Quantitative Economics 3 (2012), 1–28 1759-7331/20120001

Do disaster expectations explain household portfolios?

Sule Alan
Faculty of Economics and CFAP, University of Cambridge and College of Administrative Sciences, Koc University

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 choices of households that have less than a college degree can be partially explained by expectations of stock market disasters only if one allows for a large 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 of the stock market coupled with preference heterogeneity explain their participation and investment patterns.

Keywords. Household portfolios, disasters.

JEL classification. D91, E21.

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 respecifications (Campbell and Cochrane (1999), Bansal and Yaron (2004)), market frictions and preference heterogeneity (Constantinides, Brav, and Geczy (2002)), and model uncertainty (Weitzman (2007)). An alternative strand of the literature emphasizes the limitations of the postwar 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 postwar period).

This idea was first proposed by Reitz (1988) and extended by Barro (2006) and Barro and Ursua (2008). Barro (2006) analyzed 20th century disasters using gross domestic

Sule Alan: sule.alan@econ.cam.ac.uk

I thank Manuel Arellano, Thomas Crossley, Michael Haliassos, John Heaton, Hamish Low, Xavier Mateos-Planas, seminar participants at the University of Cambridge and University College London, and participants in the Household Finance Workshop organized by Bank of Spain for valuable comments and suggestions. I also thank Ruxandra Dumitrescu for excellent research assistance. This research was partly funded by the Center for Financial Analysis and Policy (CFAP). All errors are mine.

DOI: 10.3982/QE128

product (GDP) and stock market data from 35 countries. He suggested that a disaster probability of 1.5–2% a year, with an associated decline in per capita GDP of 15–64% from peak to trough, goes a long way to explain the equity premium puzzle. In follow-up work using aggregate consumption data from 21 countries, Barro and Ursua (2008) calibrated the disaster probability to 3.6% a year with an associated 22% decline in consumption from peak to trough. More recently, Gabaix (2008) proposed 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 microdata: most recent empirical evidence suggests that at least 50% 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), Guiso, Haliassos, and Japelli (2002), and Wachter and Yogo (2010)). Given the rather impressive equity premium in the postwar period, a particular difficulty in reconciling the standard model with observed facts is 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 that 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% 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 could 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 chose to jointly estimate disaster expectations (both probability and expected size) and preference parameters from observed portfolios, and then judge whether the estimates are plausible compared to the historically calibrated values. Moreover, I chose to use a much richer and realistic version of the consumer problem than the original Mehra–Prescott model and the one assumed by Reitz (1988) and Barro (2006). Estimating the entire structural model gives me the opportunity to test several other explanations of equity premium against an explanation based on economic disasters. If the correct quantities are not revealed in an environment that is a lot more flexible than the original, the explanation of the equity premium based on rare disasters would be significantly weakened.

The results in this paper suggest that the expectations of rare disasters can, to a certain extent, explain 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 that are believed to be the relevant portion of the population in terms of aggregate wealth and asset prices.

The structural estimation reported in this paper brings together three large surveys conducted in the United States: the Survey of Consumer Finances (SCF) (1983–2004), which contains detailed wealth and portfolio allocation information, the Consumer Expenditure Survey (CEX) (1983–2004), which contains detailed durable and nondurable expenditure information, and, finally, the Panel Study of Income Dynamics (PSID) (1983–1994), which allows me to calibrate group-specific income process parameters. Limited heterogeneity in all parameters is allowed for by estimating the structural parameters separately for four groups (two birth cohorts by two 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 (the wealthiest households in the sample). Per-period participation costs (approximately 1% of the permanent income) and heterogeneity in the coefficient of relative risk aversion (value of 4 at the 25th percentile and 9 at the 75th percentile) 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. Section 5 discusses the results. Section 6 concludes.

2. The model

I assume that the expected utility function is intertemporally additive over a finite lifetime and the subutilities are isoelastic. The problem of the generic consumer h is

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

where C is nondurable consumption, γ_h is the household-specific coefficient of relative risk aversion, and δ_h is the household-specific rate of time preference. The coefficient of relative risk aversion and the rate of time preference are assumed to be distributed log normally across households such that $\ln\gamma_h \sim N(\mu_\gamma, \sigma_\gamma)$ and $\ln\delta_h \sim N(\mu_\delta, \sigma_\delta)$, respectively.[1] The ideal setup would be to assume a joint distribution for the preference parameters and to 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

[1] The unboundedness of the discount rate and the coefficient of relative risk aversion do not pose any difficulty in estimation, because I use six-point Gaussian quadrature to approximate the distributions, which inevitably bounds possible ranges.

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 the data determine which model fits better.[2]

The end of life T is assumed to be certain. It would be straightforward to incorporate stochastic mortality into the model, but, again, this addition 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 labor income and it evolves as

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

where r^e_{t+1} is the stochastic return from the risky asset, and r is the risk-free rate, S_t is the amount of wealth invested in the risky asset, and B_t is the amount of wealth invested in the risk-free asset.

Note that housing is not included in this model of portfolio choice and consumption. There are two reasons for this exclusion (besides the additional complexity it would add to the solution). First, the purpose of the exercise reported in this paper is to determine whether the original Mehra–Prescott (1985) model augmented with disaster risk as in Barro (2006), which is argued to have solved the asset pricing puzzle, yields the correct quantities (portfolio holdings and consumption growth). Second, adding another risky asset to the portfolio choice set would necessarily lead to *smaller* estimated disaster probabilities. This is because with house price risk, the model will need smaller disaster probabilities to fit the data on quantities. Thus, if I find that the data, seen through the lens of the original model, imply small or zero disaster probabilities, this is very strong evidence against the disaster risk explanation of the asset pricing puzzle.

Turning to the model, following Carroll and Samwick (1997), Y_{t+1} is stochastic labor income, which follows the 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 is subject to multiplicative independent and identically distributed (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 value 0 with some small but positive probability, and are otherwise log normal: $\ln(U_t) \sim N(-0.5\sigma_u^2, \sigma_u^2)$. 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

[2] Alan and Browning (2010) were the first to estimate a joint distribution of the intertemporal allocation parameters using food expenditure data in the PSID. The model of consumption in that paper is much simpler than the model used here.

