Applied Economics, 2013, **45**, 4004–4021

Evaluating how predictable errors in expected income affect consumption

Luigi Giamboni[a], Emanuele Millemaci[b,*] and Robert J. Waldmann[c]

[a] *Italian Ministry of Economy and Finance, Roma, Italy*
[b] *Dipartimento di Economia, Statistica, Matematica e Sociologia (DESMaS) 'V. Pareto', Universitá di Messina, Roma, Italy*
[c] *Facoltá di Economia, Universitá di Roma 'Tor Vergata', Roma, Italy*

This article studies whether anomalies in consumption can be explained by a behavioural model in which agents make predictable errors in forecasting income. We use a micro-data set containing subjective expectations about future income. This article shows that the null hypothesis of rational expectations is rejected in favour of the behavioural model, since consumption responds to predictable forecast errors. On average, agents who we predict are too pessimistic increase consumption after the predictable positive income shock. On average, agents who are too optimistic reduce the consumption.

Keywords: behavioural economics; subjective expectations; rational expectations; consumption and saving

JEL Classification: D11; D12; D84

I. Introduction

The permanent income hypothesis implies predictable changes in income should not help explain the change in consumption. We show that an apparent economic anomaly, predictable changes in consumption, may be explained relaxing the rationality hypothesis in favour of a behavioural model where agents are irrationally optimistic or pessimistic.

Since the path of future income is uncertain, the agent makes consumption decisions based on his subjective expectation about uncertain future events. The permanent income hypothesis is our null hypothesis, the alternative is a behavioural model of consumption. In this second model, the agent aims to maximize his subjective expected utility over the life cycle, but he makes predictable (for the econometrician) systematic errors in forming subjective expectations on future income. Hence, on average, an individual who has been too pessimistic in making his prediction experiences a positive surprise when income is realized and is induced to revise his consumption decision upward. Conversely, if the agent has been overly optimistic, he experiences a negative shock on income and decides to lower consumption.[1]

Our behavioural model implies a new formulation of the Euler equation where predictable errors in income forecasts help explain the first difference of consumption. This suggests that, not properly taking irrationality into account, previous research on excess sensitivity of consumption may not be correctly specified. We will show whether the coefficient on predictable changes in income changes once the predictable forecast error is introduced in the Euler equation. In particular, irrational pessimism and optimism seem to be more statistically significant explanations of the apparent anomaly in consumption than precautionary savings and liquidity constraints.

*Corresponding author. E-mail: e.millemaci@gmail.com

[1] Brown and Taylor (2006) rely on financial expectations and realizations to link past financial optimism and pessimism with current financial prediction accuracy and determine how it affects savings and consumption. They find that past financial optimism has a positive effect on current expectations formation whilst past financial pessimism has a negative effect. Financial optimism is inversely associated with saving and that current financial expectations serve to predict future consumption. From a consumer confidence survey dataset covering 10 European countries over 22 years, Bovi (2009) compare average values of the differences between prospective views versus retrospective views as well as personal views versus general views on economic stances. Findings suggest the presence of structural psychologically driven distortions in people's judgements and expectations formation.

Applied Economics ISSN 0003–6846 print/ISSN 1466–4283 online © 2013 Taylor & Francis
http://www.tandfonline.com
http://dx.doi.org/10.1080/00036846.2012.745987

Despite the theoretical statement that actual actions depend on subjective expectations about future events, economists engaged in empirical research tend to be skeptical of the use of data on subjective expectations. The main practice has become that of inferring expectations from realizations. The attempt to infer from the distribution of realizations requires the knowledge of the information set of the agent and how he uses it. Typically the researcher imposes a model of the data-generating process, which under the assumption of rational expectations describes how individuals form their expectations. The estimation strategy is to hypothesize a stochastic process for income dynamics, estimate it and project it 1 year into the future exploiting the orthogonality condition implied by the rational expectations hypothesis (Hall and Mishkin, 1982).

In contrast, direct elicitation of subjective expectations may eliminate the need for such assumptions (Dominitz and Manski, 1997; Dominitz, 1998, 2001; Flavin, 1999; Kapteyn *et al.*, 2009; Kaufmann and Pistaferri, 2009). It allows for complete heterogeneity of income expectations formation and permits one to overcome the problem that the econometrician's information set is not rich enough to reproduce the agent's information set. Kapteyn *et al.* (2009) find evidence in support of this argument. Using data from the De Nederlandsche Bank (DNB) Household Survey (DHS), which is the same dataset we use in this article, the authors use direct subjective information on respondents' expectations in the Euler equation estimation of intertemporal consumption model. Moreover, they use information on individual preferences to estimate a welfare function of income and derive knowledge on respondents' utility function. Using such direct subjective information on respondents' expectations and preferences allows to reduce the number of assumptions needed to estimate the Euler equation of intertemporal consumption models. Kapteyn *et al.* find that welfare functions and expectations have predictive power for the variation in consumption across households. Furthermore, their estimation of the intertemporal elasticity of substitution based on the estimated welfare functions is in line with other estimates found in the literature.

We use data from the Dutch DNB Household Survey (DHS). This dataset consists of approximately 5000 individual observations for each wave in the Netherlands and contains detailed information on wealth, income, work and demographic characteristics and different kinds of subjective expectations stated by the respondents, covering the period that starts in 1993.

We calculate the forecast errors as the difference between realized family income and the mean of the elicited subjective distribution of future family income. We instrument the agent's prediction error with various specifications of the information set. The weak exogeneity of the adopted instrument is assured by the null hypothesis of agents' rationality. In the second step of the two-stage Instrumental Variable (IV) estimator, we regress the first difference of the logarithm of consumption on the fitted (hence predictable) error and find strong evidence in support of our behavioural model stating that consumption responds to predictable errors in income forecasts.

It is often argued that works on the predictability of forecast errors, either rejecting or accepting the rational expectations hypothesis, do not supply evidence to support the claim that the elicited expectations really correspond to those affecting the agent's behaviour. Hence, it is important and interesting to show that, once the rational expectations hypothesis is rejected, it is possible to explain agents' consumption decisions.

As underlined by Kapteyn *et al.* (2009), one limitation in using the DHS dataset is that one is not provided with a direct measure of consumption but needs to derive it by subtracting savings from income. Another limitation is that income and savings are reported in brackets. Moreover, savings can take only nonnegative values in the DHS.

Souleles (2004) analyses the household-level data that underlies the Michigan Index of Consumer Sentiment, a dataset called the Michigan Survey of Consumer Attitudes and behaviour (CAB). The CAB contains the answers each household gave to the five questions that comprise the Michigan Index of Consumer Sentiment. To study how sentiment is related to spending, Souleles uses the most comprehensive household-level dataset on consumer expenditures, the Consumer Expenditure Survey (CEX), which also contains a rich set of household demographic indicators. Souleles links the household-level sentiment data with consumer spending by imputing the sentiment levels of households who participated in the CEX from demographically similar households who participated in the survey of CAB. Using the imputed values of expectations, Souleles finds that they are inefficient as forecast errors are found to be correlated with demographic characteristics. Moreover, he finds that some, though not all, of the excess sensitivities are due to such systematic heterogeneity in forecast errors.

Souleles' need to construct a pseudo-panel in order to match data on elicited forecasts and data on consumption is expensive as it is not possible to use an individual forecast to predict the individual's forecast error. Since there is huge variation in the income forecasts of demographically similar individuals, a powerful candidate indicator of irrational optimism or pessimism simply cannot be used with a pseudo-panel. We find that the individual forecast is by far the most powerful instrument for the individual forecast error. Some of our more striking results are based on using the forecast as our only indicator – instrumenting irrational optimism with optimism directly.

A potential problem with panel IV estimation can be caused by the correlation of disturbances across individuals. One way to check for this is to see whether the results are changed if variables which might pick up those common shocks are included in the second stage. For example, workers in manufacturing might suffer a correlated unpredictable income shock in a recession. In this case, the apparent violation of the standard Euler equation would be eliminated if indicator variables for worker in manufacturing in period t are added to the regression. Unfortunately, this removes exactly the identifying variance which is necessary in a pseudo-panel regression. True panel data makes it easier to evaluate the relevance of concerns about cross-sectional correlations of disturbances. For these reasons, we believe that our article makes a useful addition to the project begun by Souleles.

II. The Model

The path of future income is uncertain, so individuals must make their consumption plans on the basis of their subjective expectations about uncertain future events. The conventional model of life-cycle consumption under uncertainty, with isoelastic time separable utility, consumers maximizing expected utility function and perfect credit markets, becomes

$$\mathrm{Max}_{c_1,\ldots,c_T} E_t^{su}\left[\sum_{s=0}^{\infty}(1+\delta)^{-s}\frac{c_{t+s}^{1-\gamma}}{1-\gamma}\right] \tag{1}$$

subject to

$$a_{t+s+1} = (1+r)(a_{t+s} + y_{t+s} - c_{t+s}), \quad s = 0, 1, \ldots, \infty \tag{2}$$

$$a_t \quad \text{given} \tag{3}$$

$$\lim_{s\to\infty}(1+r)^{-s}a_s = 0 \tag{4}$$

where E_t^{su} is the subjective expectations operator conditional on all information available at time t and stated at the end of the period, δ is the rate of time preference, c is the consumption, y is the total family net income, r is the real rate of interest, which is assumed to be constant, and a represents assets apart from human capital.

