Economics Division
University of Southampton
Southampton SO17 1BJ, UK

Discussion Papers in
Economics and Econometrics

Title: Subjective Return Expectations, Information and Stock Market Participation: Evidence from France

By : Luc Arrondely (CNRS-PSE-Banque de France), Hector Calvo-Pardoz (University of Southampton), Derya Tasx (University of Southampton),

No.

This paper is available on our website
http://www.southampton.ac.uk/socsci/economics/research/papers

ISSN 0966-4246

Subjective Return Expectations, Information and Stock Market Participation: Evidence from France*

Luc Arrondel† CNRS-PSE-Banque de France

Hector Calvo-Pardo‡ U. of Southampton

Derya Tas§ U. of Southampton

March 14, 2014

Abstract

Recent research has separately uncovered that stock ownership strongly correlates with both expectations and realizations of stock market returns, as well as with measures of financial literacy, ability or trust. This paper reconciles all, and reports new findings from a unique survey containing individual level data on both expectations and (knowledge of) realizations for a representative sample by age and wealth. Stock market participation monotonically increases with the conditional expectation of a positive stock market return, even among the affluent and the young. Information is very heterogeneous, increases with age and own past experience, and identifies a causal effect of expectations on stock ownership. *JEL* Codes: D12, D83, D84, G11.

*The authors thank Toke Aidt, Charles Bellemare, Alberto Bisin, Antonio Cabrales, John Campbell, Chris Carroll, Dimitris Christelis, Adeline Delavande, Dimitris Georgarakos, Michael Haliassos, Mike Hurd, Gabor Kézdi, John Knowles, Christos Koulovatianos, Hamish Low, Chuck Manski, Andre Masson, Alessandro Mennuni, Perre-Carl Michaud, Xisco Oliver, Nicola Pavoni, Hashem Pesaran, Victor Rios-Rull, Jian Tong, Giorgio Topa and Bob Willis for helpful suggestions, and seminar audiences at the NBER 2012 "Aggregate Implications of Microeconomic Consumption Behaviour" Summer Institute working group, College de France 1st meeting of the International Network on Expectational Coordination (INEXC), CIRPEE workshop on "Subjective expectations formation and revision", 2nd. Dauphine workshop on "Recent Advances on the Role of Beliefs in Decision Theory", 2012 ZEW workshop on "Expectations in Financial Markets", Nottingham University, PSE, Queen Mary 45th MMF conference, RES 2012 and at the University of Cambridge. The authors gratefully acknowledge financial support from the "Agence Nationale pour la Recherche" (ANR), Caisse des Depots et Consignations, Chaire Groupama "Les Particuliers face au Risque", ESRC (RES-061-25-0327) and the University of Cambridge for its hospitality. *The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. The authors have obtained IRB approval from their institutions.*

†Batiment A, Ecole Normale Superieure, 48 Bd Jourdan, 75014 Paris, France, arrondel@pse.ens.fr

‡Division of Economics, FSHS, University of Southampton, Southampton SO171BJ, UK, calvo@soton.ac.uk

§Division of Economics, FSHS, University of Southampton, Southampton SO171BJ, UK, tashderya@gmail.com

1 Introduction

Prior to the 2008 financial crisis, nonparticipation in risky asset markets was awarded the status of 'financial mistake', with potentially large consequences for equilibrium asset prices and the equity premium.[1] The 'nonparticipation' puzzle uncovers the fact that a significant fraction of households hold no risky assets despite of their historical excess average returns over riskless assets, or 'equity premium' (Haliassos and Bertaut, 1995), against elementary theory predictions from standard expected utility maximization models (Arrow, 1965; Merton, 1969; Samuelson, 1969). To date, and since the educated and the wealthier are more likely to participate, information and transaction costs remain the most important quantitatively (Vissing-Jorgensen, 2002; Haliassos and Michaelides, 2003). But several questions remain, like (i) the substantial heterogeneity in portfolio allocations (Curcuru, Heaton, Lucas and Moore, 2010), (ii) the nonparticipation of the wealthiest (Heaton and Lucas, 2000), or (iii) the precise nature of information costs.

Recently uncovered factors such as cognitive ability (Christelis, Jappelli and Padula, 2010; Grinblatt, Keloharju and Linnainmaa, 2011), trust (Guiso, Sapienza and Zingales, 2008), financial literacy (van Rooij, Lusardi and Alessie, 2011; Lusardi, Michaud and Mitchell, 2012), 'financial awareness' (Guiso and Jappelli, 2005), (time spent in acquiring) financial information (Guiso and Jappelli, 2007) or social interactions (Hong, Kubik and Stein, 2004) are shaping our understanding of the nature and importance of financial information costs. However the precise mechanism whereby differences in (ability to process, access to or in the actual stock of) financial information translate into differences in stock market participation remains elusive (Grinblatt, Keloharju and Linnainmaa, 2011).[2] This paper contributes to all the aforementioned open questions, by showing that differences in information[3] affect participation behaviour through their impact on stock market return expectations, and thereby reconciles investor behaviour with traditional models in financial economics.[4]

[1] The more so, the larger the fraction of the wealthiest risk-tolerant households that does not participate, since the aggregate demand for risky assets disproportionately depends on them. And even amongst the wealthiest, nonparticipation remains strikingly high. See Campbell (2006), and Guiso and Sodini (2012) for a comprehensive review of the literature.

[2] Recent theoretical contributions predict that, faced with the same set of arbitrage free asset prices, those have the most to gain from acquiring information are those who would benefit the most from participating and should therefore be more informed (Cabrales, Gossner and Serrano, 2013). Under heterogeneous beliefs however, non-participants are those who face the highest cost per unit of *expected posterior* return, and should therefore be the least well informed (Van Nieuwerburgh and Veldkamp, 2010).

[3] Our novel information measure better corresponds to the notion of 'crystallized intelligence' in cognitive psychology, than to the notion of 'fluid intelligence', both first identified by Raymond Cattell (1971). While fluid intelligence correlates with measures of abstract reasoning and puzzle solving, crystallized intelligence relies on abilities that depend on specific, acquired knowledge and is thus more amenable to change as new facts are learnt. We are grateful to Bob Willis for pointing out the distinction.

[4] Static (Arrow, 1965) or dynamic (Merton, 1969; Samuelson, 1969) elementary portfolio choice models predict that, *conditional on the investor being aware of a positive equity premium*, a positive amount of the risky asset be optimally held irrespective of their risk tolerance. See Segal and Spivak (1990) or Haliassos and Bertaut (1995) for further details.

But the crisis, that has so far lead to the "Great Recession", is also shaking the foundations of macro-economics (Hall, 2010). At the heart of the debate, is the role of expectations in state-of-the-art macro-economic models (Woodford, 2013), and in particular, in their financial counterparts (Stiglitz, 2011). The standard practice has been to adopt the rational expectations (RE) paradigm, whereby households hold a (common) statistically correct unbiased view of the future. RE have a crucial advantage: rather than attempting the difficult task of measuring expectations, they can be inferred from (past equilibrium) realizations. Because the stock market is a public non-manipulable event, under RE differences in household financial choices cannot be explained by differences in what they expect, only by differences in either what they want (preferences) or, in what they have (endowments) when participation is costly. Against this received wisdom, the importance of (heterogeneous) subjective expectations in financial markets[5] has been ascertained from evidence gathered (i) in laboratory experiments (Hommes, 2011), (ii) from agent-based computational algorithms (Arthur, 2006), or (iii) from survey data (Pesaran and Weale, 2006; Greenwood and Schleifer, 2013), among others. To measure households' stock market return expectations, here we report novel survey based evidence collected in March 2007, before the financial crisis.

An incipient strand of research in survey expectations, reviewed in Hurd (2009), uncovers that households' expectations regarding the future evolution of the stock market are: (i) for the majority, no better than a 50-50 chance that the stock market index will go up in the year ahead, albeit (ii) extremely heterogeneous (Dominitz and Manski, 2007 and 2011; Kézdi and Willis, 2011); (iii) able to explain differences in financial choices both at a point in time, and through the life-cycle (Dominitz and Manski, 2007; Hurd, van Rooij and Winter, 2011; Miniaci and Pastorello, 2010), and (iv) able to identify households' implicit risk preferences, when combined with data on financial choices (Kézdi and Willis, 2009). These novel contributions rest on the methodological corner stone put by Dominitz and Manski (1997) and on Manski (2004), who advocate for treating expectations as primitives of the model, and undertake probabilistic elicitation to obtain quantitative measures of individual expectations in surveys.[6] So far, survey data has been exploited for (i) stock market investors only (Vissing-Jorgensen, 2004), for (ii) a specific population subgroup which includes non-stockholders (by age, Dominitz and Manski, 2007, 2011; Kézdi and Willis, 2009) and for (iii) a representative (internet) sample of the population by age and wealth ((Hurd et al., 2011); Miniaci and Pastorello, 2010). Here, we exploit data from a new wave of the Taylor-Nelson Sofres French survey (TNS 2007), which contains information on attitudes, preferences, subjective expectations, (a novel proxy for) individual information and socio-economic and demographic characteristics for a representative sample of 3,826 households, by age and wealth.

We contribute to this literature methodologically, since we elicit belief probability densities both *prospectively* (expectations) and *retrospectively* (information sets) for the first time, and over

[5] The very same concern has been acknolwedged in macreocnomics since very early and until today: see for example the contributions to the volumes edited by R. Frydman and E. S. Phelps (1983, 2013).

[6] But in line with (and beyond) previous work, e.g. Keane and Runkle (1990), where exploiting data on survey expectations (and information sets) allows the researcher to overcome the difficult identification problem of unbundling the test of the model behavioural equations from the test of the model assumed expectations formation mechanism.

a longer forecasting horizon (Pesaran and Weale, 2006) for a representative sample by asset classes, age and wealth (Campbell, 2006). Extending probabilistic elicitation techniques to obtain a quantitative measure of individual information sets is crucial, because "...little is known about what kind of information rational-expectations investors should learn. Since information learned determines which assets are invested in [...]" (Van Nieuwerburgh and Veldkamp, 2010). We find (i) less 50-50 percent responses to probability questions, possibly conveying absolute uncertainty, at the expense of more answers conveying absolute certainty, i.e. 0 ('no chance') and 100 ('for sure') type of answers; that (ii) average five-year ahead probabilistic stock market forecasts appear hump-shaped in age, are higher for males and increase with total wealth, and (iii) monotonically increase with the probability of holding stocks but not with educational attainment. (iv) Our measure of information, when averaged across individuals, is also hump-shaped in age (King and Leape, 1987), is higher for males, and (unconditionally) increases with total wealth. Finally, (v) the *conditional* cross-sectional average of five-year ahead probabilistic stock market forecasts also appears hump-shaped in age. From the perspective of life-cycle (heterogeneous) portfolio choice models, these novel empirical facts appear consistent with (i) observed age-portfolio profiles at the extensive margin, uncovering (ii) heterogeneity in information sets as a novel source of heterogeneity in portfolio allocations (Curcuru, Heaton, Lucas and Moore, 2010; Malmendier and Nagel, 2011).

Conditioning on risk preferences, endowments, constraints, inertial/delegation factors and information, we find that subjective expectations determine stock market participation (i) amongst the elderly, confirming the robustness of Dominitz and Manski's (2007) findings, but (ii) not amongst the young, for whom information appears instead crucial (King and Leape, 1987; Hurd, 2009). Our measure of information is consistent with information being costly acquired (Peress, 2004; Van Nieuwerburgh and Veldkamp, 2010), gathered from social interactions (Hong *et al.*, 2004) and specialised (Cabrales and Gottardi, 2011) media (Carroll, 2003), and increases with own's past experience ('frequence of recent trades' or household trade intensity, Linnainmaa 2011; Malmendier and Nagel, 2011; Seru, Shumway and Stoffman, 2009) and age. Optimists and income constrained respondents appear worse informed, consistent with rational inattention (Sims, 2003; Huang and Liu, 2007). However, information does not (conditionally) increase with the respondents' own or parents' educational attainment, family background, total wealth or respondents' preferences for either risk (Cabrales, Gossner and Serrano, 2013) or time. Most importantly, (iii) when we interpret our novel information measure as an instrumental variable, we find evidence in support of a causal effect of expectations on participation decisions, in line with elementary portfolio choice theory predictions (Arrow, 1965; Merton, 1969; Samuelson, 1969; and under incomplete information, Genotte, 1986).[7] Hence, and although households do have limited information, they appear to act rationally upon it.

The rest of the paper is organized as follows: in section 2 we describe the methodology used to elicit expectations and individual information sets. Given the aforementioned differences, we

[7] Here we abstract from non-expected utility models (ex. Dow and Werlang, 1992), and focus only on the consistency of household choices whithin a subjective expected utility framework.

construct measures of expectations similar to Dominitz and Manski (2007) to assess the quality of our data against the 2004 wave of the Health and Retirement Study (HRS 2004), which contains a much larger sample of households. In section 3 we describe the TNS 2007 data set and provide descriptive statistics. Section 4 reports the main empirical results on stock market participation, while section 5 inspects the mechanism. Finally, section 6 concludes.

2 Measuring Expectations and Information

2.1 Survey Design

In surveys, respondents are asked to state their perception of a future event in order to understand if it determines their current behaviour. The recent literature on measuring expectations privileges the use of probability questions rather than eliciting point expectations or the traditional qualitative approach of attitudinal research (Manski, 2004). Answers to such questions are used to understand if expectations and outcomes are related, and to evaluate if individual behaviour changes in response to changes in expectations. Dominitz and Manski (2007) elicit individuals' expectations of stock market returns inquiring about how 'well' the respondent thinks the economy will do in the year ahead (Positive Nominal Return, PNR). They exploit data for a representative sample of the elderly from the 2004 wave of the U.S. Health and Retirement Study (HRS).

To validate our dataset, we build upon their work and extend it along different dimensions. First, by extending the forecasting horizon to five years, we intend to untie expectational answers from the bussiness-cycle conditions prevailing at the time of the survey (March 2007) to better capture (i) the historic average upward trend of the stock market index, and (ii) inertia in portfolio management (Bilias et al., 2010). The latter is important since it remains an open question with which horizon households invest in the stock market. Second, and to comply with limitations of survey administration (fill-in questionnaires as opposed to telephone interviews in the HRS), we extend the methodology of the Survey on Household Income and Wealth (SHIW) conducted by the Bank of Italy (Guiso et al., 1996) to the stock market. Probability densities are elicited on seven points of the outcome space, instead of just two points of the cumulative distribution functions (cdfs.), to obtain more precise individual estimates of the relevant moments. Third, we exploit data from a representative sample by age, wealth and asset classes to examine the relationship between age-portfolio profiles and subjective expectations at the extensive margin. Finally and most importantly, probabilistic elicitation of recent past stock market performance (past Positive Nominal Return, pPNR) provides a quantitative measure of households' degree of awareness regarding their investment opportunity set, to capture: (i) differences in information across households (Guiso and Jappelli, 2005), and (ii) the relationship between information and expectations (Van Nieuwerburgh and Veldkamp, 2010). Without it, households who do not invest because they expect the stock market to burst over the given forecasting horizon are indistinguishable from those who do not invest because they are (pessimistic and) unaware of the investment opportunities available in the stock market.

The new wave of the Taylor-Nelson Sofres French survey (TNS 2007) was designed by researchers at the Paris School of Economics (PSE), and administered by Taylor-Nelson Sofres, a professional agency paid with research funds from the Agence Nationale pour la Recherche (ANR). The first wave, carried in 2002, had no questions related to stock market expectations. The 2007 wave contains very detailed information on attitudes, preferences and expectations, in addition to wealth, income and socio-economic and demographic characteristics for a representative sample of French households. A questionnaire was sent to 4,000 individuals, corresponding to an equivalent number of households. Respondents had to fill the questionnaire, and return it by the post in exchange of around €25 in coupon-tickets (*bons-d'achat*). 3,826 respondents sent their questionnaires back, representing a 97% response rate.

The survey was conducted in March 2007. Figure 1 below shows that after a drop of nearly 60% in the French stock market Index (CAC-40)[8] caused by the 'dot-com crash' of 2001, by the time the survey was conducted, the stock market index had been steadily recovering since the mid 2003. In March 2007 the index was still below its 'dot-com' peak. Hence, it is likely that respondents were particularly aware of the stock market evolution regarding the past, and provide very heterogeneous and uncertain answers regarding the stock market prospects for the five years to come, given the recent experience of a bust and a boom.

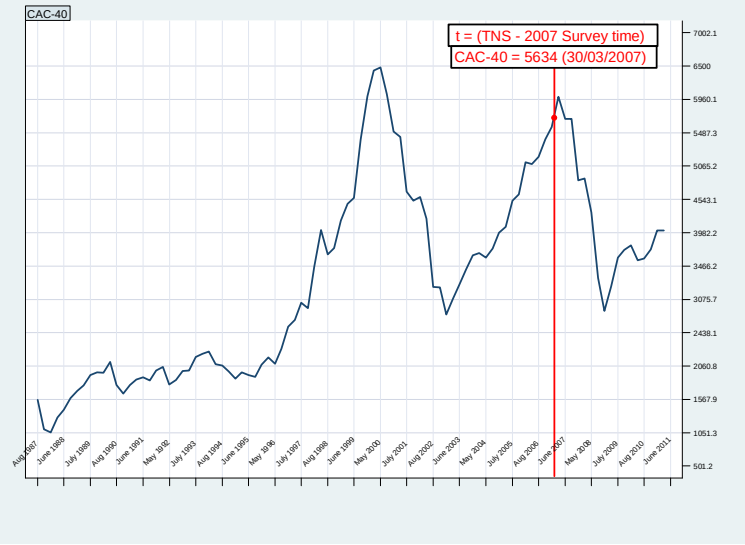


Figure 1: French Stock Market Index CAC-40 between July 1987 and July 2011. The survey was conducted in March 2007. Source: Author's own calculations from MSN Money hsitorical data.

Since the survey was conducted in March 2007, exploiting available public information from monthly data between July 1987 and March 2007 yields the relevant sample moments $\mu = 0.023$ ($\mu(5) = 0.108$) and $\sigma = 0.10$ ($\sigma(5) = 0.19$) of nominal yearly (and 5-year rolling) log returns on the

[8]The CAC-40 takes its name from the Paris Bourse's (today called Euronext Paris) early automation system "Cotation Assistée en Continu" (Continuous Assisted Quotation). Its base value of 1,000 was set on the 31^{st} of December 1987, equivalent to a market capitalisation of 370,437,433,957.70 FF.

CAC-40.[9] Conditional on (continuously compounded) returns (on a buy-and-hold portfolio tracking/mimicking the index) being normally distributed, those sample moments would characterize the subjective beliefs of those respondents who base them on the history of observed stock market index monthly closing values.[10]

2.2 Expectations

To measure expectations, we elicited households' subjective beliefs regarding the likely evolution of the stock market index five years ahead in time, I_{t+5}, relative to March 2007, I_t, from the following questions (translated wording):

C6. 'Five years from now, do you think that the stock market... -For each category write down the likelihood of occurrence assigning a value between 0 and 100. The sum of all your answers must be equal to 100-:

... will have increased by more than 25%

... will have increased by 10 to 25%

... will have increased by less than 10%

... will be the same

... will have decreased by less than 10%

... will have decreased by 10 to 25%

... will have decreased by more than 25%

C7b. 'In your opinion, if you expect the stock market to increase within the next 5 years, which would be the highest possible increase (as a percentage)?'

C8b. 'In your opinion, if you expect the stock market to decrease within the next 5 years, which would be the lowest possible decrease (as a percentage)?'

Question C6 inquires household i about the subjective relative likelihood of occurrence, $p^i_{t+1,k}$, of each of the seven alternative scenarios, $k = 1, ..., 7$. Each scenario represents a possible outcome range for the percentage change in the index between t and $t+5$, $R_{t+1}(5) \equiv \frac{I_{t+5}}{I_t} - 1$.[11] Questions

[9] The density of nominal yearly (and 5-year rolling) log returns on the CAC-40 computed from monthly data between July 1987 and July 2011 is depicted in Figure 4, panel (a) (panel (b)) in the not for publication appendix B. The distribution has sample moments $\mu = 0.034$ ($\mu(5) = 0.109$) and $\sigma = 0.093$ ($\sigma(5) = 0.188$).

[10] Those respondents are also more likely to form a rational expectation, at least from the perspective of the adaptive learning literature. See for example Evans and Honkapohja (2001), where they characterize the general conditions under which, even if individuals are initially uncertain about the underlying structure of the economy, they can end up learning it in the limit from equilibrium realizations.

[11] We follow the standard convention in finance for long-horizon returns, and let $1+R_{t+1}(s)$ denote the stock market index gross return over s *periods ahead* (hence the subindex $t+1$), which is equal to the product of the s single-period (or yearly) returns:

$$1 + R_{t+1}(s) = \prod_{f=0}^{s-1} (1 + R_{t+1+f}) = \prod_{f=0}^{s-1} \left(\frac{I_{t+1+f}}{I_{t+f}} \right)$$

Similarly, we let $1 + R_t(s)$ denote the stock market index gross return *over the most recent* s periods from date $t - s$ to date t (hence the subindex t):

$$1 + R_t(s) = \prod_{b=0}^{s-1} (1 + R_{t-b}) = \prod_{b=0}^{s-1} \left(\frac{I_{t-b}}{I_{t-1-b}} \right)$$

See Campbell *et al.* (1997) for details.