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 return process is common to all agents, and, as explained below, I allow a link between market disasters and low income realizations.

Introducing a risk of a zero income realization into the life cycle model was proposed by Carroll (1992) and adopted by many subsequent papers.[3] It is important to note that introducing a risk of a zero income realization 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, which leads to rapid wealth accumulation early in the life cycle. If the observed postwar 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.:

$$r^e_{t+1} - r = \mu + \varepsilon_{t+1}, \tag{5}$$

where μ is mean excess return and ε_{t+1} is distributed normally with mean 0 and variance σ^2_ε. 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 a correlation is very weak (see Heaton and Lucas (2000), Davis and Willen (2000)), so I set it to zero.[4]

[3]Since income realizations of zero are rarely observed in the data, it may be more realistic to assume that a labor market stress may be in the form of having to collect unemployment benefits for a given period. One of the models I test against the benchmark presented here assumes a floor above zero for minimum income realizations.

[4]This assumption, coupled with i.i.d. disasters has important implications for the structural estimates. If stock market disasters are modelled as a persistent process, they would be even more painful for the agents, leading to smaller probability estimates to fit the model. Furthermore, if disasters are modelled to be correlated with permanent income (they are correlated only with transitory income in this paper), stock markets would seem even riskier for the agents. Both these extensions would lead to smaller probability estimates in the context of this paper. Note that the evidence on the correlation between stock markets and permanent income is weak for annual frequency, as noted by Heaton and Lucas (2000), but strong for the business cycle frequency (see Lynch and Tan (2011)). Similar argument is valid, albeit weaker, if transitory shocks (unemployment spells) ware modelled more persistently.

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 warlike disaster where governments totally or partially default. Incorporating a perceived probability of government default can be done in the way Barro (2006) suggested. 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 challenging. 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.[5]

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 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

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

subject to borrowing and short-sale constraints

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

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 1, namely the ratio of cash on hand to permanent income. The Bellman equation after normalizing is

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

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. 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

[5] Barro (2006) showed that T bills did quite well in the United States during the great depression, whereas partial default on government debt occurred in Germany and Italy during WWII.

to participate in the stock market 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 optimization problems

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

where

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

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. The term $V^0(x_t, I_t)$ denotes the value the consumer gets by not participating, regardless of whether he has participated in the previous period or not, that is, exit from the stock market is assumed to be costless:[6]

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

subject to

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

where $V^1(x_t, I_t)$ is the value the consumer gets by participating, and F^c is the fixed per-period cost to permanent income ratio, which is zero if the household does not have any stock market investment and is positive if the consumer 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.[7]

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

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

[7]This assumption is fairly standard in the literature. With this simplifying but justifiable assumption, I reduce the total number of state variables to two: age (exogenous) and cash on hand (endogenous).

3. Estimation overview

3.1 *Simulating auxiliary statistics*

The structural estimation is performed for four different groups (birth year–education cohorts). Households are first grouped 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 higher are labelled as "more educated." Within these groups, households are further divided according to their birth year cohorts. Households with heads who were born before 1946 are labelled as "old"; those born after 1946 are labelled as "young." The details of the sample selection are 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.[8] For the benchmark estimation, I allow the discount rate, δ, or the coefficient of relative risk aversion, γ, to be heterogenous across groups and log normally distributed within a group. When the coefficient of relative risk aversion is assumed to be homogenous, it is still allowed to differ across the four groups. Similarly, when the discount rate is assumed to be homogenous, it is still allowed to differ across the four groups.

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

μ_γ is the mean log coefficient of relative risk aversion

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

μ_δ is the mean log-discount rate

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

p is the probability of disaster

ϕ is the size of expected loss in case of disaster

π is the probability of zero income in case of disaster

κ is the per-period stock market participation cost.

This vector of structural parameters 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 $t = 1, \ldots, T$. To perform simulations, I need two $T \times H$ matrices (for permanent and transitory income shocks) and two $H \times 1$ vectors (for initial wealth to income ratio and discount rates, or coefficient of relative risk aversion) of standard normal variables[9] in addition to actual realized stock returns from 1983 to 2004.

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

[9] If $\ln x \sim N(a, b)$, we can simulate draws from a log normal 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 $\mu_x = \exp(a)\sqrt{\exp(b^2)}$ and $\sigma_x^2 = \exp(2a)\exp(b^2)(\exp(b^2) - 1)$.

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, these *simulated* data are 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 the variance of consumption growth. Since the construction of these auxiliary parameters requires observing households for at least two periods and the CEX is repeated cross section,[10] 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.

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 $\lambda_1, \ldots, \lambda_K$. As always, we have a trade-off between the closeness of the aps to structural parameters (the "diagonality" of the binding function; see Gourieroux, Monfort, and Renault (1993) and Hall and Rust (2002)) and the necessary ability to calculate the aps quickly. Many of the aps defined below are closely related to the underlying structure, but none of the aps is a consistent estimator 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 (finw) to permanent income ratio. This helps me identify the discount rate:

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

The next six aps (λ_{02}–λ_{07}) are smoothed age profiles of participation and portfolio shares. I summarize age profiles with a quadratic polynomial, that is, I first run the regressions

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

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

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

where part is a dummy variable that equals 1 if the household owns stocks and equals 0 otherwise, and 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 subsequently becomes clear, these aps play an important role, in conjunction with consumption aps, in pinning down the coefficient of the relative risk aversion parameter and the perceived disaster probability:

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

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

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

$$\Delta \log C = \zeta_0 + \zeta_1 \Delta\text{size} + \epsilon. \tag{18}$$

Then

$$\lambda_{10} = \zeta_0, \tag{19}$$

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

The remaining two aps are the unconditional mean of the portfolio share of stocks and of the participation rate, respectively:

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

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

While the median financial wealth to permanent income ratio and the 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 a sufficiently rich and intuitive characterization of the joint distribution of parameters of interest.

It is important to emphasize that separately identifying the probability and the size of the disastrous event is difficult in this setting. Simply put, there may be many combinations of these two parameters that lead to the same auxiliary environment. However, repeated reestimation with a large set of different starting values converged to the same estimates, suggesting that the model is at least locally identified within the restricted parameter space. 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 disaster size.