Differentiating with respect to consumption and considering the first-order condition of equality of wealth's and consumption's marginal utilities at the optimum, we obtain the following Euler equation:

$$E_t^{su}(c_{t+1}^{-\gamma}) = \frac{1+\delta}{1+r}c_t^{-\gamma} \tag{5}$$

To illustrate our behavioural model, let us assume, for instance, that agents have a subjective distribution over the consumption growth rate, which is normal

$$\Delta\log c_{t+1}|\Omega_t \sim N(\mu_c; \sigma_c^2) \tag{6}$$

where Ω_t indicates all information available at time t; $\mu_c = E_t^{su}(\Delta\log c_{t+1})$ and $\sigma_c^2 = \mathrm{Var}_t^{su}(\Delta\log c_{t+1})$, for the sake of simplicity, will be assumed constant over time. We also assume that $e^{((1+\gamma)/2)\sigma_c^2+(\tilde{r}-\tilde{\delta})/\gamma} < 1+r$ and that the law of iterated expectations applies to subjective expectations, $E_t^{su}(E_{t+s}^{su}(x_{t+s})) = E_t^{su}(x_{t+s})$. Such an assumption means that agents are convinced that they are rational and that they will be rational in the future. We can write down Equation 5 as the following:

$$E_t^{su}\exp[-\gamma\Delta\log c_{t+1} + \log(1+r) - \log(1+\delta)] = 1 \tag{7}$$

which, in turn, is equal to

$$\exp[-\gamma\mu_c + (1/2)\gamma^2\sigma_c^2 + \tilde{r} - \tilde{\delta}] = 1 \tag{8}$$

where we have exploited the property that if $x \sim N(\mu; \sigma)$, then $E(e^x) = \exp[\mu + (1/2)\sigma^2]$, and to save on notation we have defined $\tilde{r} \cong \log(1+r)$ and $\tilde{\delta} \cong \log(1+\delta)$. Taking the logs, we have

$$-\gamma E_t^{su}(\Delta\log c_{t+1}) + (1/2)\gamma^2\sigma_c^2 + \tilde{r} - \tilde{\delta} = 0 \tag{9}$$

Splitting the logarithm and taking the exponential of both sides of the equation, we are left with

$$E_t^{su}(c_{t+1}) = c_t e^{((1+\gamma)/2)\sigma_c^2+(\tilde{r}-\tilde{\delta})/\gamma} \tag{10}$$

where we have again used the property of exponentials of normally distributed variables. Given the subjective expectations about future income held in period t, the individual's perceived budget constraint can be expressed as

$$\sum_{s=0}^{\infty}(1+r)^{-s}E_t^{su}(c_{t+s}) = a_t + \sum_{s=0}^{\infty}(1+r)^{-s}E_t^{su}(y_{t+s}) \tag{11}$$

where y_t is the labour income which is exogenous and is paid at the end of the period. Substituting in Equation 10 gives

$$\sum_{s=0}^{\infty}(1+r)^{-s}e^{s[((1+\gamma)/2)\sigma_c^2+(\tilde{r}-\tilde{\delta})/\gamma]}c_t = a_t + \sum_{s=0}^{\infty}(1+r)^{-s}E_t^{su}(y_{t+s}) \tag{12}$$

Using the assumption that $e^{((1+\gamma)/2)\sigma_c^2+(\tilde{r}-\tilde{\delta})/\gamma} < 1+r$, we obtain

$$c_t = \zeta[a_t + \sum_{s=0}^{\infty}(1+r)^{-s}E_t^{su}(y_{t+s})] \tag{13}$$

where we have defined $\zeta = 1 - \frac{e^{((1+\gamma)/2)\sigma_c^2+(\tilde{r}-\tilde{\delta})/\gamma}}{(1+r)}$.

Moreover, a_{t+1} is known at time t, so

$$c_{t+1} - E_t^{su}(c_{t+1}) = \zeta\sum_{s=0}^{\infty}(1+r)^{-s}[E_{t+1}^{su}(y_{t+s+1}) - E_t^{su}(y_{t+s+1})] \tag{14}$$

The assumption that consumption is log normally distributed implies that

$$\Delta\log c_{t+1} = (1/2)\gamma\sigma_c^2 + \frac{\tilde{r}-\tilde{\delta}}{\gamma} + \frac{\zeta}{E_t^{su}(c_{t+1})}\sum_{s=0}^{\infty}(1+r)^{-s}[E_{t+1}^{su}(y_{t+s+1}) - E_t^{su}(y_{t+s+1})] \tag{15}$$

Expectations are stated at the end of the period such that $E_{t+1}^{su}(y_{t+1}) = y_{t+1}$. If one assumes that the error in subjective forecasts of y_{t+1}, $y_{t+1} - E_t^{su}(y_{t+1})$, is uncorrelated with subjective expectations of subsequent periods, i.e. $E_{t+1}^{su}(y_{t+s+1}) - E_t^{su}(y_{t+s+1}) = 0$ for $s > 0$, the previous equation becomes

$$\Delta\log c_{t+1} = (1/2)\gamma\sigma_c^2 + \frac{\tilde{r}-\tilde{\delta}}{\gamma} + \frac{\zeta}{E_t^{su}(c_{t+1})}[y_{t+1} - E_t^{su}(y_{t+1})] \tag{16}$$

If agents have been too pessimistic, they revise their consumption decision upward. If they have been too optimistic, they revise their consumption decision down. More generally, if current forecast error is nonnegatively correlated with the subjective expectations of subsequent periods, $E_{t+1}^{su}(y_{t+s+1}) - E_t^{su}(y_{t+s+1}) = \rho^s[E_{t+1}^{su}(y_{t+1}) - E_t^{su}(y_{t+1})]$ and $\rho > 0$, we get the following equation:

$$\Delta\log c_{t+1} = (1/2)\gamma\sigma_c^2 + \frac{\tilde{r}-\tilde{\delta}}{\gamma} + \frac{1+r}{1+r-\rho}\frac{\zeta[y_{t+1} - E_t^{su}(y_{t+1})]}{E_t^{su}(c_{t+1})} \tag{17}$$

This is the case if the agents believe that income innovations are persistent. If $\rho = 1$, agents believe that the unpredicted disturbance to y_t is a random walk with $E_t^{su}(y_{t+s+1} - y_{t+s}) = 0$. Alternatively, if $\rho = 1$ and agents believe that the unpredicted disturbance to y_t is an Integrated Moving Average (IMA) of the first-order (IMA(1)) with a Moving Average (MA) coefficient equal to $-\theta$, we have $E_t^{su}(y_{t+s+2} - y_{t+s}) = 0$ for $s \geq 0$ and $E_t^{su}(y_{t+2} - y_{t+1}) = \theta[y_{t+1} - E_t^{su}(y_{t+1})]$. That is the case if agents experience a surprise in period $t+1$ and they are convinced that the shock will persist in the future.

In this case, Equation 15 becomes

$$\Delta \log c_{t+1} = (1/2)\gamma\sigma_c^2 + \frac{\tilde{r} - \tilde{\delta}}{\gamma} + \frac{\zeta(1 + r + \theta)}{r} \frac{[y_{t+1} - E_t^{su}(y_{t+1})]}{E_t^{su}(c_{t+1})} \tag{18}$$

Our model has the desirable feature of presenting the growth rate of consumption as the result of a precautionary saving motive plus a term that depends on agents' forecast errors. The second term is consistent with the idea of a behavioural model of consumption of irrationally optimistic agents who, having high expectations, experience a bitter surprise, once income is realized, and revise their consumption decision downward. Conversely, irrationally pessimistic agents experience a positive shock as income is realized and revise their consumption decisions upward.

If agents were rational ($E_t^{su}(y_{t+1}) - E_t(y_{t+1}) = 0$), the model would state the common result for a model without liquidity constraints that, apart from the unpredictable income innovation, excess sensitivity is due to precautionary saving, $\Delta \log c_{t+1} = (1/2)\gamma\sigma_c^2 + \frac{\tilde{r} - \tilde{\delta}}{\gamma}$. In this case, ignoring the variance term, may result in omitted variable bias. Hence, overreaction of consumption to predictable changes in income may appear because of their correlation with the error term, which depends on the variance of consumption.

That is not the case in our model, as it asserts that, if agents are irrational or myopic, consumption variation is a function of predictable forecast errors.

Muth's (1961) rational expectations hypothesis implies that expectations are unbiased and forecast errors are distributed independently of the anticipated values. This continues to be true in a model with precautionary saving or liquidity constraints. Despite the fact that consumption is a function of predictable changes in income, constrained or prudent agents, if rational, do not make systematic errors in predicting future income. So, prediction error is a nonrandom term of the Euler equation if agents are irrational or myopic.

On the contrary, if our model is valid, previous evidence of excess sensitivity may be re-interpreted not only because of the omitted variance term and for liquidity constraints, but also because of the assumption of rationality, that is, because of the omission of the predictable forecast error.[2] Thus, consideration of irrationality can help explain the anomaly of predictable changes in consumption.

III. Data

The data are taken from the DNB Household Survey (DHS) that since 1993 collects detailed information on demographics, work, health status, family composition, individual and family incomes and wealth. Moreover, the DHS is one of the few surveys that collects different kinds of subjective expectations on the future family income and inflation and information on agents' attitudes towards risk and their time preferences. Being a savings survey, the DHS panel does not collect data on consumption directly, but an estimate can be obtained by taking the difference between income and savings.

Each wave contains flow and stock information for the previous year. The period we consider in our analysis runs from 1995 to 2002, as some variables of interest were collected only in these years.

The probability distribution of next years family income

The data on expected next year income are collected by a module that is similar to the one adopted in the Survey of Economic Expectations (SEE), and discussed in Dominitz and Manski (1997).

In the DHS, the respondents are first asked to answer two questions about the range in which their family income is expected to fall in the next 12 months; the precise wording, translated into English by CentER, is the following: *What do you expect to be the lowest* (*highest*) *total net income your household may realize in the next 12 months*? After answering these questions, the interview software determines four income thresholds by means of the following algorithm: $threshold_\kappa = Y_{\min} + 0.2\kappa(Y_{\max} - Y_{\min})$ and $\kappa = 1, \ldots, 4$. Then, the respondents are asked to report the percent chance that their net family income will be between $Y_{\min}$ and each threshold. The precise wording of the question is as follows: *What do you think is the probability that the total net income of your household will be less than threshold k in the next 12 months*? *Please fill in a number between 0 to 100*.[3] After division by 100, we obtain four point values, corresponding to the thresholds, for the subjective cumulative distribution function of next year's net family income. The subjective probabilities elicited for each single respondent do not identify the distribution. As underlined by Dominitz and Manski (1997), it is possible, but cumbersome, to analyze the expectations data nonparametrically but it facilitates the analysis if we use them to fit a respondent specific probability distribution.

There is inevitably some arbitrariness in using any specific criterion to fit the expectations data to any specific parametric family of distributions. To show that the reasonableness of our findings is not driven by the chosen distribution, we will make two different assumptions on the subjective distribution of the respondents. Because of the structure of the questionnaire, we decided to use distributions with bounded support: the beta and the piecewise linear. The first is a family of distributions that is flexible enough to take into account heterogeneity in the expectation formation as it covers both symmetric and

[2] This conclusion should not be new to an economist. It was clearly an implication of the original work of Friedman on permanent income (Friedman, 1957), where agents were myopic and time horizon was shorter than the entire life. Oddly, this explanation have been forgotten by the literature that follows Hall (1978).

[3] The percent chance of $y \leq y_{\max}$ is not asked and it is implicitly assumed to be 100.

