

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

Feb 2010

¹I thank Manuel Arellano, Thomas Crossley, Michael Haliassos, John Heaton, Hamish Low, Xavier Mateos-Planas and seminar participants at the University of Cambridge and University College London, participants of Household Finance Workshop organized by Bank of Spain 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). All errors are mine.

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 portfolio choice of households that have less than a college degree can be partially explained by expectations of stock markets disasters only if one allows for a large (and implausible) probability of labor income loss at the same time. Such disaster expectations however, are not revealed in the portfolios of educated and wealthier households; simple per-period participation costs to stock market coupled with preference heterogeneity explain their participation and investment patterns.

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).

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). Put differently, perceived risk associated with a disaster in stock markets should be revealed in household portfolios. This idea could be tested 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 (both probability and expected size) and preference parameters from observed portfolios and then judge whether the estimates are plausible as compared to the historically calibrated values. Moreover, I choose to use a much richer and realistic version of the consumer problem than the original model (also the one assumed in Barro (2006) and Reitz (1988)). Estimating the entire structural model gives me the opportunity to test several other explanations of equity premium against the explanation based on economic disasters. I argue that if the correct quantities are not revealed in an environment that is a lot more flexible than the original one, the explanation of the equity premium puzzle with rare disasters would be significantly weakened. The results in this paper suggest that the expectation of a rare disaster can go some ways in explaining the portfolios of uneducated households only if it is reinforced with an extreme (and rather implausible)) labor market stress. Such expectations are not revealed in the portfolios of more sophisticated and wealthy households who are believed to be the relevant portion of the population in terms of aggregate wealth and asset prices.

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 4 groups (2 birth cohorts by 2 education levels). I also go significantly beyond the existing literature and allow for preference heterogeneity within groups. Except for the old and more educated group, the probability of a rare disaster and expected disaster size are estimated precisely. The point estimates for the perceived disaster probability range from 1% (less educated young) to 5% (more educated young). The estimated probability of a disaster is not statistically different from zero for the old and more educated households (wealthiest households in the sample). Heterogeneity in the coefficient of relative risk aversion (value of 4 at the 25th percentile and 9 at the 75th percentile) and per-period participation costs (approximately 1 percent of the permanent income) appear to be sufficient to explain the portfolios of these households.

The remainder 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 the auxiliary environment. 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

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

where C is non-durable consumption, γ_h is the household specific coefficient of relative risk aversion, δ_h is the household specific rate of time preference which are assumed to be distributed lognormally across households such that $\ln\gamma_h \sim N(\mu_\gamma, \sigma_\gamma)$, $\ln\delta_h \sim N(\mu_\delta, \sigma_\delta)$ respectively¹. The ideal would be to assume a joint distribution for the preference parameters and estimate all five distribution parameters $(\mu_\gamma, \sigma_\gamma, \mu_\delta, \sigma_\delta, \rho_{\gamma,\delta})$. However, given the core question, such an addition would increase the complexity of the problem without offering any useful insight. Here, I already go beyond what has been done in the literature in terms of preference heterogeneity and assume parameter heterogeneity one at a time. That is, when the coefficient of relative risk aversion is assumed to be heterogenous, the discount rate heterogeneity is closed down, and when discount rate heterogeneity is assumed, the heterogeneity in the coefficient of relative risk aversion is closed down. In the end, I let data determine which model fits better².

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

¹The unboundedness of the discount rate and the coefficient of relative risk aversion will not pose any difficulty in estimations because I use 6-point gaussian quadrature to approximate the distributions which inevitably bounds possible ranges.

²Alan and Browning (2009) is the first to estimate a joint distribution of the intertemporal allocation parameters using food expenditure data in the PSID. But their model of consumption is much simpler than the model used here.

evolves as follows:

$$X_{t+1} = (1 + r_{t+1}^e)S_t + (1 + r)B_t + Y_{t+1} \quad (2)$$

where r_{t+1}^e 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 income which follows the following exogenous stochastic process:

$$Y_{t+1} = P_{t+1}U_{t+1} \quad (3)$$

$$P_{t+1} = G_{t+1}P_tN_{t+1} \quad (4)$$

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 low income

realizations.

