Return Expectations and Risk Aversion Heterogeneity in Household Portfolios

First draft, preliminary and incomplete

Please, do not quote

Alessandro Bucciol (University of Verona & Netspar)

Raffaele Miniaci (University of Brescia)

Sergio Pastorello (University of Bologna)

We develop a structural econometric model within the mean variance portfolio framework to elicit household-specific expectations about future financial market returns and risk attitude, using data on observed portfolio holdings and answers to self-assessed questions on the willingness to bear financial risk. We derive an explicit solution of the model, which we analyze using a combination of tobit and ordered probit estimation methods with three latent variables. We find a positive relation of wealth and age with expectations on market returns; expectations vary with the type of portfolio, and they are the highest among those holding both bonds and stocks. In contrast wealth and age seem to have a negligible impact on risk aversion, once estimated jointly with expectations. Expectations also change over time, and fall after periods of growing markets.

NB: we are currently working along different lines: 1) we extend the data set by including data from 2010 SCF, this would allow us to investigate how households expectations and risk attitude changed during the most recent crisis, 2) we consider the role of housing as non-tradable asset

JEL codes: G11, D81, D14.

Keywords: Household finance, mean variance analysis, risk aversion, maximum likelihood.

1 Introduction

Only around 50% of the US households hold stocks, either directly or indirectly (Bucks et al. 2009); rates of participation in the stock market are even smaller in Europe (Guiso et al., 2003). There are several reasons for this low participation rate, including entry barriers and transaction costs (Haliassos and Michaelides, 2003; Vissing-Jorgensen, 2002), high borrowing costs (Davis et al., 2006), crowding-out effects due to real asset holdings (Cocco, 2005; Flavin and Yamashita, 2002) and financial illiteracy (Christelis et al., 2010; Guiso and Jappelli, 2005; van Rooij et al., 2011). In addition, life-cycle portfolio models predict that the investment in stocks correlates with age, income and entrepreneurial risk (Gomes and Michaelides, 2005; Heaton and Lucas, 2000; Viceira, 2001). However, among those who invest in stocks, large differences in portfolios held by similar households are quite common (Guiso et al., 2002) and cases of poorly diversified portfolio allocation are also frequent (Calvet et al., 2007).

Households make their own choice based on their expectations about the market return and on their financial risk attitude. The heterogeneity on these variables determines different portfolio holdings. In this paper we elicit households' expectations about future financial market performances *together with* their risk attitude by combining information from two types of data: observed financial portfolio holdings and self-assessed willingness to bear financial risks. We use a structural econometric model to investigate households' portfolio choice within the mean-variance paradigm, assuming that households cannot take short positions in risky assets, and that they have heterogeneous expectations on the distribution of asset excess returns and risk aversion.

The heterogeneity of households' expectations about the future performance of financial markets has been widely documented using probabilistic expectations data. Recent research has focused on the likely presence of different *expectation types* (Dominitz and Manski, 2011), on how expectations react to sudden downturns of the market (Hudomiet et al., 2011), on their relation with the experienced performance in the financial markets (Hurd et al., 2011) and on the quality of the information regarding past performance (Arrondel et al., 2012). Overall, expectations have been found to be only moderately correlated with the level of the financial market indexes, while they largely depend on personal investment experience and characteristics.

There is also established empirical evidence documenting the heterogeneity of risk attitude among households. Evidence comes from laboratory and field experiments (e.g., Andersen et al., 2008; Choi et al., 2007; Dohmen et al., 2010; Schubert et al., 1999; von Gaudecker et al., 2011), and more frequently from survey data. In particular, a few works derive proxies for risk attitude using different types of survey data: the observed portfolio composition, in the form of risky assets holdings (Morin and Fernandez Suarez, 1983; Siegel and Hoban, 1982; Riley and Chow, 1992) or of variance of portfolio returns (Bucciol and Miniaci, 2011), the choice in a hypothetical lottery (Donkers et al., 2001; Guiso and Paiella, 2008; Kimball et al., 2009), and the self-assessed attitude toward taking risks (whose reliability has been assessed by Dohmen et al., 2011). These works generally find that risk aversion is negatively correlated with wealth and high education, while its correlation with age is unclear – possibly because most of these studies are based on a single cross-section of data, which does not allow to properly disentangle age from time and cohort effects.

The information on expectations and risk attitude are often used as predictors in reduced form models of financial market participation (e.g. Arrondel et al., 2010, 2012; Hurd et al., 2011; Dohmen et al., 2011) but, to the best of our knowledge, they have not been linked to a structural model of household portfolio choice. Here instead, in a similar vein to Miniaci and Pastorello (2010), we use a mean-variance portfolio model where households can choose between one riskless asset (deposits) and two risky assets (bonds and stocks), and cannot hold short positions. The framework admits four possible regimes, in which households invest in no risky asset, only in bonds, only in stocks, or in both. The regime choice and the asset demands depend on the parameters of the distribution of stocks and bonds expected returns together with the risk aversion of the investors. We estimate such parameters in presence of observable and unobservable heterogeneity using a combination of tobit and ordered probit models which exploit information on actual portfolios and self-assessed willingness to bear financial risks. The model is estimated by weighted maximum likelihood using repeated survey data from the US Survey of Consumer Finances. We then simulate the model to evaluate the effects of demographic, wealth and income changes on risk attitude and expectations.

This work extends Miniaci and Pastorello (2010) in two important directions. First, we combine information on observed portfolio holdings with self-assessed willingness to take financial risk. This allows us to disentangle heterogeneity in expectations and risk attitude, something that was not possible in previous work relying on portfolio data only. In fact, risk aversion has no impact on the choice of whether to hold risky assets; in contrast it does play an important role in the decision of how much to invest in risky assets, but nevertheless cannot be separately identified if the only available information comes from portfolio data. Second, we use five repeated cross-sectional datasets (from 1995 to 2007), rather than one single survey. This allows to remove the dependency of the results on specific events occurring in a given year, and to disentangle age from time and cohort effects.

Our main results can be summarized as follows. We find a positive relation of wealth and age with expectations on market returns, which are highest among investors holding both bonds and stocks. In addition, market expectations change over time, and in particular they fall after periods of growing markets. This suggests that expectations are apparently consistent with the hypothesis of mean reversion. In contrast, they are not clearly correlated with movements in risk aversion. Moreover, once estimated together with expectations, risk aversion is found to be little influenced by some key household characteristics, such as age and wealth.

The remainder of this paper is organized as follows. Section 2 presents the framework, and Section 3 describes the data. Section 4 shows the model estimates, and discusses its implications. Section 5 concludes. A final Appendix provides details on the loglikelihood function and the identification of the parameters.

2 The framework

Our framework has two components. The first one, discussed in Section 2.1, focuses on household portfolio choice and is based on the Markowitz mean-variance setup with short-selling restrictions on risky asset shares. The second ingredient, discussed in Section 2.2, deals with the self-assessed level of risk attitude. Finally, Section 2.3 presents our assumptions on parameter heterogeneity.

2.1 Portfolio choice with short-selling restrictions

Following Miniaci and Pastorello (2010), we consider a variant of the standard Markowitz (1959) mean-variance single period framework, in which households face short-selling restrictions on the asset shares. The household chooses the optimal allocation of its wealth W_0 among one riskless asset and two risky assets. We assume that households are price takers, and that they are subject to two types of constraint: a standard budget constraint, and two non-negativity constraints corresponding to the shares of wealth allocated to the risky assets. This means that households cannot take a short position in any of the risky assets.