4. Data

4.1 *Pseudo-panel construction*

I work with two distinct repeated cross-sectional data sets to obtain the aps. One of the sets contains data on consumption and the other contains data on financial wealth. Using these data, I 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 cover 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 U.S. population. I select married households whose head was self-identified 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 medical 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. As described earlier, first I divide the sample into two broad groups by level of education: college and higher (referred to as more educated) and less than college (referred to as less educated). Then I define two birth cohorts for each education group, giving four groups in total. I restrict the age range to 25–59. 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 data.[11] The mean and the standard deviation of consumption growth over time (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 is available through the family files. The SCF contains the most comprehensive wealth data available among industrialized countries. It is a cross section that is repeated every 3 years. Note that the 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 boostrapping the variance–covariance matrix of the aps.

[11]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).

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 ratio.[12] 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 savings accounts, as they are kept mostly for household 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, certificates of deposit, 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 I calibrate the others. In principle, all the parameters could be estimated through the structural routine, including the 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 four groups. Earnings growth profiles are estimated separately for the two education levels and are 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

TABLE 1. Estimated parameters of income processes.[a]

		Estimated std of Permanent Shocks	Estimated std of Transitory Shocks
Less educated	Young	0.12	0.12
		(0.01)	(0.01)
	Old	0.15	0.13
		(0.01)	(0.01)
More educated	Young	0.11	0.10
		(0.01)	(0.004)
	Old	0.12	0.10
		(0.01)	(0.01)

[a]Standard errors are given in parentheses. Mean predictable income growth for the more and less educated are 0.018 and −0.001, respectively. Source: PSID 1983–1992.

[12]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.

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) suggested that roughly one-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% and the mean equity return is taken to be 6% with a standard deviation (std) of 20% (these values seem to be the consensus; see Mehra (2008)). I set the probability of a zero income realization to 0.00302 (as estimated by Carroll (1992)).

Since 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 to estimate the mean and standard deviation using all 25-year-olds in the data (see Gourinchas and Parker (2002), 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.[13] Unfortunately, we cannot possibly know the level of wealth the older cohorts had when they were young.

To overcome this problem, I devise a novel way to initialize 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 that 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 the 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 the head is 59. This way, I exactly replicate the age structure of the SCF, including the major shortcomings of the data (missing values, triennial structure, and absence of a panel).

5. Estimation results

The benchmark models I estimate have 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 (referred to as γ heterogeneity and δ heterogeneity, respectively, from here on). For all groups, γ heterogeneity yielded the lowest chi-squared criterion. Therefore, all further analyses in this section are based on models with γ heterogeneity (benchmark) and I do not discuss δ heterogeneity.[14]

[13] Ameriks and Zeldes (2004) pointed out the problems of estimating age profiles using cross-sectional data. As the age profiles of portfolio shares are generated by the SCF (cross-sectional data) in this paper, getting the initial wealth conditions right is crucial for the consistency of the structural estimates.

[14] Overall 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 overall fit. Homogeneity of discount rates is rejected by all groups. Full results for δ heterogeneity are available on request.

TABLE 2. Goodness of fit.

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 (nonnested)	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	40,545	8063	11,938	702.7

5.1 *Discussion of the fit*

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 with seven structural parameters and γ heterogeneity (model 1, benchmark). The overall fit is quite reasonable, even though the model is rejected for all four groups based on the chi-squared criterion. One perhaps not very surprising result is that the 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, too skewed and heterogenous for the more educated to be captured by this model. The particular effect of γ heterogeneity can be seen by examining the second row of the same table, where γ heterogeneity is closed down. Increases in chi-squared statistics are sizable enough to warrant rejection of γ homogeneity for all groups. However, the jumps in the chi-squared values are much larger for the educated group (from 705 to 1384 for the young; from 100 to 545 for the old), suggesting a higher degree of preference heterogeneity among this group. The rejection of homogeneity for the educated group is mainly because the consumption growth and wealth moments of the educated are more dispersed than those of the uneducated in the data, and the preference heterogeneity captures part of this dispersion.

The next alternative model I consider replaces the possibility of a zero income realization with the possibility of realizing a strictly positive income floor. This assumption is perhaps more realistic for the more educated households. For example, individuals may lose their jobs and settle for a small fraction of their current income for a year (collecting unemployment benefits, for example). I assume that in normal times this probability is 4% (roughly the natural rate of unemployment in the United States) and the fraction is 30%. As in the benchmark case, I let the probability of such situations arising during the disaster be a free parameter to be estimated. 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

TABLE 3. Auxiliary parameters and simulated counterparts for the less educated.[a]

	Less Educated			
	Young		Old	
Auxiliary Parameters	Data	Simulated	Data	Simulated
λ_{01}	0.08	0.08	0.10	0.10
	(0.48)		(0.17)	
λ_{02}	0.06	0.06	−0.45	−0.27
	(0.01)		(0.16)	
λ_{03}	−0.002	−0.002	0.02	0.009
	(0.01)		(0.23)	
λ_{04}	0.000	0.000	−0.00	−0.000
	(0.10)		(0.35)	
λ_{05}	0.015	0.19	−1.15	0.968
	(0.53)		(0.93)	
λ_{06}	0.019	−0.008	0.05	−0.05
	(0.54)		(1.1)	
λ_{07}	0.000	0.000	−0.00	0.00
	(0.51)		(1.2)	
λ_{08}	0.47	0.45	0.46	0.43
	(0.90)		(0.82)	
λ_{09}	0.30	0.27	0.32	0.30
	(2.7)*		(1.7)	
λ_{10}	−0.01	0.001	−0.001	0.02
	(1.3)		(0.23)	
λ_{11}	0.05	0.05	0.05	0.02
	(0.25)		(2.9)*	
λ_{12}	0.10	0.10	0.12	0.12
	(0.26)		(0.16)	
λ_{13}	0.21	0.19	0.27	0.24
	(0.64)		(0.61)	

[a]Absolute t-ratios are given in parentheses. *, significant at 5%.

group; chi-squared values go down from 1384 to 969 and from 545 to 488 for the young and the old, respectively. This suggests that the risk of a zero income realization is not a good assumption for these households. The opposite is observed for the less educated: large jumps in chi-squared values from 59.5 to 1923, and from 30.8 to 393 for the young and the old, respectively.