Introducing zero income risk into the life cycle model is proposed by Carroll (1992) and adopted by many subsequent papers³. 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 dis-

³Since zero income expectation does apply only to a small number of households, it may be more realistic to assume that a labour market stress may be in the form of having to collect unemployment benefits for a significant number of households. One of the models I test against the benchmark presented here assumes a small floor (above zero) for transitory income realization instead of zero.

aster, 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 disaster⁴.

The optimization problem involves solving the recursive Bellman equation via backward induction. I divide the life cycle problem into two main sections: The

⁴Barro (2006) shows that bills did quite well in the United States during the great depression whereas partial default on government debt occurred in Germany and Italy during WW II.

individual starts working life at the age of 25 and works until 60. He retires at 60 and lives until 80. During his retirement he receives social security income each period which is equal to a fraction τ of his permanent income at the age of 60. The recursive problem is:

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

subject to borrowing and shortsale constraints

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

where $V_t(\cdot)$ 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:

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

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:

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

where

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

subject to

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

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 costless⁵.

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

⁵It 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.

subject to

$$x_{t+1} = [(1 + r_{t+1}^e)s_t + (1 + r)b_t] / G_{t+1}N_{t+1} + U_{t+1} - F^c \quad (12)$$

$V^1(x_t, I_t)$ 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 income⁶.

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.

⁶This assumption is fairly standard in the literature. With this simplifying but justifiable assumption, I reduce the total number of state variables to two: age and cash-on-hand.

3 Estimation Overview

3.1 Simulating Auxiliary Statistics

The structural estimation is performed for four different groups. Households are grouped first, according to their broad educational attainment. Households with heads who have less than a college degree are labelled as "less educated", those who have a college degree or over are labeled as "more educated". Within these groups, households are further divided according to their birth cohorts. Households with heads who have been born before 1946 are labelled as "old", after 1946 are labelled as "young". The details of the sample selection will be given in the next section. The estimation procedure is an application of Simulated Minimum Distance (SMD) which involves matching statistics from the data and from a simulated model.⁷ For bench mark estimations, I allow the discount rate, δ , or the coefficient of relative risk aversion, γ , to be heterogenous across groups and lognormally distributed within a group. When the coefficient of relative risk aversion is assumed to be homogenous, it is still allowed to differ across four groups. Similarly, when the discount rate is assumed to be homogenous, it is still allowed to differ across four groups.

The simulation procedure takes a vector of structural parameters

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

$\Psi = \{\mu_\gamma, (\sigma_\gamma^2), \mu_\delta, (\sigma_\delta^2), p, \phi, \pi, \kappa\}$ where

(μ_γ) mean coefficient of relative risk aversion

(σ_γ^2) variance of coefficient of relative risk aversion, (set to zero if $\sigma_\delta^2 > 0$)

(μ_δ) mean discount rate

(σ_δ^2) variance of the discount rate, (set to zero if $\sigma_\gamma^2 > 0$)

(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 (or coefficient of relative risk aversion) dependent policy functions are used to simulate consumption, portfolio share and participation paths for H households for $t = 1, \dots, T$. 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, or coefficient of relative risk aversion) of standard normal variables⁸ in addition to actual realized stock returns from 1983 to 2004.

As discussed in the data section, the lack of panel data on consumption, wealth

⁸If $\ln x \sim N(a, b)$, 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 $\mu_x = \exp(a)\sqrt{\exp(b^2)}$, $\sigma_x^2 = \exp(2a)\exp(b^2)(\exp(b^2) - 1)$

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 section⁹, 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 household characteristics) and then generating consumption growth using these means. Simulated consumption data goes through the same process to replicate the actual consumption data.

