DO DISASTER EXPECTATIONS EXPLAIN HOUSEHOLD PORTFOLIOS?[1]

Sule Alan

Faculty of Economics and CFAP

University of Cambridge

Cambridge, United Kingdom

E-mail: sule.alan@econ.cam.ac.uk

December 2008

[1]I thank Thomas Crossley, Hamish Low and seminar participants at the University of Cambridge and University College London for valuable comments and suggestions. I also thank Ruxandra Dumitrescu for excellent research assistance. This research is partly funded by the Center for Financial Analysis and Policy (CFAP).

Abstract

It has been argued that rare economic disasters can explain most asset pricing puzzles. If this is the case, perceived risk associated with a disaster in stock markets should be revealed in household portfolios. That is, the framework that solves these pricing puzzles should also generate quantities that are consistent with the observed ones. This paper estimates the perceived risk of disasters (both probability and expected size) that is consistent with observed portfolios and consumption growth between 1983 and 2004 in the United States. I find that the disaster probabilities that justify life cycle profiles of stock market participation, equity shares and consumption growth are in the neighborhood of 4 percent per year. The estimated disaster probabilities vary across cohorts without any systematic pattern. Older cohorts expect such financial disasters to be more devastating. I also find that participation costs are not needed to match observed participation profiles once disaster expectations are allowed.

1 Introduction

Following Mehra and Prescott 's seminal 1985 article, a large body of research has accumulated which proposes solutions to the "equity premium puzzle." Various strands of the literature consider preference re-specifications (Campbell and Cochrane (1999), Bansal and Yaron (2004)), market frictions and preference heterogeneity (Constantinides et al. (2002)), and model uncertainty (Weitzman (2008)). An alternative strand of the literature emphasizes the limitations of the post-war historical return data. The observed equity premium can be rationalized if the standard model takes into account the possibility of rare but disastrous market events (such as occurred before the post-war period). This idea was first proposed by Reitz (1988) and extended by Barro (2006) and Barro and Ursua (2008). Barro (2006) analyses 20th century disasters using GDP and stock market data from 35 countries. He suggests that a disaster probability of $1.5-2$ percent a year, with an associated decline in per capita GDP of $15-64$ percent from peak to trough, goes a long way in explaining the equity premium puzzle. In follow up work using aggregate consumption data from 21 countries, Barro and Ursua (2008) calibrate the disaster probability to 3.6 percent a year with an associated 22 percent decline in consumption from peak to trough. More recently, Gabaix (2008) proposes a framework in which disasters have varying intensity. This framework can explain, in addition to the equity premium puzzle, many other asset pricing puzzles such as excess volatility, the value premium and the upward sloping nominal yield curve.

The equity premium puzzle has a spectacular manifestation in household micro data: most recent empirical evidence suggests that at least fifty percent of households in any developed country do not hold equities directly or indirectly (the stock market participation puzzle). Moreover, in contrast to the predictions of the standard model, we observe a great deal of heterogeneity in the share of risky assets (stocks) in household portfolios even after conditioning on stock market participation and controlling for income and wealth (see Bertaut 1998 and Guiso et al (2002)). Given the rather impressive equity premium in the post-war period, a particular difficulty in reconciling the standard model with observed facts is in explaining why younger households often hold both risk-free and risky assets. In its standard form, life cycle portfolio theory with labor income risk and return uncertainty predicts that households who are early in their life cycle should take advantage of the high equity premium and hold large positions in stocks. In fact, the model often predicts a 100 percent share of stocks in the financial portfolios of young investors (the portfolio specialization or small saver puzzle).

Like the asset pricing literature, the literature on household portfolios tries to explain these puzzles by proposing different preference specifications, transaction costs (stock market participation costs) and borrowing constraints (see Haliassos and Michaelides (2003), Cocco et al (2005), Gomes and Michaelides (2005), Alan (2006)). Unfortunately, none of these approaches provides a satisfactory explanation for household portfolio choice, consumption and financial wealth accumulation decisions jointly. Rather, they provide a patchwork, explaining some aspects of behavior (such as non-participation (Haliassos and Michaelides (2003) and Alan (2006), or portfolio shares (Gomes and Michaelides (2005)) at the expense of making counterfactual predictions about other data features[1]. The key reason for this failure is the strong functional link between risk aversion and prudence in utility functions. Specifically, the high risk aversion that is needed to make stocks unattractive (and hence to generate a high risk premium in equilibrium) triggers the precautionary motive, that in turn generates more savings and more stock market investments.

This paper is motivated by the idea that if rare economic disasters can solve the pricing puzzles they should also explain the observed quantities (household portfolio holdings). I show that this is in fact the case. Once the possibility of rare but disastrous market downturns is incorporated, observed household portfolio holdings can be reconciled with standard portfolio choice models. Note that this means dropping the usual assumption that the post-war historical premium is a good proxy for expected returns. Moreover, this reconciliation is achieved within the original CRRA framework, and without heightening risk aversion to implausible levels (and therefore without implying implausible wealth accumulation and consumption growth).

This exercise could be done in two ways. One would be to take historically calibrated values for the probability of disasters and expected size (from, for example, Barro (2006)) and apply them to a life cycle model with assumed preference parameter values to show how close one can get to observed life cycle profiles. Instead, I choose to jointly estimate disaster expectations and preference parameters from ob-

[1]The failure of these attempts is the inability to match life cycle profiles of portfolio shares conditional on stock market participation for any reasonable assumption on risk aversion. For example, using Epstein-Zin preferences one can generate low stock market participation and low unconditional portfolio shares by heightening risk aversion, but the model still generates counterfactually high portfolio shares for stock market participants early in the life cycle (the small saver puzzle).

served portfolios and then judge whether the estimates are plausible as compared to the historically calibrated values. Using a simulated minimum distance estimator (a variant of indirect inference), I estimate perceived disaster probability and size, along with intertemporal allocation parameters, for different birth cohorts and education groups. To do this, I set up a fairly rich version of a life cycle portfolio choice model where households take into account the possibility of a market disaster. Following Reitz (1988) and Barro (2006), the model assumes that disasters strike in an i.i.d fashion. Moreover, if a disaster strikes, households face significant labor market stress (in particular, a chance of zero labor income for a year). The model also incorporates a stock market participation cost and preference heterogeneity. These are ingredients proposed in the literature to solve portfolio puzzles. Incorporating these proposed remedies along with disaster expectations allows me to statistically test their relative importance in explaining household portfolio choices.

The estimation involves bringing together three large surveys conducted in the United States: the Survey of Consumer Finances (1983-2004) that contains detailed wealth and portfolio allocation information, the Consumer Expenditure Survey (1983-2004) that contains detailed durable and non durable expenditure information and finally, the Panel Study of Income Dynamics (1983-1994) that allows me to calibrate group specific income process parameters. Limited heterogeneity in all parameters is allowed for by estimating the structural parameters separately for 6 groups (3 cohorts by 2 education levels). I also allow for unlimited discount rate heterogeneity within groups.

Allowing for a possible disaster in the stock market improves the fit of the standard model substantially. Simulated household portfolios closely track their real data counterparts in most cases. The empirical estimates suggest perceived disaster probabilities in the neighborhood of 4 percent per year and a decline in stock market wealth of around 55 percent in the case of a disaster. These estimates line up well with Barro and Ursua (2008). The size estimates can be thought of as the total loss associated with a disaster, because returns are modelled as i.i.d and disasters are assumed to last only one year. There is no systematic pattern across education-birth cohort groups in estimated disaster probabilities. However, the estimates imply that older cohorts expect financial disasters to be more devastating (larger size estimates). The estimates of the coefficient of relative risk aversion range from 0.93 to 2.19. These values are in line with studies that estimate this parameter using micro data on consumption. Estimated median discount rates are around 10 percent. Finally, estimated per-period participation costs are zero for all groups. This means that household portfolio behavior is driven mainly by the perceived stock market risk rather than by participation costs.

The reminder of the paper is organized as follows: The next section presents the structural model used in the estimation. Section 3 discusses the estimation method and identification issues. Section 4 presents the data and results. Section 5 concludes.