Table 1. Number of respondents at the questions on realized family income and savings, and response rates

Year	Household	Income	%	Savings	%
1995	4055	3675	0.91	2672	0.66
1996	3384	3091	0.91	2215	0.65
1997	2660	2417	0.91	1661	0.62
1998	1365	1264	0.93	867	0.64
1999	1368	1300	0.95	937	0.68
2000	1934	1349	0.70	1002	0.52
2001	2663	2097	0.79	1624	0.61
2002	2358	1993	0.85	1560	0.66
Total	19 787	17 186	0.87	12 538	0.63

asymmetric distributions. The second assumption relaxes the hypotheses that agents are able to articulate expectations using the full theory of probability, in favour of one where they are at least able to attribute probabilities to income intervals. The beta is estimated by nonlinear least squares.[4]

Measuring consumption

An important feature of the data is the way consumption is estimated since it is not directly observed. Consumption can be defined as the difference between income and savings.

In our empirical analysis, we use respondent's answers on self-reported family savings. In particular, we refer to a pair of questions that are part of the section on psychological concepts which we report below:

> *Did you put any money aside in the past 12 months*?

If the answer is yes, the respondent is also asked the following question about the amount:

> *About how much money has your household put aside in the past 12 months*?

(0) *don't know*
(1) *less than Dfl.* 3000 (€136 134)
(2) 3000–10 000 (€136 134 *and* €453 780)
(3) 10 000–25 000 (€453 780 *and* €1 134 451)
(4) 25 000–40 000 (€1 134 451 *and* €1 815 121)
(5) 40 000–75 000 (€1 815 121 *and* €3 403 352)
(6) 75 000–150 000 (€3 403 352 *and* €6 806 703)
(7) 150 000 *or more* (€6 806 703)

Here, respondents are expected to report the amount of money put aside by choosing one of the seven predetermined classes or the noninformative 'don't know'. Out of this information, we have constructed a variable by taking the midpoints of each class. Since the last interval is right censored, no midpoint can be calculated. To overcome this problem, we assume that the highest bound corresponds to €100 000.

We construct an estimate of family income deriving it from a question where respondents are asked to indicate the interval which corresponds to the income realized over the last twelve months. The precise wording of the question is reported below:

> *Into which of the categories mentioned below did the total net income of your household go in the past 12 months*? *If you really don't know*, *use* '*don't know*'.

(0) *don't know*
(1) *less than Dfl.* 20 000 (€907 560)
(2) 20 000–28 000 (€907 560 *and* €1 270 585)
(3) 28 000–43 000 (€1 270 585 *and* €1 951 255)
(4) 43 000–80 000 (€1 951 255 *and* €3 630 242)
(5) 80 000–150 000 (€3 630 242 *and* €6 806 703)
(6) 150 000 *or more* (€6 806 703)

The estimate of income is constructed similarly to estimated savings assigning the midpoints of the intervals indicated by the respondent. For the respondents that indicate the sixth interval, as mentioned above we assign the value of €100 000 as the highest bound. We subtract subjective expected next year's income from this income estimate to calculate the error in predicting future income.

As shown in Table 1, on average 87% of all respondents answered to the questions on family income. Response rates are smaller for the modules on savings (63%). A negligible number of families does not answer to the question about savings because they spend all their income. We opt to exclude observations reporting no savings from most of our estimations in order to avoid errors from possible misreporting.

The sample used in the empirical analysis below includes only heads of households. Table 2 reports summary statistics for the variables and the observations used in the regression analysis. Forecast error (I) is positive for 333 observations and negative for 33 observations. This means that the majority of observations has higher next year's income realizations than expected. The estimated mean values of the subjective expected income distribution – expected income (I) and (II), in the case, respectively, a beta or a piecewise linear distribution is assumed – show similar means and SEs, suggesting that it is not much sensitive to the choice of the distribution.

As will be clarified below, to estimate the model, we need at least three consecutive waves of data. Since some questions of interest on subjective income were collected only from 1995 to 2002, we only consider eight waves. We do not make use of

[4]Further analyses on the subjective expectation data are available in the Data Appendix.

Table 2. Summary statistics

Variable	Mean	SD	Min.	Max.	*N*
Forecast error (I)	0.45	0.68	−3.02	6.96	426
Forecast error (II)	0.52	0.5	0	1	426
Expected income (I)	24 760.9	14 069.39	547.04	90 232.44	426
Expected income (II)	24 716.07	14 033.63	546.62	89 438	426
Consumption	30 066.69	15 931.25	1134	81 084	426
Savings	4418.51	5256.46	680.5	51 050.5	426
Income (categ.)	4.12	0.89	1	6	426
Income	34 485.2	17 235.56	4538	84 033.5	426
Expected inflation	2.79	2.17	0	15	426
Variance of income	87.25	395.89	0	6956.31	426
Pre-vocational	0.17	0.38	0	1	426
Pre-university	0.1	0.3	0	1	426
Apprentice	0.18	0.38	0	1	426
Vocational college	0.36	0.48	0	1	426
University	0.17	0.38	0	1	426
Gender	0.89	0.32	0	1	426
Jobseeker	0.01	0.12	0	1	425
Number of family members	2.57	1.27	1	7	426
Number of recipients	1.38	0.70	0	5	426
Good health	0.86	0.35	0	1	426
Employee	0.75	0.43	0	1	426
Self-employed	0.01	0.1	0	1	426
Student	0.14	0.35	0	1	426
Retired	0	0.05	0	1	426
Temporary	0.01	0.12	0	1	426
Experience	29.06	12.64	0	56	426
Age	49.3	11.33	25	70	426
Type of employer: public	0.35	0.48	0	1	426
Type of employer: nonpublic	0.5	0.5	0	1	426

Notes: Forecast error (I) is defined as the difference between realizations and expectations divided by actual income at the time t. Forecast error (II) is a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. Expected income (I) is the estimation of next year's income obtained assuming a beta distribution. Expected income (II) is the estimation of next year's income obtained assuming a piecewise linear distribution. Consumption is obtained by subtracting savings to income. Savings is calculated as the midpoints of the reported intervals. Income is calculated as the midpoints of the reported intervals (Income categ.). Variance of income is the estimated variance of expected income divided by current actual income.

imputation in the cases of item non response. Instead, we drop the families for which variables on expected and realized income are not available. Other observations are not considered due to the lack of data on relevant variables, but they are very few and substantially negligible. Finally, heads of households aged more than 70 are excluded because these individuals are mostly retired and are generally on a fixed income, other than inflation increases, and, therefore, can easily predict their future income.

Merging the data from all the DHS questionnaires produces a pooled dataset for all waves which contains 7383 individuals. However, since we use only observations that remain in the panel for at least three consecutive years, the number of available respondents is reduced to 3062. Among them 1120 remain in the panel for only three waves while 75 stay for the entire duration of the panel. Considering all observations, the mean duration is of 2.7 years with the first and third quartiles of the distribution equal to 1 and 4 years, respectively.

To deal with the fact that subjective expectations are characterized by the presence of extreme values and to take into account for heteroscedasticity, we decided to estimate robust regressions, following Flavin (1991, 1999), Browning and Lusardi (1996) and Attanasio (1998). In particular, we have calculated SEs by iteratively reweighted least squares as implemented in the STATA 11's robust estimator *rreg*.

IV. Empirical Implementation of the Model and Testing Procedure

We estimate the model presented in Section II using IVs in order to test the null of rational expectations and isoelastic separable utility. The idea is that nonrational pessimistic/optimistic agents commit systematic errors in forecasting income, which can be predicted by the econometrician. Agents that have been irrationally pessimistic experience a positive surprise when income is realized and revise their consumption decisions up. Conversely, irrationally optimistic agents experience a bitter surprise and revise their consumption decisions downward.

To implement the theoretical statement, we use a two-step procedure. In the first stage, we instrument forecast errors. That is, we run an orthogonality test regressing forecast errors on data that were in the agents' information set at the time the expectations were stated. The IV must be correlated with the endogenous explanatory variables conditional on the other covariates and cannot be correlated with the error term in the explanatory equation. The latter requirement is automatically met under the null for all the data that were part of the information set of the agent when he stated his expectations. If the null of rational expectations is rejected, we are able to predict agents' forecast errors, that is, the systematic surprises that they experience as income realizes. Thus we test our behavioural model of consumption, estimating the modified Euler equation presented in the Equations 15 and 18, as the second step of the procedure.

The first stage

Considering expectations on the growth rate of income, a general first-stage orthogonality test has the following form:

$$fe_{t+1} = X_t\beta_1 + Z_t\beta_2 + \epsilon_{t+1} \tag{19}$$

where the dependent is the forecast error ($fe_{t+1} = \frac{y_{t+1} - E_t^{su}(y_{t+1})}{y_t}$), X_t is a set of excluded variables including income expectation and Z_t is a matrix of controls including the ratio of the variance of expected income and actual income, actual income, income squared, expected inflation, change in the number of family members, dummy on whether any member is looking for a job, change in the number of members who are income recipients and time and regional dummies. Under the null of rational expectations $\beta_1 = 0$ and $\beta_2 = 0$. To estimate this model, we need to observe the same individual at least for three consequent waves. We need information on the number of family members, employment status and income recipients at time $t-1$ and t and information on actual income realization at the time $t+1$. No model that explains the alternative to the null hypothesis is specified.

Forecast errors are defined in two alternative ways in order to check that results are not driven by the choice of a specific estimate: (i) as the difference between the self-reported income realizations, calculated as the midpoints of the reported intervals at the time $t+1$ and the subjective mean of next year's family income level at the time t (Model A); and (ii) as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization (Model B).

For our purposes the main limitation of our panel remains its short time dimension, that is 8 years. The conditional expectation of the disturbance terms $E(\epsilon_{t+1})$, in accordance with permanent income hypothesis with rational expectations, must be zero. The empirical analogue of $E(\epsilon_{t+1})$ is an average calculated on a long time span. In fact, as pointed out by Chamberlain (1984), the increase of the cross-section dimension does not guarantee its convergence to zero. Even though the forecast error should be zero on average if calculated on a long time period, this may not be the case in short panels. Otherwise stated, when performed with short panels, the orthogonality test is a joint test of the orthogonality condition and of the maintained assumption that forecast errors are not correlated across households. Rejection of the null in favour of our behavioural model may be attributed to the inconsistency of the estimator. To account for macroeconomic shocks, we have included controls in both steps of the estimation procedure.[5] In particular, we allow for the presence of time and geographical dummies.

The choice between excluded variables and controls is someway arbitrary and controls cannot be used to test the null. Hence, we allow for different specifications.

As underlined above, we have information on the subjective maximum and minimum expected income and on the subjective cumulative distribution function of next year's net family income, calculated at the threshold. That makes it possible to estimate the entire distribution of income expectations without making assumptions on the shape of the loss function. Hence, the rejection of the null in our orthogonality test is never imputable to false assumptions on the loss function. The only assumption that our analysis requires is on the distribution function whose parameters have to be estimated. To understand whether this choice have an effect on our estimates, we allow for two alternative distribution functions, the beta and the piecewise uniform.