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

3.2 Choosing an Auxiliary Environment

I now need to choose statistics of the data - so called auxiliary parameters (*aps*) - that are matched in the SMD step; I denote these $\lambda_1, \dots, \lambda_K$. 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 median financial wealth to permanent income ratio. This will help me identify the discount rate.

$$\lambda_{01} = \text{median}(finw) \tag{13}$$

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:

$$share = \lambda_{02} + \lambda_{03}Age + \lambda_{04}Age^2 + \varepsilon \tag{14}$$

$$part = \lambda_{05} + \lambda_{06}Age + \lambda_{07}Age^2 + \nu \tag{15}$$

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.

$$\lambda_{08} = \text{mean}(share|part = 1) \quad (16)$$

$$\lambda_{09} = \text{std}(share|part = 1) \quad (17)$$

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,

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

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

last two aps are the unconditional mean of portfolio share of stocks and participation rate respectively:

$$\lambda_{12} = \text{mean}(share) \quad (21)$$

$$\lambda_{13} = \text{mean}(\text{part}) \tag{22}$$

While the median financial wealth to permanent income ratio and mean consumption growth rate help 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 13 *aps* to estimate 7 structural parameters, leaving me with 6 degrees of freedom. In principle, one can have many more *aps* (second, third and fourth moments, covariances etc.) but I believe that the auxiliary environment described above is sufficiently rich and intuitive characterization of the joint distribution of parameters of interest.

It is important to emphasize that separately identifying the probability and size of the disastrous event is difficult in this setting. Simply put, there may be many combinations of these two parameters leading to the same auxiliary environment. However, repeated re-estimations with a large set of different starting values indicate strongly that local identification around plausibly restricted structural parameters is feasible. These restrictions include lower and upper bounds for the preference parameters (naturally imposed by the discretization process) , positivity constraints for variances and probabilities and negativity constraint for the event size.

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 2 birth cohorts for each education group, giving four groups in total. I restrict the age range to be 25 to 59. 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 data¹⁰. 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

¹⁰The 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).

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. Finally, imputations in the SCF are taken into account when bootstrapping the variance covariance matrix of the *aps*.

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 permanent income ratio¹¹. 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.

¹¹Permanent 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.

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 4 groups. Earnings growth profiles are estimated separately for the two education levels and taken as common for both cohorts within an education level. Table 1 presents the 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 the level of wealth the older cohorts had when they were young. To overcome this problem, I devise a unique way of initializing the simulations. For each household I observe, I start the simulations using its observed wealth to permanent income ratio. For example, say I need to simulate life cycle paths of a household whom I observe at the age of 40 in year 1998, with wealth to permanent income ratio of 2.5. I start the simulations of this household by assuming that initial wealth to permanent income ratio is 2.5, using the policy functions that are relevant for 40 year-olds and actual stock market returns starting in 1998. This household's paths are simulated until he is 59. This way, I exactly replicate not only the age structure of the SCF, but also major shortcomings of the data (missing values, triennial structure and absence of a panel).

4.3 Estimation Results

The benchmark model I estimate has seven structural parameters:

$$\Psi = \{\mu_\gamma, (\sigma_\gamma^2), \mu_\delta, (\sigma_\delta^2), p, \phi, \pi, \kappa\}$$

Parameters are estimated for four groups separately assuming discount rate and coefficient of relative risk aversion heterogeneity one at a time. For all groups, γ heterogeneity yielded the lowest chi-squared criterion. Therefore, all analyses in this section are based on models with γ heterogeneity (benchmark) and I will not discuss δ heterogeneity.¹²

Before turning to the parameter estimates, I illustrate the general features of the fit. To do this, I estimate a number of restricted variants of the benchmark model. Table 2 presents my goodness of fit results. The first model is the unrestricted model (γ heterogeneity) with seven structural parameters (Model 1, benchmark). The overall fit is quite reasonable even though the model is rejected for all four groups. One perhaps not very surprising result is that fit is better for the less educated group. 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. It is on the other hand, high and more heterogenous for the more educated group leading to much more decisive rejection of the model. The particular effect of γ heterogeneity can be

¹²Overall fit and parameter estimates for δ heterogeneity are not very different from those with γ heterogeneity; see the last row of Table 3 for the over all fit. Homogeneity of discount rates are rejected by all groups. Full results for δ heterogeneity are available upon request.

seen by examining the second row of the same table where γ heterogeneity is closed down. Chi-square increases for all groups are sizable enough to warrant rejection of γ homogeneity however, jumps in the chi-square values are much larger for the educated group (from 705 to 1384 for the young, from 100 to 545 for the old) confirming a higher degree of preference heterogeneity amongst the more educated.