C7b and C8b provide subjective upper and lower bounds for the index percentage change, $R^i_{\max}$ and $R^i_{\min}$ respectively. The corresponding outcome ranges are:

$$R_{t+1} \in \left\{ \underbrace{[R^i_{\max}, 0.25)}_{k=1}, \underbrace{[0.25, 0.10]}_{k=2}, \underbrace{(0.10, 0)}_{k=3}, \underbrace{\{0\}}_{k=4}, \underbrace{(0, -0.10)}_{k=5}, \underbrace{[-0.10, -0.25]}_{k=6}, \underbrace{(-0.25, -R^i_{\min}]}_{k=7}, \right\}$$

and households' subjective likelihoods are accordingly:

$$p^i_{t+1,k} \equiv \Pr{}^i \left[R_{t+1} \in k \right] = \Pr{}^i \left[\frac{I_{t+5}}{I_t} - 1 \in k \right], \forall i$$

Out of the 3,826 sample respondents, around 63% (2,406) meaningfully answered to the expectations question. 328 cases are excluded as the sum of their answers do not round up to 100.[12] Figure 2 below depicts the histogram, which averages the individual probability density functions of those who answered.[13] On average, households appear more pessimistic and less uncertain (Table 2) than the historical record until March 2007 would predict.

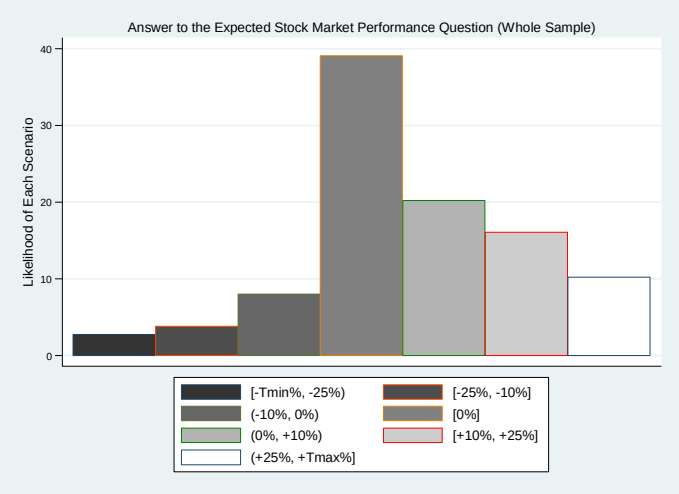


Figure 2: Histogram of average individual answers to the likelihood of the different scenarios regarding 5-year ahead stock market performance. Source: TNS 2007 survey.

To validate our novel survey data, and for comparison purposes, we construct from respondents' answers Dominitz and Manski's (2007) Positive Nominal Return (PNR) variable which captures

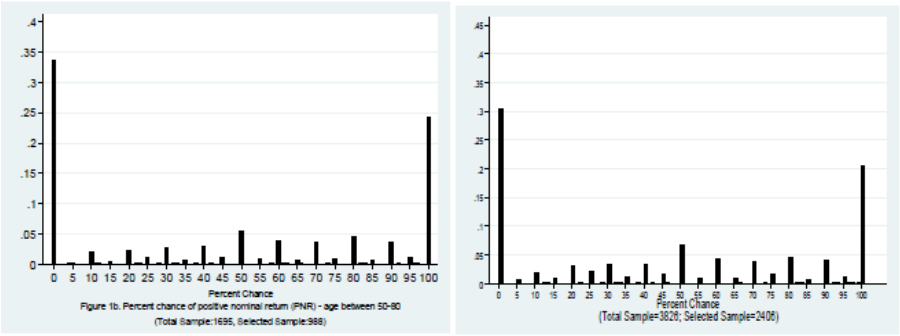
[12] Missing and erroneous answers are regressed against stockholding status, and a set of covariates (gender, marital status, education, risk preference) using a probit (Table 1 in the not for publication appendix B).

[13] The four panels ((a) - (d)) in Figure 2 in the not-for publication appendix B illustrate how do elicited probability density functions look like for a small subset of individuals.

the 5-year ahead percentage chance of a positive nominal return [14]:

$$\forall i: \quad PNR^i \equiv \Pr{}^i \left[R_{t+1}(5) > 0\right] = \Pr{}^i \left[\frac{I_{t+5}}{I_t} - 1 \in \cup_{k=1}^{3} \{k\}\right] = p^i_{t+1,1} + p^i_{t+1,2} + p^i_{t+1,3}$$

Figure 3 below, depicts the frequency distribution of responses, for ages in the 50-80 age bracket (panel (a)) and for all ages (panel (b)). Despite of the age differences, the similarities with both the HRS or the De Nederlandsche Bank Household Survey (DHS), exploited by Hurd et al. (2011), are striking: (i) there is similar hipping of responses around round numeric probability answers. (ii) For all ages, the mean response is 46.5%, while for the elderly, it is 47.1%. This compares with a 49% mean response, for the 50-80 HRS 2004 respondents, and with a 41.6% (50.1%) for a representative internet sample of the DHS 2004 (2006) respondents. However, hipping is much stronger in the {0,100} answers, than in the 50 percent chance response, indicating less epistemic uncertainty according to Bruine de Bruin *et al.* (2000).[15] For all ages, 31% (21%) gave answers consistent with absolute certainty that the index would go down (up) over the coming 5 years. Evidence from the empirical finance literature on long horizon returns suggests that the longer time horizon given to evaluate stock market performance might explain the differences, because of mean-reversion (Campbell *et al.*, 1997). In the next subsection we further examine this question.



Panel (a): Ages 50-80. Panel (b): All ages.

Figure 3: PNR by age groups, frequency distribution. Source: TNS 2007.

Consistent with the findings reported in Dominitz and Manski's (2007) Table 1, the (conditional) mean percentage chance of a positive nominal return is higher for respondents who are married and for males.[16] In line with the HRS 2002 findings by Kézdi and Willis (2009), expectations

[14] In Dominitz and Manski (2007), 15,166 HRS respondents, aged 50 to 80 in 2004, were asked:
Positive Nominal Return (PNR): We are interested in how well you think the economy will do in the next year. By next year at this time, what is the percent chance that mutual fund shares invested in blue chip stocks like those in the Dow Jones Industrial Average will be worth more than they are today?

$$\forall i: \quad PNR^i_{DM} \equiv \Pr{}^i \left[\frac{I_{t+1}}{I_t} - 1 \in \cup_{k=1}^{3} \{k\}\right]$$

[15] Kleinjans and Van Soest (2013) explicitly model (six) anomalies in the HRS 'reporting behaviour' and find that incorporating them has small effects on the estimated distribution of the genuine subjective probabilities.

[16] Table 2 in the not for publication appendix C reports the distribution of responses and the response rate conditioning on age, gender, and marital and stockholding status.

(noisily) increase with the respondent's education and households' total wealth. The most affluent households (with wealth above the 90th percentile, €413,476) appear more optimistic regarding the future evolution of the stock market. However, respondents with some college education or more are only slightly more optimistic than those having at most completed high school, while both become similarly more optimistic as they age.[17]

Finally, expectations of a positive nominal return appear roughly hump-shaped in age, as does the response rate to the probabilistic question. In Figure 4, the mean response increases until the late 40s, when expectations reach its peak, only to decline from the mid 60s, although the pattern is very noisy with potential time/cohort effects present.[18] The mean percentage chance of a positive nominal return is estimated to increase (fall) by about 8 to 12 (5 to 6) percentage points as age increases (decreases) from 20 to 50 (late 60s onwards).

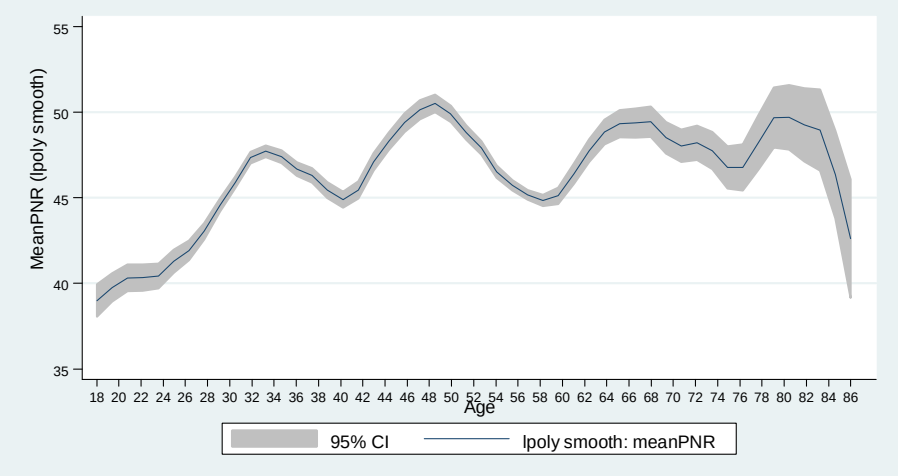


Figure 4: Mean PNR, conditional on age. Source: TNS 2007.

Standard life-cycle portfolio choice models predict that, conditional on being aware of the existence of a historical equity premium, the young should invest heavily in the stock market to take advantage and quickly accumulate wealth (Guiso et al., 2002; Ameriks and Zeldes, 2004; Gomes and Michaelides, 2005).[19] The descriptive evidence reported here suggests that expectations vary systematically with age, and that both the young and the elderly appear particularly pessimistic. Although the evidence is consistent with existing findings for the elderly in the US, in France the

[17] Figures 6 and 7 in the not for publication appendix C display kernel-smoothed estimates of the mean percentage chance of a positive nominal return conditional on total wealth and educational attainement, respectively.

[18] In line with the findings reported by Linnainmaa (2011) or Malmendier and Nagel (2011), comparison of Figure 4 (expectations) with Figure 9 (information) in the next subsection suggests that the time effects possibly correspond to respondents overweighting a particular event in the past to which they were exposed, like the burst of the technology bubble in 2002.

[19] Except if labour income and dividends were cointegrated (Benzoni *et al.*, 2007), in which case even if the young were aware of the historical equity premium, they would be willing to short stocks. Intuitively, the young would be implicitly over-exposed to stock market risk through their human capital investments (the returns of) which are non-diversifiable and highly correlated with dividends in the long run, i.e. the longer the available (life-cycle) investment horizon.

young appear pessimistic rather than optimistic (Dominitz and Manski, 2011; Vissing-Jorgensen, 2003).

2.3 Measuring Information

Another possibility is that the young are particularly unaware of the investment opportunities offered by the stock market (King and Leape, 1987; Hurd, 2009). Recent studies stress the importance of (i) coginitive ability (Christelis *et al.*, 2010; Grinblatt *et al.*, 2011), of (ii) the stock of specific knowledge, like financial literacy[20] (van Rooij *et al.*, 2011), or of (iii) measures of sources of (ex. social interactions, Hong *et al.* (2004); time spent in acquiring financial information, Guiso and Jappelli (2006); frequency of and gains/losses in recent stock market operations, Linnainmaa, 2011) or lack of access to (ex. lack of trust, Guiso *et al.*, 2008) that specific knowledge, when accounting for stock market participation decisions. Here, we are more specific and inquire respondents about the most recent stock market return realization over the relevant forecasting horizon, in line with the finance literature (e.g. Biais, Bossaerts and Spatt, 2010; Campbell *et al.*, 1996; Malmendier and Nagel, 2011; Zhang, 2006).

To that purpose, we inquired respondents about the likely evolution of the stock market index over the five years *prior* (I_{t-5}) to the time of the survey (March 2007), I_t, as follows (translated wording):

C9. 'Over the past five years, do you think that the stock market... -For each category write down the likelihood of occurrence assigning a value between 0 and 100. The sum of all your answers must be equal to 100-:

... has increased by more than 25%
... has increased by 10 to 25%
... has increased by less than 10%
... has remained the same
... has decreased by less than 10%
... has decreased by 10 to 25%
... has decreased by more than 25%

Question C9 inquires household i about the subjective relative likelihood of occurrence, $p^i_{t,k}$, of each of the seven alternative scenarios, $k = 1, ..., 7$. Each scenario represents a possible outcome range for the percentage change in the index between $t-5$ and t, $R_t(5) \equiv \frac{I_t}{I_{t-5}} - 1$. Since ranges $k = 1$ and $k = 7$ are unbounded, we set $(R_{\max}, R_{\min})$ to match observed values.The outcome ranges for R_t are therefore identical to those of question C6 described above. Accordingly, households' subjective likelihoods are given by:

$$p^i_{t,k} \equiv \Pr{}^i \left[R_t \in k\right] = \Pr{}^i \left[\frac{I_t}{I_{t-5}} - 1 \in k\right], \forall i$$

[20] Financial literacy studies ask both general numeracy questions, which proxy for individual ability, and specific questions regarding elementary notions about the working of financial markets, which measure the stock of individual knowledge. See Lusardi (2008) for a detailed explanation.

Five years prior to the time when the survey was conducted (March 2002), the stock market index was around half-way down the 'dot-com' bust. But, from the beginning of March 2002 (CAC 40 = 4688.02) until the beginning of March 2007 (CAC 40 = 5634.16), the index had increased an overall 20.2%. Therefore, the truth belongs to category $k = 2$, "[the stock market]...has increased by 10 to 25%". Figure 5 illustrates the wanderings of the CAC-40 index between 1987 and 2011.

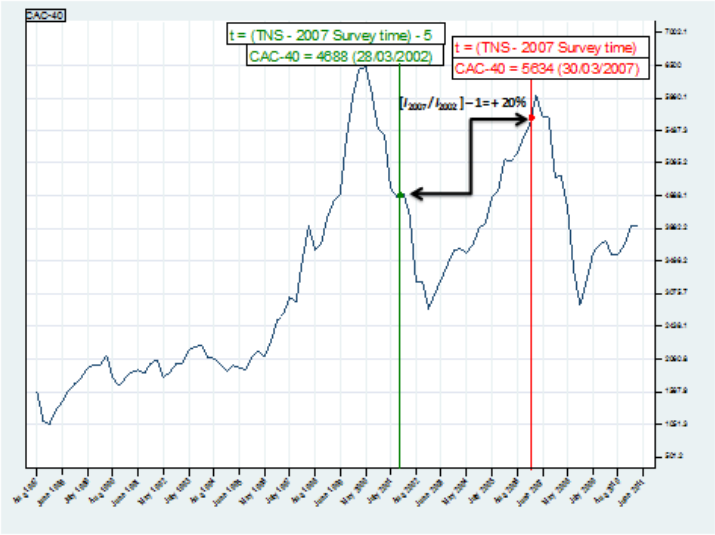


Figure 5: The French Stock Market Index CAC-40 between July 1987 and July 2011. Between March 2002 and March 2007 (5 years prior to the time of the survey) the index had increased by around 20%. Source: Author's calculations from MSN Money monthly data, available online.

To capture (i) heterogeneity in accessing/processing public information and (ii) the relationship with answers to forward-looking probability questions (Dominitz and Manski, 2011), information regarding past stock market performance is elicited as a probability density function. While the mean response conveys what most likely happened according to the respondent, the dispersion of the probability mass convenes the respondent's subjective degree of reliability in the respondent's mean response. According to Figure 5, a perfectly informed individual should allocate all probability mass (100 points) to the outcome range "...has increased by 10 to 25%" ($k = 2$).[21] Out of the 3,826 sample respondents, around 59% (2,253) provided a meaningful answer to the information question.[22] Figure 6 below depicts histogram of the average of the individual probability density functions. Surprinsingly, the modal response coincides with the truth (outcome range $k = 2$) indicating that respondents are on average well informed.

A striking finding is that households tend to be also pessimistic regarding how well has the stock market performed over the last five years. Although this might be due to imperfect recall given the unusually long horizon, it might also be related to the 'dot-com' bust being overweighted on respondents' memory (Hurd *et al.*, 2011), even if only half the bust is inside the time window

[21] The panels (a) - (d) in Figure 3 in the not for publication appendix B, illustrate how do individual information sets look like.

[22] 322 cases are excluded as the sum of their answers do not round up to 100.

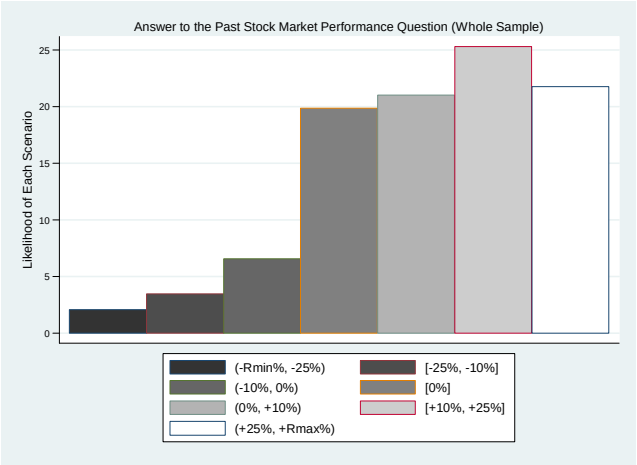


Figure 6: Histogram of average individual answers to the relative likelihood of the different scenarios regarding the stock market performance over the last 5 years. Source: TNS 2007.

spanned by the question. Table 2 reports summary sample statistics for respondents' answers regarding past and future stock market returns, imposing a uniform distribution within the different outcome ranges. Although the big spread around the sample mean realized return came as no surprise (possibly indicating ambiguity), it is remarquable that it remains smaller than the spread around the sample mean expected stock market return. Notice also that the cross-sectional dispersion of both expected and realized stock market perfomance measures is about twice the size of the respective spreads, indicating that 'disagreement' is very important, in line with the behavioural finance literature (e.g. Hong and Stein, 2007).

Table 1: Sample Moments from Questions C6 (Expectations) and C9 (Information).

Variable	No obs.	Mean	Std. Dev.	Min	Max
Expected Return (ER)	2460	0.055311	0.112602	-0.625	1.125
Std. Dev. of ER	2460	0.068028	0.07347	0	0.43056
Mean (Realized) Past Return (pR)	2231	0.11938	0.139876	-0.375	0.375
Std. Dev. of pR	2231	0.065598	0.069211	0	0.375

Source: TNS 2007.

For consistency, we construct from respondents' answers the past Positive Nominal Return (pPNR) variable which captures the percentage chance of a positive nominal return between March 2002 and March 2007:

$$\forall i: \quad pPNR^i \equiv \Pr{}^i\left[R_t(5) > 0\right] = \Pr{}^i\left[\frac{I_t}{I_{t-5}} - 1 \in \cup_{k=1}^{3}\{k\}\right] = p^i_{t,1} + p^i_{t,2} + p^i_{t,3}$$

In Figure 7, we depict the frequency distribution of responses to pPNR for all ages. As previously, there is hipping of responses around round numeric probability answers indicating that rounding is not specific to forward looking questions but rather, to respondents rounding when

confronted with the probabilistic elicitation format. For all ages, the mean response is 68%, while the true answer is a 100% chance of a positive nominal return over the last 5 years. Around 44% of sample respondents (990 individuals) gave an answer consistent with the truth.

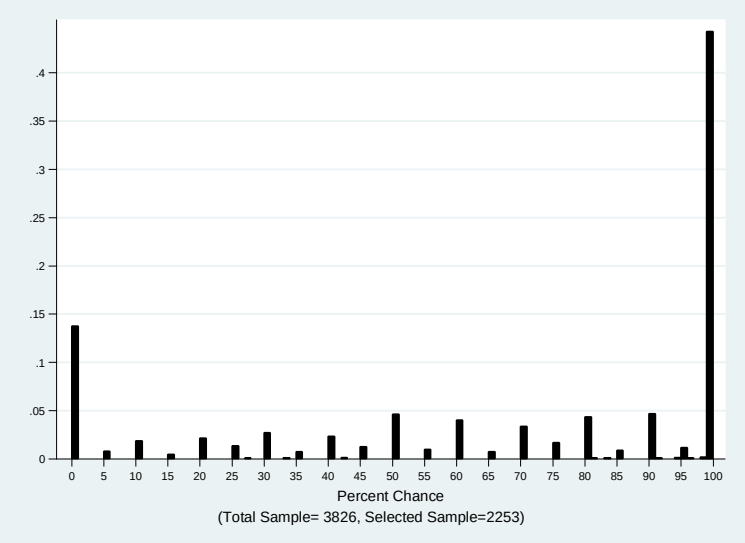


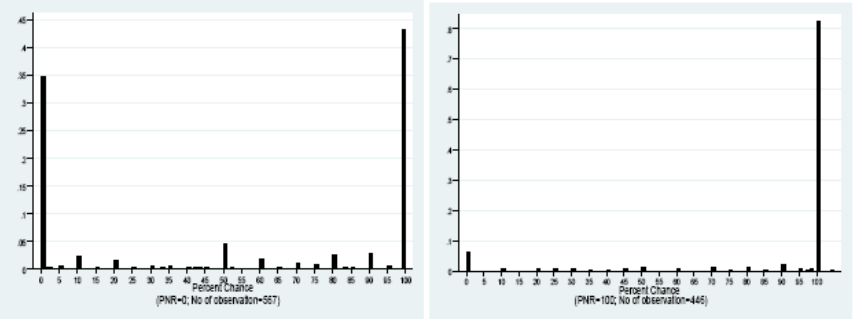
Figure 7: Percentage chance of a Positive Nominal Return over the past five years (pPNR) for all ages, frequency distribution. Source: TNS 2007.

