Entry Costs and Stock Market Participation Over the Life Cycle*

Sule Alan
Department of Economics
York University,
4700 Keele St. Toronto
Ontario, Canada M3J 1P3
salan@econ.yorku.ca

January 2005

Abstract

Several explanations for the observed limited stock market participation have been offered in the literature. One of the most promising one is the presence of market frictions mostly in the form of fixed entry and/or transaction costs. Empirical studies strongly point to a significant structural (state) dependence in the the stock market entry decision, which is consistent with costs of these types. However, the magnitude of these costs are not yet known.

This paper focuses on fixed stock market entry costs. I set up a structural estimation procedure which involves solving and simulating a life cycle intertemporal portfolio choice model augmented with a fixed stock market entry cost. Important features of household portfolio data (from the PSID) are matched to their simulated counterparts. Utilizing a Simulated Minimum Distance estimator, I estimate the coefficient of relative risk aversion, the discount factor and the stock market entry cost. Given the equity premium and the calibrated income process, I estimate a one-time entry cost of approximately 2 percent of (annual) permanent income. My estimated model matches the zero median holding as well as the hump-shaped age-participation profile observed in the data.

Keywords: Entry costs; Stock market; Structural estimation
JEL Classification: G11, D91

*I thank Thomas Crossley, John Knowles, Alex Michaelides, Gregor Smith and Mike Veall for helpful comments and suggestions. All errors are mine. This research is supported financially by the Social Sciences and Humanities Research Council of Canada.

1 Introduction

Recent empirical evidence suggests that, in any developed country, at least fifty percent of households do not hold equities directly or indirectly (through mutual funds, retirement accounts etc.)[1]. Furthermore, the median age of entry into the stock market is quite high (around forty five). Given the rather impressive equity premium over the last ninety years these facts are difficult to reconcile with the standard intertemporal portfolio choice model. Although we began to see a substantial increase in stock market participation and much more sophisticated household portfolio structures over the 1990s, the observed aversion to stockholding and differences in participation patterns across households even after controlling for age, income, wealth and education still pose a great challenge to the life cycle model[2]. Among several explanations offered in the literature, the emerging consensus seems to be some sort of perceived "stock market entry cost" typically in the form of time cost of information acquisition for the new entrants.

Most studies in the literature present evidence regarding the presence of an entry cost without inferring its magnitude (Vissing-Jorgensen 2002 and Guiso et al 2002). A few use simulation techniques to illustrate the potential size of the entry cost necessary to generate complete non-participation for different preference parameters (Haliassos and Bertaut 1997, Polkovnichenko 2001 and Haliassos and Michaelides 2003). While convincing, none of these studies attempts to quantify entry cost within a complete structural estimation framework. Naturally, the magnitude of such a cost cannot be estimated within a reduced form setting. This paper takes an important step forward in identifying fixed stock market entry costs by reconciling a fairly rich version of the standard life cycle portfolio choice model with observed participation patterns. In doing so, I estimate the stock market entry cost and intertemporal allocation parameters; the coefficient of relative risk aversion and the discount factor. In terms of novelty, this study is the first attempt to quantify the economic magnitude of the one-time stock market entry cost by accounting for observed "limited participation" within a complete structural setting under income and return uncertainty.

Costs that deter entry in the stock market may take several forms. Vissing-Jorgensen (2002)

[1] Sweden has the highest indirect stock holding (54% in 1999) followed by the U.S. (48% in 1998) See Guiso et al (2002).

[2] For instance in 1998 only 19% of the American households were holding equity directly in publicly traded corporations. This number is the highest (27%) for the UK among all developed nations. See Bertaut (1998) and Guiso et al (2002).

categorizes participation costs as fixed entry costs, fixed and variable transaction costs and per period trading costs. She points to symptoms of strong structural (state) dependence in participation and stock holding decisions as evidence of fixed entry and transaction costs and she estimates per period trading costs[3]. Structural dependence in participation manifests itself by making participation in a given period more likely if the household participated in the previous period. Using panel data on household indirect stockholding she finds that lagged participation is a very significant determinant of current participation. Another related study by Guiso et al (2002) presents cross-country evidence on the presence of participation costs. In their detailed descriptive work, they conclude that the cross-country differences in participation rates can be better justified by different institutional and informational barriers to entry across countries than differences in stock returns.

As far as quantifying fixed entry costs is concerned, a study of a particular importance is Haliassos and Michaelides (2003). Rather than estimating the magnitude of the entry cost, they simulate a stochastic portfolio choice model and then for different structural parameters they calculate the entry cost required to keep all agents out of the stock market. However, they did not consider life cycle participation profiles. Instead, they adopt an infinite life setting where the delayed entry observed in the data cannot be modelled. To my knowledge, Faria (2000) is the only study that estimates fixed entry cost. However, he uses an infinite life general equilibrium model with no equity or labor income risk. Naturally, his results are extremely sensitive to the equity premium assumed.

The entry cost considered in this paper is a one time fee; a first time investor must pay to participate in the stock market and it has a very broad definition. Simply, it can be thought of as the value of time spent to understand the basic functioning of stock markets, to learn how to follow price movements, how to trade, how to assess risk and return relationship for an optimal portfolio choice, etc. Since I think of this as a time cost incurred to acquire information, and hence related to the opportunity cost of time (the wage), it is plausible to formulate this cost as proportional to permanent income [4]. It is important to note that this fee is paid (if ever paid) only once over the

[3] Participation costs that do not create structural state dependence (per period trading costs) can be inferred within a reduced form setting. Costs, on the other hand, that create correlation of participation or stock holding decisions across periods (entry costs, fixed and variable transaction costs) can only be identified with a structural estimation. Vissing-Jorgensen (2002) concludes that a per period transaction cost of as low as \$50 can explain the choices of half of non-participants. Paiela (2001) estimates per period cost bounds in terms of forgone utility gains and finds that at least \$31 is needed to generate the observed participation pattern for a consumer with log utility.

[4] This assumption turns out to be necessary in order to make our estimation strategy feasible. With this simlifying but justifiable assumption, the number of state variables is reduced.

entire life cycle. Once it is paid, the household is free to re-enter the stock market (if exited some time after entry) without incurring any further cost; once learned, such knowledge is not forgotten for the rest of the life cycle[5]. In my empirical work below, I provide some evidence on the plausibility of this assumption.

The estimation involves matching the age-profile (corrected for cohort effects and family size) and structural dependence in participation observed in the U. S. Panel Data of Income Dynamics (PSID) with their simulated counterparts. To do so, I use a Simulated Minimum Distance (SMD) estimator. I match some carefully selected auxiliary statistics (parameters of an auxiliary model). Using a probit regression as my auxiliary model, I estimate the coefficient of relative risk aversion and subjective discount rate to be 1.625 and 0.0874 respectively. The stock market entry cost is estimated to be 2.15 percent of annual permanent income. All parameters are estimated with considerable precision. Although the overidentification restrictions are rejected, the estimates of the intertemporal allocation parameters are within the range of previous estimates in the literature. Moreover, the simulated participation profiles (based on the estimated parameters) seem to be in line with the actual profiles, and the estimated model matches the structural dependence in the data.

The reminder of the paper is organized as follows: In Section 2 I lay out a structural life-cycle model of portfolio choice. I also numerically solve and simulate the model at some illustrative parameter values, in order to demonstrate the potential effects of a stock market entry cost. I then turn to the structural estimation of the model. The data I use for estimation are described in Section 3, and the estimation method is discussed in Section 4. Section 5 presents my results. Section 6 concludes.

[5]In the standard life cycle setting the only reason to exit the stock market is to finance consumption. A buffer stock saver may have to liquidiate his shares if he gets hit by an adverse income shock that leaves him with insufficient cash on hand to afford his optimal consumption.

2 The Model

2.1 A Basic Life-Cycle Model of Portfolio Choice

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

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

where C is non-durable consumption (separable from durable consumption), γ is coefficient of relative risk aversion, δ is rate of time preference. I assume that the end of working life T is certain[6]. Following Deaton (1991), I define endogenous state variable cash on hand as the sum of financial assets and labor income and it evolves as follows:

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

where r^e_{t+1} is stochastic return from the risky asset representing the stock market, r is the risk-free rate which can be thought of as bonds, T-bills and bank accounts, 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 (1992) 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 and it is subject to multiplicative iid shocks N_t. Current income Y_t is composed of a permanent and a transitory component U_t. Further details regarding income processes is given below.

