CSEF

Centre for Studies in Economics and Finance

WORKING PAPER NO. 125

The Distribution of Gains from Access to Stocks

Yannis Bilias and Michael Haliassos

September 2004

University of Naples Federico II

University of Salerno

Bocconi University, Milan

CSEF - Centre for Studies in Economics and Finance – UNIVERSITY OF SALERNO
84084 FISCIANO (SA) - ITALY
Tel. +39 089 96 3167/3168 - Fax +39 089 96 3167 – e-mail: csef@unisa.it

WORKING PAPER NO. 125

The Distribution of Gains from Access to Stocks*

Yannis Bilias♣ and Michael Haliassos♦

Abstract

Recent market developments raise doubts regarding further spread of household stock market participation. We study, computationally and econometrically, net gains from access to stocks, and estimate the potentially changing role of their determinants across the distribution of such gains for US households. We highlight conflicting influences on net gains using a computational portfolio model, and use empirical estimates to derive differences in characteristics of potential entrants relative to marginal investors by the end of the dramatic recent expansion in the stockholder base. Findings suggest that downturns can have significant effects around the participation margin, through their influence on incomes, wealth, and employment. The role of education is found more limited than typically estimated, and confined to the low end of the gains distribution. Estimated characteristics of potential entrants relative to marginal stockholders suggest that further growth in participation poses considerable challenges, in view of more limited finances, younger age, more limited education and financial alertness, and above all significantly less self-declared willingness to assume financial risk by potential stockholders compared to marginal investors. The hurdle to financial practitioners interested in expanding the stockolder base is not estimated to be small.

JEL classification : G11; E21
Keywords : Portfolio choice, stock market participation, binary quantile regression.

* We are grateful to Carol Bertaut, Chris Carroll, Luigi Guiso, Tullio Jappelli, Christos Koulovatianos, Michael Reiter, Thomas Steinberger, and participants of the London RTN conference on the Economics of Aging and in seminar presentations at Bocconi, Ente Luigi Einaudi, University of Frankfurt, CSEF, and University of Rome Tor Vergata for very helpful discussions, comments or suggestions. Special thanks go to Chris Carroll, for his insights on computational algorithms; and to Gregory Kordas for sharing his binary quantile regression algorithm and for very helpful discussions. Dimitris Mavridis and Nayia Pospori provided excellent research assistance. We thank the Leventis Foundation and HERMES for funding. Bilias thanks CSEF for funding and for providing an excellent research environment. This work has been supported in part by the European Community's Human Potential Programme under contract HPRN-CT-2002-00235 [AGE].

♣ University of Cyprus

♦ University of Cyprus and HERMES

Contents

1 Introduction

Recent empirical studies have documented substantial increases in the proportion of households holding stocks, either directly or indirectly through mutual funds and retirement accounts.[1] This spread of 'equity culture' has been linked to investor euphoria in the 1990s and to supply-side developments such as privatization of public utilities, growth of the mutual fund industry with concomitant reductions in participation costs for investors,[2] as well as the demographic transition and consequent policies to promote individual retirement accounts. In the current decade, when about one in two US households already holds stocks and euphoria has been tempered by recession and stock market downturns, continued growth of the stockholder base cannot be taken for granted. Concerns are often voiced for a possible exodus by marginal stockholders who feel the pressure of the economic downturn. At the same time, there are continued efforts to attract potential stockholders on the part of firms, brokers, and managers of mutual funds and other managed accounts. It thus becomes crucial to understand which factors gain or lose prominence in household stock market participation decisions and how the demographic composition of the pool of most likely potential entrants changes as stock market participation spreads. For example, products and strategies designed to attract highly educated, affluent households willing to undertake financial risk are likely to be much less effective when applied to households with fewer resources, limited ability to process financial information, and considerable lack of willingness to undertake financial risk.

Stock market participation is typically regarded as based on expected gains net of any participation costs, and it has received considerable attention in recent literature.[3] Existing empirical studies employ probit or logit estimation to show that richer households are not scaled-up versions of poorer households (Carroll, 2001) and that participation is influenced by education, risk aversion, and employment status, while the role of age, if any, is less clear and less consistent across countries.[4] Though useful, these standard techniques cannot be used to uncover the potentially changing role of household characteristics as participation spreads to lower percentiles of the net gains distribution. They instead estimate the role of each factor on *mean* net gains and assume that it is the same for all households regardless of their position in the distribution of net gains. This paper applies new methods, of binary quantile regression, specifically designed to estimate coefficients at different points of the distribution of the latent variable conditional on its covariates. Our econometric findings in this paper, using pooled US data from the 1995 and 1998 Surveys of Consumer Finances, suggest that there is considerable variation in the estimated influence of each factor depending on the position of the household in the distribution of net gains from access. When we use our estimates to rank households by probability of participation based on their demographics, we find that likely potential entrants differ in important respects from current marginal stockholders, most notably in their much more limited declared willingness to undertake financial risk and to look around for the best investment opportunities, and in their more modest financial resources. Such factors are likely to make it much more difficult for the financial sector to expand the stockholding base further, e.g. by designing financial products likely to appeal to potential stockholders, than it has in the past. As an indication of how much more difficult the task will be, we estimate increases in financial assets that would be necessary to offset the deterioration in remaining characteristics of potential stockholders and to produce probabilities of participation analogous to those of current marginal stockholders. We find that such increases are far from negligible.

As a prelude to the empirical analysis, the paper employs computational methods to shed light on the often conflicting roles of important determinants of gains from access to stocks net of any costs. Existing literature has established that fixed costs are important for participation decisions.[5]

In lieu of direct observation of actual or perceived fixed costs, the literature estimates threshold cost levels sufficient to deter participation by optimizing households. Empirical estimates of such costs (Vissing-Jorgensen, 2002; Paiella, 2001), as well as estimates from infinite-horizon computational models (Haliassos and Michaelides, 2003) suggest that reasonably small costs could keep most households out of the stock market. Our analysis incorporates one-time fixed costs of entry and per-period access costs.

We solve a modern calibrated dynamic model of household portfolio choice with non-diversifiable income risk, a finite horizon, uncertain length of life, and a retirement period. Although the model fits reasonably well mean stockholdings of college graduates during working life, its purpose is not to match data on portfolios of different demographic groups, but to provide guidance in interpreting coefficient estimates on net gains. To do so, we vary one relevant factor at a time, while fixing all other factors to the same level across all demographic groups, in a manner analogous to the exercise of interpreting regression estimates. Thus, we gain insights into the direction of the effect of, say, varying educational attainment , but we do not attempt to estimate net gains for different education groups. Such welfare analysis would require an elaborate data-matching exercise beyond the scope of this paper.[6] We focus on a number of factors that have received attention in studies of portfolio choice. Education effects are introduced through the household income process and through variation in the level of access costs relative to the permanent component of income. We also examine the role of financial resources, risk aversion, age, and of two important features of the stock market environment, namely the size of the perceived equity premium and stock market volatility.

In Section 2, we describe the computational model and the concept of gains, as well as calibration of the model. Section 2.6 uses the calibrated model to discuss the role of education, age, financial resources, and degree of risk aversion for the size of net gains from stock market access. In Section 3, we describe the quantile regression approach that we follow in our econometric estimation. Section 4 describes the data and recent trends in stock market participation in the United States. Section 5 reports our empirical findings on coefficient estimates, on the implied demographics of potential stockholders versus current marginal stockholders, and on implied costs of the difference in demographics across the two groups. Section 6 concludes.

2 Gains from Access to Stocks

2.1 Definition of Gains

We model finite-horizon households that choose savings and portfolio share of the risky asset to maximize lifetime expected utility, taking as known and given asset return processes, income process, and any borrowing constraints and entry or per period access costs they face. We compare two situations: one in which the household has access to the riskless asset but no access to stocks; and another in which the household has access to stocks, as well, and can choose whether and how much to hold, after payment of any access costs. Gains from access to stocks at a point in the life cycle are defined as the difference in value functions of the two problems at that point in time and for the values of the state variable(s) faced by the household. Note that this specification of gains from stock market access is consistent with households not holding stocks when they do not find it optimal to do so. Although net gains are unobservable (a 'latent variable' in econometric specifications), and no analytical expression can be derived for them in the presence of non-diversifiable income risk, their dependence on state variables and on household characteristics can be studied by use of modern computational methods.

2.2 Value Function for a Household with Access to Stocks

We first consider a household with finite horizon but uncertain lifetime, that maximizes expected intertemporal utility faced with a menu of a risky and a riskless asset. The value of the household's problem in the first period of life, $t = 0$, is given by

$$V_0^S(X_0, P_0) = MAX_{\{C_t,\alpha_t\}_{t=0}^{T-2}} E_0 \sum_{t=0}^{T-1} \beta^t \left(\prod_{j=1}^{t} \widehat{s}_j \right) U(C_t), \tag{1}$$

subject to

$$C_t + S_t + F_t \leq X_t \tag{2}$$

$$X_{t+1} = S_t \left[R_f + \alpha_t \left(\widetilde{R}_{t+1} - R_f \right) \right] + Y_{t+1} \tag{3}$$

$$C_t \geq 0 \tag{4}$$

$$0 \leq_t \alpha \leq 1 \tag{5}$$

All variables are in real terms. S_t is the real amount of total saving between the beginning of period t and of period $t+1$, α_t is the portfolio share of the risky asset held between t and $t+1$, E_t denotes the mathematical expectation operator based on information available up to the beginning of period t, β is the discount factor that satisfies $0 < \beta < 1$, $\widehat{s}_j$ is the probability that the household is alive in period j, conditional on being alive in period $j-1$. $U(C_t)$ is the felicity derived from consumption in period t, X_t is cash on hand at the beginning of period t,defined as the sum of net wealth and labor income, $\widetilde{R}_{t+1}$ is the risky gross return on stocks held between period t and $t+1$, R_f is the gross riskless rate assumed time-invariant, Y_t is income received at the beginning of t, and P_t refers to the permanent component of income, defined below. The term "income" is used loosely to encompass all after-tax income from transfers and wages, including pension income. $F_t \geq 0$ is the exogenous, fixed per period real cost of access to the stock market. The nature of stock market access costs is discussed in Section 2.4. Note that V_0^S is a function of access costs, but this dependence is omitted for brevity. The size of cash on hand at the beginning of life, X_0, is given.

The budget constraint (2) will hold with equality in equilibrium, given non-satiation. We employ a felicity function of the constant relative risk aversion (CRRA) form

$$U(C_t) = \frac{C_t^{1-\rho} - 1}{1-\rho}, \quad \rho \neq 1, \quad \rho > 0 \tag{6}$$

Constraint (4) is never binding under CRRA utility, since $\lim_{C_t \to 0} U'(C_t) = \infty$.

Denote by $\widehat{\beta}$ the product of the discount factor with the conditional probability that the household will be alive next period, $\beta \widehat{s}_{t+1}$. First order conditions for control variable θ_t are given by:

$$U'(C_t)\frac{\partial C_t}{\partial \theta_t} + \widehat{\beta} E_t U'(C_{t+1}) \frac{\partial C_{t+1}}{\partial \theta_t} = 0 \tag{7}$$

where θ is either S or α. Now,

$$C_{t+1} = \left[R_{t+1} + (\widetilde{R}_{t+1} - R_{t+1})\alpha_t \right] S_t + Y_{t+1} - S_{t+1} - F_{t+1} \tag{8}$$

Thus, the first order condition for saving, $\theta_t \equiv S_t$, is

$$U'(X_t - S_t - F_t) = \widehat{\beta}\ E_t \left\{ \left[R_{t+1} + (\widetilde{R}_{t+1} - R_{t+1})\alpha_t \right] U'(C_{t+1}) \right\} \tag{9}$$

and for the portfolio share of stocks, $\theta_t \equiv \alpha_t$, is

$$S_t \; E_t \left\{ \left(\widetilde{R}_{t+1} - R_{t+1} \right) U'(C_{t+1}) \right\} = 0. \tag{10}$$

In the presence of constraint (5), which precludes borrowing at the riskless or the risky rate, this formulation generates ranges of cash on hand in which it is optimal to hold no stocks. Intuitively, in such ranges, households may like to transfer consumption from the future to the present borrowing at the riskless rate. Being precluded from doing so, they may be willing to borrow even at the risky rate through short sales of stock. When short sales are not allowed, they end up at a corner with zero stockholding and no other saving.[7]

The value function in period t is $V_t^S(X_t, P_t)$, a function of household age, t, of total cash on hand in period t, X_t, and of the permanent component of income, P_t:

$$V_t^S(X_t, P_t) = \underset{S_t,\ \alpha_t}{MAX} U(C_t) + \widehat{\beta} \; E_t V_{t+1}^S(X_{t+1}, P_{t+1}). \tag{11}$$

Income of household i, Y_{it}, is assumed to entail non-diversifiable risk because of moral hazard and adverse selection considerations. Observed income follows

$$Y_{it} = P_{it} U_{it}, \tag{12}$$

where U_{it} is a transitory shock. During working life, the permanent component, P_{it}, follows

$$P_{it} = G_t P_{it-1} N_{it} \tag{13}$$

and is thus subject to shocks, N_{it}. During retirement it follows the deterministic process

$$P_{it} = G_t P_{it-1} \tag{14}$$

so that retirement income is only subject to transitory shocks. We assume that $\ln U_{it}$ and $\ln N_{it}$ are each independent and identically (normally) distributed with means $\{-.5 * \sigma_u^2 \;, \; -.5 * \sigma_n^2\}$, and variances σ_u^2 and σ_n^2, respectively. The lognormality of U_{it} and the assumption about the mean of its logarithm imply that

$$EU_{it} = \exp(-.5 * \sigma_u^2 + .5 * \sigma_u^2) = 1 \tag{15}$$

and similarly for EN_{it}. Thus, the permanent component of income is the value observed when transitory shocks are at their expected value. The growth factor, G_t, is assumed to be a function of household characteristics. Variances σ_u^2 and σ_n^2 and growth factors G_t are calibrated using empirical estimates for different education categories, distinguishing between working life and retirement (see below).

To reduce the number of state variables by one, we normalize asset holdings and cash on hand by the permanent component of labor income, P_{it}. Lower case letters denote normalized variables. To simplify notation, we drop the subscript i. Recalling that marginal utility is homogeneous of degree $(-\rho)$ under CRRA preferences, we can rewrite the first-order conditions (9) and (10) as

$$U'(x_t - s_t - f_t) = \widehat{\beta} \; E_t \left\{ (G_{t+1} N_{t+1})^{-\rho} \left[R_{t+1} + (\widetilde{R}_{t+1} - R_{t+1}) \alpha_t \right] U'(c_{t+1}(x_{t+1})) \right\} \tag{16}$$

and

$$s_t \; E_t \left\{ (G_{t+1} N_{t+1})^{-\rho} \left(\widetilde{R}_{t+1} - R_{t+1} \right) U'(c_{t+1}(x_{t+1})) \right\} = 0 \tag{17}$$

where $c_{t+1}(x_{t+1})$ denotes the policy function for normalized consumption in period $t+1$. The normalized state variable x evolves according to

$$x_{t+1} = \left[R_{t+1} + (\widetilde{R}_{t+1} - R_{t+1})\alpha_t\right] \ s_t \ \left(G_{t+1}N_{t+1}\right)^{-1} + U_{it+1} \tag{18}$$

Recalling that the value function is homogeneous of degree $(1-\rho)$ when felicity exhibits constant relative risk aversion, we can rewrite the Bellman equation for the problem with access to stocks as

$$V_t^S(x_t) = \underset{s_t(x_t),\ \alpha_t(x_t)}{MAX} U(c_t(x_t)) + \widehat{\beta}\ E_t\{G_{t+1}N_{t+1}\}^{1-\rho}\ V_t^S(x_{t+1}) \tag{19}$$

2.3 Value Function for a Household without Access to Stocks

If the household has no access to stocks, it chooses only the real amount of bonds to hold. Denoting by B_t saving in the (single) riskless asset, constraints can be derived by setting $S \equiv B$ and $\alpha \equiv 0$ above, and by replacing borrowing constraint (5) with $B_t \geq 0$.Off corners, analytical first order conditions for saving are, upon normalization

$$U'(x_t - b_t) = \widehat{\beta}\ R_f\ E_t\left\{(G_{t+1}N_{t+1})^{-\rho}U'(c_{t+1}(x_{t+1}))\right\}. \tag{20}$$

Normalizing all variables by P_t, we write the Bellman equation for the value function of the problem without access to stocks as

$$V_t^B(x_t) = \underset{b_t(x_t)}{MAX} U(c_t) + \widehat{\beta}\ E_t\left\{(G_{t+1}N_{t+1})^{1-\rho}\ \ V_t^B(x_{t+1})\right\} \tag{21}$$

As analytical solutions are not available, we solve numerically the constrained portfolio and saving problems. We compute solutions in Matlab, using portfolio algorithms recently developed by Haliassos and Mavridis to exploit some clever computational shortcuts proposed by Carroll (2002) (see Appendix). We generate value functions $V_t^S(x_t)$ and $V_t^B(x_t)$ for each period t.

2.4 Gains and Costs of Stock Market Access

Subtracting the computed value function of a household without access, $V_t^B(x_t)$, from that of a household with access, $V_t^S(x_t)$, we derive gains from stock market access

$$DV_t(x_t) = V_t^S(x_t) - V_t^B(x_t) \tag{22}$$

as a function of normalized cash on hand and for given household characteristics (e.g., education level, age, risk aversion, etc.). By varying one household characteristic at a time, keeping others constant, we derive in Section 2.6 model predictions for the nature of effects of this characteristic on the systematic part of gains from access net of any costs incorporated in the computation of value functions.

In principle, one can distinguish three potentially relevant types of such costs: a fixed entry cost, a per period fixed cost, and a transactions cost that depends on the volume of stock transactions. Vissing-Jorgensen (2002) focused on stock market participation and found empirical support for the first two types of costs, but little role for the third type of costs in explaining observed participation rates.

Our calibrated model incorporates fixed entry costs and per period access costs. Fixed entry costs can be thought of as the costs that a previously uninformed household pays in order to familiarize itself with stocks and stock-related financial assets (e.g., mutual funds, retirement accounts, etc.), and with the process of using them to smooth consumption intertemporally. Once these costs are paid, the household perceives stock-related assets as part of its asset menu over its remaining horizon, and it considers in each period whether and how much to invest in stocks. Processing information necessary to figure out optimal holdings in each future period (or assigning professionals to do this) entails a per period normalized access cost f_s for $s \geq t$, that we incorporate in computation of V_t^S.[8] While this approach allows for entry and exit from *holding* stocks, it does not cover the possibility that households who have gained access to the stock market subsequently decide to ignore the presence of stockholding opportunities and resume solving the single-asset saving model. Allowing for such voluntary restriction of the asset menu complicates the optimization problem considerably more than introduction of per period participation costs that we have pursued here, and has never been done to the best of our knowledge.

Gains in the presence of any fixed (normalized) entry costs, $k(x_t)$, over and above the current per period access cost are given by

$$D\widetilde{V}_t(x_t, k(x_t)) = V_t^S(x_t - k(x_t)) - V_t^B(x_t) \tag{23}$$

Their computation is complicated by the fact that entry costs, as these are *perceived* by households making the participation choice, are essentially unobservable and could vary significantly among households without any previous stock market experience. Rather than imposing some arbitrary $k(x_t)$ and plotting (23) for that choice, we prefer to plot gains in the absence of fixed entry costs using (22), and then to solve for the $\widehat{k}(x_t)$ that makes (23) equal to zero. This $\widehat{k}(x_t)$ represents the horizontal difference between the value functions with and without stock market access, and it has two useful interpretations. One is as the monetary sum that should be taken away from a household with access to stocks in order to eliminate its gains from such access. In this sense, it is a dollar measure of how much stock market access is worth to a household with a particular set of characteristics. Clearly, a household would not be willing to pay a fixed entry cost in excess of this amount in order to gain access to stocks. Hence, its second interpretation is as the (minimum) fixed entry cost (beyond the access cost f_t paid at the time of entry) that would deter a household of given characteristics from obtaining access to the stock market.