2 The Model

I assume that the expected utility function is intertemporally additive over a finite lifetime and the sub-utilities are iso-elastic. The problem of the generic consumer h is

$$\max E_t \left[\sum_{j=0}^{T-t} \frac{(C_{h,t+j})^{1-\gamma}}{1-\gamma} \frac{1}{(1+\delta_h)^j} \right] \tag{1}$$

where C is non-durable consumption, γ the is coefficient of relative risk aversion (homogenous within a group), δ_h is the household specific rate of time preference which is assumed to be distributed lognormally across households such that $ln\delta_h \sim N(\mu_\delta, \sigma_\delta)$[2]. The end of life T is assumed to be certain. It would be straight forward to incorporate stochastic mortality into the model. This additional complexity is not likely to significantly affect the results. Following Deaton (1991), I define the endogenous state variable cash on hand as the sum of financial assets and labour income and it evolves as follows:

$$X_{t+1} = (1 + r^e_{t+1})S_t + (1+r)B_t + Y_{t+1} \tag{2}$$

where r^e_{t+1} is the stochastic return from the risky asset, r is the risk-free rate, S_t is the amount of wealth invested in the risky asset, B_t is the amount of wealth invested in the risk-free asset. Following Carroll and Samwick (1997), Y_{t+1} is stochastic labour

[2] The unboundedness of the discount rate will not pose any difficulty in estimations because I use 6-point gaussian quadrature to approximate the distribution which inevitably bounds the possible range of discount rates.

income which follows the following exogenous stochastic process:

$$Y_{t+1} = P_{t+1}U_{t+1} \tag{3}$$

$$P_{t+1} = G_{t+1}P_tN_{t+1} \tag{4}$$

Permanent income, P_t, grows at the rate G_{t+1} and it is subject to multiplicative i.i.d shocks, N_t. Current income, Y_t, is composed of a permanent component and a transitory shock, U_t. I adopt the convention of estimating the earnings growth profile by assuming $G_t = f(t, Z_t)$, where t represents age and Z_t are observable variables relevant for predicting earnings growth. I also assume that the transitory shocks, U_t, are distributed independently and identically, take the value of zero with some small but positive probability, and are otherwise lognormal: $\ln(U_t) \sim N(-0.5\sigma_u^2 , \sigma_u^2)$. Similarly, permanent shocks N_t are i.i.d with $\ln(N_t) \sim N(-0.5\sigma_n^2 , \sigma_n^2)$. By assuming that innovations to income are independent over time and across individuals I assume away aggregate shocks to income. However, aggregate shocks are not completely eliminated from the model since I assume the same return process for all agents and, as explained below, I allow a link between market disasters and the probability of zero income.

Introducing zero income risk into the life cycle model is proposed by Carroll (1992) and adopted by many subsequent papers. It is important to note that introducing zero income risk into the standard model does not by itself solve the problem of portfolio specialization or limited participation. Although it generates diversified portfolios at the low end of the wealth distribution, it also triggers prudence leading to rapid wealth accumulation early in the life cycle. If the observed post-war equity premium were the expected return, some of this wealth would be channeled into the stock market and the model would still predict counterfactually high stock market participation and large risky asset shares at young ages.

Returning to the model description, the excess return of the risky asset is assumed to be i.i.d:

$$r_{t+1}^{e} - r = \mu + \varepsilon_{t+1} \tag{5}$$

where μ is mean excess return and ε_{t+1} is distributed normally with mean 0 and variance σ_{ε}^{2}. Agents face a small but positive probability of a disastrous market downturn. When such an event occurs, a large portion of the household's stock market wealth evaporates (return of $-\phi$ percent where $\phi > 0$). Moreover, when the asset market is hit by a disaster, the probability of a zero income realization increases (from a small calibrated value to π percent). It is important to note that in the case of such a disaster, stock market participants lose ϕ percent of their stock market wealth and face a π percent chance of zero labor income for the whole year whereas nonparticipants face only the job loss risk (π percent chance of zero labor income for the whole year). I do not allow innovations to excess return to be correlated with innovations to permanent or transitory income in normal market times. Allowing for such a correlation is straightforward and would reduce the ex-ante disaster probability and disaster size needed to match the data. However, the empirical support for such correlation is very weak (see Heaton and Lucas (2000)), so I set it to zero.

One important assumption I make is that the risk-free rate is not affected by a disastrous market downturn. This may not be true as one may think that a disaster in stock markets would push down government bond yields leading to a still higher equity premium. Or, one may think of a war-like disaster where governments totally or partially default. Incorporating a perceived probability of government default can be done in the way Barro (2006) suggests. However, separately identifying such a probability (assuming the size of the default is the same as the size of the stock market decline as in Barro (2006)) from a stock market disaster probability is empirically impossible. Given that there exists no clear pattern regarding how government bonds will perform in disastrous times, I assume that the risk-free rate is not affected by a potential market disaster[3].

The optimization problem involves solving the recursive Bellman equation via backward induction. I divide the life cycle problem into two main sections: The individual starts working life at the age of 25 and works until 60. He retires at 60 and lives until 80. During his working life he contributes an (exogenous) fraction τ of his income to a personal retirement account which earns the risk-free rate. At the age of 60 his certain annual retirement income is flow value of a 20 year annuity[4]. The

[3]Barro (2006) shows that bills did quite well in the United States during the great depression whereas partial default on government debt occured in Germany and Italy during WW II.

[4]Modelling retirement in this way greatly simplifies the solution of the model. Given that the estimation is performed on households in the pre-retirement years, this simple modelling of retirement should not cause any serious bias.

recursive problem is:

$$V_t(X_t, P_t) = \max_{S_t, B_t} \left\{ \frac{(C_t)^{1-\gamma}}{1-\gamma} + \frac{1}{1+\delta} E_t V_{t+1} \left[(1 + r^e_{t+1}) S_t + (1+r) B_t + (1-\tau) Y_{t+1}, P_{t+1} \right] \right\} \tag{6}$$

subject to borrowing and shortsale constraints

$$S_t \geq 0, \;\; B_t \geq 0$$

where $V_t(.)$ denotes the value function.

The structure of the problem allows me to normalize the necessary variables by dividing them by permanent income (see Carroll 1992). Doing this reduces the number of endogenous state variables to one, namely the ratio of cash on hand to permanent income. The Bellman equation after normalizing is:

$$V_t(x_t) = \max_{s_t, b_t} \left\{ \frac{(c_t)^{1-\gamma}}{1-\gamma} + \frac{1}{1+\delta} E_t (G_{t+1} N_{t+1})^{(1-\gamma)} V_{t+1} \left[(1 + r^e_{t+1}) s_t + (1+r) b_t / N_{t+1} + (1-\tau) U_{t+1} \right] \right\} \tag{7}$$

where $x_t = \frac{X_t}{P_t}$, $s_t = \frac{S_t}{P_t}$, $b_t = \frac{B_t}{P_t}$ and $c_t = \frac{C_t}{P_t} = x_t - s_t - b_t$.

I assume away the bequest motive, therefore the consumption function c_T and the value function $V(c_T)$ in the final period are $c_T = x_T$ and $V(x_T) = \frac{x_T^{1-\gamma}}{1-\gamma}$ respectively. In order to obtain the policy rules for earlier periods I define a grid for the endogenous state variable x and maximize the above equation for every point in the grid.

When the model is augmented with a per-period participation cost, the solution requires some additional computations. Now, the optimizing agent has to decide whether to participate in the stock market or not before he decides how much to invest. This is done by comparing the discounted expected future value of participation and that of nonparticipation in every period. This results in the following optimization problems:

$$V_t(x_t, I_t) = \max_{0,1}\left(V^0(x_t, I_t), V^1(x_t, I_t)\right) \tag{8}$$

where

$$V^0(x_t, I_t) = \max_{s_t, b_t}\left\{\frac{(c_t)^{1-\gamma}}{1-\gamma} + \frac{1}{1+\delta}E_t V_{t+1}\left[x_{t+1}, I_{t+1}\right]\right\} \tag{9}$$

subject to

$$x_{t+1} = (1+r)b_t/G_{t+1}N_{t+1} + U_{t+1} \tag{10}$$

where I_t is a binary variable representing participation at time t. $V^0(x_t, I_t)$ is the value the consumer gets by not participating regardless of whether he has participated in the previous period or not, i.e. exit from the stock market is assumed to be costless[5].

$$V^1(x_t, I_t) = \max_{s_t, b_t}\left\{\frac{(c_t)^{1-\gamma}}{1-\gamma} + \frac{1}{1+\delta}E_t V_{t+1}\left[x_{t+1}, I_{t+1}\right]\right\} \tag{11}$$

subject to

$$x_{t+1} = \left[(1+r^e_{t+1})s_t + (1+r)b_t\right]/G_{t+1}N_{t+1} + (1-\tau)U_{t+1} - F^c \tag{12}$$

[5] It is plausible to assume that the agent incurs some transaction cost by exiting the stock market. Considering different types of transaction costs associated with the stock market participation would make estimation infeasible and it does not add any insight to the point made in the paper. See Vissing-Jorgensen (2002) for a detailed treatment of stock market participation costs.