Let r_0 denote the interest rate on the riskless asset, and $\mathbf{e}=(e_1,e_2)'$ be the random vector of the excess returns on the risky assets. If $\mathbf{a}=(a_1,a_2)'$ is the vector of shares of financial wealth invested in the risky assets, the random wealth at the end of the period is given by $W_1=W_0(1+r_0+\mathbf{a}'\mathbf{e})$. Investors' preferences are described by a standard mean-variance expected utility function, $E\left[U(\mathbf{a})\right]=E(W_1)-\frac{\eta}{2}V(W_1)$, where $\eta>0$ is a risk aversion index varying across households. By construction, the moments of W_1 depend on those of excess returns vector $\mathbf{e}$. If we let

$$\boldsymbol{\mu}=\begin{bmatrix}\mu_1\\ \mu_2\end{bmatrix}=E(\mathbf{e});\quad \boldsymbol{\Omega}=\begin{bmatrix}\omega_1^2 & \omega_{12}\\ \omega_{12} & \omega_2^2\end{bmatrix}=V(\mathbf{e}),$$

we can rewrite the mean-variance expected utility function as:

$$E\left[U(\mathbf{a})\right] = (1+r_0)W_0 + W_0\left(\mathbf{a}'\boldsymbol{\mu} - \frac{\eta W_0}{2}\mathbf{a}'\boldsymbol{\Omega}\mathbf{a}\right).$$

The household portfolio selection problem can then be formally stated as follows:

$$\max_{\mathbf{a}\in\mathbb{R}_+^N}\left\{\mathbf{a}'\boldsymbol{\mu} - \frac{\eta W_0}{2}\mathbf{a}'\boldsymbol{\Omega}\mathbf{a}\right\}. \tag{1}$$

Before solving the constrained problem (1), let us define $\rho = \omega_{12}/\left(\omega_1\omega_2\right)$ the correlation coefficient between the two excess returns, and $\pi_1 = \mu_1/\omega_1$ and $\pi_2 = \mu_2/\omega_2$ the associated Sharpe performances. With two assets there are four possible regimes, denoted $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$, where a 0 in the first (second) position indicates that the first (second) asset is not included in the portfolio, while 1 indicates inclusion.

Consider first the solution of the unconstrained problem for the set of two assets, i.e. regime $(1,1)$. The optimal shares are given by:

$$\mathbf{a}_*^{(1,1)} = \frac{1}{\eta W_0(\omega_1^2\omega_2^2 - \omega_{12}^2)}\begin{bmatrix}\omega_2^2\mu_1 - \omega_{12}\mu_2 \\ -\omega_{12}\mu_1 + \omega_1^2\mu_2\end{bmatrix} = \frac{1}{\eta W_0(1-\rho^2)}\begin{bmatrix}(\pi_1 - \rho\pi_2)/\omega_1 \\ (\pi_2 - \rho\pi_1)/\omega_2\end{bmatrix}. \tag{2}$$

The (1,1) regime is the solution of the constrained problem (1) if the two above quantities are strictly positive. In terms of expected Sharpe performances, this is equivalent to requiring $\pi_1 > \rho\pi_2;\quad \pi_2 > \rho\pi_1$. Notice that these two conditions cannot hold at the same time if both π_1 and π_2 are negative, irrespective of the sign of ρ.

Consider now the $(1,0)$ regime. In this case,

$$\mathbf{a}_*^{(1,0)} = \begin{bmatrix}\dfrac{\mu_1}{\eta W_0\omega_1^2} \\ 0\end{bmatrix} = \begin{bmatrix}\dfrac{\pi_1}{\eta W_0\omega_1} \\ 0\end{bmatrix}.$$

This regime is the solution of the constrained problem if $\pi_2 < \rho\pi_1$. and $\pi_1 > 0$ (that is, $\mu_1 > 0$). A similar argument can be made to the analysis of the $(0,1)$ regime, for which:

$$\mathbf{a}_*^{(0,1)} = \begin{bmatrix} 0 \\ \dfrac{\mu_2}{\eta W_0 \omega_2^2} \end{bmatrix} = \begin{bmatrix} 0 \\ \dfrac{\pi_2}{\eta W_0 \omega_2} \end{bmatrix}.$$

This regime is the solution of (1) if $\pi_1 < \rho\pi_2$. and $\pi_2 > 0$. Finally, the $(0,0)$ regime holds whenever both π_1 and π_2 are negative.

It is worth noticing that in this framework the choice of the portfolio regimes conveys information on the heterogeneity of return expectations, not on the heterogeneity in risk aversion. More precisely, the regime choice is fully determined by the relation between the Sharpe ratios (π_1 and π_2) and the correlation (ρ). The variances (ω_1 and ω_2) and the investor's risk aversion η are only relevant for the portfolio shares in each regime (except the (0,0) case). The expressions for $\mathbf{a}_*^{(1,1)}$, $\mathbf{a}_*^{(1,0)}$ and $\mathbf{a}_*^{(0,1)}$ suggests that, even if we knew π_1, π_2 and ρ, ω_1, ω_2 and η could not be separately identified without further additional information. Our strategy is to use the self-assessed willingness to bear financial risk provided by the US *Survey of Consumer Finances* as valuable source of information on the risk aversion parameter η.

2.2 Risk attitude

Nowadays, a few household surveys include hypothetical lottery questions trying to elicit the attitude toward financial risk directly from the respondents, and some works already use these responses to infer the correlation between risk attitude and socio-demographic and economic characteristics (see, e.g., Barsky et al., 1997; Kimball et al., 2008, on the US *Health and Retirement Study*; Kimball et al., 2009, on the US *Panel Study of Income Dynamics*; Guiso and Paiella, 2008, on the Italian *Survey of Household Income and Wealth*; and Donkers et al., 2001, on a precursor of the *Dutch Household Survey*). An alternative strategy is to directly ask the households how much risk they are willing to bear in financial investments. In the *Survey of Consumer Finances* (SCF) the question reads as follows:

"Which of the following comes closest to describing the amount of financial risk that you [and your husband/ wife/ partner] are willing to take when you save or make investments?"

1. *Take substantial financial risks expecting to earn substantial returns*
2. *Take above average financial risks expecting to earn above average returns*
3. *Take average financial risks expecting to earn average returns*
4. *Not willing to take any financial risks*

This is a general question on the self-assessed level of risk aversion in the financial domain, and the answers are clearly influenced by the personal beliefs of the respondent on what "substantial", "above average" and "average" financial risks mean. Although very simple, questions like this proved to deliver information consistent with the one derived from more sophisticated elicitation methods as the paid lottery choices (see Dohmen et al., 2011) and the answers are usually found to be highly correlated with a wide range of objective measures of risk-related behavior (see, e.g., Bucciol and Miniaci, 2011, for the SCF case).

In what follows we assume that households declaring a higher willingness to bear financial risks are consistently less risk averse than the others, that is, households choosing a higher value in the SCF question are more risk averse. More precisely, let us denote with c the option chosen by the household when asked about its willingness to bear financial risk, with $c=1,\dots,L$ where a higher value of c corresponds to a lower willingness to bear financial risks. We assume that there exists a continuous latent variable $z \in \mathbb{R}$ such that a specific value of c is observed if $\delta_{c-1} < z \le \delta_c$ where $-\infty < \delta_1 < \ldots < \delta_{L-1} < +\infty$ are $L-1$ unknown threshold parameters to be estimated. The household reported c is related to its risk aversion η by $z = \log \eta + \epsilon_z$, where ϵ_z is a zero mean random error independent of η.