The 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 Table 2 (model 4). This variant of the model is estimated by closing down the disaster possibility while keeping γ heterogeneity.[15] In fact, for these households, even the simplest model with no heterogeneity, no disaster expectations, and no entry cost does not lead to a huge jump in the chi-squared criterion (model 5, $\chi^2_{11} = 702.7$), while such a variant makes the fit hopeless for all other groups; see the last row. The essential point from this

[15]The chi-squared increment between the benchmark and model 4 is $\chi^2_3 = 104.9 - 100.2 = 4.7$. Given that the critical value for χ^2_3 is 7.81 at 95%, the model 4 restrictions are not rejected.

table is that the standard model has serious difficulties explaining household portfolios and this difficulty cannot be overcome by assuming expectations of a market disaster. Although this explanation seems to go some distance to explain the behavior of households with very little financial wealth, keep in mind that these are not the individuals who are relevant for prices.

The natural question to ask is, Why does the model need nonzero disaster probabilities for the uneducated to fit the data, while they are not needed for the educated? Technically, the probability of a painful crash in the stock market and its correlation with unemployment (zero income spell) help depress portfolio shares and participation for all agents, but this is particularly needed to match the portfolios of the uneducated group. The standard model implies almost 100% shares in stocks at the low saving levels, so that the estimation procedure needs to depress portfolios of the uneducated a lot more. This can easily be done via disaster expectations, but the real action comes from the hedging demand created by the correlation between the crash and unemployment. Implications of this correlation are very painful for the uneducated, since labor income is the most important lifetime resource to finance consumption for this group. Given that the stock market participation of the educated group is generally higher and they hold most of the shares in the economy, such probabilities are not needed to fit their portfolios (after adding some participation costs to account for some of the nonparticipation).

An economically meaningful 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 3 for the less educated and in Table 4 for the more 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 comes from the first ap (λ_1)—the median financial wealth to permanent income ratio. As can be seen in the first row of Table 4, the model persistently generates higher aps than the data.

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 5) 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 worsen 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 counterparts 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 generates a hump shape for shares and participation that does not exist in the data.

Figure 3 tell us exactly where each 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

TABLE 4. Auxiliary parameters and simulated counterparts for the more educated.[a]

Auxiliary Parameters	More Educated			
	Young		Old	
	Data	Simulated	Data	Simulated
λ_{01}	0.28	0.55	0.64	0.98
	(13)*		(3.7)*	
λ_{02}	−0.18	−1.76	−0.50	−2.21
	(9.4)*		(1.2)	
λ_{03}	0.01	0.09	0.02	0.09
	(9.1)*		(1.2)	
λ_{04}	−0.00	−0.001	−0.00	−0.001
	(8.8)*		(1.3)	
λ_{05}	−0.45	−2.32	−0.86	−3.09
	(7.3)*		(0.80)	
λ_{06}	0.03	0.12	0.05	0.13
	(6.8)*		(0.71)	
λ_{07}	−0.00	−0.00	−0.00	−0.001
	(7.1)*		(0.61)	
λ_{08}	0.53	0.64	0.52	0.63
	(11)*		(7.8)*	
λ_{09}	0.30	0.25	0.31	0.30
	(11)*		(1.6)	
λ_{10}	0.007	0.009	0.01	0.02
	(0.35)		(1.7)	
λ_{11}	0.039	0.014	0.06	0.09
	(2.2)*		(1.4)	
λ_{12}	0.29	0.24	0.33	0.39
	(4.9)*		(2.3)*	
λ_{13}	0.54	0.38	0.63	0.60
	(11)*		(0.74)	

[a]Absolute t-ratios are given in parentheses. *, significant at 5%.

diminishes, especially for the less educated, when we allow for the possibility of disasters. Models 1 and 2 deliver lower portfolio shares in earlier life, and so are much more congruent with the data. This is obviously not the case for the more educated. Note that the main reason for the decisive rejection of the model for the more educated (large chi-squared 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.

5.2 *Disaster expectations*

I now turn to the structural estimates based on the benchmark model. Table 5 presents the estimates for all four groups.[16] Except for the old and more educated group, the

[16]Although the asymptotic standard errors are unreliable for these types of models, I still report them. The precision can be judged in an economically more meaningful way by considering the proximity of the aps generated from the data and from the simulated data at the estimated values; see Tables 3 and 4.