$V^1(x_t, I_t)$ is the value the consumer gets by participating. F^c is the fixed per-period cost to permanent income ratio which is 0 if the household does not have any stock market investment and it is positive if he has some stock market investments. The per-period cost considered here is not a one-time fee. It has to be paid (annually in this framework) as long as the household holds some stock market wealth. It can be thought of as the value of time spent to follow markets and price movements in addition to actual trading fees. Since it is related to the opportunity cost of time it is plausible to formulate it as a ratio to permanent income[6].

In each time period, given his current participation state, the household first decides whether to invest in the stock market or not (or stay in it if he is already in) by comparing the expected discounted value of each choice. Then, conditional on participation he decides how much wealth to allocate to the risky asset. If he chooses not to participate, the only saving instrument is the risk-free asset which has a constant return r. Further details of the solution method are given in Appendix A.

3 Estimation Overview

3.1 Simulating Auxiliary Statistics

The estimation procedure is an application of Simulated Minimum Distance (SMD) which involves matching statistics from the data and from a simulated model.[7] I allow

[6] This assumption is fairly standard in the literature. With this simlifying but justifiable assumption, I reduce the total number of state variables to three: age, cash-on-hand and participation status.

[7] A description of the general SMD procedure is given in Appendix B.

the discount rate, δ, to be heterogenous across cohorts and lognormally distributed within a cohort. The coefficient of relative risk aversion is common for everyone within a cohort but allowed to differ across cohorts.

The simulation procedure takes a vector of structural parameters $\Psi = \{\gamma, \mu_\delta, \sigma_\delta^2, p, \phi, \pi, \kappa\}$ where

(γ) coefficient of relative risk aversion

(μ_δ) mean discount rate

(σ_δ^2) variance of the discount rate

(p) probability of the event

(ϕ) size of the expected loss in case of the event

(π) probability of zero income in the case of the event

(κ) per-period participation cost to stock market

and solves the underlying dynamic program described in the previous section. The resulting age and discount rate dependent policy functions are used to simulate consumption, portfolio share and participation paths for H households for $t = 1, ...T$. To perform simulations, I need two T by H matrices (for permanent and transitory income shocks), and two H by 1 vectors (for initial wealth to income ratio and discount rates) of standard normal variables[8] in addition to realized stock returns from

[8] If $lnx \sim N(a, b)$, we can simulate draws from a lognormal by taking $x \sim \exp(a + bN(0, 1))$ where $N(0, 1)$ denotes the standard Normal. The mean and variance of x are given by $\mu_x = \exp(a)\sqrt{\exp(b^2)}$, $\sigma_x^2 = \exp(2a)\exp(b^2)(\exp(b^2) - 1)$

1983 to 2004. Although households ex-ante differ only in their discount rate, their ex-post life cycle outcome will differ due additionally to different income and initial wealth realizations.

As discussed in the data section, the lack of panel data on consumption, wealth and income forces me to use some complementary data techniques. This means having to replicate the limitations of the actual data in the simulated data to obtain consistent estimates. To do this, the procedure first simulates the balanced panel of consumption, portfolio shares and participation for all households and then selects observations to replicate the structure of the cross section data. For example, suppose we have 234 25-year-olds and 567 26-year-olds in the youngest cohort in the SCF. The procedure will pick 234 25 year old households from the simulated paths, then will pick 567 26 year olds (different households as we are creating a cross section to imitate the data) and so on. In the end this *imitated* data is used to calculate all wealth related auxiliary parameters (described below).

For consumption, the process is more involved. As described below, natural auxiliary parameters to describe consumption behavior are the mean and variance of consumption growth. Since the construction of these auxiliary parameters requires observing households for at least two periods and CEX is repeated cross section[9], I use the quasi-panel methods developed by Browning, Deaton and Irish (1985) and used by many other researchers. This method amounts to taking the cross section averages of consumption within a given cohort (controlling for some time-invariant house-

[9] The CEX has a rotating quarterly panel dimension that I do not use here. This is explained in the data section.

hold characteristics) and then generating consumption growth using these means. Simulated consumption data goes through the same process to replicate the actual consumption data.

3.2 Choosing Auxiliary Parameters: Identification

I now need to choose statistics of the data - so called auxiliary parameters (*aps*) - that are matched in the SMD step; I denote these $\lambda_1, ..\lambda_K$. As always, we have a trade-off between the closeness of the *aps* to structural parameters (the 'diagonality' of the binding function, see Gouriéroux *et al* (1993) and Hall and Rust (1999)) and the need to be able to calculate the *aps* quickly. It should be noted that many of the *aps* defined below are closely related to the underlying structure but none of the *aps* are consistent estimators of any parameter of interest; rather, they are chosen to give a good, parsimonious description of the joint distribution of consumption, financial wealth and stock market returns across cohorts.

The first *ap* relates to the total financial wealth: it is the mean financial wealth to permanent income ratio. This will help me identify the mean discount rate.

$$\lambda_{01} = mean\,(finw) \tag{13}$$

The next six *aps* ($\lambda_{02}-\lambda_{07}$) are smoothed age profiles of participation and portfolio shares. I summarize age profiles with a quadratic polynomial, i.e., I first run the following two regressions:

$$
\begin{aligned}
share &= \lambda_{02} + \lambda_{03} Age + \lambda_{04} Age^2 + \varepsilon && (14) \\
part &= \lambda_{05} + \lambda_{06} Age + \lambda_{07} Age^2 + \nu && (15)
\end{aligned}
$$

where *part* is a dummy variable that equals 1 if the household owns stocks and zero otherwise. *Share* is the portfolio share of stocks in the household's financial portfolio. The next two *aps* are the mean and standard deviation of the portfolio share of stocks conditional on participation. As will subsequently become clear, these *aps* play an important role, in conjunction with consumption *aps*, in pinning down the coefficient of relative risk aversion parameter and the perceived disaster probability.

$$
\begin{aligned}
\lambda_{08} &= mean(share|part = 1) && (16) \\
\lambda_{09} &= std(share|part = 1) && (17)
\end{aligned}
$$

The last two *aps* relate to consumption; they are the mean and standard deviation of consumption growth. The effect of family size changes ($\Delta size$) on consumption growth is removed via an initial regression:

$$\Delta \log C = \zeta_0 + \zeta_1 \Delta size + \epsilon \quad (18)$$

Then,

$$\lambda_{10} = \zeta_0 \quad (19)$$

$$\lambda_{11} = std(\epsilon) \tag{20}$$

While the mean consumption growth rate helps to identify the mean discount rate, the variation in consumption growth helps to identify the elasticity of intertemporal substitution (the reciprocal of the coefficient of relative risk aversion). Thus, I have 11 *aps* to estimate 7 structural parameters. In principle, one can have many more *aps* (second, third and forth moments, covariances etc.) but these statistics are sufficient (and intuitive) to estimate the parameters of interest.

4 Estimation

4.1 Data and Pseudo-Panel Construction

The ideal data set for this task would be a panel that contains both consumption and portfolio allocation information, covers a large number of households, and spans a sufficiently long time period. Unfortunately, such data do not exist and, to further impede my analysis, neither do cross section data for the U.S. that contain both consumption and wealth information. Therefore, I resort to complementary data techniques to generate the auxiliary parameters and, of course, I process the simulated data as if they contain the same limitations. I work with two distinct repeated cross-sectional data sets, one containing data on consumption and the other data on financial wealth. I use these two data sets to create a pseudo-panel following Browning, Deaton and Irish (1985). This technique involves defining cells based on birth cohorts, and other time invariant or perfectly predictable characteristics (typically education, sex and race), and then following the cell mean of any given variable of interest over time.

I use the American Consumer Expenditure Survey (CEX) for consumption expenditure information. The data covers the period between 1983 and 2004. The expenditure information is recorded quarterly with approximately 5000 households in each wave. Every household is interviewed five times, four of which are recorded (the first interview is practice). Although the attrition is substantial (about 30% at the end of the fourth quarter), the survey is considered to be a representative sample of the US population. I select married households whose head identified himself as white. Households that do not report nondurable consumption for all four quarters are excluded as I use annual nondurable consumption expenditure to generate my consumption *aps*. My nondurable consumption measure excludes medicare and education expenditures and all durable expenditures. Annual nondurable consumption for each household is obtained by aggregating over four quarters.