2.3 Heterogeneity assumption

The above results show that the shares invested in the risky assets depend on π_1, π_2, ρ, η, ω_1 and ω_2, while the reported risk attitude depends on z, and hence on η. As households differ in various respects, it seems natural to introduce heterogeneity in the model through these quantities. In this paper we make the assumption that π_1, π_2, ρ and η are heterogeneous across households. For tractability reasons we consider ω_1 and ω_2 unknown parameters to be estimated, but we do not allow them to be heterogeneous across households. More specifically, let $y = \tan\left(\frac{\pi}{2}\rho\right)$ and assume that

$$\begin{bmatrix} \pi_1 \\ \pi_2 \\ y \\ z \end{bmatrix} \Bigg| \mathbf{x} \sim N\left(\begin{bmatrix} \mathbf{x}_1'\boldsymbol{\beta}_1 \\ \mathbf{x}_2'\boldsymbol{\beta}_2 \\ \mathbf{x}_y'\boldsymbol{\beta}_y \\ \mathbf{x}_z'\boldsymbol{\beta}_z \end{bmatrix}, \begin{bmatrix} \sigma_1^2 & r_{12}\sigma_1\sigma_2 & r_{1y}\sigma_1\sigma_y & 0 \\ r_{12}\sigma_1\sigma_2 & \sigma_2^2 & r_{2y}\sigma_2\sigma_y & 0 \\ r_{1y}\sigma_1\sigma_y & r_{2y}\sigma_2\sigma_y & \sigma_y^2 & 0 \\ 0 & 0 & 0 & \sigma_z^2 \end{bmatrix} \right) \tag{3}$$

where $\mathbf{x} = \left[\mathbf{x}_1', \mathbf{x}_2', \mathbf{x}_y', \mathbf{x}_z'\right]'$. Notice that (3) implies $\ln(\eta) = \mathbf{x}_z'\boldsymbol{\beta}_z$, i.e. that, contrary to the Sharpe performances π_1 and π_2 and the correlation coefficient ρ, heterogeneity in η is only due to the observable households characteristics, rather than to some unobserved heterogeneity component. This choice helps keeping the model estimation manageable. The zero correlation between (π_1, π_2, y) and z also contributes to the tractability of the model, but we feel that this assumption is fairly reasonable. In fact ϵ_z cannot be interpreted as unobserved heterogeneity in the risk aversion parameter η, but rather as heterogeneity in answer styles and in the relation between households' understanding of the "financial risk" and their risk aversion.

We estimate the model via weighted maximum likelihood. To attain identification, however, we set to one the variance of the random component of π_1 ($\sigma_1 = 1$) and to zero the lowest threshold in the ordered probit part of the model (i.e., $\delta_1 = 0$). These constraints are sufficient to identify all the remaining parameters, but in order to improve the accuracy of the estimates we also set σ_z at 1 as it is usually done in ordered probit analysis. Further details on the identification and estimation issues are provided in the Appendix.

3 The data

Our analysis is based on data from the US *Survey of Consumer Finances* (SCF). The SCF is a repeated cross-sectional survey of households conducted every three years since 1983 on behalf of the Federal Reserve Board. Its purpose is to collect detailed information on assets and liabilities, together with income and the main socio-demographic characteristics of a sample of US households. In this work we will consider the five waves from 1995 to 2007, i.e., the most recent waves available at the moment of this study. We neglect previous waves mainly to analyze observations with similar investment options.[1] Our final sample consists of 19,408 observations, roughly equally distributed over the five waves.

Given its characteristics, the SCF has been widely used to investigate asset allocation. In our model of portfolio choice the endowment available to households is made of one risk free asset (*deposits*), which includes checking, savings and brockerage accounts, and two risky assets. The first risky asset (*bonds*) includes the market value of corporate and government bonds, bond mutual funds, half the value of balanced mutual funds, certificates of deposits, savings bonds, and life insurances; the second risky asset (*stocks*) includes directly held stocks, stock mutual funds, and half the value of balanced mutual funds. Respondents also report the investment allocation of any composite asset holdings (IRA-KEOGH accounts, retirement accounts, annuities, and trust-managed accounts), which allows us to split them across our asset categories accordingly. We define financial wealth W as the sum of risk free and risky asset holdings.

The size of financial wealth almost doubled over the period under investigation, going from a median 19,649 USD in 1995 to 33,350 USD in 2007 (see Table 1). About 24% of the households

[1] For instance, 401(k) plans and other retirement assets, nowadays important in household portfolios, were not widespread until the mid-90s.

held neither stocks nor bonds in 1995. This fraction went down to 22% in 1998 and 2001 but climbed up to 30% in 2004 and 2007. Households exited both the bond and the stock markets after 2001: bond holders fell from 58.5% in 2001 to 54.1% in 2007, stock holders from 61.2% to 53.4%, mainly in the period 2001 to 2004. On average 39.4% of the households held a financial portfolio which included both stocks and bonds. Conditional on participation, the share of financial wealth held in stocks ranged between 43% (2007) and 57.3% (2001).

Table 1. Median financial wealth (2007 prices), household portfolio regimes, household portfolio shares conditional on holding the asset, willingness to bear financial risk above or at the average level. All the statistics are computed using population weights.

	Median financial wealth	Portfolio regime (percent)			Conditional portfolio shares (percent)		Willing to bear financial risk (percent)	
	(USD)	(1,0)	(0,1)	(1,1)	Bonds	Stocks	Above average	Average
1995	19,649	25.71	13.63	36.33	47.22	48.92	17.39	39.76
1998	31,397	20.68	16.64	40.97	44.07	53.96	24.11	40.80
2001	36,254	16.87	19.50	41.66	38.77	57.31	24.00	39.72
2004	31,897	15.08	15.19	39.67	29.65	42.25	19.98	40.75
2007	33,350	15.69	15.01	38.43	32.66	43.00	21.21	40.02
Total	29,932	18.54	15.99	39.44	38.37	49.08	21.36	40.22

For our purpose, the other important piece of information conveyed in the SCF is the self-assessed question on willingness to bear financial risks described in Section 2.2. Most frequently households responded category 3 (willing to *"Take average financial risks expecting to earn average financial returns",* 40.22% of the sample) or 4 (*"Not willing to take any financial risk"*, 38.42%). Given the few households willing to "*take substantial financial risks expecting to earn substantial returns"* (label 1, 3.92% overall) we merge the two most risk prone categories together. In the end, we have three possible levels of stated willingness to bear risk. Its distribution is relatively stable across the waves, although we observe moderately higher willingness to take risks in the years 1998 and 2001 in correspondence with the highest stock market participation rates during the period under analysis.

Households not willing to bear any financial risk were substantially less wealthy than the remaining 60 percent of the sample: their median financial wealth (7,330 USD) was about one tenth of the median financial wealth of the most risk prone households (see Table 2). About 78% of the households in this latter category held stocks, and only 13.5% did not hold any risky asset. On the contrary, among households declaring not to be willing to bear any risk, 41.9% were consistent with their stated preferences and held only deposits, but a noticeable 30.7% held some stocks. Their self-assessed attitude toward financial risk is associated to a higher fraction of wealth invested in bonds.