Second stage: the Euler equation

If the hypothesis of rational expectations is rejected, we can test our behavioural model of consumption estimating the following empirical specification of the Euler equation:

$$\Delta \log c_{t+1} = \widehat{fe}_{t+1}\omega_1 + Z_t\omega_2 + \eta_{t+1} \tag{20}$$

where the dependent is the log of consumption change, $\widehat{fe_{t+1}}$ is the predicted forecast error as obtained from Equation 19, Z_t contains the same set of variables as in first step and $+\eta_{t+1}$ is an error term. The matrix Z_t contains the conditional variance term to allow for the fact that if utility exhibits decreasing absolute risk aversion, prudent consumers, to an extent that depends on prudence, reduce consumption with respect to future as reaction to an increase in consumption risk. Ludvingson and Paxson (2001) and Jappelli and Pistaferri (2000) have pointed out that the failure to properly taking into account consumption risk will bias the coefficient of the intertemporal elasticity of substitution and will generate spurious evidence of excess sensitivity. The same reasoning applies to our behavioural model.

The Z_t also includes the expected inflation, $E_t^{su}\pi_{t+1}$. Theoretically, the expected values of the real interest rate should enter the Euler equation as a relevant variable in saving decision. Our dataset does not collect subjective expectations about next year's real interest rate, but it is possible to proxy it by using expected inflation. This approximation is exact if financial market is perfect. In this case, there is only one interest rate and subjective expected real interest rates differ only because of inflation expectations.

The main limitation of our panel continues to be its short-time dimension that makes it susceptible of the Chamberlain's (1984) critique. As underlined by Jappelli and Pistaferri (2000), the excess sensitivity test when performed on a short panel is a

[5] A macroeconomic shock occurring in the observed years and potentially affecting the Dutch household consumption behaviour is the final constitution of the European Monetary Union (EMU) and the consequent adoption of a single currency. Jappelli and Pistaferri (2011) study whether financial integration and liberalization brought about by the introduction of the euro has affected the sensitivity of consumption with respect to income shocks in Italy. The authors do not find a significant effect on consumption smoothing.

joint test of the null and of an assumed structure of the disturbance term, η_{t+1}. Apparent excess sensitivity may arise as the result of not properly taking into account the cross-correlation of disturbances. To control for evenly and unevenly distributed macroeconomic shocks, we have included controls in both steps of the estimation procedure. In particular, we allow for the presence of time dummies and geographical dummies.

Another problem may arise because of the failure of the separability assumption. If consumption and leisure are not separable, today's decision will be affected by predictable changes in households' labour supply. This implies that consumption is correlated with hours of work, which are in turn correlated with income growth. Failure to consider for nonseparability may bring us to spurious evidence of excess sensitivity. Therefore, among the controls at the second step, we have explicitly included variables describing variations in the number of family members, members who are looking for a job and income recipients.

V. Results

In this section, we present the empirical evidence concerning the model presented in Section II. As already underlined, to perform our test we need observations that stay in the panel for, at least, 3 consecutive years.

To deal with the noise contained in the measured income and savings, and hence in measured consumption, and with the extreme values contained in the subjective expectations, we have run the STATA 11's robust estimator (*rreg*) using default parameters.[6] Such an estimator is robust with respect to outliers either in the space of the regressors and in the space of residuals.[7]

For Model A, the null hypothesis of rational expectations is rejected with both OLS and the robust estimator. We use the robust estimates as our linear prediction of the systematic error component to use in the second step. For Model B, we have performed the standard *logit* estimator in first steps.

The assumption of rational expectations implies that our instruments are weakly exogenous as long as we use instruments that were in the agents' information sets. In order to show that our results are not due to a particular set of instruments, we use two alternative specifications. In the first specification, the matrix X_t consists of a large set of variables containing information on the household head and household characteristics. The second specification differs from the first in the fact that income expectation remains the only component of X_t in Equation 19 and income and income squared are dropped from Z_t at both steps. The reason for eliminating all these variables is to avoid over-prediction in the IV estimator. If that were the case, our predicted forecast error may capture events that were genuinely unpredictable, resulting in spurious evidence in favour of our behavioural model. Both the specifications pass the heteroscedasticity-robust tests of overidentifying restrictions (the Anderson–Rubin Wald tests and the Stock–Wright Lagrange Multiplier (LM) *S*-statistic) and the weak instruments tests (Kleibergen–Paaprk Wald LM and *F*-statistics).[8]

Given that we choose a logarithmic transformation for our dependent variable (consumption change), that the second-stage independent variable of interest (predicted forecast error) is expressed in terms of the actual level of income at the time t, and that we use a robust estimator, we are not very concerned about heteroskedasticity and reasonably confident that the reported t- and F-statistics are valid.[9]

The reported p-values in first-step equations suggest the rejection of the hypothesis of rational expectations at any conventional significance level. Results for the estimation of the corresponding second-stage Euler equations show that predictable forecast errors help explain consumption change, which is evidence in favour of our behavioural model.

Let us look at the reported first stages. Table 3 reports those results obtained regressing the both definitions of forecast error on a large set of excluded variables X_t and a set of controls Z_t, using a beta distribution function for inferring the mean of income expectation and the variance of expected income. The set of controls is compounded by the same variables we allow at the second stage. The matrix X_t includes variables on the household structure and variables describing the head of the household. The reported F and Chi-squared tests are based on the set of excluded variables but not on the controls.

There is a significant negative coefficient on expected income with all specifications, which may reflect the fact that people who have been too optimistic are going to experience a bitter surprise in the realization and the converse if they have been too pessimistic. Furthermore, for the Model A, we find significant coefficients on the variance of expected income, income, income squared, education, self-employed, good health, employed on a temporary basis, working in public institution and temporal and regional dummies. For the Model B, we find that only coefficients on the variables change in the number of family members, working in public institution and time dummies are statistically significant.

Considering the specification with fewer instruments, p-values of first-step equations continue to suggest the rejection of the hypothesis of rational expectations (Table 4). The coefficients on income expectation, subjective inflation, subjective variance of expected income, working in public institution, time and regional dummies are significant at 1%. The choice between excluded variables and controls is someway arbitrary, so we have calculated the F-test on different sub-samples of the excluded variables. For instance, we have considered the hypothesis that the stated expectation were the only excluded variables completely immune to the influence of

[6]Moreover, because of the pooling of the data into a single cross section, we have calculated SEs corrected to take into account those observations that belong to the same respondent.
[7]The *rreg* procedure first performs an initial screening based on Cook's distance >1 to eliminate gross outliers prior to calculating starting values and then performs Huber iterations followed by biweight iterations with tuning constant of 7 (Li, 1985). A more detailed description of *rreg* and some Monte Carlo evaluations are provided by Hamilton (1991).
[8]All these tests were performed by means of the STATA 11's routines '*ivreg2*' and '*robivreg*'.
[9]Moreover, a visual analysis of error versus fitted values seemed to give support to this view.

Table 3. Estimation results assuming a beta distribution: more instruments

First stage	(1)	(2)	(3)	(4)
$E_t^{su} y_{t+1}$	−0.0000221*	(−21.50)	−0.0000762*	(−11.59)
$E_t^{su} \pi_{t+1}$	−0.00313	(−0.95)	−0.00981	(−0.45)
Δ*jobseek*	0.0835	(0.96)	0.0793	(0.13)
Δ*n_fam*	−0.0392	(−0.84)	−0.538***	(−1.79)
Δ*recipient*	−0.00812	(−0.47)	−0.0149	(−0.15)
$\frac{\mathrm{Var}_t^{su}(y_{t+1})}{y_t}$	−0.0000687**	(−2.35)	−0.000115	(−0.64)
Instit	0.0296	(1.20)	0.21	(1.43)
Public	0.0705*	(2.58)	0.351**	(2.18)
y_t	0.0000147*	(5.56)	0.000017	(1.31)
y_t^2	−5.74e−11**	(−2.02)	3.38e−11	(0.28)
Primary	−0.379**	(−2.36)	−0.168	(−0.19)
Pre-vocational	−0.137***	(−1.73)	0.0563	(0.12)
Pre-university	−0.0605	(−0.73)	0.00992	(0.02)
Apprentice	−0.115	(−1.45)	−0.0678	(−0.15)
Vocational college	−0.076	(−0.99)	−0.0247	(−0.06)
University	−0.0575	(−0.71)	−0.0991	(−0.22)
Gender	−0.0295	(−0.79)	0.301	(1.44)
Good health	0.0787**	(2.53)	0.204	(1.10)
Poor health	0.0588	(0.34)	1.08	(0.78)
Employee	0.0507	(1.25)	−0.197	(−0.82)
Self-employed	0.141***	(1.90)	0.0712	(0.17)
Student	0.00429	(0.09)	−0.405	(−1.31)
Retired	−0.0387	(−0.28)	−1.07	(−1.25)
Temporary	0.257**	(2.39)	0.0329	(0.06)
Experience	−0.00223	(−0.72)	0.00757	(0.42)
Age	−0.000145	(−0.01)	−0.0234	(−0.41)
Age^2	0.0000252	(0.26)	0.000228	(0.40)
R-squared	0.522			
	F-test: $F(18, 848) = 28.20$	Pr > *F*: 0.000	$\mathrm{Chi}^2(18) = 147.05$	Pr > Chi^2: 0.000
Observations	884		1514	
Second stage	**(1)**	**(2)**	**(3)**	**(4)**
$\widehat{fe}$	0.0489*	(2.63)	0.0742*	(3.93)
y_t	−6.31e−07	(−0.70)	−6.96e−07	(−0.94)
y_t^2	1.28e−11	(1.33)	1.08e−11	(1.64)
$E_t^{su} \pi_{t+1}$	−0.000191	(−0.10)	−0.0000826	(−0.04)
$\mathrm{Var}_t^{su}(y_{t+1})$	5.14e−11	(0.02)	5.17e−11	(0.02)
Δ*n_fam*	0.00496	(0.16)	0.00582	(0.15)
Δ*jobseek*	−0.0797	(−1.20)	−0.0772	(−1.16)
Δ*recipient*	0.0000654	(0.01)	0.000647	(0.09)
Public	0.0121	(1.44)	0.0113	(1.31)
Instit	−0.00989	(−1.04)	−0.0119	(−1.23)
Reps	10 000		10 000	
Observations	559		562	

Notes: Columns (1–2) report results when forecast-error is defined as the difference between realizations and expectations divided by actual income at time *t*; Columns (3–4) report results when forecast error is defined as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. *First stage*: $E_t^{su} y_{t+1}$ is the income expectation calculated assuming a beta distribution function. $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\mathrm{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. *Gender* is an indicator variable that takes value 1 if respondent is male. *Temporary* is an indicator that takes value 1 if employed on a temporary basis. *Experience* is years of work since the first occupation. Δ*n_fam* controls for the variation in family composition. Δ*jobseek* controls for the variation in the number of family members who declare to be looking for a job. Δ*recipient* controls for the variation in the number of income recipients in the family. *Primary*, *Pre-vocational*, *Pre-university*, *Apprentice*, *Vocational college* and *University* are dummies indicating the educational level. *Public* is an indicator for employed by the government. *Instit* is an indicator for employed by another public institution. Controls not allowed to perform *F*-test and Chi-squared test. *Second stage*: The dependent variable is the consumption change. $\widehat{fe}$ is the predicted forecast error. In both steps, year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. Reported second-stage SEs result from 10 000 bootstrap replications of the entire two steps estimation procedure.