In Figure 8 we examine what type of information had those respondents who were absolutely certain regarding the future evolution of the stock market (see panel (b), Figure 3). Panel (a) shows that amongst those who were absolutely certain that the stock market would go down (24%, 567 answered PNR=0%), around 35% were absolutely certain that it had not increased over the last 5 years, while 43% were absolutely certain that it had gone up. Around 5% gave a 50 percent chance of either going up or down. This contrasts with panel (b), for respondents who were absolutely certain that the stock market would go up (21%, 446 answered PNR=100%): 83% were absolutely certain that the stock market had gone up, while only 6% gave answers consistent with absolute certainty about the stock market having gone down. Only 2% gave a 50 percent response. Hence, individual information regarding the past contributes to (i) significantly unbundle the 'hipping' on responses conveying absolute certainty regarding the future, i.e. $PNR = \{0, 100\}$, thereby (ii) capturing a novel source of heterogeneity amogst respondents (Curcuru *et al.*, 2010).

Table 4 in the not for publication appendix D reports the distribution of responses and the response rate conditioning on age, gender and stockholding status[23]. In accordance with the findings reported in Lusardi (2008) on the financial literacy of US adults, male respondents who are older, single and stockholders report higher mean (and a lower standard deviation of) percentage chances of a past positive nominal return. Although information broadly increases with age irrespective of gender, males are better informed than females on average.[24] In addition, information broadly

[23] Stockholders report a higher mean by about 10 percentage points, and are around 6 percentage points more likely to give a response.

[24] See Figure 8 in the not for publication appendix D.



Panel (a): How much do pessimists know? pPNR conditional on PNR=0%.

Panel (b): How much do optimists know? pPNR conditional on PNR=100.

Figure 8: pPNR conditional on PNR, for pessimists (a) and optimists (b), frequency distribution. Source: TNS 2007.

increases until the 50th percentile of wealth (€118,792), remains constant until the 90th percentile (€413,476), only to increase again albeit very heterogeneously.[25] The richest households (with wealth above the 90th percentile), may thus be more optimistic (and disagree more) regarding the future investment opportunities because they are better (albeit more heterogeneously) informed.

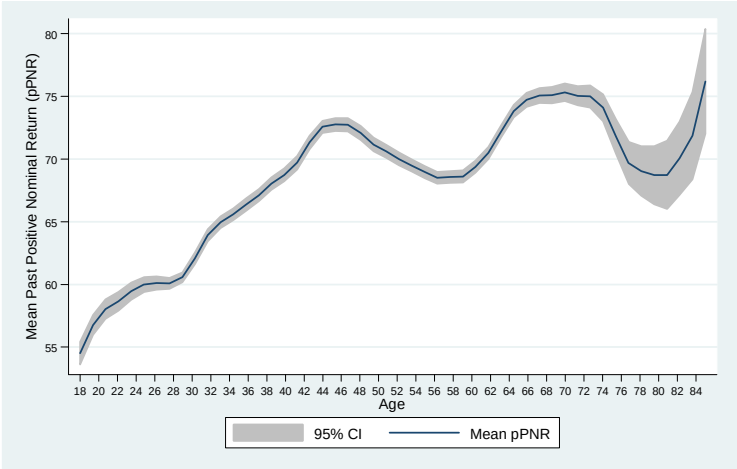


Figure 9: Mean percentage chance of a positive nominal return over the last 5 years (pPNR) by age. Source: TNS 2007.

Figure 9 presents kernel-smoothed estimates of the mean percentage chance of a past positive nominal return conditional on age. Information about past stock market performance sharply increases with age until the mid 40s, mildly increases until the mid 70s, and then decreases, although the point estimates are much noisier.[26] In line with King and Leape's (1987) conjecture and Lusardi

[25] However, the increased heterogeneity in information in the top decile of the wealth distribution might just be a small sample problem, since only 272 respondents within that decile answered to the information question. See Figure 9 for total household wealth by age in the not for publication appendix D for additional evidence.

[26] The mean percentage chance of a past positive nominal return is estimated to increase (fall) by about 20 to 22

et al.'s (2012) structural estimations, figure 9 provides evidence in support of an 'informational stock' being slowly built through respondents' life-cycle, which (slightly) depreciates by the end of it. Importantly, this process of life-cycle information accumulation may explain the identified pessimism regarding the future performance of the stock market amongst the young and the elderly (see Figure 4), providing support to the relationship between expectations and information sets further examined below.[27]

Surprisingly, information about past stock market performance does not increase with own's educational attainement, although it broadly increases with age within educational groups.[28] This finding is confirmed in Table 2 below, which reports marginal effects of estimating a two-way censored Tobit specification for answers to question C9 (censored below by '0' and above by '100'). The likelihood of being informed is specified to be a function of age, gender, sources of advice (friends, family, professional, broad media, specialised media) and of information (TV, economics/finance emissions), own and parents' educational attainment, family background (middle/lower/other class), endowments (income and wealth), financial decision taking (no/partial/complete delegation of financial decisions), own past experience (proxied by the 'frequency of recent trades'), preferences (risk aversion and impatience), constraints in either accessing information ('online banking') or related to inertia in informational sources (parents' stockownership status, 'parents own stocks'), and of the tightness of households' budget constraint ('importance of money in life').

Categorical answers to frequency, variety and access specialised media, advice from professionals, as well as age or the number of stock market transactions carried over the last year, increase the likelihood of being informed.[29] Interestingly, parents' stockownership status and educational attainment or family background (omitted from the table) do not increase the odds of being informed, and actually significantly decreases them for those who follow 'family advice'. Since those who follow 'friends' advice' are more likely to be informed, we interpret it as being consistent with social interactions being instrumental in gathering information (Hong et al., 2004): while friends are a product of respondents' choices, the family in which they are born is not. Alternatively, advice from the family could be capturing trust, and thereby rationalize Guiso, Sapienza and Zingales' (2008) finding regarding the negative impact of lack of trust on stock holdings. On the other hand, a measure of optimism ('being lucky in life') has a negative impact on being informed, indicating that an 'overconfidence bias' is not present once gender is conditioned upon: although males appear

(10 to 12) percentage points as age increases until (decreases after) 75.

[27] As with expectations, similar event-related time effects seem to be present (Malmendier and Nagel, 2011). For example, if the information accumulation process peaks at the age of 48, a less well informed respondent aged 53 in 2007 was 48 years old in 2002 when the 'dot-com' bubble burst. Hence, the apparent decline in the reported mean realized return of the average respondent in the 48-60 age bracket appears consistent with respondents overweighting the 'dot.com' bust (and hence being rationally inattentive; Sims, 2003), and also might explain why they appear more pessimistic regarding the future (Figure 4) if they base their expectations on such an information set (Van Nieuwerburgh and Veldkamp, 2010).

[28] See Figure 10 in the not for publication appendix D.

[29] The positive and significant effect of access to specialised media and professionals' advice on households' information is consistent with the epidemiological model of Chris Carroll (2003), as well as with the existence of a market for financial information where quality is priced, i.e. the vertical differentiation component in Cabrales and Gottardi (2011).

Table 2: The Determinants of Information (pPNR)

	(1)	(2)	(3)	(4)	(5)
Male	23.73***	21.91***	22.28***	19.57***	17.72***
	(3.631)	(3.624)	(3.618)	(3.744)	(3.782)
Age	2.521***	1.978***	1.920***	1.284*	1.267*
	(0.643)	(0.65)	(0.65)	(0.690)	(0.712)
Age squared	-0.0196***	-0.0146**	-0.0132**	-0.00840	-0.00849
	(0.0065)	(0.00652)	(0.00656)	(0.00683)	(0.00702)
Friends advice		9.551**	8.791**	9.175**	8.357*
		(4.427)	(4.418)	(4.431)	(4.450)
Professional advice		3.609	3.459	2.069	-0.419
		(4.091)	(4.079)	(4.101)	(4.178)
Family advice		-14.62***	-14.40***	-15.29***	-13.93***
		(4.162)	(4.15)	(4.167)	(4.204)
Media advice		0.211	-0.469	0.150	-3.743
		(5.251)	(5.239)	(5.223)	(5.210)
Specialised media advice		19.81***	18.55***	16.09***	11.01**
		(5.194)	(5.186)	(5.191)	(5.260)
Information from TV		-8.264*	-7.293	-7.192	-7.624*
		(4.528)	(4.52)	(4.515)	(4.507)
Information from economics emissions		-1.52	-0.286	-1.373	-0.625
		(4.03)	(4.03)	(4.025)	(4.036)
Education (Ref. category: High school or less)					
Less than college			8.047	4.086	-1.265
			(6.329)	(6.361)	(6.473)
College or more			22.34***	14.68**	6.443
			(6.935)	(7.172)	(7.456)
Log (Income)				8.797***	8.431***
				(3.262)	(3.265)
Log (Total wealth)				1.377	0.312
				(1.269)	(1.315)
Self account management				5.892	1.457
				(3.654)	(3.709)
Frequency of recent trades					13.97***
					(2.349)
Risk aversion (CARA)					27.71
					(22.68)
Mother education (Ref. category: Less than secondary)					
Secondary					-2.140
					(7.169)
Higher than secondary					-1.731
					(6.692)
Father education (Ref. category: Less than secondary)					
Secondary					-0.404
					(6.897)
Higher than secondary					-3.839
					(7.484)
Importance of money in life					-4.548***
					(1.065)
Online banking					5.502
					(4.643)
Optimism					-6.622*
					(3.684)
Parents own stocks					0.495
					(4.192)
Temporal preference					-0.818
					(0.792)
Constant	30.70**	44.63***	32.50**	-47.72	-8.927
	(14.91)	(15.47)	(16.35)	(29.68)	(32.38)
Log-likelihood	-6530	-6507	-6499	-6530	-6507
No. of observations	2,253	2,253	2,253	2,200	2,133

Note: (i) Robust standard errors are in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

better informed, supporting more optimistic forward looking expectations, optimists appear consistently worse informed.[30] We do not find evidence of either temporal or risk preferences determining information sets, in line with Van Nieuwerburgh and Veldkamp's (2010) theoretical prediction that risk aversion does not determine the demand for information.[31] Although total wealth does not increase the odds of being informed, income does (even if we condition on the number of stock market operations carried over the last year), in line with a costly information acquisition interpretation (Peress, 2004).[32]

Overall, these findings are consistent with financial information being slowly acquired through the life-cycle (King and Leape, 1987; Lusardi et. al., 2012) from own past experience (Linnainmaa, 2011; Malmendier and Nagel, 2011), from others (Hong et al., 2004) and from specialised media (Carroll, 2003). On the other hand, the negative effect of the 'importance of money in life', which scores higher the poorer and the more financially constrained the respondent is, reveals a novel aspect in information acquisition: disenfranchisement. Notice that the negative impact of both disenfranchisement and optimism is consistent with rational inattention theory (Sims, 2003), since both decrease the expected returns of costly gathering/processing publicly available relevant information.

3 Expectations, Information and Stock Market Participation

An important puzzle in the literature is why so few households hold stocks (Haliassos and Bertaut, 1995). Table 3 reports the frequency of *non*-stock holders by motive for *not* holding stocks, as well as their relative incidence by age groups.[33] Quantitatively, the percentages reported suggest borrowing/liquidity constraints (Guiso *et al.*, 1996), being 'too risky' and 'having other priorities' as the most important, closely followed by not trusting the stock market (Guiso *et al.*, 2008), being uninformed and entry/management costs being too high (Vissing-Jorgensen, 2002) in accounting for overall non-participation amongst non-participants.[34]

[30] This may help reconciling Bilias et al.'s (2010) findings consistent with inertia in households' portfolios (linked to limited ressources and limited education), with Guiso and Jappelli's (2006) fidings consistent with excess trading even amongst the general population, if those who overtrade are also optimists (and hence, worse informed).

[31] Our measure of temporal preference is inversely proportional to 'impatience', measuring how much the respondent cares about the future (rather than a preference for an early resolution of uncertainty). This empirical finding does not contradict Cabrales, Gossner and Serrano's (2013) theoretical prediction of a positive relationship between risk tolerance and the value of information, because their prediction holds under the restriction of a common prior (pinned-down by the set of arbitrage free asset market prices) while here we report empirical evidence consistent with heterogeneous priors.

[32] In unreported regressions, we estimated an ordered probit on the number of stock market operations carried over the previous year (grouped in 5 categories), and find evidence consistent with wealth and income increasing the number of trades, but also with liquidity and borrowing constraints forcing households to trade at the extensive margin. These results are available upon request.

[33] Question C18 inquires non-stockholders about the reasons for not holding stocks, and the following options were given: (1) 'I do not have enough liquidity', (2) 'It is too risky', (3) 'I am poorly informed', (4) 'I do not trust the stock market', (5) 'Fixed entry costs are too high', (6) 'Management costs are too high', (7) 'I have other priorities'.

[34] Notice that the relative importance of the different motives changes by age group: having other priorities and lack of trust in the stock market slightly *increases* with age, while being uninformed significantly *decreases* with age, with the elderly being around 40% less likely not to participate than the young, amongst non-participants.

Table 3: Reasons for not Holding Stocks, by Age (%)

	All Ages		Age<50		Age≥50	
	Whole sample	Respondents	Whole sample	Respondents	Whole sample	Respondents
I am liquidity constrained	24.9	23.6	24.7	24	25	22.8
It is too risky	20.7	20.8	19.9	20.6	21.9	21.5
I am uninformed	12.8	12.9	14.5	14.9	9.4	9.3
I don't trust stock market	15.8	15.9	14.8	14.7	17.2	17.9
Entry costs are too high	3.9	4.2	6.2	4.2	4	4.2
Management costs are too high	4.6	4.8	4.2	4.6	4.8	5.2
I have other priorities	17.3	17.8	15.7	17	17.7	19.1

Source: TNS 2007.

Although in an expected utility framework, the standard two-asset model predicts that decision takers invest in the risky asset if and only if its expected return exceeds the return of the riskless asset (Arrow, 1965; Merton, 1969; Samuelson, 1969), only recently have researchers started to collect data on subjective expectations of stock market returns (Dominitz and Manski, 2007; Kézdi and Willis, 2009; Hurd et al., 2011).[35] Here we go one step further and examine the extent to which *conditional* subjective expectations determine households' stock ownership decision, *conditioning on what they know* (Merton, 1987; O'Hara, 2003; Biais et al., 2010).

To obtain a measure of stock ownership, question C19 in the TNS 2007 inquires respondents about the different types of financial instruments and accounts they hold, and in particular whether they invest in the stock market either directly or indirectly. We define direct stockholdings as the sum of stocks of privatised public companies, listed stocks of private companies and stocks of foreign firms held. Indirect stockholdings are those held through mutual funds and managed investment accounts.[36] The proportion of households who hold stocks directly is 22%, and 37% either directly or indirectly. Although low, the participation rates are slightly higher than those obtained from previous past surveys[37] and similar to the figures reported by Haliassos (2008) for other countries at that time. In Figure 10, stock market participation amongst respondents displays a clear hump-shaped pattern by age.

The literature on household finance has found that those who are better educated, older and wealthier, are more likely to hold stocks. Dominitz and Manski (2007) report that the probability of holding stocks monotonically increases with the perceived chance of a positive return of investing in the stock market amongst the elderly.[38] Table 4 reproduces their Table 2, albeit exploiting data

[35] In particular, the decision to invest in the stock market does not depend on either preferences, endowments or information about past stock market performance: just on the subjective expected return. See Arrow (1965) and Segal and Spivak (1990) or Haliassos and Bertaut (1996) for a proof and clarifications.

[36] We exclude both government bonds and homeownership from the risky asset category, even if the latter are highly illiquid and indivisible (and therefore risky), because French households mostly buy houses for the flow of services they provide rather than as a financial investment. Still, in the estimation we control for the level of total wealth (real plus financial) and include a dummy variable that takes value one when home-ownership status is observed.

[37] For the 35-55 year-olds corresponding subsample in the Patrimoine 1998 INSEE survey, the proportion of households holding risky assets 'directly' is 21.6 and either directly or through mutual funds, 32.4.

[38] Table 3 in the not for publication appendix C shows a similar qualitative pattern emerges for the elderly (all ages) in the TNS 2007, consistent with their findings and theoretical predictions.

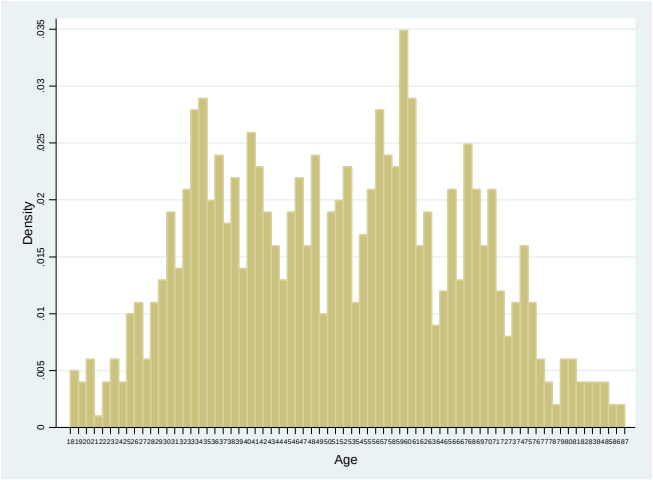


Figure 10: The probability of holding stocks and stock mutual funds, by age. Source: TNS 2007.

for all ages from the TNS 2007. Despite our sample size being substantially smaller and deploying a different elicitation methodology, the monotonic positive relationship between expectations and stockholdings conditional on gender and marital status first uncovered by them appears strikingly robust.[39]

Table 4: Probability of Holding Stocks or Stock Mutual Funds Conditional on Percent Chance of Positive Nominal Return, Gender, Age and Marital Status.

	Married or living with a partner				*NOT* Married or living with a partner			
Percent chance	Male		Female		Male		Female	
of positive	Point	Standard	Point	Standard	Point	Standard	Point	Standard
nominal return	Estimate	Error	Estimate	Error	Estimate	Error	Estimate	Error
0	0.31	(0.02)	0.28	(0.02)	0.25	(0.02)	0.22	(0.02)
1-10	0.39	(0.06)	0.35	(0.06)	0.33	(0.06)	0.29	(0.06)
11-20	0.40	(0.05)	0.36	(0.05)	0.34	(0.05)	0.30	(0.05)
21-30	0.28	(0.04)	0.25	(0.04)	0.23	(0.04)	0.20	(0.03)
31-40	0.48	(0.05)	0.44	(0.05)	0.42	(0.05)	0.37	(0.05)
41-49	0.41	(0.08)	0.37	(0.07)	0.35	(0.07)	0.30	(0.07)
50	0.42	(0.04)	0.38	(0.04)	0.35	(0.04)	0.31	(0.04)
51-59	0.24	(0.09)	0.20	(0.08)	0.19	(0.08)	0.16	(0.07)
60-69	0.49	(0.05)	0.45	(0.05)	0.42	(0.05)	0.38	(0.05)
70-79	0.56	(0.04)	0.52	(0.05)	0.49	(0.05)	0.44	(0.05)
80-89	0.55	(0.05)	0.50	(0.05)	0.48	(0.05)	0.44	(0.05)
90-99	0.52	(0.05)	0.47	(0.05)	0.45	(0.05)	0.40	(0.05)
100	0.48	(0.02)	0.44	(0.03)	0.42	(0.03)	0.37	(0.03)
All	0.40	(0.02)	0.38	(0.02)	0.34	(0.02)	0.32	(0.02)

Source: TNS 2007.

Since subjective expectations have been found to systematically vary with risk preferences, information, and demographic and socio-economic characteristics (Hurd, 2009), here we estimate the

[39] However, since the forecasting horizon is much longer (5 years instead of 1), the relationship appears less pronounced quantitatively. In addition, the probability of stockholding is much larger for those who anticipate a 0% chance of a positive return, and (also strangely) recedes as it increases beyond 80% (beyond 90% in Dominitz and Manski's findings, 2007).

conditional effect of the percentage chance of a positive nominal return on stockholdings. Conditioning on individual information is important for both theoretical and empirical reasons. Theoretically, households form their expectations conditioning upon their individual information sets, Θ^i. Although in the rational expectations tradition Θ^i would contain "all relevant information", here we proxy it by the individually elicited degree of knowledge of the most recent stock market return realization over the relevant horizon, $\Theta^i = \{R_t(5)\}$.[40] Although admittedly simplistic, minimal information survey elicitation provides results beyond individuals overweighting more recent stock market return realizations because of learning from experience (Linnainmaa, 2011; Malmendier and Nagel, 2011, 2013), while uncovering a novel age-dependent imperfect knowledge fact, broadly consistent with recent findings from the financial literacy literature (Lusardi *et al.*, 2011, 2012).

Empirically, Dominitz and Manski (2011) or Hurd and Rohwedder (2012) conjecture that differences in the way people discount publicly available information may explain much of the observed heterogeneity in subjective expectations. Our information measure, $pPNR^i$, precisely measures (probabilistically) the extent to which respondents know about the most recent realized stock market realization.[41] Most importantly, the availability of observable measures of both expectations and information sets allows us to examine whether households' behaviour is consistent with the theoretical prediction of elementary (static/dynamic) portfolio choice models without having to worry about the rationality of their stock market return forecasts.