Table 2. Median financial wealth (2007 prices), age, household portfolio regimes, household portfolio shares conditional on holding the asset. All the statistics are computed using population weights.

Willingness to bear financial risk	Median financial wealth	Age	Portfolio regime (percent)			Conditional portfolio shares (percent)	
	(USD)		(1,0)	(0,1)	(1,1)	Bonds	Stocks
Above average	70,100	44.98	8.90	22.57	55.07	27.15	56.61
At average	53,756	49.65	15.20	17.34	49.94	34.76	47.58
Not willing to bear any risk	7,330	55.63	27.40	10.93	19.75	52.07	41.93
Total	29,932	50.95	18.54	16.00	39.44	38.37	49.08

The analysis in Section 4 uses information on portfolio choice and the willingness to bear financial risk, taking as control variables further household information included in the SCF. Specifically, we consider age, gender, race, education level, marital status, and employment status of the household head; and financial wealth and total annual income of the household. To disentangle time and cohort effects from age effects our specification includes a second-order polynomial in age, year dummies and for the cohort effects we follow Ameriks and Zeldes (2004) and control for the excess return of the stock market when the head of the household was 20 to 24 years old. This variable should be able to control for the residual heterogeneity across cohorts with respect to their expectations and risk attitudes. We conjecture that, *ceteris paribus*, most of this heterogeneity is related to how good or bad were the market conditions at the age at which many individuals start taking independent financial decisions. This early experience can affect the way people learn how to process the available information and elaborate their own expectations as well as their perception of risk.

4 Estimation results

We present the weighted-ML estimates of the parameters values with their asymptotic standard errors in Table 3 and the conditional and unconditional predictions for π_1 and π_2 in Table 4.

The parameters in β_1 and β_2 describe the relation between the covariates and the marginal expected values for the bonds and stocks Sharpe performances ($E[\pi_1]$ and $E[\pi_2]$ respectively; we omit conditioning on **x** whenever it does not cause confusion). Given the restriction we imposed on some parameters, the fact that the average predicted unconditional Sharpe performance is equal to 0.76 for both assets is not particularly informative. It is more interesting to notice that the estimated parameters show that some covariates have similar effects on bonds and stocks expected Sharpe ratios, while others differ. Married couples, households headed by a retiree and/or white person, or by a female, all tend to be more optimistic on both assets returns, while self employed workers are more pessimistic than the others. Other variables affect the two unconditional expectations differently: expectations on bond returns always increase with age, while those on stock returns increase only up to age 48; graduates are positive with respect to the future performances of stocks but not of bonds.

Over time, people revise their expectations in an "asset specific" way: *ceteris paribus*, $E[\pi_1]$ is decreasing over the entire period while expectations on stocks were at their maximum in 2001. If we look at historical performances of bond and stock markets (Figure 1), it is worth noticing that only the 2001 data refer to a period with a negative trend for the stock index; the other surveys were carried out after periods of growing prices. The estimated time effects for stock returns suggest that investors tend to have lower expectations after periods of expanding markets, consistently with a mean reverting process of expectation formation for equity returns (see Dominitz and Manski, 2011). The result is also consistent with the evidence in Hudomiet et al. (2011), who show that the 2008 market crash caused an increase in the population average of expectations.

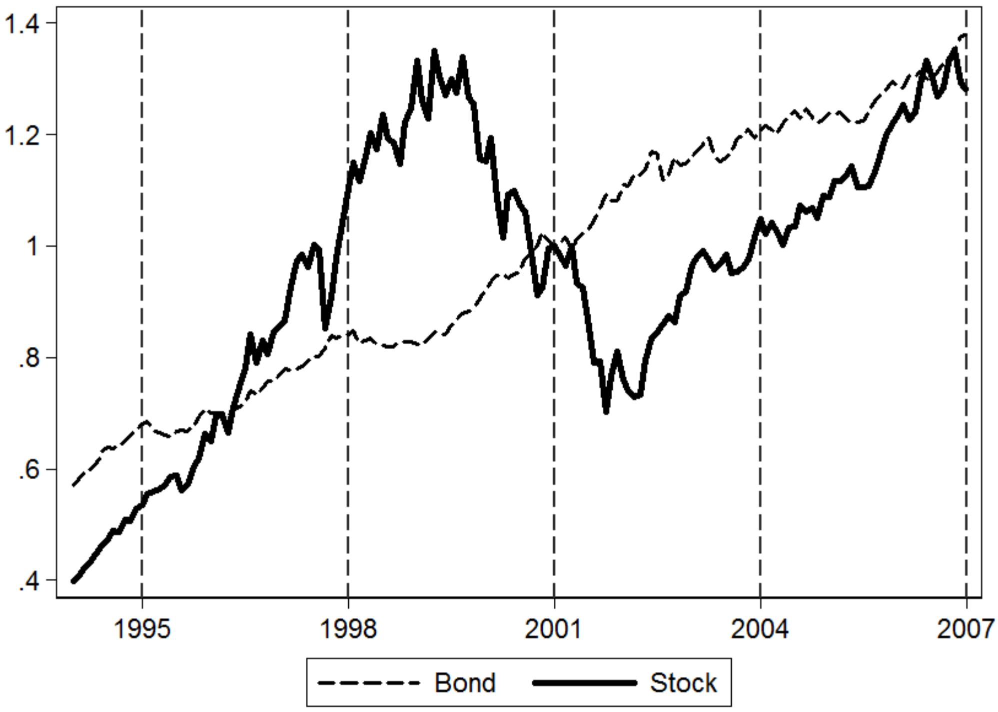

Figure 1. Historical US bond and stock indexes, standardized to 1 in December 2001, measured by the "Merrill Lynch US Corporate & Government Master Index" (bonds) and the "MSCI USA Stock Index" (stocks), respectively. Source: Datastream..

Wealth and income of investors are correlated with their views: a 10% increase in wealth brings an average 4% increase in the unconditional expected bonds Sharpe ratio and a 8% rise for stocks; a similar income increase raises $E[\pi_1]$ by 16% on average and $E[\pi_2]$ by 32%. The corresponding variations on the correlation $E[\rho]$ are instead negligible (0.8% for the wealth and 1% for income). The positive correlation between expected returns and investors' endowment has also been documented in probabilistic expectations data by Hurd et al. (2011).

We are cautious in interpreting the estimated relation between wealth and income and the expected returns. It is probably the case that being wealthier makes people more optimistic and therefore investing more in risky assets. However, the inverse relation may also be at work: being more optimistic (for any reason) might make people looking for higher (although riskier) returns, which makes them richer on average. It is difficult to assess to what extent this potential endogeneity issue affects our results. The use of predetermined levels of wealth and income might help to solve the problem, but this information is not available since the SCF is a repeated cross sectional survey.

Table 3. Weighted maximum likelihood of model (6). Wealth in million USD and income in hundred thousand USD (2007 prices). Max loglikelihood: -43056.4.