2.5 Calibration

2.5.1 Income Processes

Income is defined as after-tax, non-asset income. It includes not only labor income, but also government transfers, bequests, and lump-sum windfalls. We distinguish between three education categories, based on the educational attainment of the household head: less than high-school education, high-school graduates, and college graduates (or more). Income processes consist of (deterministic) age-income profiles and of stochastic shocks. Income during working life is assumed to follow equation (12), and its permanent component is modeled as

$$P_{it} = G(t, Z_{it})\ P_{it-1} N_{it}, \tag{24}$$

where $G(t, Z_{it})$ denotes the growth factor of the permanent component of income derived from regressions of the logarithm of non-asset, after-tax income on age variables, separately on the number of head and spouse, number of dependent adults, and number of children, on cohort dummies (based on year of birth and split into five-year cohorts), and on the unemployment rate in the household's state of residence (as a proxy for time effects). The permanent component of income during retirement is assumed to take the form

$$P_{it} = G^R(t, Z_{it})\ P_{it-1}, \tag{25}$$

i.e. retirement income is assumed to be subject only to transitory shocks. Regressions for log income during working life allow for a third-order polynomial in age, while those for retirement assume a linear age term. Coefficient estimates are based on data from PSID 1983-1990 and are taken from Laibson et al. (2000, Tables 3 and 4). The retirement age for high-school dropouts is set to 61, for high-school graduates to 63, and for college graduates to 65, based on mean ages observed in the data. Estimated age income profiles are hump-shaped during working age for all education categories, but with different peaks prior to retirement.

Education groups also face differences in the process followed by income shocks. We employ the estimates of Carroll and Samwick (1997) regarding variances of transitory and permanent shocks, (σ_u^2, σ_n^2), for different education categories during working life.[9] For high-school dropouts, we use the Carroll-Samwick estimates for those who had completed between 9 and 12 grades: $(0.0658, 0.0214)$; for high-school graduates, we use $(0.0431, 0.0277)$; and for college graduates $(0.0385, 0.0146)$. These and other estimates in the literature suggest a decreasing pattern of risk associated with temporary earnings shocks as education increases. There is a less clear pattern with respect to the variance of permanent income shocks.

We follow Laibson et al. (2000) in modeling (logarithms of) shocks to retirement income as the sum of a household fixed effect and of a purely transitory shock:

$$u_{it}^R = \vartheta_i + \nu_{it}^R \tag{26}$$

They estimate the variance of ν^R using residuals from the income regressions during retirement years. Estimates of $\sigma_{\nu^R}^2$ for high-school dropouts, high-school graduates, and college graduates are, respectively, 0.077, 0.051, 0.042. In simulations, household fixed effects are set to zero.

2.5.2 Parameters

Conditional probabilities of survival are calculated from the 1998 United States Life Tables (National Vital Statistics Report, 2001). We set the rate of time preference, $\delta \equiv \beta^{-1} - 1$, equal to 0.05. The expected rate of return on equity, μ_r, is set to 0.06 in the benchmark runs, and the constant real interest rate, r, to 0.01. Understating the historical equity premium is an often used shortcut to introducing unaccounted for proportional costs, but we also consider higher values. The assumed standard deviation of the equity premium is set at its historical value of 18 percent in the benchmark, and raised to twice this level in sensitivity analysis. The benchmark value for risk aversion is $\rho = 2$, but we also consider values up to $\rho = 8$.

Perceived access costs or participation costs are unobservable. Vissing Jorgensen (2002) has estimated that per period participation costs of about 260 dollars (in 2000 prices) can account for the behavior of 75 percent of non-participants in the stock market. Paiella's (2001) estimates were not far from those. We use a real amount of 250 dollars to calibrate the per period access cost, normalized by permanent income derived from the age income profile for each education category. Assuming the same real cost of access across education categories is a useful benchmark, consistent with the notion that less educated households face bigger access costs relative to permanent income. Still, it is conservative in the sense that low-education households may indeed face higher real costs than their more educated counterparts, leading to lower gains, both because they find it more costly to acquire and process such specialized financial information, and because they do not typically benefit from the lower fees available to bigger investors.

2.6 Determinants of Gains from Stock Market Access

In this section, we study model predictions regarding the influence of household-specific variables as well as of stock market factors, namely the equity premium and its volatility, on gains from stock market access. We vary one factor at a time, holding all others constant, in order to isolate effects, and we differentiate between working life and retirement. We find that even these fairly stylized models uncover a rich pattern of often conflicting and sometimes surprising effects.

2.6.1 Effects of Education

A standard result in existing empirical literature based on probit or logit is that, controlling for other characteristics, education increases the probability of participation. Usually, this is justified by reference to higher fixed entry and participation *costs* relative to permanent income for less educated households. Higher costs can be attributed to greater difficulty in processing financial information, reduced or costlier access to sources such as the internet, and more limited access to financial advisors because of more limited scale of investments.[10] Less attention is paid to the size of potential *gains* from access to stocks, often presumed to be lower for the less educated, controlling for other characteristics.

Figure 2 (top row) plots predicted net gains from access (22) against normalized cash on hand, for different education categories but for given other characteristics. The left column refers to working households aged 45, while the right to retirees aged 70. The four lower panels show policy functions for the portfolio share of stocks, real stock amounts, riskless asset holdings, and consumption, all against normalized cash on hand. For given cash on hand and other characteristics at age 45, increases in educational attainment reduce net gains from access to stocks over the remaining lifetime.

The source of this result lies in the saving motives generated by the income processes of the three categories.[11] Age-income profiles by education are shown in Figure 1, and variances of income shocks were presented in Section 2.5. Less educated households face higher variances of transitory and permanent income shocks, earlier retirement, and worse prospects for income growth over their working lives. Limited prospects for income growth create incentives for working high-school dropouts to save more for the future, while higher income risk increases their precautionary saving (see bottom row). Greater saving motives dominate lower (unconstrained) portfolio shares of stocks for lower education households (second row) and deliver higher net gains.[12] By age 70, the ratio of per period access costs to permanent income is much smaller for college graduates than for lower-education households, and this reverses education effects among retirees. However, when we set the per period participation cost to zero, gains continue to be inversely related to education even at age 70 (see below).[13]

In principle, the effect of income processes on gains could be reversed by higher entry and per period access costs for lower education households. Since such costs are unobservable, we examine the sensitivity of net gains to two different assumptions. Figure 3 (top row) plots normalized monetary equivalents of gains from stock market access against normalized cash on hand, if all education categories pay the same absolute per period cost of $250. This cost will, of course, be higher relative to permanent income for lower-education households. The gains of 45-year-old college graduates over their remaining lifetime, net of such per period access costs, are of the order of a quarter of their annual permanent income component or less. They are about double if no per period access costs are paid (Fig. 4). Any fixed entry costs, likely to be smaller for more educated households, should be subtracted from these figures. It is also possible to see the effects of differential absolute per period access costs, by combining the two Figures (e.g. no cost measures for college graduates with positive cost measures for high-school dropouts).[14]

Note that, as we vary education levels keeping all other characteristics fixed, we can examine if gains go up or down with education, but we should not necessarily interpret the levels of monetary measures we obtain for lower education categories as estimates of actual net gains in this category. The reason is that low-education households may differ from college graduates in their remaining characteristics, which are kept fixed for purposes of our analysis. In unreported simulations, we found that benchmark parameter settings approximate reasonably well stockholding by college graduates during working life, but they tend to overstate financial assets for other education categories.

All in all, results highlight two conflicting effects of education on net gains: a negative one from more limited need to save for the future; and a positive one resulting from lower access costs relative to permanent income associated with higher educational attainment.[15] Our results suggest that a positive education effect on net gains, controlling for other characteristics, requires higher entry/access costs for low-education households, not only relative to permanent income, but also in absolute terms. We will see below that quantile regression limits positive education effects to households at the bottom of the distribution of net gains.

2.6.2 Effects of Risk Aversion

Survey data include questions on willingness to undertake financial risk. Here we explore the link between demand for stocks and the unobservable degree of risk aversion, ρ, of the felicity function, $U(C_t)$. Figure 5 plots predicted net gains for college graduates under different ρ. The second row of Figures 3 and 4 presents monetary equivalents of gains, with and without access costs, respectively. We find that, at working age 45 (left column), higher values of ρ are associated with larger gains from stock market access. At retirement age 70 (right column), differences are eliminated in all but the lowest range of cash on hand, where the ranking is actually reversed. Implied differences in monetary equivalents are small (of the order of 5,000 dollars), again about half of what they would be without access costs.

Risk aversion affects gains through two conflicting channels. Unconstrained portfolio shares of stocks are lower at any normalized cash on hand (Fig. 5, second panel); but greater risk aversion also implies higher prudence (precautionary motive) and thus larger precautionary wealth accumulation (bottom row). The stockholding policy function (third row) shows that aversion to riskiness of stocks dominates at high levels of cash on hand and in most of the range with positive stockholding during retirement. However, there is a wide range of cash on hand, where it is optimal for the more risk averse to hold more stocks. Payment commitments in the form of per-period access costs are not sufficient to reverse this dominance of the precautionary effect. By the age of 70, future income risk is significantly reduced and stock demand drops with risk aversion among savers (panel 5). This produces different rankings of net gains during retirement compared to working life.[16] Thus, higher risk aversion, ρ, is not necessarily associated with lower net gains from stock market access nor with lower demand for stocks. A systematically positive empirical relationship between survey responses and probability of stockholding would seem to support their interpretation as reflecting risk aversion of the value function rather than ρ.

2.6.3 Age Effects

Age effects on net gains result mainly from two factors. First, as the household ages, fewer periods are left in which to benefit from stocks and pay the access cost. Second, since age effects are computed at given normalized cash on hand, gains at older ages incorporate revised expectations regarding future resources. Fig. 6 shows that predicted age effects are generally negative and non-linear over the life cycle, for given cash on hand and other characteristics.[17] Downward shifts of the net gains schedule are bigger after the middle of working life (e.g., between 45 and 55), and much smaller during retirement.[18]

Age effects are small at the beginning of working life because stockholding tends to be limited early on and there are still many future periods of life ahead. As working life progresses, fewer years of high stockholding lie ahead. Unconstrained portfolio shares (second row) are predicted to fall with age,[19] though not if borrowing constraints are binding,[20] while assets increase. The net age effect on stockholding (third row) is positive where borrowing constraints are binding, but negative otherwise. Moreover, a given level of cash on hand implies progressively worse future prospects, and thus lower stockholding. Age effects are more substantial in the latter part of working life, for both reasons. During retirement, precautionary and other saving motives weaken. Moreover, since retired households face little income uncertainty, their expectations of cash on hand tend to cluster around current levels, and future stockholding is expected to follow closely the drops of the stockholding schedule. These factors generate negative net gains from gaining access in retirement. In the econometric part below, we allow for a non-linear (quadratic) age effect on net gains.

2.6.4 Equity Premium and Volatility

Effects of the equity premium or its volatility on net gains from access can help us understand year effects (1995 versus 1998) in the econometric section. Comparing 1995 and 1998, perhaps the two most relevant features are the continuing increase in stockholding participation and the decline in access costs throughout the 1990s (see Section 4). While a decline in costs clearly increases net gains from access, the limited theoretical literature seems to imply negligible or negative but small effects of increased participation on the equity premium.[21] There seems to be more empirical support for the idea that volatility increased, at least over the longer haul.[22] However, general equilibrium literature is still inconclusive regarding effects of increased participation on stock market volatility[23].

Fig. 7 confirms that a reduction in the (perceived) equity premium lowers net gains from access. Results are shown for expected equity returns of 6 and 12 percent and for college graduates. Faced with a lower equity premium, the household tilts its portfolio away from stocks (second to fourth panels), but it also reduces assets held (bottom panel) in view of worse investment prospects. Effects are similar but much smaller during retirement.

Fig. 8 shows drops in net gains when the benchmark (perceived) standard deviation of the equity premium (18 percent) is doubled. While the consumption rule remains largely unaffected (bottom panel), increased volatility causes a portfolio switch away from stocks (second row). Households hold smaller stock amounts, except for low cash on hand where they either do not save or they save little in stocks. Effects are qualitatively similar across working and retirement periods.[24]

3 Econometric Approach

The conceptual link between a computational portfolio model and empirical studies of stock market participation, whether in the traditional probit/logit or in the binary quantile regression mode, is provided by gains from involvement in the stock market net of any associated costs. The computational model can be used to predict effects on net gains of varying different factors while keeping the rest fixed at some level; and the econometric model to compute estimates of such effects, controlling for other characteristics. Econometric estimates of effects refer either to mean net gains (probit/logit) or to any percentile of net gains (binary quantile regression). It is useful to see if effects are at least qualitatively consistent across these two net gains constructs, even if neither the computational nor the econometric model fits portfolio holdings perfectly. The computational model can shed light on what possibly lies behind effects of each factor, while the econometric model can additionally yield estimates of demographics of groups at different ranges of participation probabilities and of its implications.

The link between theoretical and empirical modeling is not seamless. Since analytical expressions for net gains cannot be derived, the latent variable specification in the econometric model cannot inherit a particular functional form from theory, although it certainly benefits from qualitative results on effects of various factors. Moreover, use of a portfolio model that is rich enough to allow for zero optimal stock holdings introduces a slight discrepancy between the notion of access in the theoretical model and survey data where current stock market participation rather than access is observed. Thus, we implicitly identify those in the sample who held no stocks with households perceiving negative net gains from access. This may create problems if there are households that paid the costs of figuring out whether they should invest in stocks but found it optimal not to do so. Although such cases are unobservable, we doubt that they are quantitatively important in our data. The reason is that we are focusing on a period of stock market euphoria and of spreading equity culture, in which households that paid the costs of figuring out optimal stockholding tended to include stocks in their portfolios.

We apply recent binary quantile regression techniques to study stock market participation probabilities and the distribution of net gains from such participation. As far as we know, this is the first application of such techniques to stock market participation. Binary quantile regression is a natural approach to the study of participation and has advantages over the more standard probit and logit techniques. For households with given observable characteristics, binary quantile regression makes it possible to describe the role of a given factor at any percentile of the (conditional) distribution of net gains from access, as opposed to merely at the mean of net gains. Secondly, the technique allows estimation of participation probabilities on the basis of household characteristics and of coefficient estimates most appropriate for households in the relevant position of the underlying distribution of net gains. By contrast, standard probit and logit estimation assigns the same coefficient estimates to all households and results in different probability rankings, as illustrated below. Finally, binary quantile regression makes it possible to quantify more accurately how much harder it is for the financial sector to draw into the stockholder pool the next group of potential stockholders, compared to most likely current marginal stockholders. We conduct all such excercises below.

Formally, existing analyses of stock market participation have relied on standard discrete choice models of the following general structure:

$$y^* = x'\beta + u \tag{27}$$

$$y = I\left(y^* \geq 0\right) \tag{28}$$

where y^* is a latent response variable (in this case, the utility gain from access to stocks net of entry or participation costs), y is a binary dependent variable observable by the researcher (participation or non-participation in the stock market), x is a vector of covariates (typically household demographic characteristics, financial variables, and reported attitudes), and u is specified to follow either a standard normal or a logistic distribution.

The usual maximum likelihood methods, probit and logit, allow estimation of the mean of the distribution of the latent variable y^* and of covariate effects on this mean alone. Thus, probit or logit estimation uncovers effects only on the average utility gain from access to stocks in the population, and it does not allow for differences of covariate effects among households at different points in the distribution, e.g. for marginal investors, for those most likely to be the next entrants, and those that stand to gain a lot or very little from stocks. Moreover, although parametric distributional assumptions on the error term u usually allow estimation of β by maximum likelihood methods, excessive use of such assumptions entails the risk of inconsistent estimation in case they are violated.

By contrast, quantile regression originally proposed by Koenker and Basset (1978, 1982) allows estimation, in principle, of all the percentiles of the distribution of the dependent variable, while its semiparametric nature greatly reduces the risk of inconsistent estimation. In particular, quantile regression can provide valid statistical inference in the presence of heterogeneity of unknown form. As a result, the quantile regression methodology offers a considerably enhanced picture of the effects of the conditioning variables x across the entire spectrum of net utility gains from access to stocks.

Formally, consider a particular percentile, $0 < \tau < 1$, of the distribution of the utility gains from access to stocks, y^*, and assume the following linear model for the conditional quantile function $Q_{y^*}(\tau|x)$:

$$Q_{y^*}(\tau|x) \equiv F_{y^*}^{-1}(\tau|x) = x'\beta(\tau) \tag{29}$$

The model for the τth percentile of the latent variable y^* can yield a model for the τth percentile of the binary observable variable y using the *equivariance* property of quantile functions to monotonic transformations: for any monotone function $h(\cdot)$, it holds that

$$Pr(Y < y|x) = Pr(h(Y) < h(y)|x). \tag{30}$$

Then it is easily shown that

$$Q_y(\tau|x) \equiv Q_{I(y^* \geq 0)}(\tau|x) = I\left(Q_{y^*}(\tau|x) \geq 0\right) \tag{31}$$

Thus

$$Q_y(\tau|x) = I\left(x'\beta(\tau) \geq 0\right) \tag{32}$$

is our proposed model for the τth percentile of the observable binary variable of stock market participation y.

Following the formulation of Koenker and Basset (1978), the Binary Quantile Regression (BQR) program is to find the estimate $b^*(\tau)$ for each chosen $0 < \tau < 1$ that solves

$$\min_{\{b:||b||=1\}} n^{-1} \sum_{i=1}^{n} \rho_\tau \left[y_i - I\left(x_i'b \geq 0\right)\right] \tag{33}$$

where $\rho_\tau(\nu) = [\tau - I(\nu < 0)]\,\nu$ is the *check function* of Koenker and Basset (1978). The estimator is normalized to have Euclidean norm equal to unity for purposes of identification. This problem and estimator were first studied by Manski (1975, 1985) under the name *maximum score estimator*. Manski proposed to find the estimate $b^*(\tau)$ that solves

$$\max_{\{b:||b||=1\}} n^{-1} \sum_{i=1}^{n} [y_i - (1-\tau)] I\left(x_i'b \geq 0\right). \tag{34}$$

Manski's estimator does not specify the form of relation between x and u; it only requires that the τth quantile of u conditional on x is zero. Based on this assumption, the model can be identified up to scale. A similar identification problem arises in the case of parametrically specified probit and logit. The fact that the error distribution in BQR is left unspecified forces us to place the identification restriction on the beta coefficients by normalizing their euclidean norm $\|b\| = 1$ (Manski, 1985).[25]

Manski (1985) has shown that his maximum score estimator of β is consistent under the identifying assumption $Q_u(\tau|x) = 0$. However, the estimator is not $n^{1/2}$-consistent and, asymptotic normality and subsequent statistical inference cannot be realized. Horowitz (1992) proposed a modification to the maximum score estimator by smoothing the objective function (called 'score function' in this literature). Horowitz (1992), utilizing ideas from nonparametric density and regression estimation, essentially replaces the indicator $I(x'b \geq 0)$ in the objective function by an integral of kernels. As a result, Horowitz's *smoothed maximum score estimator* solves a sufficiently smoothed objective function whose difference from Manski's objective function tends to 0 almost surely as $n \to \infty$.

Formally, the smoothed maximum score estimator $b^*(\tau)$ solves the program

$$\max_{\{b:||b||=1\}} n^{-1} \sum_{i=1}^{n} [y_i - (1-\tau)] K\left(x_i' b/\sigma_n\right), \tag{35}$$

where $\{\sigma_n\}$ is a sequence of positive real numbers (bandwidths) that converges to zero as $n \to \infty$. $K(\cdot)$ is a continuous bounded function that bears the properties of a cumulative distribution function; that is, *(i)* $|K(v)| < M$ for some finite M and all v in $(-\infty, \infty)$ and, *(ii)* $\lim_{v\to-\infty} K(v) = 0$, $\lim_{v\to\infty} K(v) = 1$.