Accordingly, we estimate households' probability of holding stocks $\Pr(\, s_t^i = 1 \big|\, p_{t+1}^i; p_t^i; \mathbf{x}_i)$ as a function of the percentage chance of a positive nominal return ($p_{t+1}^i \equiv PNR^i$),[42] conditioning on information ($p_t^i \equiv pPNR^i$), and a vector of observables $\mathbf{x}_i$, containing measures of (rate of) time and risk (aversion) preference, endowments (income and total wealth), household constraints (being

[40] In the empirical finance literature, it is typically assumed to be $\Theta^i = \{R_t(5), R_{t-1}(5), ..., R_0(5)\} = \Theta, \forall i$. In the rational expectations tradition, Θ^i *includes* also knowledge of the data generating process of stock market return realizations, as well as of the true economic model compatible with such data generating process. However, Guesnerie (1992) shows that common knowledge of rationality, of the model and of the data generating process are *not sufficient* to form a rational expectations within a game-theoretic epistemic formulation encompassing Muth's original model.

[41] Instead of attributing differences in information sets to differences in private information (e.g. amongst professional forecasters, as in Keane and Runkle, 1990), the previous section points towards rational inattention as being instrumental (e.g. Sims, 2003).

[42] To obtain subjective expectations from answers to the probability question (PNR^i), Dominitz and Manski (2007) show that if (i) stock market returns are normally distributed, with cdf. $\Phi(.)$:

$$PNR^i = \Pr{}^i \left[R_{t+1}(5) > 0\right] = \Pr{}^i \left[\frac{R_{t+1}(5) - \mu_i}{\sigma_i} > -\frac{\mu_i}{\sigma_i}\right] = 1 - \Phi\left(-\frac{\mu_i}{\sigma_i}\right)$$

and if (ii) a common variance is assumed, $\sigma_i = \sigma$ (for example, equal to the value obtained from historical records), then:

$$\mu_i = -\sigma\Phi^{-1}\left(1 - PNR^i\right)$$

meaning that respondents reporting a higher percentage chance that the stock market will increase over the next five years (PNR^i), have a higher subjective mean return expectation (μ_i), and should then be more likely to invest in the stock market. Since we inquire about a longer investment horizon, we exploited monthly data on the CAC 40 stock market index between July 1987 and March 2007 (228 observations) to compute the standard deviation of five-year log returns to be 0.188. When inserted into the above expression, the sample average percentage chance of a positive nominal return of 46 percent (reported in the appendix, Table 11) corresponds to a sample mean expected return of 0.019, about five times smaller than the historical sample mean of 0.109. A respondent reporting a value of $PNR^i = 70$ percent, would match the historical sample mean of 0.109.

liquidity constrained, access to online banking), demographics (age, gender and marital status) or inertial factors (who takes financial decisions, stocks in pay, parents own stocks or trust[43]) previously found in the literature to matter at the extensive margin:

$$\Pr(s_t^i = 1 \big| p_{t+1}^i, p_t^i; \mathbf{x}_i) = \mathbf{\Phi}(\delta_{t+1} p_{t+1}^i + \delta_t p_t^i + \boldsymbol{\delta}' \mathbf{x}_i) \tag{1}$$

where $\mathbf{\Phi}(.)$ denotes the standard normal cumulative distribution function, since we assume that there is a normally distributed unobserved error term e_t^i. Table 11 in the appendix reports descriptive statistics for the main variables, for the whole and selected samples.

Table 5 reports the marginal effects of the probit estimation, for all ages. The variables have the expected signs with minor differences across columns[44], confirming the robust effect of subjective expectations on the probability of holding stocks, *conditional on heterogeneity in individual information*. Column (5) reports that an increase of 1 percentage point (pp.) in the percent chance of a positive nominal return increases the probability of holding stocks by around 9 percent, corresponding to an increase of 28.3% in the unconditional probability (from 32.1% to 41.2%).[45] The effect of information is also sizable: comparing those who were certain that the stock market had (indeed) increased between March 2002 and March 2007 ($pPNR = 100$) with those who were certain that it had not ($pPNR = 0$), raises the probability of holding stocks by 7.5% (relative to the unconditional probability of 31.4%). Notice that the positive effect of information remains statistically significant once heterogeneity in either preferences, expectations, decision taking or constraints is taken into account, but that when omitted (column 6), it biases upwards the estimated effect of expectations on stockownership.

In the presence of transaction costs in capital markets, households' endowments (proxied by income and total wealth) influence portfolio choice.[46] The empirical analysis reveals that their effect is best captured by a second order polynomial, which facilitates the comparison with existing results in the literature (Calvet and Sodini, 2013; Guiso *et al.,* 2003; King and Leape, 1998).[47] Both have a positive effect on participation and are statistically significant at the 1 percent level. An increase of 100,000 euros in mean total wealth (233,757 euros) increases the probability of participation by 9 pps., corresponding to an increase of 26% in the unconditional probability, while an increase of 10,000 euros in mean income (19,634 euros) increases the probability of stockownership by 5 pps.

[43] Our measure of trust differs from Guiso et al.'s (2008) in that TNS survey question I25 inquires the respondent about 'the extent to which s/he trusts online payment systems'. Hence, and although our results are not directly comparable, our measure of trust captures still the extent to which respondents 'trust others'.

[44] Results in Table 5 refer to both direct and indirect stockownership. But the sign and magnitude of the reported estimates are robust to changes in the definition of stockownership (only direct stockholders, Table 9 below), and to a semi-log specification in income and total wealth (unreported, but available upon request).

[45] Although the effect appears too important quantitatively, it is consistent with the results reported by Hurd *et al.* (2011), Kézdi and Willis (2009) or Arrondel *et al.* (2013) who, instead of working with the subjective probability of a positive nominal return (PNR^i), introduce as regressors the first and second moments (μ^i, σ_i) of the individually elicited distributions, reported in Table 1.

[46] See King and Leape (1998) and simulated results by Cocco *et al.* (2005) or Haliassos and Michaelides (2003).

[47] In unreported regressions we replaced the second order polynomials in income and wealth for their quartiles/deciles with no significant change in the estimated coefficients of expectations and information. Results are available upon request.

Table 5: Probability of Holding Stocks or Stock Mutual Funds (All Ages)

	(1)	(2)	(3)	(4)	(5)	(6)	2SLS (7)
Positive nominal return (PNR)	0.00185***	0.00146***	0.00139***	0.00111***	0.000908***	0.00116***	0.00737***
	(0.000279)	(0.000299)	(0.000301)	(0.000310)	(0.000316)	(0.000295)	(0.002067)
Male	0.0112	0.00108	-0.000288	-0.0177	-0.0232	-0.0166	-0.0639
	(0.0220)	(0.0223)	(0.0224)	(0.0241)	(0.0248)	(0.0246)	(0.0657)
Married/living with a partner	0.0572**	0.0585**	0.0688***	-0.0182	-0.0177	-0.0189	-0.0497
	(0.0238)	(0.0238)	(0.0239)	(0.0267)	(0.0273)	(0.0273)	(0.0708)
Age	0.00789*	0.00711*	0.00603	-0.00585	0.000417	0.000944	0.00153
	(0.00424)	(0.00426)	(0.00433)	(0.00461)	(0.00472)	(0.00471)	(0.01215)
Age squared	-2.67e-05	-2.10e-05	1.34e-06	8.64e-05*	4.15e-05	3.82e-05	1.05e-04
	(4.24e-05)	(4.25e-05)	(4.34e-05)	(4.59e-05)	(4.68e-05)	(4.68e-05)	(1.05e-04)
Past positive nominal return (pPNR)		0.00115***	0.00105***	0.000853**	0.000756**		
		(0.000321)	(0.000323)	(0.000333)	(0.000338)		
Education							
(Ref. category: High school or less)							
Less than college			0.115***	0.0450	0.0358	0.0360	0.1120
			(0.0396)	(0.0417)	(0.0423)	(0.0423)	(0.1167)
College or more			0.256***	0.113**	0.0710	0.0727	0.1961
			(0.0438)	(0.0484)	(0.0494)	(0.0494)	(0.1299)
Income (10E-6)				7.204***	5.027**	5.071**	11.70*
				(2.300)	(2.455)	(2.456)	(6.323)
Income squared (10E-11)				-6.760**	-5.531*	-5.443*	-12.89
				(2.989)	(3.154)	(3.156)	(7.928)
Total wealth (10E-7)				12.94***	9.211***	9.271***	24.37***
				(2.035)	(2.103)	(2.103)	(5.5814)
Total wealth squared (10E-13)				-7.975**	-4.960	-5.062	-14.21
				(3.262)	(3.335)	(3.336)	(8.979)
Self account management					-0.00761	-0.00708	-.0279
					(0.0237)	(0.0237)	(0.631)
Risk aversion (CARA)					-0.160	-0.155	-0.3566
					(0.146)	(0.145)	(0.365)
Liqudity constrained					-0.133***	-0.134***	-0.345***
					(0.0243)	(0.0243)	(0.0684)
Firm shares in remuneration					0.0844	0.0859*	0.194
					(0.0521)	(0.0520)	(0.1311)
Temporal preference					0.0146***	0.0145***	0.0380***
					(0.00526)	(0.00525)	(0.0146)
Online banking					0.103***	0.105***	0.260***
					(0.0297)	(0.0297)	(0.0771)
Irregular income risk					-0.0107	-0.00907	-0.0286
					(0.0424)	(0.0424)	(0.1097)
Parents own stocks					0.161***	0.159***	0.384***
					(0.0265)	(0.0265)	(0.0700)
Trust					0.0363	0.0390	0.0911
					(0.0277)	(0.0276)	(0.0738)
Wald test of exogeneity: Chi-sq(1)							4.72
Pseudo R-squared/(P-value Chi-sq(1))	0.0426	0.0472	0.0639	0.1191	0.1601	0.1582	(0.0298)
Chi-squared	118.1	131.0	177.3	320.0	429.9	424.9	389.4
Log-likelihood	-1327	-1321	-1298	-1183	-1128	-1131	-11205
No of observations	2,066	2,066	2,016	2,016	2,016	2,016	2,016

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

(corresponding to an increase of 12.5 percent in the unconditional probability). Given that we condition on our measure of information, and that wealth does not increase the conditional odds of being better informed, the identified positive and statistically significant effects are consistent with the presence of fixed/transaction costs of accessing the stock market.

Although previous empirical studies also find that education increases the probability of participation, most of them interpret its effect as a proxy variable for information. The results in Table 5, column (5) reveal that, conditioning on our individual measure of the most recent past stock market performance, holding a college degree (or further) does not increase the probability of participation at a statistically significant level, relative to those who hold only a high school diploma or less. (Not even when we exclude it: column (6), Table 5).

Management variables also appear important: although those who take financial decisions by themselves are less likely to participate relative to those who totally or partially delegate in a financial advisor, the effect is not statistically significant.[48] On the other hand, those who manage their accounts online ('online banking') are around 11 percent more likely to be stock owners, and if respondents' parents are stock owners themselves ('parents own stocks'), they are 16 percent more likely to own stocks themselves. Since Table 2 reports that neither significantly increases the odds of being informed, we interpret these effects as inertial factors in stockownership.

Also, measures of preference heterogeneity are important and consistent with recent theoretical contributions. Individuals who have a long planning horizon (temporal preference) are 1.4 pps. more likely to participate than those who are impatient, in line with the empirical results obtained by Donkers and van Soest (1999) for The Netherlands. More risk averse individuals also have a lower probability of participation, although the effect is not statistically significant [49].[50] Overall, these effects are consistent with Van Nieuwerburgh and Veldkamp (2010), who model optimal information acquisition within the standard two-asset two-period portfolio choice model. They find that at the extensive margin, preference for an early resolution of uncertainty matters since it determines the optimal amount of information, while risk aversion does not. Intuitively, within an expected utility framework, risk aversion has a second order effect, affecting only the intensive margin, and subjective stock market return expectations are formed conditional on respondents' individual information sets.

Constraints are very significant, in line with existing empirical results in the literature. Households who have been liquidity constrained or who think that they will be so in the future are less

[48] In the next subsection, Table 8, we examine the conditional effect of conditional expectations on stockownership for respondents who partially or totally delegate financial decition taking to a professional.

[49] Since we have found evidence consistent with the presence of fixed/transaction costs to access the stock market, more risk averse respondents would find it less worth to pay them today. However, since within an expected utility framework, they are also more prudent, they are more likely to save more, and invest part of the additional savings in the stock market. Hence, the overall effect of risk aversion depends on the strength of the two opposing forces. See Haliassos and Michaelides (2003) for additional details.

[50] The results are robust to an alternative measure of risk aversion: the coefficient of relative risk aversion for preferences in the constant relative risk aversion class (CRRA), advanced by Barsky *et al.* (1997) and available in the TNS 2007 survey wave. In addition, Kimball *et al.* (2008) show that the CRRA measure is robust to survey measurement error. The results are available from the authors upon request.

likely to participate (around 14 percent less likely). Deaton (1992) explains how the expectation of being liquidity constrained in the future leads prudent households to save more ('buffer stock'), which results in an overall reduction in stock ownership for those households whose preferences display both decreasing risk aversion (DARA) and prudence (DAP) –Elmendorf and Kimball (2000).[51] Although income risk has a negative effect on stock ownership, in line with theoretical predictions (and simulations) that treat it as exogenous, the effect is not statistically significant.[52]

Finally, the age variables indicate that the probability of owning risky assets is lower for younger households, although it has a hump-shaped effect reaching its maximum at the age of 50. In Tables 6 and 7 below, we decompose the effect of expectations on stock market participation by age groups.

4 Inspecting the Mechanism

In Table 6 we report the maginal effects of the estimation for respondents in the same age bracket (50-80) as those exploited by Dominitz and Manski (2007) from the HRS 2004. Two main messages emerge: (i) Dominitz and Manski's (2007) results are robust to the inclusion of previously identified covariates in the empirical literature of household finance, like transaction costs (proxied by wealth), liquidity constraints, income risk, access to the internet or preferences. (ii) Subjective stock market return expectations of the elderly determine their financial investment decisions at the extensive margin. However, information regarding past stock market performance does not because, as Figure 9 illustrates, they are overall (similarly) well informed.

In Table 7, we complete the life-cycle picture and report the estimated marginal effects of the probit specification only for the young (18-49 age bracket). The main message conveyed is that although subjective expectations determine the decision to enter the stock market by the young *unconditionally*, once we condition on individual information sets, the effect of expectations becomes statistically insignificant.

To (i) further inquire into the conditional effect of subjective expectations on stockownership decisions amongst the young, and (ii) because the precise mechanism whereby differences in (ability to process, access to or in the actual stock of) financial information translate into differences in stock market participation remains elusive (Lusardi et al., 2011; Grinblatt *et al.*, 2011), we postulate a two-step econometric specification whereby information (p_t^i) affects the individual decision to hold stocks *only* through its effect on subjective expectations (p_{t+1}^i):

$$\left.\begin{array}{l} p_{t+1}^i = \alpha_t p_t^i + \alpha_O Opt_i + \boldsymbol{\alpha}'\mathbf{x}_i + u_{t+1} \\ \Pr(\, s_t^i = 1 \big|\, p_{t+1}^i; \mathbf{x}_i) = \boldsymbol{\Phi}(\delta_{t+1} p_{t+1}^i + \boldsymbol{\delta}'\mathbf{x}_i) \end{array}\right\} \tag{2}$$

[51] See also Gollier (2001), who proves that the willigness to take risk is reduced in the presence of a liquidity constraint if absolute risk tolerance is increasing and convex in wealth.

[52] Tables 6 and 7 in the next subsection reveal that there is an age-composition effect of income risk on stock ownership: it has a negative and significant effect amongst the elderly (for whom it is either non-diversifiable or non-existent), while a positive non-significant one amongst the young (for whom it is endogenous while non-diversifiable). See Arrondel and Calvo-Pardo (2012) for a model that rationalizes this conclusion.

Table 6: Probability of Holding Stocks or Stock Mutual Funds ($50 \leq$ Age ≤ 80)

							2SLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Positive nominal return (PNR)	0.00194***	0.00179***	0.00179***	0.00144***	0.00134***	0.00143***	0.0055*
	(0.00043)	(0.000457)	(0.000459)	(0.000481)	(0.000494)	(0.000465)	(0.0033)
Male	0.0195	0.0144	-8.16E-05	-0.0328	-0.0399	-0.0366	-0.0994
	(0.0359)	(0.0363)	(0.0366)	(0.0404)	(0.0419)	(0.0415)	(0.1057)
Married/living with a partner	0.0542	0.0542	0.0685*	-0.0117	-0.0110	-0.0113	-0.0445
	(0.0392)	(0.0393)	(0.0396)	(0.0441)	(0.0455)	(0.0455)	(0.1153)
Age	0.0763**	0.0758**	0.0840**	0.0647*	0.0719*	0.0719*	0.1851*
	(0.0352)	(0.0353)	(0.0356)	(0.037)	(0.0380)	(0.0380)	(0.0949)
Age squared	-0.000567**	-0.000564**	-0.000620**	-0.000472	-0.000528*	-0.000528*	-0.00136*
	(0.000279)	(0.000279)	(0.000282)	(0.000292)	(0.000300)	(0.000300)	(0.00074)
Past positive nominal return (pPNR)		0.0005	0.000507	0.000304	0.000309		
		(0.000503)	(0.000505)	(0.000529)	(0.000547)		
Education							
(Ref. category: High school or less)							
Less than college			0.135***	0.0799	0.0741	0.0743	0.1964
			(0.0486)	(0.0513)	(0.0525)	(0.0525)	(0.1390)
College or more			0.204***	0.0568	0.0362	0.0356	0.0993
			(0.0548)	(0.0645)	(0.0678)	(0.0678)	(0.1705)
Income (10E-6)				8.739**	5.440	5.418	12.551
				(3.905)	(4.156)	(4.152)	(9.485)
Income squared (10E-11)				-8.942*	-6.387	-6.305	-14.790
				(5.426)	(5.688)	(5.674)	(11.25)
Total wealth (10E-7)				14.62***	12.38***	12.37***	31.19***
				(3.37)	(3.491)	(3.491)	(8.742)
Total wealth squared (10E-13)				-10.73**	-9.604*	-9.597*	-24.49*
				(5.048)	(5.200)	(5.200)	(13.24)
Self account management					-0.0579	-0.0580	-0.1549
					(0.0388)	(0.0388)	(0.1006)
Risk aversion (CARA)					-0.0327	-0.0339	-0.0681
					(0.209)	(0.209)	(0.503))
Liquidity constrained					-0.147***	-0.147***	-0.371***
					(0.0415)	(0.0415)	(0.1090)
Firm shares in remuneration					-0.0209	-0.0193	-0.0450
					(0.0990)	(0.0989)	(0.239)
Temporal preference					0.0158*	0.0161*	0.0409*
					(0.00909)	(0.00907)	(0.0235)
Online banking					0.174***	0.175***	0.436***
					(0.0538)	(0.0537)	(0.1404)
Irregular income risk					-0.153*	-0.153*	-0.401*
					(0.0838)	(0.0838)	(0.2263)
Parents own stocks					0.182***	0.180***	0.444***
					(0.0445)	(0.0445)	(0.1142)
Trust					0.0102	0.0128	0.0242
					(0.0431)	(0.0428)	(0.1123)
Wald test of exogeneity: Chi-sq(1)							0.37
Pseudo R-squared (P-value Chi-sq(1))	0.0298	0.0306	0.042	0.0989	0.1438	0.1435	(0.542)
Chi-squared	34.89	35.88	49.23	110.9	161.2	160.9	137.16
Log-likelihood	-568.6	-568.1	-561.4	-505.1	-480.0	-480.1	-4579.2
No of Observations	847	847	813	813	813	813	813

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

Table 7: Probability of Holding Stocks or Stock Mutual Funds (Age< 50)

	(1)	(2)	(3)	(4)	(5)	(6)	2SLS (7)
Positive nominal return (PNR)	0.00177***	0.00118***	0.00106***	0.000750*	0.000482	0.000890**	0.00956***
	(0.000365)	(0.000392)	(0.000393)	(0.000403)	(0.000408)	(0.000382)	(0.00270)
Male	0.0124	0.000113	0.0115	0.000977	-0.00559	0.00338	-0.0226
	(0.0277)	(0.0279)	(0.0281)	(0.0302)	(0.0308)	(0.0306)	(0.0877)
Married/living with a partner	0.0322	0.0370	0.0491	-0.0379	-0.0322	-0.0352	-0.0497
	(0.031)	(0.031)	(0.031)	(0.0348)	(0.0356)	(0.0356)	(0.0998)
Age	0.0453***	0.0436***	0.0335**	2.45E-02	0.0292*	0.0296*	0.0790*
	(0.0146)	(0.0146)	(0.015)	(0.0158)	(0.0163)	(0.0163)	(0.0461)
Age squared	-0.000584***	-0.000573***	-0.000420**	-0.000371*	-0.000388*	-0.000386*	-0.00107*
	(0.000206)	(0.000207)	(0.000212)	(0.000223)	(0.000230)	(0.000229)	(0.00064)
Past positive nominal return (pPNR)		0.00165***	0.00143***	0.00122***	0.00119***		
		(0.000411)	(0.000414)	(0.000423)	(0.000428)		
Education							
(Ref. category: High school or less)							
Less than college			0.168*	0.0892	0.0421	0.0439	0.132
			(0.0874)	(0.0909)	(0.0932)	(0.0928)	(0.2943)
College or more			0.346***	0.197*	0.107	0.115	0.295
			(0.0988)	(0.103)	(0.103)	(0.103)	(0.303)
Income (10E-6)				5.721*	4.430	4.521	11.54
				(3.052)	(3.225)	(3.224)	(9.143)
Income squared (10E-11)				-5.042	-5.043	-4.892	-13.36
				(3.973)	(4.117)	(4.116)	(11.80)
Total wealth (10E-7)				11.16***	6.765**	6.987***	18.88**
				(2.556)	(2.644)	(2.645)	(7.533)
Total wealth squared (10E-13)				-4.845	-1.088	-1.376	-5.229
				(4.472)	(4.575)	(4.589)	(12.87)
Self account management					0.0280	0.0298	0.0800
					(0.0296)	(0.0296)	(0.0823)
Risk aversion (CARA)					-0.234	-0.201	-0.530
					(0.214)	(0.212)	(0.581)
Liquidity constrained					-0.133***	-0.134***	-0.352***
					(0.0298)	(0.0298)	(0.0910)
Firm shares in remuneration					0.116*	0.116*	0.243
					(0.0611)	(0.0610)	(0.1588)
Temporal preference					0.0147**	0.0137**	0.038**
					(0.00643)	(0.00640)	(0.0193)
Online banking					0.0563*	0.0595*	0.1436
					(0.0336)	(0.0337)	(0.0948)
Irregular income risk					0.0407	0.0428	0.1134
					(0.0482)	(0.0482)	(0.1256)
Parents own stocks					0.144***	0.144***	0.3417***
					(0.0319)	(0.0318)	(0.0904)
Trust					0.0557	0.0547	0.1498
					(0.0350)	(0.0350)	(0.1025)
Wald test of exogeneity: Chi-sq (1)							6.93
Pseudo R-squared/(P-value Chi-sq (1))	0.0315	0.0422	0.657	0.1247	0.1710	0.1658	(0.0085)
Chi-squared	47.65	63.78	99.36	185.4	254.2	246.5	241.21
Log-likelihood	-732.4	-724.3	-706.5	-650.8	-616.4	-620.3	-6439.61
No of Observations	1,188	1,188	1,174	1,174	1,174	1,174	1,174

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

where Opt_i denotes the 'optimism' of the respondent, treated as a fixed personality trait –see Table 4. (u_{t+1}, e_t) are assumed to follow a zero mean, bivariate normal distribution, independent of $(\mathbf{x}_i, Opt_i)$.

The two-step econometric specification (2) is more efficiently estimated jointly by maximum likelihood. It effectively assumes that "*agents solve the investment decision problem in two stages: derivation of the vector of conditional expected returns, and choice of an optimal portfolio of assets using estimated expected returns*" (Separation Theorem in Gennotte, 1986).[53] The advantage of the two-step estimation strategy is that it allows us to consider individual information as an instrumental variable for subjective expectations and test its statistical validity, in the sense that the exclusion restriction requires that information has no direct effect on stockownership other than through expectations. Table 7, column (7) reports statistical evidence in support of a causal effect of subjective expectations: they have a stastitically significant and quantitatively important positive effect on stockownership decisions amongst the young. The Chi-square statistic of 6.93 reported at the bottom, has an associated P-value of 0.0085, confirming that there is statistical evidence in support of the null hypothesis of exogeneity.

Similarly, below column (7) in Tables 5 and 6, we report the maximum likelihood coefficient estimates of specification (2) for all ages and the elderly, respectively. For all, the reported quantitative effect corresponds to a 115.1 percent increase in the probability of holding stocks, when comparing the conditional probability of detention amongst those who are certain that the stock market will not increase (23.9%) with the probability of those who are sure that it will increase over the next 5 years (51,4%). The Chi-square statistic of 4.72 reported at the bottom of Table 5 column (7), has an associated P-value of 0.0298, confirming that there is weaker statistical evidence in support of the null hypothesis of exogeneity for all ages than there is amongst the young. The Amemiya-Lee-Newey Chi-squared statistic for overidentifying restrictions with 1 degree of freedom (when optimism is interpreted as a valid additional instrument; Kézdi and Willis, 2008) reports a value of 1.014, with an associated P-value of 0.3140, confirming the validity of information as an instrument, conditional on optmism being also a valid instrument. However, and in contrast to the young, we are unable to reject the null of exogeneity of subjective expectations on the stockownership decisions amongst the elderly, and hence the results under specification (1) remain valid (and causal) for them.

To further inquire into the causal effect of conditional sujective expectations on stockownership, we conduct a counter-factual test, reported in Table 8 by age groups. The results show that respondents' conditional expectations who have totally or partially delegated financial decision taking to a professional (i.e. those who have signed a 'mandat de gestion'), do *not* (consistenly) determine their stockownership decisions. In unreported regressions, we further inquired into the robustness of our findings to unobserved state dependence, measured by the number of stock market

[53] The crucial assumption for the separation between the estimation of future stock market returns and the selection of the optimal portfolio to hold is that the distribution of realized returns does not depend on the level of investment. Otherwise, the quality of information depends on how much was invested in the past, and hence on past return realizations. See Genotte (1986), pp. 742-3, for further details.

Table 8: Probability of Holding Stocks or Stock Mutual Funds: 2SLS by Age and Financial Advisor

	All Ages	$50 \leq$ Age ≤ 80	Age < 50
Positive nominal return (PNR)	0.007	0.0094	0.00782
	(0.0499)	(0.0102)	(0.00562)
Male	-0.211	-0.2997	-0.107
	(0.132)	(0.2051)	(0.188)
Married/living with a partner	0.0451	0.0525	-0.0546
	(0.127)	(0.2104)	(0.181)
Age	0.0373	0.4569**	0.201*
	(0.0248)	(0.1856)	(0.115)
Age squared	-0.00016	-0.0035**	-0.00247
	(0.00024)	(0.0015)	(0.00156)
Past positive nominal return (pPNR)	-	-	-
Education			
(Ref. category: High school or less)			
Less than college	0.120	0.0197	0.576
	(0.228)	(0.3356)	(0.477)
College or more	0.252	0.2545	0.629
	(0.246)	(0.3636)	(0.495)
Income (10E-6)	23.45*	-35.15	11.93
	(13.44)	(25.16)	(21.85)
Income squared (10E-11)	-32.10*	-35.147	-22.47
	(18.02)	(25.16)	(28.38)
Total wealth (10E-7)	18.26*	38.94**	2.777
	(10.143)	((15.238))	(15.54)
Total wealth squared (10E-13)	-9.40	-40.85*	23.08
	(15.36)	(22.934)	(25.97)
Self account management	-	-	-
Risk aversion (CARA)	-1.318	-2.219*	-0.514
	(0.952)	(1.286)	(1.999)
Liquidity constrained	-0.293**	-0.3688	-0.296*
	(0.1304)	(0.2353))	(0.170)
Firm shares in remuneration	0.325	0.203	0.388
	(0.280)	(0.6354)	(0.300)
Temporal preference	0.0566**	-0.0021	0.101**
	(0.0284)	(0.047))	(0.0396)
Online banking	-0.0935	-0.2615	-0.0353
	(0.154)	(0.2978)	(0.188)
Irregular income risk	-0.349	-1.164***	-0.0359
	(0.218)	(0.3973)	(0.266)
Parents own stocks	0.537***	0.3709*	0.583***
	(0.133)	(0.2114)	(0.175)
Trust	-0.0048	0.0495	-0.148
	(0.142)	(0.2122)	(0.211)
Wald test of exogeneity: Chi-sq (1)	0.11	0.01	1.14
P-value Chi-sq (1)	0.7430	0.9120	0.2866
Chi-squared	117.39	60.0	55.44
Log-likelihood	-3156	-1461.67	-1589.11
No of Observations	570	265	290

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

operations carried over the previous year. Amongst respondents that have *not* recently traded in the stockmarket (inertial non-/stockholders), we find that their conditional expectations do *not* (consistently) determine their stock ownership decisions.[54]

Figure 11 reports kernel-smoothed estimates of the mean reported realized nominal return conditional on age and stockownership status: stockholders appear better informed than non-stockholders between the late 20s and the early 50s, being statistically indistinguishable at earlier and later stages. The displayed patterns are not completely consistent with "learning from experience" (Linnainmaa, 2011; Malmendier and Nagel, 2011, 2013), whereby young stockholders should be much better informed about the most recent stockmarket return realization than old stockholders, while the opposite holds in Figure 11. As well, they are difficult to reconcile with a pure life-cycle "financial literacy" argument (Lusardi *et al.*, 2012), since even non-stockholders seem to accumulate relevant financial information over their life-cycles (although they start at a later stage than stockholders).

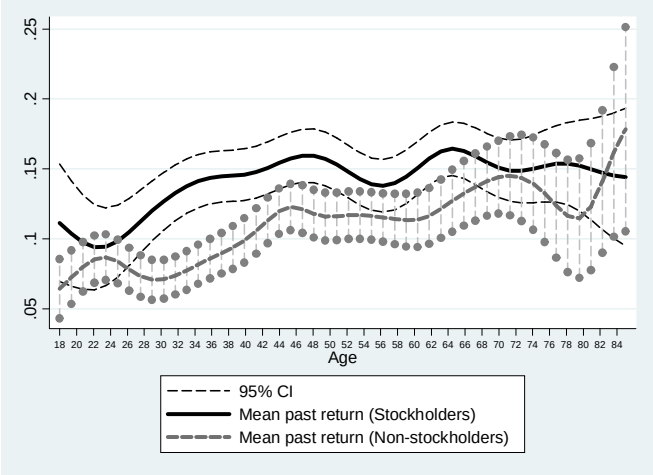


Figure 11: Information (mean realized return) by age, for stockholders (black) and non-stockholders (grey). Source: TNS 2007.

These results, together with Figure 9, lend support to King and Leape (1987) and Hurd's (2009) conjectures: what determines stock market participation amongst the young is their degree of awareness regarding the investment opportunities the stock market offers, and we add *'through* their impact on subjective stock market return expectations'.[55]

Table 9 reports estimated marginal effects for the same specification as in Table 5, but on a narrower definition of stockholdings (direct stockholders only), by age groups (18-49, 50-80 and all

[54] The results are available from the authors upon request.

[55] In the not for publication appendix A, we extract the information component of expectations using a simple linear regression. The predicted OLS residual is the conditional probability of a positive nominal return over the next five years. It is plotted in Figure 1 in the not for publication appendix A, which broadly displays a hump-shaped pattern by age. In unreported estimations, we replaced PNR and pPNR in the main probit specification by the conditional probability thus computed. It has a significant positive effect on stockholdings amongst the young, consistent with the expectations of the informed being instrumental for the decision to own stocks and against a persistence interpretation. These results are available upon request.

ages). Since mutual funds are professionally managed, narrowing the definition to exclude them should lead to respondents' decisions being more responsive to their expectations. For each age group, we examine the within effect that conditioning on information (*even* numbered columns) has on the unconditional effect (*odd* numbered columns) of expected stock market performance on the decision to invest in shares of national and foreign firms. For all age groups, the main message remains although quantitatively the estimated coefficients are roughly similar. Importantly, comparison of the last two columns of the table lends support to Dominitz and Manski's (2011) conjecture about the importance of heterogeneity in the amount of public information when accounting for heterogeneity in subjective expectations: the upwards bias in the unconditional marginal effect of expectations on decisions (all ages, column (5)) due to the effect of information on expectations (all ages, column (5)) is larger with a narrower definition of stock ownership.[56]

Finally, Table 10 reports the estimated marginal effects by respondents' (top) quartile of total wealth (under columns (3) and (4)), and by respondents' information (last two columns). For each subgroup, we examine the effect of conditioning on information (even numbered columns) on the unconditional effect (odd-numbered columns) of expectations on the decision to own stocks (directly and indirectly). Amongst those who are in the top quartile of the total wealth distribution (stock of total wealth above €236,000), the conditional marginal effect of expectations on stockownership roughly doubles in size relative to those in the lower quartiles. Hence, the effect of conditional subjective expectations is strenghtened amongst the rich, contrary to the effect of participation/transaction costs. Also, and although our measure of trust is not comparable to Guiso *et al.* (2008), its positive effect on stockownership becomes statistically insignificant for the wealthier. Therefore, when accounting for non-participation amongst the most affluent, heterogeneity in conditional return expectations appears as a different alternative to better private investment opportunities (Heaton and Lucas, 2000) or taxation (Poterba, 2002), while encompassing the ability to process it (Grinblatt *et al.*, 2011), its sources (social interactions, Hong *et al.*, 2004; trust, Guiso *et al.*, 2008) or more broadly, financial literacy (van Rooij *et al.*, 2011) through our novel information measure.

Comparing columns (5) and (6) of Table 10, notice that even amongst the uninformed, the effect of expectations on stock market participation is quantitatively very important and statistically very significant (column 2, $pPNR < 100$), in line with elementary portfolio choice predictions (Arrow, 1965; Merton, 1969) and with extensions to incorporate incomplete information (Merton, 1987; O'Hara, 2003; Biais *et al.*, 2010). Compared to the informed (column 1, $pPNR = 100$), expectations matter even more.[57] We rationalize this last finding as follows: if information affects decisions

[56] Qualitatively similar conclusions follow when we examine the conditional effect of expectations on decisions by financial decision taker: conditioning on information, the effect of expectations on decisions is quantitatively more important for those who delegate than for those who do not. In addition, and consistent with Table 4 results, we find that (i) amongst those who do not delegate financial decisions partially or totally, males are about 9% more likely to be stock owners than females, and about 11% less likely amongst those who delegate. Similarly, (ii) having access to online banking increases the probability of owning stocks by 12% amongst those who do not delegate, while it has no effect amongst those who delegate. These results are unreported to save on space, but are available upon request.

[57] This effect appears very robust to relaxing the definition of "informed". For example, when classifying respondents as informed if "pPNR>50" (those who gave more than a 50-50 percent chance of the stock market index going up

Table 9: Probability of Holding Stocks (Only)

	Age< 50		50 ≤ Age ≤ 80		All ages	
	(1)	(2)	(3)	(4)	(5)	(6)
Positive nominal return (PNR)	0.000658**	0.000305	0.000962***	0.000656	0.000829***	0.000525**
	(0.000261)	(0.000307)	(0.000347)	(0.000415)	(0.000208)	(0.000246)
Male	-0.001852	-0.00843	0.00722	0.000795	0.00146	-0.00338
	(0.0208)	(0.0230)	(0.0309)	(0.0355)	(0.0173)	(0.0194)
Married/living with a partner	-0.0109	-0.0147	-0.0136	-0.00595	-0.0131	-0.0116
	(0.0244)	(0.0269)	(0.0342)	(0.0388)	(0.0196)	(0.0217)
Age	-0.00802	-0.00479	0.0177	0.00344	0.000878	-0.000246
	(0.0110)	(0.0122)	(0.0286)	(0.0319)	(0.00337)	(0.00374)
Age squared	0.000116	6.08e-05	-0.000131	-1.81e-05	1.27e-05	2.42e-05
	(0.000155)	(0.000171)	(0.000225)	(0.000252)	(3.32e-05)	(3.67e-05)
Past positive nominal return (pPNR)		0.00105***		0.000507		0.000799***
		(0.000328)		(0.000461)		(0.000268)
Education						
(Ref. category: High school or less)						
Less than college	0.150*	0.146*	0.0368	0.0408	0.0313	0.0349
	(0.0780)	(0.0880)	(0.0408)	(0.0468)	(0.0305)	(0.0346)
College or more	0.221	0.200	0.0393	0.0447	0.0445	0.0453
	(0.135)	(0.139)	(0.0539)	(0.0609)	(0.0375)	(0.0418)
Income (10E-6)	3.546*	2.988	5.891*	6.107*	4.062**	3.932**
	(2.096)	(2.352)	(3.287)	(3.596)	(1.658)	(1.861)
Income squared (10E-11)	-1.077	-1.454	-6.512	-6.691	-2.674	-3.125
	(2.553)	(2.910)	(4.540)	(4.834)	(2.010)	(2.282)
Total wealth (10E-7)	8.305***	7.349***	9.623***	10.69***	8.517***	8.393***
	(1.713)	(1.891)	(2.572)	(2.961)	(1.443)	(1.616)
Total wealth squared (10E-13)	-9.175***	-6.987**	-8.716**	-10.06**	-8.128***	-7.541***
	(2.739)	(3.028)	(3.788)	(4.280)	(2.196)	(2.430)
Self account management	0.0139	0.00338	0.0373	0.0329	0.0223	0.0159
	(0.0202)	(0.0221)	(0.0294)	(0.0330)	(0.0169)	(0.0187)
Risk aversion (CARA)	-0.0153	-0.207	0.0759	0.0527	0.0319	-0.0535
	(0.134)	(0.157)	(0.154)	(0.184)	(0.0982)	(0.115)
Liquidity constrained	-0.0600***	-0.0582**	-0.0954***	-0.111***	-0.0784***	-0.0817***
	(0.0211)	(0.0230)	(0.0304)	(0.0344)	(0.0174)	(0.0194)
Firm shares in remuneration	0.0341	0.0398	-0.0111	-0.0196	0.0285	0.0339
	(0.0405)	(0.0452)	(0.0744)	(0.0798)	(0.0375)	(0.0416)
Temporal preference	0.00165	0.00257	0.0107	0.0116	0.00552	0.00585
	(0.00438)	(0.00486)	(0.00692)	(0.00812)	(0.00378)	(0.00428)
Online banking	0.0908***	0.0919***	0.238***	0.254***	0.142***	0.147***
	(0.0243)	(0.0262)	(0.0462)	(0.0497)	(0.0230)	(0.0249)
Irregular income risk	0.0234	0.0119	-0.00830	-0.00461	0.0226	0.0179
	(0.0343)	(0.0366)	(0.0693)	(0.0792)	(0.0327)	(0.0360)
Parents own stocks	0.104***	0.112***	0.132***	0.140***	0.115***	0.123***
	(0.0232)	(0.0250)	(0.0371)	(0.0407)	(0.0203)	(0.0221)
Pseudo R-squared	0.1765	0.1742	0.1649	0.1671	0.1705	0.1698
Chi-squared	230.3	200.3	180.7	159.9	418.3	366.9
Log-likelihood	-537.2	-474.9	-457.6	-398.6	-1017	-896.8
No of Observations	1,355	1,174	962	813	2,349	2,016

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

Table 10: Probability of Holding Stocks or Stock Mutual Funds by Wealth and Information

	Wealth<75^{th} percentile		Wealth≥75^{th} percentile		pPNR=100	pPNR< 100
	(1)	(2)	(3)	(4)	(5)	(6)
Positive nominal return (PNR)	0.000937***	0.000663*	0.00131**	0.00115**	0.000613	0.00115***
	(0.000316)	(0.000342)	(0.000532)	(0.000557)	(0.000429)	(0.000431)
Male	-0.000865	-0.00746	-0.0437	-0.0498	0.0288	-0.0699**
	(0.0261)	(0.0263)	(0.0452)	(0.0456)	(0.0391)	(0.0305)
Married/living with a partner	-0.0419	-0.0393	0.0708	0.0673	-0.0958**	0.0369
	(0.0276)	(0.0276)	(0.0592)	(0.0593)	(0.0434)	(0.0330)
Age	-0.000549	-0.00125	0.000374	0.000566	-0.00585	0.00285
	(0.00481)	(0.00482)	(0.0105)	(0.0105)	(0.00755)	(0.00592)
Age squared	4.79e-05	3.82e-05	3.57e-05	-4.95e-07	8.14e-05	2.72e-05
	(4.82e-05)	(4.82e-05)	(0.000101)	(0.000101)	(7.35e-05)	(5.94e-05)
Past positive nominal return (pPNR)		0.000745**		0.000593		
		(0.000364)		(0.000604)		
Education						
(Ref. category: High school or less)						
Less than college	0.0255	0.0231	0.0556	0.0600	0.0111	0.0412
	(0.0431)	(0.0432)	(0.0837)	(0.0836)	(0.0665)	(0.0530)
College or more	0.0477	0.0451	0.117	0.118	-0.00455	0.115*
	(0.0528)	(0.0528)	(0.0878)	(0.0876)	(0.0748)	(0.0657)
Income (10E-6)	2.814	2.530	5.075	5.254	5.934*	6.328
	(3.037)	(3.053)	(4.210)	(4.233)	(3.548)	(4.359)
Income squared (10E-11)	0.0822	0.405	-7.567	-7.895	-5.938	-10.38
	(4.601)	(4.647)	(5.128)	(5.174)	(4.074)	(7.697)
Total wealth (10E-7)	24.22***	24.30***	20.92**	20.42**	11.24***	7.831***
	(6.584)	(6.584)	(10.17)	(10.18)	(3.313)	(2.623)
Total wealth squared (10E-13)	-87.57***	-87.83***	-17.05	-16.48	-5.845	-4.784
	(29.13)	(29.11)	(11.00)	(11.01)	(5.305)	(4.138)
Risk aversion (CARA)	-0.0333	-0.0357	-0.375	-0.373	-0.431*	-0.0244
	(0.151)	(0.152)	(0.299)	(0.299)	(0.238)	(0.181)
Liquidity constrained	-0.145***	-0.145***	-0.0448	-0.0428	-0.140***	-0.125***
	((0.0254)	(0.0254)	(0.0498)	(0.0499)	(0.0396)	(0.0299)
Firm shares in remuneration	0.0776	0.0764	0.0811	0.0785	0.0727	0.0727
	(0.0570)	(0.0570)	(0.0882)	(0.0888)	(0.0736)	(0.0723)
Temporal preference	0.0104*	0.0104*	0.0194*	0.0192*	0.0230***	0.00922
	(0.00533)	(0.00533)	(0.0108)	(0.0109)	(0.00836)	(0.00651)
Online banking	0.104***	0.103***	0.0847*	0.0824*	0.0996**	0.0923**
	(0.0345)	(0.0345)	(0.0475)	(0.0477)	(0.0438)	(0.0396)
Irregular income risk	-0.0276	-0.0307	0.0104	0.0116	-0.00311	-0.00983
	(0.0437)	(0.0435)	(0.0789)	(0.0790)	(0.0699)	(0.0514)
Parents own stocks	0.134***	0.134***	0.171***	0.173***	0.182***	0.140***
	(0.0312)	(0.0312)	(0.0422)	(0.0423)	(0.0396)	(0.0348)
Self account management	-0.0174	-0.0173	0.0111	0.00953	0.0202	-0.0312
	(0.0254)	(0.0254)	(0.0421)	(0.0421)	(0.0374)	(0.0294)
Trust	0.0527*	0.0495*	0.00830	0.00669	0.0223	0.0492
	(0.0291)	(0.0292)	(0.0511)	(0.0512)	(0.0433)	(0.0346)
Pseudo R-squared	0.1130	0.0889	0.0926	0.0938	0.1500	0.1602
Chi-squared	199.8	204.0	76.42	77.38	184.8	224.6
Log-likelihood	-742.9	-740.8	-374.4	-373.9	-523.6	-588.7
No of Observations	1,402	1,402	614	614	891	1,124

Note: (i) Robust standard errors in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

only through expectations, in accordance with specification (2), within the most homogeneous information group one would expect expectations to differ less, and hence to be less important empirically in accounting for differences in investment decisions. In the limit, perfectly informed individuals would all form rational expectations, and cross-sectional differences in decisions could not be explained by differences in expectations, as implicitly assumed in the literature until recently. Dominitz and Manski (2011) report evidence consistent with persistence in the modal type of revision of expectations with new public information. Hence, if the informed process information similarly, they will revise their expectations similarly, and there will be less cross-sectional variation in their expectations than there will be amongst the uninformed.

5 Conclusion

Elementary static (Arrow, 1965) and dynamic (Merton, 1969; Samuelson, 1969) models of portfolio choice put emphasis on the importance of individuals' expectations to explain stock market participation. However, it has been pervasive in the empirical literature on household portfolios to adopt the rational expectations assumption, thus neglecting a potential source of heterogeneity that, in addition to heterogeneity in preferences, endowments and constraints, could help reconcile economic theory predictions with empirically observed low participation rates.

In line with some recent efforts in the literature (Dominitz and Manski, 2007; Hurd *et al.*, 2011; Kézdi and Willis, 2009; Miniaci and Pastorello, 2010), here we have collected novel data on households' expectations and, for the first time, on households' information sets (TNS 2007). To validate our novel data set, and for comparison purposes, we have adopted Dominitz and Manski's (2007) methodology. Our results confirm that the novel effect of expectations on the decision to participate in the stock market first identified by them amongst the elderly, is robust to the inclusion of measures of information, risk and time preference, endowments and constraints. Crucially, it extends to a representative sample by age and wealth. Most of the included factors had been previously identified in the literature as important determinants of age-portfolio profiles at the extensive margin. Similar conclusions ground the works by Hurd *et al.* (2011), Kézdi and Willis (2009) or Miniaci and Pastorello (2010) on either selected samples or with a different set of covariates.

But, taking advantage of our novel information measure, we are able to empirically (i) confirm King and Leape (1987) - Hurd's (2009) conjecture for the first time, on the importance of being aware of the investment opportunities offered by the stock market (the 'investor recognition hypothesis' of Merton, 1987), specially to account for low participation rates of the young; (ii) confirm the basic principle of elementary portfolio choice models, since even amongst the uninformed and the affluent, subjective stock market expectations determine their decision to participate (Arrow, 1965; Merton, 1969; Samuelson, 1969), and when individually elicited information is exploited as an instrumental variable, we identify a causal effect of expectations on stock market participation

over the last 5 years), and uninformed otherwise, the marginal effect of expectations on decisions for the informed was 0.0021, whereas for the uninformed was 0.0057.

conditional on wealth, income, and measures of attitudes, preferences and socio-economic and demographic characteristics; (iii) confirm the effect of social interactions, professional advice, past own experience and specialised media access, as relevant sources of information at the individual level, thereby contributing to the literature on financial literacy, trust and adaptive expectations. However, (iv) although males appear better informed and are also more likely to take financial decisions by themselves instead of delegating; and conditioning on gender, optimists appear less well informed, at odds with a behavioural overconfidence bias interpretation (Guiso and Jappelli, 2006), but in line with inertia amongst the general population (Bilias *et al.*, 2010). Finally, (v) poor or constrained households, for whom stock market information is useless, appear less well informed in line with rational inattention theory (Sims, 2003).

Our results suggest novel and important policy implications, that deepen and complement those already identified: (i) in the financial literacy literature (e.g. Christelis et al., 2010; van Rooij et al., 2011), i.e. absence of financial specific knowledge is pervasive amongst the young, irrespective of their level of educational attainement, with potentially dramatic consequences for wealth accumulation over the life-cycle and hence, for wealth inequality (Lusardi et al., 2012); (ii) in macroeconomics, i.e. publicly available information appears costly to acquire in terms of money, effort or time, explaining information stickiness and hence stickiness in expectations (e.g. Carroll, 2003); (iii) in finance, i.e. recent contributions examining the asset pricing consequences of heterogeneous return expectations, since expectational heterogeneity apears to be fuelled by heterogeneity in public information gathering/processing (e.g. Allen, Morris and Shin, 2006).

However, many questions remain that would require further data collection and analysis.[58] Perhaps the most important one is that much observed heterogeneity remains unexplained at the extensive margin. In light of our results and the recent crisis, attempts to understand Dominitz and Manski's (2011) conjecture about heterogeneity in processing public information feeding heterogeneity in subjective expectations, is likely to be the most promising and challenging one (Frydman and Phelps, 2013). In that respect, we welcome the recent launch of the Survey of Consumer Expectations by the Federal Reserve Bank of New York[59], and hope that the results reported convey the complementary usefulness of including also retrospective probabilistic assessments of public information events.

[58] Data collection is crucial if one aims at understanding the macroeconomic implications and derive policy reccomendations using realistically calibrated dynamic macroeconomic models (ex. Alan, 2010).

[59] For further information, see http://libertystreeteconomics.newyorkfed.org/2013/12/introducing-the-frbny-survey-of-consumer-expectations-survey-goals-design-and-content.html

6 References

Alan, S. (2012), 'Do Disaster Expectations Explain Household Portfolios?', *Quantitative Economics* 3(1): 1-28.

Allen, F., S. Morris and H. S. Shin (2006), 'Beauty Contests and Iterated Expectations in Asset Markets,' *Review of Financial Studies* 19(3): 719-52.

Ameriks, J. and S.P. Zeldes (2004), 'How Do Household Portfolios Vary with Age?,' mimeo.

Arrondel, L., and H. Calvo-Pardo (2012), 'Endogenous Non-Tradable Earnings, and Households' Risky Assets Demand?', mimeo.

Arrondel, L., Calvo-Pardo, H., and X. Oliver (2013), 'Subjective Stock Market Return Expectations, Information and Portfolio Choice', mimeo.

Arrow, K. J. (1965). *Aspects of the Theory of Risk Bearing*, Yrjo Jahnsson Lectures, The Academic Book Store, Helsinki.

Arthur, W. B. (2006), 'Out-of-Equilibrium Economics and Agent-Based Modeling', in K. Judd and L. Tesfatsion, eds, *Handbook of Computational Economics*, Vol. 2: Agent-Based Com-putational Economics, Elsevier/North-Holland Publishers.

Barsky, R. B., Juster, T. F., Kimball, M. S. and Shapiro, M. D. (1997), 'Preference Parameters and Behavioral Heterogeneity: an Experimental Approach in the Health and Retirement Study,' *Quarterly Journal of Economics*, CXII: 537-580.

Benzoni, L., P. Collin-Dufresne and R. S. Goldstein (2007), 'Portfolio Choice over the Life-Cycle when the Stock and Labor Markets are Cointegrated,' *The Journal of Finance* 62(5): 2123-2167.

Bilias, Y., D. Georgarakos and M. Haliassos (2010), 'Portfolio Inertia and Stock Market Fluctuations,' *Journal of Money, Credit and Banking*, 42(4): 715-742.

Biais, B., P. Bossaerts and C. Spatt (2010), 'Equilibrium Asset Pricing and Portfolio Choice Under Asymmetric Information,' *Review of Financial Studies* 23(4): 1503-1543.

Brandt, M. W. (2010), 'Portfolio Choice Problems,' in Y. Ait-Sahalia and L.P. Hansen, eds., *Handbook of Financial Econometrics, Vol. 1: Tools and Techniques*, Elsevier North-Holland: Amsterdam.

Bruine de Bruin, W., B. Fischhoff, S. Millstein and B. Halpbern-Felscher (2000), 'Verbal and numerical expressions of probability: It's a Fifty-Fifty Chance,' *Organizational Behavior and Human Decision Processes*, 81(1): 115-131.

Cabrales, A. and P. Gottardi (2011), 'Markets for Information: On Inefficient Firewalls and Efficient Monopolies,' mimeo.

Cabrales, A., O. Gossner and R. Serrano (2013), 'Entropy and the Value of Information for Investors,' *American Economic Review* 103: 360-377.

Calvet, L. and P. Sodini (2013), 'Twin Picks: Disentangling The Determinants of Risk Taking in Household Portfolios,' fortcoming *Journal of Finance*.

Campbell, J. Y. (2006), 'Household Finance,' *The Journal of Finance*, 61: 1553-1604.

Campbell, J. Y., Lo, A. W., and A. C. MacKinlay (1997), *The Econometrics of Financial Markets*, Princeton University Press, Princeton, New Jersey.

Carroll, C. D. (2003), 'Macroeconomic Expectations of Households and Professional Forecasters,' *Quarterly Journal of Economics* 118(1): 269-98.

Cattell, R. (1971), *Abilities: Their structure, growth, and action.* New York: Houghton Mifflin Ed.

Christelis, D., T. Jappelli and M. Padula (2010), 'Cognitive Abilities and Portfolio Choice', *European Economic Review* 54: 18-38.

Curcuru, S., J. Heaton, D. Lucas and D. Moore (2010), 'Heterogeneity and Portfolio Choice: Theory and Evidence', in Y. Ait-Sahalia and L.P. Hansen, eds., *Handbook of Financial Econometrics, Vol. 1: Tools and Techniques*, Elsevier North-Holland: Amsterdam.

Deaton, A. (1992), *Understanding Consumption*, Oxford: Clarendon Press.

Dominitz, J. and C. Manski (1997), 'Using Expectations Data to Study Income Expectations,' *Journal of the American Statistical Association*, 92: 855-67.

Dominitz, J. and C. Manski (2007), 'Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study,' *Journal of the European Economic Association* 5: 369-79.

Dominitz, J. and C. Manski (2011), 'Measuring and Interpreting Expectations of Equity Returns,' *Journal of Applied Econometrics*, 26: 352–370.

Donkers, B. and A. van Soest (1999), 'Subjective measures of household preferences and financial decisions,' *Journal of Economic Psychology* 20: 613-642

Dow, J., and S. R. da Costa Werlang (1992), 'Uncertainty Aversion, Risk Aversion, and the Optimal Choice of Portfolio,' *Econometrica*, 60: 197-204.

Elmendorf, Douglas W. and Miles S. Kimball (2000), 'Taxation of labor income and the demand for risky assets,' *International Economic Review*, 41: 801-832.

Evans, G. W., and S. Honkapohja (2001), *Learning and Expectations in Macroeconomics*, Princeton U. Press, Princeton, NJ.

Frydman, R. and E.S. Phelps (1983), *Individual Forecasting and Aggregate Outcomes. "Rational Expectations" Examined*, Cambridge, Cambridge University Press.

Frydman, R. and E.S. Phelps (2013), *Rethinking Expectations: The Way Forward for Macroeconomics*, Princeton, Princeton University Press.

Gennotte, G. (1986), 'Optimal Portfolio Choice Under Incomplete Information', *Journal of Finance* 61: 733-46.

Gollier, C. (2001), *The Economics of Risk and Time*, Cambridge: MIT Press.

Gourieroux, C., Monfort, A., and Trognon, A. (1987). 'Simulated Residuals,' *Journal of Econometrics*, 34: 201-252.

Gomes, F. and A. Michaelides (2005), 'Optimal Life-cycle Asset Allocation: Understanding the Empirical Evidence,' *The Journal of Finance* 60(2): 869-904.

Greenwood, R. and A. Schleifer (2013), 'Expectations of Returns and Expected Returns,' NBER WP No. 18686.

Grinblatt, M., M. Keloharju and J. Linnainmaa (2011), 'IQ and Stock Market Participation,' *The Journal of Finance* 66(6): 2121-64.

Guesnerie, R. (1992), "An Exploration of the Eductive Justifications of the Rational Expectations Hypothesis", *American Economic Review* 82 :1254-1278.

Guiso, L. and T. Jappelli (2005), 'Awareness and Stock Market Participation,' *Review of Finance*, 9: 537-567.

Guiso, L. and T. Jappelli (2007), 'Information Acquisition and Portfolio Performance,' EUI WP ECO 2007/45.

Guiso, L. and M. Paiella (2008), 'Risk Aversion, Wealth and Background Risk,' *Journal of the European Economic Association*, 6: 1109-1150.

Guiso, L. and P. Sodini (2012), 'Household Finance. An Emerging Field,' *forthcoming* in the Handbook of the Economics of Finance, edited by Constandinides, G., M. Harris and R. Stulz, Elsevier Science.

Guiso, L., T. Jappelli and D. Terlizzese (1996), 'Income risk, borrowing constraints and portfolio choice,' *American Economic Review*, 86: 158–172.

Guiso, L., M. Haliassos and T. Jappelli (2002), *Household Portfolios*, Cambridge: MIT Press.

Guiso, L., Sapienza, P., and L. Zingales (2008), 'Trusting the stock market,' *The Journal of Finance*, 63: 2557-2600.

Haliassos, M. (2008), 'Household Portfolios,' in *The New Palgrave Dictionary of Economics*, Eds. Steven N. Durlauf and Lawrence E. Blume, Palgrave Macmillan, and in The New Palgrave Dictionary of Economics Online, Palgrave Macmillan.

Haliassos, M. and C. Bertaut (1995), 'Why Do So Few Hold Stocks?,' *Economic Journal*, 105: 1110-29.

Haliassos, M. and A. Michaelides (2003), 'Portfolio Choice and Liquidity Constraints,' *International Economic Review*, 44: 143-78.

Hall, R. E. (2010), 'Why Does the Economy Fall to Pieces after a Financial Crisis?,' *Journal of Economic Perspectives*, 24(4): 3-20.

Heaton, J., and D. Lucas (2000), 'Asset pricing and portfolio choice: the importance of entrepreneurial risk', *The Journal of Finance*, 55: 1163-1198.

Hommes, C. (2011), 'The Heterogeneous Expectations Hypothesis: Some Evidence from the Lab', *Journal of Economic Dynamics and Control*, 35(1): 1-24.

Hong, H. and J.C. Stein (2007), 'Disagreement and the Stock Market', *Journal of Economic Perspectives*, 21(2): 109-128.

Hong, H., Kubik, J.D., and J.C. Stein (2004), 'Social Interaction and Stock Market Participation,' *The Journal of Finance*, 59: 137-163.

Huang, L. and H. Liu (2007), 'Rational Inattention and Portfolio Selection', *The Journal of Finance*, 62: 1999-2040.

Hurd, M. D. (2009), 'Subjective Probabilities in Household Surveys', *Annual Review of Economics*, 1: 543-62.

Hurd, M. D., and S. Rohwedder (2012), 'Stock Price Expectations and Stock Trading,' NBER WP No. 17973.

Hurd, M. D., M. van Rooij and J. Winter (2011), 'Stock Market Expectations of Dutch Households', *Journal of Applied Econometrics* 26(3): 416-436.

Keane, M.P., and D. E. Runkle (1990), 'Testing the Rationality of Price Forecasts: New Evidence from Panel Data,' *The American Economic Review* 80(4): 714-35.

Kézdi, G. and R. J. Willis (2009), 'Stock Market Expectations and Portfolio Choice of American Households,' mimeo.

Kézdi, G. and R. J. Willis (2011), 'Households Stock Market Beliefs and Learning,' NBER WP No. 17614.

Kimball, M. S., C. R. Sham and M. D. Shapiro (2008), 'Imputing Risk Tolerance from Survey Responses,' *Journal of the American Statistical Association*, 103 (483): 1028-38.

King, M.A., and J. I. Leape (1987), 'Asset accumulation, information and the life cycle,' NBER wp No. 2392.

King, M.A., and J. I. Leape (1998), 'Wealth and portfolio composition: Theory and Evidence,' *Journal of Public Economics* 69: 155-93.

Kleinjans, K. J. and Arthur Van Soest (2013), 'Rounding, Focal Point Answers and Nonresponse to Subjective Probability Questions,' forthcoming *Journal of Applied Econometrics*.

Linnainmaa, J. T. (2011), 'Why Do (Some) Households Trade so Much?,' *Review of Financial Studies* 24(5), 1630-1666.

Lusardi, A. (2008), 'Financial Literacy: An Essential Tool for Informed Consumer Choice?,' mimeo.

Lusardi, A., P.-C. Michaud and O. Mitchell (2012), 'Optimal Financial Knowledge and Wealth Inequality,' mimeo.

Malmendier, U. and S. Nagel (2011), 'Depression Babies: Do Macroeconomic Experiences Affect Risk Taking?,' *Quarterly Journal of Economics* 126: 373-416.

Malmendier, U. and S. Nagel (2013), 'Learning from inflation Experiences,' *Quarterly Journal of Economics* 126: 373-416.

Manski, C. (2004), 'Measuring Expectations,' *Econometrica*, 72: 1329-76.

Merton, R. C. (1969), 'Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case,' *Review of Economics and Statistics*, 51(3): 247–257.

Merton, R. C. (1987), 'A Simple Model of Capital Market Equilibrium with Incomplete Information,' *Journal of Finance*, 42: 483–510.

Miniaci, R. and S. Pastorello (2010), 'Mean–Variance Econometric Analysis of Household Portfolios,' *Journal of Applied Econometrics* 25: 481–504.

O'Hara, M. (2003), 'Presidential Address: Liquidity and Price Discovery', *Journal of Finance* 58(4): 1335-54.

Paiella, M. (2001), 'Transaction Costs and Limited Stock Market Participation to Reconcile Asset Prices and Consumption Choices', IFS Working Paper.

Peress, J. (2004), 'Wealth, Information Acquisition and Portfolio Choice,' *Review of Financial Studies* 17(3): 879-914.

Pesaran, M. H., and M. Weale (2006), 'Survey Expectations,' in G. Elliott, C.W.J. Granger and A. Timmermann, eds., *Handbook of Economic Forecasting*, North-Holland.

Poterba, J. M. (2002), 'Taxation and Portfolio Structure: Issues and Implications,' in L. Guiso, M. Haliassos and T. Jappelli Eds., *Household Portfolios*, Cambridge: MIT Press.

Samuelson, P. A. (1969), 'Lifetime Portfolio Selection by Dynamic Stochastic Programming,' *Review of Economics and Statistics*, 51(3): 239–246.

Segal, U. and A. Spivak (1990), 'First versus Second Order Risk Aversion,' *Journal of Economic Theory* 51(1): 111-125.

Seru, A., T. Shumway and N. Stoffman (2009), 'Learning by Trading,' forthcoming *Review of Financial Studies*.

Sims, C. (2003), 'Implications of Rational Inattention', *Journal of Monetary Economics* 50(3): 665-90.

Stiglitz, J. E. (2011), 'Rethinking Macroeconomics: What Went Wrong and How to Fix it,' *Journal of the European Economic Association*, 9(4): 594-645.

Van Nieuwerburgh, S. and L. Veldkamp (2010), 'Information Acquisition and Under-Diversification,' *Review of Economic Studies*, 77(2): 779-805.

van Rooij, M., A. Lusardi and R. Alessie (2011), 'Financial Literacy and Stock Market Participation,' *Journal of Financial Economics* 101: 449-472.

Vissing-Jorgensen, A. (2002), 'Towards and Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures,' NBER WP No. 8884.

Vissing-Jorgensen, A. (2004), 'Perspectives on Behavioural Finance: Does Irrationality Disappear with Wealth? Evidence from Expectations and Actions,' in: M. Gertler and K. Rogoff eds., *The NBER Macroeconomics Annual 2003*, Cambridge: MIT Press.

Woodford, M. (2013), 'Macroeconomic Analysis without the Rational Expectations Hypothesis,' forthcoming *Annual Review of Economics*.

Zhang, X.F. (2006), 'Information Uncertainty and Stock Market Returns,' *The Journal of Finance*, 61(1): 105-137.

Appendix

A. Variable Definitions

Endowments

Total wealth: In the survey, the respondent is asked in which of the eight predefined available brackets is her household's situation. As we are interested in a continuous measure, we have used the method of simulated residuals (Gourieroux et al. 1987). We have regressed an ordered probit of the respondents' total wealth (bracket) on demographic and socio-economic household characteristics. Once we have the estimated total wealth, a normally distributed error is added. We then check if the value falls inside the bracket originally chosen by the individual. If not, another normal error is added and so on until we predict the true interval. Doing so allows us to overcome the non-response problem for some households. If there is a missing value, the predicted value plus a normal error is directly used. Total wealth is given in Euros.

Income: For the income of the household, the survey asks the respondent which of the 7 predefined available brackets better corresponds to her situation. We obtained a continuous measure in the same way we did for total wealth. Income refers to the individual's household annual income in Euros.

Preferences

Absolute risk aversion (CARA): The following question is asked to the respondent: 'If someone suggests that you make an investment, $\widetilde{S}_i$, whereby you have one chance out of two win 5000 euros and one chance out of two of losing the capital invested, how much (as a maximum) will you invest?' The question aims at eliciting the taste for risk from each respondent i, with preferences $u^i(.)$, from the following equality:

$$u^i(w_i) = \frac{1}{2}u^i(w_i + 5,000) + \frac{1}{2}u^i(w_i - Z_i) \equiv Eu^i(w_i + \widetilde{S}_i)$$

The coefficient of absolute risk aversion (CARA) can be then obtained from a second order Taylor expansion, as

$$A_i(w_i) = 2(5000 - Z_i)/(5000^2 + Z_i^2)$$

where Z_i is the amount that the respondent declares to be willing to invest. Those who declare $Z_i < 5000$ are risk-averse $Z_i = 5000$, are risk-neutral and $Z_i > 5000$ are risk-lovers. The outcome range for the coefficient of absolute risk aversion $A_i(w_i)$ is $[0, 40]$. 3,343 respondents answered the question, with a mean response of 39.11. In the TNS 2007, the histogram of responses is very skewed to the left. Further details regarding the measure of absolute risk aversion (CARA) can be found in Guiso and Paiella's (2008) work.

Relative risk aversion (CRRA): To obtain a measure of risk aversion, we asked individuals about their willingness to gamble on lifetime income according to the methodology of Barsky et al. (1997). The "game" resides in determining sequentially whether the interviewee would accept to give up his present income and to accept other contracts, in the form of lotteries: he has one chance in two to double his income, and one chance in two for it to be reduced by one third (contract A), by one half (contract B), and by one fifth (contract C). More precisely, the question in the survey was:

* 'Suppose that you have a job which guarantees for life your household's current income R. Other companies offer you various contracts which have one chance out of two (50%) to provide you with a higher income and one chance out of two (50%) to provide you with a lower income.

* Are you prepared to accept Contract A which has 50% chances to double your income R and 50% chances that your income will be reduced by one third?

* For those who answer YES : the Contract A is no longer available. You are offered Contract B instead which has 50% chances to double your income R and 50% chances that it will be reduced by one half. Are you prepared to accept?

* For those who answer NO : you have refused Contract A. You are offered Contract C. which has 50% chances to double your income R and 50% chances that it will be reduced by 20%. Are you prepared to accept?'

This allows us to obtain a range measure of relative risk aversion under the assumption that preferences are strictly risk averse and utility is of the CRRA type. The degree of relative risk aversion is less than 1 if the individual successively accepts contracts A and B; between 1 and 2 if he accepts A but refuses B; between 2 and 3.76 if he refuses A but accepts C; and finally more than 3.76 if he refuses both A and C.

Temporal preference: It is a numerical scale from 0 to 10. The survey asks the respondent about her attitude regarding life: 0 represents living the present (impatience) and 10 only caring about the future (extreme patience).

Demographics

Gender: is a dummy variable equal to 1 if the household head is a male, and is equal to 0, if a female.

Marital status: Marital status is based on current legal marital status. Respondents who are married or/and living with a partner are coded as 1, and 0 otherwise.

Constraints

Liquidity constrained: Respondents are asked if they ever had to struggle to balance their household budget. It is a dummy variable that takes value 1 if the respondent answers the question in the categories 'very often' or 'often', and value 0 otherwise.

Online banking: It is a dummy variable that takes value 1 if the respondent uses the internet for managing her financial accounts, and 0 otherwise.

Importance of Money in Life: Respondents are asked about the relative importance of money in life. It is a quantitative variable on a discrete scale from 0 to 10.

Income Risk: Question qa16 in the survey asks respondents about the regularity of household's income (wages, retirement income...), providing three categories: 'regular, certain'; 'irregular, random' and 'partly certain, partly random'. Income risk is defined as a dummy variable that takes value 1 if the respondent answers 'irregular, random', and zero otherwise.

Delegation/Inertia/Trust

Self portfolio management: The survey asks the respondent who takes household's financial decisions (stocks, SICAV/FCP bonds, life insurance contracts, saving accounts). Respondents who answer 'themselves' or 'them with their partners' are coded as 1, and 0 otherwise (which includes sharing some decisions with a financial advisor, or the financial advisor taking all decisions on the households' behalf).

Firm shares in remuneration: It is a dummy variable that takes value 1 if the respondent receives shares of the firm he works in as part of her compensation package/remuneration, and 0 otherwise.

Frequency of recent trades: Respondents are asked about the number of stock market operations closed over the year prior to the date in which the survey was conducted (March 2006-March 2007). The answers are categorical: no operations, 1-2 operations, 3-5 operations, 6 or more operations.

Optimism: Respondents are inquired 'whether they esteem that they have been lucky in life'. It is a discrete variable, that takes value 1 if they answer 'yes', and 0 otherwise.

Trust: Respondents are inquired 'whether they trust online payment systems'. It is a discrete variable that takes value 1 if they answer either 'yes' or 'rather yes', and 0 if they either answer 'rather no' or 'absolutely not'.

Parents own stocks: Respondents are inquired 'whether their parents invest/ed in the stock market either directly or indirectly'. It is a discrete variable that takes value 1 if they answer either 'yes', and 0 if they either answer 'no'.

Information

Education: is a categorical variable, grouped into three broad categories: High school or less (primary and secondary), some college (technical degrees beyond high school but below college, including professional and vocational degrees) and college or more (BAs, BScs, MScs, MBAs, professional certifications, PhDs and postdoctoral students).

Sources of Information variables:

* Respondents are inquired, for each alternative source of information (Friends, family, financial advisors, general media and specialised media), about the relative frequency of consultation (often, sometimes or never). For each information source, a dummy variable is created which takes value 1 if the answer is 'often', and 0 otherwise.

* Respondents are inquired, for each alternative source of TV information (General information and economics emissions) , about the relative frequency of consultation (very often, often, occasionally, sometimes or never). For each information source, a dummy variable is created which takes value 1 if the answer is 'often' or 'very often', and 0 otherwise.

Mother's education: The respondent is inquired about the educational attainment of her mother. Three categories are available: less than High school, completed High school and more than High school. The reference category is 'less than High school'.

Father's education: The respondent is inquired about the educational attainment of her father. Three categories are available: less than High school, completed High school and more than High school. The reference category is 'less than High school'.

Family background: The respondent is inquired about her family background. Three categories are available: Middle class, low class and neither middle nor low class. The reference category is 'neither middle nor low class'.

B. Descriptive Statistics

Table 11: Descriptive Statistics

	Whole Sample		Respondents	
	Mean	Std. Dev.	Mean	Std. Dev.
Positive nominal return (PNR)	46.45	39.68	46.45	39.68
Past positive nominal return (pPNR)	68.07	37.7	68.44	37.73
Risk aversion (CARA)	39.11	3.58	39	3.77
Temporal preference	6.73	2.33	6.79	2.25
Age	48	17	47	16
		% of the sample		
Male	45.6		49.1	
Married/living with a partner	63.2		66.6	
Education:				
High School or less	14.8		10.2	
Less than college	62.2		62.3	
College or more	23		27.5	
Income (in €/year):				
Less than 8,000	18.3		15.2	
8,000-15,999	32.1		28.8	
16,000-29,999	38.4		41.6	
More than 30,000	11.2		14.4	
Total wealth (in €):				
Less than 39,999	21.3		23.3	
40,000-149,999	29.1		28.9	
150,000-449,999	30.4		24.2	
More than 450,000	19.2		23.6	
Self management account	51.5		52	
Financial advisor or other	6.1		5.9	
Firm shares in remuneration	4.7		5.6	
Parents own stocks	31		33	
Online banking	19.7		24.5	
Liquidity constrained	43.6		41.1	
Irregular income risk	10.2		10.1	
Trust	5.57		5.55	
Stockownership	30.7		37.5	
No. of Observations	3,826		2,406	

Source: TNS 2007.

Not for Publication Appendix to: Subjective Return Expectations, Information and Stock Market Participation: Evidence from France*

Luc Arrondel† CNRS-PSE-Banque de France

Hector Calvo-Pardo‡ U. of Southampton

Derya Tas§ U. of Southampton

December 13, 2013

A. Elementary Theory

The standard two-asset model, in either its static (Arrow, 1965) or dynamic (Merton, 1969; Samuelson, 1969) version[1], predicts that a necessary and sufficient condition for investing in the risky asset ($\alpha_t^{i*} > 0$) is that its expected return, $E_t^i r_{t+1}$, exceeds the return of the riskless asset, r :

$$\max_{\alpha_t^i \in [0, w_t^i]} \int_{R_{\min}^i}^{R_{\max}^i} u_i[(1+r)w_t^i + (r_{t+1} - r)\alpha_t^i] p^i(r_{t+1} | r_t) dr_{t+1}$$

$$\text{FOC:} \qquad E_t^i\{(r_{t+1} - r)u_i'[(1+r)w_t^i + (r_{t+1} - r)\alpha_t^{i*}]\} = 0$$

$$\text{Participation Condition:} \qquad E_t^i r_{t+1} - r > 0$$

$$\text{Conditional Demand Equation:} \qquad \alpha_t^{i*} \cong \frac{E_t^i r_{t+1} - r}{A_u^i(w_t^i)\sigma_{it}^2}$$

*The authors gratefully acknowledge financial support from the "Agence Nationale pour la Recherche" (ANR), Caisse des Depots et Consignations, Chaire Groupama "Les Particuliers face au Risque", ESRC (RES-061-25-0327) and the University of Cambridge for its hospitality. *The authors declare that they have no relevant or material financial interests that relate to the research described in this paper. The authors have obtained IRB approval from their institutions.*

†Batiment A, Ecole Normale Superieure, 48 Bd Jourdan, 75014 Paris, France, arrondel@pse.ens.fr

‡Division of Economics, FSHS, University of Southampton, Southampton SO171BJ, UK, calvo@soton.ac.uk

§Division of Economics, FSHS, University of Southampton, Southampton SO171BJ, UK, tashderya@gmail.com

[1]The dynamic (life-cycle) versions that produce a myopic (independent of the investment horizon) optimal share of wealth invested in risky assets, necessitate in addition, that log-returns are normally distributed, independent and identically distributed through time, and investors' preferences in the constant relative risk aversion (CRRA) class. See Brandt (2010) for a detailed explosition.

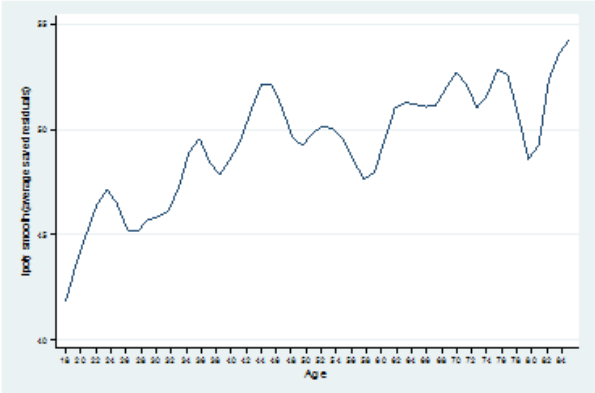


Figure 1: Conditional Subjective Expectation by Age, TNS 2007.

where $E_t^i\{.\}$ denotes the subjective expectation of the decision maker i conditional on his individual information set, I_t^i. Individuals are fully characterized by their preferences, u_i, initial wealth endowments, w_t^i, and beliefs, $E_t^i\{.\} \equiv E\{.\,|\,I_t^i\}$, over the available investment opportunities rewarding time, r, and risk and time, r_{t+1}. Conditional on the reward for bearing risk being positive ($E_t^i r_{t+1} - r > 0$), investors then allocate more or less of their wealth to the risky asset (α_t^{i*}) depending on how much do they distaste risk, represented by the absolute risk aversion coefficient, $A_u^i(w_i) \equiv -\frac{u_i''(w_t^i)}{u_i'(w_t^i)}$. Risk itself is summarized by the subjective variance of the risky asset return, σ_{it}^2.

In this work we are only interested in the participation condition. For an estimation of the demand for risky assets at both the intensive and extensive margins, see Arrondel et al. (2013).

The conditional subjective expectation

We extract the information component from subjective expectations, by regressing expectations (PNR^i) on information (pPNR^i), and recover the predicted residuals, $\widehat{\epsilon}^i$, which are plotted by age in Figure 1.

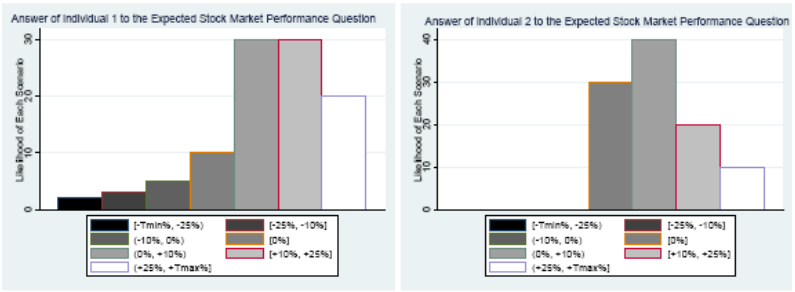
For all ages, the regression summary statistics are:

$$PNR^i = \underset{(1.65)}{21.57} + \underset{(0.021)}{0.40}\, pPNR^i + \epsilon^i; \qquad \underset{N=2{,}066}{R^2 = 0.15}$$

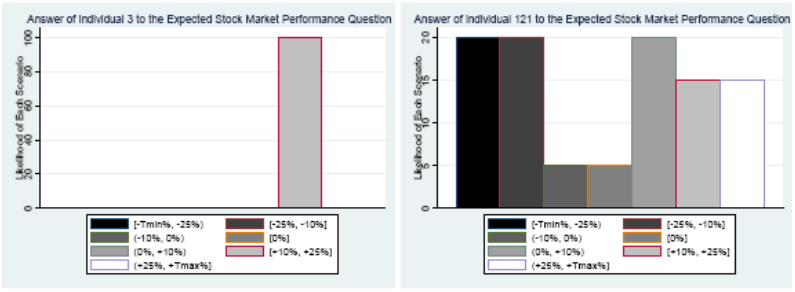
If we replace PNR and pPNR in the main equations by the predicted residuals, $\widehat{\epsilon}^i$, they are statistically significant only for the young, conditional on all other covariates. These results are available upon request.

B. Inspecting the Data

How do answers to the forwad-looking return expectations (Figure 2) and return realizations question (Figure 3) compare –for the same four respondents?



Panel (a): Individual 1. Panel (b): Individual 2.



Panel (c): Individual 3. Panel (d): Individual 121.

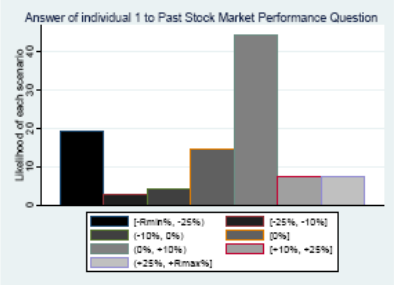
Figure 2: Individual histograms of respondents' answers to survey question qc6 (PNR). Source: TNS 2007.

Histograms of CAC 40 index historical log returns

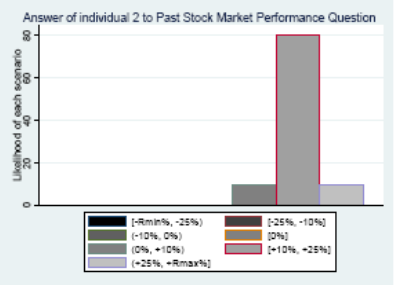
The density of nominal yearly (and 5-year rolling) log returns on the CAC-40 computed from monthly data between July 1987 and July 2011 is depicted in Figure 4, panel (a) (panel (b)). The distribution has sample moments $\mu = 0.023$ ($\mu(5) = 0.108$) and $\sigma = 0.10$ ($\sigma(5) = 0.19$). The densities depicted in Figure 4 can be thought as representing the subjective beliefs of those respondents who base them on the history of observed stock market index monthly closing values.

Missing/Erroneous answers to the expectations question

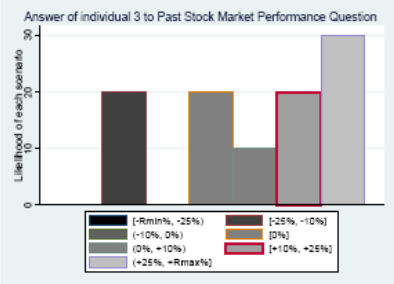
In Table 1, we estimate a Probit specification for erroneous or missing answers (1 if answer to qc6 missing or does not add up to 100) as a function of stockholding, and covariates (gender, marital status, education, risk preference).



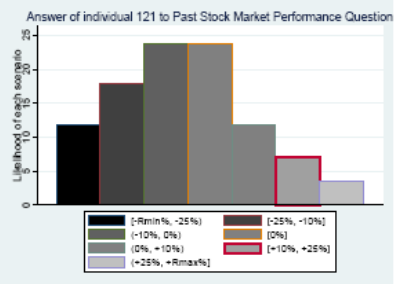
Panel (a): Individual 1.



Panel (b): Individual 2.

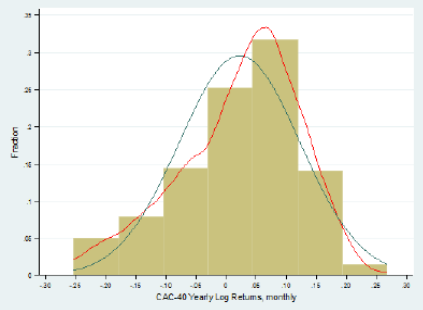


Panel (c): Individual 3.

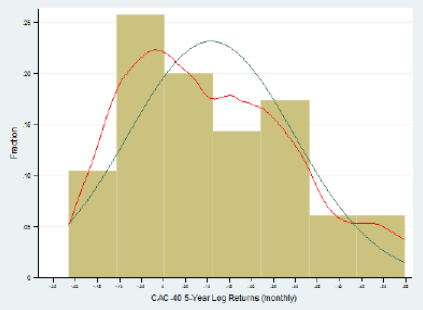


Panel (d): Individual 121.

Figure 3: Individual histograms of responses to the survey question C9 (pPNR). Source: TNS 2007.



Panel (a): 1-year log-returns.



Panel (b): 5-year log-returns.

Figure 4: Histogram of CAC-40 index log-returns, computed at 1-year (panel a) and 5-year (panel b) rolling window frequencies. Source: Author's own calculations using monthly data between July 1987 and July 2011, available online from MSN Money.

Table 1: Probability of Missing or Erroneous Answers - Marginal Effects

	(1)	(2)	(3)	(4)	(5)
Stockholding	-0.202***	-0.198***	-0.193***	-0.170***	-0.158***
	(0.0156)	(0.0157)	(0.0158)	(0.0164)	(0.0167)
Male		-0.0792***	-0.0697***	-0.0712***	0.0774***
		(0.0157)	(0.0159)	(0.0161)	(0.0161)
Married/living with a partner			-0.0668***	-0.0718***	-0.0614***
			(0.0167)	(0.0169)	(0.0170)
Education (Ref. category: High school or less)					
Less than college				-0.168***	-0.172***
				(0.0228)	(0.0229)
College or more				-0.256***	-0.258***
				(0.0208)	(0.0208)
Temporal preference					-0.0149***
					(0.00296)
Pseudo R-squared	0.0293	0.0343	0.0375	0.0601	0.0651
Chi-squared	148.1	173.2	189.4	303.3	328.8
Log-likelihood	-2449	-2437	-2429	-2372	-2359
No of observations	3,826	3,826	3,826	3,826	3,826

Note: (i) Standard errors are in parentheses. (ii) *** p<0.01, ** p<0.05, * p<0.1. Source: TNS 2007.

C. Additional Data Validation Results

Table 2 reproduces Table 1 in Dominitz and Manski (2007) for all ages, reporting the distribution of responses and the response rate conditioning on age, gender, and marital and stockholding status. Consistent with their findings, the mean percentage chance of a positive nominal return is higher for respondents who are married. The differential is 1.4 percentage points for men (50.1% versus 48.7%), while for females, it is only 1.1 percentage points (43.7% versus 42.6%). Figure 5 shows that also in France, males are more optimistic than females. Even when conditioning on marital and stockholding status and for all age brackets, men give more optimistic reports than do women, and are 6-7 percentage points more likely to give a response.