Following Gourinchas and Parker (2002) I define a retirement value function so that the consumption rule at the time of retirement is

$$C_T = \lambda_1(X_{T+1} + H_{T+1}) \tag{5}$$

where λ_1 is marginal propensity to consume out of wealth, and H_{T+1} is exogenously accumulated illiquid wealth, which is modelled proportional to the last permanent income hP_T. This assumption

[6]It would be straight forward to incorporate a stochastic mortality into the model. This additional complexity though is not likely to contribute significantly to the estimation results. This argument is especially stronger considering I am interestred only in working life at this point.

leads to positive wealth at the time of retirement. I do not consider retirement years or explicitly model a bequest motive.

The excess return on the risky asset is assumed to be iid:

$$r_{t+1}^{e} - r = \mu + \varepsilon_{t+1} \tag{6}$$

where μ is the mean excess return and ε_{t+1} is distributed normally with mean 0 and variance σ_{ε}^{2}. A positive correlation between return and permanent income shocks along with shortsale constraints also generates delayed entry since such a correlation leads to a hedging demand for the riskless asset. However, the empirical evidence for the existence of such a correlation is rather weak. Heaton and Lucas (2000a) find a positive significant correlation between stock returns and intrapreneurial income (around 0.2) but small and insignificant values for other occupation groups. Estimates obtained by Davis and Willen (2000) range between 0.1 to 0.3 for a college educated group and significantly negative for a lower education group. Since such a correlation is not strongly evident in the data I choose to set it to zero, both in my illustrative simulations and in my estimation procedure.

The maximization problem involves using the Bellman equation and solving the recursive equation via backward induction. The problem is:

$$V_t(X_t, P_t, H_t) = \max_{S_t, B_t} \left\{ \frac{(C_t)^{1-\gamma}}{1-\gamma} + \beta E_t V_{t+1}\left[(1 + r_{t+1}^{e})S_t + (1+r)B_t + Y_{t+1}, P_{t+1}, H_{t+1}\right] \right\} \tag{7}$$

subject to shortsale and borrowing constraints,

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

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

In order to make the estimation computationally feasible I normalize the necessary variables by dividing them by permanent income (see Carroll 1992). By doing this, I reduce the number of endogenous state variables to one, namely, ratio of cash on hand to permanent income. The resulting Bellman equation after normalizing is as follows:

$$V_t(x_t) = \max_{s_t, b_t} \left\{ \frac{(c_t)^{1-\gamma}}{1-\gamma} + \beta E_t (G_{t+1}N_{t+1})^{(\gamma-1)} V_{t+1}\left[(1 + r_{t+1}^{e})s_t + (1+r)s_t/G_{t+1}N_{t+1} + U_{t+1}\right] \right\} \tag{8}$$

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

Normalized consumption at the final period is:

$$c_T = \lambda_0 + \lambda_1 x_T \tag{9}$$

where $\lambda_0 = h\lambda_1$. In order to obtain the policy rules for the earlier periods I define a grid for the endogenous state variable x and maximize the above equation for every point in the grid. Value and policy functions are approximated with cubic splines.

2.2 Adding a Stock Market Entry Cost

I now assume that participating in the stock market requires an entry fee, whereas investing in the risk-free asset is costless. When augmented with the fixed entry cost, the solution of the above model requires additional computations. The optimizing agent now has to decide whether to enter into stock market or not before he decides how to allocate his wealth. This is done by comparing the discounted expected future value of participation and that of nonparticipation in every period. This results in the following optimization problems:

$$V_t(x_t, I_t) = \max_{0,1} \left(V^0(x_t, I_t), V^1(x_t, I_t)\right) \tag{10}$$

where

$$V^0(x_t, I_t) = \max_{s_t, b_t} \left\{ \frac{(c_t)^{1-\gamma}}{1-\gamma} + \beta E_t V_{t+1}\left[x_{t+1}, I_{t+1}\right] \right\} \tag{11}$$

subject to

$$x_{t+1} = (1+r)b_t / G_{t+1} N_{t+1} + U_{t+1} \tag{12}$$

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

$$V^1(x_t, I_t) = \max_{s_t, b_t} \left\{ \frac{(c_t)^{1-\gamma}}{1-\gamma} + \beta E_t V_{t+1}\left[x_{t+1}, I_{t+1}\right] \right\} \tag{13}$$

subject to

$$x_{t+1} = \left[(1 + r^e_{t+1})s_t + (1+r)b_t - E^c\right] / G_{t+1} N_{t+1} + U_{t+1} \tag{14}$$

$V^1(x_t, I_t)$ is the value consumer gets by participating. The entry cost is proportional to current permanent income ($P_t E^c$). The parameter E^c is fixed and it is 0 if the agent is already participated once in the stock market and it is positive if he has never participated before. It is important to note that, like exit, re-entry is costless.

In each time period, given his current participation state, the agent first decides whether to enter the stock market or not (or stay if he is already in) by comparing the expected discounted value of each decision. 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. The details of the solution method are given in Appendix A.

2.3 Illustrative Simulations

Solution and simulation of the standard portfolio choice model under labor income uncertainty (i.e., without a participation cost) is well described in the literature[7]. To illustrate the potential effect of a fixed stock market entry cost on consumption, investment and participation decisions, I solve the model with and without an entry cost for 50 years. I then simulate the life cycle paths of consumption, investment in risky and riskless assets and participation. The parameter values assumed for this simulation exercise are given in Table 1.

To solve and simulate the model, I need the income process. For the purpose of these simulations, I assume that the growth rate of income is nonstochastic and $G = 1$[8]. 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 otherwise lognormal such that $\ln(U_t) \sim N(-0.5\sigma_u^2 \ , \sigma_u^2)$. Similarly, permanent shocks N_t are iid and $\ln(N_t) \sim N(-0.5\sigma_n^2 \ , \sigma_n^2)$. Assuming that the 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 all agents face the same return process.

Introducing zero income risk into the life cycle model is motivated by Carroll (1992) and adapted by Gourinchas and Parker (2002). This assumption has important implications for optimal behavior. Given the fact that iso-elastic utility function yields infinite marginal utility of consumption at zero consumption, backward induction dictates that a consumer who faces such a risk optimally chooses never to borrow. Thus, consumer saves at every level of wealth and more importantly, the Euler

[7] See Cocco et al (1997) and Haliassos and Michaelides (2002).

[8] Note that tis assumption is only for these illustrative simulations. When I turn to structural estimation of the model, although I continue to assume nonstochastic income growth, I use estimated age-growth rate profiles.

equation is always satisfied. Although the characterization of the model is presented with this zero income probability assumption, estimation of the structural parameters is performed by assuming a Deaton-type explicit borrowing constraint.

The panels in Figure 1 display the policy function differences for different stages of life. The figures are obtained by simply subtracting policy functions from the model with an entry cost from policy functions from the model without an entry cost. The first two panels show the effect of entry cost on optimal stock holding at very old and very young ages respectively[9]. At young ages, optimal stock holding in the presence of an entry cost is lower for all cash on hand levels. The difference is quite large at the lower wealth levels where the consumer is still out of the stock market as he cannot afford to pay the entry cost. The wedge becomes much smaller after stock market entry occurs although the amount of wealth invested in the stock market is still lower than it is without entry cost. Interestingly, at older ages entry does not take place as soon as it is affordable. The consumer decides to participate in the stock market only if he has considerable accumulated wealth since his investment horizon is not long enough to take advantage of the equity premium. As can be seen in panel 1 of Figure 1, entry takes place only at a very high cash on hand level (note the sharp drop in differences) and stock holding after the entry is lower than it is without entry cost.

Not surprisingly, policy function differences for bond holding are qualitatively the reverse of that for stock holding (panels 3 and 4). At very young ages, optimal bond holding in the presence of entry cost is much higher at low wealth levels since it is the only investment tool available to a consumer who cannot yet afford the entry cost. With an entry cost, bond holding remains slightly higher even after the stock market entry takes place. At old ages total bond holding at the very high wealth levels is very large under an entry cost - so much so that total investment (stocks plus bonds) under an entry cost is higher.

The last two panels illustrate the effect of entry cost on consumption functions. At very young ages, optimal consumption is lower (reflecting higher saving) if the consumer does not face any entry cost (difference is negative). After the entry takes place the difference is not as big although it is still negative. The difference between consumption functions presents a very interesting pattern for older ages (panel 5). At the very low wealth levels, we observe no difference, at middle levels (still no

[9] Note that this a finite life model where policy rules are functions of age as well as cash on hand. The solution of the model for 50 periods results in 50 different policy rules for stocks, bonds and consumption. See Appendix A for the details of the solution method.

entry) consumption under entry cost is higher (or total investment is lower). As wealth accumulates, although stock holding is still lower, bond holding becomes so high that total investment in the model with an entry cost becomes higher. This in turn leads to higher consumption growth.