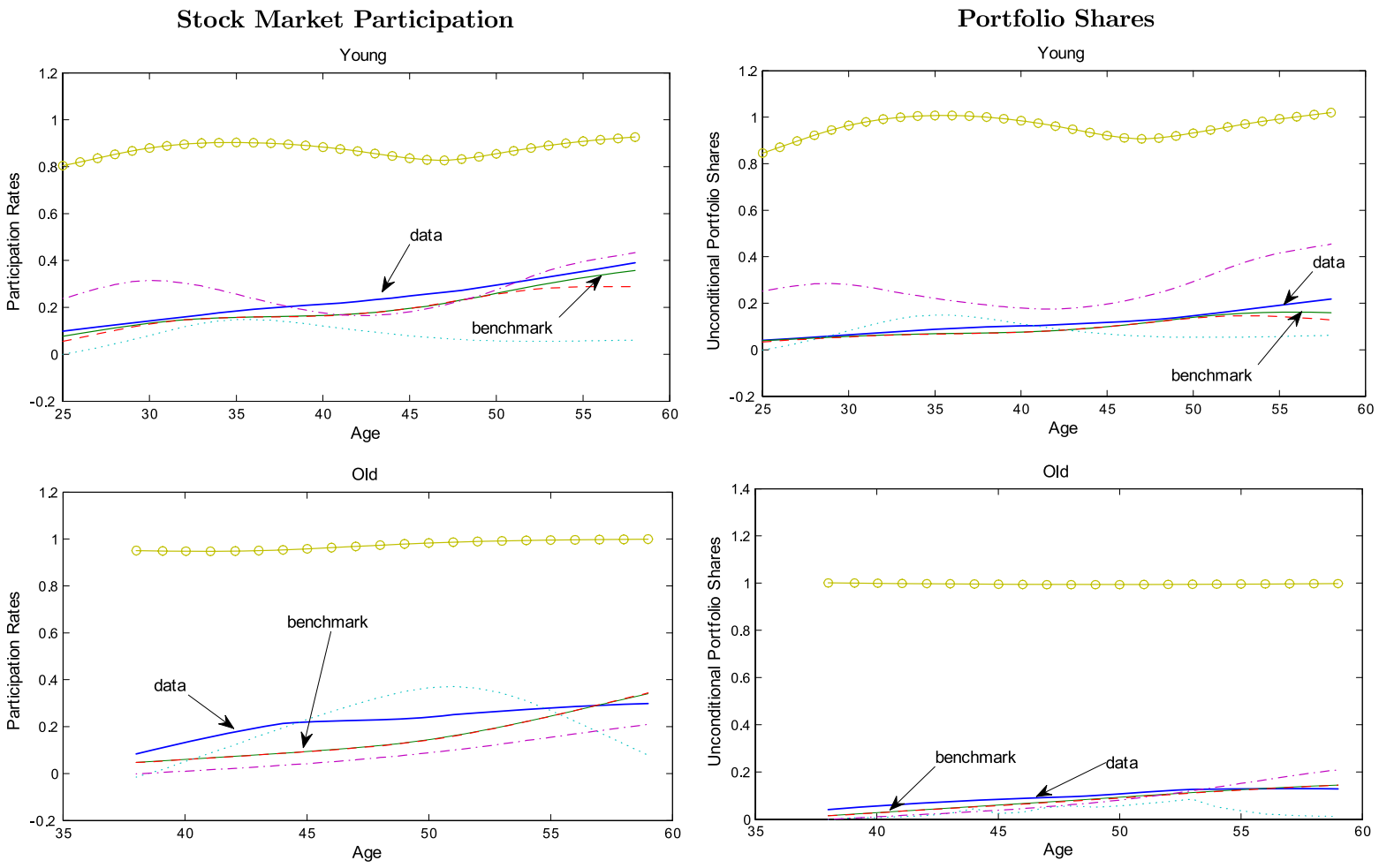

FIGURE 1. Portfolios of less educated. Notation: dashed line, model 2; dotted line, model 3; dotted–dashed line, model 4; circle marks, model 5.

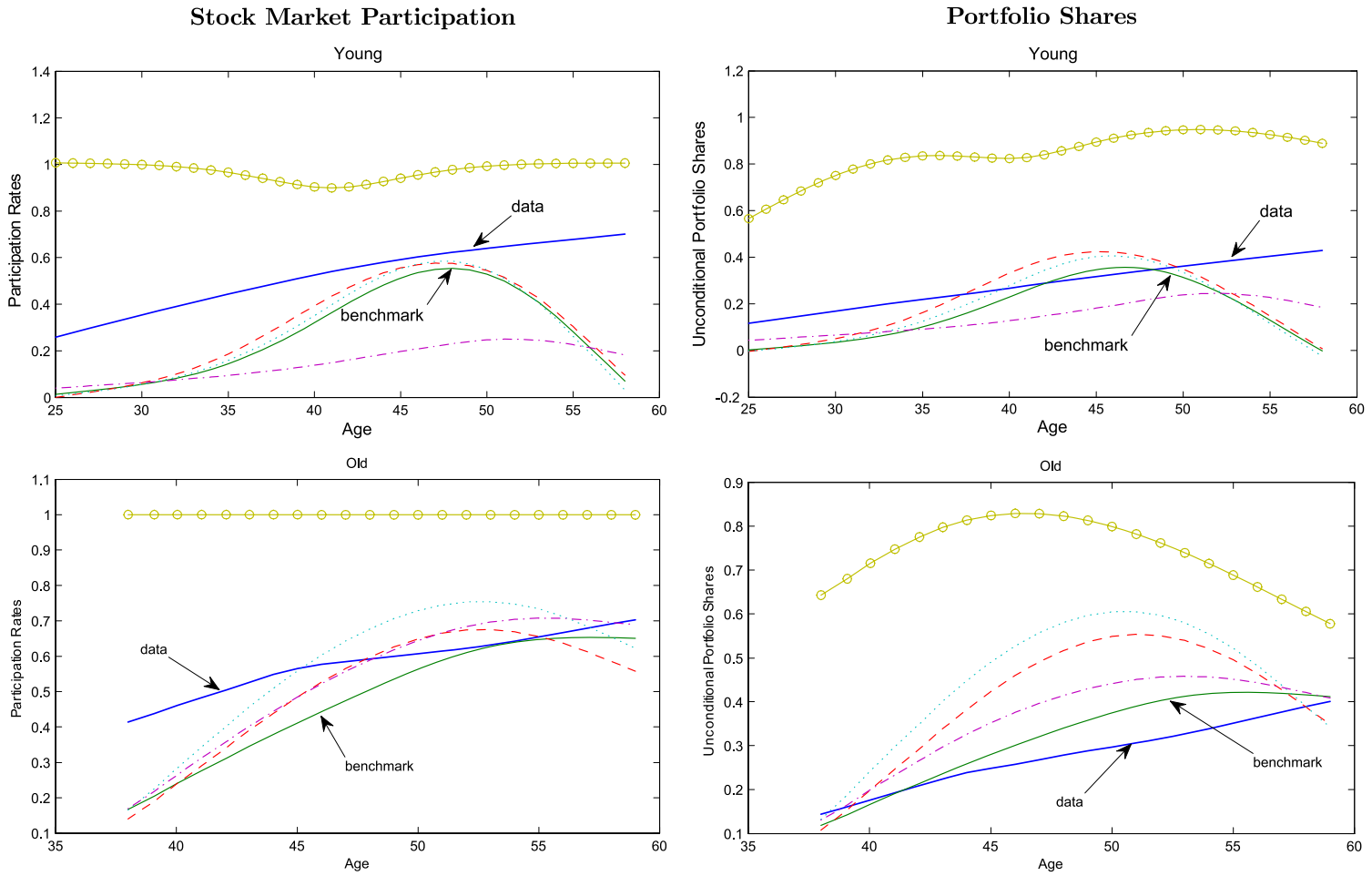

FIGURE 2. Portfolios of more educated. Notation: dashed line, model 2; dotted line, model 3; dotted–dashed line, model 4; circle marks, model 5.