Another interesting variant of the model is estimated by assuming a possibility of a large income drop bounded above zero instead of a possibility of a zero income realization. This assumption is perhaps more realistic for older and wealthier households. The situation in mind is that an individual may lose his job and settle for a small fraction of his current income for a year (collecting unemployment benefit for example). I assume that in normal times this probability is 4% (roughly the natural rate of unemployment in the U.S.) and the fraction is 30%. Similar to the benchmark case, I let the probability of such situation arising during the disaster be a free parameter to estimate. I estimate this model by closing down preference heterogeneity so the fair comparison would be against model 2 where γ heterogeneity is closed down. Note also that this model is not nested in the benchmark model and should be viewed as an alternative instead of a restricted variant. As can be seen in the third row of the table, the fit for this model is much better for the educated group; chi-square values go down from 1384 to 969 and from 545 to 488 for the young and the old respectively. This suggests that zero income realization is not a good assumption for these households. The opposite is observed for the less educated; large jumps in chi square values from 59.5 to 1923 and from 30.8 to 393 for the young and the old

respectively.

Similarly, a possibility of a disaster does not seem to be a good assumption for older and more educated households as suggested by the statistics in the fourth row of the Table 2. This variant of the model is estimated by closing down the disaster possibility while keeping γ heterogeneity¹³. In fact, for these households, even the simplest model with no heterogeneity, no disaster expectations and no entry cost do not lead to a huge jump in the chi-squared criterion (model 5, $\chi_{11}^2 = 702.7$) while such a variant makes the fit hopeless for all other groups; see the last row. The take away from this table is that the standard model has serious difficulties to match household portfolios and this difficulty cannot be overcome by assuming a market disaster. Although this explanation seems to go some ways to explain the behavior of the households with very little financial wealth, one should keep in mind that these are not the individuals who are relevant for prices.

Where does the fit fail? A good and economically meaningful way to see this 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 5 for the more educated and Table 4 for the less educated. For the less educated, only a couple of the t -ratios point to rejection, whereas for the more educated most of the simulated *aps* do not come close to their data counterpart. The biggest failure come from the first *ap* (λ_1), median financial wealth to permanent income ratio. As can be seen in the first row of Table 5 model persistently generate higher *aps* than the data.

¹³The chi-squared increment between the benchmark and model 4 is $\chi_3^2 = 104.9 - 100.2 = 4.7$. Given the critical value for χ_3^2 is 7.81 at 95%, model 4 restrictions are not rejected.

I now turn to the structural estimates based on the benchmark model. Table 3 presents the estimates for all four groups¹⁴. Except for the old and more educated group, the probability of a disaster and the expected size are estimated precisely. The point estimates for the perceived disaster probability range from 1% (less educated young) to 5% (more educated young). For the less educated, the expected size estimates are very large (80% and 70% for the young and the old respectively). The probability of zero income realization in the case of a disaster is also quite high for the less educated young (38%), although not as large for the old (16%). The estimated probability of a disaster is not statistically different from zero for the old and more educated. Consistent with the earlier discussion on goodness of fit, more educated older cohort (wealthiest households in the sample) do not appear to think that such market disasters are relevant for their portfolio choice. The very fact that it is these households that drive the aggregate wealth casts serious doubt on explanation of the equity premium with rare disasters.

The question now is how my 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. Since disasters are assumed to strike in an iid fashion (as in Reitz and Barro), the size estimates are not directly comparable but can be interpreted as *total* expected wealth loss in the event of a disaster. Estimate of 80% loss seem to be too

¹⁴Although the asymptotic standard errors are unreliable for these types of models, I still report them. The precision can be judged in economically more meaningful way by the proximity of the *aps* generated by the data and the simulated data at the estimated values; see Tables 4 and 5.

disastrous especially since it is coupled with 38% probability of zero income realization for the less educated young. For the more educated young, the expected size estimate seem reasonable (41%) although it is coupled with 5% perceived probability of disaster which is too high compared to the calibrated values in Barro (2006) and Barro and Ursua (2008).