The implications of entry cost for consumption and saving can be illustrated more clearly comparing the life cycle paths of consumption, bond and stock holding. Panels 1, 2 and 3 in Figure 2 display the difference in life cycle paths. Using the policy functions and random income and return draws, $10,000$ ex-ante identical paths are generated with and without entry cost. Figures are obtained by subtracting average paths from the model with an entry cost from those from model with no entry cost. The last panel shows the life cycle participation paths with and without an entry cost (note the difference between them).

At young ages, consumption without entry cost is naturally low as consumers accumulate wealth through aggressive stock holding. As wealth is accumulated, higher consumption is enjoyed. However at very old ages and high wealth levels, higher stock holding in the absence of entry cost leads to lower consumption levels. At older ages, consumers are better off with lower stock holding and higher bond holding since that means a lower correlation between aggregate risk and consumption. Life cycle path of stock holding/bond holding under entry cost is strictly lower/higher. The differences are sharper at older ages leading to a hump shaped difference in life cycle consumption paths.

The last panel of Figure 2 shows the life cycle path of participation with and without entry cost. It is obvious that the standard model without entry cost has no hope of matching the participation pattern observed in the data since the model predicts participation at every stage of the life cycle. The implied mean participation is much higher than those observed in the data. More specifically, the model without entry cost implies 100% participation rate at all ages while the model augmented with only 1 percent entry cost implies a hump shaped participation profile with about 70 percent mean participation rate at prime ages.

I now turn to the structural estimation of the model with an entry cost, beginning with a discussion of the data.

3 Data

The proposed estimation procedure necessitates the use of panel data on portfolio composition. I use the PSID wealth supplements conducted in 1984 and 1989. To make the sample more representative, the original census sample and the Latino sample are excluded.[10] Nevertheless, the sample I use for estimation is not fully representative of the US population since I exclude split up families to create a two-period panel. Moreover, the sample includes only households that reported all necessary demographic information; family size, race, age, and marital status as well as their portfolio choices for both years. Female headed households and black households are excluded. Because I do not explicitly model retirement years, a bequest motive or educational choices, I take only households whose heads were older than 24 in 1984 and younger than 60 in 1989. Students and retirees are also excluded. The final sample has 1294 households. Participation statistics for the full and final (estimation) samples are presented in Table 1. More general facts regarding household portfolios in the U.S. are well documented by a number of researchers including Vissing-Jorgensen (2002) and Guiso et al (2002).

Because the estimation method requires simulating data from the underlying structural model and since the structural model I use is rather time consuming to solve, I focus on estimating only the entry cost and intertemporal allocation parameters. The parameters of the income and asset return processes used in the estimations are calibrated.

Income process parameters required to implement my estimation procedure are the average age-income growth profile, and the variances of the permanent and transitory components of labor income. To estimate these parameters for each education group I follow Carroll and Samwick (1997). I estimate income process parameters for the sample period covering 1981-1992 (12 years). Income data in the PSID refers to previous year's income. I define non-financial income broadly enough to account for possible insurance schemes available to households, such as unemployment insurance and social assistance. Total household nonfinancial income is total labor income plus unemployment insurance, workers compensation, social security, supplemental social security, child support, the value of food stamps and some other transfers. Real income data are calculated using the Consumer Price Index. Following Carroll and Samwick (1997) I assume an income process that

[10] Although wealth data is also available in 1994, my structural estimation only employs the 1984 and 1989 data because I have access to income data only until 1992.

can be decomposed into permanent and transitory components. The logarithm of permanent income p_t^i for each household follows random walk with drift:

$$p_t^i = g_t^i + p_{t-1}^i + z_t^i \tag{15}$$

where p_t^i is the logarithm of permanent income of ith individual in period t, g_t^i is income growth (likely to be a function of individual characteristics and demographics) and z_t^i is mean zero iid shocks with variance σ_z^2.

Then, the logarithm of current income y_t^i evolves as:

$$y_t^i = g_t^i + p_{t-1}^i + z_t^i + \varepsilon_t^i \tag{16}$$

where the ε_t^i are mean zero iid transitory shocks with variance σ_ε^2.

First, I estimate the average age-income growth profile by simply regressing income growth on occupation, industry dummies, age, age- squared, age cubed, race, marital status, and family size. Estimated average age-income growth profiles are obtained by taking the predicted values from this regression and calculating age-specific averages. Predictable income growth has important implications for household portfolio composition. If an individual expects high income growth, depending on the other parameters of his utility function, he may want to borrow against his future income when young if he is not facing borrowing constraints. In my setting, borrowing against future labor income is not allowed.

Second, I estimate the error structure of the income process described above. For this, first, I regress the logarithm of real income on age dummies, marital status, family size and race. Then I construct differenced regression residuals. I define (following the notation and the procedure of Carroll and Samwick 1997)

$$r_d^i = Y_{t+d}^i - Y_t^i \tag{17}$$

where Y_t is the residuals obtained from the log real income regression. Assuming a constant growth rate

$$r_d^i = (z_{t+1}^i + z_{t+2}^i + ... + z_{t+d}^i) + \varepsilon_{t+d}^i - \varepsilon_t^i \tag{18}$$

Given

$$Var(r_d) = d\sigma_z^2 + 2\sigma_\varepsilon^2 \tag{19}$$

I combine all possible series of $Var(r_d)$ and d and regress $Var(r_d)$ on a constant and d for each education group. The results are presented in Table 3[11].

I use value weighted and dividend adjusted annual returns on NYSE and AMEX between 1950 – 1992. Return and risk-free rate data were obtained from the Center for Research in Security Prices (CRSP) at the University of Chicago. Annual inflation series (to calculate real returns) were obtained from the U. S. Department of Labor Bureau of Labor Statistic. Years were chosen so that the oldest household head in 1984 data (56 years of age) was at the age of 22 in 1950. The mean equity premium calculated is 6% with the standard deviation of 17%. The model is solved using these two empirical moments of the return distribution. In order to generate simulated paths I use realized returns. The risk-free rate is calculated by taking the mean of the real annual 3-month T- Bill rate (3%).

Following Gourinchas and Parker (2002), the retirement value function parameters λ_0 and $\lambda_1$are set to be 0.001 and 0.71 respectively.

The model presented in the previous section assumes that the stock market entry cost is paid, at most, once over the life-cycle: stock market re-entry is costless. Moreover, if the entry cost is a one time cost, it is most natural to think of it as an information cost, which in turn makes modelling it as proportional to permanent income attractive (because the opportunity cost of information acquisition depends on the wage.) Fortunately, I can perform a simple empirical test to asses the plausibility of costless re-entry. It is now possible to observe the portfolios of the PSID households in 3 different time periods (1984,1989 and 1994). Participation in 1989 is a good predictor of participation in 1994 (this is the structural dependence reported in the literature.) However, if the entry cost is a one time cost, then participation in 1989 should have no effect on participation in 1994 among those households that participated in 1984. Empirically, this turns out to be the case.[12]

4 Estimation Method

4.1 Simulated Minimum Distance

I employ a simulation based estimation technique. Hall and Rust (1999) refer to the general technique as Simulated Minimum Distance (SMD) since it is based on matching (minimizing the distance

[11] In Table 3 I only report average income growth. For the estimations, the model is solved and simulated with the age specific income growth rates.

[12] Full results on of this analysis are available upon request.

between) statistics from the data with statistics from a simulated model. The class of SMD estimators includes the EMM procedure of Gallant and Tauchen (1996) and the Indirect Inference methods of Gouriéroux, Monfort and Renault (1993). Here I present a short account of the method as applied generally to panel data; see Hall and Rust (1999) and Alvarez et al (2003) for further details.

Suppose that we observe $h = 1, 2..H$ cross section units over $t = 1, 2...T$ periods and we wish to model variable Y with a set of explanatory variables X. Thus we have panel data on H agents. In my case Y contains the participation indicator. 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.[13] If this distribution is tractable enough we could derive a likelihood function and use either maximum likelihood estimation or simulated maximum likelihood estimation (if the distribution is difficult to integrate). Alternatively, with some moment conditions of this distribution for observables we use GMM to recover estimates of θ. In the cases where integration is analytically infeasible we can use SMD if we can simulate Y_h given the observed X_h and parameters for the model. Thus we choose a integer S for the number of replications and then generate $S*H$ simulated outcomes $\left\{(Y_1^1, X_1), ...(Y_H^1, X_H), (Y_1^2, X_1), ...(Y_H^S, X_H)\right\}$; these outcomes, of course, depend on the model chosen ($F(.)$) and the value of θ taken in the model.

