.031***	0.020	0.079***
	(0.023)	(0.012)	(0.014)	(0.021)
Wealth (in logs)	0.010*	0.001	-0.001	-0.010*
	(0.005)	(0.003)	(0.003)	(0.005)
Second income quintile	0.058	0.005	0.024	-0.087
	(0.075)	(0.057)	(0.035)	(0.077)
Third income quintile	0.055	-0.003	0.033	-0.084
	(0.066)	(0.052)	(0.031)	(0.071)
Fourth income quintile	0.037	0.009	0.025	-0.071
	(0.076)	(0.054)	(0.036)	(0.079)
Fifth income quintile	-0.105	0.155	0.054	-0.104
	(0.129)	(0.114)	(0.050)	(0.128)
Self employed	-0.068	-0.018	0.049**	0.036
	(0.046)	(0.023)	(0.024)	(0.042)
Other employment categories	0.029	-0.032*	0.026	-0.023
	(0.036)	(0.017)	(0.018)	(0.034)
Has a permanent contract	0.045	-0.054***	-0.013	0.022
	(0.036)	(0.016)	(0.017)	(0.034)
Constant	0.314***	0.088	0.073	0.525***
	(0.095)	(0.063)	(0.051)	(0.099)
Observations	9,597	9,597	9,597	9,597
R-squared	0.061	0.012	0.017	0.059

Notes: Weighted regressions. Wave dummies included. SE clustered at HH level. Back

Correlations with characteristics, HH heads 25-65

	(1)	(2)	(3)	(4)
VARIABLES	Bunching	Drop	Increase	Uncertain
Reference person is a woman	0.084***	0.003	-0.020	-0.066***
	(0.020)	(0.011)	(0.012)	(0.019)
Reference person is bet 35 and 45	0.055	0.032**	-0.002	-0.085*
	(0.052)	(0.016)	(0.035)	(0.051)
Reference person is bet 45 and 55	0.036	0.058***	-0.025	-0.069
	(0.050)	(0.017)	(0.033)	(0.050)
Reference person is above 55	0.119**	0.078***	-0.047	-0.150***
	(0.051)	(0.017)	(0.033)	(0.051)
Medium levels of education	-0.048**	-0.034***	0.039***	0.044*
	(0.024)	(0.013)	(0.015)	(0.023)
High levels of education	-0.066**	-0.046***	0.017	0.095***
	(0.027)	(0.014)	(0.017)	(0.025)
Wealth (in logs)	0.012*	-0.001	-0.002	-0.009
	(0.006)	(0.003)	(0.004)	(0.006)
Second income quintile	0.086	0.015	0.010	-0.110
	(0.085)	(0.071)	(0.047)	(0.097)
Third income quintile	0.085	0.007	0.041	-0.133
	(0.074)	(0.064)	(0.042)	(0.089)
Fourth income quintile	0.102	0.019	0.037	-0.158
	(0.084)	(0.067)	(0.048)	(0.096)
Fifth income quintile	-0.052	0.170	0.061	-0.179
	(0.131)	(0.122)	(0.058)	(0.137)
Self employed	-0.046	-0.028	0.045*	0.029
	(0.047)	(0.024)	(0.024)	(0.043)
Other employment categories	0.002	-0.026	0.050***	-0.026
	(0.038)	(0.018)	(0.019)	(0.036)
Has a permanent contract	0.057	-0.059***	-0.019	0.021
	(0.036)	(0.016)	(0.017)	(0.035)
Constant	0.251**	0.112	0.070	0.568***
	(0.103)	(0.076)	(0.059)	(0.115)
Observations	5,941	5,941	5,941	5,941
R-squared	0.045	0.023	0.015	0.055

Notes: Weighted regressions. Wave dummies included. SE clustered at HH level. ▸ Back