TABLE 5. Structural estimates.[a]

	Less Educated		More Educated	
	Young	Old	Young	Old
Mean of log coefficient of relative risk aversion (μ_γ)	0.42*	0.31	1.17*	1.80*
	(0.12)	(0.20)	(0.11)	(0.06)
Std of log coefficient of relative risk aversion (σ_γ)	0.020	0.009	0.167*	0.601*
	(0.07)	(0.10)	(0.09)	(0.18)
Discount rate (δ)	0.17*	0.19*	0.06*	0.28*
	(0.06)	(0.04)	(0.02)	(0.08)
Probability of disaster (p)	0.007*	0.024*	0.051*	0.0004
	(0.001)	(0.008)	(0.002)	(0.014)
Size of disaster (ϕ)	0.81*	0.70*	0.41*	0.0001
	(0.11)	(0.13)	(0.03)	(0.013)
Probability of zero income increase of disaster (π)	0.38*	0.16*	0.004	0.0004
	(0.056)	(0.04)	(0.005)	(0.06)
Per-period participation cost (κ)	0.00	0.00	0.004*	0.007*
	(0.00)	(0.01)	(0.001)	(0.003)
χ^2_6 for γ heterogeneity	38.6**	24.7**	705.0**	100.2**
χ^2_6 for δ heterogeneity	44.3	25.8	1008	468.5

[a]Asymptotic standard errors are given in parentheses. **, preferred model; *, significant at 5%.

probability of a disaster and the expected size of the disaster 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 a zero income realization in the case of a disaster is implausibly high for the less educated young (38%). It is not as large for the less educated 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, the more educated older cohort (the wealthiest households in the sample) do not appear to expect such disasters. The very fact that these households drive aggregate wealth casts serious doubt on an explanation of the equity premium based rare disasters.

How do my estimates compare with Barro's (2006) 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% decline in the stock market wealth in 4 years. Since disasters are assumed to strike in an i.i.d. fashion (as in Reitz (1988) and Barro (2006)), the size estimates are not directly comparable, but can be interpreted as *total* expected wealth loss in the event of a disaster. On the other hand, I can directly compare my estimated disaster probabilities with Barro's calibrated values. An estimate of 80% loss seems to be too large, especially since it is coupled with 38% probability of zero income realization for the less educated young. For this group, the estimated disaster probability is about 1%. For the old and less educated, this parameter is estimated to be around 2%. These estimates are perfectly in line with Barro's calibrated values (1.5–2%). However, for the more educated young, although the expected size estimate seems reasonable (41%), the esti-

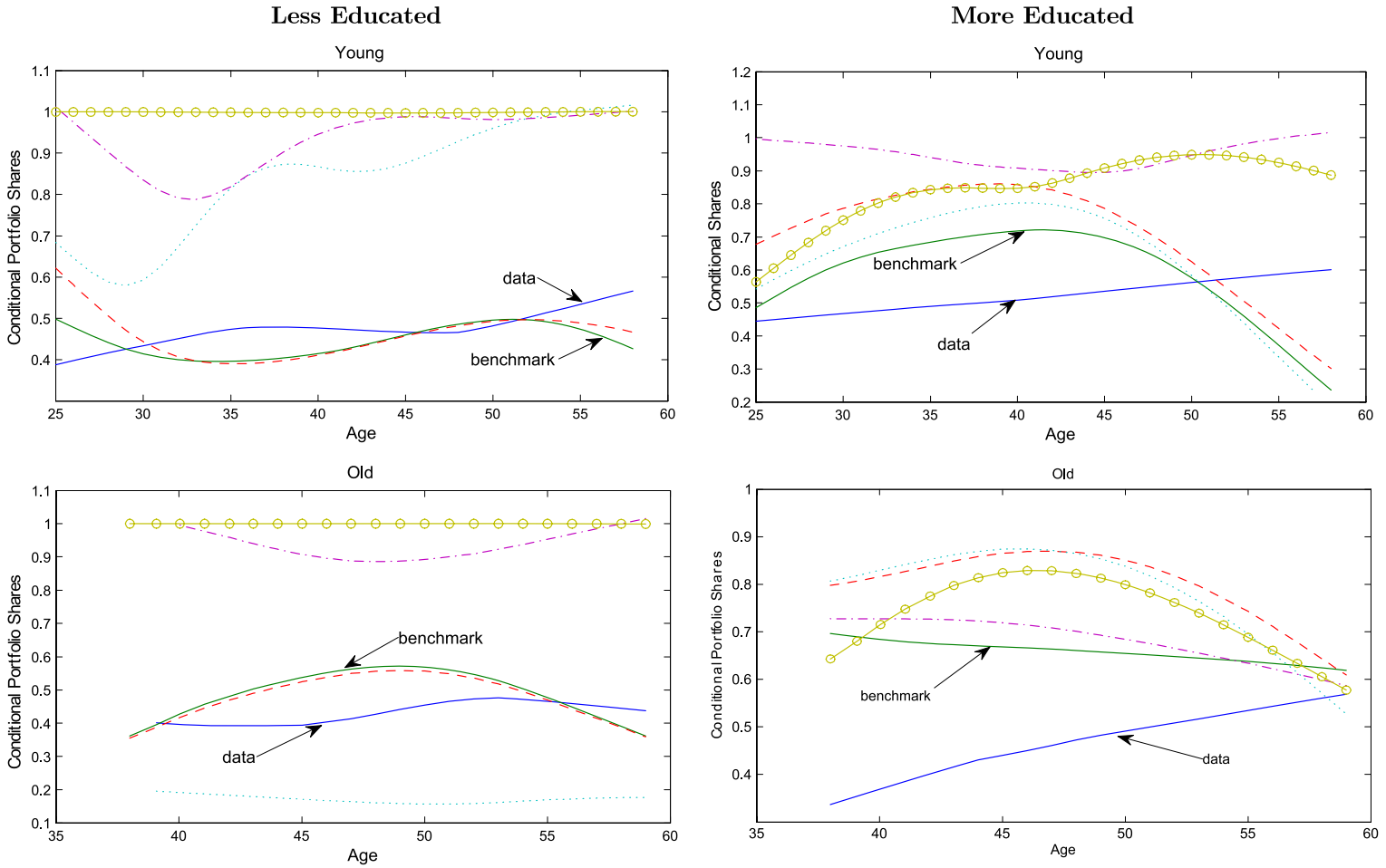

FIGURE 3. Portfolio shares conditional on participation. Notation: dashed line, model 2; dotted line, model 3; dotted–dashed line, model 4; circle marks, model 5.

mated disaster probability is around 5%, which is too high compared to the calibrated values in Barro (2006) and Barro and Ursua (2008).