After generating the real annual consumption measure for each household, I create a pseudo-panel for nondurable consumption. First, I divide the sample into two broad groups by level of education: college and over (referred to as more educated) and less than college (referred to as less educated). Then I define 3 birth cohorts for each education group, giving six groups in total. The first group contains more educated households who were 25 to 27 years old in 1983 and 26 to 28 in 1984 and so on. The second group contains more educated households who were 28 to 32 years old in 1983 and 29 to 33 in 1984 and so on. The third group contains more educated households who were 33 to 35 years old in 1983 and 34 to 36 in 1984 and so on. Groups 4, 5 and 6 are defined analogously for the less educated households. I restrict the age range to be 25 to 54. The reason, as explained in the results section, is that it becomes increasingly difficult to model portfolio holdings as households approach retirement age. I calculate the mean of the logarithm of real annual consumption for each group for each year I have data[10]. The mean and standard deviation of consumption growth (after removing family size effect) constitute my consumption *aps*.

For asset information I use the American Survey of Consumer Finance (SCF) which covers the same time period as the CEX. The information on financial wealth and portfolio allocation is recorded at the household level and it is available through the family files. The SCF details the most comprehensive wealth data available among industrialized countries. It is a cross section that is repeated every three years. Note that CEX provides annual expenditure information whereas wealth information is available triennially in the SCF. This limitation is also replicated in the simulated data. It is important to note that wealth *aps* are generated using SCF weights as SCF oversamples wealthy households.

I restrict the sample from the SCF in the same way that I restricted the CEX, and define the same groups. Variables of interest from this data source are the share of stocks in households' financial portfolios (portfolio share), stock market participation indicator, portfolio shares conditional on participation and financial wealth to per-

[10]The fact that one can control the order of aggregation is one of the great advantages of the pseudo-panel technique. Since I have to generate a consumption growth measure later on, I first take logs of household consumption and then calculate the mean. Related studies using aggregate data lack this luxury (as the sum of logs does not equal the log of sums).

manent income ratio[11]. A household's financial portfolio is defined as the sum of all bonds, stocks, certificate of deposits and mutual funds. Assets such as trust accounts and annuities are excluded as they are not incorporated in my life cycle model. I also exclude checking and saving accounts as they are kept mostly for households' transactional needs, and my model abstracts from liquidity issues. Risky assets are defined as all publicly and privately traded stocks as well as all-stock mutual funds. Bonds, money market funds, certificate of deposits and bond funds altogether constitute the risk-free asset.

4.2 Initial Conditions and Other Parameters

Following standard practice in the literature, I restrict the number of structural parameters that I estimate and calibrate the others. In principle, all the parameters could be estimated through the structural routine, including income process parameters. However, this extra complication does not add any insight to the point made in the paper as the real issue is to estimate the perceived disaster parameters that justify observed household portfolios. I use the Panel Studies of Income Dynamics (PSID) to calibrate the parameters of income processes (1983-1992). The variances of innovations to permanent income and transitory income are estimated separately for all 6 groups. Earnings growth profiles are estimated separately for the two education levels and taken as common for all 3 cohorts within an education level. Table 1 presents the

[11] Permanent income for each household is the predicted values obtained from the regression of labor income on age, occupation and industry dummies. This estimation (although imperfect) is quite standard in the literature.

estimates. It has been argued that the ex-post variation in individual income may not accurately represent the true uncertainty that the individual is facing. In particular, households may have several informal ways to mitigate idiosyncratic background risk that an econometrician cannot observe. If this is the case, we tend to overestimate actual income variances. Bound and Krueger (1991) and Bound (1994) suggest that roughly a third of estimated variance is due to mismeasurement. Therefore I use two thirds of the estimated value of the permanent income variance and use the actual estimated value for the transitory income variance.

I set the risk-free rate to 2%, the mean equity return is taken to be 6% with a standard deviation of 20% (these values seem to be the consensus, see Mehra (2008)). I set the probability of zero income to 0.00302 (as estimated by Carroll 1992). Because I do not observe all households at the beginning of their life cycle, i.e. at age 25, I need to estimate an initial wealth distribution to initiate simulations. One approach is to assume that initial assets to permanent income ratios are drawn from a log normal distribution and estimate the mean and standard deviation using all 25-year-olds in the data (see Gourinchas and Parker (2002) and Alan (2006)).

The immediate objection to this approach is that it is unrealistic to think that older cohorts started out with the same level of initial wealth as younger cohorts. Unfortunately, we cannot possibly know what the older cohorts had when they were young. Therefore, instead, I fit a lognormal distribution to the earliest observed financial wealth to permanent income ratios for each group, that is at the youngest age I observe them, and then initiate simulations at this age by drawing from these distributions. For example, the initial value for the financial wealth to permanent income ratio for a household in group 1 (the youngest cohort) is drawn from the lognormal distribution whose parameters are estimated using financial wealth to permanent income ratios of this cohort when households were at the age of 25-27. Similarly, the initial value for a household in group 3 (oldest cohort) is drawn from the lognormal distribution whose parameters are estimated using data of this cohort when households were at the age of 33-35 (first time they were observed). Finally, retirement tax τ is set so that the replacement ratio is 50%.

4.3 Estimation in Practice

Steps of the estimation procedure are as follows:

- Obtain necessary auxiliary parameters from the data $(\lambda_{01} - \lambda_{11})$.

- Obtain variance-covariance matrix of auxiliary parameters through a boostrap procedure. This is done by estimating the *aps* $(\lambda_{01} - \lambda_{11})$ repeatedly by sampling from the main sample with replacement.

- Solve the underlying structural model for given structural parameters (Ψ) and simulate portfolio paths that imitate the data patterns. Obtain simulated *aps*: $[\lambda_{01}(\Psi) - \lambda_{11}(\Psi)]$.

- Minimize the distance between the simulated and data auxiliary parameters with respect to the seven structural parameters using the variance-covariance matrix as the weighting matrix.

This procedure is repeated for all six groups. After obtaining the base estimates, I use the same procedure to re-estimate the model under several alternative sets of parameter restrictions. These restricted estimates are used to assess which features of the model are crucial to its ability to fit the data.

4.4 Estimation Results

The model I estimate has seven structural parameters ($\Psi = \{\gamma, \mu_\delta, \sigma_\delta^2, p, \phi, \pi, \kappa\}$). Table 2 presents my baseline estimates of these parameters for all six groups[12]. The perceived probability of disaster estimates are in the neighborhood of 4 percent per annum which is comparable to values obtained by Barro and Ursua(2008). If we compare the two education groups, it appears that the youngest more educated group attach a larger probability to a market disaster than the youngest less educated group (4.4 percent versus 3.9 percent). These numbers do not tell us much about disaster expectations without the expected disaster size (expected annual percentage loss of stock market wealth). Overall, estimated disaster sizes imply that more than half the stock market wealth is expected to evaporate if a disaster strikes. Older and less educated households are found to expect larger decline (around 60 percent). The estimated probability of zero income in the case of a disaster ranges from 8 to 20 percent; not surprisingly less educated households picture a gloomier labor market experience in such a case (a 20 percent chance of zero income for the youngest less

[12] As the asymptotic standard errors are unreliable for these types of models, I do not report them. Instead, I discuss the precision in terms of the proximity of the *aps* generated by the data and the simulated data at the estimated values later in this section.

educated). How do the size estimates compare with Barro's calibrated values? The real stock market return was -16.5% per year between the years 1929 and 1932 in the United States, implying over a 50 percent decline in the stock market wealth in four years. As disasters do not last more than one period (year) in my model, the estimated size (around 55% for the more educated and 60 % for the less educated) can be interpreted as *total* expected wealth loss in the event of a disaster, and thus my estimates are broadly comparable to Barro (2006) and Barro and Ursua (2008).

One of the most striking results in this paper is that participation cost estimates are zero for all groups. There is now a sizeable body of research promoting transaction cost based explanations of the portfolio and equity premium puzzles (see for example Alan (2006) and other references therein). The idea is that households face costs associated with participating and trading in the stock market. The definition of these costs is usually very broad; it incorporates a range of things from simple trading fees to the opportunity cost of time spent on portfolio management. The main difficulty regarding this explanation is related to portfolio shares conditional on participation. For example, stock market participation can be delayed by introducing an entry cost. However, upon entry, portfolio shares implied by the standard model, which uses ex-post return as a proxy for expected returns, are implausibly high (usually 100% at young ages) with any reasonable coefficient of relative risk aversion coefficient[13].

[13] The reason for this result is that at young ages, households make decisions mainly based on their labor income. Even though the labor income is risky and not fully insurable, since its innovations are not correlated with stock returns, it still acts like a bond. As the household ages, accumulation of financial wealth reduces the importance of labor income in decision making (unless the household has a very high discount rate and does not accumulate much) and then we observe some diversification (lowering the weight of stocks in the portfolio).

When we reformulate the risk associated with investing in the stock market by allowing for the possibility of a disaster (affecting labor earnings as well as stock market wealth), participation, portfolio shares and shares conditional on participation come down to reasonable levels making the participation cost assumption unnecessary.