*, ** and *** indicate coefficients are significant at the 1, 5 and 10%, respectively.

Table 4. Estimation results assuming a beta distribution: income expectation as the only instrument

First stage	(1)	(2)	(3)	(4)
$E^{su}(y_{t+1})$	−0.0000144*	(−17.76)	−0.0000542*	(−11.91)
$E_t^{su}\pi_{t+1}$	−0.00792**	(−2.27)	−0.0402**	(−2.08)
$\frac{\text{Var}_t^{su}(y_{t+1})}{y_t}$	−0.000125*	(−5.09)	−0.00016	(−1.08)
Δ*jobseek*	0.00995	(0.14)	−0.59	(−1.56)
Δ*n_fam*	−0.0364	(−0.77)	−0.42***	(−1.72)
Δ*recipient*	−0.00846	(−0.51)	0.0962	(1.19)
Instit	0.0564**	(2.41)	0.188	(1.62)
Public	0.0927*	(3.67)	0.293**	(2.35)
1997	−0.176*	(−5.16)	−0.712*	(−3.98)
1998	−0.181*	(−5.29)	−0.249	(−1.33)
1999	0.735*	(17.81)	0.751*	(4.47)
2000	0.267*	(5.01)	2.38*	(8.71)
2001	0.22*	(6.24)	2.33*	(12.42)
West	0.0819*	(2.77)	0.32**	(2.19)
East	0.0751**	(2.28)	0.126	(0.77)
South	0.0837*	(2.68)	0.372**	(2.42)
R-squared	0.423			
	$F(1, 1283) = 315.37$	Pr > F: 0.000	$\text{Chi}^2(1) = 141.86$	Pr > Chi^2: 0.000
Observations	1300		2207	
Second stage	(1)	(2)	(3)	(4)
$\widehat{FE}$	0.0263	(1.60)	0.0455**	(2.32)
$E_t^{su}\pi_{t+1}$	−0.000677	(−0.53)	−0.000523	(−0.41)
$\text{Var}_t^{su}(y_{t+1})$	1.12E−10	(0.13)	1.98E−11	(0.04)
Δ*n_fam*	0.0127	(0.54)	0.0157	(0.69)
Δ*jobseek*	−0.0692	(−1.50)	−0.0658	(−1.44)
Δ*recipient*	−0.00236	(−0.55)	−0.0028	(−0.65)
Public	0.00258	(0.39)	0.00255	(0.38)
Instit	−0.00823	(−1.13)	−0.00893	(−1.21)
1997	0.00589	(0.63)	0.00645	(0.68)
1998	0.00117	(0.12)	−0.00086	(−0.09)
1999	−0.0238	(−1.39)	−0.00885	(−0.58)
2000	0.00174	(0.09)	−0.0131	(−0.66)
2001	0.00669	(0.68)	−0.00877	(−0.67)
West	0.00806	(1.04)	0.00795	(1.03)
East	0.00472	(0.59)	0.00587	(0.74)
South	0.00178	(0.22)	0.00176	(0.23)
Reps	10 000		10 000	
Observations	824		828	

Notes: Columns (1–2) report results when forecast error is defined as the difference between realizations and expectations; columns (3–4) report results when forecast error is defined as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. *First stage*: $E_t^{su}y_{t+1}$ is the income expectation calculated assuming a beta distribution function. $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\text{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. Δ*n_fam* controls for the variation in family composition. Δ*jobseek* controls for the variation in the number of family members who declare to be looking for a job. Δ*recipient* controls for the variation in the number of income recipients in the family. *Primary*, *Pre-vocational*, *Pre-university*, *Apprentice*, *Vocational college* and *University* are dummies indicating the educational level. *Public* is an indicator for employed by the government. *Instit* is an indicator for employed by another public institution. Controls not allowed to perform prediction and the *F*-test. *Second stage*: The dependent variable is the consumption change. $\widehat{fe}$ is the predicted forecast error. In both steps, year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. Reported second-stage SEs result from 10 000 bootstrap replications of the entire two-steps estimation procedure.

* and ** indicate coefficients are significant at the 1 and 5%, respectively.

macroeconomic shocks and all the other variables as controls. In this case, the orthogonality test reduces to a t-test. Results continue to support the rejection of the null.

Second step estimates reported in Tables 3 and 4 support our behavioural model. Predictable errors in forecasting income $\widehat{fe}$ explain consumption variation, confirming that irrational pessimistic/optimistic consumers upward/downward revise their consumption decision as income realizes. The omission of all components of X_t except income expectation in the first step gives smaller estimated coefficients on predictable forecast errors but still significant at the 5% level in one of the two cases. The estimated significant coefficients on predictable forecast errors range between 0.045 and 0.074. Coefficients on forecast errors are higher with the second definition (Model B), but still quite similar to coefficients obtained using the first definition of forecast errors

Table 5. Estimation results assuming a piecewise linear distribution

More instruments Second stage	(1)	(2)	(3)	(4)
$\widehat{fe}$	0.0554*	(3.05)	0.0813*	(4.28)
y_t	−6.09E−07	(−0.63)	−5.72e-07	(−0.77)
y_t^2	1.34E−11	(1.30)	1.03E−11	(1.54)
$E_t^{su}\pi_{t+1}$	−0.000667	(−0.37)	−0.000375	(−0.20)
$\mathrm{Var}_t^{su}(y_{t+1})$	−1.34E−10	(−0.04)	−1.46E−10	(−0.05)
Δn_fam	0.00933	(0.30)	0.0103	(0.19)
$\Delta jobseek$	−0.0801	(−1.22)	−0.0776	(−1.10)
$\Delta recipient$	0.00072	(0.10)	0.00171	(0.24)
Reps	10 000		10 000	
Observations	551		552	
Fewer instruments Second stage	(1)	(2)	(3)	(4)
$\widehat{fe}$	0.0335***	(1.94)	0.0522*	(2.70)
$E_t^{su}\pi_{t+1}$	−0.000743	(−0.51)	−0.000803	(−0.54)
$\mathrm{Var}_t^{su}(y_{t+1})$	5.41E−10	(0.40)	1.28E−10	(0.10)
Δn_fam	0.0135	(0.47)	0.0182	(0.45)
$\Delta jobseek$	−0.0701	(−1.47)	−0.0685	(−1.47)
$\Delta recipient$	−0.000381	(−0.08)	−0.00084	(−0.17)
Reps	10 000		10 000	
Observations	778		781	

Notes: First stages are not reported. Columns (1–2) report second stage results where forecast error, in the first stage, is defined as the difference between realizations and expectations; columns (3–4) report results where forecast error, in first stage, is defined as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. The dependent variable is the consumption change. $\widehat{fe}$ is the predicted forecast error. $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\mathrm{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. Δn_fam controls for the variation in family composition. $\Delta jobseek$ controls for the variation in the number of family members who declare to be looking for a job. $\Delta recipient$ controls for the variation in the number of income recipients in the family. In both steps, year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. SEs result from 10 000 bootstrap replications of the entire two steps estimation procedure. * and *** indicate coefficients are significant at the 1 and 10%, respectively.

(Model A). These estimated parameters suggest that systematic errors explain on average 4.3% of the variation in consumption with our sample.

It is hard to compare this percentage value with other similar results in the literature. To our knowledge, one exception is the paper of Souleles (2004). Looking at his estimated parameters, we conclude that the effect to consumption change has quite similar magnitude compared to our estimates. As the forecast error may be the cause or at least one of the causes for the excess sensitivity of consumption to expected income change (we will focus on it in the folllowing section), we could additionally compare our result with existing estimates of excess sensitivity of consumption to predicted income. As documented in a clear and complete review by Jappelli and Pistaferri (2010), results reported from the existing literature are mixed, ranging from some studies not finding any statistically significant effect to others reporting significant large impacts on consumption decisions. Therefore, our result is also compatible with such findings.

The coefficients on the other variables included at the second stage are never statistically significant for any specification of forecast errors and instruments. Despite being significant predictors of the forecast errors, the variables included to control for nonseparability of consumption and leisure do not seem to be important in the consumption decision. Variations in the number of job seekers, $\Delta jobseek$ and members, Δn_fam, in a household are never significant. Also, precautionary savings and interest rates appear not to be important in determining consumption changes. Despite the lack of statistical evidence, the inclusion of those variables in the second stage was intended to reduce risk of omitted variables bias in the second stage (inclusion of a variable with true coefficient zero reduces the precision of the estimate of the coefficient of interest but does not create bias). Determining whether or not alternative theories of excess sensitivity still represent valid explanations of the issue is out of the scope of the exercise. To show that our results are not driven by the choice of the subjective expectations distribution function, the second step's results referring to the same models described above are reported in Table 5 for the case of a piecewise linear distribution function.

The estimated coefficient of predictable forecast error is positive and statistically significant at the 5% level, with values from 0.052 to 0.081. It is smaller when we consider the specification with only one excluded variable.