The smoothed maximum score estimator was shown by Horowitz (1992) to be consistent and, after centering and suitable normalization, to be asymptotically normal. Its convergence rate is at least $n^{-2/5}$ and can become arbitrarily close to $n^{-1/2}$. Asymptotic normality allows statistical inference to be carried out in the usual way. However, the finite sample properties can be very different from the behavior claimed by asymptotic theory. In this paper we derive confidence intervals by use of bootstrap methodology. Experience with Monte Carlo experiments indicated that bootstrap can reduce substantially the difference between exact and nominal coverage levels. Equally important is the indication that bootstrap-based inference is not very sensitive to the selection of bandwidth parameter.[26]

Although the theoretical literature (Manski, 1975; Horowitz, 1992) usually concentrates on the median case, $\tau = 0.50$, extension to other quantiles is extremely useful when complete characterization of the distribution of the response variable is needed, as in our case, and essential for theoretical reasons as well (e.g. identification, efficiency, testing of various hypotheses). Kordas (2001) provides a useful treatment of these issues and calls this estimator *Smoothed Binary Quantile Regression* (SBQR).

Unlike probit and logit, binary quantile regression does not estimate probabilities directly. However, using the estimated coefficients for each quantile, we can characterize the probability of participation in the stock market for a given composition of covariates. Since the assumed model for gains y^* is $Q_{y^*}(\tau|x) = x'\beta(\tau)$, we can use the basic definition of a quantile to write

$$\Pr(y^* \leq x'\beta(\tau)) = \tau \tag{36}$$

In view of (36), the following statements hold, conditional on the value of quantile index $x'\beta(\tau)$:

$$\left.\begin{array}{l} \Pr(y = 1|x'\beta(\tau) > 0) > (1-\tau) \\ \Pr(y = 1|x'\beta(\tau) = 0) = (1-\tau) \\ \Pr(y = 1|x'\beta(\tau) < 0) < (1-\tau) \end{array}\right\} \tag{37}$$

Intuitively, (37) indicates a clear relationship between net gains from participation and probability of participation. Households of given characteristics x can have any level of net gains out of a distribution generated by unobserved heterogeneity among them. However, households characterized by a given x have higher probability of participation, $1-\tau$, if net gains are zero among x-households at a lower percentile, τ, of the net gains distribution generated by unobservables. Crudely speaking, this makes it likely that unobservables will push more x-households into positive net gains and into participation.[27]

We can use (37) to describe the characteristics of household groups ranked by their probability of participation, $1 - \tau$. As we consider higher probabilities of participation, we are dealing with observable characteristics that, combined with unobservables, tend to generate greater proportions of households with positive net gains and hence stock market participation. For instance, while the unconditional probability that any US household participates is about 50%, three quarters of those whose characteristics imply 75% probability of participation tend to exhibit positive net gains and participation. We define 'marginal stockholders' and 'potential entrants' as those whose conditional probability of participation correspondingly exceeds or falls short of the unconditional probability (50%) by up to 25 percentage points. We describe below the demographic and other characteristics of these two groups and of the remaining two quartiles that are the most and the least likely to participate, respectively.

This can be implemented as follows. Consider a particular household with characteristics x_0. In principle, we can compute $x_0'\beta(\tau)$ for all $0 < \tau < 1$. These represent different percentiles of the distribution of net gains generated by unobserved heterogeneity among households sharing characteristics x_0. If we find that for, say, $\tau = \tau_0$ it holds that $x_0'\beta(\tau_0) = 0$, then it follows from (37) that a household with observables x_0 has probability of participating in the stock market equal to $(1 - \tau_0)$. By repeating this for all households in the sample, we can rank them in terms of participation probabilities.

In practice, estimates are not computed for all τ, but only for a subset of percentiles to limit computational costs. We obtain estimates for 19 equally spaced percentiles, from $\tau = 0.05$ to $\tau = 0.95$, at increments of 0.05. For each x, we find the τ that makes $x'\beta(\tau)$ closest to 0 from below. Having thus assigned households to each of the τ for which we have coefficient estimates, we group them further into quartiles. We derive the demographic characteristics of these quartiles, using population weights. We then compare them to those derived by ranking households according to their probabilities of participation implied by probit or logit estimates of mean net gains.

Besides ranking households in terms of participation prospects, we also compute increases in household financial assets that would raise the probability of participation to some desired level, keeping the rest of household characteristics fixed. This gives a monetary measure of how much harder it is for the financial sector to attract the likely potential entrants, given their characteristics, compared to marginal stockholders. This can be done as follows. Suppose that the participation probability of households with demographics x_0 is $1-\tau_0$. Now consider a $\tau_1 < \tau_0$, that would imply a higher probability of participation. Construct a new vector x_1, fixing all elements to their values under x_0, except for (log) financial assets. Solve for the level of log financial assets that would bring about $x_1'\beta(\tau_1) = 0$. From this we can derive the increase in financial assets that would bring about the higher probability of participation $1 - \tau_1$, given the remaining characteristics of the household. Results of all such exercises are presented in Section 5.3.

4 Portfolio Data and Participation Trends

We use pooled data from the most detailed sources on household portfolios, the United States Surveys of Consumer Finances (SCF) for 1995 and 1998. The SCF provides information on whether households hold stocks directly, on whether they hold stocks indirectly (e.g., through mutual funds or defined contribution pension funds), and on a range of demographic characteristics as well as on attitudes towards risk taking, borrowing, saving, etc. Although participation in direct stockholding by households can also be observed in some European countries, the US SCF allows a much more precise estimate of participation in indirect stockholding, since it also asks respondents whether their mutual funds invest in stocks.

Bertaut and Starr-McCluer (2001) provide a comprehensive account of stockholding trends. They report an overall (direct and indirect) stockholding participation rate among households of 48.9 percent in 1998, compared to 40.4 in 1995 and to only 31.6 in 1989. This big spread of equity culture cannot be traced to an increase in direct participation in stocks. Indeed, by 1998, direct participation in stockholding was at 19.2 percent, roughly where it was in 1983, bouncing back from 15.2 percent in 1995. By contrast, participation in indirect stockholding rose dramatically. Mutual fund ownership rose from 4.5 percent in 1983 to 12.3 percent in 1995, to 16.5 percent in 1998. The demographic transition to an aging population and the concomitant policy measures induced a move away from defined-benefit pension plans and towards defined-contribution plans. The share of households having a tax-deferred retirement account –either IRA or 401(k)-type– rose from about 31 percent in 1983 to about 48 percent in 1998.

These increases in participation were occurring at a time when stock returns were rising and costs of stock market participation were falling. The S&P 500 stock price index rose from 165 in 1983 to 600 in 1995 and to 1,100 in 1998. At the same time, the number of mutual funds rose from 564 in 1980 to 6,778 in 1998 (Investment Company Institute, 1998). Rea and Reid (1998) report figures for "total shareholder cost", that includes fund operating expenses plus distribution costs, expressed as a percentage of the amount invested in the fund. The sales-weighted average of such cost ratios for different equity funds was 2.25 in 1980, 2.17 in 1988, and only 1.49 by 1997.[28]

5 Econometric Results

We consider the choice to participate in the stock market either directly or indirectly, through mutual funds or retirement accounts. The latent variable in the econometric specification represents gains net of any entry or participation costs perceived by the household. We control for a number of household characteristics and for year effects, using the discussion of determinants of gains in Section 2.6 to interpret relevant findings. In presenting our results, we compare smoothed binary quantile regression estimates to those obtained from traditional probit and logit estimation.

5.1 Regressors

Household resources (cash on hand) are measured by the logarithms of financial assets and of total household income. We control for the education level of the household head by including a dummy variable for high-school dropouts and one for heads with college education or more. The omitted dummy is for high-school graduates. In Section 2.6, we found conflicting effects of education on gains from access net of access costs.

We allow for a possibly non-linear effect of age on the latent variable by including a linear and a quadratic term in the age of the household head. Age was shown in subsection 2.6 to have sizeable effects on gains from stock market access. On the other hand, existing empirical evidence on the importance of age for participation has been mixed.[29]

Measuring the degree of risk aversion in Survey data is difficult.[30] The Survey of Consumer Finances tries to elicit instead information on willingness to undertake financial risk (such as stockholding risk) by asking households to put themselves in one of four categories, ranging from "not willing to take any financial risk" to "willing to take above average risk for above average expected return". We introduce a dummy variable identifying households that declare lack of willingness to undertake any financial risk. We also control for whether the household head is non-white or Hispanic, is not working, and is self-employed. Table 1 presents some descriptive statistics for the included regressors and for some additional variables in our data set, distinguishing between the 1995 and 1998 samples.

5.2 Standard Approaches: Probit and Logit

The standard approach to estimation of participation equations, namely probit or logit regression, provides a useful benchmark. Results from these two approaches (Table 2) are quite similar and consistent with the empirical literature cited in the Introduction.

Education is estimated to have positive and statistically significant effects on mean gains from access net of entry and participation costs, and hence on participation, as is typical in existing literature. Age is estimated to have a nonlinear, hump-shaped, effect on participation. Thus, controlling for other factors, probit and logit regressions imply that participation is most likely in the middle range of ages, somewhat less likely for the very young and considerably less likely for older households. Hump-shaped patterns of participation also appear in simple graphs of participation data in the US, though Guiso et al. (2003) have pointed out that this robustness of the hump-shaped age-participation pattern to econometric controls is not necessarily observed across European countries.

The dummy variable declaring lack of willingness to undertake financial risk exhibits statistically significant negative effects on participation in any kind of stockholding, controlling for other factors. Being more affluent, in terms of current income or financial assets, makes a household more likely to participate in stocks, consistent with intuition and with the computational model. Self employment actually reduces the probability of participation. It is possible that this is due to the small probabilities of income disasters that face the self-employed but are less likely to influence salaried employees. Similarly, not working reduces the likelihood that the household will decide to put money in the stock market. The race variable is typically negative and statistically significant in participation probits, controlling for incomes and wealth. It is hard to know what causes this effect that appears also in many other studies, but it may be due to the practice of the financial services industry in the US to target to a lesser extent minorities for activity related to the stock market.

The year dummy for the 1998 Survey relative to the 1995 Survey is positive and statistically significant, suggesting that conditions were overall more conducive to stock market participation at the end of the decade than in the middle. Combined with our computational results, the positive estimate suggests that adverse effects of increased overall participation on equity premia and volatility were either negligible or, at any rate, small relative to the beneficial effects of lower access costs in mutual funds and in retirement accounts that were observed as the decade progressed.

5.3 Quantile Regression

5.3.1 Coefficient Estimates

As explained in Section 3, quantile regression (QR) estimates the influence of each covariate, on net gains of households with given characteristics, at different percentiles of the net gains distribution, rather than simply at the mean. In this Section, we discuss our estimates and how they differ from standard participation probit estimates. Figures 9a and 9b plot binary quantile regression estimates for each fifth percentile and regressor. Probit estimates are plotted as a straight, dotted line, for comparison. The solid line in each panel is the zero line. The light- and dark-shaded areas represent the 90 percent and 95 percent confidence intervals, respectively. Confidence intervals were obtained by percentile bootstrap. In the range of percentiles for which the zero line lies within the chosen confidence interval, the coefficient estimates obtained by quantile regression methods are not statistically different from zero at the corresponding confidence level.

It is worthwhile to stress the meaning of a coefficient estimate at a given quantile τ. This estimate shows the effect of the corresponding covariate on net gains for households who are at the τth percentile of the distribution of gains conditional on their characteristics, x. In addition to this, it shows the effect of the covariate on net gains for households that satisfy $x'\beta(\tau) = 0$, i.e. for households with probability of participation conditional on their characteristics equal to $1 - \tau$. This latter group of households is a subset of the former, as indicated by (37).

QR supports the importance of current cash on hand for participation, but shifts some weight from financial assets to household income, and emphasizes the role of income around the participation margin, relative to probit. Fig. 9b (LOGINC) shows that income has a statistically significant, positive, nonlinear effect virtually throughout the distribution of net gains conditional on characteristics. Estimated effects are larger between the 40th and 60th percentiles, and larger than probit estimates. These refer to households with net gains around the median of their conditional distribution. They also apply to households with estimated conditional probabilities of participation between 60% and 40%, around the unconditional probability for US households. QR point estimates are below probit estimates for most other income quantiles.[31] Financial assets (Fig. 9b, LOGFIN) have consistently positive, significant effects on net gains, regardless of the household's position in the conditional distribution. For the most part, i.e., between the 30th and the 80th percentile, QR estimates are fairly constant and close to probit estimates, though somewhat smaller. All in all, changes in current resources, and especially in household income, caused by recession or other factors can be important for maintaining or expanding the participation frontier.

Age has statistically significant, nonlinear effects on net gains, virtually throughout the conditional distribution (see Fig. 9a, AGE and AGESQ for the linear and quadratic age term, respectively). Figure 10 plots the age profile of net gains from for different percentiles using QR point estimates of coefficients on the linear and squared terms and setting other variables to their sample means.The age profile of net gains is concave and downward sloping at all percentiles. It is estimated to be higher than implied by probit around the 50th percentile of the conditional distribution of gains, and for the participation margin; lower for high percentiles, and much lower for low percentiles.

Perhaps the most striking difference with probit refers to the much weaker role of educational attainment in QR (Fig. 9b). In standard probit (and logit) regressions, there is a clear-cut, positive role of education. In QR, having a college degree (COLL) has a positive but statistically insignificant effect on net gains throughout their conditional distribution. Being a high-school dropout (LTHS) has a statistically significant negative effect on net gains, but only for those in the bottom thirty five percent of the distribution. This result also implies that low education has a negative effect only on households with conditional participation probability in excess of 65%. The QR estimate suggests that the negative probit estimate on low education may be caused largely by a large negative effect of this factor at the lower end of the gains distribution. One possible explanation for such mixed findings on the role of educational attainment may be the tension between lower entry/access costs and more limited incentive to benefit from the equity premium that was uncovered in the computational section. In this light, QR estimates for households at the bottom 35% of the conditional distribution suggest heavier entry/access fees for high-school dropouts by amounts sufficient to offset the extra benefits provided by stockholding to households with the age-income profile and income variances of high-school dropouts. In the remaining 65% of the conditional net gains distribution, effects of cost increases seem to be roughly in line with the additional benefits from stockholding.

The importance of attitudes to risk found in standard regressions is confirmed by our findings (Fig. 9a, NORISK). The estimated coefficient is negative and statistically significant for virtually all percentiles of the net gains distribution conditional on characteristics. QR point estimates suggest bigger effects on net gains for households between the 30th and 50th percentiles than the probit estimate, though the latter is within the confidence intervals for QR estimates in this range. These estimates also imply bigger effects for households with conditional participation probabilities between 70% and 50%. QR estimates are smaller (in absolute value) than the probit estimate (and the latter lies outside the confidence intervals) for percentiles above the 60th.

Although QR estimates for the influence of the overall economic environment, as proxied by the 1998 dummy, are consistently positive across quantiles, they are statistically significant only for low percentiles (up to the 35th) and marginally significant for households between the 40th and the 55th percentiles (see Fig. 9b, D1998). This implies that such effects are statistically significant for households with conditional participation probability in excess of 65% and marginally significant for households with probability between 60% and 45%. This picture is quite different from the positive dummy coefficient estimate in probit regression. It suggests that supply-side developments between 1995 and 1998, such as drops in participation costs and general euphoria about the stock market, may have influenced participation choices of households with high conditional probabilities of participation, but their role close to the participation margin is less likely to be significant. It also suggests that reversal of the overall positive climate to stockholding could have significant effects on participation of current stockholders, especially if it is reinforced by a worsening of their own household circumstances, such as income, wealth, attitude to financial risk taking or employment status.[32]

Race and ethnic origin (Fig. 9a, NONWHITE) have negative effects, but statistically significant only up to the 25th percentile and marginally significant between the 30th and the 55th percentile. The fact that minority effects are not found to be significant at other points in the distribution suggests that the usual explanation for the negative race coefficient in probits, namely more limited targeting of minorities by the financial industry, is likely to be confined to those with higher conditional probabilities of participation than 50%. Finally, households with self-employed heads (Fig. 9a, SELFEMP) tend to experience lower net gains from stock market access. Effects of self-employment are estimated to be bigger (in absolute value) for those around and beyond the participation margin. This suggests that self-employment is likely to become a progressively more important deterrent to participation as the stockholder base progressively expands. The next Section estimates the demographic composition and other relevant characteristics of households at different points in the distribution of gains, with emphasis on marginal stockholders and on most likely future entrants.

5.3.2 Demographics of Marginal Stockholders and of Potential Entrants

We follow the approach described in Section 3 to rank households in terms of participation probabilities conditional on their characteristics, divide them up into four quartiles, and compute group demographics using population weights. Results are shown in Table 3.a, for 1995 in the top panel and for 1998 at the bottom. The first column shows the share of each group in the population, while the second to fifth columns show the corresponding shares in each quartile of the population ranked by participation probabilities. Of particular interest are the two middle quartiles, of 'marginal stockholders' as defined above, i.e. those with conditional probability of participation between 50-75%; and of 'potential entrants' with respective probability between 25-50%.

The top and bottom quartiles can be used to gain more confidence in the results, since we find that they accord well with established notions about the characteristics of those most and least likely to participate. In 1995, the quartile with the highest conditional participation probabilities is comprised, almost exclusively, by the richest 50% of the population (in terms of financial assets or of income). Only 4 percent of the bottom quartile consists of those rich in financial assets and 19 percent of those who are income-rich. Households aged between 35 and 64 are overrepresented in the top quartile, compared to their population share, while this is true of the youngest and oldest ages in the bottom quartile. Minorities are underrepresented in the top and overrepresented in the bottom quartile of the distribution of gains. Non-working households and those reporting lack of willingness to undertake any financial risk are under-represented in the top quartile and seriously over-represented in the bottom. Representation of college graduates in the top quartile is more than twice their population share, while in the bottom it is about one third. High school dropouts are absent from the top quartile, but they are seriously over-represented in the bottom. High-school graduates are under-represented in the top quartile, but their share of the bottom quartile is about as much as in the population. Results for 1998 are qualitatively similar and numerically quite close to those for 1995.

Comparison of the two intermediate quartiles can shed light on the obstacles that further spread of equity culture is likely to meet, in view of differences in the profile of potential entrants compared to marginal stockholders. Such comparison shows that, in 1995, about half of potential future entrants had below-median financial assets and income, almost 90 percent belonged to the middle two quartiles of financial assets, and about two thirds to the middle two quartiles of income. This outlook changed little between 1995 and 1998, with the exception that in 1998 two thirds of potential entrants were in the middle two quartiles of income. The majority of potential entrants were below 49 years old in 1995, dropping to 44 percent in 1998. Minorities were a slightly greater proportion of potential entrants than of marginal stockholders in 1995, but this was reversed in 1998. The employment status of potential entrants was similar across 1995 and 1998, and quite similar to the pool of marginal stockholders.

Perhaps the most striking difference between potential entrants and marginal stockholders lies in their declared attitudes towards financial risk. In 1998, almost two thirds of potential entrants reported lack of willingness to undertake any financial risk, compared to only one quarter of marginal stockholders. Moreover, while the incidence of such responses dropped significantly in the population between 1995 and 1998, from 46 to 34 percent, it slightly increased among those most likely to be next in line for entering the stock market. Another interesting finding is that the two lower education strata are more heavily represented among potential entrants than among marginal stockholders, comprising three quarters in both years. This happens, despite a fall in their proportion in the 1998 population.

These findings can be compared to those based on probit and logit estimation. Although probit and logit demographics almost coincide (see Tables 3.b and 3.c), there are some differences with the demographics based on QR, especially in the bottom two quartiles that refer to future entrants.[33] Still, the overall impression from these demographics is the same, regardless of the method used.[34]

Table 3d sheds further lights on the characteristics of potential entrants and of marginal stockholders, by reporting some additional characteristics not included as covariates in the regressions. Median net worth drops by about $ 18,000 as we compare these two groups in 1995, but by half of that in 1998. Homeownership may affect willingness to enter the stock market through various channels with conflicting implications. Homeowners with mortgages have monthly payment commitments that may discourage them from holding their savings in risky form (Fratantoni, 2001). On the other hand, they also have access to home equity loans not available to renters. Renters accumulating down payments for a home may be similarly unwilling to assume financial risk and jeopardize the value of their down payment accumulation. Due to investment in housing, younger and poorer investors have limited financial wealth to invest in stocks, which reduces the benefits of equity market participation (Cocco, 2004). About three quarters of marginal stockholders in 1995 and in 1998 are homeowners, and these percentages drop by 7-8 percentage points among potential entrants. However, dramatic changes in homeownership profiles (by more than 20 percentage points) are not observed until we get to the bottom quartile of households in terms of participation probabilities.

Ownership of private business equity often takes up a large proportion of a household's net worth, and it may discourage introduction of additional, stockholding risk into the portfolio. Still, Carroll (2001) found that rich households are both much more likely to hold private equity and riskier financial portfolios than the rest. Business equity owners are underrepresented among potential stockholders and overrepresented among marginal stockholders in both years, with a difference of 4 percentage points, or more than one third of such owners in the population.

The SCF includes a question on whether households shop around for the best saving or investment opportunities, which can be used as an indicator of financial alertness. The proportion who declare that they do drops quite dramatically, from about one third of marginal stockholders to about a quarter of potential entrants, well below their overall share in the population. Finally, potential entrants have substantial under-representation of households at the two extreme quartiles of the overall distribution of net worth, and over-representation of households in the middle two quartiles. By contrast, marginal stockholders have over-representation of the top two net worth quartiles.

The more limited willingness of potential entrants to undertake financial risk, their lower financial alertness when it comes to saving and investment opportunities, the worsening of their educational background, their younger age, and their more limited overall resources are likely to pose challenges for the financial sector as it attempts to expand the stockholding frontier further.

As described in Section 3, we can also use quantile regression estimates to compute a monetary measure of how much harder it is likely to be for the financial sector to attract this pool of potential stockholders relative to its experience with marginal stockholders. Since we have coefficient estimates for every fifth percentile, we compute these measures by grouping households over every five percentiles, retaining the same degree of precision as above. Table 4 presents increases in household financial assets that would raise the probability of participation of potential stockholders by 25 percentage points, to make up for their remaining characteristics and reach corresponding probabilities of participation of current marginal stockholders. The first column presents the current probability of participation, aggregated over each five percentiles of the distribution of households ranked by participation probabilities. The third column shows median increases in financial assets required for households in the corresponding five percentiles, while the second and the fourth columns show the 25th and 75th percentiles of such increases in financial assets. Median increases in financial assets necessary to move potential entrants to their corresponding position in the distribution of marginal stockholders range from about \$ 33,000 to about \$ 150,000. No matter which statistic one uses, the general impression from Table 4 is that considerable increases in financial assets would be needed to make up for shortcomings in the remaining characteristics of potential stockholders and induce them to be as receptive to stockholding opportunities as current marginal stockholders.

6 Concluding Remarks

In this paper, we estimated the potentially changing role of factors that influence gains from access to stockholding opportunities net of any entry and access costs, across the distribution of such gains conditional on household characteristics. We highlighted potentially conflicting influences on net gains with the help of a computational portfolio model, and we used empirical estimates to derive differences in characteristics of potential entrants relative to marginal investors by the end of the 1990s, when most of the dramatic recent expansion in the stockholder base had already taken place.

We found clear nonlinear effects of financial resources (income and financial assets) on net gains from stock market access. Estimated effects of income and financial assets, as well as of switches to the status of not working were found to be significant for households around the participation margin, suggesting that economic downturns could seriously affect participation decisions of such households. Although standard econometric techniques find strong positive effects of education on participation, quantile regression estimates suggest that education effects refer mainly to having less than high-school education and are significant only for the bottom thirty percent of the conditional distribution of net gains and for households with conditional probabilities of participation in excess of 70%. Through computations, we highlighted two conflicting possible effects of education on net gains, controlling for other characteristics: a positive effect from a reduction in access costs for more educated households, and a negative one from a reduced need to build up assets due to smaller variance of income shocks and steeper age-income profile. Willingness to undertake financial risk is estimated to influence participation across the entire distribution of net gains, while statistically significant positive year effects between 1995 and 1998 are observed only at the bottom of the distribution and for households with high conditional participation probabilities.

Our findings on characteristics of those estimated to be marginal stockholders suggest, when compared to those of potential entrants, that extensive further growth in stock market participation is likely to pose considerable challenges, in view not only of recessions and stock market downturns, but also of more limited finances, younger age, more limited education and financial alertness, and above all significantly less self-declared willingness to assume financial risk by potential stockholders compared to marginal investors. This poses a hurdle to financial practitioners interested in expanding the stockolder base, whose magnitude is not estimated to be small.

A Appendix: Numerical Dynamic Programming

In this Appendix, we describe briefly the algorithm Haliassos and Mavridis developed for solving finite-horizon portfolio choice models under short sales constraints. The corresponding single-asset saving models needed for computation of gains from stock market access were solved by converting into MATLAB Chris Carroll's Mathematica code described in Carroll (2002). Our MATLAB code for the portfolio model benefits from some of Carrol's proposed computational shortcuts for the saving model, but extends them to multi-asset models in a somewhat different direction from that proposed by Carroll.

The presence of fixed entry and per-period access costs can render solution of the optimization problem quite complicated. Value functions are not in general differentiable, the problem is not convex, and first-order conditions are not sufficient to determine the optimum. We are not aware of existing computational studies that handle both types of costs in general form. We handle the problem posed by fixed entry costs by computing separate value functions under access and no access, and computing gains as the difference between the two. As explained in the text, we introduce one simplification: we abstract from people who, after having gained access, decide to remove stocks from their perceived asset menu and not even consider investing in stocks. We assume instead that, once access is obtained, households consider the possibility of stock investments from that time on. Each period, they engage in costly decision making (paying the per-period access cost) to figure out whether their cash on hand merits positive stockholding or not. This simplification allows us to focus on first-order conditions in deriving optimal solutions. Finally, we make sure that households with access are always able to pay the per-period access fee without violating their constraints. Essentially, households with access to stocks and current cash on hand X_t are programmed to know that they have at their disposal only $X_t - F_t$, that their choices will result in having $X_{t+1} - F_{t+1}$ tomorrow, and that they should not violate short sales constraints or nonnegativity of consumption.

Policy functions for the portfolio model with two short sales constrains contain three regions: a low range of normalized cash on hand in which both constraints are binding; an intermediate range in which the household faces a binding constraint that prevents it from borrowing at the riskless rate; and an open-ended range of normalized cash on hand in which neither short sales constraint is binding. We handle these ranges in reverse order. We solve first for the penultimate period of life, $T-2$, and then go backwards in time.

A.1 Unconstrained Range

In this range of cash on hand, the FOC for the portfolio share in the second-to-last period of life, α_{T-2}, is

$$s_{T-2}\ E_{T-2}\left\{G_{T-1}^{-\rho}\left(\widetilde{R}_{T-1} - R_{T-1}\right)U'(x_{T-1})\right\} = 0 \tag{38}$$

since the household consumes all cash on hand in the last period of life, $c_{T-2} = x_{T-2}$, and since there are no shocks to the permanent component of income during retirement ($N \equiv 1$). Using the transition equation for cash on hand, this can be written as

$$s_{T-2}\ E_{T-2}\left\{(G_{T-1})^{-\rho}\left(\widetilde{R}_{T-1} - R_{T-1}\right)U'\left(\left[R_{T-1} + \left(\widetilde{R}_{T-1} - R_{T-1}\right)\alpha_{T-2}\right]s_{T-2}\frac{1}{G_{T-1}} + U_{T-1}\right)\right\} = 0 \tag{39}$$

Carroll's proposed computational shortcut is based on the observation that a FOC such as this is an equilibrium relationship between optimal choices (here α_{T-2} and s_{T-2}). Generalized to many controls, the approach is to specify a grid of possible optimal values for one control (here: savings values $s^i_{T-2},\ i = 1,...,n$) and find the corresponding optimal values for the remaining controls (portfolio shares), using grids for transitory income shocks, U_{T-1}, shocks to the permanent component of income, N_{T-1}, and stock returns, $\widetilde{R}_{T-1}$ obtained by discretizing their corresponding distributions. In our paper, we used 100 grid points for s, and 50 grid points for each of U, N, and $\widetilde{R}$. In Caroll's consumption-saving case with only one control, consumption can be found by mere evaluation of the FOC. When controls are more than one (as here), one can solve the FOCs using a non-linear solver.

The minimum admissible value in the savings grid is (slightly above) what would be required to produce zero consumption in the final period, c_T, under the worst possible configuration of income and stock return shocks, and no less than zero (since borrowing is not allowed):

$$s^{Min}_{T-2} \geq \max\left(\frac{-U^{Min}_{T-1} G_{T-1}\ N^{Max}_{T-1}}{\widetilde{R}^{Min}_{T-1}} + \varepsilon, 0\right) \tag{40}$$

where ε is small. For each savings value in the grid, we use MATLAB's nonlinear equation solver to solve the FOC for the optimal portfolio share that would correspond to it. We thus come up with a set of pairs $(\alpha^i_{T-2}, s^i_{T-2})$. For the time being, we consider only the pairs that do not violate either short sales constraint, namely those that satisfy $0 < \alpha_{T-2}(s_{T-2}) < 1$. For such pairs, the FOC for saving

$$\begin{aligned} U'(c_{T-1}) &= \widehat{\beta}\ E_{T-2}\{(G_{T-1})^{-\rho}\left[R_{T-1} + (\widetilde{R}_{T-1} - R_{T-1})\alpha_{T-2}\right] \cdot \\ &\quad U'\left(\left[R_{T-1} + \left(\widetilde{R}_{T-1} - R_{T-1}\right)\alpha_{T-2}\right] s_{T-2}\frac{1}{G_{T-1}} + U_{T-1}\right)\} \end{aligned} \tag{41}$$

and the assumption of CRRA felicity allow evaluation of the corresponding optimal consumption level for each $(\alpha^i_{T-2}, s^i_{T-2})$:

$$\begin{aligned} c^i_{T-2} &= [\widehat{\beta}\ E_{T-2}\{(G_{T-1})^{-\rho}\, R_{T-1} + (\widetilde{R}_{T-1} - R_{T-1})\ \alpha^i_{T-2}] \cdot \\ &\quad U'\left(\left[R_{T-1} + \left(\widetilde{R}_{T-1} - R_{T-1}\right)\ \alpha^i_{T-2}\right]\ s^i_{T-2}\frac{1}{G_{T-1}} + U_{T-1}\right)\}]^{-\frac{1}{\rho}} \end{aligned} \tag{42}$$

This determines triplets of optimal consumption, savings, and portfolio shares $(c^i_{T-2},\ s^i_{T-2},\ \alpha^i_{T-2})$, from which the corresponding normalized cash on hand level can easily be computed by noting that $x^i_{T-2} = c^i_{T-2} +\ s^i_{T-2} +\ f_{T-2}$, where f denotes the fixed access cost. Consumption, saving, and portfolio shares are thus mapped to the levels of normalized cash on hand for which they would be optimal. Policy rules $c_{T-2}(x_{T-2}),\ s_{T-2}(x_{T-2})$, and $\alpha_{T-2}(x_{T-2})$ for the range of normalized cash on hand where the short sales constraints are not binding are computed through (linear) interpolation between these points.

A.2 Binding Constraint on Borrowing at the Riskless Rate

In principle, there are two types of portfolio share-savings pairs that violate the borrowing constraints. One exhibits $\alpha^i_{T-2}(s^i_{T-2}) < 0$. This type is not relevant for our problem, as we are assuming zero correlation between stock returns and income shocks and it will not be optimal for households to short stocks in order to invest in the riskless asset. The second type, $\alpha^i_{T-2}(s^i_{T-2}) > 1$, is relevant and represents cases where the household would like to borrow at the riskless rate to invest in stocks. To solve for policy rules in this range, we first use the procedure described in the previous subsection to compute the normalized cash on hand that would correspond to each such pair in the absence of the borrowing constraint. This pinpoints the levels of normalized cash on hand for which constrained solutions should be sought.

Now, we know that constrained solutions involve saving (exclusively) in stocks and that the short sales constraint is not binding, unlike the borrowing constraint at the riskless rate. Thus the FOC for stocks is:

$$U'(x_{T-2} - s_{T-2}) = \widehat{\beta}\ E_{T-2}\left\{(G_{T-1})^{-\rho}\,\widetilde{R}_{T-1}\ U'\left(\widetilde{R}_{T-1}\ s_{T-2}\frac{1}{G_{T-1}} + U_{T-1}\right)\right\}.$$

Using MATLAB's nonlinear solver, we solve this FOC for the optimal saving (in stocks), s^i_{T-2}, that corresponds to the x^i_{T-2} computed in the previous step. This allows us also to compute consumption as $c^i_{T-2} = x^i_{T-2} - s^i_{T-2} - f_{T-2}$. Since the constrained portfolio share is 1, this completes the solution for the particular x^i_{T-2}. We can repeat this procedure for a number of levels of normalized cash on hand, but it is advisable to find first the lower limit of this constrained range. This is the minimum level of normalized cash on hand, x^*_{T-2}, for which the FOC for investment in stocks still holds, while the constraint on borrowing at the riskless rate is binding. Below this level, both constraints are binding.

To determine x^*_{T-2}, we use the FOC for investment in stocks, since households simply choose to hold no stocks (in the limit) as opposed to being prevented from short-selling stock:

$$U'(X^*_{T-2} - F_{T-2}) = \widehat{\beta}E_{T-2}\left\{\widetilde{R}_{T-1}U'(Y_{T-1})\right\}. \tag{43}$$

Upon normalization, this can be used to evaluate x^*_{T-2} as follows:

$$x^*_{T-2} = \left[\widehat{\beta}E_{T-2}\left\{\widetilde{R}_{T-1}U'(G_{T-1}U_{T-1})\right\}\right]^{-\frac{1}{\rho}} + f_{T-2} \tag{44}$$

A.3 Both Constraints Binding

In the range of low cash on hand up to x^*_{T-2}, the household ends up consuming its resources. This pinpoints the solution fully, as $c_{T-2} = x_{T-2} - f_{T-2}$, $s_{T-2} = 0$, and the portfolio share is strictly not defined for non-savers. The policy rule in this range can be plotted simply by drawing the 45-degree line between the point (x^*_{T-2}, x^*_{T-2}) and the origin (or, more strictly, point $(U^{Min}$, $U^{Min})$, since normalized cash on hand cannot follow below U^{Min} in this model). This extends a suggestion to this effect by Carroll for the saving model.

Thus, a complete solution for all three relevant ranges of normalized cash on hand has been derived for period $T-2$. Solutions for periods prior to this can be obtained recursively in a similar fashion, with one modification. Instead of imposing that $c_{t+1} = x_{t+1}$, which can be assumed for the entire range of normalized cash on hand only for $t = T-2$, we must use instead the policy function computed in the previous round, $c_{t+1}(x_{t+1})$, interpolating between points as needed. Value functions are then computed using the Bellman equation and the policy rules. Given that we need to compute differences between value functions from the portfolio and from the single-asset models, we adopt a relatively fine grid of 1000 points for evaluating (and comparing) value functions.

References

[1] Aizcorbe, Ana M., Arthur Kennickell, and Kevin B. Moore (2003). "Recent Changes in U.S. Family Finances: Evidence from the 1998 and 2001 Survey of Consumer Finances", *Federal Reserve Bulletin*, January, 1-32.

[2] Allen, Franklin and Douglas Gale (1994). "Limited Market Participation and Volatility of Assets Prices", *The American Economic Review*, 84, 933-55.

[3] Arrow, Kenneth (1974). *Essays in the Theory of Risk Bearing.* Amsterdam: North Holland.

[4] Basak, S. and D. Cuoco (1998). "An equilibrium model with restricted stock market participation", *The Review of Financial Studies*, 11, 309–41.

[5] Bertaut, Carol C. and Martha Starr-McCluer (2001). "Household Portfolios in the United States" in Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (eds.) (2001a). *Household Portfolios*, Cambridge, MA: MIT Press.

[6] Calvet, Laurent, Martin Gonzalez-Eiras, and Paolo Sodini (2001). "Financial Innovation, Market Participation and Asset Prices", Harvard Institute of Economic Research Discussion Paper No. 1928.

[7] Campbell, John, Martin Lettau, Burton Malkiel, and Yexiao Xu (2001). "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *The Journal of Finance*, 56, 1-44.

[8] Campbell, John and Luis Viceira (2002). *Strategic Asset Allocation: Portfolio Choice for Long-term Investors*, Oxford: Oxford University Press.

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

[10] Carroll, Christopher D. (2001). "Portfolios of the Rich" in Guiso, L., M. Haliassos, and T. Jappelli (eds.), *Household Portfolios*, Cambridge, MA: MIT Press.

[11] Carroll, Christopher D. (2002). "Lecture Notes On Solution Methods for Microeconomic Dynamic Stochastic Optimization Problems", mimeo, Johns Hopkins University.

[12] Carroll, Christopher D., and Andrew A. Samwick (1997). "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40, 41-71.

[13] Cocco, Joao (2004). "Portfolio Choice in the Presence of Housing", IFA Working Paper 410, London Business School.

[14] Cocco J., F. Gomes, and P.J. Maenhout (1997). "Consumption and Portfolio Choice over the Life Cycle," Working Paper, Harvard University.

[15] Davis, Steven J. and Paul Willen (2000). "Occupation Level Income Shocks and Asset Returns: Their Covariance and Implications for Portfolio Choice", NBER Working Paper No. 7905.

[16] Davis, Steven J., Felix Kubler, and Paul Willen (2004). "Borrowing Costs and the Demand for Equity over the Life-Cycle", Mimeo, Graduate School of Business, University of Chicago.

[17] Deaton, Angus (1991). "Saving and Liquidity Constraints," *Econometrica*, 59, 1221-48.

[18] Fratantoni, Michael C (2001). "Homeownership, Committed Expenditure Risk, and the Stockholding Puzzle", *Oxford Economic Papers*, 53, 241-59.

[19] Gollier, Christian (2001). *The Economics of Risk and Time*, Cambridge, MA: MIT Press.

[20] Gomes, Francisco and Alexander Michaelides (2004). "Optimal Life-Cycle Allocation: Understanding the Empirical Evidence", forthcoming, *Journal of Finance.*

[21] Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (eds.) (2001a). *Household Portfolios*, Cambridge, MA: MIT Press.

[22] Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (2001b). "Household Portfolios: An International Comparison", in Guiso, L., M. Haliassos, T. Jappelli (eds.). *Household Portfolios*, Cambridge, MA: MIT Press.

[23] Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (eds.) (2002). *Stockholding in Europe*, Houndmills, UK and New York, USA: Palgrave Macmillan.

[24] Guiso, Luigi, Michael Haliassos, and Tullio Jappelli (2003). "Stockholding in Europe: Where Do We Stand and Where Do We Go?" *Economic Policy*, 36, 117-64.

[25] Guiso, Luigi and Monica Paiella (2001). "Risk Aversion, Wealth, and Background Risk," CEPR Discussion Paper No. 2728.

[26] Haliassos, Michael (2002). "Stockholding: Recent Lessons from Theory and Computations" in Guiso, L., M. Haliassos, and T. Jappelli (eds.), *Stockholding in Europe*, Houndmills and New York: Palgrave Macmillan Publishers.

[27] Haliassos, Michael and Carol C. Bertaut (1995). "Why Do So Few Hold Stocks?" *The Economic Journal,* 105, 1110-1129.

[28] Haliassos, Michael and Alexander Michaelides (2003). "Portfolio Choice and Liquidity Constraints," *International Economic Review*, 44, 143-78.

[29] Heaton John, and Deborah Lucas (1997). "Market Frictions, Savings Behavior, and Portfolio Choice," *Macroeconomic Dynamics*, 1, 76-101.

[30] Heaton, John and Deborah Lucas (1999). "Stock Prices and Fundamentals," *NBER Macroeconomics Annual*, 213-42.

[31] Heaton John, and Deborah Lucas (2000). "Portfolio Choice in the Presence of Background Risk," *The Economic Journal*, 110, 1-26.

[32] Herrera, Helios (2001). "Participation Externalities and Stock Market Volatility", mimeo, New York University.

[33] Horowitz, J. L. (1992). "A Smoothed Maximum Score Estimator for the Binary Response Model", *Econometrica*, 60, 505-31.

[34] Horowitz, J. L. (1998). *Semiparametric Methods in Econometrics*, New York: Springer-Verlag.

[35] Horowitz, J.L. (2002). "Bootstrap Critical Values for Tests Based on the Smoothed Maximum Score Estimator", *Journal of Econometrics*, 111, 141-67.

[36] Investment Company Institute (1998). *Mutual Fund Fact Book*, Washington, D.C.: The Investment Company Institute.

[37] Kezdi, Gabor and Robert J. Willis (2003). "Who Becomes A Stockholder? Expectations, Subjective Uncertainty, and Asset Allocation", Working Paper, University of Michigan Retirement Research Center.

[38] King, M. and J. Leape (1984). "Wealth and Portfolio Composition: Theory and Evidence," NBER Working Paper No. 1468.

[39] Koenker, R. and G. Basset (1978). "Regression Quantiles", *Econometrica*, 46, 33-50.

[40] Koenker, R. and G. Basset (1982). "Robust Tests for Heteroskedasticity Based on Regression Quantiles", *Econometrica*, 50, 43-61.

[41] Kordas, Gregory (2001). "Smoothed Binary Regression Quantiles", Unpublished PhD dissertation, University of Illinois.

[42] Laibson, David, Andrea Repetto, and Jeremy Tobacman (2000). "A Debt Puzzle", NBER Working Paper No. 7879, forthcoming in Philippe Aghion, Roman Frydman, Joseph Stiglitz, and Michael Woodford (Eds), *Knowledge, Information, and Expectations in Modern Economics: In Honor of Edmund S. Phelps.*

[43] Luttmer, Erzo G.J. (1999). "What Level of Fixed Costs Can Reconcile Consumption and Stock Returns?" *Journal of Political Economy*, 107, 969-97.

[44] Mankiw, N.G. and S. Zeldes (1991). "The Consumption of Stockholders and Non-Stockholders," *Journal of Financial Economics*, 29, 97-112.

[45] Manski, C.F. (1975). "Maximum score Estimation of the Stochastic Utility Model of Choice", *Journal of Econometrics*, 3, 205-28.

[46] Manski, C. F. (1985). "Semiparametric Analysis of Discrete Response: Asymptotic Properties of the Maximum Score Estimator", *Journal of Econometrics*, 27, 313–33.

[47] Pagano, Marco (1989). "Endogenous Market Thinness and Stock Price Volatility," *The Review of Economic Studies*, 56, 269-87.

[48] Paiella, Monica (2001). "Limited Financial Market Participation: a Transaction Cost-Based Explanation," IFS Working Paper 01/06.

[49] Peress, Joel (2002). "Wealth, Information Acquisition, and Portfolio Shares," mimeo, INSEAD.

[50] Polkovnichenko, Victor (2000). "Heterogeneous Labor Income and Preferences: Implications for Stock Market Participation," mimeo, University of Minnesota.

[51] Rea, John D., and Brian K. Reid (1998). "Tends in the Ownership Cost of Equity Mutual Funds", *Investment Company Institute Perspective*, 4(3), 2-15.

[52] National Vital Statistics Report, 2001, Vol. 48, No. 18, February 7, 2001.

[53] Storesletten, K., C. Telmer and A. Yaron, "Persistent Idiosyncratic Shocks and Incomplete Markets," mimeo, Carnegie Mellon University, 1998.

[54] Vayanos, Dimitri (1998). "Transaction Costs and Asset Prices: A Dynamic Equilibrium Model." *Review of Financial Studies*, 11, 1-58.

[55] Vissing-Jorgensen, Annette (2002). "Towards an Explanation of Household Portfolio Choice Heterogeneity: Non-financial Income and Participation Cost Structures," NBER Working Paper No. 8884.

Notes

[1]See Guiso, Haliassos, and Jappelli (2001a, b) and the contributions therein. Among them, Bertaut and Starr-McCluer (2001) focus especially on the US, and Carroll (2001) on portfolios of the rich in the US.

[2]For an excellent study of trends in the United States mutual fund industry, see Rea and Reid (1998).

[3]The starting point for this literature is the stockholding participation puzzle, based on the theoretical point made by Arrow (1974) that an expected utility maximizer will find it optimal to invest an amount ε in the asset with the expected return premium. Limited stockholding participation in the early to mid 1980s was documented in US data by King and Leape (1984), Mankiw and Zeldes (1991), and Haliassos and Bertaut (1995). The factors that determine whether participation occurs or not have been explored by a number of authors. See, for example, Haliassos and Bertaut (1995), Cocco, Gomes and Maenhout (1997), Heaton and Lucas (1997, 2000), Storesletten, Telmer, and Yaron (1998), Gollier (2001), Campbell and Viceira (2002), Haliassos and Michaelides (2003). International comparisons are to be found in Guiso, Haliassos, and Jappelli (2001b).

[4]See the contributions in Guiso et al., 2001a.

[5]On the potential role of participation costs, see Haliassos and Bertaut (1995), Luttmer, (1996), Vayanos (1998). Guiso, Haliassos and Jappelli (2003) utilize cross-country variation in participation rates in household-level data from 6 European countries and the US, and produce evidence in favor of a cost-based explanation of participation patterns internationally.

[6]Good examples of recent papers aimed at matching features of stockholding data include Gomes and Michaelides (2004), and Davis, Kubler, and Willen (2004).

[7]For infinite-horizon examples of zero stockholding generated by short sales constraints, see Heaton and Lucas (1997) and Haliassos and Michaelides (2003). The classic paper in the single-asset saving literature that generated zero saving by use of a no-borrowing constraint is Deaton (1991).

[8]Although V_t^S is obviously a function of the assumed sequence of per period participation costs, $\{f_s\}_{s\geq t}$, we will drop this sequence from the list of arguments for notational simplicity.

[9]Carroll (1992) estimates variances of idiosyncratic shocks for total US household population using data from the *Panel Study of Income Dynamics*, and finds 0.01 percent per year both for σ_u^2 and for σ_n^2.

[10]The conflicting effect of higher opportunity cost of time for the more educated households is normally controlled for by allowing explicitly for household income.

[11]Education may also affect household preferences and constraints in various other, less observable ways, such as making them less averse to undertaking financial risk. The effects of risk aversion are discussed separately below.

[12]Results shown are for the benchmark of zero correlation between stock returns and income shocks. Davis and Willen (2000) found that aggregate equity returns are uncorrelated with occupation-level income innovations, but some industry-level portfolios and a portfolio formed on firm size are correlated with income innovations for a number of occupations. Their findings suggest that further empirical study of such correlations is warranted. To the extent that correlations are found to be low or negative for low-education groups and positive for high-education groups, they raise further the gains from access for low-education relative to high-education households, and thus reinforce our findings.

[13]Per period access costs payable to ensure access through the remaining lifetime can result in negative net gains. This would not be possible in the absence of per period access costs, as households always have the option not to hold any stocks.

[14]Note that curves depicting monetary equivalents start at higher levels of normalized cash on hand for high-school dropouts than for the other categories. For levels below the starting point of each of these curves, the household would be willing to borrow in order to pay the fixed entry cost to the stock market, but is unable to do so because of borrowing constraints.

[15]Here we have followed the standard practice of assuming the same survival probabilities across education groups. Since there is some evidence of a link between income and life expectancy, a further factor that might limit saving incentives, and hence net gains, is more limited life expectancy for less educated households. However, there is also evidence of a positive link between wealth and health that could work in the opposite direction.

[16] Although a positive relationship between risk aversion and stockholding has also been found in infinite-horizon models, the age-variations uncovered by a model with finite lives and retirement periods are quite striking. For infinite-horizon models with this property, see Heaton and Lucas (1997) and Haliassos and Michaelides (2003).

[17] An exception to negativity is shown in comparing gains at age 25 and 35. The cause for higher gains at age 35 is that during the period 25 to 35, stockholding is expected to be low relative to the per period access cost, so committing early results in lower gains compared to gaining access later.

[18] Note that it is conceivable that fixed costs of entry also drop with age, as the household becomes more knowledgeable about its work and financial environment in the course of its working life and thus finds it easier to figure out how to begin investing in stocks.

[19] The reason for the portfolio shift away from stocks as the household ages is that it relies more heavily on its portfolio of accumulated assets and less heavily on its human wealth to finance future consumption. As a result, it becomes progressively less willing to invest a large portion of its wealth in the risky asset. (See also Cocco, Gomes, and Maenhout, 1999.) Interestingly, this model prediction is in accordance to the advice given to households by financial advisors, not necessarily with the same motivation.

[20] Thus, borrowing constraints may contribute towards explaining the absence of clear negative age effects on portfolio shares noted in empirical studies. See, for example, Guiso et al. (2003) and the country studies in Guiso et al. (2001a).

[21] A class of general equilibrium models that compare steady states with limited stock market participation to those with full participation imply that asset returns are not likely to change much in response to more extensive participation (see Basak and Cuoco, 1998; Heaton and Lucas, 1999; Polkovnichenko, 2000). This is mainly because marginal stockholders tend to invest much less in stocks than incumbents, since marginal stockholders tend to exhibit higher risk aversion, more limited wealth and other characteristics that limit their demand for stocks. There are also models in which increased participation is associated with reductions in the equity premium (e.g., Peress, 2001; Calvet et al., 2001). However, differences in characteristics of newcomers relative to incumbents could even cause a rise in the premium, e.g. if marginal stockholders are on average more risk averse than incumbents.

[22] Campbell et al. (2001) find that the idiosyncratic volatility of single stocks in the United States has increased significantly over the past 30 years. Over the same time span, also the volatility of the price/earnings ratio of the Standard&Poor 500 index has increased (Herrera, 2001).

[23] New entrants increase market liquidity by bringing previously untapped funds into the stock market. In equilibrium, higher liquidity implies that sellers short of cash can more easily trade with buyers with excess cash. This tends to reduce market volatility (Pagano, 1989; Allen and Gale, 1994). Herrera (2001) points out that if new stockholders are more risk averse than previous stockholders, then their stock demand is less responsive to current stock prices and this can lead to higher price volatility. Peress (2002) points to two conflicting considerations. Although the entry of new investors spreads risks across a bigger pool and this enhanced risk sharing by itself tends to lower volatility, it also reduces incentives to acquire costly information, which exerts upward pressure on volatility. The net effect depends on the number of shareholders. Interestingly, if there is an exogenous reduction in the entry cost to the market for widely held stocks, then volatility increases, because new entrants do not purchase information and they also reduce the incentives of incumbents to purchase information.

[24] Results in this Section can also be put to a different use, namely to show effects of heterogeneous perceptions regarding the equity premium and its volatility among households. Indeed, Kezdi and Willis (2003) provide interesting preliminary evidence that heterogeneous expectations, in terms both of levels and of degree of precision, help predict stockholding participation.

[25] The normalization that the coefficient of a regressor has absolute value equal to 1 is used by Horowitz (1992). It is assumed that this regressor has probability distribution conditional on the other regressors that is absolutely continuous with respect to the Lebesgue measure. The two normalizations are basically equivalent.

[26] See Horowitz (1998, 2002).

[27] Note also that, for given $\beta(\tau)$, the probability statements (37) show that the τth quantile model for y^* implies some relationships between observed data (y, x). Indeed, this was used by Manski (1985) for the identification of parameters $\beta(\tau)$.

[28] Although most of the dramatic increase in stock market participation was completed by the end of the 1990s, participation in 2001 was still slightly higher than in 1998, despite the reversal in stock market performance.

Aizcorbe et al. (2003) report a total stockholding participation rate of 51.9 percent in 2001, up by three percentage points compared to 1998. Direct participation rose to 21.3 versus 19.2 percent in 1998; mutual fund participation rose to 17.7 percent versus 16.5 percent in 1998; and participation in retirement accounts rose from 48.9 percent in 1998 to 52.2 percent in 2001.

[29] See the contributions in Guiso et al. (2001a).

[30] See, for example, Guiso and Paiella (2001) for an attempt to measure risk aversion by reference to a hypothetical lottery offered to respondents.

[31] Probit estimates, however, often lie within the confidence bands of QR estimates, though not below the 20th percentile.

[32] This last point is supported by our findings in Fig. 9a (NOTWORK), where significant negative effects are found for households above the 35th percentile.

[33] QR-based demographics imply more limited representation of households with below-median financial assets among potential entrants than probit/logit. They also imply more limited presence of very young households and of those not willing to undertake any financial risk in this category. As for the bottom quartile, QR implies more limited presence of the poorest in financial assets or income, of whites, and of those willing to take financial risk in 1995. In 1998, QR implies greater presence of minorities, of those not willing to take any risk, and of college graduates than probit or logit.

[34] Any differences between demographics implied by QR and by standard methods result from differences in ranking of households in terms of participation probabilities. A visual representation of such differences is given in Figs. 11a, and b. We split each of the two samples into 20 groups of equal (weighted) size in terms of participation probabilities derived from QR estimates, and compute for each group the (population-weighted) average of probabilities of participation implied by probit (or logit) estimates. These are then compared to probabilities implied by QR estimates, derived as described in Section 3. With one small exception, participation probabilities from all three methods increase together. Probit and logit average probabilities typically fall somewhat below those implied by quantile regression. Consistent overstatement of participation probabilities by the QR method is due to the economizing choice not to estimate coefficients for all percentiles of net gains but to set slightly negative estimates of net gains to zero, thus increasing somewhat the implied probability of participation. Logit average probabilities are below those from probit for those with less than fifty percent participation probability, but this reverses for those with more than fifty percent probability of participation.

Variable	Description	Mean	Std. Dev.	Percentiles 25th	median	75th	95th
DSTOCKS	Indicator for holding stocks directly	0.15	0.006	0	0	0	1
STOCKS	Level of direct stock holding	15103.60	1336.473	0	0	0	21358
DEQUITY	Indicator for direct or indirect stockholding	0.40	0.01	0	0	1	1
EQUITY	Level of direct or indirect stockholding	39442.98	2146.273	0	0	8543	128146
DINDHOL	Indicator for indirect stockholding	0.35	0.009	0	0	1	1
INDHOL	Level of indirect stockholding	24339.38	1477.633	0	0	5339	101449
NONWHITE	Indicator for being non-white or hispanic	0.22	0.008	0	0	0	1
NOTWORK	Indicator for not working	0.06	0.004	0	0	0	1
SELFEMP	Indicator for being self-employed	0.10	0.005	0	0	0	1
NORISK	Indicator for reporting refusal to undertake any financial risk	0.46	0.011	0	0	1	1
AGE	Age of household head	48.46	0.667	35	45	62	79
LTHS	Household head is a high-school dropout	0.19	0.008	0	0	0	1
COLL	Household head has college degree or more	0.31	0.009	0	0	1	1
LOGINC	Log of total income	10.18	0.107	9.709	10.403	10.913	11.683
LOGFIN	Log of financial assets	8.47	0.106	7.004	9.306	10.927	12.685

Variable	Description	Mean	Std. Dev.	Percentiles 25th	median	75th	95th
DSTOCKS	Indicator for holding stocks directly	0.19	0.39	0	0	0	1
STOCKS	Level of direct stock holding	30169.09	558472.9	0	0	0	60000
DEQUITY	Indicator for direct or indirect stockholding	0.48	0.5	0	0	1	1
EQUITY	Level of direct or indirect stockholding	70738.62	714088.7	0	0	22000	235000
DINDHOL	Indicator for indirect stockholding	0.44	0.5	0	0	1	1
INDHOL	Level of indirect stockholding	40569.52	306418.8	0	0	14900	161000
NONWHITE	Indicator for being non-white or hispanic	0.23	0.42	0	0	0	1
NOTWORK	Indicator for not working	0.05	0.22	0	0	0	1
SELFEMP	Indicator for being self-employed	0.11	0.31	0	0	0	1
NORISK	Indicator for reporting refusal to undertake any financial risk	0.39	0.49	0	0	1	1
AGE	Age of household head	48.71	17.32	35	46	61	80
LTHS	Household head is a high-school dropout	0.21	0.41	0	0	0	1
COLL	Household head has college degree or more	0.33	0.47	0	0	1	1
LOGINC	Log of total income	10.21	1.54	9.7	10.42	11	11.75
LOGFIN	Log of financial assets	8.9	3.46	7.31	9.7	11.29	13.05

Table 1: Descriptive statistics of the 1995 (N=4299) and 1998 (N=4305) subsamples, computed using SCF survey weights.

	Probit Model			
Variable	Coefficient	Normalized Coef.	Std. Err	t-stat
Intercept	-0.0429		0.00196	-21.9
Household head is not white or Hispanic	-0.225	-0.229	0.0494	-4.56
Household head is not working	-0.497	-0.506	0.101	-4.92
Household head is self-employed	-0.468	-0.476	0.0459	-10.2
Not willing to take any financial risk	-0.481	-0.489	0.041	-11.7
Age of household head	0.0198	0.0202	0.00651	3.05
Age-squared	$-3.55e-4$	$-3.62e-4$	$6.08e-5$	-5.85
Less-than high-school education	-0.165	-0.168	0.0609	-2.71
College degree or more	0.1	0.102	0.0394	2.54
Log of total income	0.0669	0.0681	0.0137	4.89
Log of financial wealth	0.385	0.391	0.011	34.9
Dummy for 1998 survey	0.168	0.171	0.0356	4.73
	Logit Model			
Variable	Coefficient	Normalized Coef.	Std. Err	t-stat
Intercept	-0.076		0.0035	-21.7
Household head is not white or Hispanic	-0.369	-0.217	0.0858	-4.3
Household head is not working	-0.872	-0.513	0.179	-4.86
Household head is self-employed	-0.805	-0.474	0.0814	-9.89
Not willing to take any financial risk	-0.819	-0.482	0.0705	-11.6
Age of household head	0.0302	0.0178	0.0114	2.66
Age-squared	$-5.82e-4$	$-3.42e-4$	$1.07e-4$	-5.46
Less-than high-school education	-0.279	-0.165	0.106	-2.64
College degree or more	0.167	0.098	0.0687	2.42
Log of total income	0.126	0.0743	0.0242	5.2
Log of financial wealth	0.682	0.401	0.0203	33.5
Dummy for 1998 survey	0.288	0.17	0.0625	4.61

Table 2: Probit and Logit Results for Participation in Direct or Indirect Stockholding

		Ranking by Quantile Regression Estimates			
Household Characteristic	1995 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.65	0.00	0.00	0.00
Between 25% and 50% of financial assets		0.31	0.50	0.24	0.00
Between 50% and 75% of financial assets		0.04	0.37	0.54	0.25
Upper 25% of financial assets		0.00	0.12	0.22	0.75
Lower 25% of income		0.49	0.18	0.04	0.01
Between 25% and 50% of income		0.32	0.37	0.20	0.07
Between 50% and 75% of income		0.16	0.30	0.45	0.25
Upper 25% of income		0.03	0.15	0.31	0.67
Of age less than 34	0.248	0.30	0.21	0.24	0.18
Age between 35 and 49	0.335	0.25	0.31	0.43	0.43
Age between 50 and 64	0.211	0.18	0.18	0.21	0.26
Age more than 65	0.206	0.27	0.31	0.12	0.13
Household head is unemployed	0.064	0.14	0.02	0.01	0.01
Household head is self-employed	0.102	0.10	0.10	0.10	0.11
Household head is an employee	0.834	0.76	0.88	0.89	0.88
Household head is white, non-hispanic	0.776	0.68	0.76	0.80	0.96
Non-white or hispanic	0.224	0.32	0.24	0.20	0.04
Willing to undertake financial risk	0.544	0.30	0.38	0.75	0.98
Not willing to undertake any financial risk	0.456	0.70	0.62	0.25	0.02
Less than high school education	0.186	0.35	0.16	0.09	0.00
High-school education	0.507	0.53	0.60	0.52	0.36
College degree or more	0.307	0.13	0.25	0.39	0.64

		Ranking by Quantile Regression Estimates			
Household Characteristic	1998 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.68	0.00	0.00	0.00
Between 25% and 50% of financial assets		0.28	0.47	0.30	0.00
Between 50% and 75% of financial assets		0.04	0.40	0.47	0.28
Upper 25% of financial assets		0.00	0.13	0.23	0.72
Lower 25% of income		0.57	0.25	0.05	0.03
Between 25% and 50% of income		0.27	0.41	0.26	0.11
Between 50% and 75% of income		0.13	0.27	0.37	0.26
Upper 25% of income		0.03	0.08	0.32	0.60
Of age less than 34	0.233	0.31	0.19	0.24	0.15
Age between 35 and 49	0.334	0.26	0.25	0.39	0.43
Age between 50 and 64	0.230	0.17	0.20	0.24	0.27
Age more than 65	0.203	0.26	0.36	0.14	0.14
Household head is unemployed	0.051	0.12	0.02	0.01	0.02
Household head is self-employed	0.256	0.11	0.09	0.08	0.14
Household head is an employee	0.693	0.78	0.89	0.90	0.84
Household head is white, non-hispanic	0.824	0.64	0.78	0.74	0.95
Non-white or hispanic	0.176	0.36	0.22	0.26	0.05
Willing to undertake financial risk	0.659	0.33	0.36	0.75	0.97
Not willing to undertake any financial risk	0.341	0.67	0.64	0.25	0.03
Less than high school education	0.125	0.41	0.22	0.16	0.00
High-school education	0.432	0.46	0.54	0.53	0.39
College degree or more	0.443	0.13	0.23	0.32	0.61

Table 3.a: Demographic composition of quartiles of the population ranked by participation probabilities. Quartiles were constructed using estimates from quantile regression, along with population weights, for each of the 1995 and 1998 samples. For method of construction, see text.

		Ranking by Probit Estimates			
Household Characteristic	1995 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.81	0.13	0.00	0.00
Between 25% and 50% of financial assets		0.18	0.57	0.25	0.00
Between 50% and 75% of financial assets		0.01	0.25	0.52	0.25
Upper 25% of financial assets		0.00	0.04	0.23	0.75
Lower 25% of income		0.61	0.20	0.07	0.01
Between 25% and 50% of income		0.28	0.39	0.24	0.06
Between 50% and 75% of income		0.10	0.29	0.42	0.26
Upper 25% of income		0.01	0.12	0.27	0.67
Of age less than 34	0.248	0.28	0.27	0.25	0.17
Age between 35 and 49	0.335	0.24	0.31	0.38	0.45
Age between 50 and 64	0.211	0.17	0.18	0.20	0.26
Age more than 65	0.206	0.31	0.24	0.17	0.12
Household head is unemployed	0.064	0.15	0.06	0.01	0.01
Household head is self-employed	0.102	0.09	0.11	0.12	0.09
Household head is an employee	0.834	0.76	0.83	0.87	0.90
Household head is white, non-hispanic	0.776	0.61	0.77	0.84	0.94
Non-white or hispanic	0.224	0.39	0.23	0.16	0.06
Willing to undertake financial risk	0.544	0.20	0.44	0.71	0.96
Not willing to undertake any financial risk	0.456	0.80	0.56	0.29	0.04
Less than high school education	0.186	0.43	0.17	0.08	0.01
High-school education	0.507	0.50	0.61	0.52	0.37
College degree or more	0.307	0.07	0.22	0.40	0.62

		Ranking by Probit Estimates			
Household Characteristic	1998 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.86	0.21	0.00	0.00
Between 25% and 50% of financial assets		0.14	0.52	0.33	0.00
Between 50% and 75% of financial assets		0.01	0.23	0.46	0.26
Upper 25% of financial assets		0.00	0.03	0.21	0.74
Lower 25% of income		0.68	0.31	0.11	0.02
Between 25% and 50% of income		0.23	0.37	0.30	0.11
Between 50% and 75% of income		0.08	0.24	0.35	0.26
Upper 25% of income		0.01	0.08	0.24	0.61
Of age less than 34	0.233	0.29	0.27	0.27	0.14
Age between 35 and 49	0.334	0.26	0.26	0.34	0.45
Age between 50 and 64	0.230	0.17	0.21	0.19	0.29
Age more than 65	0.203	0.29	0.26	0.21	0.13
Household head is unemployed	0.051	0.13	0.05	0.03	0.01
Household head is self-employed	0.256	0.09	0.12	0.11	0.12
Household head is an employee	0.693	0.78	0.83	0.86	0.87
Household head is white, non-hispanic	0.824	0.55	0.77	0.81	0.91
Non-white or hispanic	0.176	0.45	0.23	0.19	0.09
Willing to undertake financial risk	0.659	0.22	0.40	0.73	0.95
Not willing to undertake any financial risk	0.341	0.78	0.60	0.27	0.05
Less than high school education	0.125	0.50	0.27	0.11	0.03
High-school education	0.432	0.43	0.52	0.57	0.37
College degree or more	0.443	0.08	0.21	0.32	0.60

Table 3.b: Demographic composition of quartiles of the population ranked by participation probabilities. Quartiles were constructed using participation probability estimates from our probit runs, along with population weights, for each of the 1995 and 1998 samples.

		Ranking by Logit Estimates			
Household Characteristic	1995 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.81	0.13	0.00	0.00
Between 25% and 50% of financial assets		0.18	0.58	0.25	0.00
Between 50% and 75% of financial assets		0.01	0.25	0.53	0.25
Upper 25% of financial assets		0.00	0.04	0.23	0.75
Lower 25% of income		0.62	0.20	0.07	0.01
Between 25% and 50% of income		0.27	0.40	0.24	0.06
Between 50% and 75% of income		0.10	0.29	0.42	0.25
Upper 25% of income		0.01	0.12	0.27	0.67
Of age less than 34	0.248	0.28	0.28	0.25	0.17
Age between 35 and 49	0.335	0.24	0.31	0.38	0.44
Age between 50 and 64	0.211	0.17	0.18	0.20	0.26
Age more than 65	0.206	0.31	0.24	0.17	0.12
Household head is unemployed	0.064	0.15	0.06	0.01	0.01
Household head is self-employed	0.102	0.09	0.11	0.12	0.09
Household head is an employee	0.834	0.76	0.83	0.87	0.90
Household head is white, non-hispanic	0.776	0.62	0.77	0.83	0.94
Non-white or hispanic	0.224	0.38	0.23	0.17	0.06
Willing to undertake financial risk	0.544	0.20	0.44	0.71	0.96
Not willing to undertake any financial risk	0.456	0.80	0.56	0.29	0.04
Less than high school education	0.186	0.43	0.17	0.08	0.01
High-school education	0.507	0.50	0.61	0.52	0.37
College degree or more	0.307	0.07	0.22	0.41	0.62

		Ranking by Logit Estimates			
Household Characteristic	1998 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Lower 25% of financial assets		0.86	0.21	0.00	0.00
Between 25% and 50% of financial assets		0.14	0.53	0.32	0.00
Between 50% and 75% of financial assets		0.01	0.23	0.47	0.26
Upper 25% of financial assets		0.00	0.03	0.20	0.74
Lower 25% of income		0.68	0.31	0.11	0.02
Between 25% and 50% of income		0.23	0.37	0.30	0.11
Between 50% and 75% of income		0.08	0.24	0.35	0.26
Upper 25% of income		0.01	0.08	0.24	0.61
Of age less than 34	0.233	0.29	0.27	0.26	0.14
Age between 35 and 49	0.334	0.26	0.26	0.34	0.45
Age between 50 and 64	0.230	0.17	0.21	0.19	0.29
Age more than 65	0.203	0.29	0.26	0.21	0.13
Household head is unemployed	0.051	0.14	0.05	0.03	0.01
Household head is self-employed	0.256	0.09	0.12	0.11	0.12
Household head is an employee	0.693	0.77	0.83	0.86	0.87
Household head is white, non-hispanic	0.824	0.55	0.77	0.82	0.91
Non-white or hispanic	0.176	0.45	0.23	0.18	0.09
Willing to undertake financial risk	0.659	0.22	0.40	0.73	0.95
Not willing to undertake any financial risk	0.341	0.78	0.60	0.27	0.05
Less than high school education	0.125	0.49	0.27	0.11	0.03
High-school education	0.432	0.43	0.52	0.57	0.37
College degree or more	0.443	0.08	0.21	0.32	0.60

Table 3.c: Demographic composition of quartiles of the population ranked by participation probabilities. Quartiles were constructed using participation probability estimates from our logit runs, along with population weights, for each of the 1995 and 1998 samples.

		Ranking by Quantile Regression Estimates			
Household Characteristic	1995 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Median Net Worth	60282	8853	67063	85335	260564
Homeownership Indicator	0.65	0.47	0.69	0.76	0.84
Business Equity Ownership Indicator	0.11	0.07	0.09	0.13	0.20
Shops around for investment opportunities	0.30	0.27	0.25	0.33	0.39
Lower 25% of Net Worth		0.57	0.13	0.04	0.02
Between 25% and 50% of Net Worth		0.32	0.35	0.27	0.07
Between 50% and 75% of Net Worth		0.10	0.40	0.42	0.26
Upper 25% of Net Worth		0.01	0.13	0.27	0.65

		Ranking by Quantile Regression Estimates			
Household Characteristic	1998 sample	(0%, 25%)	[25%,50%)	[50%,75%)	[75%,100%)
Median Net Worth	70500	6250	71910	80800	253450
Homeownership Indicator	0.66	0.44	0.67	0.75	0.87
Business Equity Ownership Indicator	0.11	0.06	0.08	0.12	0.17
Shops around for investment opportunities	0.30	0.22	0.25	0.31	0.42
Lower 25% of Net Worth		0.51	0.06	0.02	0.01
Between 25% and 50% of Net Worth		0.34	0.33	0.24	0.06
Between 50% and 75% of Net Worth		0.14	0.43	0.42	0.26
Upper 25% of Net Worth		0.02	0.18	0.32	0.67

Table 3.d: Additional demographic characteristics, not included as covariates in regressions, of quartiles of the population ranked by participation probabilities. Household rankings are the same as in Table 3.a, for each of the 1995 and 1998 samples.

Current Participation Probability	Increases in Financial Assets		
	25th Quantile	Median	75th Quantile
0.5	48513	145589	404161
0.45	32099	102543	212927
0.4	27878	120436	283384
0.35	20189	56812	203589
0.3	7521	33095	117580

Table 4. Increases in financial assets that would move households currently classified as potential stockholders, to the corresponding point in the distribution of marginal stockholders. Statistics are weighted using population weights. Household classification is done using quantile regression estimates. For method of classification and of computation of these wealth increases, see text.

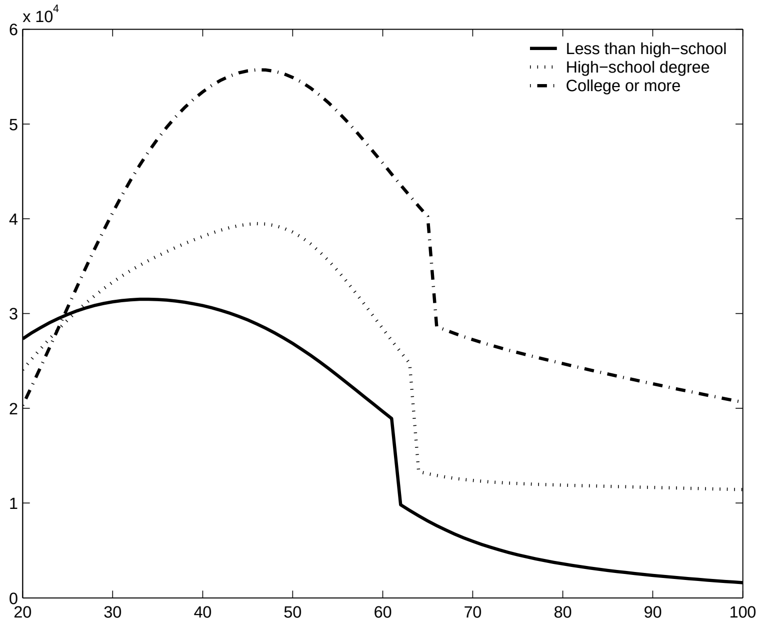

Figure 1: Age income profiles by education category.

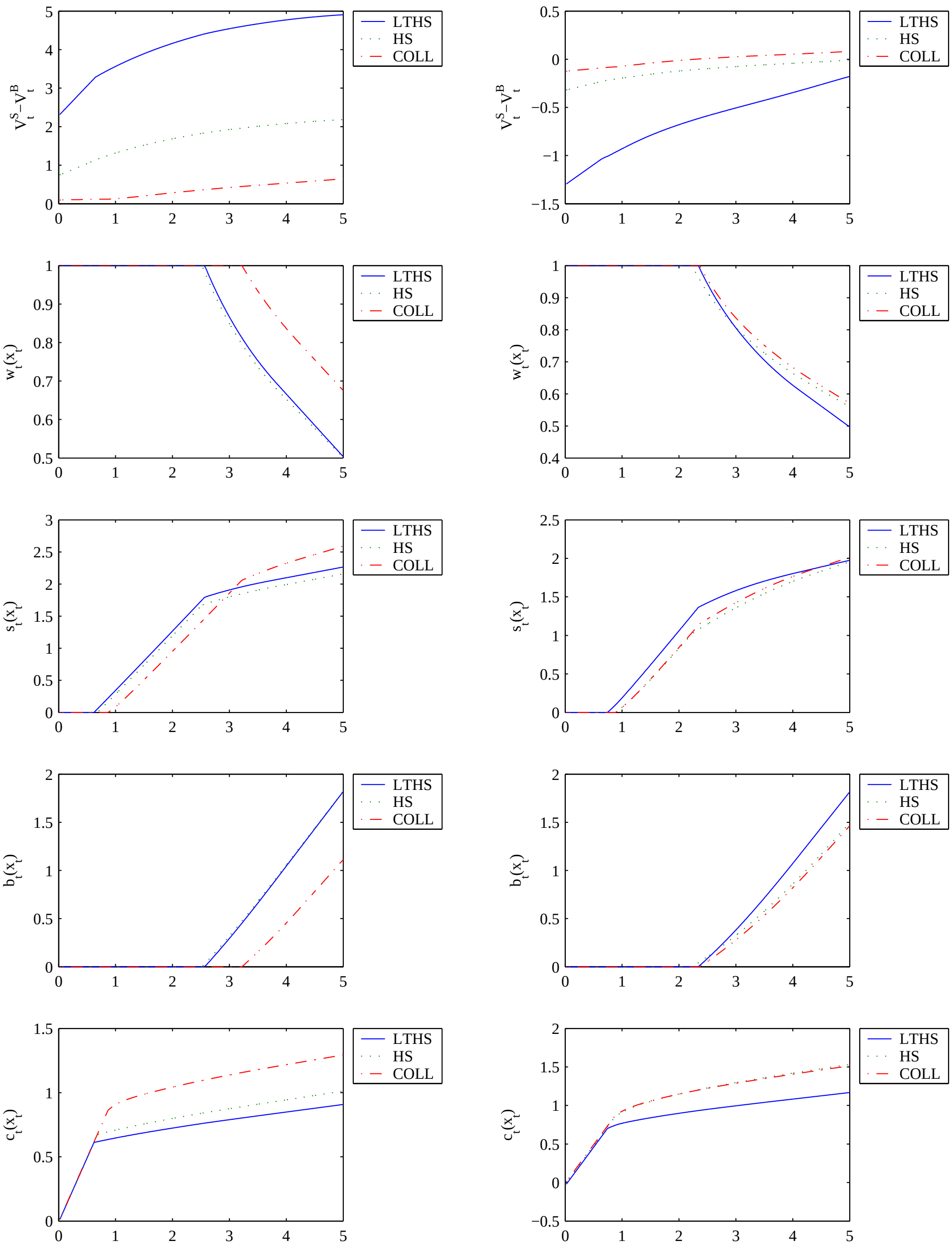

Figure 2: Effects of education on gains from stock market access and on underlying policy rules. Education levels are high-school dropouts (LTHS), high-school graduates (HS), and holders of College degree or more (COLL). Policy rules are for the risky portfolio share, the amount of risky asset holdings, the amount of riskless asset holdings, and consumption. The solution on the left is for (working) age 45, and on the right for (retirement) age 70.

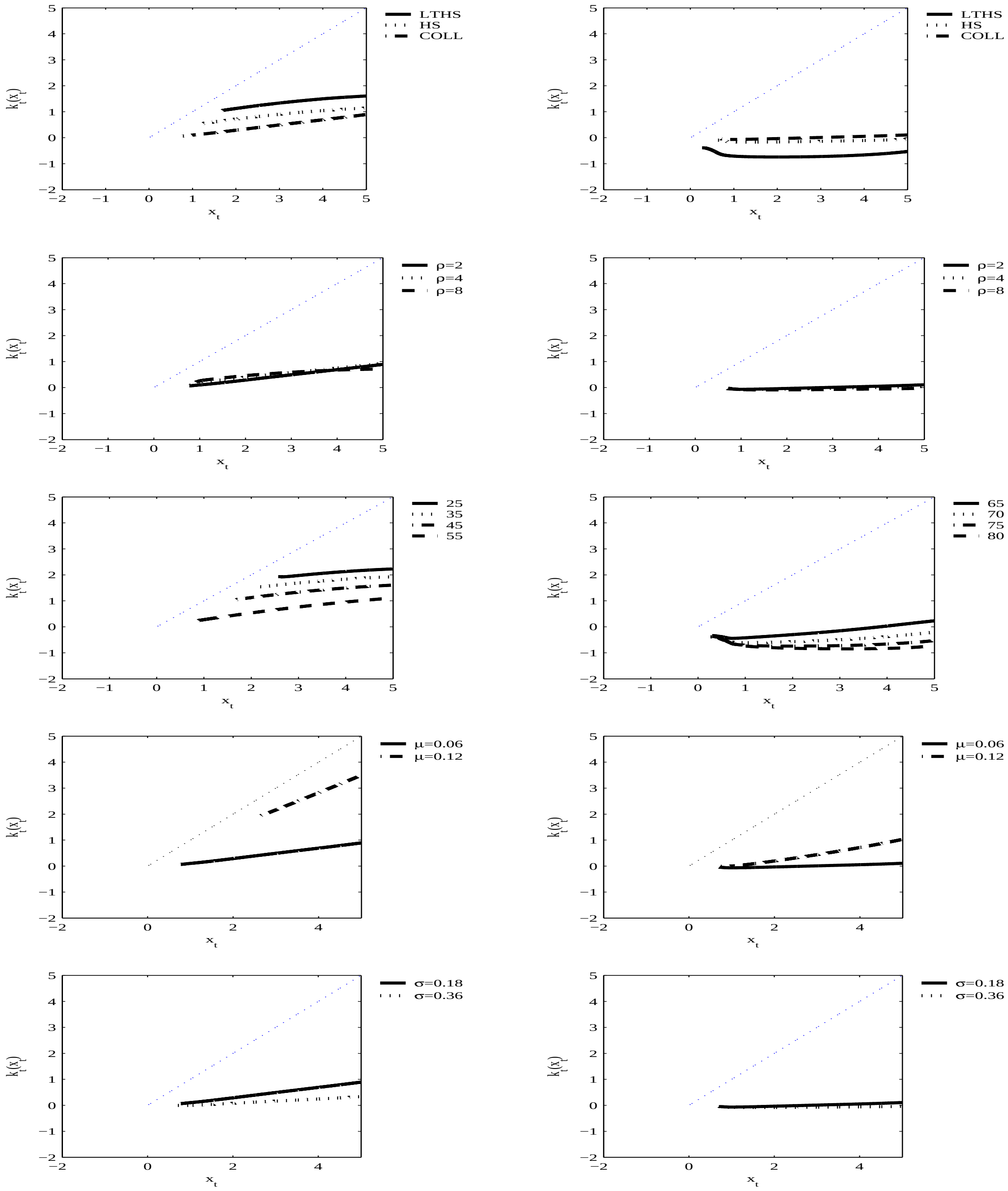

Figure 3: The monetary equivalent of gains from stock market access as a function of normalized cash on hand, *with embedded fixed per-period participation costs*. The effect of education, risk aversion of the felicity function, and age are summarized in the top three rows, with the effect of an equity premium and volatility increase to follow in the fourth and fifth row, respectively.