The coefficient of relative risk aversion estimates are in line with estimates based on micro data on consumption (see Attanasio et al (1999), Gourinchas and Parker (2002)). In general, estimates based on consumption data generate a lower coefficient of relative risk aversion compared to estimates based on wealth data (see Cagetti (2003)). Overall, consumption based estimates of the coefficient of relative risk aversion range between unity and 3. The range I estimate is similar; the coefficient of relative risk aversion for the oldest less educated cohort is estimated to be 0.93 (my lowest estimate), and that for the youngest more educated cohort is estimated to be 2.19 (my highest estimate). The more educated appear to be more risk averse than the less educated[14]. Moreover, younger cohorts appear to be more risk averse than older ones.

Of all the features that empirical analyses using household consumption and saving data have to address, discount rate heterogeneity is perhaps the most important. Alan and Browning (2008) document marked differences among households in their consumption growth that cannot be explained by a simple measurement error argument. A model that does not incorporate discount rate heterogeneity will have great difficulty in explaining financial wealth and allocation statistics. Table 2 presents

[14]The consumption based estimates of Alan and Browning (2008) point to the same conclusion although their estimated values are much higher.

the median and the standard deviation of the estimated discount rate distribution. These estimates are generally high compared to those obtained by Alan and Browning (2008). They found a median discount rate of 7.5 % for the less educated and 4.2 % for the more educated. My estimates range from 10% to 12.1%. Dispersion within a group is usually quite large; for example, the estimated standard deviation of 5.2% implies a 25th percentile of 7.7% and a 75th percentile of 13.8%.

In order to illustrate the relative importance of different model features in fitting life cycle profiles, I estimate a number of restricted variants of the main model. Table 3 presents my goodness of fit results. The first model is the unrestricted one with seven structural parameters (Model 1). The overall fit is quite reasonable even though the model is rejected for all six groups. A good way to see where the fit fails is to look at the t-ratios for the difference between data *aps* and their simulated counterparts calculated at estimated structural parameters. This is shown in Table 4 for the more educated and Table 5 for the less educated. For the less educated, only a couple of the t-ratios point to rejection (both in cohort 3), whereas for the more educated, portfolio share profiles of the oldest cohort (cohort 3) and the *ap* based on the financial wealth to permanent income ratio (for cohort 1 and 2) are rejected. This is consistent with the fact that overall fit is better for the less educated group and it is worse for older households. The likely reason for a better fit for the less educated is that financial wealth is more homogenous (as well as low) and much less skewed for this group, whereas it is too heterogenous to be captured by the standard model for the more educated. Similar reasoning applies to the relative poor fit of older cohorts as they tend to become more heterogeneous than younger ones. In addition, behavior becomes harder to characterize as households approach retirement age[15].

It is informative to see how far we can go in explaining portfolios without appealing to preference heterogeneity and participation costs. The next row in Table 3 gives the version of the model estimated while discount rate variation and per-period participation costs are set to zero (Model 2). As can be seen, chi-square statistics jump up quite significantly. Given estimated transaction costs are zero for all cohorts in the unrestricted model, these results suggest that discount rate heterogeneity is essential to match observed portfolios and consumption growth (Alan and Browning (2008) reached the same conclusion using only consumption data). The third row gives the chi-square statistics when the probability of disaster and disaster size are set to zero, implying also that the probability of a zero income realization is set back to 0.32 percent (Model 3). This experiment is performed to show that it is not possible to even come close to observed profiles by simply introducing discount rate heterogeneity and participation costs. As can be seen, chi-square statistics become massive. For the youngest more educated cohort the chi-square statistic jumps from 83.7 to 7,436 and for the youngest less educated cohort it jumps from 27 to 2379. This is the household data manifestation of the equity premium puzzle. Keep in mind that this row still allows for heterogenous discount rates and per-period participation costs (remedies proposed in the literature to address the household portfolio puzzles).

[15]Retirement is modelled in the simplest possible way here and the bequest motive is assumed away. Given the rather simplisticly modelled post-retirement income and abstracting from possible health shocks (which becomes increasingly important as households age), I chose to restrict my estimation sample to households whose head is less than 55 years of age.

The fourth row is simply the standard model without heterogeneity or transaction costs and without disaster expectations (Model 4). The huge chi-square values tell the story. It is clear that there is no hope for the standard model to match household portfolios (participation, shares and shares conditional on participation) without re-specifying the risk associated with investing in the stock market. It is important to emphasize that there are two main culprits behind the bad fit when disaster expectations are not allowed: the distribution of portfolio shares of stocks conditional on participation and consumption growth. Models with participation costs and heterogeneity (along with implausibly high risk aversion) can mimic observed participation profiles, but such models cannot resolve the small saver puzzle or provide an adequate fit for the distribution of consumption growth.

How do the simulated life-cycle profiles of portfolio holdings look compared to the data? Figures 1 and 2 depict life cycle stock market participation and portfolio share profiles calculated at the estimated structural parameters (see Table 2) superimposed on their data counterparts. Profiles obtained from restricted models (see Table 3) are also superimposed for a more general comparison. As can be seen from these figures, simulated participation and portfolio share paths from the unrestricted model (Model 1) closely track their data counterparts for most groups. Shutting down discount rate heterogeneity and participation costs (Model 2) worsens the fit, suggesting that discount rate heterogeneity is essential. Allowing for discount rate heterogeneity and participation costs but shutting down the possibility of a disaster (Model 3) generates an even worse fit than Model 2. Finally, the standard model (Model 4) is absolutely hopeless.

What do the structural parameters imply for the life-cycle profiles of portfolio shares of stocks conditional on participation? Previous research on household portfolios has not been successful in matching this important statistic (as illustrated by Model 2, 3 and 4). Figure 3 depicts simulated age profiles of conditional portfolio shares (at the estimated parameter values) and their data counterparts. In order to put the results into perspective, I also present conditional shares simulated from Model 3 (the model which includes participation costs and discount rate heterogeneity but not disaster expectations.) The first and most important thing to note is that the small saver puzzle disappears when we allow for the possibility of disasters. Model 1 delivers lower portfolio shares in earlier life, and so is much more congruent with the data. In many, but not all groups, the simulated profiles closely track their data counterparts. It should be mentioned that the simulated mean and standard deviation of conditional portfolio shares are not statistically different from their data counterparts for Model 1, and are significantly different for Model 3 (see the t-ratios for λ_{08} and λ_{09} in Table 4 and Table 5 for Model 1) [16]. It is clear that a model with participation costs and discount rate heterogeneity, like Model 3, can go some way to explaining the life cycle profiles of participation and, to some extent, unconditional portfolio shares, but cannot account for the profiles of portfolio shares conditional on participation. With any sensible value of the coefficient of relative risk aversion, such

[16]The t-tests for the difference between mean conditional portfolio shares (standard deviations) and their simulated counterparts calculated at estimated structural parameters using model 3 result in rejection for all six groups. Results are available upon request.

a model will imply 100 percent equity holding, conditional on participation.

Overall, the results suggest that once we depart from using the post-war historical premium as a proxy for ex-ante return expectations and allow for a small probability of a disaster in the stock market (and allow the probability of a bad labour market outcome to rise during a stock market disaster) standard models fit the household portfolio data quite well. How would the results change if some key calibrated parameters were changed? One could easily allow for a positive correlation between returns and permanent income shocks (Gomes and Michaelides (2005) assume a 0.15 correlation even though the empirical evidence suggest no particular pattern). This would introduce a strong hedging motive away from stocks and so lower the perceived disaster probability and size necessary to fit the portfolio data. Increasing the permanent income variance would imply a stronger precautionary motive leading to offsetting reductions in the coefficient of relative risk aversion and increases in discount rate estimates. One could allow for committed household expenditures such as housing and rent to better characterize the young households. This would reduce disposable income which in turn would likely to reduce my median discount rate estimates.

4.5 A Disaster Scenario

Note that even though ex-ante expectations include a disaster possibility (in Model 1), the simulations reported up to this point were performed using the realized returns from the 1983-2004 period, a period that did not contain a disaster. In this section I present a counterfactual experiment to assess the reaction of aggregate consumption to an actual disaster. The experiment involves the counterfactual assumption that 2004 is a disaster year. The stock market index goes down by 50% and 15% of the working population experience a zero labor income realization. Note that in the case of such a disaster, stock market participants lose 50% of their stock market wealth and face a 15% chance of zero labor income for the whole year whereas nonparticipants face only the 15% chance of a zero labor income realization.

Table 6 presents the percentage fall in aggregate consumption that results from this disaster. The first row compares total consumption of more educated groups in 2003 with that of 2004, the second row repeats the same exercise for the less educated groups. The last row presents the aggregate consumption fall due to the disaster. As seen from the table, the estimated parameters imply an annual decline in aggregate consumption of around 10%. Using a sample of 21 countries (including the United States) Barro and Ursua (2008) calculate that stock market disasters are associated with a mean decline in aggregate consumption of 22 percent from peak to trough. Thus, the 10 percent decline in aggregate consumption, which Model 1 implies will occur when a disaster is realized, seems quite reasonable.

5 Conclusion

This paper builds upon the argument that rare economic disasters, once taken into account, can solve asset pricing puzzles. I show that actual quantities in the micro data support this argument. If ex-ante expectations include a small probability of a disastrous market event, observed household portfolio holdings and consumption growth can be reconciled with the standard intertemporal model. Moreover, this can be done without abandoning the original CRRA framework or introducing transaction costs.

Instead of applying calibrated disaster probabilities and sizes to the standard model, I estimate these parameters using the micro data on wealth allocation and consumption. The estimated disaster probabilities and disaster sizes that justify the life cycle profiles of stock market participation and portfolio shares are in line with historically calibrated values. Moreover, the estimates of the coefficient of relative risk aversion parameters are in the range obtained by studies that match only consumption growth. Finally, estimated per-period participation costs are zero for all groups, confirming that household portfolio demand is driven mainly by perceived stock market risk rather than by participation costs.

I do not test this explanation directly against explanations based on preference re-specifications. Such explanations include the internal and external habit models proposed by Constantinides (1990), Campbell and Cochrane (1999) and Abel (1990). The common feature of these preference re-specifications is that they increase effective risk aversion. In terms of the implied life cycle paths of portfolios, such models behave similarly to models with extreme uninsurable income risk. In both cases, the marginal utility of consumption can become extremely high (near zero consumption, the subsistence level, or the habit level.) The limitation of all of these explanations (uninsurable income risk and internal and external habits) is that when the effective risk aversion is high, so is prudence. This implies counterfactually high financial wealth accumulation and consequently counterfactually high stock market participation over the life cycle. Even though one can match overall mean conditional and unconditional portfolio shares with such models, the implied life cycle profiles will not look anything like their data counterparts. Thus in my view, disaster expectations seem a more promising solution.

References

Abel, A. B., "Asset Prices under Habit Formation and Catching up with the Joneses", *American Economic Review, Papers and Proceedings,* (1990), 80, 38-42

Alan, S. "Entry Costs and Stock Market Participation over the Life Cycle", *Review of Economic Dynamics* (2006), 9(4).

Alan, S. and M. Browning, "Estimating Intertemporal Allocation Parameters using Simulated Residual Estimation", (2008) CAM Working Paper 2003-03.

Attanasio,O., J. Banks, C. Meghir and G. Weber, "Humps and bumps in Lifetime Consumption" *Journal of Business and Economic Statistics*, 17 (1999), 22-35.

Bansal R., A. Yaron, "Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles", *Journal of Finance*, (2004), no 4, 1481-1509.

Barro R. J., "Rare Disasters and Asset Markets in the Twentieth Century", *Quarterly Journal of Economics*, (2006), 803-866.

Barro R. J. and J. Ursua, "Rare Disasters and Asset Markets" *American Economic Review: Papers and Proceedings*, (2008), 98:2, 58-63.

Bertaut, C. C., " Stockholding Behavior of U.S. Households: Evidence from the 1983-89 Survey of Consumer Finances" *Review of Economics and Statistics* 80 (1998), 263-75.

Bound J., "Evidence on the Validity of Cross-sectional and longitudinal Labor Market Data", *Journal of Labor Economics,* (1994), 12, 345-368.

Bound J. and A. B. Krueger, "The Extend of Measurement Error in Longitudinal Earnings Data: Do Two Wrongs make a Right?", *Journal of Labor Economics,* (1991), 9, 1-24.

Browning M., A. Deaton and M. Irish, "A Profitable Approach to Labor Supply and Commodity Demands Over the Life Cycle, *Econometrica* 53(1985), 503-543.

Browning M., M. Ejrnaes and J. Alvarez, "Modelling Income Processes with Lots of Heterogeneity", (2002), CAM working paper.

Cagetti, M., "Wealth Accumulation Over the Life Cycle and Precautionary Savings", *Journal of Business and Economic Statistics*, (2003), 21(3), 339-353.

Campbell, J. Y., and J. H. Cochrane, "By Force of Habit: A Consumption Based Explanation of Aggregate Stock Market Behavior", *Journal of Political Economy,* (1999), 107, 205-251.

Carroll, C. D., "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence", *Brookings Papers on Economic Activity,* 2 (1992), 61-156.

Carroll, C. D. and A. Samwick, "The Nature of Precautionary Wealth" *Journal of Monetary Economics* 40 (1997), 41-72.

Cocco, F. J., F. J. Gomes and P. J. Maenhout, "Consumption and Portfolio Choice Over the Life Cycle" *Review of Financial Studies,* (2005), (18(2), 491-533.

Constantinides, G. M., "Habit Formation: A Resolution of the Equity Premium Puzzle", *Journal of Political Economy,* (1990), 98,519-543.

Deaton, A., "Saving and Liquidity Constraints" *Econometrica* 59 (1991), 1221-1248.

Gabaix X., "Variable Rare Disasters: A Tractable Theory of Ten Puzzles in Macro-Finance", *American Economic Review: Papers and Proceedings*, (2008).

Gomes, F., A. Michaelides, "Optimal Life Cycle Asset Allocation: Understanding Empirical Evidence", *Journal of Finance*, (2005), no:2, 869-904.

Gourieroux, C., A. Monfort and E. Renault, (1993) “Indirect Inference”, *Journal of Applied Econometrics*, 1993, 8:S85-S118.

Gourinchas, P. O. and J. A. Parker, "Consumption over the Life Cycle" *Econometrica* 70 (2002), no:1, 47-89.

Guiso, L., M. Haliassos and T. Japelli, "Household Portfolios" MIT Press 2002.

Haliassos, M., A. Michaelides, "Portfolio Choice and Liquidity Constraints" *International Economic Review*, 44 (2003) no:1, 143-177.

Hall, G. and J. Rust, “Econometric Methods for Endogenously Sampled Time Series: The Case of Commodity Price Speculation in the Steel Market” (2002), Cowles Foundation discussion papers 1376, University of Yale.

Heaton, J., and D. Lucas, "Portfolio Choice in the Presence of Background Risk", *The Economic Journal* 110 (2000a), 1-26.

Mehra, R., "The Equity Premium Puzzle: A Review", Foundations and Trends in Finance, (2006), vol(2), no:1, 1-81.

Mehra, R. and E. C. Prescott, "The Equity Premium: A Puzzle", *Journal of Monetary Economics* (1985), 15, 145-161.

Reitz, T. A., "The Equity Risk Premium: A Solution", *Journal of Monetary Economics,* (1988), 117-131.

Vissing-Jorgensen, A., "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures" mimeo, University of Chicago, 2002

Weitzman M. L., "Subjective Expectations and Asset-Return Puzzles" *American Economic Review*, (2008), Forthcoming.

A Solution and Simulation Methods

The standard life cycle model for portfolio choice described in Section 2 is solved via backward induction by imposing a terminal wealth condition. Simply, in the last period of life all accumulated wealth has to be consumed so the policy rule for consumption is

$$c_T = x_T$$

and for stocks and bonds

$$s_T = 0, \; b_T = 0$$

Therefore the last period's value function is the indirect utility function:

$$V_T(x_T) = \frac{x_T^{1-\gamma}}{1-\gamma}$$

In order to solve for the policy rules at $T-1$, I discretize the state variable cash on hand to permanent income ratio x. The algorithm first finds the investment in the risky and risk-free assets that maximizes the value function for each value in the grid of x. Then, another optimization is performed where the generic consumer has only the risk-free asset to invest in. Values of both optimizations are compared and the rule that results in a higher value is picked. The value function at $T-1$ is the outer envelope of the two value functions. Since I use a smooth cubic spline to approximate value functions, nonconvexities due to taking the outer envelope of two functions do not pose any numerical difficulty.