	π_1 (β_1)		π_2 (β_2)		$\tan(\rho/2)$ (β_y)		$\ln(\eta)$ (β_z)	
	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.	Coeff.	Std. Err.
Constant	-0.3932	0.0006	-0.7060	0.0006	1.2491	0.0014	0.2835	0.0002
Age/100	0.5789	0.0013	1.4507	0.0014	-1.1309	0.0055	0.5846	0.0008
$(\text{Age}/100)^2$	0.4568	0.0013	-1.5115	0.0019	1.5087	0.0055	-0.1110	0.0007
Married	0.1881	0.0003	0.0769	0.0002	-0.0868	0.0005	0.0915	0.0001
Graduate	-0.0111	0.0003	0.1853	0.0003	-0.0914	0.0004	-0.0354	0.0001
Self employed	-0.1082	0.0003	-0.3347	0.0003	-0.1223	0.0004	-0.0599	0.0001
Retired	0.0977	0.0004	0.0286	0.0005	-0.0748	0.0006	0.0141	0.0001
White	0.1397	0.0003	0.3288	0.0004	-0.0861	0.0005	0.0018	0.0001
Female	0.1594	0.0003	0.0502	0.0003	-0.0967	0.0005	0.0937	0.0001
Year = 1995	0.1566	0.0003	-0.1244	0.0003	0.1265	0.0005	0.0486	0.0001
Year = 1998	0.1080	0.0003	-0.0137	0.0003	0.0863	0.0005	-0.0019	0.0001
Year = 2004	-0.5079	0.0003	-0.5749	0.0003	-0.6782	0.0005	0.0381	0.0001
Year = 2007	-0.4790	0.0003	-0.5989	0.0003	-0.6851	0.0005	0.0090	0.0001
Cohort/100	-0.1441	0.0016	-0.1503	0.0017	-0.2013	0.0025	0.0507	0.0004
ln(1+wealth)	3.6127	0.0015	5.1089	0.0020	2.3222	0.0015	0.7357	0.0005
$\ln(1+\text{wealth})^2$	-1.1641	0.0008	-1.6401	0.0011	-0.7882	0.0008	-0.2419	0.0003
ln(1+income)	0.9356	0.0016	1.8956	0.0019	0.6076	0.0020	0.2351	0.0005
$\ln(1+\text{income})^2$	-0.4369	0.0008	-0.7982	0.0010	-0.2922	0.0011	-0.0911	0.0002
δ_2							1.0818	0.0001
ω_1	1.0555	0.0001						
ω_2	1.4419	0.0002						
σ_2	1.3409	0.0001	r_{12}	0.6995	0.0001	r_{2y}	0.5498	0.00005
σ_y	1.5500	0.0001	r_{1y}	0.7860	0.00005			

The ratio of the estimated return standard deviations ($\hat{\omega}_2 / \hat{\omega}_1$) is equal to 1.42, lower than the historical one (about 2.5 for the period considered). The variances of the unobserved heterogeneity component for π_2 and ρ are of comparable magnitude. The average unconditional predicted value of the correlation between bond and stock returns is 0.31, which makes the returns expected by the households more correlated than those observed in the market (-0.17 in our sample period). Finally, the correlations between the random components of π_1, π_2 and ρ are all estimated to be positive: the unobservable factors that make an investor positive with respect to the bond returns make her optimistic on the stock market as well.

The unconditional expectations $E[\pi_1]$ and $E[\pi_2]$ do not give us insights on the difference between the expectations of households who have chosen different types of portfolios. Further understanding can be gained looking at the conditional expectations. The model imposes (see section 2.1) $E[\pi_1 | (1,0)], E[\pi_2 | (0,1)] > 0$, $E[\pi_1 | (0,0)], E[\pi_2 | (0,0)] < 0$, and if $\rho > 0$ then $E[\pi_1 | (1,1)], E[\pi_2 | (1,1)] > 0$. In contrast $E[\pi_1 | (0,1)]$ and $E[\pi_2 | (1,0)]$ are unconstrained. In addition, we can use the model to predict the conditional expectations out of the portfolio regimes in which we observe the households. For a household which currently holds only riskless assets we can for instance assess how its expectations on bonds would be if it decided to hold only bonds (that is, if it were in the regime (1,0) instead of being in (0,0)). The sample averages of these prediction are reported in Table 4.

The values along the diagonal describe the average predictions conditional on the portfolio regimes in which we observe the households. Here we see that – as expected – households holding only deposits have negative expectations on both stocks (-1.3) and bonds (-0.79). Households investing only in bonds also have pessimistic views on the stock returns (with an average estimated $E[\pi_2 | (1,0)]$ equal to -0.23). In contrast, those investing only in stocks have positive expectations also on bonds (mean estimated $E[\pi_1 | (0,1)]$ equal to 0.43), but not sufficiently high to make them choose a complete portfolio. The conditional expected value of the correlation between stock and bond returns varies remarkably across portfolio regimes; in particular the means of the predicted values are always positive except for the households not participating to the risky financial market.

From the values out of the diagonal we learn that the households with riskless portfolios would have lower expectations than the other households even when moved in similar regimes.

Households holding only bonds or only stocks have instead expectations whose means are very close to each other. Finally, the investors with complete portfolios (1,1) have on average more optimistic views than the others. It is also interesting to notice that on average $E\left[\pi_1 \mid (1,0)\right] < E\left[\pi_1 \mid (1,1)\right]$, that is, that expectations on bond returns when holding only bonds are on average lower than when the portfolio includes also stocks, while the reverse is true for stocks ($E\left[\pi_2 \mid (0,1)\right] > E\left[\pi_2 \mid (1,1)\right]$): if the only risky investment is in stocks, then their returns are expected to be higher than in the case of complete portfolios.

Table 4. Conditional and unconditional predictions for bonds and stocks Sharpe performances.

Observed regime	$E\left[\pi_1 \mid (0,0)\right]$	$E\left[\pi_1 \mid (1,0)\right]$	$E\left[\pi_1 \mid (0,1)\right]$	$E\left[\pi_1 \mid (1,1)\right]$
$d=(0,0)$	-0.7889	0.6919	0.1546	0.8759
$d=(1,0)$	-0.6923	0.8890	0.4054	1.1081
$d=(0,1)$	-0.7345	0.8734	0.4253	1.1009
$d=(1,1)$	-0.6390	1.2124	0.8280	1.4827
	$E\left[\pi_2 \mid (0,0)\right]$	$E\left[\pi_2 \mid (1,0)\right]$	$E\left[\pi_2 \mid (0,1)\right]$	$E\left[\pi_2 \mid (1,1)\right]$
$d=(0,0)$	-1.3003	-0.3992	1.1594	1.1567
$d=(1,0)$	-1.2291	-0.2338	1.4183	1.3320
$d=(0,1)$	-1.0563	-0.1240	1.5565	1.4435
$d=(1,1)$	-0.9450	0.1757	2.0600	1.8248

We now focus our attention on the risk aversion parameter η. Table 3 confirms most of the established results in the literature: considering a single non-white male employee as reference individual, risk aversion is higher for women (+9%) and married couples (+9%), and lower for self employed workers (-6%) and graduates (-3.5%). Ageing does not seem to have a great impact on risk aversion: adding 5 years to all the individuals in the sample (i.e. moving from the average of 51 years to 56) raises the average estimated η by only 0.05 from the original 2.23 (with an average percentage difference of 2.57%). Wealth and income effects are also negligible: in both cases a 10% increase generates an average percentage difference of 0.4%. It is relevant to notice that, if we estimated the parameters of $\ln(\eta)$ using a standard ordered probit model, we would obtain a negative and significant effect of income and wealth. The difference is therefore due to the fact that we are jointly estimating (and disentangling) the determinants of the risk aversion and those of the expectations. The unclear relation between risk, income and wealth is also documented, in a large sample and in a completely different framework, by von Gaudecker et al. (2011). Risk aversion instead seems to vary over time. *Ceteris paribus* in 1995 and in 2004 the average risk aversion was respectively 4.8% and 3% higher than in 2001, without a clear association with the movement of the expectations on the stock markets.