Thus we have some actual data on H agents for T periods and some simulated data on $S*H$ units that have the same form. The next step 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, we first choose a set of auxiliary statistics that are used for matching and obtain them using the data in hand. The natural question is: which statistics? Gallant and Tauchen (1996) suggest first finding a 'score generator' (flexible quasi-likelihood function) which nests the true model, and then using the score vector from this as auxiliary parameters. 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 DGP. Both of these approaches are motivated by attempts to derive estimators that have efficiency properties that are close to MLE. In Hall and Rust (1999), the auxiliary parameters are simply statistics that describe important aspects of the data. For now I disregard the efficiency issues and follow this approach.

[13] This could be generalised to allow for dependence on the initial values of the Y variables, as in Alvarez *et al.* (2003).

The estimation method amounts to choosing J sample auxiliary parameters $\alpha_j^D(Y_h, X_h)$ where $J \geqslant m$ so that we have at least as many auxiliary parameters as structural parameters. These are simple statistics of the data. Denote J-vector of auxiliary parameters derived from data by α^D. The general assumption is that sample auxiliary parameters have a mean and covariance under the true distribution. The vector of population auxiliary parameters are defined as

$$\alpha^0(\theta) = E_0(\alpha^D(Y_h, X_h)) \tag{20}$$

where E_0 denotes expectation with respect to the true distribution. Assuming independence across cross-section units we have

$$p\lim \alpha^D(Y_h, X_h) = \alpha^0(\theta) \tag{21}$$

so that $\alpha^D(Y_h, X_h)$ is a consistent estimator for $\alpha^0(\theta)$. The true covariance matrix for the auxiliary parameters is denoted by $V_0(\theta)$. In the estimations I shall use a bootstrap estimator for this and denote it by V_b. The assumption here is that the bootstrap gives a consistent estimate of V_0 as $H \to \infty$.

Same statistics can be calculated S times using the simulated data where the DGP is the stochastic life cycle model of portfolio choice. Denote the corresponding vector by $\alpha^S(\theta)$ and

$$\alpha^S(\theta) = \frac{1}{S}\sum_{s=1}^{S} \alpha(Y_h^s, X_h) \tag{22}$$

where (Y_h^s, X_h) is a set of H simulated path conditional on X_h and on the initial values in the observed data. Assume that $\alpha^S(\theta)$ is twice continuously differentiable and the model is well-specified so that simulated auxiliary parameters converge to a deterministic vector as S becomes large

$$\lim_{S\to\infty} \alpha^S(\theta) = \alpha^\infty(\theta) \tag{23}$$

Assume that $\alpha^\infty(\theta)$ is one-one (for global identification of θ) and continuous and if we have cross-section independence

$$p\lim_{H\to\infty} \alpha^S(\theta) = \alpha^\infty(\theta) \tag{24}$$

Identification requires that the Jacobian of the mapping from model parameters to auxiliary parameters has full rank:

$$rank\left(\nabla_\theta \alpha^S(\theta)\right) = m \text{ with probability } 1 \tag{25}$$

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

$$\hat{\theta}_{SMD} = \arg\min_{\theta} \left(\alpha^S(\theta) - \alpha^D(Y_h, X_h)\right)' A \left(\alpha^S(\theta) - \alpha^D(Y_h, X_h)\right) \tag{26}$$

under the assumption that θ is contained in a compact set and the differentiability assumption on $\alpha^S(\theta)$, the SMD estimator always exists. The estimate is locally unique if $\left(\nabla_\theta \alpha^S\left(\theta_{smd}\right)\right)$ is full rank and globally unique if $\alpha^S(\theta)$ is one-one.

4.2 Choice of Auxiliary Parameters

In general, the number of potential auxiliary statistics can be larger than the model parameters. For instance, it is plausible to match second, third, even fourth moments, and all cross moments, in which case the estimation procedure provides an opportunity to test the overidentifying restrictions. Naturally, the use of the correct weighting matrix becomes relevant if one proceeds with this strategy. However, the monte carlo experiments performed by Alvarez et al. (2003), suggest that SMD estimators do not perform well in environments where we have large numbers of overidentifying restrictions.

I use a simple auxiliary model to generate auxiliary statistics. In particular, my auxiliary statistics are the constant and coefficients on age, age- squared, and lagged participation, in a linear probability model or probit for stock market participation. I chose this approach for two reasons. First, it connects my auxiliary statistics to the existing reduced form literature. Second, with this approach I have only 1 overidentifying restriction (whereas if I matched the entire age-profile, I would have many over-identifying restrictions).

The specification of the linear probability model is:

$$I_t = \alpha_0 + \alpha_1 Age_t + \alpha_2 Age_t^2 + \alpha_3 I_{t-1} + \alpha_4 FS_t + \sum_{i=1}^{8} \beta_i Cohort_i + \epsilon_t \tag{27}$$

where I_t denotes participation status, FS_t denotes family size in period t. The index of the probit model is specified in the same way.

When constructing age-participation profiles it is particularly important to adjust for cohort effects for two reasons. First, different cohorts may have different participation attitudes and this will not be represented in the simulated data. For example if the earlier cohorts did not know much about the stock market or if they had a particular dislike for risky investments, the estimated age-participation profile will have a spurious decline at older ages. Second, earlier cohorts may have had lower initial wealth and, consequently, lower stock market participation rate. This may also cause a bias in the estimated profiles by, again, pushing the profiles down at older ages. Note that the omitted cohort dummy is for the middle (5th of 9) cohort.

After estimating the equation above I take the coefficients $\alpha_0,\ \alpha_1,\ \alpha_2,\ \alpha_3$ as auxiliary parameters. Since the focus of this paper is to estimate the entry cost, the choice of the auxiliary parameter that identifies it deserves a particular attention. It is now well known that the presence of a fixed entry cost leads to structural dependence in participation decision. Hence, a significant coefficient on past participation (α_3) seems to be a natural choice to identify the entry cost. If the entry cost were zero, participation in any given period would be independent of participation in the previous period and this coefficient would be statistically not different from zero. I performed a simulation check to establish the relationship between this coefficient and the structural entry cost parameter and found a strong monotone relationship over a wide range of entry cost values[14].

A possible concern here is that the structural dependence in the data is the spurious result of unobserved heterogeneity.[15] I require a method deal with these effects in the auxiliary estimations since the same effects are not present in the simulated data. For dynamic panel data models Heckman (1981) proposes an approximate random effect estimator to remove (or diminish) the effect of unobserved heterogeneity. This estimator can be used for $T = 2$ or higher and its bias is not as big as the fixed effects estimator. The estimator assumes that individual effects are not correlated with the right hand side variables other than the lagged variable. This is a plausible assumption for the auxiliary model I use in the paper since I do not include moments of labor income in the participation equation. These variables would be very likely to be correlated with unobserved individual effects[16]. The estimation method amounts to specifying participation for both periods 1984 and 1989 separately where the equation for 1989 has 1984's participation as an additional regressor,

[14] The relationship displays a concave structure in general. However, it is not possible to obtain a sensible relationship for very large entry cost values due to computational reasons. At very large values, there is simply no entry and the relationship is not defined.

[15] Note that the inclusion of cohort dummies in the specification means that the coefficients on age and age-squared are estimated using longitudinal variation in age (because cohort is collinear with cross-sectional variation in mean age.) Thus concern arising from the possibility of unobserved heterogeneity is limited to the coefficient on lagged participation.

[16] Since the purpose of the paper is to match the age profiles and the structural dependence in stock market participation, these variables can be left out as long as they are left out in both actual and the simulated data.

and allowing for correlation in errors across two equations.[17],[18]

4.3 Initial Conditions

Because I do not observe all households at the beginning of their life cycle (24 years of age) I need to estimate an initial wealth distribution to initiate simulations. For this, I use a sample of households whose head is between 24 and 28 years of age and fit a lognormal distribution to the empirical distribution of the ratio of wealth to permanent income. Wealth is the sum of financial and real assets and permanent income is the predicted value obtained from the regression of labor income on household characteristics, and occupation, education and industry dummies. The mean of the logarithm of the wealth to permanent income ratio is estimated to be -3.15 and the standard deviation 1.96. In the simulation part of the estimation procedure every household begins its working life by drawing an initial wealth to permanent income ratio from this distribution. Furthermore, initial permanent income is normalized to 1. Although it is quite common for young households to start their working life with debt, I do not allow for borrowing in the current setting. Thus the lognormal initial wealth distribution is appropriate.

4.4 Estimation in Practice

The structural estimation involves a grid search over three parameters: the coefficient of risk aversion, discount rate and the entry cost. Other parameters: income variance, income growth, return process parameters and the risk free rate are calibrated. Steps in the estimation procedure are as follows:

1. Obtain the necessary sample auxiliary parameters from the data: These are the first four coefficients of the participation regression (equation 27).

2. Obtain the variance-covariance matrix of the auxiliary parameters through a panel data bootstrap procedure.

[17] See Vissing-Jorgensen 2002 for the details of this estimator.