B Simulated Minimum Distance

Here I present a short account of the Simulated Minimum Distance (SMD) method as applied generally to panel data (see Hall and Rust (2002) and Browning, Ejrnæs and Alvarez (2006) for details). Suppose that we observe $h = 1, 2..H$ units over $t = 1, 2...T$ periods recording the values on a set of Y variables that we wish to model and a set of X variables that are to be taken as conditioning variables. Thus we record $\{(Y_1, X_1), ...(Y_H, X_H)\}$ where Y_h is a $T \times l$ matrix and X_h is a $T \times k$ matrix. For modelling we assume that Y given X is identically and independently distributed over units with the parametric conditional distribution $F(Y_h|X_h;\theta)$, where θ is an m-vector of parameters. If this distribution is tractable enough we could derive a likelihood function and use either maximum likelihood estimation or simulated maximum likelihood estimation. Alternatively, we might derive some moment implications of this distribution for observables and use GMM to recover estimates of a subset of the parameter vector. Sometimes, however, deriving the likelihood function is extremely onerous; in that case, we can use SMD if we can simulate Y_h given the observed X_h and parameters for the model. To do this, we first choose an integer S for the number of replications and then generate $S * H$ simulated outcomes $\left\{(Y_1^1, X_1), ...(Y_H^1, X_H), (Y_1^2, X_1), ...(Y_H^S, X_H)\right\}$; these outcomes, of course, depend on the model chosen ($F(.)$) and the value θ takes in the model.

Thus we have some data on H units and some simulated data on $S * H$ units that have the same form. The obvious procedure is to choose a value for the parameters which minimizes the distance between some features of the real data and the same features of the simulated data. To do this, define a set of *auxiliary parameters* that are used for matching. In the Gouriéroux *et al.* (1993) Indirect Inference procedure, the auxiliary parameters are maximizers of a given data dependent criterion which constitutes an approximation to the true data generating process. In Hall and Rust (2002), the auxiliary parameters are simply statistics that describe important aspects of the data. I follow this approach. Thus I first define a set of J auxiliary parameters:

$$\gamma_j^D = \frac{1}{H}\sum_{h=1}^{H} g^j\left(Y_h, X_h\right), \; j = 1, 2...J \tag{21}$$

where $J \geq m$ so that I have at least as many auxiliary parameters as model parameters. The J-vector of auxiliary parameters derived from the data is denoted by $\boldsymbol{\gamma}^D$. Using the same functions $g^j\left(.\right)$ I can also calculate the corresponding values for the simulated data:

$$\gamma_j^S = \frac{1}{S * H}\sum_{s=1}^{S}\sum_{h=1}^{H} g^j\left(Y_h^s, X_h\right), \; j = 1, 2...J \tag{22}$$

and denote the corresponding vector by $\boldsymbol{\gamma}^S\left(\theta\right)$. Identification follows if the Jacobian of the mapping from model parameters to auxiliary parameters has full rank:

$$rank\left(\nabla_\theta \boldsymbol{\gamma}^S\left(\theta\right)\right) = m \text{ with probability 1} \tag{23}$$

This effectively requires that the model parameters be 'relevant' for the auxiliary parameters.

Given sample and simulated auxiliary parameters, I take a $J \times J$ positive definite matrix W and define the SMD estimator as:

$$\hat{\theta}_{SMD} = \arg\min_{\theta} \left(\boldsymbol{\gamma}^{S}(\theta) - \boldsymbol{\gamma}^{D}\right)' W \left(\boldsymbol{\gamma}^{S}(\theta) - \boldsymbol{\gamma}^{D}\right) \tag{24}$$

The choice I adopt is the (bootstrapped) covariance matrix of $\boldsymbol{\gamma}^{D}$. Typically we have $J > m$; in this case the choice of weighting matrix gives a criterion value that is distributed as a $\chi^2(J-m)$ under the null that we have the correct model.

		Estimated variance of permanent component	**Estimated variance of transitory component**
	cohort 1	.016 (.001)	.010 (.002)
More Educated	cohort 2	.017 (.002)	.007 (.001)
	cohort 3	.015 (.002)	.011 (.002)
	cohort 1	.018 (.005)	.026 (.004)
Less Educated	cohort 2	.027 (.005)	.012 (.003)
	cohort 3	.024 (.003)	.017 (.002)

Standard errors in parentheses. Mean predictable income growth for the more and less educated are 0.018 and -0.001 respectively. Source PSID 1983-1992

Table 1: Estimated Parameters of Income Processes

Parameter	More Educated			Less Educated		
	Cohort 1	Cohort 2	Cohort 3	Cohort 1	Cohort 2	Cohort 3
CRRA (γ)	2.19	2.16	1.84	1.26	1.35	0.93
Median discount rate (%)	10.3	11.4	11.1	12.1	11.8	10.4
Standard dev. of discount rate (%)	5.2	8.9	6.9	7.0	6.8	5.5
Probability of the event (%)	4.4	4.2	3.7	3.9	4.0	5.3
Probability of zero income in the case of the event (%)	10.5	7.9	14.9	20.1	18.9	14.2
Per-period participation cost (%)	0.00	0.00	0.00	0.00	0.00	0.00
Size of the expected loss in case of the event (%)	53.7	55.8	59.8	65.1	66.2	51.8
χ^2	83.7	186.1	294.6	27.0	25.5	131.9

Note: critical value for $\chi^2_4 = 9.5$ for (95% confidence)

Table 2: Structural Estimation Results

Model	**Parameter restrictions**	**Degrees of freedom**	**More educated** (χ^2) **Cohort 1**	**Cohort 2**	**Cohort 3**	**Less educated** (χ^2) **Cohort 1**	**Cohort 2**	**Cohort 3**
1	–	4	83.7	186.1	294.6	27.0	25.5	131.9
2	$\sigma_\delta^2 = \kappa = 0$	6	172.7	334.3	592.4	117.1	105.2	153.2
3	$p = \phi = 0,\ \pi = 0.032$	7	7,436	9,620	5,655	2,379	2,244	5,080
4	$p = \phi = \sigma_\delta^2 = \kappa = 0,\ \pi = 0.032$	9	11,863	14,701	13,650	7,969	8,441	9,511

Notes: Number of *aps* =11

Table 3: Goodness of fit

Auxiliary Parameters	More Educated Cohort 1 Data	 Simulated	 Cohort 2 Data	 Simulated	 Cohort 3 Data	 Simulated
λ_{01}	.94	.81	.98	1.1	1.2	1.1
		$(2.8)^*$		$(2.19)^*$		(1.8)
λ_{02}	$-.53$	$-.76$	$-.23$	$-.07$	.45	-1.5
		(.70)		(1.8)		$(2.1)^*$
λ_{03}	.03	.03	.05	.02	$-.02$	.06
		(.50)		(1.1)		$(2.6)^*$
λ_{04}	$-.00$	$-.00$	$-.00$	.00	.00	$-.00$
		(.21)		(0.8)		(.88)
λ_{05}	$-.19$	$-.17$	-1.2	$-.66$	$-.37$	-2.1
		(.62)		(.63)		$(2.3)^*$
λ_{06}	.01	.05	.11	.03	$-.01$	.08
		(.88)		(.95)		(.87)
λ_{07}	.00	$-.00$	$-.00$	$-.00$	.00	$-.00$
		(.86)		(.98)		(.77)
λ_{08}	.53	.57	.59	.53	.54	.61
		(1.81)		$(2.18)^*$		$(2.35)^*$
λ_{09}	.30	.29	.27	.29	.30	.27
		(1.48)		(1.32)		(1.22)
λ_{10}	.001	$-.003$	$-.000$	.004	.001	$.-000$
		(.25)		(.17)		(.34)
λ_{11}	.049	.021	.016	.034	.04	.021
		(1.31)		(.43)		(.56)

Note: absolute $t-$ratios (testing the equality of data and simulated data *aps*) in parentheses.

Table 4: Auxiliary Parameters and Simulated Counterparts, More Educated

Auxiliary Parameters	Less Educated Cohort 1 Data	Simulated	Cohort 2 Data	Simulated	Cohort 3 Data	Simulated
λ_{01}	.33	.33	.44	.45	.45	.40
		(.02)		(.13)		(1.78)
λ_{02}	−.37	−.26	−.48	−.52	.32	.47
		(.98)		(.18)		(.40)
λ_{03}	.03	.02	.02	.02	−.02	−.02
		(1.07)		(.17)		(.20)
λ_{04}	−.00	−.00	−.00	−.00	.00	−.00
		(1.7)		(.42)		(2.4)*
λ_{05}	−.36	−.50	−.63	−.72	−.29	−.94
		(.18)		.08		(.44)
λ_{06}	.02	.02	.03	.03	.01	.03
		(.02)		(.15)		(.83)
λ_{07}		−.00	−.00	−.00	.00	−.00
		(.38)		(.27)		(.80)
λ_{08}	.53	.54	.45	.52	.48	.50
		(.40)		(1.0)		(2.13)*
λ_{09}	.31	.30	.28	.26	.31	.31
		(.85)		(1.3)		(.21)
λ_{10}	−.006	−.001	−.004	−.002	−.003	−.001
		(.47)		(.17)		(.11)
λ_{11}	.036	.029	.047	.033	.075	.019
		(1.0)		(.61)		(1.3)

Note: absolute $t-$ratios (testing the equality of data and simulated data *aps*) in parentheses.