Finally, our control variable for the cohort effects sort out to be significant: one percentage point more when aged 20-24 years decreases the unconditional expected Sharpe ratios by about -0.15 (the average predicted value is 0.75) and increases the risk aversion parameter by 5%. Apparently positive initial conditions of the financial market make the investors more cautious later in their lives.

5 Conclusions

In this paper we develop a structural econometric model within the mean-variance portfolio framework to elicit US household-specific expectations about future financial market returns *and* risk attitude, using data on observed portfolio holdings and answers to self-assessed questions on the willingness to bear financial risk. The investment set is made of one riskless asset (deposits) and two risky assets (bonds and stocks); we assume that households cannot take short positions in risky assets. We derive an explicit solution of the model characterized by four possible portfolio regimes and three levels of self-assessed risk attitude. We then analyze the model using a combination of structural tobit and ordered probit estimation methods with three latent variables.

Our findings reveal a positive relation between wealth and optimistic expectations on market returns and a negligible association between investors' endowment and their risk aversion. In this respect, the use of a structural model leads to conclusions on expectations and risk preferences which are shared by works based on probabilistic expectations data and actual lotteries. The use of repeated cross sections of the SCF allows us to show that equity market expectations are apparently consistent with the hypothesis of mean reversion and are not clearly correlated with movements in risk preferences.

Future research will aim at enriching the current framework by including real assets in the definition of portfolio. These assets are usually a relevant component of household portfolios as shown, for instance, in Bucciol and Miniaci (2011). Even if one treats them as illiquid, their holding still influences in a non trivial way the portfolio decisions on financial assets.

References

Ameriks, J., and S. Zeldes (2004), "How Do Household Portfolio Shares Vary with Age?," mimeo, Columbia University.

Andersen, S., G.W. Harrison, M.I. Lau, and E.E. Rutstrom (2008), "Eliciting Risk and Time Preferences," *Econometrica*, 76(3), 583–618.

Arrondel, L., H. Calvo-Pardo, and X. Oliver (2010), "Temperance in Stock Market Participation: Evidence from France," *Economica*, 77(306), 314-333.

Arrondel, L., H. Calvo-Pardo, and D. Tas (2012), "Subjective Return Expectations, Information and Stock Market Partcipation: Evidence from France," mimeo, University of Southampton.

Barsky, R.B., F.T. Juster, M.S. Kimball, and M.S. Shapiro (1997), "Preference Parameters and Behavioral Heterogeneity: an Experimental Approach in the Health and Retirement Study," *Quarterly Journal of Economics*, 112(2), 537-579.

Bucciol, A., and R. Miniaci (2011), "Household Portfolios and Implicit Risk Preference," *The Review of Economics and Statistics*, 93, 1235-1250.

Bucks B.K., A.B. Kennickell, T.L. March and K.B. Moore (2009), "Changes in U.S. Family Finances from 2004 to 2007: Evidence from the Survey of Consumer Finances," *Federal Reserve Bulletin*, 95, A1-A55.

Calvet, L.E., J.Y. Campbell, and P. Sodini (2007), "Down or Out: Assessing the Welfare Costs of Household Investment Mistakes," *Journal of Political Economy*, 115(5), 707-747.

Choi, S., R. Fisman, D. M. Gale, and S. Kariv (2007), "Revealing Preferences Graphically: An Old Method Gets a New Tool Kit,", *American Economic Review*, 97(2), 153-158.

Christelis, D., T. Jappelli, and M. Padula (2010), "Cognitive Abilities and Portfolio Choice," *European Economic Review*, 54(1), 18-38.

Cocco, J.F. (2005), "Portfolio Choice in the Presence of Housing," *The Review of Financial Studies*, 18(2), 535-567.

Davis, S.J., F. Kubler, and P. Willen (2006), "Borrowing Costs and the Demand for Equity over the Life Cycle," *The Review of Economics and Statistics*, 88(2), 348-362.

Dohmen, T., A. Falk, D. Huffman, and U. Sunde (2010), "Are Risk Aversion and Impatience Related to Cognitive Ability?," *American Economic Review*, 100(3), 1238-60.

Dohmen, T., A. Falk, D. Huffman, U. Sunde, J. Schupp, and G. Wagner (2011), "Individual Risk Attitudes: Measurement, Determinants, and behavioral Consequences", *Journal of the European Economic Association*, 9(3), 522-550.

Dominitz, J., and C.F. Manski (2011), "Measuring and Interpreting Expectations on Equity Returns," *Journal of Applied Econometrics*, 26, 352-370.

Donkers, B., B. Melenberg, and A. Van Soest (2001), "Estimating Risk Attitudes Using Lotteries: A Large Sample Approach," *Journal of Risk and Uncertainty*, 22(2), 165-195.

Flavin, M., and T. Yamashita (2002), "Owner-Occupied Housing and the Composition of the Household Portfolio," *American Economic Review*, 92(1), 345-362.

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

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

Guiso, L., M. Haliassos, and T. Jappelli (2003), "Household Stockholding in Europe: Where Do We Stand and Where Do We Go?," *Economic Policy*, 18(36), 123-170.

Guiso, L., and T. Jappelli (2005), “Awareness and Stock Market Participation, “ *Review of Finance*, 9(4), 537-567.

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

Haliassos, M., and A. Michaelides (2003), “Portfolio Choice and Liquidity Constraints,” *International Economic Review*, 44(1), 143-177.

Heaton, J., and D. Lucas (2000), “Portfolio Choice and Asset Prices: The Importance of Entrepreneurial Risk,” *Journal of Finance*, 55(3), 1163-1198.

Hudomiet, P., G. Kedzi, and R. Willis (2011), “Stock Market Crash and Expectations of American Households,” *Journal of Applied Econometrics*, 26, 393-415.

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

Kimball, M.S., C.R. Sahm, and M.D. Shapiro (2008), “Imputing Risk Tolerance from Survey Responses,” *Journal of the American Statistical Association*, 103(483), 1028-1038.

Kimball, M.S., C.R. Sahm, and M.D. Shapiro (2009), “Risk Preferences in the PSID: Individual Imputations and Family Covariation,” *American Economic Review: Papers & Proceedings*, 99 (2), 363-368.

Markowitz, H. (1959), Portfolio Selection: Efficient Diversification of Investments. Wiley: New York.

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

Morin, R.A., and A. Fernandez Suarez (1983), “Risk Aversion Revisited,” *Journal of Finance*, 38(4), 1201-1216.

Riley, W.B., and K.V. Chow (1992), “Asset Allocation and Individual Risk Aversion,” *Financial Analysts Journal*, 48(6), 32-37.

Schubert, R., M. Brown, M. Gysler, and H.W. Brachinger (1999), “Financial Decision-Making: Are Women Really More Risk-Averse?,” *American Economic Review,* 89(2), 381-385.

Siegel, F.W., and J.P. Hoban (1982), “Relative Risk Aversion Revisited,” *Review of Economics and Statistics,* 64(3), 481-487.

Van Rooij, M., A. Lusardi, and R. Alessie (2011), “Financial Literacy and Stock Market Participation,” *Journal of Financial Economics*, 101(2), 449-472.

Viceira, L.M. (2001), “Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income,” *Journal of Finance*, 56(2), 433-470.