[18] Alternatively, one could adopt a more structural approach and estimate a heterogenous model where typically, a parametric distribution is assumed for the coefficient of relative risk aversion instead of imposing a single parameter for everyone. In such a case we would estimate the moments of the assumed distribution. This approach is used in Alan and Browning (2003) for the estimation of heterogenous discount factor. Unfortunately, I cannot follow the same route in this paper since I do not have sufficiently long panel data on portfolio composition. Such estimation requires construction of auxiliary statistics for each individual in the actual data. The resulting empirical distribution of individual statistics would be used to identify the structural distribution of the parameter of interest. With 2 observations per individual, this is not possible.

3. Solve the underlying structural model and simulate participation paths that imitate the data patterns (the age composition and the panel structure of the sample). Obtain simulated auxiliary parameters by estimating a participation equation on the simulated data (a probit relating current participation to age, age- squared and participation lagged five periods).

4. Minimize the distance between the simulated and actual data auxiliary parameters using the bootstrapped variance-covariance matrix as the weighting matrix.

Exact imitation of the panel data in hand when simulating fictitious households is extremely crucial in the estimation procedure. Remember two important features of the real data. First, households are observed at different ages in only two points in time. A 35 year old observed in 1984 is re-observed in 1989 as a 40 year old. Since I do not observe his wealth at the age of 24, I draw a random number from the estimated initial wealth distribution and simulate his participation path up until he is 40. Then, I take his participation status for the age of 35 and 40 only. I repeat this for all 35 year olds (however many they are) in the sample. The procedure performs the same simulation and selection technique for all age groups. At the end, I obtain the exact age composition of the actual data with 5-year apart panel structure.

Second, the observed data does not reveal the actual participation path[19]. For example, a household who did not participate in 1984 may have participated in 1985 and if he did not participate in 1989 (when I re-observe him) he is recorded as a complete non-participant. Situations similar to this may cause underestimation of the entry cost. By replicating the exact structure of the data, I hope to address this problem. Note that the same situation can arise in the simulated data. Depending on the initial condition and income realizations, a household may: never participate; participate sometime but not be observed while participating; participate sometime and be observed participating in both years or in one of the years; or it may participate all along. It should be reemphasized that as long as the problems of the real data are replicated in the simulated data SMD estimator is consistent[20]. This ability of the SMD estimator to overcome complicated sampling and selection issues in the real data simply by replicating the sampling and selection procedures on the simulated data is very big feature of this method.

[19] I thank Gregor Smith for raising this important point.

[20] The implicit assumption here is that the probability of each path occuring is the same in both simulated and the actual data. Unfortunately, this assumption is not possible to test.

Turning to estimation, after defining grids for all three parameters I calculate the criterion function for every point in the grid keeping other parameters fixed. I initially define 20 grid points for the risk aversion coefficient, 20 for the discount rate and 30 points for the entry cost. The entire procedure is repeated for every defined combination in this three dimensional grid. After narrowing down the possible parameter values I perform a finer grid search to make the criterion function as close as possible to zero. After estimating the parameters, the variance covariance matrix is calculated at the estimated parameter values. It is important to note that the variance-covariance matrix is calculated using the boostrapped variance-covariance matrix of the auxiliary parameters so that precision of the structural estimates takes into account the precision of the auxiliary estimates. Finally, since the model is overidentified, an overidentification test is also performed.

5 Results

Three auxiliary regression models are estimated. For identification, a heterogeneity corrected probit is estimated by restricting age coefficients across equations to be the same. Table 4 presents the auxiliary parameter estimates for linear probability, maximum likelihood probit and heterogeneity corrected probit models. All models display a significant concave age profile even after controlling for cohort effects and family size. Moreover, lagged participation seems to be a very significant determinant of current participation even after controlling for unobserved heterogeneity suggesting a significant "true" structural dependence due to entry cost. Note that the estimated correlation coefficient across residuals for heterogeneity corrected probit is not statistically significant suggesting that heterogeneity is not a serious problem for this particular sample. For this reason, structural estimation was performed using only the parameters of the first two auxiliary models.

Table 5 reports the results of the SMD estimation of the coefficient of relative risk aversion, the discount rate and the entry cost to permanent income ratio for the linear probability and maximum likelihood probit models. The estimates do not seem to be very different in magnitude across auxiliary models. Moreover, all structural parameters are estimated with a remarkable precision for both auxiliary models. The entry cost is estimated to be approximately 2.1% of annual permanent income. Zero median participation is matched precisely.

The entry cost estimates are quite small compared to the values used in Haliassos and Michaelides

(2003). Remember that their work involves experimenting with different values of entry cost to generate nonparticipation. The obvious reason why I am able to generate observed participation pattern with a much smaller entry cost is that my underlying model is a finite life model where agents may not be able to accumulate enough wealth to make it worthwhile to participate in the stock market over their entire life cycle. Note on the other hand, that in an infinite life setting all agents eventually participate if the entry cost is not sufficiently high. In the finite life case, the participation decision depends also on the investment horizon. For instance, an agent who is at the age of $T-1$ and who never participated before will not find it worthwhile to pay the cost and invest in the stock market for a wide range of wealth. Thus, a tiny entry cost will suffice to discourage him. As investment horizon becomes longer, the magnitude of the cost necessary to deter entry becomes larger. At the very extreme, infinite life case, the required cost to keep all agents out of the stock market will naturally be much higher.

The estimates of the coefficient of relative risk aversion are 1.625 and 1.61 for probit and linear models respectively. Both estimates are statistically significant and perfectly in line with the previous estimates based on consumption data. Based on an Euler equation estimation, Attanasio et al (1999) estimate the coefficient of relative risk aversion to be around 1.5. Among the studies which perform structural estimation, the estimates of Gourinchas and Parker (2002) range between 0.28 and 2.29, and the estimates of Alan and Browning (2003) range between 1.2 and 1.95. The estimates of Gakidis (1998) however, are somewhat higher (around 3). The discount rate estimates are also reasonable and precise although somewhat higher than the estimates obtained by Alan and Browning (2003) and Gourinchas and Parker (2002).

Not surprisingly, both estimations resulted in massive a overidentification rejection. Admittedly, the chi square values are too large to be attributed to approximation error due to discretization and the inevitable coarseness of the grid search. However, the simulated auxiliary parameters at the estimated values are very close to their actual data counterparts (see Table 6). This is especially striking for the structural dependence parameter; estimation using the data yields the value 1.272 while the simulations at the estimated structural parameters result in the value 1.246. The estimated structural parameters also do a good job in matching the predicted age-participation profile observed in the data. Figure 3 depicts the simulated and actual predicted age-participation profiles for both auxiliary models. As shown in the first picture, the probit specification match is very good for the early ages. The real success of the model is that given estimated structural parameters, it is able to generate the observed humped shape age-participation profile and structural dependence in the data.

The fact that the structural parameters are estimated with a remarkable precision calls for a discussion of identification. It appears that a small change in parameter values results in considerable changes in simulated profiles and structural dependence. Table 6 presents three counterfactual experiments where each counterfactual represents a small deviation of the structural parameters from the optimum (i.e., from the actual estimates). For these experiments, I use the baseline estimates derived from the probit auxiliary model. The experiments involve computing predicted mean participation at the ages of 30 (early in the life cycle), 45 (about peak participation age) and 59 (just before retirement). The first two rows of the table present predicted mean participation and the structural dependence parameter for the actual data and at the actual structural parameter estimates respectively. Even though the participation at early ages are well matched, the height of the predicted age profile is lower than that of the actual data: At the age of 45, estimated mean participation is .634 while the actual data suggest such value to be .754. The linear model seems to do better job in capturing the height of the age profile.

In the first experiment, the coefficient of relative risk aversion is increased from 1.625 to 1.645. A higher risk aversion coefficient is expected to push the estimated peak participation. However, as evident in the table, such a move results in higher early participation and lower structural dependence due to fast wealth accumulation and consequently better affordability of the entry cost; mean participation at the age of 30 increased from .527 to .534 and structural dependence parameter fell from 1.246 to 1.235, both further away from the values obtained from the data.

Increasing the discount rate from .0874 to .088 resulted in values that have quite similar interpretation. Higher impatience prevents the early participation and naturally slows down the wealth accumulation leading to a flatter age-participation profile with higher structural dependence. A higher structural dependence is the obvious artifact of the low wealth accumulation and consequently a higher effective entry cost.

The final experiment is conducted by increasing entry cost from .0215 to .0225. Not surprisingly, a higher entry cost leads to a higher structural dependence parameter (from 1.246 to 1.282). Furthermore, it deters participation in early years leading to slightly lower life cycle wealth accumulation.