Table 5: Auxiliary Parameters and Simulated Counterparts, Less Educated

	Percentage Fall in Consumption
More Educated	9.5%
Less Educated	11.1%
Aggregate	10.1%

2004 is the disaster year. 50% of the stock market wealth is lost and 15% of the population receive zero labor income for the entire year.

Table 6: Aggregate consumption fall in case of a disaster

FIGURE 1
Portfolios of More Educated

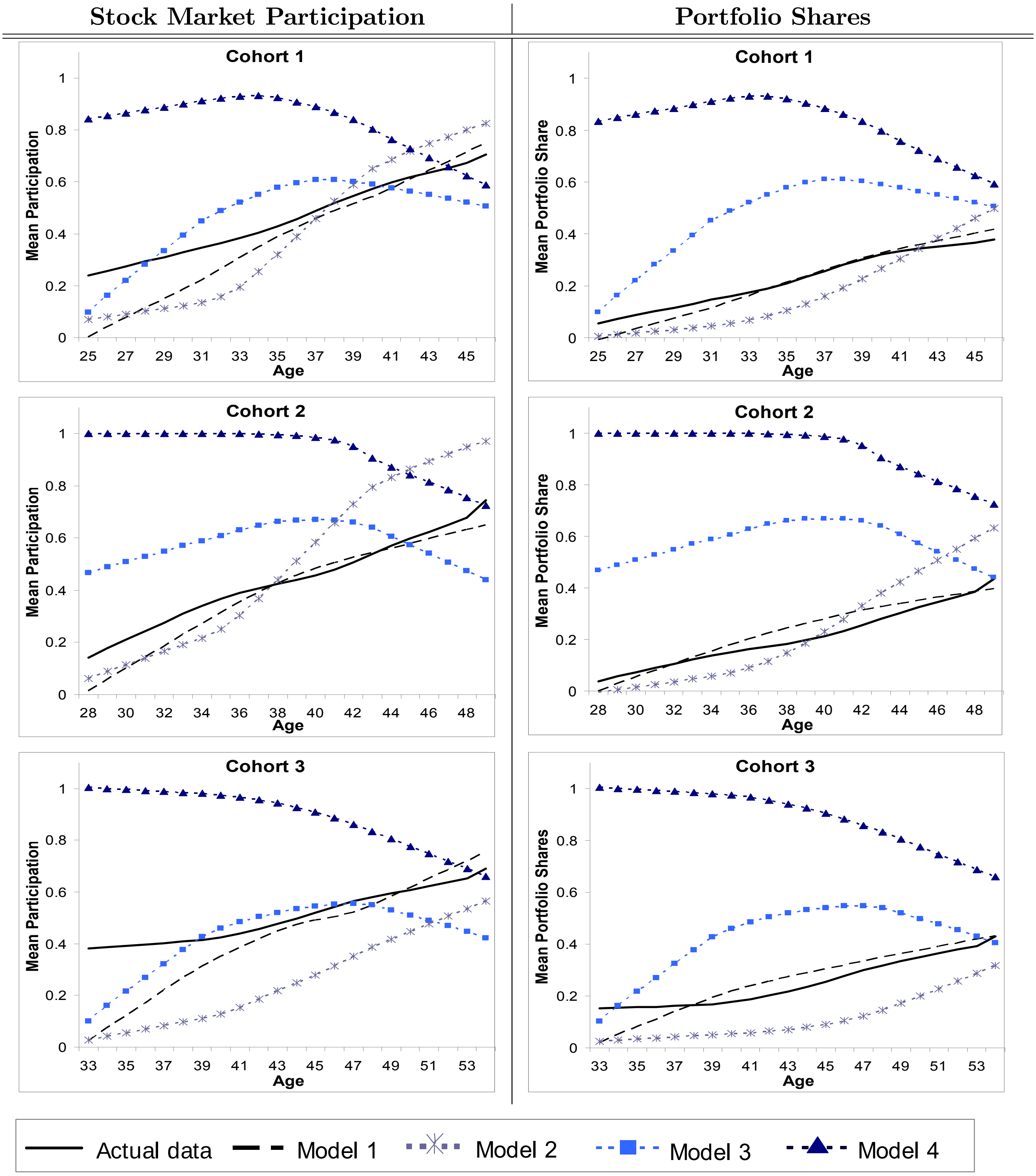

FIGURE 2
Portfolios of Less Educated

Stock Market Participation	Portfolio Shares

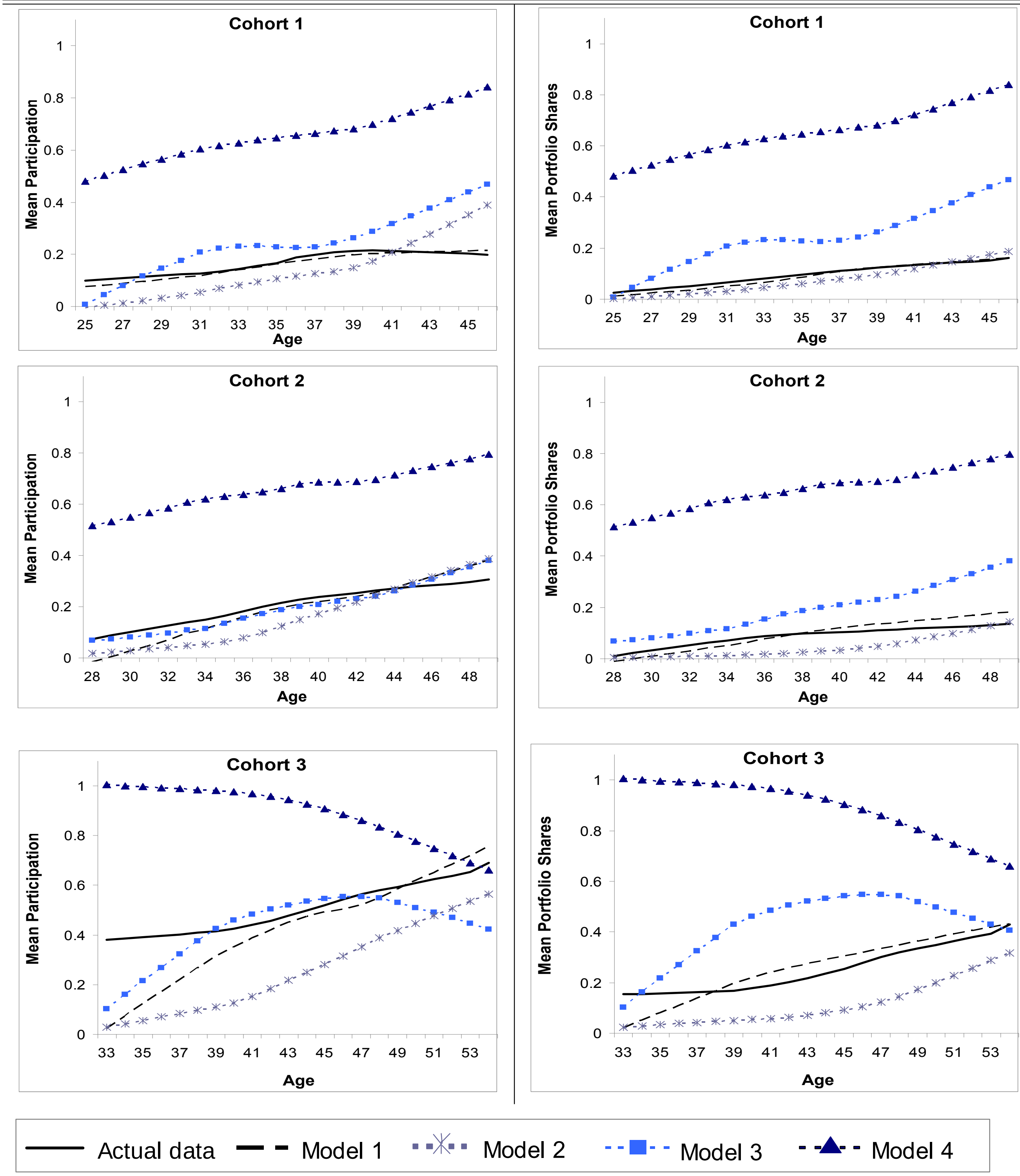

FIGURE 3
Portfolio Shares Conditional on Participation