Robustness analysis

As shown in Fig. 1, we observe significant shifts to upper classes in the reported income categories between 1999 and 2000, while the distribution of answers is stable along the other years. The magnitude of this change is huge, as the mean

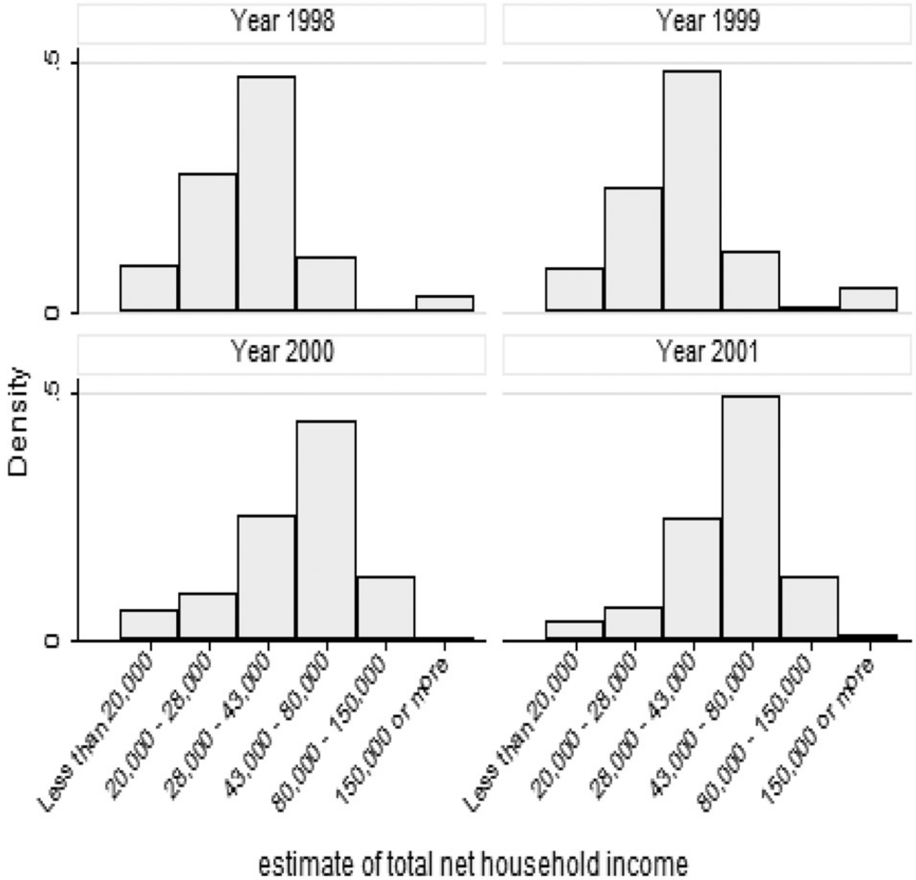

Fig. 1. Income distribution by years

of household's income level jumps from €25 310 in 1999 to €42 193 in 2000 (Fig. 2). As pointed out by Kapteyn *et al.* (2009), a possible explanation for this anomalous change in the distribution of answers on income categories may be that the technology used for the interviewing of respondents was thoroughly modernized in 2000. In order to understand whether and how this anomalous shock influences our findings, we drop all observations of the year 1999, with which the change from the 1999 to 2000 is associated, and replicate all regressions. We perform this for all forecast errors and instruments specifications, and distribution functions. Results, reported in Tables 6 and 7, confirm our previous findings, showing again an estimated coefficient of predictable forecast error as positive and significant at 1% or 5% in six out of eight cases with significant values between 0.05 and 0.082.

Results are still in line with our hypothesis when we eliminate the subjective income variance from both steps of all specifications, with or without observations of 1999 and using both distributions for inferring the subjective mean of income expectations. For the ease of exposition, all these results have not been reported.[10]

As a further robustness check, we separate households with positive versus negative income growth to have some information on possible asymmetry. Table 8 shows Model A's results where second step Equation 20 is modified as follows. We create two dummy variables. The first dummy takes 1 whether income of the household increases (*pos*), while the second takes 1 whether income of the household decreases (*neg*). We consider two specifications. Column (1) reports

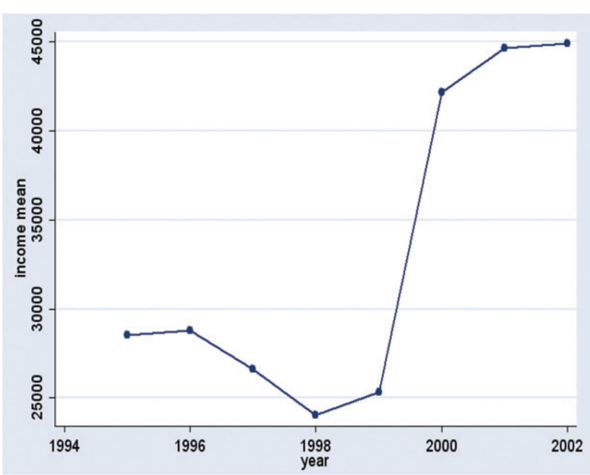

Fig. 2. Sample income mean by years

[10]These results are available from the authors upon request.

Table 6. Estimation results assuming a beta distribution: year 1999 excluded

More instruments Second stage	(1)	(2)	(3)	(4)
$\widehat{fe}$	0.0584*	(2.68)	0.0828*	(4.10)
y_t	−4.69E−07	(−0.47)	−2.01E−07	(−0.25)
y_t^2	1.17E−11	(1.1)	6.25E−12	(0.89)
$E_t^{su}\pi_{t+1}$	−0.000546	(−0.28)	−0.000335	(−0.17)
$\mathrm{Var}_t^{su}(y_{t+1})$	9.01E−11	(0.04)	−1.24E−11	(−0.01)
Δn_fam	0.0111	(0.23)	0.00901	(0.13)
$\Delta jobseek$	−0.0816	(−1.21)	−0.0783	(−1.04)
$\Delta recipient$	−0.0000365	(−0.01)	0.000017	(0.00)
Reps	10 000		10 000	
Observations	481		479	
Fewer instruments Second stage	(1)	(2)	(3)	(4)
$\widehat{fe}$	0.0311	(1.62)	0.0493**	(2.41)
$E_t^{su}\pi_{t+1}$	−0.00104	(−0.65)	−0.00106	(−0.66)
$\mathrm{Var}_t^{su}(y_{t+1})$	1.07E−10	(0.12)	7.54E−12	(0.01)
Δn_fam	0.014	(0.44)	0.0186	(0.41)
$\Delta jobseek$	−0.0658	(−1.42)	−0.0657	(−1.40)
$\Delta recipient$	−0.0013	(−0.25)	−0.00229	(−0.44)
Reps	10 000		10 000	
Observations	676		678	

Notes: First stages are not reported. Columns (1–2) report second stage results where forecast error, in the first stage, is defined as the difference between realizations and expectations; columns (3–4) report results where forecast error, in first stage, is defined as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. The dependent variable is the consumption change. $\widehat{fe}$ is the predicted forecast error. $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\mathrm{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. Δn_fam controls for the variation in family composition. $\Delta jobseek$ controls for the variation in the number of family members who declare to be looking for a job. $\Delta recipient$ controls for the variation in the number of income recipients in the family. In both steps – other than year 1999 – year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. SEs result from 10 000 bootstrap replications of the entire two-steps estimation procedure.
* and ** indicate coefficients are significant at the 1 and 5%, respectively.

results from a specification where, other than $\widehat{fe}$, also the dummy (*pos*) and the interaction term between $\widehat{fe}$ and the dummy (*pos*) are included. Column (2) reports similar results using the dummy (*neg*). While the parameter on $\widehat{fe}$ remain statistically significant in five out of six regressions, the interaction terms are never statistically significant, suggesting that there is not a significant association between, on one side, predictable forecast error and, on the other side, magnitude and sign of income shocks.

Finally, we check whether our results are driven from not properly taking into account the nonseparability of consumption and leisure. To do this, we restrict estimations on those individuals with stable (i) number of family members, (ii) family employment status and (iii) number of family members recipient of income. Results, which are reported in columns (3–5) of Table 8, show that the parameter on $\widehat{fe}$ is still statistically significant, suggesting that previous estimates were not biased from the above-mentioned potential source of spurious correlation.

VI. Irrationality and Excess Sensitivity

In this section, we investigate the relative importance of irrationality, liquidity constraints and precautionary saving in explaining excess sensitivity.

Theoretically, the rejection of the hypothesis that consumption is a random walk can be attributed to the presence of liquidity constraints, precautionary savings and irrationality or myopia. Oddly, in the extensive literature on testing the permanent income hypothesis, the possibility that rejection is due to predictable forecast errors is rarely mentioned, let alone explored. From Hall's (1978) article onwards, all the effort in testing the Euler equation and excess sensitivity of consumption to predictable income changes have concentrated on liquidity constraints[11] and precautionary saving, although, as pointed out by Carroll (1992), it is very hard to distinguish empirically between precautionary saving and liquidity constraints as households may increase saving today if they expect to be liquidity constrained in the future.[12]

[11]For instance, see Hall and Mishkin (1982), Runkle (1991), Garcia *et al.* (1997) and Jappelli *et al.* (1998). More recently, Johnson and Li (2010) distinguish between a household with low liquid assets (liquidity-constrained household) and a household without ready access to credit (borrowing-constrained household) and find that only the consumption growth of households that are both liquidity- and borrowing-constrained is excessively sensitive to lagged income.
[12]See also Jappelli and Pistaferri (2010) for an extended and updated review on empirical approaches and evidence on the sensitivity of consumption to predicted income changes, including works combining realizations and expectations of income or consumption in surveys, in which data on subjective expectations are available.

Table 7. Estimation results assuming a piecewise linear distribution: year 1999 excluded

More instruments Second stage	(1)	(2)	(3)	(4)
$\widehat{fe}$	0.0576*	(2.99)	0.0794*	(4.22)
y_t	−5.01E−07	(−0.51)	−2.70E−07	(−0.35)
y_t^2	1.20E−11	(1.15)	6.91E−12	−0.99
$E_t^{su}\pi_{t+1}$	−0.000873	(−0.47)	−0.000676	(−0.36)
$\mathrm{Var}_t^{su}(y_{t+1})$	9.92E−11	(0.04)	−1.31E−10	(−0.05)
Δn_fam	0.0129	(0.35)	0.0124	(0.19)
$\Delta jobseek$	−0.0783	(−1.19)	−0.0782	(−1.06)
$\Delta recipient$	−0.000658	(−0.09)	−0.000776	(−0.11)
Reps	10 000		10 000	
Observations	501		498	
Fewer instruments **Second stage**	**(1)**	**(2)**	**(3)**	**(4)**
$\widehat{fe}$	0.0311***	(1.72)	0.0518*	(2.67)
$E_t^{su}\pi_{t+1}$	−0.00133	(−0.83)	−0.00138	(−0.86)
$\mathrm{Var}_t^{su}(y_{t+1})$	1.14E−10	(0.12)	2.28E−11	(0.03)
Δn_fam	0.0161	(0.51)	0.0201	(0.46)
$\Delta jobseek$	−0.066	(−1.41)	−0.0649	(−1.37)
$\Delta recipient$	−0.00148	(−0.29)	−0.00264	(−0.52)
Reps	10 000		10 000	
Observations	704		706	

Notes: First stages are not reported. Columns (1–2) report second stage results where forecast error, in the first stage, is defined as the difference between realizations and expectations; columns (3–4) report results where forecast error, in first stage, is defined as a binary variable taking 1 whether income expectation is less than the minimum of the income category containing the self-reported actual realization. The dependent variable is the consumption change. $\widehat{fe}$ is the predicted forecast error. $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\mathrm{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. Δn_fam controls for the variation in family composition. $\Delta jobseek$ controls for the variation in the number of family members who declare to be looking for a job. $\Delta recipient$ controls for the variation in the number of income recipients in the family. In both steps – other than year 1999 – year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. SEs result from 10 000 bootstrap replications of the entire two steps estimation procedure.
* and *** indicate coefficients are significant at the 1 and 10%, respectively.