Vissing-Jorgensen, A. (2002), “Towards an Explanation of Household Portfolio Choice Heterogeneity: Non-financial Income and Participation Cost Structures”, NBER Working Paper No. 8884.

Von Gaudecker, H., A. van Soest, and E. Wengstrom (2011), “Heterogeneity in Risky Choice Behavior in a Broad Population,” *American Economic Review,* 101, 664-694.

Appendix A

A.1 Loglikelihood

Under assumption (3), the total loglikelihood is the sum of two components: the loglikelihood of the specification studying the portfolio data, and the loglikelihood of the ordered probit model for the risk attitude. The latter is standard: for any observed self-reported risk attitude $c = l$, the probability of observing it is given by

$$q^l(\boldsymbol{\theta}) = \Pr(c = l \mid \mathbf{x}) = \begin{cases} \Phi\left(\dfrac{\delta_1 - \mathbf{x}_z'\boldsymbol{\beta}_z}{\sigma_z}\right) & \textit{for} \quad l = 1 \\ \Phi\left(\dfrac{\delta_l - \mathbf{x}_z'\boldsymbol{\beta}_z}{\sigma_z}\right) - \Phi\left(\dfrac{\delta_{l-1} - \mathbf{x}_z'\boldsymbol{\beta}_z}{\sigma_z}\right) & \textit{for} \quad l = 2, \ldots, L-1 \\ 1 - \Phi\left(\dfrac{\delta_{L-1} - \mathbf{x}_z'\boldsymbol{\beta}_z}{\sigma_z}\right) & \textit{for} \quad l = L-1 \end{cases}$$

where $\boldsymbol{\theta} = \left[\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \boldsymbol{\beta}_y', \boldsymbol{\beta}_z', \sigma_1, \sigma_2, \sigma_y, \sigma_z, r_{12}, r_{1y}, r_{2y}, \omega_1, \omega_2, \boldsymbol{\delta}'\right]'$ is the vector collecting all the parameters, and $\Phi(\cdot)$ is the cdf of a standard Normal distribution; for later use we also denote with $\phi(\cdot)$ the corresponding pdf. Here and in the following discussion we simplify the notation by omitting the conditioning on $\mathbf{x}$ when this is not cause of confusion.

Portfolio data could lend themselves to either a probit analysis, exploiting only the qualitative dimension in the data (holding /not holding of a given asset), or a tobit one, exploiting also the quantitative information (portfolio share held in a given asset). As mentioned at the end of Section 2.1, however, risk aversion is irrelevant in the choice of the portfolio regime. If we limited the analysis to the qualitative portfolio data, then, the portfolio choice parameters $\left[\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \boldsymbol{\beta}_y', \sigma_1, \sigma_2, \sigma_y, r_{12}, r_{1y}, r_{2y}, \omega_1, \omega_2\right]'$ and the risk aversion parameters $(\boldsymbol{\beta}_z', \sigma_z, \boldsymbol{\delta}')'$ would be estimated separately. In contrast, with the quantitative portfolio data the $\boldsymbol{\beta}_z$ parameters defining η enter both loglikelihood components. For this reason in this paper we focus on the tobit specification exploiting the quantitative portfolio information.

To evaluate the likelihood contributions of the portfolio data under assumption (6), it is convenient to first compute them conditionally on y (or, equivalently, on ρ), and then integrate the resulting expressions relative to the marginal distribution of y, which is easily derived from (6). Notice however that this is not necessary for the probability of regime $(0,0)$. As the region over which density (6) has to be integrated ($\pi_1 < 0, \pi_2 < 0$) is independent on ρ, it is possible to directly compute the desired probability using the joint marginal cdf of (π_1, π_2). The result is then given by

$$l^{(0,0)}(\boldsymbol{\theta}) = \Phi_2\left(-\frac{\mathbf{x}_1'\boldsymbol{\beta}_1}{\sigma_1}, -\frac{\mathbf{x}_2'\boldsymbol{\beta}_2}{\sigma_2}; r_{12}\right).$$

For the remaining three regimes we proceed by first computing the relevant probabilities/densities conditional on y, and then dropping the conditioning by integrating the resulting expressions relative to the marginal distribution of y. For the (1,0) regime, and conditional on y, the portfolio likelihood contribution is the product of the density of the optimal share invested in the first asset, denoted by $a_{*1}^{(1,0)}$, evaluated at the observed share a_1, and the probability that $\pi_2 < \rho\pi_1$, conditional on $a_{*1}^{(1,0)} = a_1$. To evaluate these two quantities, notice that $a_{*1}^{(1,0)} = \pi_1 / (\eta W_0 \omega_1)$ according to Section 2.1; hence, using (6), we derive the joint distribution of $(a_{*1}^{(1,0)}, \pi_2)'$ as follows:

$$\left[\begin{array}{c} a_{1*}^{(1,0)} \\ \pi_2 \end{array}\right] \Bigg| y \sim N\left(\left[\begin{array}{c} \dfrac{\mu_{1|y}}{\eta W_0 \omega_1} \\ \mu_{2|y} \end{array}\right], \left[\begin{array}{cc} \dfrac{\sigma_{1|y}^2}{\eta^2 W_0^2 \omega_1^2} & \dfrac{r_{12|y}\sigma_{1|y}\sigma_{2|y}}{\eta W_0 \omega_1} \\ \dfrac{r_{12|y}\sigma_{1|y}\sigma_{2|y}}{\eta W_0 \omega_1} & \sigma_{2|y}^2 \end{array}\right]\right).$$

This result allows to derive the two elements of the contribution of the $(1,0)$ regime to the loglikelihood. The pdf of $a_{*1}^{(1,0)}$ evaluated at a_1 can be expressed as:

$$f^{(1,0)}(a_1, y; \boldsymbol{\theta}) = \frac{1}{\sigma_{1|y} / (\eta W_0 \omega_1)} \phi\left(\frac{a_1 - \dfrac{\mu_{1|y}}{\eta W_0 \omega_1}}{\dfrac{\sigma_{1|y}}{\eta W_0 \omega_1}}\right),$$

while the probability that $\pi_2 < \rho\pi_1 = \rho\eta W_0 \omega_1 a_1$, conditional on $a_{*1}^{(1,0)} = a_1$ and y, is given by:

$$p^{(1,0)}(a_1, y;\boldsymbol{\theta}) = \Phi\left(\frac{\rho\eta W_0\omega_1 a_1 - (\mu_{2|y} + \eta W_0\omega_1 \frac{r_{12|y}\sigma_{2|y}}{\sigma_{1|y}}(a_1 - \frac{\mu_{1|y}}{\eta W_0\omega_1}))}{\sigma_{2|y}(1-r_{12|y}^2)^{1/2}}\right).$$

To drop the conditioning on y, we must take the expectation of the product of these two quantities relative to the marginal distribution of y, which leads to the following likelihood contribution:

$$l^{(1,0)}(\boldsymbol{\theta}; a_1) = E^y[p^{(1,0)}(a_1, y;\boldsymbol{\theta}) f^{(1,0)}(a_1, y;\boldsymbol{\theta})].$$