A striking result of the estimation is that there is a substantial variation in preference parameter estimates across education groups but not so much across cohorts within education groups. Consistent with Alan and Browning (2009) the less educated seem to have a lower relative risk aversion. Discount rate estimates seem very high especially for the older cohorts (28% for the old and more educated). 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)) especially for the less educated. 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 much wider; the median coefficient of relative risk aversion for the oldest less educated cohort is estimated to be 1.36 (my lowest estimate), and that for the oldest more educated cohort is estimated to be 6.1 (my highest estimate). Note that the more educated group is the most heterogenous (consistent with the goodness of fit tests). Not surprisingly, the old and more educated group is the most heterogenous with the coefficient of relative risk aversion of 4 and 9 at the 25th and 75th percentiles. The same estimates

for the young and more educated are 2.9 and 3.6¹⁵.

One of the most striking results in this paper is that participation cost estimates are zero for the less educated but positive and significant for the more educated. 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 coefficient¹⁶. 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 for the less educated. But

¹⁵Table 3 reports the mean log coefficient of relative risk aversion and its standard deviation. The median values and percentiles that I am reporting here come from the simulation of the relevant log-normal distribution (at the estimated parameters) for 100,000 households for each group.

¹⁶The 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).

these costs seem to be still important for the more educated, especially for the older cohort where the per-period participation cost is estimated to be approximately 1 percent of the permanent income.

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 ??) superimposed on their data counterparts. Profiles obtained from restricted models (see Table 2) 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 the less educated groups and shutting down γ heterogeneity does not visibly worsens the fit; see Figure 1. Note also that the standard model (Model 5) is absolutely hopeless. The life cycle profiles do not seem to track their data counterpart as closely for the more educated group, consistent with estimation results; see Figure 2. What is particularly disturbing in this figure is that the model persistently generate a hump shape for shares and participation which does not exist in the data.

Figure 3 tell us exactly where the model fails. It depicts simulated age profiles of conditional portfolio shares (at the estimated parameter values) and their data counterparts. The first and most important thing to note is that the degree of small saver puzzle diminishes especially for the less educated when we allow for the possibility of disasters. Model 1 and 2 deliver lower portfolio shares in earlier life, and so is much more congruent with the data. This is obviously not the case for the more educated.

It should be noted that the main reason for the decisive rejection of the model for the more educated (large chi-square criterions) is the fact that conditional shares are low and very precisely estimated in the data. Such low conditional shares are hard to match given the financial wealth of this group.

Overall, the results suggest that allowing for rare disasters does lead standard model to fit the household portfolio data well for some households albeit not the ones that are driving the aggregate wealth. Preference heterogeneity and participation costs appear to be better explanations for the portfolio decisions of wealthier households. If we are to accept the explanation of equity premium by appealing to economic disasters, we should, at the very least, be able to infer the expectation of such disasters from the quantities held by the wealthy households. We do not see this in a clear cut way in the data.

5 Conclusion

This paper builds upon the argument that rare economic disasters, once taken into account, can solve asset pricing puzzles. This idea naturally calls for an exercise to assess whether correct quantities can be obtained from a framework that claims to yield correct prices. 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. I show that it is difficult to reconcile actual quantities in the micro data with this explanation. If return 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 only for households that possess very little wealth. Even for these households such reconciliation is not possible without assuming a serious labor market stress at the time of the stock market disaster. Such probabilities are not statistically different from zero for wealthier households. Portfolio decisions of such households can be better explained by preference heterogeneity and transaction 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. Explanations based on business cycle risk (a way of correlating stock returns with labor earnings indirectly) may be a more promising route as in Lynch and Tan (2009). However, need to reconcile other

aspects of intertemporal behavior such as consumption and savings within the same framework remains crucial.

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", (2010), forthcoming, *Review of Economic Studies*.

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.

Lynch, A. W., and S. Tan, "Labor Income Dynamics at Business-cycle Frequencies: Implications for Portfolio Choice", forthcoming, *Journal of Financial Economics*.