6 Conclusion

In the presence of entry costs, stockholding is concentrated at the upper end of the wealth distribution. Such costs discourage small savers by making stock holding not worthwhile for them. These are the investors for whom the entry costs exceeds the optimal value of stock investment. With a small entry cost, these savers are left with only low risk-low return saving tools, such as bank accounts, money market funds and bonds. Naturally, a reduction in the entry costs would result in an increase in the number of stockholders, to whom consequently, more consumption smoothing tools are available. Such an improvement in the capital markets may very well contribute to reducing the need for some public insurance schemes that are designed to help smooth consumption such as unemployment insurance and publicly provided health insurance. For example, in a recent paper, Lentz (2003) emphasizes that the optimal unemployment insurance benefit rate in a search model with savings is quite sensitive to the rate of return on savings. A high rate of return makes it attractive to hold wealth and hence self-insurance is not as costly.

Going beyond elaborating on their symptoms, knowing the magnitude of entry costs is important whether their reduction calls for a public policy or if such action should simply be left to publicly traded corporations and financial intermediaries.

In this paper I set up a structural estimation procedure which involves solving and simulating an intertemporal portfolio choice model augmented with a fixed stock market entry cost. Important statistics of portfolio data (from the PSID) were matched with their simulated counterparts. The latter were obtained from the numerical solution of the model. Utilizing a Simulated Minimum Distance estimator, I estimated the coefficient of relative risk aversion, the discount factor and the stock market entry cost. Given the equity premium and the calibrated income process, I estimated an one-time entry cost of approximately 2 percent of (annual) permanent income. My estimated model matches the zero median holding as well as the hump-shaped age-participation profile observed in the data..

Matching participation rate statistics with their simulated counterparts is a challenging task. It is well documented that as soon as we start operating under labor income uncertainty, the solution of the intertemporal model requires numerical methods. Although the solution methods for these types of models are now standard, tractability can easily disappear with a seemingly small addition to the model. Adding a participation decision to the problem with two controls (risky and risk-free asset holdings), one endogenous state variable (cash on hand) and three stochastic state variables (shocks to risky asset return and shocks to permanent and transitory income) makes the solution fairly complicated and time consuming.

Clearly, there are several different participation costs (whether fixed or variable) that a trader incurs not only upon entry but also over the course of participation. Brokerage commissions (fixed and variable) have to be paid every time a transaction takes place. No transaction, whether it is re-balancing of a portfolio or simply exiting the stock market, is costless. Transaction costs directly affect the frequency of portfolio re-balancing leading to a structural dependence in the share of stocks in the financial portfolio. This has particularly serious implications for the optimal portfolio of a small saver who happens to be in the stock market. In the presence of transaction costs, it may not be worthwhile for him to re-balance his portfolio for a long period of time over the life cycle. In principle, it is possible to identify fixed transaction costs within the estimation framework used in this paper. Unfortunately, this additional complication makes the solution of the model even more time consuming.

Another obvious direction in which to extend this work is the simultaneous modelling of stock market participation and other aspects of intertemporal allocation behavior: portfolio shares, consumption and wealth levels. As Browning and Crossley (2001) emphasize, a great virtue of the life-cycle framework is the coherence it brings to thinking about different aspects of intertemporal allocation behavior, and the discipline it demands: it should be possible to reconcile different aspects intertemporal allocation behavior with the same set of parameter values. Because my model is a essentially a buffer stock savings model (impatient agents facing labour income certainty) it cannot match the wealth distribution. The same is true of the Gourinchas and Parker's (2002) structural model of life-cycle consumption profiles. However, because my estimates of the rate of time preference and inter-temporal substitution elasticity are not too dissimilar to Gourinchas and Parker's (2002), it seems that it might be possible to reconcile both consumption profiles and stock market participation profiles with a model of the type I have presented. With respect to portfolio shares, preliminary work suggests that it will only be possible to match the age profile of portfolios shares with a model of this type if another risk (for example a depression or a calamity risk) is added to the model. These extensions are left for future research.

References

ALAN, S. and M. BROWNING (2003). "Estimating Intertemporal Allocation Parameters using Simulated Residual Estimation" Unpublished manuscript.

ALVAREZ, J., M. BROWNING and M. EJRNAES (2003). "Modelling Income Processes with Lots of Heterogeneity", mimeo, University of Copenhagen.

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.

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, 263-75.

BERTAUT, C. C. and M. HALIASSOS (1997). "Precautionary Portfolio Behavior from a Life Cycle Perspective" *Journal of Economic Dynamics and Control* 21, 1511-42.

BROWNING, M., and T. F. CROSSLEY (2001). "The Life-cycle Model of Consumption and Savings", *Journal of Economic Perspectives*, 15(3), 3-22.

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

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

COCCO, F. J., F. J. GOMES and P. J. MAENHOUT (1997). "Consumption and Portfolio Choice Over the Life Cycle" Working Paper, Harvard University.

DAVIS, S., and P. WILLEN (2000). "Occupation-Level Income Shocks and Asset Returns: Their Covariance and Implications for Portfolio Choice" NBER Working Paper No. 7905.

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

FARIA, A. L. (2000). "Stock Market Participation and Wealth Distribution" The University of Chicago.

GALLANT, A. R. and G. TAUCHEN (1996) "Which Moments to Match?", *Econometric Theory*, 12:657-681.

GAKIDIS, H. (1998). "Stocks for the Old? Earnings Uncertainty and Life Cycle Portfolio Choice" Ph.D. Dissertation, MIT.

GOURIEROUX, C., A. MONFORT and E. RENAULT (1993) "Indirect Inference", *Journal of*

Applied Econometrics, 8:S85-S118.

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

GUISO, L., M. HALIASSOS and T. JAPELLI (2002). "Household Stock Holding in Europe: Where Do We Stand and Where do We Go?", University of Cyprus Discussion paper.

HALIASSOS, M., A. MICHAELIDES (2002). "Calibration and Computation of Household Portfolio Models", *Household Portfolios*, Cambridge,MA: MIT Press, 2002, pp. 55-102.

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

HALL, G. and J. RUST (1999). "Econometric Methods for Endogenously Sampled Time Series: The Case of Commodity Price Speculation in the Steel Market", Unpublished Manuscript.

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

HECKMAN, J. J. (1981). "The Incidental Parameters Problem and the Problem of Initial Condition in Estimating a Discrete Time-Discrete Data Stochastic Process" *Structural Analysis of Discrete Data with Econometric Applications.* MIT Press. Cambridge.

LENTZ, R. (2003) "Optimal Unemployment Insurance in an Estimated Job Search Model with Savings" Boston University, Unpublished manuscript.

PAIELLA, M. (2001). "Limited Financial Market Participation: A Transaction Cost- based Explanation." The Institute for Fiscal Studies, UK.

POLKOVNICHENKO, V. (2001). "Limited Stock market Participation and the Equity Premium" Working paper, University of Minnesota.

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

A Appendix: Solution and Simulation Methods

The standard life cycle model for portfolio choice described in Section 2 is solved via backward induction by imposing an exogenous illiquid wealth accumulation function at the final period T. Simply, in the last period of working life the policy rule for normalized consumption is

$$c_T = \lambda_0 + \lambda_1 x_T \tag{28}$$

In order to solve for the policy rules at $T-1$ I discretize the state variable x (cash on hand to permanent income ratio) by defining an exogenous grid $\{x_j\}_{j=1}^{J}$ $j = 1...50$. The control space is also discretized such that normalized stock holding is $\{s_i\}_{i=1}^{I}$ $j = 1...100$ and normalized bond holding is $\{b_j\}_{i=1}^{I}$ $i = 1...100$. For the estimation, the borrowing constraint is assumed to be explicit so I set the lowerbound for cash to 0.1 and upperbound to 10. Since borrowing is not allowed, the possible range for cash on hand is always positive therefore it is not necessary to adjust the grid as the solution goes back in time[21]. For the illustration of the model in Section 2, a positive zero income probability is assumed and the control space is not discretized.

The algorithm first finds the investment on risky and riskless assets that maximize the value function for each value in the grid of x. Then, another optimization is performed where the generic consumer has only risk free asset to invest. Values of both optimizations are compared and the rule that results in 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 earlier value functions, nonconvexities due to taking outer envelope of two functions do not pose any numerical difficulty.

For illustration of the model, first, I generate $10,000$ income shocks for 60 years using the income process described in Section 2. 60 years of returns are generated in the similar fashion. The probability of zero income shocks is generated using a uniform random number generator. After generating all the necessary shocks, I simulate life cycle paths of consumption, stock and bond holding for $10,000$ agents and take cross-section averages.