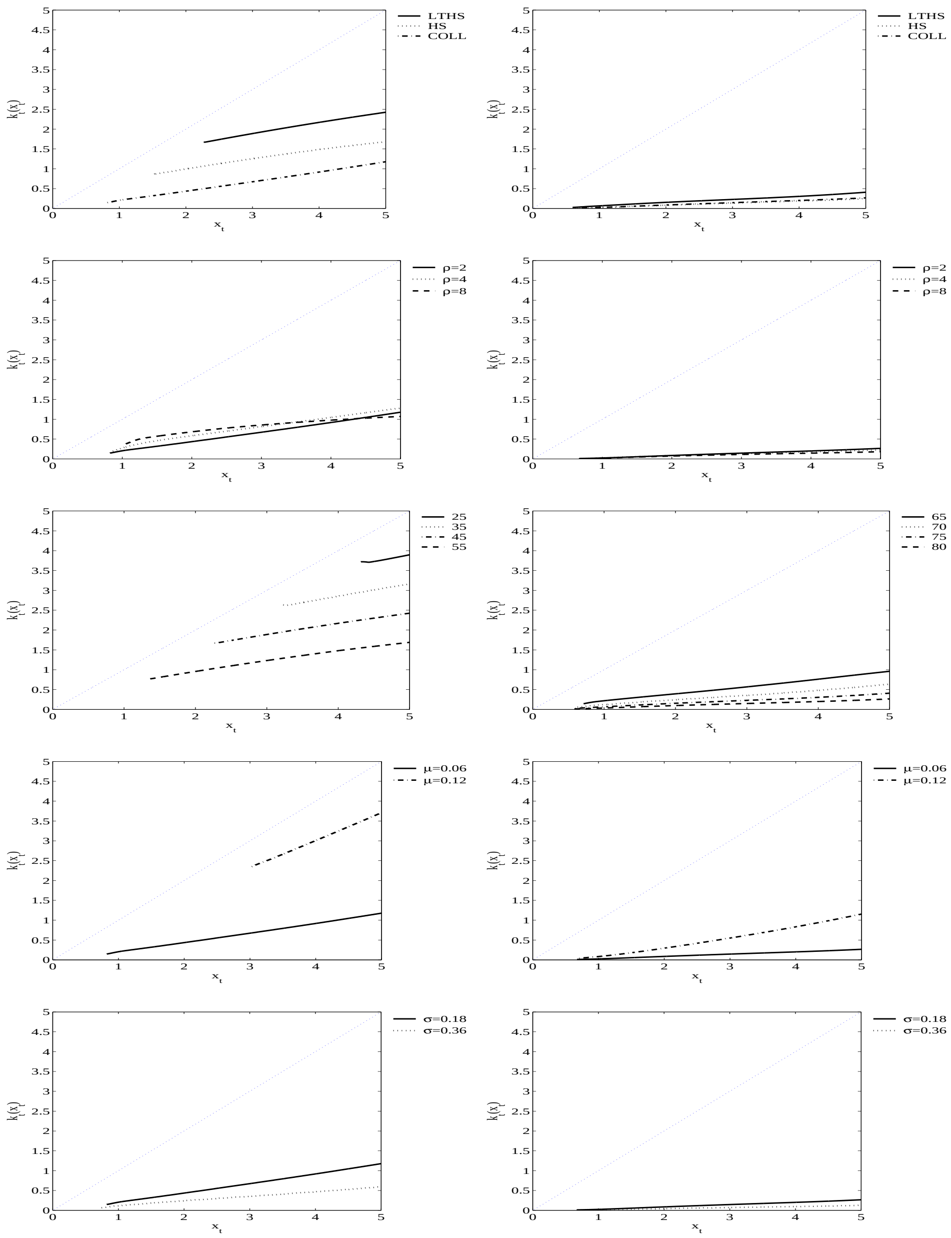

Figure 4: The monetary equivalent of gains from stock market access as a function of normalized cash on hand *without embedding any fixed per-period participation costs.* The effect of education, risk aversion of the felicity function, and age are summarized in the top three rows, with the effect of an equity premium and volatility increase to follow in the fourth and fifth row, respectively.

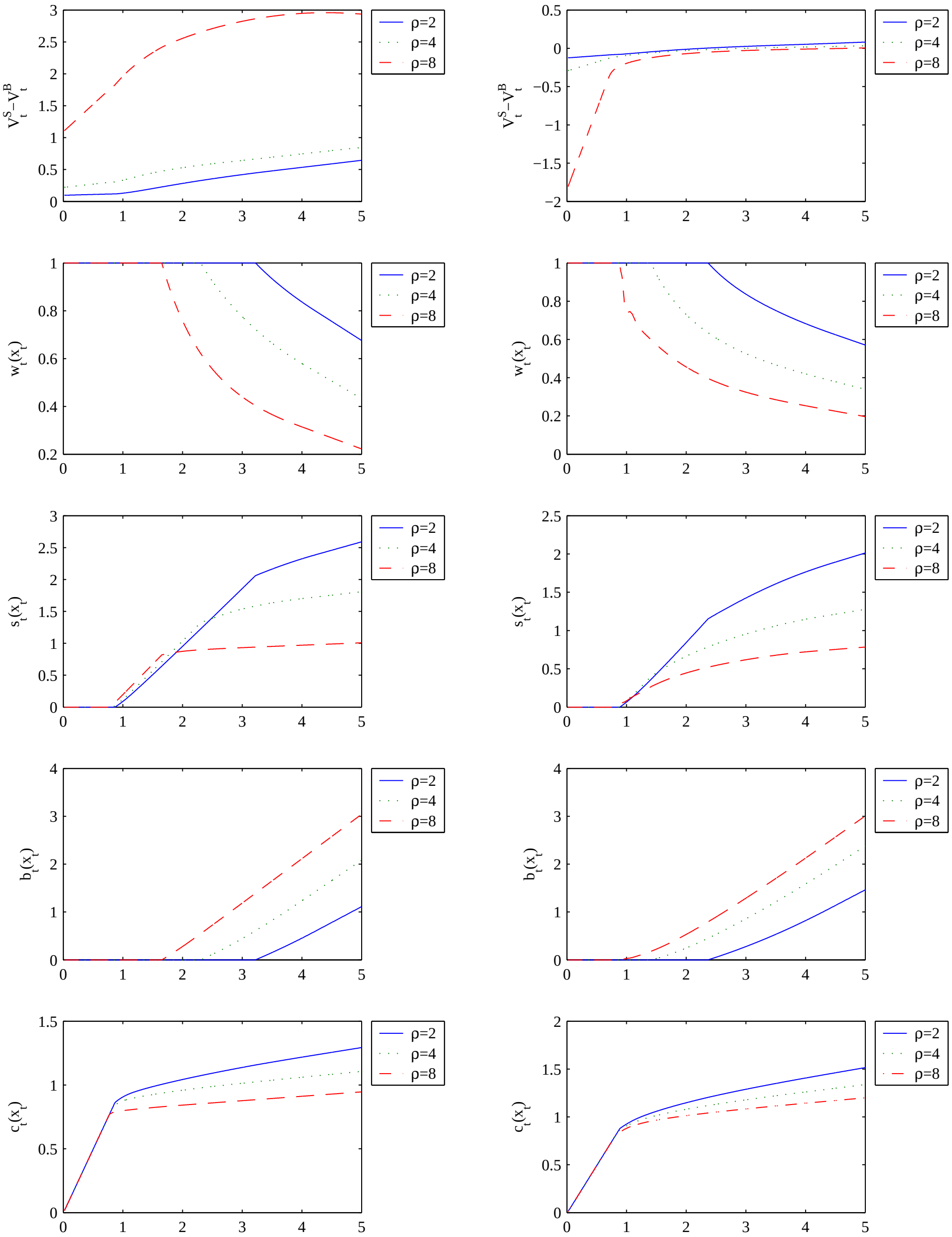

Figure 5: Effects of risk aversion of felicity function on gains from stock market access and on underlying policy rules. Policy rules in order of appearance are for the risky portfolio share, the size of risky asset holdings, the size of riskless asset holdings, and consumption. The solution depicted on the left is for age 45, and on the right for (retirement) age 70. Results are for the education group of College graduates.

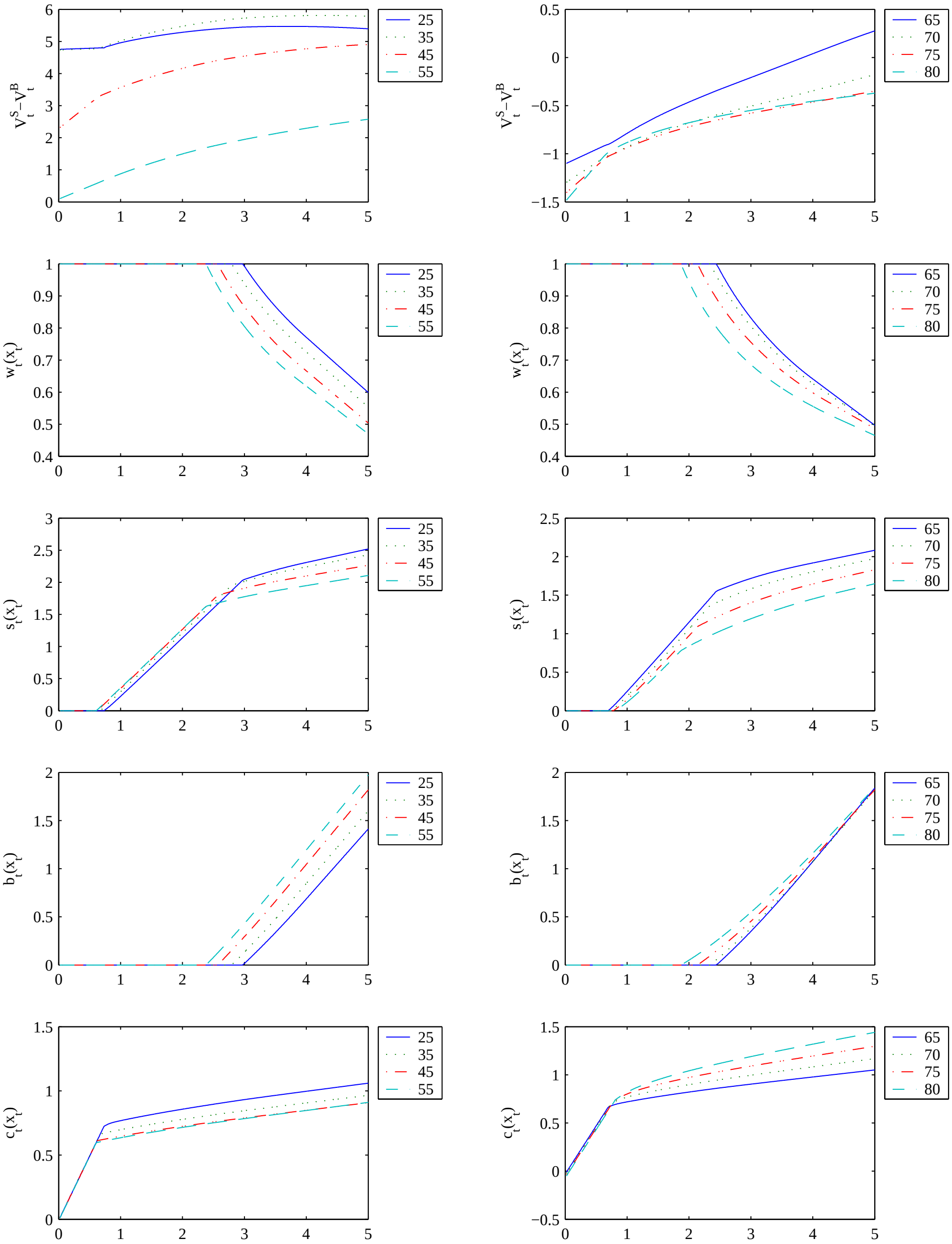

Figure 6: Age effects on gains from stock market access and on underlying policy rules. Policy rules in order of appearance are for the risky portfolio share, the size of risky asset holdings, the size of riskless asset holdings, and consumption. On the left panel we depict ages from working life, whereas on the right panel a selection of retirement ages is shown. Results are for households with less than high-school education.

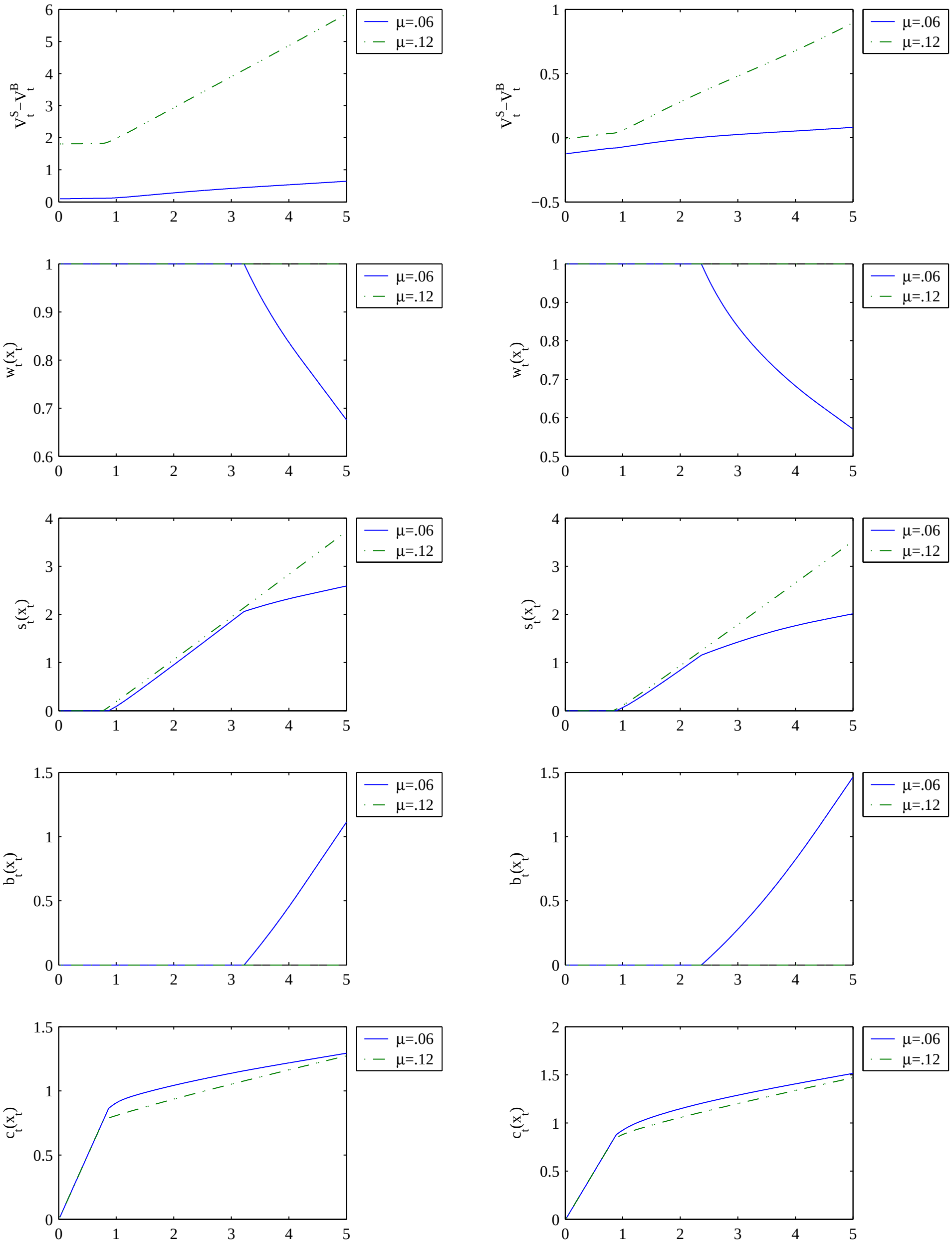

Figure 7: Effects of the size of equity premium on gains from stock market access and on underlying policy rules. Policy rules in order of appearance are for the risky portfolio share, the size of risky asset holdings, the size of riskless asset holdings, and consumption. The solution depicted is for (working) age 45 and (retirement) age 70, on the left and on the right panel, respectively. Results shown are for College graduates.

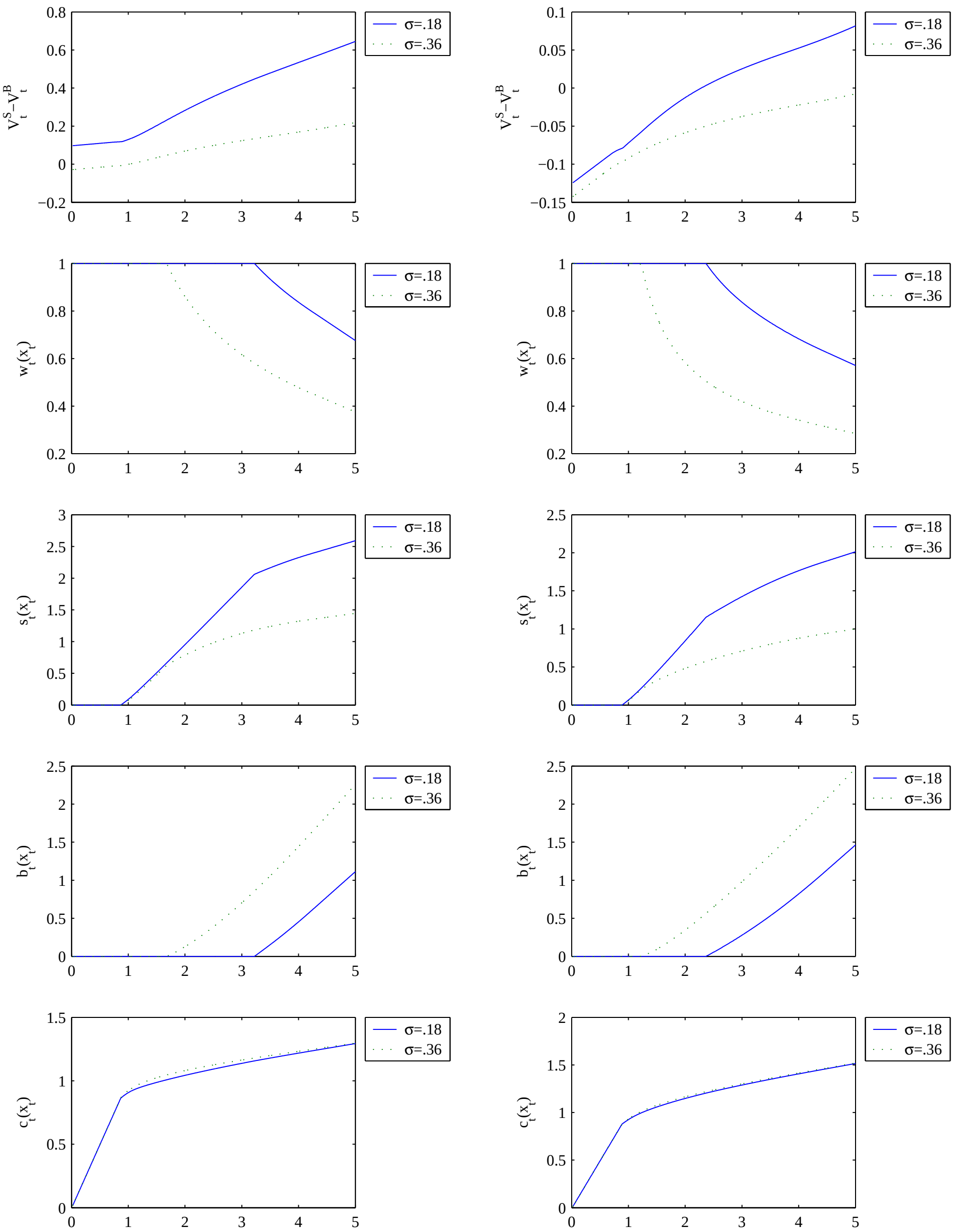

Figure 8: Effects of an increase in volatility of risky asset return on gains from stock market access and on underlying policy rules. Policy rules in order of appearance are for the risky portfolio share, the size of risky asset holdings, the size of riskless asset holdings, and consumption. The solution depicted is for (working) age 45 and (retirement) age 70, on the left and on the right panel, respectively. Results shown are for education group of high-school dropouts.