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, Ejrnaes and Alvarez (2006) for details). Suppose that we observe $h = 1, 2, \dots, H$ units over $t = 1, 2, \dots, T$ 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 $\{(Y_1, X_1), \dots, (Y_H, X_H)\}$ 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 $\{(Y_1^1, X_1), \dots, (Y_H^1, X_H), (Y_1^2, X_1), \dots, (Y_H^S, X_H)\}$; these outcomes, of course, depend on the model chosen ($F(\cdot)$) 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:

$$\gamma_j^D = \frac{1}{H} \sum_{h=1}^H g^j(Y_h, X_h), \quad j = 1, 2, \dots, J \quad (23)$$

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 γ^D . Using the same functions $g^j(\cdot)$ I can also calculate the corresponding values for the simulated data:

$$\gamma_j^S = \frac{1}{S_* H} \sum_{s=1}^S \sum_{h=1}^H g^j(Y_h^s, X_h), \quad j = 1, 2, \dots, J \quad (24)$$

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

$$\text{rank}(\nabla_{\theta} \gamma^S(\theta)) = m \text{ with probability } 1 \quad (25)$$

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:

$$\hat{\theta}_{SMD} = \arg \min_{\theta} (\boldsymbol{\gamma}^S(\theta) - \boldsymbol{\gamma}^D)' W (\boldsymbol{\gamma}^S(\theta) - \boldsymbol{\gamma}^D) \quad (26)$$

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 shocks	Estimated variance of transitory shocks
Less Educated	Young	.12 (.01)	.12 (.01)
	Old	.15 (.01)	.13 (.01)
More Educated	Young	.11 (.01)	.10 (.004)
	Old	.12 (.01)	.10 (.01)

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

Model	Parameter restrictions	Degrees of freedom	Less educated (χ^2)		More educated (χ^2)	
			Young	Old	Young	Old
1 (benchmark)	—	6	38.6	24.7	705.0	100.2
2	$\sigma_\gamma^2 = 0$ (nested in 1)	7	49.5	30.8	1384	544.6
3	$\sigma_\gamma^2 = 0$, income floor (non-nested)	7	1923	393.0	969.2	488.4
4	$p = \phi = 0$, $\pi = 0.0032$ (nested in 1)	9	947.5	764.0	6056	104.9
5	$p = \phi = \sigma_\gamma^2 = \kappa = 0$, $\pi = 0.0032$ (nested in 1)	11	40545	8063	11938	702.7

Table 2: Goodness of fit

	Less Educated		More Educated	
	Young	Old	Young	Old
Mean of Log Coefficient of Relative Risk Aversion(μ_γ)	.42*	.31	1.17*	1.80*
	(.12)	(.20)	(.11)	(.06)
Std of Log Coefficient of Relative Risk Aversion(σ_γ)	.020	.009	.167*	.601*
	(.07)	(.10)	(.09)	(.18)
Discount Rate(δ)	.17*	.19*	.06*	.28*
	(.06)	(.04)	(.02)	(.08)
Probability of Disaster (p)	.007*	.024*	.051*	.0004
	(.001)	(.008)	(.002)	(.014)
Size of disaster (ϕ)	.81*	.70*	.41*	.0001
	(.11)	(.13)	(.03)	(.013)
Probability of zero income increase of disaster (π)	.38*	.16*	.004	.0004
	(.056)	(.04)	(.005)	(.06)
Per-period participation cost (κ)	.00	.00	.004*	.007*
	(.00)	(.01)	(.001)	(.003)
χ^2 for γ heterogeneity	38.6**	24.7**	705.0**	100.2**
χ^2 for δ heterogeneity	44.3	25.8	1008	468.5

Table 3: Structural Estimates

Less Educated				
Auxiliary Parameters	Young		Old	
	Data	Simulated	Data	Simulated
λ_{01}	.08	.08 (.48)	.10	.10 (.17)
λ_{02}	.06	.06 (.01)	-.45	-.27 (.16)
λ_{03}	-.002	-.002 (.01)	.02	.009 (.23)
λ_{04}	.000	.000 (.10)	-.00	-.000 (.35)
λ_{05}	.015	.19 (.53)	-1.15	.968 (.93)
λ_{06}	.019	-.008 (.54)	.05	-.05 (1.1)
λ_{07}	.000	.000 (.51)	-.00	.00 (1.2)
λ_{08}	.47	.45 (.90)	.46	.43 (.82)
λ_{09}	.30	.27 (2.7)*	.32	.30 (1.7)
λ_{10}	-.01	.001 (1.3)	-.001	.02 (.23)
λ_{11}	.05	.05 (.25)	.05	.02 (2.9)*
λ_{12}	.10	.10 (.26)	.12	.12 (.16)
λ_{13}	.21	.19 (.64)	.27	.24 (.61)