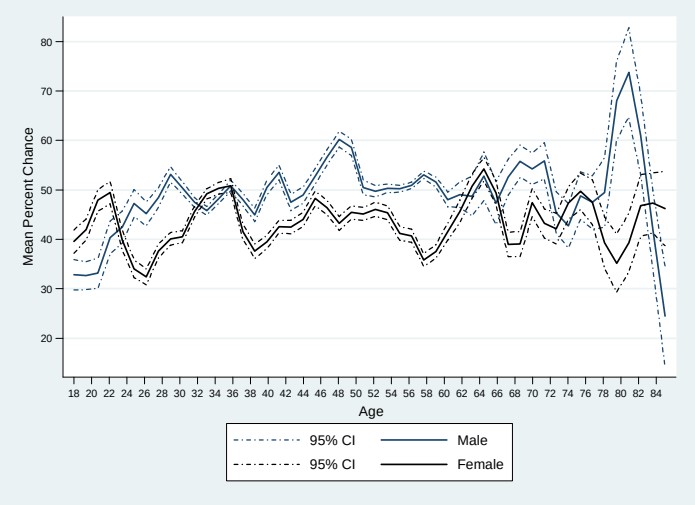


Figure 5: Mean PNR by age and gender. Source: TNS 2007.

In line with the HRS 2002 findings by Kézdi and Willis (2009), expectations (noisily) increase with the

Table 2: Expectations of positive nominal return (PNR), by attribute; TNS 2007

	Male							Female						
Attribute to PNR	Number of respondents to PNR	Mean	St. Dev.	Quantile 0.25	Quantile 0.50	Quantile 0.75	Rate of response to PNR	Number of respondents to PNR	Mean	St. Dev.	Quantile 0.25	Quantile 0.50	Quantile 0.75	Rate of response to PNR
All Respondents	1,169	49.7	40.2	0	50	95	0.77	1,205	43.2	39.1	0	40	80	0.66
Married or living with a partner														
No	322	48.7	38.2	0	50	90	0.76	471	42.6	38.9	0	40	80	0.60
Yes	847	50.1	40.9	0	50	97	0.78	734	43.7	39.2	0	40	83	0.72
Age														
Under 30	150	42.6	36.4	0	38	70	0.71	193	40.1	37.1	0	35	75	0.68
30-39	242	47.2	38.3	0	50	80	0.81	280	45.4	38.5	0	45	80	0.75
40-49	252	53.7	40.2	0	60	100	0.81	236	43.5	39.0	0	40	82	0.71
50-59	240	51.4	41.1	0	58	100	0.81	243	41.5	39.2	0	40	80	0.68
60-69	166	50.1	42.8	0	55	100	0.77	145	45.0	40.1	0	45	90	0.64
70 and over	119	51.6	42.3	0	55	100	0.68	108	44.4	43.1	0	40	98	0.46
Holds stocks or mutual funds														
No	709	44.0	40.2	0	40	90	0.71	777	39.2	38.8	0	30	75	0.61
Yes	460	58.6	38.6	20	70	100	0.90	428	50.5	38.6	5	50	90	0.81

Note: Sample restricted to those with own or spouse/partner report of whether or not household holds 'stocks or stock mutual funds'.

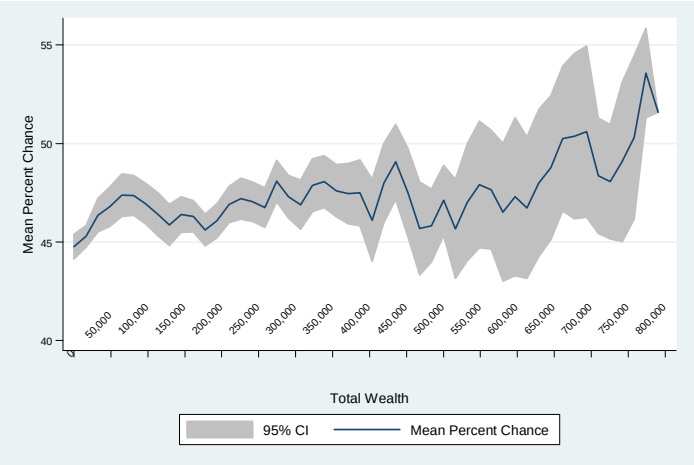


Figure 6: Mean percentage chance of a positive nominal return over the next 5 years (PNR) by total wealth. Source: TNS 2007.

respondent's education and households' total wealth. Figure 6 (7) displays kernel-smoothed estimates of the mean percentage chance of a positive nominal return conditional on total wealth (educational attainement). The solid curve depicts point estimates and the grey area around it represents (bootstrap) 95% confidence intervals. The median of total wealth is €118,792, and the 90th percentile €413,476. We excluded 42 households with wealth above €800,000. The richest households (with wealth above the 90th percentile), appear more optimistic regarding the future evolution of the stock market. An increase in wealth from the 10th to the 90th percentile, is estimated to increase the mean percentage chance of a positive nominal return by about 2 to 3 percentage points.

Figure 7 compares by age, the mean percentage chance of a positive nominal return of respondents with some college education or more relative to those having at most completed high school. Broadly, the former seem to be slightly more optimistic than the latter, although both tend to similarly become more optimistic as they age.

Table 3 reports both table 2 in Dominitz and Manski (2007) and our replication using instead data from the TNS 2007 survey, for respondents in the same age bracket (50-80 years old). Figure shows that among the 50-80 year-olds, the probability of holding stocks is increasing in the percent chance of a positive Stock Market return, albeit in a more volatile way than in the US, since we have less observations.

Table 3: Comparing Dominitz and Manski's (2007) HRS 2004 with TNS 2007 for 50 $\leq$Age$\leq$ 80

Percent chance of positive nominal return	Married or living with a partner: Male Point Estimate	Married or living with a partner: Male Standard Error	Married or living with a partner: Female Point Estimate	Married or living with a partner: Female Standard Error	*NOT* Married or living with a partner: Male Point Estimate	*NOT* Married or living with a partner: Male Standard Error	*NOT* Married or living with a partner: Female Point Estimate	*NOT* Married or living with a partner: Female Standard Error
0	0.16	(0.02)	0.25	(0.03)	0.08	(0.03)	0.08	(0.02)
1-10	0.27	(0.03)	0.31	(0.02)	0.16	(0.04)	0.20	(0.02)
11-20	0.30	(0.03)	0.34	(0.03)	0.16	(0.05)	0.14	(0.03)
21-30	0.29	(0.03)	0.35	(0.02)	0.19	(0.05)	0.23	(0.03)
31-40	0.33	(0.04)	0.37	(0.03)	0.16	(0.05)	0.18	(0.03)
41-49	0.22	(0.14)	0.18	(0.12)	0.50	(0.25)	0.33	(0.14)
50	0.37	(0.01)	0.40	(0.01)	0.25	(0.02)	0.25	(0.02)
51-59	0.50	(0.14)	0.63	(0.17)	0.20	(0.18)	0.20	(0.18)
60-69	0.48	(0.03)	0.50	(0.03)	0.30	(0.06)	0.31	(0.03)
70-79	0.48	(0.02)	0.50	(0.02)	0.38	(0.04)	0.41	(0.03)
80-89	0.52	(0.02)	0.52	(0.03)	0.42	(0.05)	0.30	(0.04)
90-99	0.48	(0.03)	0.49	(0.05)	0.24	(0.07)	0.43	(0.07)
100	0.43	(0.03)	0.45	(0.04)	0.25	(0.05)	0.23	(0.04)
All	0.40	(0.01)	0.40	(0.01)	0.25	(0.01)	0.24	(0.01)

Source: Dominitz and Manski (2007).

Percent chance of positive nominal return	Married or living with a partner: Male Point Estimate	Married or living with a partner: Male Standard Error	Married or living with a partner: Female Point Estimate	Married or living with a partner: Female Standard Error	*NOT* Married or living with a partner: Male Point Estimate	*NOT* Married or living with a partner: Male Standard Error	*NOT* Married or living with a partner: Female Point Estimate	*NOT* Married or living with a partner: Female Standard Error
0	0.37	(0.03)	0.34	(0.03)	0.33	(0.04)	0.30	(0.03)
1-10	0.49	(0.10)	0.45	(0.10)	0.44	(0.10)	0.40	(0.10)
11-20	0.51	(0.09)	0.47	(0.09)	0.46	(0.09)	0.42	(0.09)
21-30	0.29	(0.07)	0.26	(0.07)	0.24	(0.07)	0.22	(0.06)
31-40	0.50	(0.08)	0.46	(0.08)	0.44	(0.08)	0.41	(0.08)
41-49	0.32	(0.13)	0.29	(0.12)	0.28	(0.12)	0.25	(0.11)
50	0.50	(0.07)	0.46	(0.07)	0.45	(0.07)	0.41	(0.07)
51-59	0.37	(0.17)	0.33	(0.16)	0.32	(0.16)	0.29	(0.15)
60-69	0.58	(0.07)	0.55	(0.08)	0.53	(0.08)	0.50	(0.08)
70-79	0.70	(0.07)	0.67	(0.07)	0.65	(0.07)	0.62	(0.08)
80-89	0.56	(0.07)	0.52	(0.07)	0.51	(0.08)	0.48	(0.07)
90-99	0.65	(0.07)	0.61	(0.07)	0.60	(0.08)	0.56	(0.07)
100	0.53	(0.04)	0.50	(0.04)	0.48	(0.05)	0.45	(0.04)
All	0.47	(0.02)	0.44	(0.03)	0.43	(0.04)	0.41	(0.03)

Source: TNS 2007.

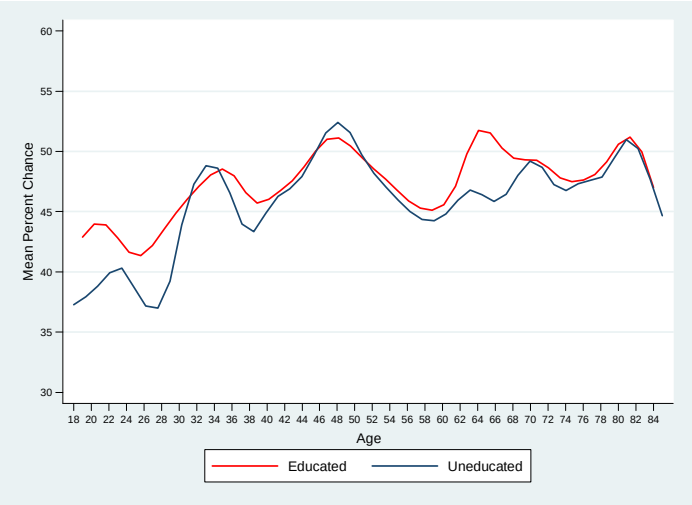
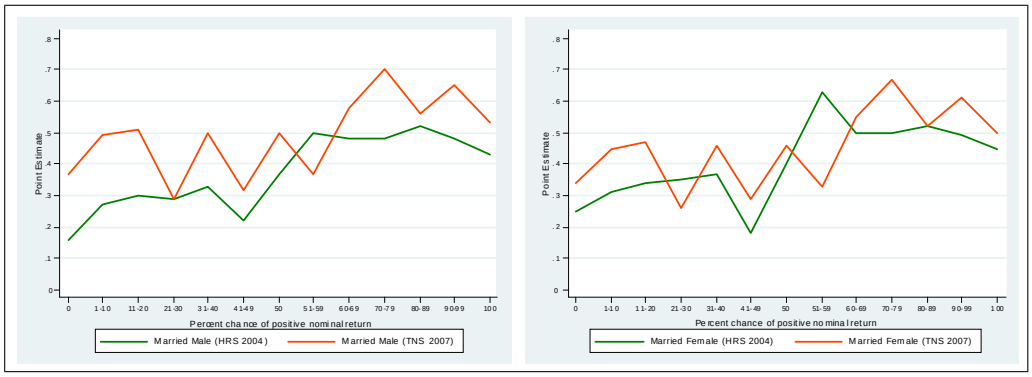


Figure 7: Mean percentage chance of a positive nominal return over the next 5 years (PNR) by educational attainement. Source: TNS 2007.



The Probability of Holding Stocks Conditional on the Percent Chance of a Positive Nominal Return (PNR), Gender, and Marital Status. Source (left panel) Dominitz and Manski (2007) and (right panel) TNS 2007.

D. Additional Descriptive Evidence on Information (pPNR)

Table 4 reports the distribution of responses and the response rate conditioning on age, gender and stockholding status. In accordance with the findings reported in Lusardi (2008) on the financial literacy of US adults, male respondents who are older, single and stockholders report higher mean (and a lower standard deviation of) percentage chances of a past positive nominal return. The differential is 11.4 percentage points higher for men than for women. Although information broadly increases with age, irrespective of gender, the uncertainty of the reports decrease with age for males, while for females, it remains broadly constant. Stockholders report a higher mean by about 10 percentage points, and are around 6 percentage points more likely to give a response. Figure 8 shows that males are broadly better informed than females.

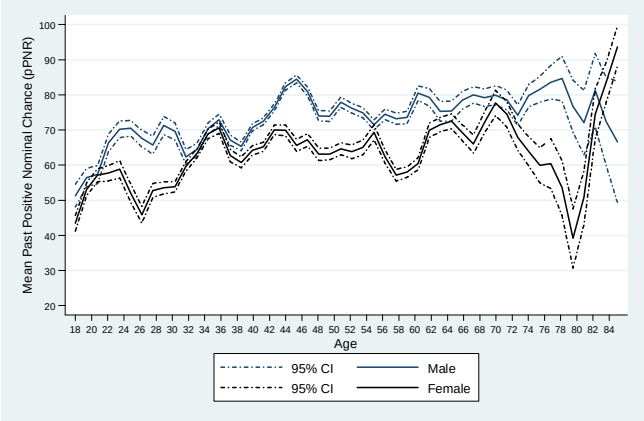


Figure 8: Mean percentage chance of a positive nominal return over the last 5 years (pPNR) by age and gender. Source: TNS 2007.

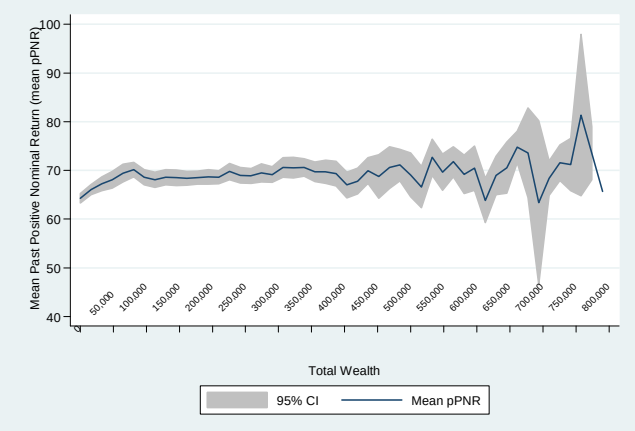


Figure 9: Mean percentage chance of a positive nominal return over the last 5 years (pPNR) by wealth. Source: TNS 2007.

Figure 9 shows that information broadly increases until the 50th percentile of wealth (€118,792), remains constant until the 90th percentile (€413,476), only to increase again albeit very heterogeneously. The richest households (with wealth above the 90th percentile), may thus be more optimistic (and disagree more)

Table 4: Distribution of Responses to Past Positive Nominal Return (pPNR), by Attribute; TNS 2007

	Male							Female						
	Number of respondents			Quantile			Rate of response to	Number of respondents			Quantile			Rate of response to
Attribute	to pPNR	Mean	St. Dev.	0.25	0.50	0.75	pPNR	to pPNR	Mean	St. Dev.	0.25	0.50	0.75	pPNR
All Respondents	1,024	74.2	36.2	50	100	100	0.72	1,009	62.8	38.5	0	30	100	0.88
Married or living with a partner														
No	286	76.8	34.5	60	100	100	0.95	386	59.3	38.6	25	70	100	0.86
Yes	738	73.2	36.9	50	100	100	0.91	623	64.9	38.3	30	80	100	0.87
Age														
Under 30	132	66.6	36.8	40	82.5	100	0.94	156	52.5	38.2	17.5	50	90	0.83
30-39	215	66.8	39.3	30	80	100	0.93	235	63.1	38.4	25	75	100	0.87
40-49	223	77.9	33.6	60	100	100	0.93	208	65.9	36.7	40	80	100	0.92
50-59	219	74.7	37.2	50	100	100	0.87	201	69.3	39.0	25	75	100	0.94
60-69	137	80.2	33.2	10	75	100	0.89	123	68.3	38.6	30	90	100	0.90
70 and over	98	82.9	31.8	20	80	100	0.87	84	67.7	38.7	42.5	85	100	0.83
Holds stocks or mutual funds														
No	604	68.8	37.5	45	90	100	0.90	627	59.4	38.5	25	70	100	0.85
Yes	420	81.9	32.8	80	100	100	0.96	380	68.8	37.6	42.5	90	100	0.93

Note: Sample restricted to those with own or spouse/partner report of whether or not household holds 'stocks or stock mutual funds'.

regarding the future investment opportunities because they are better (and more heterogeneously) informed. Relative to non-participation amongst the most affluent, and in line with Guiso and Jappelli's (2005) findings, heterogeneity in stock market information appears as a different alternative to better private investment opportunities (Heaton and Lucas, 2000) or taxation (Poterba, 2002), while encompassing the ability to process it (Grinblatt et al., 2011), its sources (social interactions, Hong et al., 2004; trust, Guiso et al., 2008) or more broadly, financial literacy (van Rooij et al., 2011). Figure 10 shows that information about past stock market performance does not increase with own's educational attainement, although it broadly increases with age within educational groups.

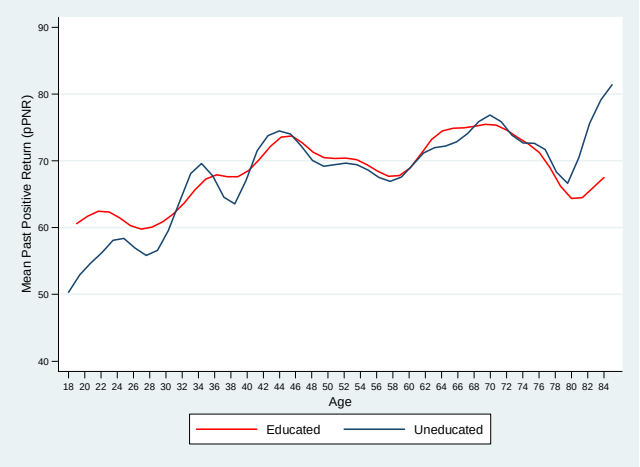


Figure 10: Mean percentage chance of a positive nominal return over the last 5 years (pPNR), by education. Source: TNS 2007.

1 References

Arrondel, L., Calvo-Pardo, H., and X. Oliver (2013), 'Subjective Stock Market Return Expectations, Information and Portfolio Choice', mimeo.

Arrow, K. J. (1965). *Aspects of the Theory of Risk Bearing*, Yrjo Jahnsson Lectures, The Academic Book Store, Helsinki.

Brandt, M. W. (2010), 'Portfolio Choice Problems,' in Y. Ait-Sahalia and L.P. Hansen, eds., *Handbook of Financial Econometrics, Vol. 1: Tools and Techniques*, Elsevier North-Holland: Amsterdam.

Dominitz, J. and C. Manski (2007), 'Expected Equity Returns and Portfolio Choice: Evidence from the Health and Retirement Study,' *Journal of the European Economic Association* 5: 369-79.

Grinblatt, M., M. Keloharju and J. Linnainmaa (2011), 'IQ and Stock Market Participation,' *The Journal of Finance* 66(6): 2121-64.

Guiso, L. and T. Jappelli (2005), 'Awareness and Stock Market Participation,' *Review of Finance*, 9: 537-567.

Guiso, L., Sapienza, P., and L. Zingales (2008), 'Trusting the stock market,' *The Journal of Finance*, 63: 2557-2600.

Heaton, J., and D. Lucas (2000), 'Asset pricing and portfolio choice: the importance of entrepreneurial risk', *The Journal of Finance*, 55: 1163-1198.

Hong, H., Kubik, J.D., and J.C. Stein (2004), 'Social Interaction and Stock Market Participation,' *The Journal of Finance*, 59: 137-163.

Kézdi, G. and R. J. Willis (2009), 'Stock Market Expectations and Portfolio Choice of American Households,' mimeo.

Lusardi, A. (2008), 'Financial Literacy: An Essential Tool for Informed Consumer Choice?,' mimeo.

Merton, R. C. (1969), 'Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case,' *Review of Economics and Statistics*, 51(3): 247–257.

Poterba, J. M. (2002), 'Taxation and Portfolio Structure: Issues and Implications,' in L. Guiso, M. Haliassos and T. Jappelli Eds., *Household Portfolios*, Cambridge: MIT Press.

Samuelson, P. A. (1969), 'Lifetime Portfolio Selection by Dynamic Stochastic Programming,' *Review of Economics and Statistics*, 51(3): 239–246.

van Rooij, M., A. Lusardi and R. Alessie (2011), 'Financial Literacy and Stock Market Participation,' *Journal of Financial Economics* 101: 449-472.