In the same way, we obtain the expression of the likelihood contribution of the (0,1) observed regimes:

$$l^{(0,1)}(\boldsymbol{\theta}; a_2) = E^y[p^{(0,1)}(a_2, y;\boldsymbol{\theta}) f^{(0,1)}(a_2, y;\boldsymbol{\theta})],$$

where a_2 denotes the observed invested share in asset 2, and where

$$f^{(0,1)}(a_2, y;\boldsymbol{\theta}) = \frac{1}{\sigma_{2|y}/(\eta W_0\omega_2)}\phi\left(\frac{a_2 - \frac{\mu_{2|y}}{\eta W_0\omega_2}}{\frac{\sigma_{2|y}}{\eta W_0\omega_2}}\right)$$

$$p^{(0,1)}(a_2, y;\boldsymbol{\theta}) = \Phi\left(\frac{\rho\eta W_0\omega_2 a_2 - (\mu_{1|y} + \eta W_0\omega_2 \frac{r_{12|y}\sigma_{1|y}}{\sigma_{2|y}}(a_2 - \frac{\mu_{2|y}}{\eta W_0\omega_2}))}{\sigma_{1|y}(1-r_{12|y}^2)^{1/2}}\right)$$

respectively denote the (conditional on y) density of $a_{*2}^{(0,1)}$ evaluated at a_2, and the probability (conditional on y and $a_{*2}^{(0,1)} = a_2$) that $\pi_1 < \rho\pi_2$.

It remains to derive the likelihood contribution of the $(1,1)$ regime. In this case, two strictly positive invested shares are observed. Since their relation with π_1 and π_2 is given by (3), we get

$$\mathbf{a}_*^{(1,1)} \mid (\mathbf{x}, y) \sim \mathcal{N}\left(\begin{bmatrix} \dfrac{\mu_{1|y} - \rho\mu_{2|y}}{\eta W_0 \omega_1 (1-\rho^2)} \\ \dfrac{-\rho\mu_{1|y} + \mu_{2|y}}{\eta W_0 \omega_2 (1-\rho^2)} \end{bmatrix}, \begin{bmatrix} \dfrac{\sigma_{1|y}^2 + \rho^2\sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}{\eta^2 W_0^2 \omega_1^2 (1-\rho^2)^2} & \dfrac{r_{12|y}\sigma_{1|y}\sigma_{2|y}(1+\rho^2) - \rho(\sigma_{1|y}^2 + \sigma_{2|y}^2)}{\eta^2 W_0^2 \omega_1\omega_2 (1-\rho^2)^2} \\ \dfrac{r_{12|y}\sigma_{1|y}\sigma_{2|y}(1+\rho^2) - \rho(\sigma_{1|y}^2 + \sigma_{2|y}^2)}{\eta^2 W_0^2 \omega_1\omega_2 (1-\rho^2)^2} & \dfrac{\rho^2\sigma_{1|y}^2 + \sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}{\eta^2 W_0^2 \omega_2^2 (1-\rho^2)^2} \end{bmatrix}\right).$$

For the (1,1) regime, and conditional on y, the likelihood contribution is given by this bivariate gaussian pdf evaluated at the observed shares invested, $(a_1, a_2)'$:

$$f^{(1,1)}(a_1, a_2, y; \boldsymbol{\theta}) = \phi_2\left(\frac{\eta W_0\omega_1(1-\rho^2)a_1 - \mu_{1|y} + \rho\mu_{2|y}}{\sqrt{\sigma_{1|y}^2 + \rho^2\sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}}, \frac{\eta W_0\omega_2(1-\rho^2)a_2 + \rho\mu_{1|y} - \mu_{2|y}}{\sqrt{\rho^2\sigma_{1|y}^2 + \sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}}; \frac{r_{12|y}\sigma_{1|y}\sigma_{2|y}(1+\rho^2) - \rho(\sigma_{1|y}^2 + \sigma_{2|y}^2)}{\sqrt{\sigma_{1|y}^2 + \rho^2\sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}\sqrt{\rho^2\sigma_{1|y}^2 + \sigma_{2|y}^2 - 2\rho r_{12|y}\sigma_{1|y}\sigma_{2|y}}}\right).$$

Integration of this expression leads to the likelihood contribution of the observed $(1,1)$ regimes:

$$l^{(1,1)}(\boldsymbol{\theta}; a_1, a_2) = E^y[f^{(1,1)}(a_1, a_2, y; \boldsymbol{\theta})].$$

All calculations in this paper were made in Fortran. The numerical evaluation of the different likelihood contributions involves the evaluation of a bivariate Gaussian cdf, and one-dimensional quadrature techniques. To this end we use Fortran routines based on high-order polynomial approximation to Mills ratio, and 64 (the maximum allowed) Gauss-Hermite weights and abscissae. To maximize the loglikelihoods we used the Nelder and Mead derivative-free simplex algorithm. This method shows to better handle the high dimension of the parameter vector than the derivative-based quasi Newton algorithm.

A.2 Parameters identification

Inspection of the above expressions shows that all the probabilities and densities involved in both likelihood components remain unchanged if $\boldsymbol{\beta}_1$, $\boldsymbol{\beta}_2$, σ_1, σ_2, ω_1 and ω_2 are multiplied by the same constant. For this reason, we set $\sigma_1 = 1$. The same identification issue arises when any constant is added to $\boldsymbol{\delta}$ and to the intercept in $\boldsymbol{\beta}_z$, and multiplied after exponentiation to $\boldsymbol{\beta}_1 / \omega_1$, $\boldsymbol{\beta}_2 / \omega_2$, σ_1 / ω_1 and σ_2 / ω_2. To attain identification, we set $\delta_1 = 0$. Under these constraints, all the remaining parameters in $\boldsymbol{\theta}$ are identified; in particular, contrary to the standard ordered probit model, σ_z is in principle uniquely determined by the nonlinear occurrence of η in the portfolio component of the loglikelihood. In practice, however, some preliminary results suggested that the accuracy of the estimates could be greatly increased by fixing σ_z at 1 as it is usually done in ordered probit analysis. We also think that this constraint does not pose any relevant limitation in terms of the interpretability of the results, as σ_z is not a central parameter in our analysis. Taken all things together, we collect in $\boldsymbol{\theta} = \left[\boldsymbol{\beta}_1', \boldsymbol{\beta}_2', \boldsymbol{\beta}_y', \boldsymbol{\beta}_z', \sigma_2, \sigma_y, r_{12}, r_{1y}, r_{2y}, \omega_1, \omega_2, \delta_2, \ldots, \delta_{L-1}\right]'$ the set of free parameters to be estimated.

Let M denote the number of households in the sample, and let:

$$\mathbb{I}_m^{(i,j)} = \begin{cases} 1 & \text{if household } m \text{ chooses regime } (i, j) \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{J}_m^{(l)} = \begin{cases} 1 & \text{if household } m \text{ reports } c_m = l \\ 0 & \text{otherwise} \end{cases}$$

for $(i, j) \in \{(0,0), (1,0), (0,1), (1,1)\}$, and $l \in \{1, \ldots, L\}$. Using this notation, the loglikelihood function can be written as

$$\ell(\boldsymbol{\theta}) = \sum_{m=1}^{M} \left(\sum_{i=1}^{2} \sum_{j=1}^{2} \mathbb{I}_m^{(i,j)} \log l^{(i,j)}(\boldsymbol{\theta}; \mathbf{a}_m) + \sum_{l=1}^{L} \mathbb{J}_m^{(l)} \log q^{(l)}(\boldsymbol{\theta}) \right).$$

For details on the regime probabilities and the expected asset demands please refer to the appendix in Miniaci and Pastorello (2010).