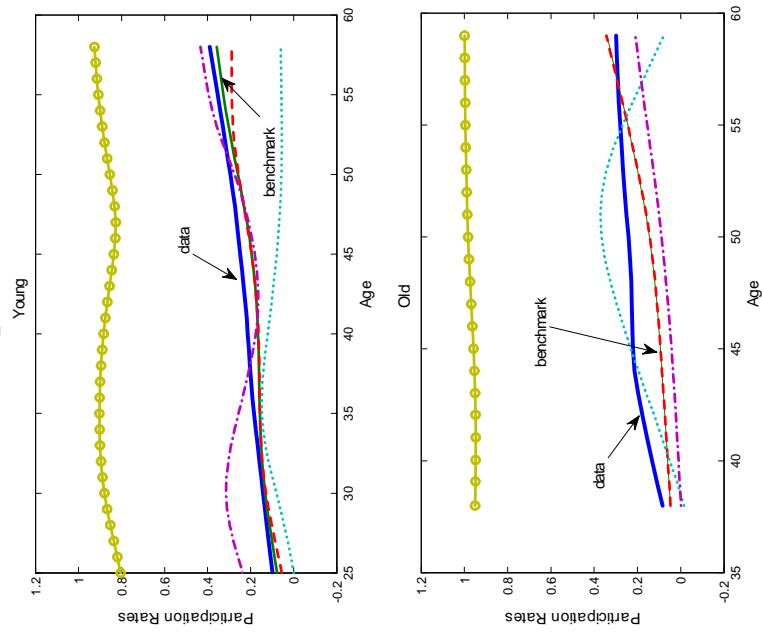
Table 4: Auxiliary Parameters and Simulated Counterparts, Less Educated

Auxiliary Parameters	More Educated			
	Young Data	Young Simulated	Old Data	Old Simulated
λ_{01}	.28	.55 (13)*	.64	.98 (3.7)*
λ_{02}	-.18	-1.76 (9.4)*	-.50	-2.21 (1.2)
λ_{03}	.01	.09 (9.1)*	.02	.09 (1.2)
λ_{04}	-.00	-.001 (8.8)*	-.00	-.001 (1.3)
λ_{05}	-.45	-2.32 (7.3)*	-.86	-3.09 (.80)
λ_{06}	.03	.12 (6.8)*	.05	.13 (.71)
λ_{07}	-0.00	-.00 (7.1)*	-.00	-.001 (.61)
λ_{08}	.53	.64 (11)*	.52	.63 (7.8)*
λ_{09}	.30	.25 (11)*	.31	.30 (1.6)
λ_{10}	.007	.009 (.35)	.01	.02 (1.7)
λ_{11}	.039	.014 (2.2)*	.06	.09 (1.4)
λ_{12}	.29	.24 (4.9)*	.33	.39 (2.3)*
λ_{13}	.54	.38 (11)*	.63	.60 (.74)

Table 5: Auxiliary Parameters and Simulated Counterparts, Less Educated

FIGURE 1
Portfolios of Less Educated

Stock Market Participation



Portfolio Shares

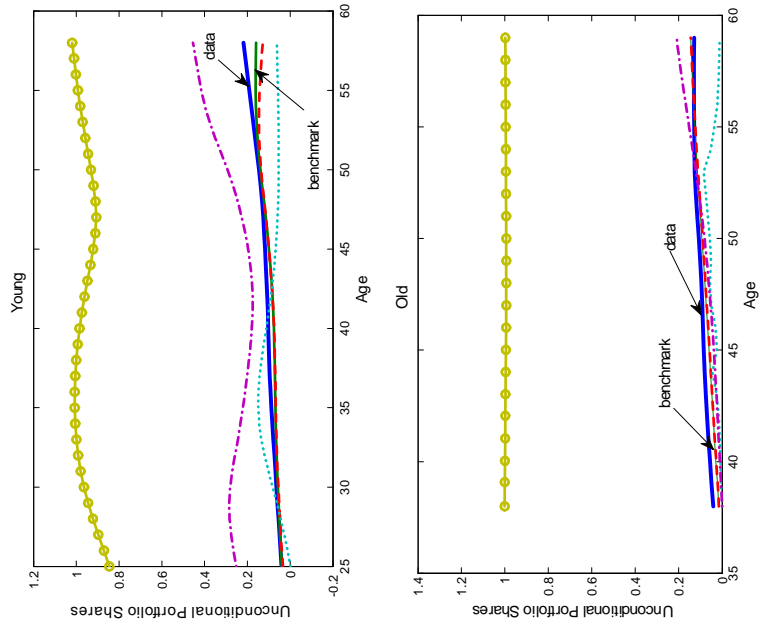
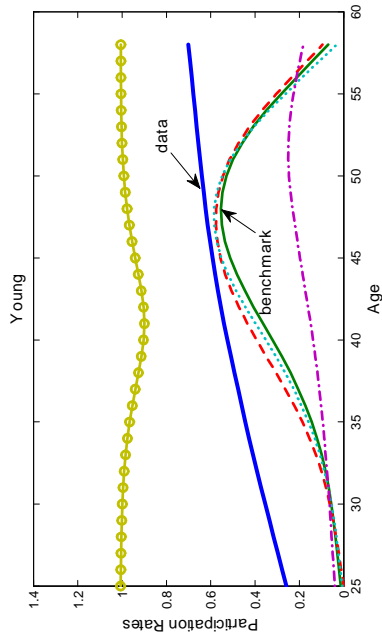


FIGURE 2
Portfolios of More Educated

Stock Market Participation



Portfolio Shares

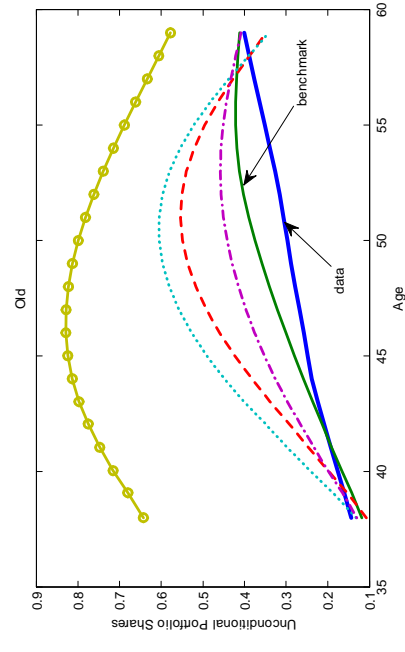
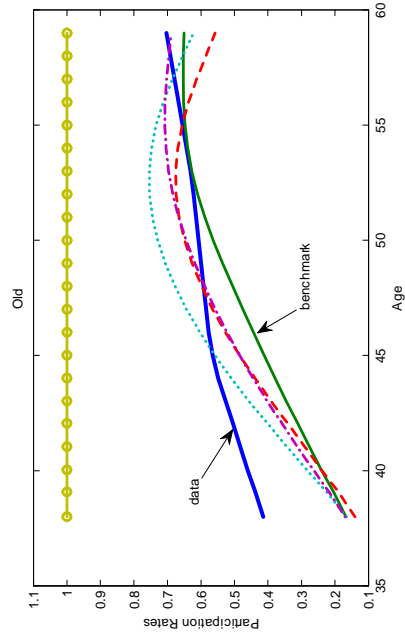
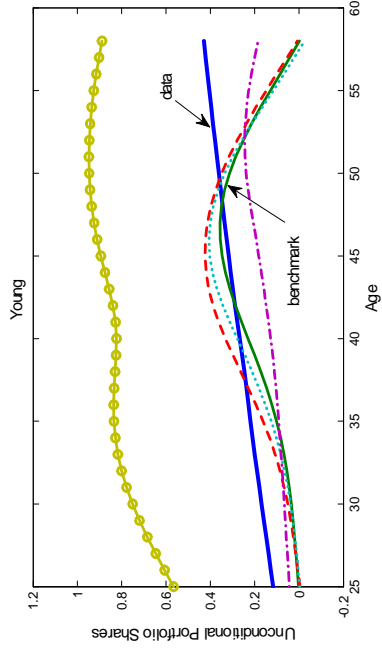


FIGURE 3
Portfolio Shares Conditional on Participation
Less Educated **More Educated**