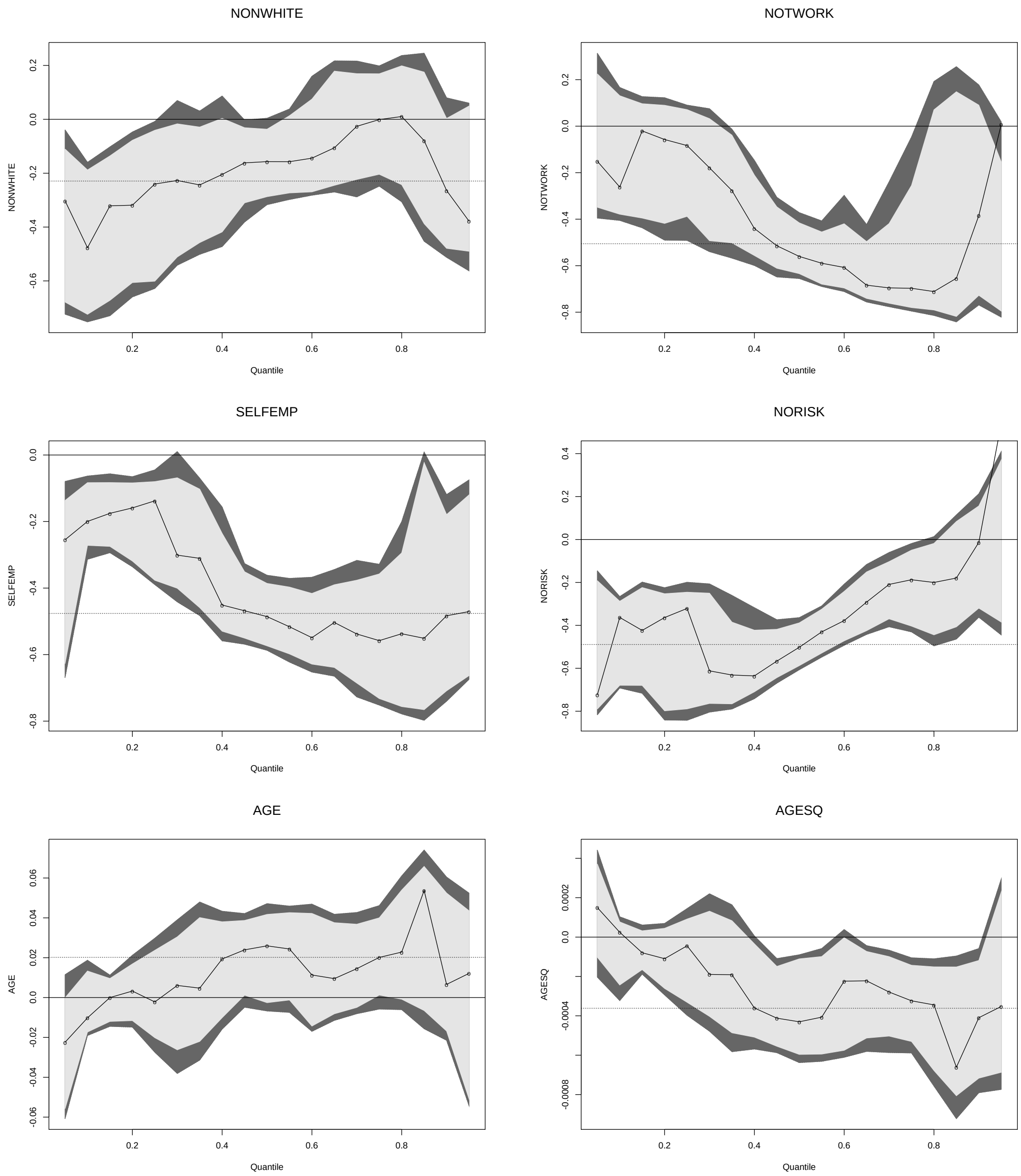

Figure 9a: Smoothed Binary Quantile Regression. The light- and dark-shaded areas represent the 90% and 95% confidence intervals, respectively, and the line portrays the smoothed binary quantile regression coefficient. The zero level is indicated with a continuous line, whereas the normalized Probit coefficient with a dotted line. The Probit and Logit normalized coefficients virtually coincide. (figure continued next page).

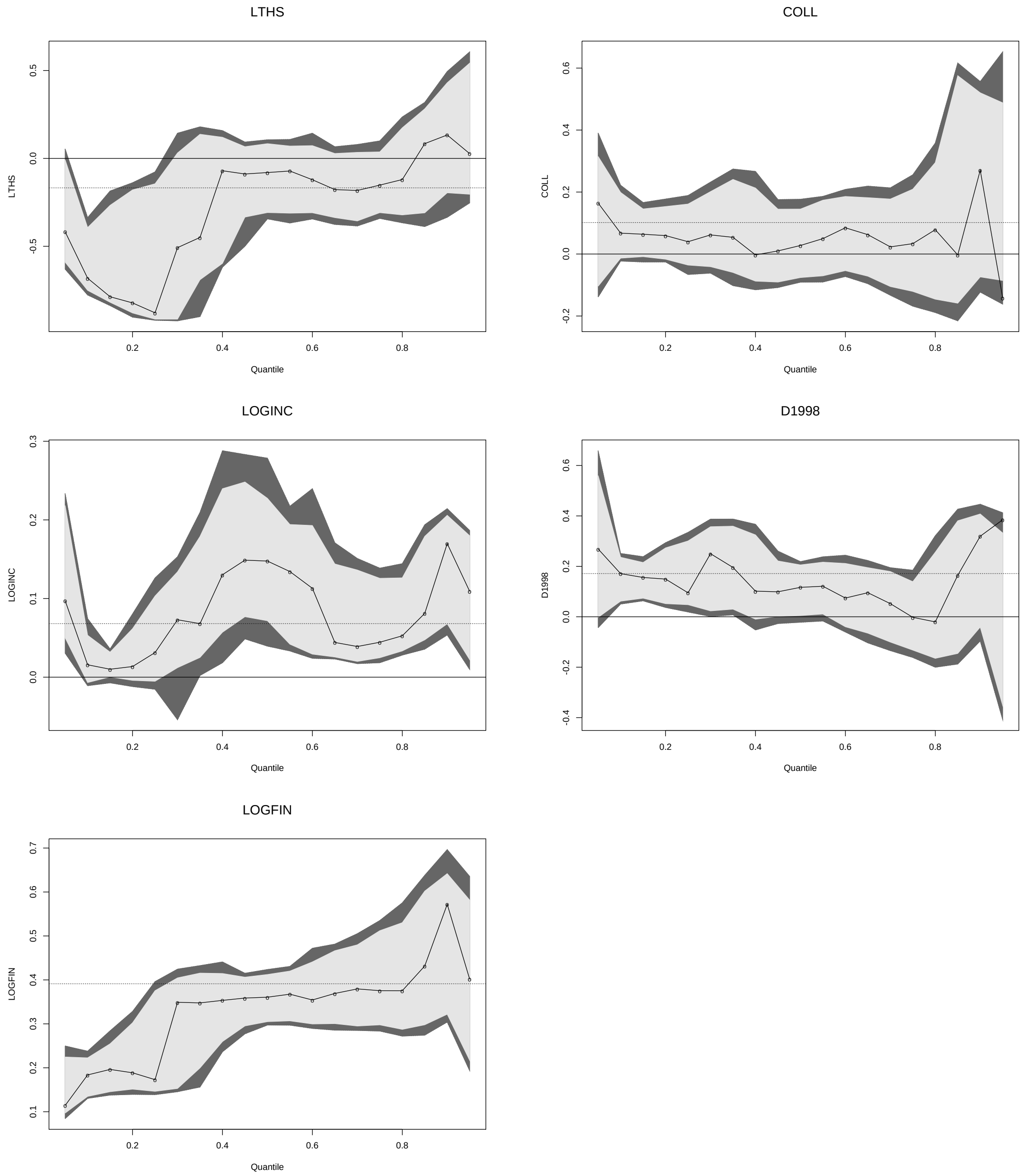

Figure 9b: Smoothed Binary Quantile Regression (cont'd.)

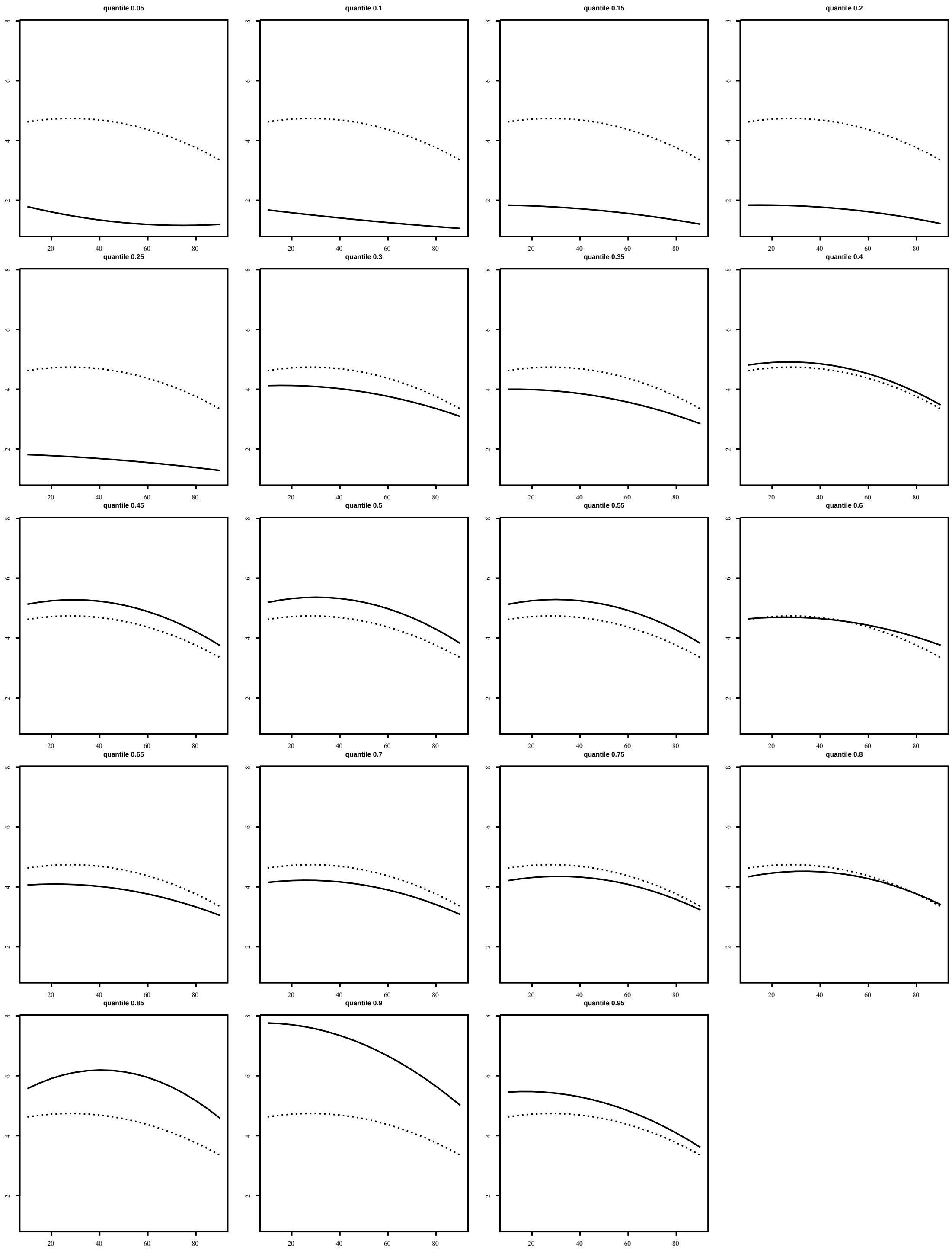

Figure 10: Estimated Effects of Age on Gains from Stockmarket Participation Across Quantiles. The dotted line represents the Probit estimated profile, whereas the continuous line is the Smoothed Binary Quantile Regression estimated profile for that particular quantile.

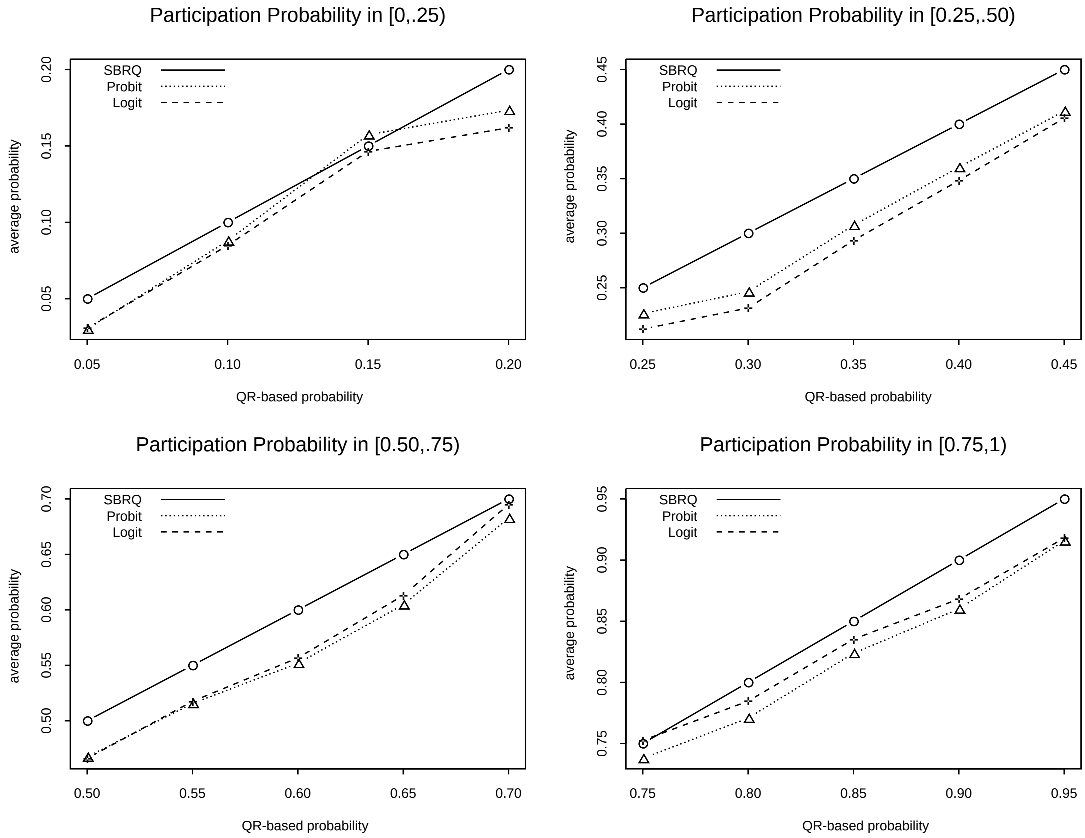

Figure 11a: Comparison of Quantile Regression Participation Probabilities and of the mean Probit and Logit probabilities within groups sharing the same QR participation probability. 1995 Sample. For method of construction, see text.

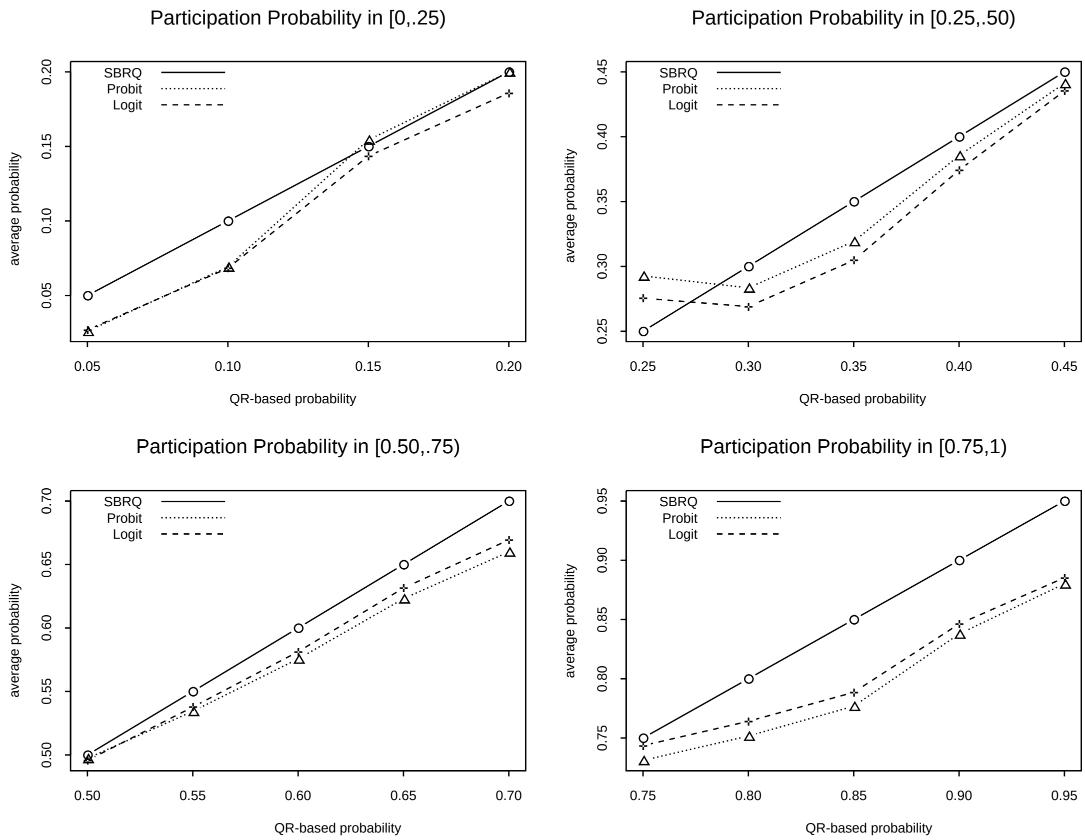

Figure 11b: See Notes to Figure 11a. 1998 Sample.