5.3 *Other findings*

A striking result of the estimation is that there is substantial variation in preference parameter estimates across education groups, but not so much across birth cohorts within education groups. Consistent with Alan and Browning (2010), the less educated seem to have a lower relative risk aversion.[17] 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 microdata on consumption (see Attanasio, Banks, Meghir, and Weber (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). In terms of heterogeneity within cohort–education cells, 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 coefficients 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 3.6.[18] It is important to note that rejection of preference homogeneity does not necessarily mean that there is a genuine underlying preference heterogeneity. In fact, it is extremely hard to identify preference heterogeneity using the cross-sectional dispersion of consumption growth. For example, measurement error in consumption, taste shocks, and heterogeneity in income processes can generate significant dispersion in consumption growth across agents, even with homogenous preference parameters. Such dispersion in consumption growth can be observationally equivalent to the dispersion generated by the genuine preference heterogeneity.

Another interesting result in this paper is that participation cost estimates are zero for the less educated, but are positive and significant for the more educated. There is now a sizable 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

[17] This result contradicts Calvet, Campbell, and Sodini (2007); it is mainly the outcome of the constant relative risk aversion specification. Prudence is linked to risk aversion in this specification, and the fact that educated people hold stocks and accumulate wealth is consistent with the high coefficient of relative risk aversion (and high prudence); the negative effect of high risk aversion on stock holding is dominated by the positive effect of prudence on wealth accumulation.

[18] Table 5 reports the mean log coefficient of relative risk aversion and its standard deviation. The median values and percentiles that I report here come from the simulation of the relevant log-normal distribution (at the estimated parameters) for 100,000 households for each group.

TABLE 6. Composition of the structural estimates of the benchmark and restricted model 4.[a]

	Less Educated		More Educated	
	Young	Old	Young	Old
Coefficient of relative risk aversion, benchmark (μ_γ)	1.52	1.36	3.22	6.05
Coefficient of relative risk aversion, restricted (γ)	1.66	2.04	2.22	4.21
Discount rate, benchmark (δ)	0.17	0.19	0.06	0.28
Discount rate, restricted (δ)	0.15	0.14	0.16	0.25
Per-period participation cost, benchmark (κ)	**0.0%**	**0.0%**	**0.4%**	**0.7%**
Per-period participation cost, restricted (κ)	**1.0%**	**2.15%**	**1.5%**	**2.94%**

[a]The restricted model (model 4 in Table 2) sets disaster probabilities to zero.

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. While such transaction costs go some way to reconcile observed patterns of stock market participation, they are not sufficient to explain other observed portfolio features, particularly shares conditional on participation. When I 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), then participation, portfolio shares, and shares conditional on participation come down to reasonable levels, making the participation cost assumption unnecessary for the less educated. However, these costs still seem to be important for the more educated, especially for the older cohort, where the per-period participation cost is estimated to be approximately 1% of permanent income.

It is worthwhile to note that participation cost seems to act counterintuitively in the model at first: It is not required to fit the portfolios of the uneducated, but seems to be important for the educated. The reason behind this is that it is modelled as a fraction of permanent income. This makes participation cost important for the high permanent income group. After depressing the portfolio shares and participation through hedging demand generated by disaster expectations and unemployment, portfolios of the uneducated do not need participation cost to fit the model. To illustrate this point, in Table 6 I present the structural estimation results of the restricted model (model 4 in Table 2) where disaster expectations are closed down. One can immediately see that in the absence of disaster expectations, the standard model tries to get a better fit by using the participation cost and preference heterogeneity. Here, the estimated participation costs are all positive for all groups (even for the uneducated), but the chi-squared values are very large, except for the old and educated, for whom we did not detect disaster expectations in the first place. It is also important to note that if, instead, a constant per-period cost were assumed for all agents (like monetary trading costs), it would have less of an effect on the wealthy households' portfolios unless it was implausibly high.

Overall, the results suggest that allowing for rare disasters does lead the life cycle portfolio choice 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 based on 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. The message from the data is mixed at best.

6. Conclusion

This paper evaluates the argument that rare economic disasters, once taken into account, can solve asset pricing puzzles. It is natural to assess whether correct quantities can be obtained from a framework that claims to yield correct prices. I show that it is difficult to reconcile actual quantities in the microdata 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 possessvery 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. Portfolio decisions of wealthier households can be better explained by a combination of preference heterogeneity and transaction costs. I estimate virtually zero probability of disaster for these households. One could add housing to the model estimated in this paper. However, as noted above, adding another risky asset to the portfolio choice set would necessarily lead to *smaller* estimated disaster probabilities. This is because with house price risk, the model will need smaller disaster probabilities to fit the data on quantities. Thus the disaster risk explanation of the asset pricing puzzle cannot be rescued by adding housing to the model.

I do not test the disaster explanation directly against explanations based on preference respecifications. 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 respecifications 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 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 in other dimensions. Explanations based on business cycle risk (a way to correlate stock returns with labor earnings indirectly) may be a more promising route as in Lynch and Tan (2011). However, the need to reconcile other aspects of intertemporal behavior such as consumption and savings within the same framework remains crucial.

Appendix 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 put, 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 is

$$s_T = 0, \qquad b_T = 0.$$

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

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

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 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.

Appendix 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 (2010) for details). Suppose that we observe $h = 1, 2, \ldots, H$ units over $t = 1, 2, \ldots, T$ 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 $\{(Y_1, X_1), \ldots, (Y_H, X_H)\}$, 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 the generalized method of moments 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 $\{(Y_1^1, X_1), \ldots, (Y_H^1, X_H), (Y_1^2, X_1), \ldots, (Y_H^S, X_H)\}$; 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 that 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 Gourieroux, Monfort, and Renault (1993) indirect inference procedure, the auxiliary parameters are maximizers of a given data-dependent criterion that 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,

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

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,

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

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:

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

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

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

The choice I adopt is the (bootstrapped) covariance matrix of γ^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.

References

Abel, A. B. (1990), "Asset prices under habit formation and catching up with the Joneses." *American Economic Review, Papers and Proceedings*, 80, 38–42. [24]

Alan, S. (2006), "Entry costs and stock market participation over the life cycle." *Review of Economic Dynamics*, 9 (4), 588–611. [13, 22]

Alan, S. and M. Browning (2010), "Estimating intertemporal allocation parameters using simulated residual estimation." *Review of Economic Studies*, 77 (4), 1231–1261. [4, 22]

Ameriks, J. and S. P. Zeldes (2004), "How do household portfolio shares vary with age?" Working paper, Columbia University. [13]

Attanasio, O., J. Banks, C. Meghir, and G. Weber (1999), "Humps and bumps in lifetime consumption." *Journal of Business and Economic Statistics*, 17, 22–35. [22]

Bansal, R. and A. Yaron (2004), "Risks for the long run: A potential resolution of asset pricing puzzles." *Journal of Finance*, 4, 1481–1509. [1]

Barro, R. J. (2006), "Rare disasters and asset markets in the twentieth century." *Quarterly Journal of Economics*, 121, 823–866. [1, 2, 4, 6, 20, 22]

Barro, R. J. and J. Ursua (2008), "Rare disasters and asset markets." *American Economic Review: Papers and Proceedings*, 98 (2), 58–63. [1, 2, 22]

Bertaut, C. C. (1998), "Stockholding behavior of U.S. households: Evidence from the 1983–89 survey of consumer finances." *Review of Economics and Statistics*, 80, 263–275. [2]

Bound, J. (1994), "Evidence on the validity of cross-sectional and longitudinal labor market data." *Journal of Labor Economics*, 12, 345–368. [13]

Bound, J. and A. B. Krueger (1991), "The extend of measurement error in longitudinal earnings data: Do two wrongs make a right?" *Journal of Labor Economics*, 9, 1–24. [13]

Browning, M., A. Deaton, and M. Irish (1985), "A profitable approach to labor supply and commodity demands over the life cycle." *Econometrica*, 53, 503–543. [9, 11]

Browning, M., M. Ejrnaes, and J. Alvarez (2010), "Modelling income processes with lots of heterogeneity." *Review of Economic Studies*, 77 (4), 1353–1381. [25]

Cagetti, M. (2003), "Wealth accumulation over the life cycle and precautionary savings." *Journal of Business and Economic Statistics*, 21 (3), 339–353. [22]

Calvet, L. E., J. Campbell, and P. Sodoni (2007), "Down or out: Assessing the welfare costs of household investment mistakes." *Journal of Political Economy*, 115, 707–747. [22]

Campbell, J. Y. and J. H. Cochrane (1999), "By force of habit: A consumption based explanation of aggregate stock market behavior." *Journal of Political Economy*, 107, 205–251. [1, 24]

Carroll, C. D. (1992), "The buffer-stock theory of saving: Some macroeconomic evidence." *Brookings Papers on Economic Activity*, 2, 61–156. [5, 6, 13]

Carroll, C. D. and A. Samwick (1997), "The nature of precautionary wealth." *Journal of Monetary Economics*, 40, 41–71. [4]

Constantinides, G. M. (1990), "Habit formation: A resolution of the equity premium puzzle." *Journal of Political Economy*, 98, 519–543. [24]

Constantinides, G. M., A. Brav, and C. Geczy (2002), "Asset pricing with heterogeneous consumers and limited participation: Empirical evidence." *Journal of Political Economy*, 110, 793–824. [1]

Davis, S. and P. Willen (2000), "Occupation-level income shocks and asset returns: Their covariance and implications for portfolio choice." Working Paper 7905, NBER. [5]

Deaton, A. (1991), "Saving and liquidity constraints." *Econometrica*, 59, 1221–1248. [4]

Gabaix, X. (2008), "Variable rare disasters: A tractable theory of ten puzzles in macro-finance." *American Economic Review: Papers and Proceedings*, 98, 64–67. [2]

Gourieroux, C., A. Monfort, and E. Renault (1993), "Indirect inference." *Journal of Applied Econometrics*, 8, S85–S118. [9, 26]

Gourinchas, P. O. and J. A. Parker (2002), "Consumption over the life cycle." *Econometrica*, 70 (1), 47–89. [13, 22]

Guiso, L., M. Haliassos, and T. Japelli (2002), *Household Portfolios*. MIT Press, Cambridge, Massachusetts. [2]

Hall, G. and J. Rust (2002), "Econometric methods for endogenously sampled time series: The case of commodity price speculation in the steel market." Cowles Foundation Discussion Papers 1376, University of Yale. [9, 25, 26]

Heaton, J. and D. Lucas (2000), "Portfolio choice in the presence of background risk." *The Economic Journal*, 110, 1–26. [5]

Lynch, A. W. and S. Tan (2011), "Labor income dynamics at business-cycle frequencies: Implications for portfolio choice." *Journal of Financial Economics*, 101, 333–359. [5, 24]

Mehra, R. (2008), "The equity premium puzzle: A review." *Foundations and Trends in Finance*, 2 (1), 1–81. [13]

Mehra, R. and E. C. Prescott (1985), "The equity premium: A puzzle." *Journal of Monetary Economics*, 15, 145–161. [1, 4]

Reitz, T. A. (1988), "The equity risk premium: A solution." *Journal of Monetary Economics*, 22, 117–131. [1, 2, 20]

Vissing-Jorgensen, A. (2002), "Towards an explanation of household portfolio choice heterogeneity: Nonfinancial income and participation cost structures." Mimeo, University of Chicago. [7]

Wachter, J. A. and M. Yogo (2010), "Why do household portfolios shares rise in wealth?" *Review of Financial Studies*, 23, 3929–3965. [2]

Weitzman, M. L. (2007), "Subjective expectations and asset-return puzzles." *American Economic Review*, 97 (4), 1102–1130. [1]

Submitted January, 2011. Final version accepted November, 2011.