For the estimations, the model is solved given the calibrated income and return processes for 37 years (ages 24 to 60). Using the resulting policy functions and the realized returns simulated

[21] In genearal, when borrowing is allowed, cash on hand in any given period (except for the last period) can be negative. It is then crucial to adjust the grid since the possible ranges for cash on hand are different at different stages of life. For instance, if one wants to impose a borrowing constraint such that all debt must be paid before death, then possible lower bound for cash on hand at time $T-1$ is minus the minimum possible income realization divided by gross risk-free rate.

data for 1294 households are generated. Age composition and the panel structure of the actual data are exactly replicated in the simulated data. Due to the extreme complexity of the solution of the underlying model I set the number of simulations to 1[22].

B Appendix: Asymptotic Distribution of the SMD Estimator

Following Hall and Rust (1999) and Alvarez, Browning and Ejrnaes (2003), consistency is established by using

$$p\lim \alpha^S(\theta) = \alpha^\infty(\theta) \tag{29}$$

and

$$p\lim \alpha^D(Y_h, X_h) = \alpha^0(\theta) \tag{30}$$

If the model is well-specified $\alpha^S(\theta)$ is converges in probability to $\alpha^D(Y_h, X_h)$, α^D hereafter. Then,

$$p\lim \theta_{SMD} = \theta_0 \tag{31}$$

Assuming the weighting matrix A converges to a non-stochastic matrix

$$\begin{aligned} \alpha^S(\theta) - \alpha^D &= \alpha^S(\theta) - \alpha^D + \alpha^\infty(\theta_0) - \alpha^\infty(\theta_0) + \alpha^0(\theta_0) - \alpha^0(\theta_0) && (32) \\ &= \left(\alpha^S(\theta) - \alpha^\infty(\theta_0)\right) + \left(\alpha^0(\theta_0) - \alpha^D\right) + \left(\alpha^\infty(\theta_0) - \alpha^0(\theta_0)\right) && (33) \end{aligned}$$

If the model is well specified the last term disappears. Applying the Central Limit Theorem and evaluating $\theta = \theta_0$

$$\sqrt{H}\left(\alpha^S(\theta_0) - \alpha^D\right) = \sqrt{H}\left(\alpha^S(\theta_0) - \alpha^\infty(\theta_0)\right) + \sqrt{H}(\alpha^0(\theta_0) - \alpha^D) \tag{34}$$

And

$$\sqrt{H}\left(\alpha^S(\theta_0) - \alpha^D\right) \to^d N\left(0, \frac{S+1}{S} V_0\right) \tag{35}$$

Note that if $S = 1$, variance is twice a large as the variance obtained from the analytical solution.

To find the variance of θ_{SMD} we take Taylor series expansion of the first order condition

$$\left[\alpha^S(\theta) - \alpha^D\right]' A \nabla \alpha^S(\theta) = 0 \tag{36}$$

[22] Efficiency gain of and extra simulation is not very large considering the CPU time of each iteration (approximately 20 hours).

Expanding $\alpha^S(\theta)$ around θ_0 we have

$$\alpha^S(\theta) = \alpha^S(\theta_0) + \nabla\alpha^S(\theta_0)(\theta - \theta_0) \tag{37}$$

Substituting this into the first order condition and solving for $(\theta - \theta_0)$ we have

$$(\theta - \theta_0) = -\left[\nabla\alpha^S(\theta)' A \nabla\alpha^S(\theta)\right]^{-1} \nabla\alpha^S(\theta)' A \left[\alpha^S(\theta_0) - \alpha^D\right] \tag{38}$$

Using this result and $\sqrt{H}\left(\alpha^S(\theta_0) - \alpha^D\right) \to^d N\left(0, \frac{S+1}{S} V_0\right)$ we have

$$\sqrt{H}\left(\theta_{SMD} - \theta_0\right) \to^d N\left(0, \frac{S+1}{S} G_1^{-1} G_2 G_1^{-1}\right) \tag{39}$$

where

$$G_1 = \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right)' A_\infty \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right) \tag{40}$$

$$G_2 = \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right)' A_\infty V_0 A_\infty \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right) \tag{41}$$

So, the optimal weight matrix $A = V_0^{-1}$. Then, the asymptotic distribution of the SMD estimator which uses a consistent estimator of V_0^{-1} as weight matrix is

$$\sqrt{H}\left(\theta_{SMD} - \theta_0\right) \to^d N\left(0, \frac{S+1}{S} G^{-1}\right) \tag{42}$$

where

$$G = \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right)' V_0^{-1} \left(p\lim\left[\nabla\alpha^S(\theta_0)\right]\right) \tag{43}$$

Parameter	Value
CRRA (γ)	2
Discount Rate (δ)	0.1
risk-free rate (r)	0.03
mean excess return on risky asset (μ)	0.06
std of risky asset (σ_ε)	0.20
std of transitory income shocks (σ_u)	0.14
std of permanent income shocks (σ_n)	0.0
Fixed entry cost ratio (E^c)	0 and 0.02
probability of zero income	0.01
Retirement function parameters	$\lambda_0 = 0,\ \lambda_1 = 1$

Table 1: Parameters for Simulations

Year		Full Sample	Estimation Sample
1984	# of observations	7241	1294
	pct. of hh participating	14.8	27.5
	median participation	0	0
1989	# of observations	5921	1294
	pct. of hh participating	17.8	36.4
	median participation	0	0
	correlation btw 84 and 89 participation	.41	.45

Table 2: Summary Statistics of portfolio holdings. Values are in 82-84 prices (Dollars). Financial wealth is stocks+bonds+cash

.

	Estimates
g	.024
	(.001)
σ_z^2	.005
	(.001)
σ_ε^2	.041
	(.011)

Table 3: Variance Decomposition and Growth of Income. Standard erros in parantheses.

Parameter	**Probit**	**Linear**	**Heterogeneity Corrected Probit**
constant	-3.276	$-.708$	-8.48
	(1.03)	(.306)	(2.14)
Age	.179	.058	.442
	(.052)	(.015)	(.119)
Age^2	-0.0024	$-.0007$	$-.006$
	(.0006)	(.0002)	(.0016)
lagged participation	1.272	.465	1.897
	(.086)	(.027)	(.449)
ρ			$-.410$
			(.331)

Table 4: Auxiliary Estimates from the PSID. Standard errors in parantheses. Right hand size variables also include 8 cohort dummies (4-year intervals) and family size.

Structural Parameter	**Auxiliary Model**	
	Probit	**Linear**
Coefficient of RRA (γ)	1.625	1.610
	(.1748)	(.1911)
Discount Rate (δ)	.0874	.0861
	(.0069)	(.0071)
Entry Cost (E^c - % of Permanent Iincome)	.0215	.0210
	(.0047)	(.0044)
Simulated Median Participation	0	0
Overidentifying Restriction (χ_1)	22.11	17.86
Simulated Auxiliary Parameter	**Probit**	**Linear**
constant	-3.17	.446
	(.987)	(.237)
age	.142	.042
	(.048)	(.011)
age^2	$-.0018$	.0005
	(.0005)	(.0001)
lagged participation	1.246	.460
	(.055)	(.018)

Table 5: Structural Estimation Results. Standard errors in parantheses.

	Parameter			Mean Predicted Participation			Coef. on lagged Participation
	γ	δ	E^c	*age* 30	*age* 45	*age* 59	
Actual Data				.511	.754	.302	1.272
Estimates	1.625	.0874	.0215	.527	.634	.287	1.246
Counterfactual 1	1.645			.534	.642	.282	1.235
Counterfactual 2		.0880		.523	.631	.288	1.267
Counterfactual 3			.0225	.520	.633	.287	1.281

Note: Actual data values are predicted mean participation obtained after the probit regression of 89 participation on a constant, age, age squared, 84 participation, cohort dummies and family size.

Table 6: Sensitivity of Age Profiles and Structural Dependence (Probit Model)

.