Here, we are not interested in discerning between the two classical sources of excess sensitivity. We aim at demonstrating the importance of irrationality as an alternative source of excess sensitivity. We estimate the following Euler equation, modified to allow for irrationality.

$$\begin{aligned}\Delta\log c_{t+1} &= \phi_1\Delta D_{t+1} + \rho^{-1}(E(r_{t+1}|\Omega_t) - \delta) \\ &\quad + \frac{\rho}{2}\mathrm{var}_t(\Delta\log c_{t+1} - \rho^{-1}(r_{t+1})) \\ &\quad + \phi_2 E\Delta\log(y_{t+1}|\Omega_t) + \phi_3 E[y_{t+1} - E_t^{su}(y_{t+1})] + \varepsilon_{t+1}\end{aligned} \quad (21)$$

where i is an household index, $c_{i,t+1}$ is our estimate of consumption, $D_{i,t+1}$ is a matrix that includes our controls for households' preferences, nonseparability between consumption and leisure, and macroeconomic shocks, $r_{i,t+1}$ is the real after tax rate of interest, δ is the rate of time preferences and ρ^{-1} is the inter-temporal elasticity of substitution. Predicted income growth, $E\Delta\log(y_{i,t+1}|\Omega_t)$, and predicted forecast error, $E[y_{t+1} - E_t^{su}(y_{t+1})|\Omega_t)]$, are added to the Euler equation in order to test the orthogonality condition, i.e. $\phi_2 = 0$ and $\phi_3 = 0$. We choose a log specification for income growth and instrument it with the same set of variables that we use to instrument the forecast error.

Table 9 shows the estimated coefficients of predictable forecast errors, predictable changes in income, subjective variance and expected rate of inflation. We consider the Model A, where forecast errors are given by comparing income expectations and realizations of families, with the specification of a larger set of excluded variables defined in the previous section. Expectations and subjective variances have been calculated using the beta distribution.

The first column shows that when the excess sensitivity test is performed the coefficients on the predictable forecast error remain large and significant. This suggests that irrationality is still a possible explanation for excess sensitivity of consumption, even when other explanations are considered. The second column presents results for the equation without considering predictable changes in income. The estimated coefficient for the forecast error is significant and similar to the one reported in column 1. This is an evidence of the fact that irrationality is an explanation that stands on its own. Hence, the coefficient on predictable seems not to be biased much if precautionary savings and liquidity constraints are not properly taken into account. The third column shows the results of the excess sensitivity test under the rational expectations hypothesis. A higher and statistically significant coefficient of the predictable changes in income could be interpreted as an evidence of the fact that not taking into account irrationality may bias upwards the coefficient of the predictable changes in income. However, this parameter remains statistically insignificant. Hence, with this dataset, not taking into account irrationality

Table 8. Robustness analysis

More instruments	(1)	(2)	(3)	(4)	(5)
Dummy income gr. (pos)	−0.0037004 (−0.28)				
Dummy income gr. (neg)		−0.047166 (−1.63)			
$\widehat{fe}$	0.0584* (3.61)	0.057* (3.73)	0.0619* (3.37)	0.0600* (3.68)	0.0513** (2.18)
Interaction with (pos)	−0.020 (−0.85)				
Interaction with (neg)		0.065623 (1.09)			
y_t	−6.91E−07 (−0.83)	−6.50E−07 (−0.76)	−1.12E−06 (−0.97)	−7.13E−07 (−0.68)	8.00E−07 (0.53)
y_t^2	1.36E−11 (1.50)	1.44E−11 (1.52)	1.95E−11 (1.43)	1.45E−11 (1.17)	1.58E−12 (0.09)
$E_t^{su}\pi_{t+1}$	−0.0002475 (−0.15)	−0.0001159 (−0.06)	−0.0004079 (−0.17)	−0.0000659 (−0.03)	−0.0018759 (−0.55)
$\text{Var}_t^{su}(y_{t+1})$	−9.68E−11 (−0.22)	−1.30E−10 (−0.28)	−3.42E−10 (−0.52)	−1.82E−11 (−0.03)	−1.06E−09 (−0.86)
Δn_fam	0.0064955 (0.44)	0.015933 (0.99)		0.0024885 (0.14)	0.023337 (0.88)
$\Delta jobseek$	−0.089174* (−3.02)	−0.075389** (−2.37)	−0.080443** (−2.11)		−0.11548** (−2.19)
$\Delta recipient$	0.0002469 (0.04)	0.0016262 (0.27)	0.0044674 (0.57)	0.0069289 (0.96)	
Observations	529	529	364	390	198
Fewer instruments	(1)	(2)	(3)	(4)	(5)
Dummy income gr. (pos)	−0.023443** (−2.04)				
Dummy income gr. (neg)		0.018188 (1.19)			
$\widehat{fe}$	0.027447*** (1.89)	0.02713** (2.06)	0.042566** (2.46)	0.040276** (2.46)	0.021561 (0.88)
Interaction with (pos)	0.0056109 (0.27)				
Interaction with (neg)		−0.029746 (−0.80)			
$E_t^{su}\pi_{t+1}$	−0.0008435 (−0.60)	−0.0007178 (−0.50)	−0.0010662 (−0.53)	−0.0008725 (−0.46)	−0.0009022 (−0.31)
$\text{Var}_t^{su}(y_{t+1})$	8.89E−11 (0.46)	−2.34E−11 (−0.11)	6.47E−10 (1.28)	6.92E−10 (1.41)	−1.02E−09 (−0.85)
Δn_fam	0.0102 (0.82)	0.013829 (1.11)		0.0089384 (0.59)	0.029795 (1.35)
$\Delta jobseek$	−0.071589* (−3.21)	−0.071454* (−3.19)	−0.071279* (−2.69)		−0.073071*** (−1.85)
$\Delta recipient$	−0.0021805 (−0.49)	−0.0005353 (−0.12)	0.0012441 (0.21)	0.0033368 (0.57)	
Observations	748	748	521	556	295

Notes: First stages are not reported. Forecast error, in the first stage, is defined as the difference between realizations and expectations. The dependent variable is the consumption change. *Dummy income gr.* (*pos*) is a dummy that takes 1 whether individuals reported their income belonging to a higher income category at the time $t+1$ than they reported at the time t and 0 otherwise. *Dummy income gr.* (*neg*) is a dummy that takes 1 whether individuals reported their actual income belonging to a lower income category at the time $t+1$ than they reported at the time t and 0 otherwise. $\widehat{fe}$ is the predicted forecast error. *Interaction with* (*pos*) is the interaction between $\widehat{fe}$ and (*pos*). *Interaction with* (*neg*) is the term of interaction between $\widehat{fe}$ and (*neg*). $E_t^{su}(\pi_{t+1})$ is the inflation expectation (point expectation). $\frac{\text{Var}_t^{su}(y_{t+1})}{y_t}$ is the ratio of variance of expected income and actual income. Δn_fam controls for the variation in family composition. $\Delta jobseek$ controls for the variation in the number of family members who declare to be looking for a job. $\Delta recipient$ controls for the variation in the number of income recipients in the family. In both steps, year 1996, no education and northern region are the dummies excluded to avoid perfect collinearity. SEs are those obtained from using the STATA 11's *rreg* estimator.
*, ** and *** indicate coefficients are significant at the 1, 5 and 10%, respectively.

Table 9. Irrationality and excess sensitivity: expectations calculated assuming a beta distribution function – more instruments

	Without families saving zero			With families saving zero		
	(1)	(2)	(3)	(4)	(5)	(6)
$\widehat{fe_t}$	0.060048	0.055689		0.029034	0.032419	
	(2.76)	(2.71)		(1.97)	(2.25)	
$\ln(\widehat{\frac{y_{t+1}}{y_t}})$	0.026989		−0.030568	−0.01601		−0.042625
	(0.51)		(−0.63)	(−0.53)		(−1.40)
$E_t^{su}\pi_{t+1}$	−0.0001495	−0.0003649	−0.0003306	−0.0006197	−0.0005003	−0.0008722
	(−0.08)	(−0.20)	(−0.17)	(−0.84)	(−0.71)	(−1.18)
$\mathrm{Var}_t^{su}(y_{t+1})$	−1.36E–10	−1.12E–10	−9.13E–10	4.01E–10	3.83E–10	3.33E–10
	(−0.04)	(−0.03)	(−0.28)	(0.33)	(0.33)	(0.3)

Notes: Forecast errors are defined as the difference between the self reported income realizations, calculated as the midpoints of the reported intervals, at the time $t+1$ and the subjective mean of next year's family income level at the time t calculated assuming a beta distribution function. Forecast errors and predictable income growth instrumented with the larger set of instruments. Models (1)–(3): families that reported not to have put money aside last 12 months are dropped. Models (4)–(6): families that reported not to have put money aside last 12 months are included. t-statistics from bootstrapped SEs with 1000 replications are shown in parentheses. Regression in column (1) is performed over 529 observations, while regression reported in column (4) uses 719 observations.

does not give biased evidence in favour of liquidity constraints. All other variables including the subjective variance of income show insignificant parameters. A similar investigation using data where there exists statistically significant evidence of excess sensitivity of consumption to predicted income growth would help to better understand the relative importance of irrationality/myopia explanation on the one side and liquidity constraints and precautionary saving on the other side.

One final remark on sample composition should be done. Because of the way we have built up consumption, starting from those who declared to have put money aside in the last 12 months, we could have induced some form of selection in the sample. In particular, as consumption has been calculated only for those with positive savings, the sample could have been selected against liquidity constrained families. Hence reported evidence from Table 9 could be biased in favour of our model. In particular, estimated coefficients of the predictable changes in income and of the subjective variance, among the others, could be biased and not statistically significant.

To avoid the selection problem we have decided to include in the sample also the respondents who declared that they have not been able to put money aside during the last 12 months. For those respondents saving has been considered equal to 0. Results remain in line with those obtained previously (Table 9).[13] Estimated coefficients on predictable forecast error are a little smaller but significant in all specifications. Moreover, and more importantly for the sample selection issue, also the coefficients on predictable income growth continue not to be significant. Estimated coefficients for the subjective variance term are still nonsignificant, confirming that our results are not induced by sample selection.

VII. Conclusions

We have presented evidence that suggest anomalies in consumption, here the fact that consumption reacts to predictable changes in income, can be explained by a behavioural model in which agents do not have rational expectations and make predictable errors in forecasting income. We have tested and rejected the null of rational expectations.

This adds to the literature on testing rational expectations with self-reported expectations, because we have demonstrated a connection between predictable forecast errors and actual economic behaviour. It is often argued that earlier contributions do not supply evidence to support the claim that the elicited expectations really correspond to those affecting the agent's behaviour. Our result that it is possible to partially explain agents consumption decisions using predictable forecast errors should therefore be of interest.

Moreover, we find that irrationality is an important and autonomous source of the excess sensitivity of consumption, even when precautionary savings and liquidity constraints are considered.

Acknowledgements

The authors thank Luigi Guiso and Tullio Jappelli for useful comments, as well as seminar participants at the University of Urbino. We are responsible for any errors in the manuscript. The views expressed in this article are those of the authors and do not necessarily reflect those of the Ministry of Economy and Finance. The usual disclaimer applies.

References

Attanasio, O. (1998) A cohort analysis of saving behaviour by US households, *Journal of Human Resources*, **33**, 575–609.

Bovi, M. (2009) Economic versus psychological forecasting: evidence from consumer confidence surveys, *Journal of Economic Psychology*, **30**, 563–74.

Brown, S. and Taylor, K. (2006) Financial expectations, consumption and saving: a microeconomic analysis, *Fiscal Studies*, **27**, 313–38.

[13]We have also replicated estimations reported in the previous section including those observations declaring not to be able to put money aside and have found similar results.

Browning, M. and Lusardi, A. (1996) Household saving: micro theories and micro facts, *Journal of Economic Literature*, **34**, 1797–855.

Carroll, C. D. (1992) The Buffer–Stock theory of saving: some macroeconomic evidence, *Brookings Papers on Economic Activity*, **2**, 61–156.

Chamberlain, G. (1984) Panel data, in *Handbook of Econometrics*, Vol. II (Eds) Z. Griliches and M. D. Intrilligator, North-Holland, Amsterdam, pp. 1247–318.

Dominitz, J. (1998) Earnings expectations, revisions, and realizations, *Review of Economics and Statistics*, **80**, 374–88.

Dominitz, J. (2001) Estimation of income expectations models using expectations and realization data, *Journal of Econometrics*, **102**, 165–95.

Dominitz, J. and Manski, C. (1997) Using expectations data to study subjective income expectations, *Journal of the American Statistical Association*, **92**, 855–67.

Flavin, M. (1991) The joint consumption/asset demand decision: a case study in robust estimation, NBER Working Paper No. 3802.

Flavin, M. (1999) Robust estimation of the joint consumption/asset demand decision, NBER Working Paper No. 7011.

Friedman, M. (1957) *A Theory of the Consumption Function*, Princetown University Press, Princetown.

Garcia, R., Lusardi, A. and Ng, S. (1997) Excess sensitivity and asymmetries in consumption: an empirical investigation, *Journal of Money, Credit and Banking*, **29**, 154–76.

Hall, R. (1978) Stochastic implications of the life cycle-permanent income hypothesis: theory and evidence, *Journal of Political Economy*, **86**, 971–87.

Hall, R. and Mishkin, F. (1982) The sensitivity of consumption to transitory income: estimates from panel data on households, *Econometrica*, **50**, 461–77.

Hamilton, L. C. (1991) How robust is robust regression?, *Stata Technical Bulletin*, **2**, 21–6.

Jappelli, T., Pischke, J. and Souleles, N. S. (1998) Testing for liquidity constraints in Euler equations with complementary data sources, *Review of Economics and Statistics*, **80**, 251–62.

Jappelli, T. and Pistaferri, L. (2000) Using subjective income expectations to test for excess sensitivity of consumption to predicted income growth, *European Economic Review*, **44**, 337–58.

Jappelli, T. and Pistaferri, L. (2010) The consumption response to income changes, *Annual Review of Economics*, **2**, 479–506.

Jappelli, T. and Pistaferri, L. (2011) Financial integration and consumption smoothing, *The Economic Journal*, **121**, 678–706.

Johnson, K. W. and Li, G. (2010) The debt-payment-to-income ratio as an indicator of borrowing constraints: evidence from two household surveys, *Journal of Money, Credit and Banking*, **42**, 1373–90.

Kapteyn, A., Kleinjans, K. and Van Soest, A. (2009) Intertemporal consumption with directly measured welfare functions and subjective expectations, *Journal of Economic behaviour and Organization*, **72**, 425–37.

Kaufmann, K. and Pistaferri, L. (2009) Disentangling insurance and information in intertemporal consumption choices, *American Economic Review: Paper and Proceedings*, **99**, 387–92.

Li, G. (1985) Robust regression, in *Exploring Data Tables, Trends, and Shapes* (Eds) D. C. Hoaglin, F. Mosteller and J. W. Tukey, John Wiley and Sons, New York, NY, pp. 281–340.

Ludvigson, M. and Paxson, C. (2001) Approximation bias in linearized Euler Equations, *Review of Economics and Statistics*, **83**, 242–56.

Muth, F. (1961) Rational expectations and the theory of price movements, *Econometrica*, **29**, 315–35.

Runkle, E. D. (1991) Liquidity constraints and the permanent-income hypothesis: evidence from panel data, *Journal of Monetary Economics*, **27**, 73–98.

Souleles, N. S. (2004) Expectations, heterogeneous forecast errors, and consumption: micro evidence from the Michigan consumer sentiment surveys, *Journal of Money, Credit and Banking*, **36**, 39–72.

Data Appendix

The data used in this article are taken from the DNB Household Survey (DHS) that since 1993 has been part of a project started and administered by CentER, a research institute at the University of Tilburg. Since 2003, the project is managed in collabouration with De Nederlandsche bank. The DHS is an unbalanced panel. As reported in Table A1, when the survey started, it consisted of two panels: one representative of the Dutch population (representative, RE), covering 1760 households, and the other representative of the top 10% of the income distribution (high income, HI), encompassing approximately 900 families, with a share of 66% and 34%, respectively. The latest wave of the panel (collected in 2011) consists of 1761 households in the RE panel and only 20 in the HI panel. The severe reduction in the HI panel is due to the fact that since 1997 new families have not been recruited for the HI panel, so it quickly shrank as the higher income families exited the panel.

The DHS consists of six questionnaires, presented to all the people aged 16 or over within the family. The survey method is completely computerized. Each household is provided with a personal computer, receives the questionnaires by modem, answers the questionnaires on its home computer and returns the answers to the CentER by modem again. This means that the questionnaires are self-administered and the respondents can answer the questionnaires at a time that is convenient for them.

Our analysis is based on data from most of the questionnaires of the DHS panel. In particular, it draws heavily upon the part on health and income, where subjective expectations on next year's income were collected, and upon the part on psychological concepts where subjective inflation forecasts and self-reported previous years realized income and savings were collected. The questionnaire on health and income was presented to a decreasing number of respondents during the period that goes from 1995 to 2000 and to around 2000 individuals in the subsequent years. As shown in Table A2, 72% of the respondents stated at least $Y_{\min}$ and $Y_{\max}$. It should be underlined that they were not asked the subsequent questions if the difference between $Y_{\max}$ and $Y_{\min}$ was smaller than a fixed amount which corresponds to 5 Dutch florins (dfl.) until 2002 and 5 euros for the following years. This is the case for 2277 observations (10%).

The DHS suffers a problem of nonmonotonicity in the stated subjective cumulative distribution function. The cases which present this problem are 2251 (10%). A brief analysis of the answers reveals that some people are not able to articulate their expectations using the theory of probability and/or commit typing and recording errors. The final response rate is around half (48%) of the respondents. It is small for the first 2 years (35%), but increases over time to 63%.

The analysis of the lowest and highest possible incomes reveals that 64 respondents have declared a highest possible income inferior to 100 euros and 14 far superior to 500 000 euros. These values seem implausible to us and we decide to drop the corresponding observations. The mean value of the lowest possible income is €18 587 with stated values that vary from 0 to 385 900, while the mean value of the highest possible income is €23 176 in a range that goes from 100 to 500 000.

Table A1. Number of households by panel type and year

Year	Representative (RE)	%	High income (HI)	%	Total
1993	1760	66.2	899	33.8	2659
1994	2174	71.8	852	28.2	3026
1995	2083	74.9	697	25.1	2780
1996	2005	79.0	533	21.0	2538
1997	1920	85.0	338	15.0	2258
1998	1686	95.0	88	5.0	1774
1999	1504	95.7	67	4.3	1571
2000	1738	97.5	45	2.5	1783
2001	2094	97.9	44	2.1	2138
2002	1954	98.2	36	1.8	1990
2003	1914	98.5	29	1.5	1943
2004	1842	98.5	29	1.5	1871
2005	1965	98.6	28	1.4	1993
2006	1904	98.7	26	1.3	1930
2007	1749	98.5	26	1.5	1775
2008	1632	98.5	25	1.5	1657
2009	1680	98.7	22	1.3	1702
2010	1864	98.9	21	1.1	1885
2011	1761	98.9	20	1.1	1781

Notes: Column RE reports summary statistics for the panel representative of the Dutch population. Column HI reports summary statistics for the panel representative of the top 10% of the income distribution of the Dutch population.

Table A2. Response rates on the income expectations variables

Year	1995	1996	1997	1998	1999	2000	2001	2002	Pooled
Household	4854	4250	3447	2392	2250	1055	2075	2139	22 462
$Y_{\max}$, $Y_{\min}$	2335	2035	2847	1966	1863	1037	2043	2095	16 221
%	0.48	0.48	0.83	0.82	0.83	0.98	0.98	0.98	0.72
$Y_{\max} - Y_{\min} < 5$	323	293	339	239	245	135	338	365	2277
%	0.07	0.07	0.10	0.10	0.11	0.13	0.16	0.17	0.10
Probab.	2010	1741	2195	1483	1372	899	1709	1732	13 141
%	0.41	0.41	0.64	0.62	0.61	0.85	0.82	0.81	0.59
No monoton.	307	295	311	212	202	184	352	388	2251
%	0.06	0.07	0.09	0.09	0.09	0.17	0.17	0.18	0.10
Final	1703	1446	1884	1271	1170	715	1357	1344	10 890
%	0.35	0.34	0.55	0.53	0.52	0.68	0.65	0.63	0.48

Notes: Number of respondents at the questions on lowest and highest possible income, cumulative subjective probability distribution and response rates.