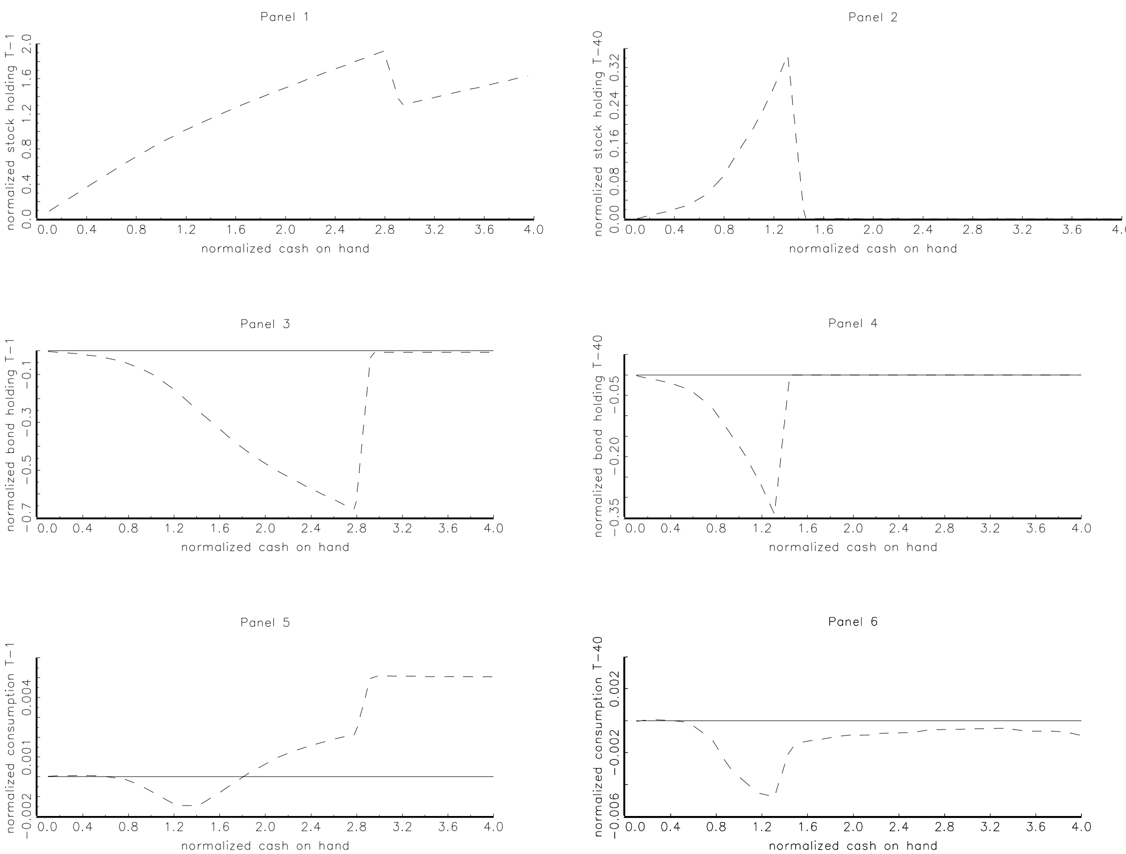


Figure 1: Policy Function Differences

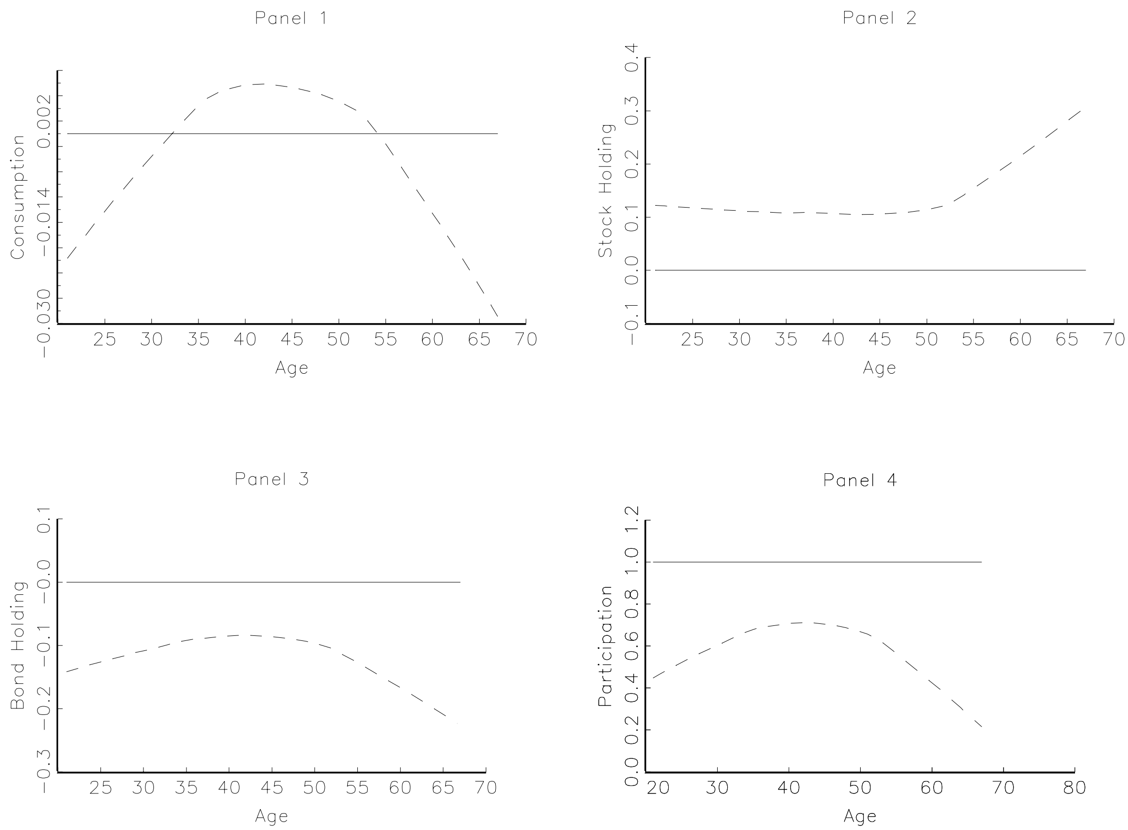
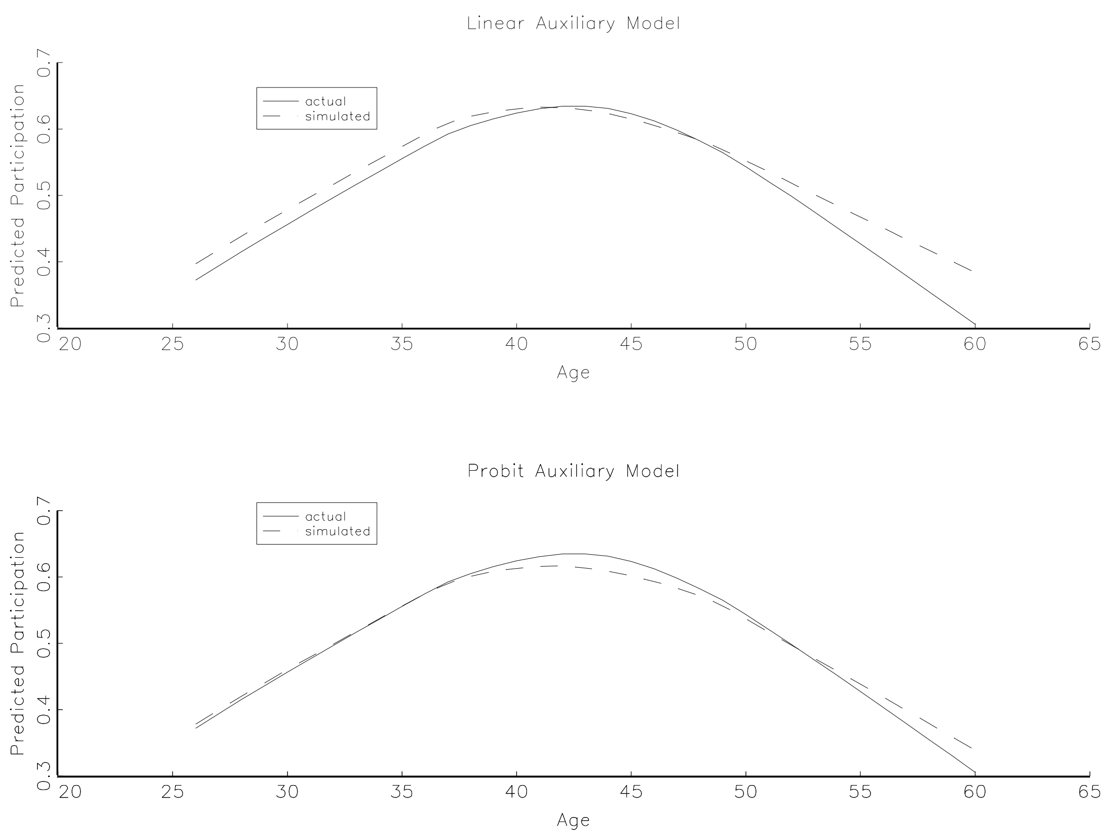


Figure 2: Life cycle path differences



Linear Auxiliary Model
actual
simulated
Predicted Participation
0.3
0.4
0.5
0.6
0.7
20
25
30
35
40
45
50
55
60
65
Age
Probit Auxiliary Model
actual
simulated
Predicted Participation
0.3
0.4
0.5
0.6
0.7
20
25
30
35
40
45
50
55
60
65
Age

Figure 3: