Stock Market Mean Reversion and Portfolio Choice over the Life Cycle*

Alexander Michaelides†

Imperial College London, University of Cyprus, CEPR and NETSPAR

Yuxin Zhang

Imperial College London‡

January 26, 2015

*This version is preceded by an older working paper entitled "Portfolio Choice, Liquidity Constraints and Stock Market Mean Reversion." We would like to thank, without implicating, Nicholas Barberis, John Campbell, John Cochrane, George Constantinides, Steve Davis, Christian Gollier, Francisco Gomes, Michael Haliassos, James Kahn, Martin Lettau, Sydney Ludvigson, Annette Vissing-Jorgensen, Paul Willen and seminar participants at the Athens University of Economics and Business, the University of Chicago Graduate School of Business, the University of Cyprus, and the New York Federal Reserve Bank for many helpful comments.

†Address: Department of Finance, Imperial College Business School, South Kensington Campus, SW7 2AZ: a.michaelides@imperial.ac.uk.

‡Address: Department of Finance, Imperial College Business School, South Kensington Campus, SW7 2AZ: y.zhang11@imperial.ac.uk.

Stock Market Mean Reversion and Portfolio Choice over the Life Cycle

Abstract

We solve for optimal consumption and portfolio choice in a life-cycle model with short-sales and borrowing constraints, undiversifiable labor income risk and a predictable, time-varying, equity premium. The investor pursues aggressive market timing strategies, and quantitatively substantial hedging demands are found for risk averse investors, despite the presence of frequently binding liquidity constraints. Importantly, in the presence of stock market predictability, the model casts doubt on the conventional financial advice that households should reduce their exposure to the stockmarket as they approach retirement.

JEL Classification: E21, G11.

Key Words: Portfolio Choice over the Life Cycle, Stock Market Mean Reversion, Stock Market Predictability, Hedging Demands, Lifestyle funds.

1 Introduction

How does the presence of stock market predictability, undiversifiable labor income risk and liquidity constraints affect optimal consumption and portfolio choice for a stockholder over the life cycle? Various papers have analyzed the implications of stock market predictability[1] for consumption and/or portfolio choice while ignoring labor income risk; Kim and Omberg (1996), Brennan, Schwartz and Lagnado (1997), Brandt (1999), Campbell and Viceira (1999), Balduzzi and Lynch (1999), Barberis (2000), Campbell et. al. (2002, 2003) and Wachter (2002) show that stock market exposure varies substantially as a response to the predictive factor(s). The effect of background labor income risk on portfolio choice while ignoring stock market predictability has been analyzed numerically by Heaton and Lucas (1996, 1997, 2000), Gomes and Michaelides (2005), Cocco, Gomes and Maenhout (2005), Haliassos and Michaelides (2003) and analytically by Viceira (2001). This paper jointly models stock market predictability and non-diversifiable background labor income risk and analyzes the normative implications for optimal consumption and portfolio choice over the life cycle.

A potentially binding liquidity constraint in both the risky and riskless asset markets is an important component of the current model for a number of reasons. First, in the absence of borrowing restrictions, households with long horizons facing nontradable labor income risk that is only weakly correlated with stock returns would borrow to invest in the stock market, given the equity premium (Viceira, 2001).[2] This theoretical prediction would not only contradict directly the observed zero stockholding puzzle (Mankiw and Zeldes, 1991, and Haliassos and Bertaut, 1995) but would make the equity premium puzzle even harder to resolve as demand for the risky asset would rise relative to a model with borrowing constraints. Second, a recent literature on portfolio selection has stressed the importance of borrowing and short sales constraints in understanding observed portfolio choice patterns. Cocco, Gomes and Maenhout (2005), for instance, solve numerically a model with short sales and borrowing constraints over the life cycle in the presence of undiversifiable labor income risk and show that households should invest a larger proportion of their savings in the stock market when young because the future labor income they will receive (against which they cannot borrow) acts as a risk free asset that crowds out the accumulation of riskless assets[3]. This prediction resembles the advice given by financial planning consultants in recommending lifestyle funds that reduce exposure to stocks as retirement approaches. Finally, the presence of borrowing constraints is an important component of the buffer stock saving model (Deaton (1991) and Carroll (1997)) that has been proposed as the leading alternative to the classic Permanent Income Hypothesis (PIH) or Life Cycle model in an effort to explain the observed "excess smoothness"[4] and "excess sensitivity" puzzles.[5]

The second important component of the model is undiversifiable labor income risk. In related papers, Barberis (2000) and Campbell et. al. (2002) analyze the portfolio choice implications of stock market predictability but in the absence of consumption choice and labor income risk. The buffer stock saving literature (Deaton (1991) and Carroll (1997), for instance) has shown, however, that nontradable labor income risk is an important factor that must be taken into account by households making optimal savings plans. The importance of undiversifiable labor income risk has also been stressed by Viceira (2001), who has shown that higher nontradable labor income risk can affect positively the level of savings through a precautionary savings channel and negatively the share of savings invested in the risky asset through a temperance channel (Gollier and Pratt (1996)). Integrating stock market predictability with labor income risk in a single model can potentially yield further insights on the effects of both labor income risk and predictability on optimal consumption and portfolio choice.

Some recent papers have analyzed models incorporating some form of stock market or labor income predictability in dynamic or life-cycle settings. Benzoni, Collin-Duffresne and Goldstein (2007) investigate the implications of a cointegrating relationship between labor income and stock returns to show how stock demand for young investors can be reduced relative to the absence of this type of long-run risk. Lynch and Tan (2011) generate a similar result by focussing on the implications of time-variation in the mean and, in particular, the variance of labor income. Munk and Sørensen (2010) focus instead on time variation in interest rates and expected income growth to illustrate the effects on portfolio choice, while Koijen, Nijman and Werker (2009) focus on the effects of bond risk premia predictability on optimal life cycle asset allocation. Brennan and Xia (2002) instead focus on the effects of inflation on dynamic asset allocation. We differ from all papers by introducing Epstein-Zin

(1989) and Weil (1990) preferences in an explicit life-cycle setting with a factor predicting stock returns, so that stock market mean reversion exists in the model. Taking the extreme (but instructive) view that there is no uncertainty about the model predicting stock returns, we focus on emphasizing the differences between the i.i.d. stock returns model and the one with stock market predictability. We use the stockholder preference specification in Gomes and Michaelides (2005) that matches some of the stockholder wealth data in the Survey of Consumer Finances (SCF). The i.i.d. stock returns model arises as a special case in our setup and this allows us to quantitatively evaluate hedging demands by comparing the mean reversion with the i.i.d. stock returns model.

We rely on numerical techniques and calibration to draw out the implications of the model for optimal consumption and portfolio choice. Optimal consumption is shown to be a concave function of liquid wealth and has a similar shape to that found in the buffer stock saving literature. Furthermore, for plausible parameters of the coefficient of relative risk aversion, stock market predictability generates a speculative increase in savings when the excess return of stocks over the riskless asset is expected to be high and conversely a decrease in savings when the excess stock return is expected to be low. On the other hand, changing the correlation structure between the different innovations (stock return, factor predicting returns or labor income) does not substantially affect total savings.

With regards to portfolio choice, the consumer/investor is shown to be an aggressive market timer in the presence of stock market predictability. Relative to the i.i.d. returns model, high expected future returns generate a higher allocation of stocks in the portfolio for a given level of saving (when constraints are not binding), while low expected future returns decrease the exposure in the stock market and can cause complete portfolio specialization in the riskless asset. This translates to large variations over the life cycle depending on the realization of the factor rather than the level of financial wealth. This result substantially alters one of the main insights of life-cycle models with i.i.d. stock returns, namely that financial wealth tends to be the main predictor of life-cycle portfolio choice. With stock market predictability, the persistent factor predicting returns can take center stage and outweigh the effect of financial wealth on portfolio choice. Average life-cycle asset allocation profiles hover around 50% of financial wealth, reflecting the oscillation between the extreme bounds of the liquidity constraints (zero and full investments in the stock market). Aggressive market timing behavior is similar to the behavior predicted in Brennan, Schwartz and Lagnado (1997) and Barberis (2000), models that do not feature undiversifiable labor income uncertainty.

Hedging demands are evaluated by comparing average life cycle asset allocation choices in the mean reversion model relative to the i.i.d. stock returns model. Interestingly, when there is no negative correlation between the innovation in the predictability factor and the stock return innovation, the average portfolio shares become almost identical to the i.i.d. model. This is an interesting finding because there is a substantial number of factors used to predict stock returns, yet not all of them have this negative correlation in common. In contrast, in the baseline factor model that features a correlation equal to -0.8, there is a substantial divergence of average life cycle portfolios between the i.i.d. and mean reversion model.

The positive correlation between permanent earnings shocks and the stock market innovation also generates substantial hedging demands when that correlation matters, namely in the working part of the life cycle. This occurs when that correlation is increased from 0.15 to 0.5, an increase that is not unrealistic given the recent empirical evidence by Bonaparte, Korniotis and Kumar (2014) who find that the distribution of this correlation can vary up to 0.6 for certain households. Changing the correlation between the permanent labor income shock and the factor innovation (from zero to 0.15) also generates a substantial change in average portfolio allocations over the lifecycle: these changes are quite substantial given the focus on computing average changes over the life cycle.

What does the model imply about lifestyle funds? The average life cycle asset allocation profile for the mean reversion model is around 0.5 and does not feature any resemblance to lifestyle funds. This average profile masks substantial underlying heterogeneity, however, with investors almost always found between a full or a zero asset allocation in stocks based on the realization of the factor. Given the strong persistence of the factor, initial realizations of the factor (even in the early part of the life cycle) can determine allocations for many periods into the future. This can generate average portfolio shares that might be rising over the life cycle, contrary to popular financial advice (and what the i.i.d. model predicts) that households should reduce their exposure to the stock market as retirement approaches. This finding casts some doubt on the idea that lifestyle funds are always and everywhere optimal for all types of investors, all investor expectations and all market conditions.

The paper is organized as follows. Section 2 describes the theoretical model, outlines the numerical solution algorithm and discusses the parameter choices for the calibration. Section 3 discusses the optimal consumption-saving policy functions and life-cycle simulation profiles and presents results when varying the elasticity of intetemporal substitution and the risk aversion coefficient. Section 4 discusses the effects of stock market mean reversion by comparing the benchmark results to the i.i.d. stock returns model. Section 5 discusses hedging demands and how different correlation changes also affect wealth accumulation. Section 6 discusses the implications of the model for lifestyle funds and Section 7 concludes.

2 The Model

Time is discrete, there is one non-durable good, one riskless financial asset and a risky time varying investment opportunity. The riskless asset yields a constant gross after tax real return, R_f, while the gross real return on the risky asset is denoted by $\widetilde{R}$. At time t, the agent enters the period with invested wealth in the stock market S_{t-1} and the bond market B_{t-1} and receives Y_t units of the non-durable good. Following Deaton (1991), cash on hand in period t is denoted by $X_t = S_{t-1}\widetilde{R_t} + B_{t-1}R_f + Y_t$. The investor then chooses savings in the bond (B_t) and stock (S_t) market to maximize welfare. The particular assumptions made about the economic environment are as follows:

2.1 Preferences

Preferences separate the elasticity of intertemporal substitution from risk aversion as in Epstein and Zin (1989) and Weil (1990). Specifically, they are given by

$$V_t = \left\{(1-\beta)C_t^{1-1/\psi} + \beta\left(E_t(p_{t+1}V_{t+1}^{1-\gamma} + b(1-p_{t+1})X_{t+1}^{1-\gamma})\right)^{\frac{1-1/\psi}{1-\gamma}}\right\}^{\frac{1}{1-1/\psi}}$$

where β is the time discount factor, b is the strength of the bequest motive, ψ is the elasticity of intertemporal substitution (EIS) and γ is the coefficient of relative risk aversion. The conditional probability of surviving next period conditional on having survived until period t is given by p_{t+1}.

2.2 Labor Income Process

Following a relatively standard specification in the literature (as used by Cocco, Gomes and Maenhout (2005), for example), the labor income process before retirement is given by

$$Y_{it} = Y_{it}^{p} U_{it} \tag{1}$$

$$Y_{it}^{p} = \exp(f(t, Z_{it})) Y_{it-1}^{p} N_{it} \tag{2}$$

where $f(t, Z_{it})$ is a deterministic function of age and household characteristics Z_{it}, Y_{it}^{p} is a permanent component with innovation N_{it}, and U_{it} a transitory component of labor income, where $\ln U_{it}$ and $\ln N_{it}$ are independent and identically distributed with mean $\{-.5 * \sigma_u^2, -.5 * \sigma_n^2\}$, and variances σ_u^2 and σ_n^2, respectively. The log of Y_{it}^{p} evolves as a random walk with a deterministic drift, $f(t, Z_{it})$. For simplicity, retirement is assumed to be exogenous and deterministic, with all households retiring in time period K, corresponding to age 65 ($K = 46$). Earnings in retirement ($t > K$) are given by $Y_{it} = \lambda Y_{iK}^{p}$, where λ is the replacement ratio ($\lambda = 0.68$).

Durable goods, and in particular housing, can provide an incentive for higher spending early in life. We exogenously subtract a fraction of labor income every year allocated to durables (housing). This empirical process is taken from Gomes and Michaelides (2005) and is based on Panel Study Income Dynamics (PSID) data.

2.3 Liquidity Constraints

Borrowing and short sales of stocks are not allowed. Specifically, $B_t \geqq 0$ and $S_t \geqq 0$ as has been assumed in the recent life cycle literature to avoid the counterfactual implication that households in the model lever up to invest in the stock market.

2.4 Mean Reversion

We follow Campbell and Viceira (1999) in assuming that there is a single factor that can predict future excess returns. Letting $\{r_f, r_t\}$ denote the net risk free rate and the net stock market return respectively and f_t the factor that predicts future excess returns, we have

$$r_{t+1} - r_f = f_t + z_{t+1} \tag{3}$$

$$f_{t+1} = \mu + \phi(f_t - \mu) + \varepsilon_{t+1} \tag{4}$$

where the two innovations $\{z_{t+1}, \varepsilon_{t+1}\}$ are i.i.d. Normal random variables with mean equal to zero and variances σ_z^2 and σ_ε^2, respectively. Contemporaneous correlation between these innovations is allowed, while correlation between the permanent earnings innovation ($\ln N_t$) and $\{z_t, \varepsilon_t\}$ can also exist. Mean reversion in the stock market is captured by the autoregressive nature of the factor (f_t) predicting stock market returns ($\phi > 0$). Negative correlation between the excess stock market return innovation (z_{t+1}) and the innovation to the factor (ε_{t+1}) is documented by Campbell and Viceira (1999). One of the key contributions of the paper is to understand how changing these correlations affects saving and portfolio choice decisions over the life cycle.

We will also be reporting results from a model with i.i.d. excess returns; in that case $r_{t+1} - r_f = \mu + z_{t+1}$. In order for the i.i.d. model to be comparable to the factor model, the first two unconditional moments of returns are set to be equal in both cases. This specification of the model is the one found in recent papers with either CRRA preferences (Cocco, Gomes and Maenhout (2005)) or Epstein-Zin-Weil preferences (Gomes and Michaelides (2005) or Cooper and Zhu (2014)).

2.5 Normalized Value Function

The unit root process for labor income is convenient because it allows the normalization of the problem by the permanent component of labor income (Y_{it}^p). Letting lower case letters denote variables normalized by the permanent component of labor income (Y_{it}^p), the evolution of the single endogenous state variable is then given by

$$x_{it+1} = \frac{Y_{it}^p}{Y_{it+1}^p}(r_{t+1}s_{it} + r_f b_{it}) + U_{it+1} \tag{5}$$

Letting $v_{it}^j \equiv V_{it}/Y_{it}^p$ be the normalized value of individual i at age t, the representation of consumer preferences in terms of stationary (normalized) units is then given by:

$$v_{it}(x_{it}, f_t) = \Bigg[(1-\beta)c_{it}^{1-1/\psi} + \beta\left\{E_t\left[\begin{array}{c} p_{t+1}(v_{it+1}(x_{it+1}, f_{t+1}))^{1-\gamma}(Y_{it+1}^p/Y_{it}^p)^{1-\gamma}+ \\ b(1-p_{t+1})(x_{it+1})^{1-\gamma}(Y_{it+1}^p/Y_{it}^p)^{1-\gamma} \end{array}\right]\right\}^{\frac{1-1/\psi}{1-\gamma}}\Bigg]^{\frac{1}{1-1/\psi}}$$

The continuous state is x_{it} (normalized cash on hand) and its evolution is given by (5).

When $f_t = \mu$ and $\sigma_\varepsilon^2 = 0$ this simplifies to the life cycle model in Cocco, Gomes and Maenhout (2005) for CRRA preferences and Gomes and Michaelides (2005) for Epstein-Zin-Weil preferences. In the absence of a risky investment alternative we have the Deaton (1991) and Carroll (1997) specifications for power utility. Appendix A details the numerical solution technique and the numerical accuracy in the implementation of the Tauchen and Hussey (1991) approximation procedure for a vector autoregression. Numerically, this proves to be a substantial challenge because of the strong persistence in the factor f_t that requires a substantial number of grid points to retrieve the posited parameters with the desired accuracy. The appendix provides a detailed discussion of the choices made to satisfy a reasonable computational speed-accuracy tradeoff without having to resort to a supercomputer for the solution.

2.6 Parameter Choice

The model is solved for a set of baseline parameters at an annual frequency. The net constant real interest rate, r_f, equals 0.02. Carroll (1997) estimates the variances of the idiosyncratic shocks using data from the *Panel Study of Income Dynamics*, and the benchmark simulations use values close to those: 0.1 percent per year for σ_u and 0.1 percent per year for σ_n[6]. The deterministic component of labor income is identical to the one used by most life cycle papers in this literature (Cocco, Gomes and Maenhout (2005)), and this also facilitates comparisons between this model and its counterpart with i.i.d. stock returns.

The baseline preference specification is taken to capture the observed behavior of stockholders. Gomes and Michaelides (2005) argue that this is well achieved when using a discount factor (β) equal to 0.95, a coefficient of relative risk aversion (γ) equal to 5, and an elasticity of intertemporal substitution (ψ) equal to 0.5. These choices are consistent with the empirical estimates for the elasticity of intertemporal substitution in Vissing-Jorgensen (2002) and the empirical preference parameter estimates in Gomes, Michaelides and Polkovnichenko (2009). The bequest parameter is set to 2.5 to capture the empirical observation that few rich stockholders die with zero financial assets. To understand the implications of the model, we then present results by changing the preference parameters sequentially to $\gamma = \{2, 8\}$ and $\psi = \{0.2, 0.8\}$.

The parameters describing the evolution of stock market returns are selected mostly from Campbell and Viceira (1999). We use numbers roughly corresponding to their estimates but it should be noted that no estimate of the correlation between the innovation in the factor predicting stock returns and permanent, idiosyncratic earnings shocks ($\rho_{n,\varepsilon}$) exists in the literature. Moreover, given that there are three correlations to be calibrated ($\rho_{z,\varepsilon}, \rho_{n,\varepsilon}$ and $\rho_{n,z}$), there is a constraint that needs to be satisfied by these correlations so that the variance-covariance matrix of these innovations is positive definite. Campbell and Viceira (1999) estimate $\rho_{z,\varepsilon}$ to equal -0.91, while Davis, Kubler and Willen (2006) use previous estimates of $\rho_{n,\varepsilon}$ that can vary between -0.2 and 0.3 over different occupation and education groups. Angerer and Lam (2009) note that the transitory correlation between stock returns and labor income shocks does not empirically affect portfolios and this is consistent with simulation results in life cycle models (Cocco, Gomes and Maenhout, 2005, for instance). We therefore set the correlation between transitory labor income shocks and stock returns equal to zero. The baseline correlation between permanent labor income shocks and stock returns is set equal to 0.15, consistent with the mean estimates in most empirical work (Campbell et. al. (2002) for example), but this can vary and be higher across heterogeneous occupations

(Angerer and Lam (2009)) and workers (Bonaparte, Korniotis and Kumar (2014)). In order to implement a range of comparative statics exercises, we set $\rho_{z,\varepsilon}$ to -0.8 so that $\rho_{n,\varepsilon}$ and $\rho_{n,z}$ can be varied between zero and 0.5 while respecting the constraint posed by the positive definiteness of the variance covariance matrix. The unconditional stock market volatility is given by the unconditional standard deviation of stock returns and is set equal to 0.18. The benchmark parameters for the generation of stock market returns are, therefore, $\mu = .04$, $\phi = .91$, $\sigma_z^2 = 0.18^2 - \sigma_f^2$, $\sigma_\varepsilon^2 = .000034$, $\rho_{z,\varepsilon} = -0.8$, $\rho_{n,\varepsilon} = 0.0$, $\rho_{n,z} = 0.15$. Hedging demands will be evaluated by varying the correlations and also by comparing this model to the special case of the i.i.d. model ($\phi = 0.0$ and $\sigma_\varepsilon = 0.0$).

3 Optimal Consumption and Portfolio Choice

3.1 Consumption and Portfolio Choices in Baseline Model

The consumption policy functions for ages 25, 55 and 75 are plotted in figure 1 (top row). A few observations can be made about the shape of the policy functions. First, the consumption policy rule has the familiar shape from the buffer stock saving literature without risky asset choice (Deaton (1991) and Carroll (1997)); below a cutoff point x^* no saving takes place, while the marginal propensity to consume falls quickly beyond x^*. Second, a low current factor realization signifying higher future stock returns induces an increase in saving to take advantage of more favorable future investment opportunities, while a very high factor realization makes saving less desirable and induces an increase in consumption over the relevant parts of the cash on hand state space (below around 5 units of normalized cash on hand).

The portfolio policy functions for ages 25, 55 and 75 are plotted in figure 1 (bottom row). For the higher factor realizations that predict a one-period ahead stock market return that is lower than the risk free rate, the policy rule does not change; the investor wishes to short the stock market, the short selling constraint becomes binding and all saving is allocated in the riskless asset market. For intermediate factor realizations (like the median factor plotted in the Figure), there is a co-existence between bonds and stocks in the financial portfolios. For the lower factors, the investor wants to borrow to invest in the stock market. Given that no borrowing is allowed, all savings is allocated in the stock market and the share of wealth in stocks is one hundred percent, the maximum possible limit.

The optimal portfolio choice policy rule is of the yo-yo type for most factor realizations. Specifically, either the investor is fully invested in the stock market or does not participate at all in the stock market. This type of policy rule has an interesting implication; for the factors that generate positive stock holdings, the share of wealth invested in the stock market very often stays the same as in the i.i.d. model due to the presence of the binding borrowing constraint. Moreover, the factor can determine an exit from the stock market regardless of the level of financial wealth, indicating a break from the i.i.d. returns model that provides a tight link between wealth levels and portfolio choices.

Figure 2 plots the results from simulating the model starting with an initial wealth distribution for stock market participants from the SCF in 2001. The model is simulated 200 times for 500 individual life histories and the results report the averages from these simulations. Panel A is the standard life cycle wealth accumulation profiles in the presence of a bequest motive. Panel B plots the average share of wealth in stocks over the life cycle and illustrates that the model generates an average profile that is substantially away from one, and at the same time only mildly generates a movement away from stocks as the household approaches retirement.

The fact that the average share of wealth in stocks is never close to one might be surprising given the results in the i.i.d. version of the model where the share of wealth in stocks is close to, or equal to, one. This arises here because we are simulating based on different initial realizations of the factor and then averaging over them many different times. For most of these factors, as the policy functions have illustrated, the investor either invests zero or 100% of their financial wealth in stocks. Moreover, these factors are persistent, and therefore the share of wealth in stocks remains at these levels for substantial parts of the life cycle. Averaging over these experiences generates the average share of wealth in stocks depicted in Figure 2, Panel B.

3.2 Consumption and Portfolio Choices for different preference specifications

There are two main preference parameters that are important in this setup: the risk aversion coefficient (γ) and the elasticity of intertemporal substitution (ψ). The two columns in Figure 3 plot the consumption policy functions for ages 25, 55 and 75 evaluated at the median factor. The first column keeps risk aversion (γ) constant and varies the elasticity of intertemporal substitution (ψ) from 0.2 to 0.5 (baseline) and to 0.8. The second column does the reverse by keeping the elasticity of intertemporal substitution constant at ψ and increasing the risk aversion coefficient from $\gamma = 2$ to $\gamma = 5$ (baseline) and to $\gamma = 8$. For a given risk aversion coefficient, Figure 3 shows that a higher elasticity of intertemporal substitution increases saving. This is as expected since the response of saving to the intertemporal elasticity of substitution depends on the difference between the (endogenous) expected return on the stock market and the discount rate. For the median factor this difference is positive and therefore generates higher saving from higher intertemporal substitution (as noted in Campbell and Viceira (1999) and Gomes and Michaelides (2005)). The precautionary saving effect from a higher risk aversion on the right column is also well understood and as expected.

Figure 4 depicts the counterpart of figure 3 but for the share of wealth in stocks invested in the stock market and evaluated at the median factor for ages 25, 55 an 75. The main lessons from these graphs is that the risk aversion has a larger effect on portfolio allocations than the elasticity of intertemporal substitution.

Figure 5 performs the simulations associated with these policy functions. As anticipated from the policy function discussion, a higher elasticity of intertemporal substitution, for a given risk aversion coefficient, increases wealth accumulation. As a result, there is a lower average allocation in the stock market (figure 5 panel C) due to the standard intuition of the lower implicit value of human capital in this model when wealth accumulation is higher. This is not completely monotonic here, however, since the averaging over different factors might generate some non-monotonicities (as in the early part of the life cycle). Changing risk aversion also affects wealth accumulation in expected ways and generates safer portfolios, as expected from the policy function discussion.

4 The Effect of Stock Market Mean Reversion

How does the presence of a factor predicting returns affect saving and portfolio choice behavior relative to the i.i.d. model? This is one fundamental question that needs to be addressed in the context of this setup.

The saving effects are not substantial even though there is a slight increase in saving for the low factors that predict high stock returns for a substantial range of cash on hand. The differences between the i.i.d. model and the lowest factor are not substantial and therefore we focus our attention on the share of wealth in stocks that is more substantially affected by the factor. The policy functions for the share of wealth in stocks illustrate the large dependence of portfolios on the factor (Figure 6 for ages 25, 55 and 75 (Panels A, B and C, respectively). Portfolios can shift from zero to one and vice versa depending on factor realizations, something that does not happen in the i.i.d. model which has a more stable share of wealth in stocks. The median factor tends to generate a more balanced allocation between bonds and stocks and is closer to the i.i.d. case. The policy function illustrates clearly that portfolio allocations will be a lot more volatile in the mean reversion than in the i.i.d. model.

Figure 7 illustrates clearly the main differences between the i.i.d. and mean reversion model. Wealth accumulation is higher in the i.i.d. model (Figure 7, Panel A) because the share of wealth invested in the stock market is higher almost throughout the life cycle (Figure 7, Panel B). In the early parts of the life cycle in the mean reversion model, due to the presence of the factor that can shift the share of wealth in stocks between zero and one, on average the share of wealth in stocks is around one half, whereas in the i.i.d. model the share of wealth in stocks is 100%. Given the equity premium, the expected porftolio return in the i.i.d. model is higher and therefore total wealth accumulation is higher in the i.i.d. model than in the model with predictability.

The effect of the factor on portfolios can be more clearly seen by tracking certain simulations starting from different initial factor realizations. Figure 8 depicts the pictures for the lowest (ft=1), sixth, tenth and highest (ft=15) factor over the life cycle. For the lowest factor that predicts higher returns in the future the share of wealth in stocks starts at 100%. For the lowest factor it starts at zero investments in the stock market. Because the factor is persistent, it takes a substantial amount of time for a change to happen: when it does the portfolio moves very quickly from one to zero and vice versa. This illustrates the bang-bang movement in the share of wealth in stocks, reminiscent of the allocations in Brennan, Schwartz and Lagnado (1997).

5 Hedging Demands

How do these results change when the correlations between the different innovations vary? Changing correlations generate naturally hedging demands and the model is able to quantitatively assess the magnitude of these demands.

5.1 Variation in Correlations

Negative correlation between the stock market innovation and the factor innovation gives rise to a type of hedging demand arising from a deterioration of future investment opportunities when current stock market returns are high (Merton, 1973). This hedging demand differs from market timing since the former arises as protection from unfavorable shifts in the investment opportunity set, reflecting an attempt to minimize (unanticipated) consumption variability. On the other hand, market timing demand arises from the desire to take advantage of current information regarding future returns. To investigate the importance of hedging demand due to $\rho_{z,\varepsilon}$, the correlation between the factor and the stock market innovation $\rho_{z,\varepsilon}$ is set equal to zero. It is perhaps useful to point out that even though a lot of forecasting variables that have stock market prices in the denominator generate a strongly negative magnitude for this correlation empirically (justifying the calibration of this parameter to equal -0.8 in the benchmark case), other non-price based forecasting variables need not generate such a prediction. For instance, using the consumption-asset-income variable constructed by Lettau and Ludvigson (2001), one finds that this correlation is insignificant from zero. This correlation might change, therefore, depending on the investor's preferred model. Setting it to zero is an extreme change that will be useful in assessing the possible range of hedging demands that might be generated from varying this parameter.

We also evaluate hedging demands when changing the correlation between permanent earnings shocks and stock market innovations ($\rho_{z,n}$) and the correlation between the factor innovation and the permanent labor income shock ($\rho_{n,\varepsilon}$). For the former, there are some empirical estimates offered by Davis, Kubler and Willen (2006) and more recently by Bonaparte, Korniotis and Kumar (2014) who find that this correlation can vary for different households from -0.6 to 0.6 and therefore can have a substantial effect on portfolio decisions depending on its value. Bagliano, Fugazza and Nicodano (2014) also emphasize this correlation in combination with uncertainty about retirement pensions replacement rates which can have a similar effect across heterogeneous workers, while Angerer and Lam (2009) estimate correlations that can be higher than 0.15 among different occupations. In our baseline model we use 0.15 for this correlation, a value that reflects the substantial idiosyncratic risk that exists in labor income data. Nevertheless, one cannot deny that there are some households for whom this correlation is substantially higher. We therefore use 0.5 to investigate how our results change.

There is no known empirical estimate for $\rho_{n,\varepsilon}$ in the literature. There are potentially some a priori reasons to expect it not to be statistically different from zero since earnings shocks at the household level have a large idiosyncratic variance component, yet the component of this variance that can be attributed to aggregate shocks is generally quite small (Pischke, 1995, for instance). Nevertheless, we consider the potential effects of this parameter by increasing it to a high enough bound that can simultaneously maintain the positive definiteness of the variance covariance matrix given the other chosen parameters. We therefore use $\rho_{n,\varepsilon} = 0.15$ for this parameter in our comparative statics results.

The wealth accumulation and mean shares of wealth in stocks over the life cycle are depicted in Figure 9, Panel A and Figure 9, Panel B, respectively. We note that both wealth accumulation and portfolio shares are substantially affected over the life cycle. The most dramatic effect arises when the correlation between the factor and the stock return innovation is set to zero ($\rho_{z,\varepsilon} = 0$). In this instance, the share of wealth in stocks becomes very similar to the i.i.d. model illustrating the importance of this correlation in the model. As a result of the higher asset allocation in stocks almost throughout the life cycle, and the lack of correlation between the factor innovation and stock returns, the wealth accumulation is now substantially higher than in the benchmark model. Since this wealth accumulation divergence takes a while to show up (it takes place beyond age 35) we can infer that this effect is driven by the higher expected equity return from the higher asset allocation in stocks.

Increasing the correlation between permanent labor income shocks and stock returns reduces the share of wealth in stocks, especially in the early part of the life cycle when these shocks are more permanent in the sense that they affect labor income over many more periods (Figure 9, Panel B). The higher wealth accumulation from the early part of the life-cycle shows that there is a slightly higher saving when the correlation between permament earnings shocks and stock returns increases, leading to a slightly higher wealth accumulation by retirement.

Finally, the positive correlation between the factor innovation and the permanent income shock tends to increase the share of wealth in stocks. This is explained by the lower wealth accumulation (Figure 9, Panel A) that reflects a higher consumption policy function early in life when this correlation is positive. The lower wealth accumulation translates to a higher average share of wealth in stocks as riskless assets in the form of human capital are now a lower percentage of total accumulated wealth.

5.2 Evaluating hedging demands

Campbell and Viceira (1999) quantify hedging demands by comparing hedging demands from a model with a factor predicting returns relative to the myopic model with a constant share of wealth in stocks. We consider the i.i.d. model as the equivalent of the myopic model in our case in the sense that the portfolios in the i.i.d. model do not exhibit any time variation in response to the factors. We first simulate all models of interest to compute the average shares of wealth in stocks over the life cycle. We then define hedging demands for every year of the life cycle as the percentage change in the average share of wealth allocated to stocks that arises from varying a specific parameter relative to the i.i.d. model. For instance, the hedging demand component that arises when the benchmark mean reversion model is simulated with $\rho_{z,\varepsilon} = -0.8$ relative to the i.i.d. model is equal to

$$hedg(\rho_{z,\varepsilon} = -0.8) = 100 * \{\frac{\alpha(\rho_{z,\varepsilon} = -0.8) - \alpha(I.I.D.)}{\alpha(I.I.D.)}\}$$

This allows us to quantify the hedging demand arising from the intertemporal variation in hedging demands relative to the model with i.i.d. returns under different parameterization of the correlations of interest.

Figure 10 illustrates that using this definition generates a substantial hedging demand component even when comparing average profiles. The variations are substantial and are a mirror image of the large variations in average portfolios over the life cycle. The baseline case generates a substantial change in the average hedging demand component, as the cases when the correlations are different (the correlations between permanent labor income shocks and stock returns). The only exception is the case when the correlation between the factor innovation and the stock return innovation is zero: in that case there are almost no hedging demands relative to the i.i.d. model. It is interesting that large intertemporal hedging demands arise even though we are comparing averages over the life cycle and averages mask an even greater variation that takes place at the individual level.

We also use a second definition of hedging demands to understand how a change in a particular correlation affects the share of wealth in stocks relative to the baseline factor model. In this case, defining $\alpha(\rho_{z,\varepsilon} = -0.8)$ to equal the share of wealth in stocks in the benchmark case and $\alpha(\rho_{z,\varepsilon} = 0.0)$ as the share of wealth in stocks when all parameters stay the same except for $\rho_{z,\varepsilon}$ being set to equal zero, the fraction of hedging demand from changing $\rho_{z,\varepsilon}$ from -0.8 to 0.0 in the factor model is given by

$$hedg(\rho_{z,\varepsilon} = -0.8)_{\text{factor}} = 100 * \{\frac{\alpha(\rho_{z,\varepsilon} = 0.0) - \alpha(\rho_{z,\varepsilon} = -0.8)}{\alpha(\rho_{z,\varepsilon} = -0.8)}\}$$

The results from these two calculations are depicted in Figure 10 B and show a dramatic increase in the share of wealth in stocks relative to the baseline model when this correlation is set to zero.

What happens when the correlation between the permanent earnings shock and the stock market innovation is raised from zero to 0.5? When increasing the correlation of permanent labor income shocks and stock returns from 0.15 to 0.5, we can compare the cases when the correlation in the mean reversion model is 0.5 versus the case in the i.i.d. model when the correlation is 0.5. Figure 11, Panel A plots the average life cycle hedging demands for the mean reversion model against the i.i.d. model when this correlation is 0.5. There is a substantial difference across the results and this is illustrated in Panel B that plots the mean difference across the two average asset allocation profiles. With this correlation we have even larger hedging demands because in that case some of the portfolio allocations in the i.i.d. model fall to zero.

We conclude that the correlation structure between the innovations generates quantitatively significant intertemporal hedging demands, both when comparing the baseline mean reversion with the i.i.d. model but also when changing the correlations within the context of the mean reversion model. The fact that even when comparing the average responses generates large variation in intertemporal asset allocations is interesting and should be the subject of further research to better understand the time variation in these correlation structures at different investment horizons and decision-making intervals.

6 Are Lifestyle funds Optimal?

Financial advisors argue that the share of wealth in stocks should decrease as the investor approaches retirement and qualitatively this is what the i.i.d. model predicts as well. Nevertheless, we have seen that a factor model will generate substantial variation in the share of wealth in stocks over the life cycle based on the realization of the factor. The intuitive argument is that households approaching retirement in 2008-2009 when the stock market had lost a substantial percentage of its value should not have followed blindly the rule followed by life style funds.

In this section we evaluate how important this intuition might be. We start a simulation from the beginning of life but assume different initial realizations of the factor. As we have seen before, this implies that households faced with the low factor (predicting higher future stock returns) should be fully invested in the stock market while households faced with a high factor (predicting low future stock returns) should not invest anything in the stock market. Because the factor is persistent, it will take time before these initial decisions get reversed. Figure 12, Panel A, produces such a diagram. Even if an investor participates in the stock market from the beginning of their life cycle, the investor might find it optimal to actually increase their share of wealth in stocks as investment opportunities improve, because the factor is persistent.

We can again compute the hedging demands that arise from this setup by comparing these portfolios relative to their i.i.d. counterparts. Specifically, we compute

$$hedg(f(t) = i) = 100 * \{\frac{\alpha(f(t) = i) - \alpha(I.I.D.)}{\alpha(I.I.D.)}\}$$

where the initial realization of the factor occurs in the beginning of the life cycle. Figure 12, Panel B, illustrates that such hedging demands can be substantial and the persistent nature of the factor can generate changes relative to the i.i.d. case that might last up to thirty years.

Given that most households will probably start actively participating in the stock market at some later point in their life cycle, we interpret these results as casting substantial doubt on the conventional wisdom that lifestyle funds are optimal all the time and for all households.

7 Conclusion

In the presence of stock market predictability, undiversifiable labor income risk and exogenously imposed liquidity constraints, the consumption policy rule has a similar shape with consumption functions derived in the buffer stock saving literature. Specifically, the consumption function is concave with a marginal propensity to consume out of liquid wealth equal to one for low levels of liquid wealth but with the marginal propensity falling very quickly beyond a certain consumption level. Nevertheless, stock market predictability generates one important quantitative difference on the optimal level of saving, conditional on the factor realization. Specifically, high future expected stock returns generate an increase in the speculative demand for saving while low future expected stock returns generate a decrease in saving.

Optimal portfolio choice is shown to be heavily dependent on the realization of the factor predicting future returns. Consistent with Barberis (2000) and Brennan, Schwartz and Lagnado (1997) who study a similar problem without labor income, stock market predictability implies that portfolio holdings will very often be either completely allocated in the stock market or in the riskless asset market when no borrowing and no short selling are allowed. Hedging demands can be substantial over the life cycle and doubts can be cast on the conventional wisdom that lifestyle funds that require a lower allocation to stocks as retirement approaches are always optimal.

A Appendix

A.1 Accurately approximating a VAR

There are four exogenous variables that need to be discretized to compute expectations. The four variables are the exogenous factor predicting returns (f_t), the stock market return (r_t), the innovation to the permanent component of labor income ($\ln N_{it}$) and the innovation to the transitory component of labor income ($\ln U_{it}$). The factor that predicts future stock returns, f_{t+1}, follows the AR(1) process (4). The quadrature methods proposed by Tauchen and Hussey (1991) are used to compute expectations numerically after stacking all four exogenous variables in a vector autoregression of order 1. The method allows us to use arbitrary correlations through a Choleski decomposition approach as in Burnside (1999).

A key methodological innovation in the paper is to determine the number of quadrature points needed to accurately capture the dynamics of the VAR without causing a computational intractability as the problem suffers from the curse of dimensionality. We report extensive experiments with the Tauchen-Hussey parameterization when there are four different variables. The first variable (f_t) is the factor predicting stock market returns and that is the main exogenous, persistent factor that needs to be well approximated. The stock market returns (r_t) is the second variable. We add the two other labor income variables to this VAR, even though they are not persistent, because we will investigate how changing correlations between these variables affects hedging demands. Adding them as part of the VAR can help us ascertain what number of grid points is required for the numerical accuracy of the discretization to be acceptable (there will always be an accuracy – computational time trade off in this work). The log of the innovation to the permanent income component of labor income is $ln(N_t)$ and the log of the transitory labor income shock innovation is $ln(U_t)$. We start with the following VAR model:

$$
\begin{bmatrix} f_{t+1} \\ r_{t+1} \\ lnU_{t+1} \\ lnN_{t+1} \end{bmatrix} = \begin{bmatrix} 0.0036 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix} + \begin{bmatrix} 0.91 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} f_t \\ r_t \\ lnU_t \\ lnN_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \\ u_{t+1} \\ n_{t+1} \end{bmatrix}
$$

Denote the coefficient matrix as follows:

$$
A1 = \begin{bmatrix} 0.0036 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix}, \; A2 = \begin{bmatrix} 0.91 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}
$$

and the variance-covariance matrix of the innovations as follows:

$$
\Omega = \begin{bmatrix} 0.000034 & -0.0008397 & 0 & 0 \\ & 0.0324 & 0 & 0.0027 \\ & & 0.01 & 0 \\ & & & 0.01 \end{bmatrix}
$$

To test the accuracy of the approximation method, we simulate based on a Tauchen-Hussey discretization for a given number of grid points. We simulate based on this discretization and then perform a Monte Carlo analysis to investigate how close the estimated parameters are to the actual parameters. Specifically, we generate 100 simulation paths over 52,000 periods. In each simulation, the first 2000 periods are discarded. The VAR is then estimated and the coefficients averaged over the 100 trials and reported below. The numbers in parentheses are the standard deviations of the reported coefficients.

Experiment 1: The number of grid points for f_t is 10, the number of grid points for r_t is 20, the number of grid points for $ln(N_t)$ is 5 and the number of grid points for $ln(U_t)$ is 3.

$$
\hat{A1} = \begin{bmatrix} \underset{(0.00009)}{0.0041} \\ \underset{(0.0027)}{-0.0098} \\ \underset{(0.0015)}{-0.0049} \\ \underset{(0.0013)}{-0.0056} \end{bmatrix} = \begin{bmatrix} A1_1 \\ A1_2 \\ A1_3 \\ A1_4 \end{bmatrix}
$$

$$\hat{A}2 = \begin{bmatrix} \underset{(0.0019)}{0.897} & \underset{(0.00016)}{0.000036} & \underset{(0.00029)}{0.0000039} & \underset{(0.00025)}{-0.000078} \\ \underset{(0.06)}{1.26} & \underset{(0.0046)}{-0.00109} & \underset{(0.0084)}{0.00052} & \underset{(0.0072)}{0.0022} \\ \underset{(0.03)}{0.00043} & \underset{(0.0027)}{0.00033} & \underset{(0.0045)}{0.00057} & \underset{(0.0041)}{0.00034} \\ \underset{(0.032)}{0.018} & \underset{(0.0027)}{-0.00012} & \underset{(0.0049)}{0.00048} & \underset{(0.0048)}{0.00028} \end{bmatrix} = \begin{bmatrix} A2_{11} & A2_{12} & A2_{13} & A2_{14} \\ A2_{21} & A2_{22} & A2_{23} & A2_{24} \\ A2_{31} & A2_{32} & A2_{33} & A2_{34} \\ A2_{41} & A2_{42} & A2_{43} & A2_{44} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000002)}{0.000033} & \underset{(0.000006)}{-0.00082} & \underset{(0.0000026)}{-0.0000001} & \underset{(0.0000027)}{0.0000006} \\ & \underset{(0.00021)}{0.032} & \underset{(0.000073)}{0.0000059} & \underset{(0.000082)}{0.0027} \\ & & \underset{(0.000066)}{0.01} & \underset{(0.00005)}{0.0000009} \\ & & & \underset{(0.00006)}{0.00997} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.000000007983 \\ 0.000007549 \\ 0.000002263 \\ 0.000001913 \end{bmatrix}$. The average is 0.0000029.

MSE of elements of $A2$:

$$\begin{bmatrix} 0.00000348 & 0.00000002557 & 0.00000008176 & 0.0000000621 \\ 0.003505 & 0.00002077 & 0.00007021 & 0.00005277 \\ 0.0009682 & 0.000007453 & 0.00002009 & 0.00001666 \\ 0.001078 & 0.000007486 & 0.00002378 & 0.00002317 \end{bmatrix}$$

The average is 0.00036.

We can see that the estimations of $A1_2$ and $A2_{21}$ are not accurate. We infer from this result that we can improve accuracy by increasing the number of grid of r_t. This leads to experiment 2.

Experiment 2: Relative to experiment 1, the number of grid points for r_t is 30.

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0009)}{0.00409} \\ \underset{(0.0027)}{0.0105} \\ \underset{(0.0015)}{-0.0051} \\ \underset{(0.00138)}{-0.0055} \end{bmatrix}, \hat{A}2 = \begin{bmatrix} \underset{(0.002)}{0.897} & \underset{(0.00015)}{0.00006} & \underset{(0.00024)}{0.000034} & \underset{(0.00024)}{-0.000022} \\ \underset{(0.065)}{1.24} & \underset{(0.0042)}{-0.0014} & \underset{(0.0077)}{0.00012} & \underset{(0.0083)}{0.00012} \\ \underset{(0.036)}{0.0024} & \underset{(0.0026)}{0.00023} & \underset{(0.005)}{-0.00042} & \underset{(0.0049)}{-0.0004} \\ \underset{(0.033)}{0.0013} & \underset{(0.0026)}{-0.000035} & \underset{(0.0042)}{0.0005} & \underset{(0.0049)}{-0.00008} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000001)}{0.000033} & \underset{(0.0000056)}{-0.00082} & \underset{(0.0000028)}{0.0000003} & \underset{(0.0000025)}{-0.000000} \\ & \underset{(0.000189)}{0.032} & \underset{(0.00009)}{-0.000005} & \underset{(0.00008)}{0.0027} \\ & & \underset{(0.000056)}{0.01} & \underset{(0.000045)}{0.000008} \\ & & & \underset{(0.000065)}{0.00996} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.000000007615 \\ 0.000007731 \\ 0.000002373 \\ 0.000001908 \end{bmatrix}$. The average MSE is 0.000003.

MSE of elements of $A2$:

$$\begin{bmatrix} 0.000004143 & 0.00000002334 & 0.0000000591 & 0.0000000582 \\ 0.004309 & 0.00001779 & 0.00005878 & 0.0000687 \\ 0.00132 & 0.0000071 & 0.0000252 & 0.00002402 \\ 0.001113 & 0.000006973 & 0.00001821 & 0.00002402 \end{bmatrix}$$

The average is 0.0004375.

This does not improve accuracy. Since r_t is closely related to f_t, we then increase the number of grid points of f_t.

Experiment 3: Relative to experiment 1, the number of grid points for f_t is 15.

$$\hat{A1} = \begin{bmatrix} \underset{(0.00008)}{0.00395} \\ \underset{(0.0027)}{0.0197} \\ \underset{(0.00154)}{-0.0049} \\ \underset{(0.0016)}{-0.0062} \end{bmatrix} \quad \hat{A2} = \begin{bmatrix} \underset{(0.002)}{0.901} & \underset{(0.0002)}{0.00015} & \underset{(0.00032)}{-0.000013} & \underset{(0.0003)}{-0.00008} \\ \underset{(0.0634)}{1.01} & \underset{(0.0049)}{-0.0005} & \underset{(0.0094)}{0.00017} & \underset{(0.00927)}{0.00039} \\ \underset{(0.0367)}{0.0004} & \underset{(0.0023)}{0.00023} & \underset{(0.0043)}{0.00023} & \underset{(0.0044)}{-0.00017} \\ \underset{(0.0368)}{0.032} & \underset{(0.0031)}{-0.00044} & \underset{(0.005)}{-0.0005} & \underset{(0.0046)}{-0.00094} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000002)}{0.000034} & \underset{(0.0000061)}{-0.000837} & \underset{(0.0000029)}{-0.0000003} & \underset{(0.0000029)}{0.0000007} \\ & \underset{(0.00021)}{0.0323} & \underset{(0.00008)}{0.000016} & \underset{(0.00009)}{0.00267} \\ & & \underset{(0.00006)}{0.00999} & \underset{(0.00005)}{0.00001} \\ & & & \underset{(0.000066)}{0.0099} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.00000000737 \\ 0.000007507 \\ 0.000002373 \\ 0.00000259 \end{bmatrix}$. The average is 0.0000031.

MSE of elements of $A2$:

$$\begin{bmatrix} 0.00000386 & 0.00000002635 & 0.000000101 & 0.0000000859 \\ 0.004027 & 0.00002398 & 0.0000877 & 0.0000859 \\ 0.00135 & 0.00000537 & 0.00001848 & 0.00002 \\ 0.001355 & 0.000009968 & 0.00002494 & 0.0000211 \end{bmatrix}$$

The average is 0.0004397.

The higher number of grid points of f_t increases the estimation accuracy for $A1_2$ and $A2_{21}$. An interesting question is whether we can further improve estimation accuracy by increasing the number of grid points of f_t and r_t. Experiments 4 and 5 are intended to answer this question.

Experiment 4: Relative to experiment 1, the number of grid points for f_t is 20.

$$\hat{A1} = \begin{bmatrix} \underset{(0.00007)}{0.00398} \\ \underset{(0.0026)}{0.01987} \\ \underset{(0.0016)}{-0.0049} \\ \underset{(0.0017)}{-0.0052} \end{bmatrix} \quad \hat{A2} = \begin{bmatrix} \underset{(0.0017)}{0.9005} & \underset{(0.00014)}{0.00014} & \underset{(0.0003)}{0.000032} & \underset{(0.00025)}{-0.000012} \\ \underset{(0.06)}{1.0008} & \underset{(0.00436)}{00069} & \underset{(0.0093)}{0.000022} & \underset{(0.0079)}{-0.000311} \\ \underset{(0.038)}{-0.00083} & \underset{(0.0026)}{-0.00023} & \underset{(0.0049)}{0.000036} & \underset{(0.0044)}{0.00022} \\ \underset{(0.04)}{0.0077} & \underset{(0.00298)}{0.00021} & \underset{(0.0042)}{-0.00031} & \underset{(0.0043)}{-0.000087} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000002)}{0.000034} & \underset{(0.0000062)}{-0.000839} & \underset{(0.0000003)}{-0.00000015} & \underset{(0.0000029)}{-0.00000054} \\ & \underset{(0.0002)}{0.00324} & \underset{(0.00009)}{0.000012} & \underset{(0.000088)}{0.00267} \\ & & \underset{(0.000073)}{0.00998} & \underset{(0.0000435)}{0.0000065} \\ & & & \underset{(0.000066)}{0.00995} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.000000005105 \\ 0.000006839 \\ 0.00000268 \\ 0.000002842 \end{bmatrix}$. The average is 0.0000031.

MSE of elements of $A2$:

$$\begin{bmatrix} 0.00000292 & 0.0000000206 & 0.000000086 & 0.0000000666 \\ 0.003548 & 0.00002897 & 0.0000859 & 0.0000623 \\ 0.001466 & 0.00000674 & 0.00002428 & 0.00002 \\ 0.0016 & 0.00000891 & 0.00001818 & 0.0000189 \end{bmatrix}$$

The average is 0.00043.

We conclude from these results that accuracy does not increase too much compared to experiment 3 (the MSEs are almost identical) and therefore start using 15 gridpoints for f_t.

Experiment 5: The number of grid points for f_t is 15, the number of grid points for r_t is 25, the number of grid points for $ln(N_t)$ is 5 and the number of points $ln(U_t)$ for is 3.

$$\hat{A}1 = \begin{bmatrix} \underset{(0.00007)}{0.00395} \\ \underset{(0.0028)}{0.02} \\ \underset{(0.00149)}{-0.0049} \\ \underset{(0.0015)}{-0.006} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.0016)}{0.901} & \underset{(0.00014)}{0.000097} & \underset{(0.00028)}{0.0000003} & \underset{(0.00027)}{-0.00003} \\ \underset{(0.061)}{0.989} & \underset{(0.0049)}{0.00055} & \underset{(0.0084)}{0.0012} & \underset{(0.0085)}{-0.00088} \\ \underset{(0.0345)}{-0.00118} & \underset{(0.0028)}{0.000058} & \underset{(0.00475)}{0.00102} & \underset{(0.005)}{-0.00059} \\ \underset{(0.035)}{0.027} & \underset{(0.00262)}{-0.00043} & \underset{(0.0041)}{0.0000065} & \underset{(0.0041)}{-0.00076} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000002)}{0.000034} & \underset{(0.0000056)}{-0.000838} & \underset{(0.0000028)}{-0.000000005} & \underset{(0.0000028)}{-0.0000003} \\ & \underset{(0.00021)}{0.03234} & \underset{(0.000081)}{-0.0000029} & \underset{(0.0000878)}{-0.00269} \\ & & \underset{(0.0000605)}{0.01} & \underset{(0.000051)}{-0.00000007} \\ & & & \underset{(0.000068)}{0.00996} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.00000000556 \\ 0.0000076 \\ 0.00000221 \\ 0.00000213 \end{bmatrix}$. The average is 0.000003.

MSE of elements of $A2$:

$$\begin{bmatrix} 0.0000027 & 0.00000002 & 0.000000078 & 0.000000076 \\ 0.0037 & 0.000024 & 0.000071 & 0.000073 \\ 0.0012 & 0.0000079 & 0.0000226 & 0.000025 \\ 0.00123 & 0.00000689 & 0.0000171 & 0.0000168 \end{bmatrix}$$

The average is 0.0004.

In this experiment, accuracy is only slightly better than the one in experiment 3.

Experiment 6: The number of gridpoints for f_t is 15, the number of grid points for r_t

is 30, the number of grid points for $ln(N_t)$ is 5 and the number of points $ln(U_t)$ for is 3.

$$\hat{A}1 = \begin{bmatrix} \underset{(0.00008)}{0.00397} \\ \underset{(0.0028)}{0.02} \\ \underset{(0.0016)}{-0.0048} \\ \underset{(0.0015)}{-0.0046} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.0018)}{0.90} & \underset{(0.00015)}{0.00017} & \underset{(0.00026)}{-0.000034} & \underset{(0.00025)}{-0.00007} \\ \underset{(0.065)}{0.99} & \underset{(0.0044)}{-0.0008} & \underset{(0.0084)}{0.0011} & \underset{(0.0075)}{0.001375} \\ \underset{(0.038)}{-0.004} & \underset{(0.0027)}{0.00001} & \underset{(0.0043)}{0.0001} & \underset{(0.0048)}{0.00024} \\ \underset{(0.0337)}{-0.0083} & \underset{(0.0027)}{0.00035} & \underset{(0.0045)}{0.00084} & \underset{(0.0049)}{0.00022} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.00000021)}{0.000034} & \underset{(0.0000056)}{-0.00083765} & \underset{(0.0000028)}{0.00000011} & \underset{(0.0000028)}{-0.0000005} \\ & \underset{(0.00019)}{0.03233} & \underset{(0.000081)}{-0.0000018} & \underset{(0.0000886)}{0.0027} \\ & & \underset{(0.00007)}{0.01} & \underset{(0.000042)}{0.0000054} \\ & & & \underset{(0.00006)}{0.00996} \end{bmatrix}$$

MSE of elements of $A1$: $\begin{bmatrix} 0.0000000068 \\ 0.0000084 \\ 0.0000027 \\ 0.0000022 \end{bmatrix}$. The average is 0.0000033

MSE of elements of $A2$:

$$\begin{bmatrix} 0.0000033 & 0.000000024 & 0.00000007 & 0.000000065 \\ 0.0037 & 0.000019 & 0.0000698 & 0.000057 \\ 0.0015 & 0.0000073 & 0.0000186 & 0.000023 \\ 0.00113 & 0.0000073 & 0.00002 & 0.000024 \end{bmatrix}$$

The average is 0.00045.

After increasing the number of gridpoints for r_t, accuracy does not improve too much compared with the last experiment.

Based on these reported (and other similar unreported) experiments, in all results reported in this paper we use 15 gridpoints for f_t, 20 gridpoints for r_t, 5 gridpoints for $ln(N_t)$ and 3 gridpoints for $ln(U_t)$.

A.2 Value Function

We use value function iteration to solve a life cycle model with Epstein-Zin-Weil preferences that separate risk aversion from the elasticity of intertemporal substitution. The expectation operator E_t is conditional on all information known at time t (x_t, f_t). Conditional on this information we have the following problem

$$v_{it}(x_{it}, f_j) = \Bigg[(1-\beta)c_{it}^{1-1/\psi} + \beta \left\{ \sum_{k=1}^{15} \sum_{U,N,R} \pi_{jk}\pi^{'} \left[\begin{array}{c} p_{t+1}(v_{it+1}\left(x_{it+1}, f_k\right))^{1-\gamma}(Y_{it+1}^{p}/Y_{it}^{p})^{1-\gamma}+ \\ b(1-p_{t+1})(x_{it+1})^{1-\gamma}(Y_{it+1}^{p}/Y_{it}^{p})^{1-\gamma} \end{array} \right] \right\}^{\frac{1-1/\psi}{1-\gamma}} \Bigg]^{\frac{1}{1-1/\psi}}$$

while π_{jk} is the probability that the factor moves from the current period value of f_j to the next period value equal to f_k. The other probability is $\pi^{'}$ and denotes the probabilities for next period's labor income shocks (U, N) and stock returns (R). We discretize the state variable x by dividing it into 251 grid points, with a larger number of grid points for low levels of cash on hand. We use a grid search approach to maximize backwards the value function and obtain the optimal policy functions.

From the Bellman equation the optimal decisions are given as current utility plus the discounted expected continuation value $(E_t v_{t+1}^{j}(.))$, which we can compute since we have just obtained v_{t+1}^{j}. The VAR approximation discussed previously is used to compute expected values and value function interpolation is done with cubic splines.

A.3 Simulation

After policy functions are computed we perform a simulation based on 500 individual life histories that are averaged over 200 simulated factor draws.

References

[1] Abowd, John and David Card, 1989, "On the Covariance Structure of Earnings and Hours Changes," *Econometrica* 57: 411-45.

[2] Angerer, Xiaohong and Pok-Sang Lam, 2009, "Income Risk and Portfolio Choice: An Empirical Study," *Journal of Finance*, LXIV, 2, 1037-55.

[3] Bagliano, Fabio, Carolina Fugazza and Giovanna Nicodano, 2014, "Optimal Life-Cycle Portfolios for Heterogeneous Workers," *Review of Finance*, 18: 2283-2323.

[4] Balduzzi, Pierluigi and Anthony W. Lynch, 1999, "Transaction Costs and Predictability: Some Utility Cost Calculations," *Journal of Financial Economics* 52: 47-78.

[5] Barberis, Nicholas, 2000, "Investing for the Long Run when Returns are Predictable," *Journal of Finance*, Vol. 55 (1).

[6] Benzoni, Luca, Pierre Collin-Dufresne and Robert Goldstein, 2007, "Portfolio choice over the life-cycle when the stock and labor markets are cointegrated," *Journal of Finance*, 62(5), 2123–2167.

[7] Bodie, Zvi, Robert C. Merton and William F. Samuelson, 1992, "Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model," *Journal of Economic Dynamics and Control*, 16, 427-449.

[8] Bonaparte, Yosef, George Korniotis and Alok Kumar, 2014, "Income Hedging and Portfolio Decisions," *Journal of Financial Economics*, 113, 300-324.

[9] Brandt, Michael, 1999, "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach," *Journal of Finance*, 54, 1999, 1609-1646.

[10] Brennan, Michael, Eduardo Schwartz and R. Lagnado, 1997, "Strategic Asset Allocation," *Journal of Economic Dynamics and Control*, 21, 1377-1403.

[11] Brennan, Michael, and Yihong Xia, 2002, "Dynamic Asset Allocation under Inflation," *Journal of Finance,* 57:1201–38.

[12] Burnside, Craig. 1999, "Dicrete state-space methods for the study of dynamic economies." in *Computational Methods for the Study of Dynamic Economies*, edited by Ramon Marimon and Andrew Scott, Oxford University Press.

[13] Campbell, John Y., and Angus Deaton, 1989, "Why is consumption so smooth?" *Review of Economic Studies,* 56, 357-74.

[14] Campbell, John Y., Andrew Lo, and C. MacKinlay, 1997, "The Econometrics of Financial Markets," *Princeton University Press.*

[15] Campbell, John Y. and Luis Viceira, 1999, "Consumption and Portfolio Decisions When Expected Returns are Time Varying," *Quarterly Journal of Economics*, 114, 433-495.

[16] Campbell, J. Y., J. Cocco, F. Gomes, P. Maenhout, L. Viceira, 2002, "Stock Market Mean Reversion and the Optimal Allocation of a Long Lived Investor," *European Finance Review*, Vol. 5, Nr 3.

[17] Campbell, John, Y. Lewis Chan and Luis Viceira, 2003, "A Multivariate Model of Strategic Asset Allocation," *Journal of Financial Economics*, January, Vol. 67 (1), pp. 41-80.

[18] Carroll, Christopher, 1997, "Buffer Stock Saving and the Life Cycle / Permanent Income Hypothesis'. *Quarterly Journal of Economics* CXII no. 1: 3-55.

[19] Carroll, Christopher and Andrew Samwick, 1997, "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1).

[20] —— 1998, "How important is precautionary saving?" *Review of Economics and Statistics,* 80 no. 3: 410-19.

[21] Cocco, Joao, Francisco Gomes and Pascal Maenhout, 2005, "Consumption and Portfolio Choice over the Life-Cycle," *The Review of Financial Studies*, 18 (2): 491-533.

[22] Cochrane, John, 1999, "New Facts in Finance," *Economic Perspectives*, Federal Reserve Bank of Chicago, pp. 36-58.

[23] Cooper, Russell and Guozhong Zhu, 2014, "Household Finance over the Life-Cycle: What does Education Contribute?" Penn State working paper.

[24] Davis, Steven, Felix Kubler and Paul Willen, 2006, "Borrowing Costs and the Demand for Equity over the Life Cycle," *The Review of Economics and Statistics,* 88: 348–62.

[25] Deaton, Angus,1991, "Saving and Liquidity Constraints," *Econometrica* 59 no.5: 1221-48.

[26] Epstein, Larry, and Stanley Zin, 1989, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57, 937-969.

[27] Gollier Christian, and John W. Pratt, 1996, "Risk Vulnerability and the Tempering Effect of Background Risk," *Econometrica*, 64: 1109-1124.

[28] Gomes, Francisco and Alexander Michaelides, 2005, "Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence," *Journal of Finance*, 60:869–904.

[29] Gomes, Francisco, Alexander Michaelides and Valery Polkovnichenko, 2009, "Optimal Savings with Taxable and Tax-Deferred Accounts," *Review of Economic Dynamics*, 12: 718-735.

[30] Gourinchas Pierre-Olivier and Jonathan Parker, 2002, "Consumption over the Life-Cycle," *Econometrica*, Vol 70, 1, 47-89.

[31] Haliassos, Michael and Carol C. Bertaut, 1995, "Why do so few hold stocks?" *The Economic Journal*, 105, 1110-1129.

[32] Haliassos, Michael and Alexander Michaelides, 2003, "Portfolio Choice and Liquidity Constraints," *International Economic Review*, Vol. 44, No.1, pp. 144-177.

[33] Heaton John, and Deborah Lucas, 1996, "Evaluating the Effects of Incomplete Markets on Risk Sharing and Asset Pricing," *Journal of Political Economy* 104: 443-487.

[34] —— 1997. "Market Frictions, Savings Behavior, and Portfolio Choice," *Macroeconomic Dynamics* 1: 76-101.

[35] —— 2000. "Portfolio Choice in the Presence of Background Risk," *The Economic Journal*, 110, 1-26.

[36] Hubbard, Glenn, Jonathan Skinner, and Stephen Zeldes, 1995, "Precautionary Saving and Social Insurance," *Journal of Political Economy*, 103: 360-399.

[37] Kim, Tong Suk, and Edward Omberg, 1996, "Dynamic Nonmyopic Portfolio Behavior," *Review of Financial Studies*, 9, 141-61.

[38] Koijen, Ralph, Theo Nijman, and Bas Werker, 2010, "When Can Life Cycle Investors Benefit from Time-Varying Bond Risk Premia?" *The Review of Financial Studies*, 23 (2): 741-780.

[39] Lettau, Martin and Sydney Ludvigson, 2001, "Consumption, Aggregate Wealth and Expected Stock Returns," *Journal of Finance,* LVI (3), pp.815-849.

[40] Lynch, Anthony, and Sinan Tan, 2011, "Labor Income Dynamics at Business-Cycle Frequencies: Implications for Portfolio Choice," *Journal of Financial Economics*, 101: 333-359.

[41] Mankiw, N. Gregory and Stephen P. Zeldes, 1991, "The Consumption of Stockholders and Non-Stockholders," *Journal of Financial Economics,* 29, 97-112.

[42] Merton, Robert, 1969, "Lifetime Portfolio Selection under Uncertainty: The Continuous Time Case," *Review of Economics and Statistics*, 51, 247-57.

[43] —— 1971, "Optimum Consumption and Portfolio Rules in a Continuous Time Model," *Journal of Economic Theory*, 3, 373-413.

[44] —— 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867-87.

[45] Munk, Claus, and Carsten Sørensen, 2010, "Dynamic asset allocation with stochastic income and interest rates," *Journal of Financial Economics*, 96: 433–462.

[46] Pishcke Jörn-Steffen, 1995, "Individual Income, Incomplete Information and Aggregate Consumption," *Econometrica,* 63, 4: 805-40.

[47] Samuelson, Paul, 1969, "Lifetime Portfolio Selection by Dynamic Stochastic Programming," *Review of Economics and Statistics*, 51, 239-46.

[48] Tauchen, George, and R. Hussey, 1991, "Quadrature-Based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models," *Econometrica*, 59, 371-396.

[49] Viceira, Luis, 2001, "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income," *The Journal of Finance*, 55, 1163-98.

[50] Vissing-Jorgensen, Annette, 2002, "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution," *Journal of Political Economy*, vol. 110, no.4, pp.825-853.

[51] Wachter, Jessica, 2002, "Optimal Consumption and Portfolio Allocation with Mean-Reverting returns: An exact solution for complete markets," *Journal of Financial and Quantitative Analysis*, 37, pp. 63-91.

[52] Weil, Philippe, 1990, "Non-Expected Utility in Macroeconomics," *Quarterly Journal of Economics*, 2 91): 29-42.

[53] Xia, Yihong, 2001, "Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation," *Journal of Finance*, 56, 205-246.

Notes

[1]Many financial economists nowadays consider some predictability of stock returns over longer horizons as a stylized fact in finance (for an extensive list of references, see Campbell, Lo and MacKinlay (1997) and Cochrane (1999)).

[2]Viceira (2001) rigorously verifies the popular advice in the financial management industry that higher exposure in the stock market be taken during working life with a shift towards safe assets after retirement. It is perhaps useful to recall that the infinite horizon models of portfolio choice (Merton 1969, 1971 and Samuelson 1969) that assume fully tradable human capital and a constant investment opportunity set, predict that the share of wealth invested in the risky asset should be constant.

[3]Viceira (2001) finds similar results in a stylized life-cycle model with no explicitly imposed constraints. Bodie, Merton and Samuelson (1992) also show that it is optimal for employed investors to hold proportionately more stocks in their portfolios than retired investors in the presence of nontradable certain future labor income.

[4]The "excess smoothness" puzzle arises in the context of the PIH; given the observed positive serial correlation of labor income growth in aggregate data, the representative agent PIH predicts that consumption growth should be more volatile than aggregate income growth (Campbell and Deaton (1989)).

[5]The importance of buffer stock saving has been stressed by Carroll (1997), Carroll and Samwick (1997, 1998), Hubbard, Skinner and Zeldes (1995) and Gourinchas and Parker (2002).

[6]These values generate an autocovariance structure for the growth rate of labor income that is almost identical to the one used by Deaton (1991), who in turn deflates previous estimates to take into account the effects of measurement error.

Figure 1: Policy Functions: Benchmark case (γ=5,ψ=0.5)

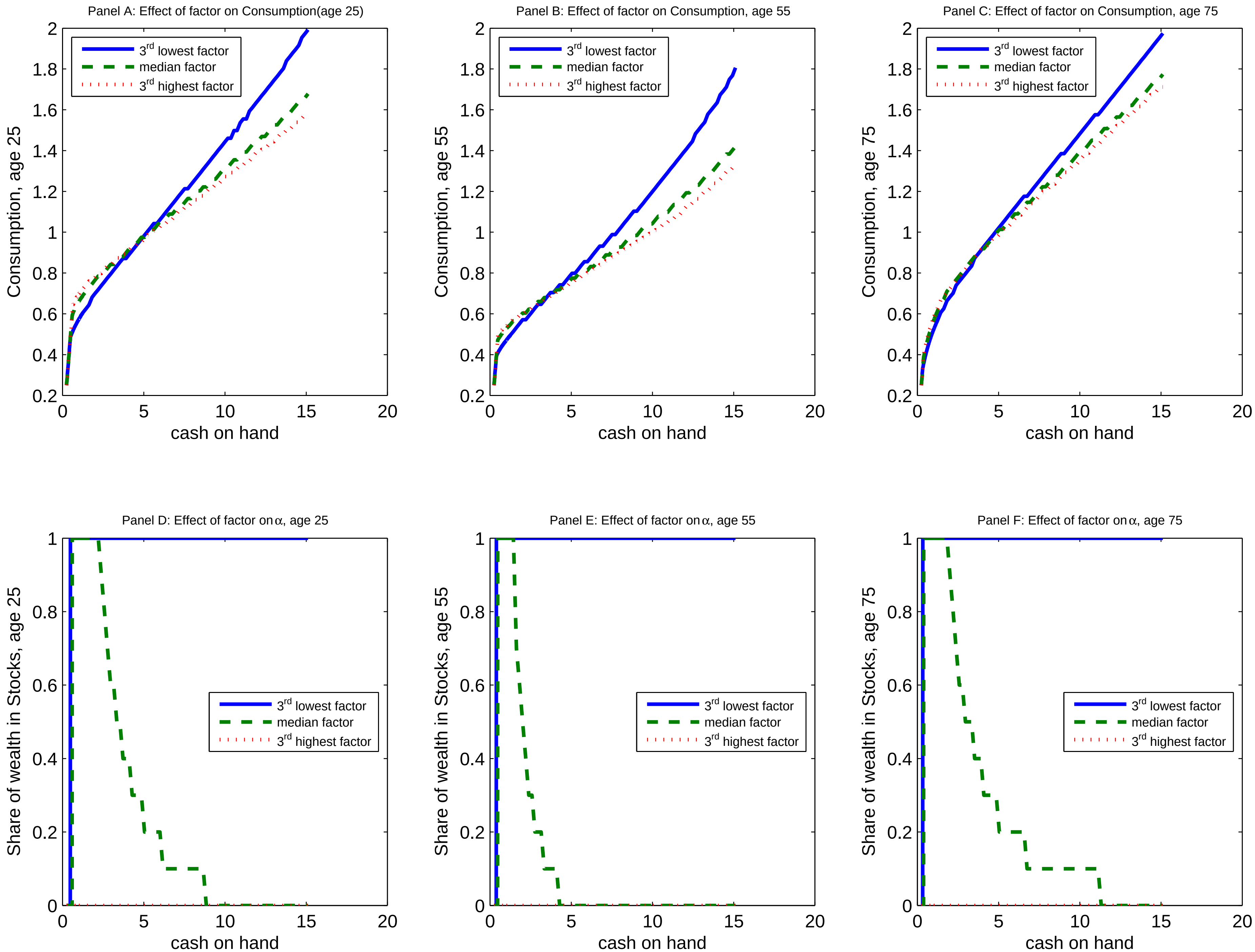


Figure 2: Benchmark

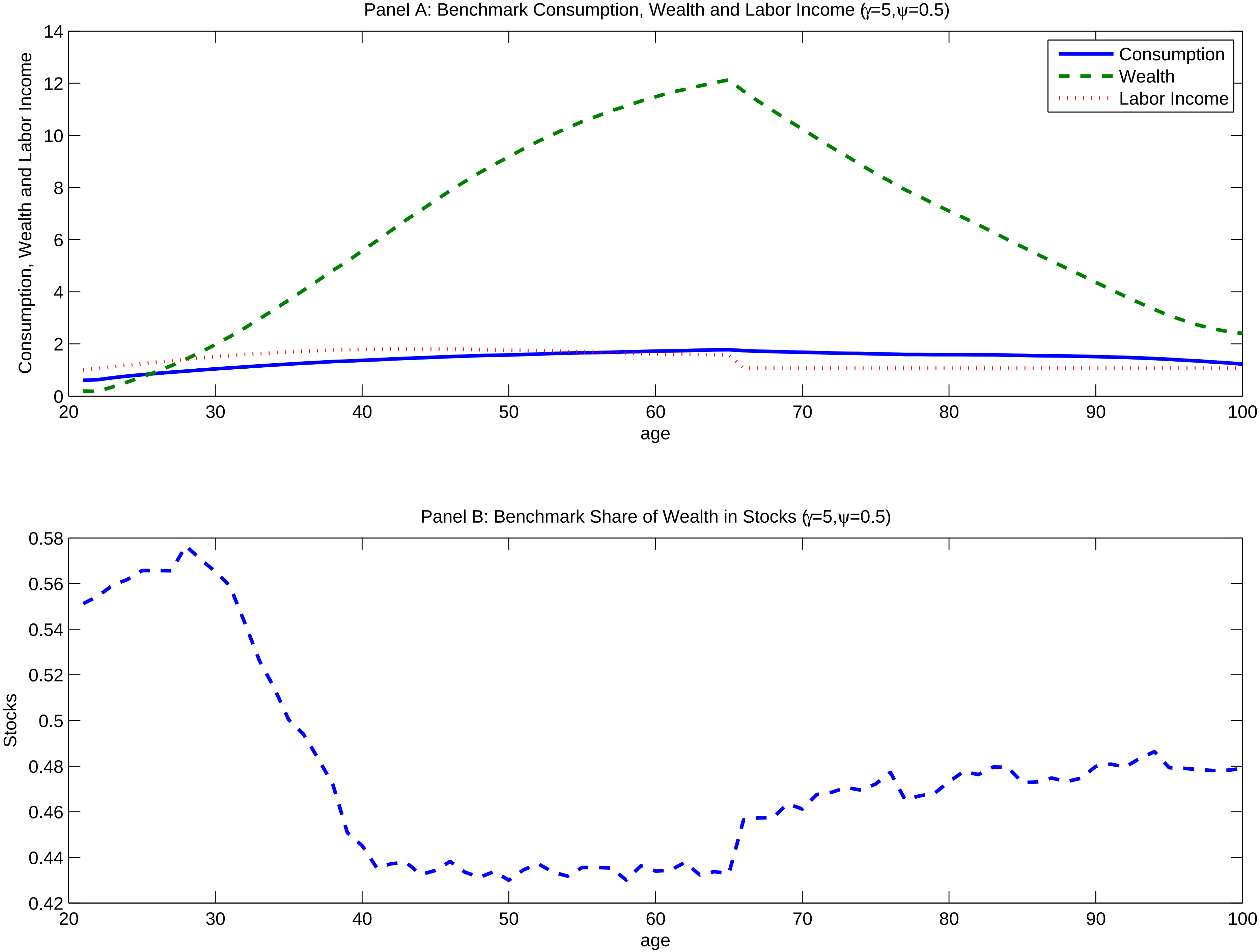


Figure 3: Consumption policy functions for different preference parameters

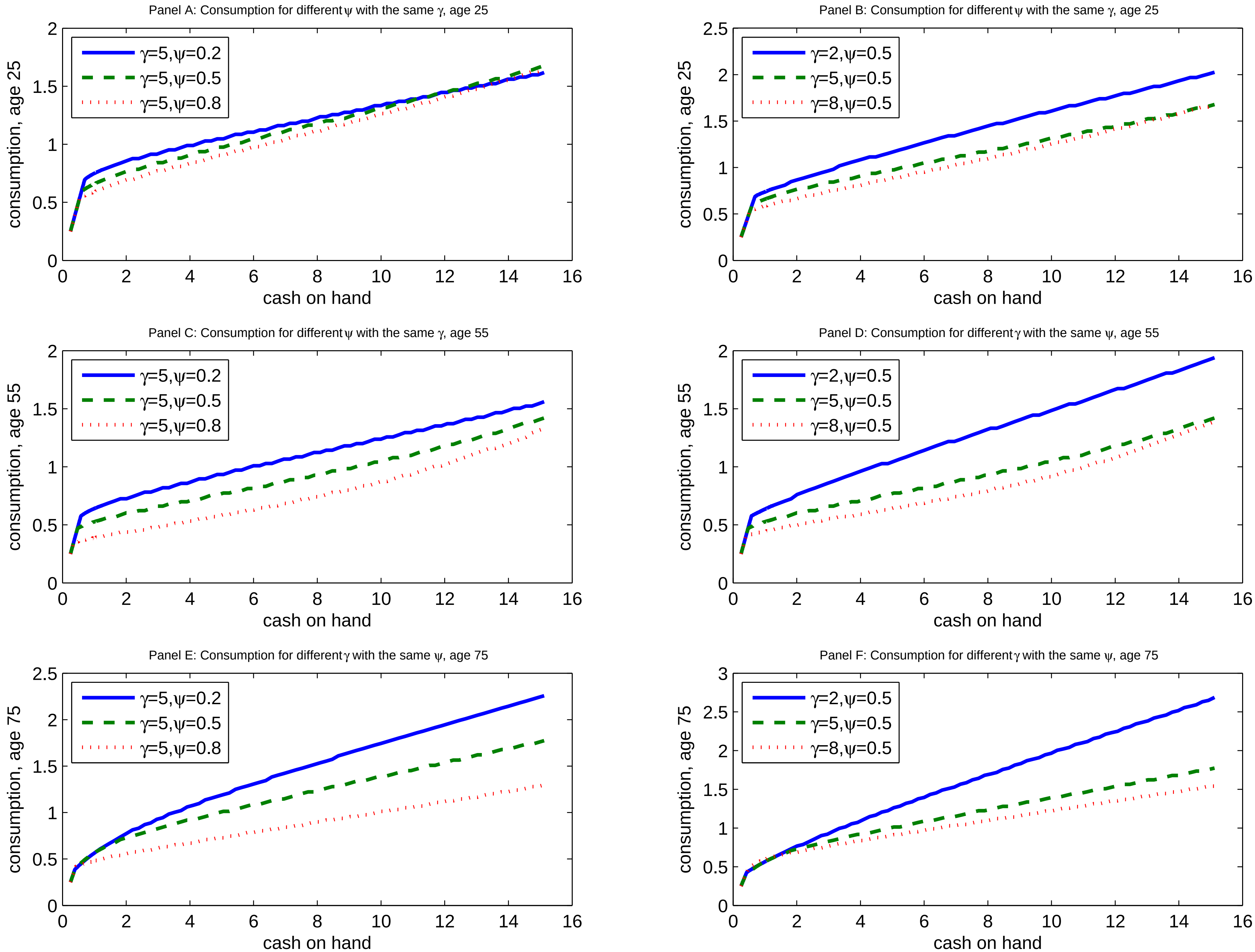


Figure 4: Share of wealth in stocks for different preference parameters

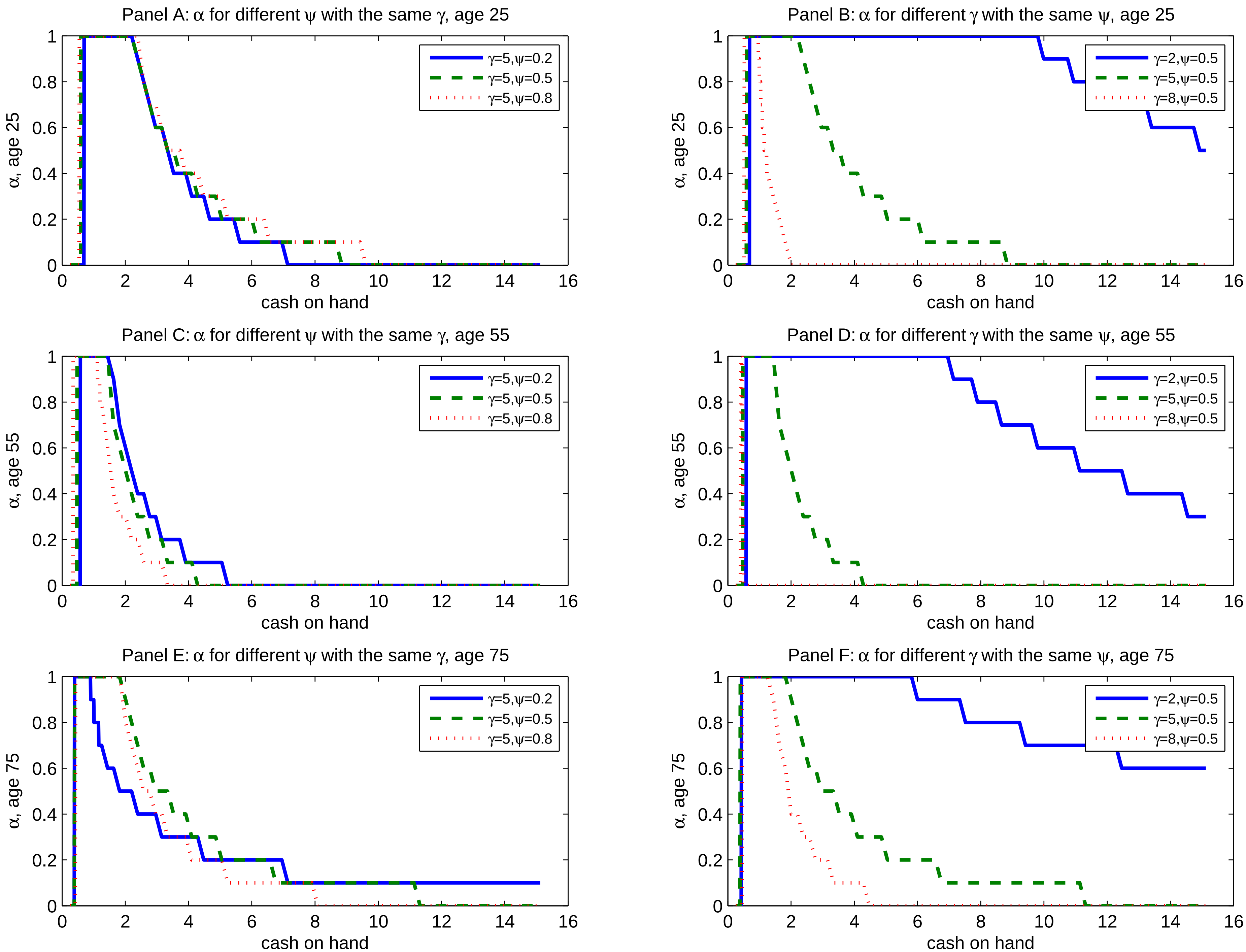


Figure 5: Benchmark with different parameters

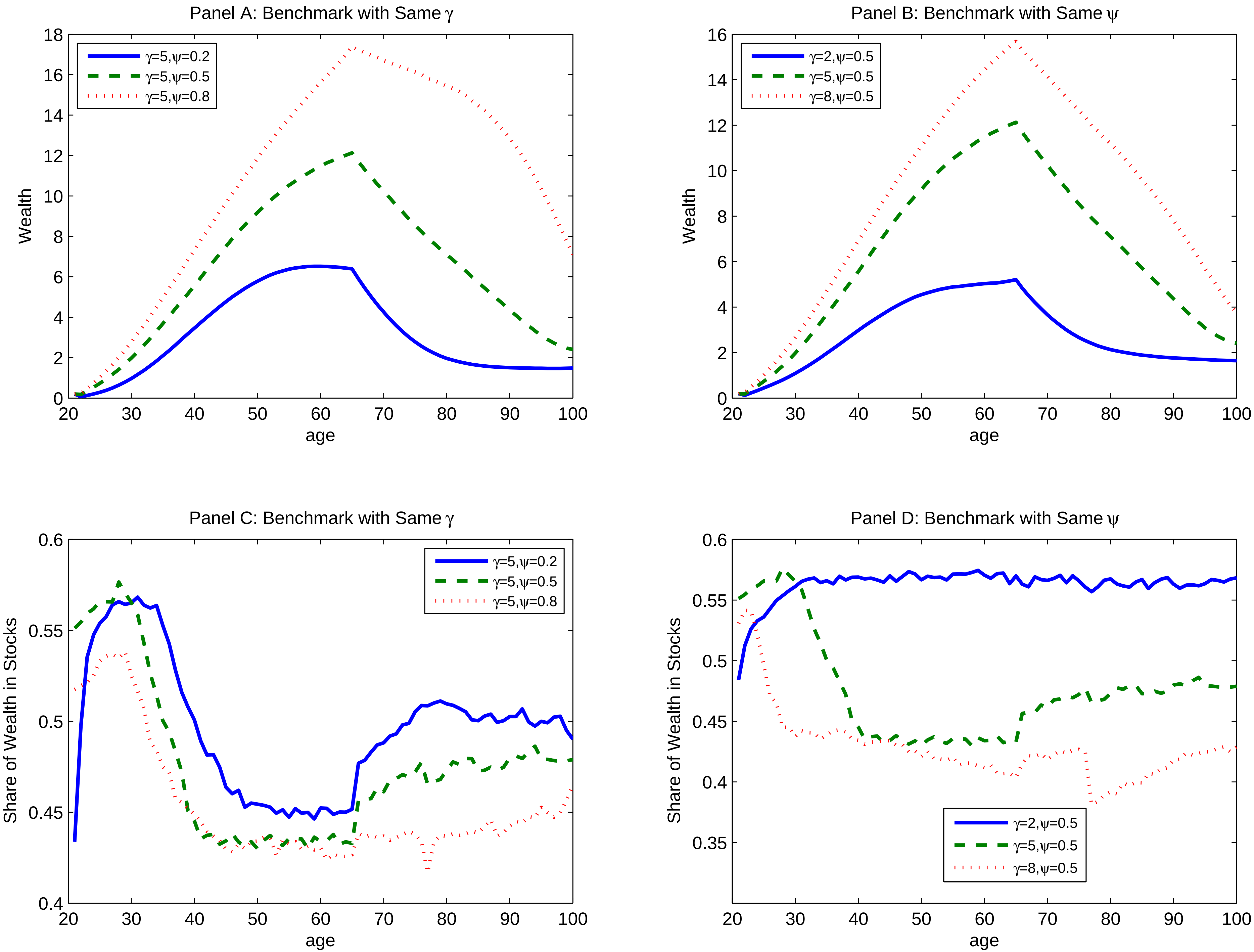


Figure 6: Share of Wealth in Stocks Policy Functions: Mean Reversion vs I.I.D. case

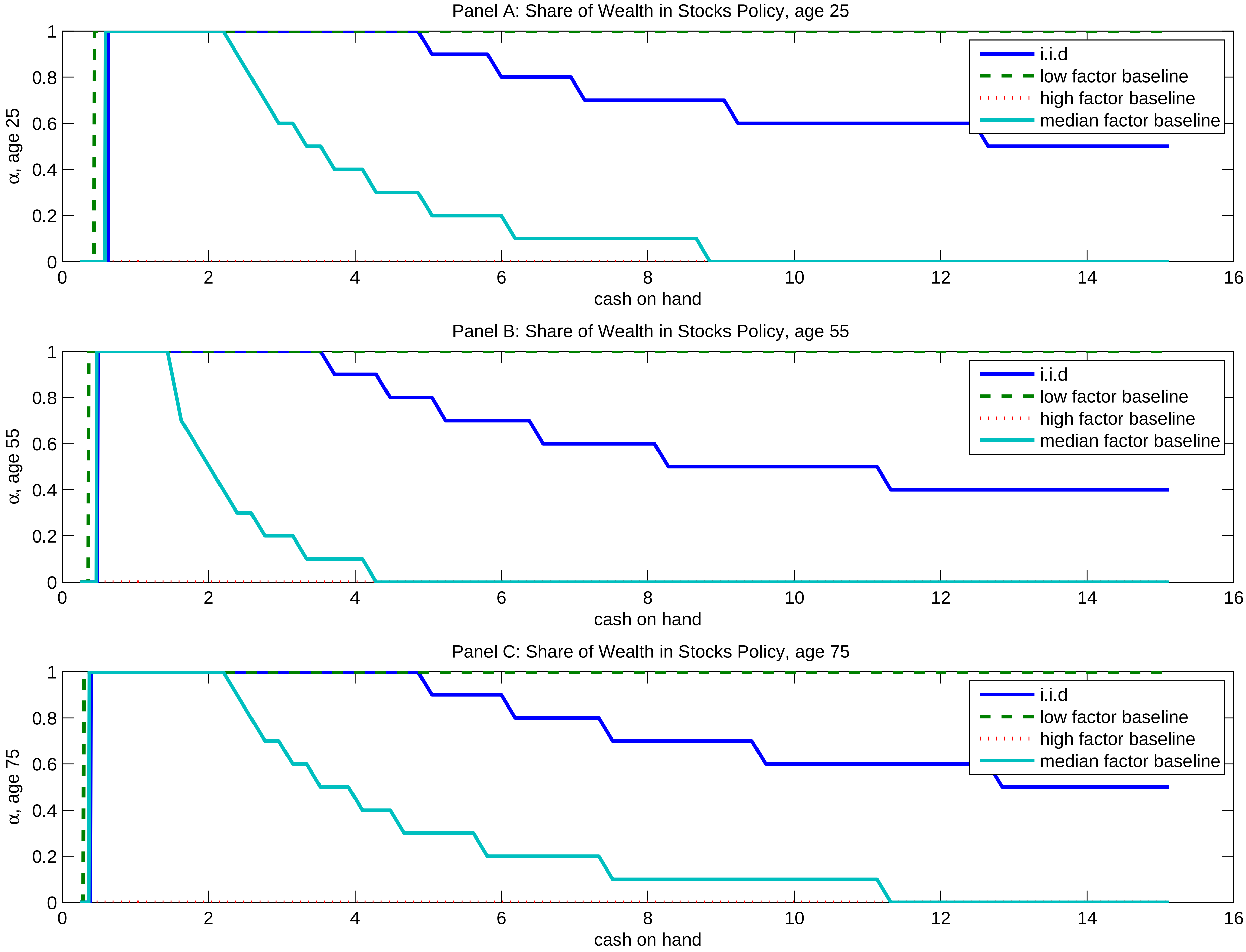


Figure 7: Wealth and Portfolio Shares Comparison

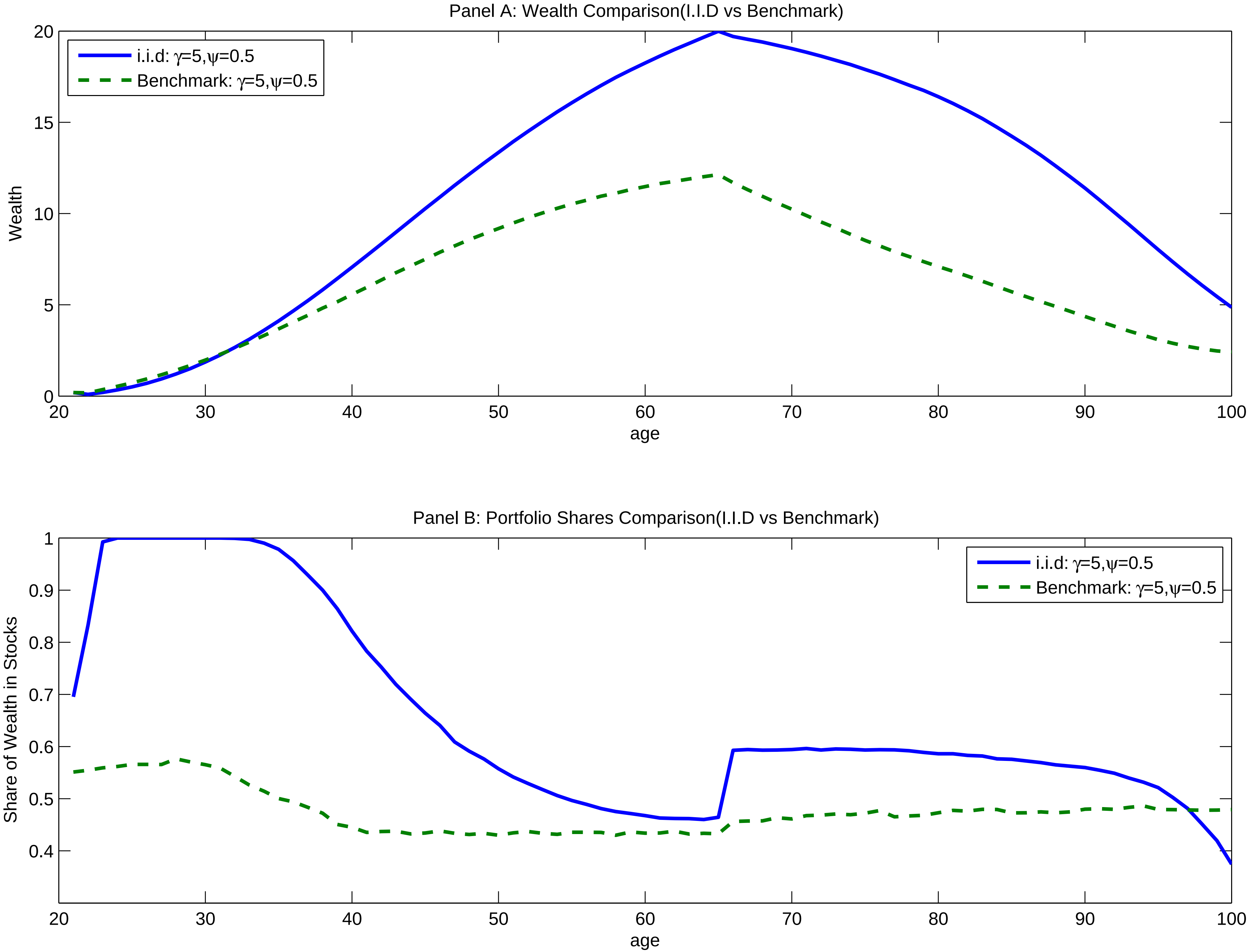


Figure 8: Life−cycle simulated share of wealth in stocks for different initial factor

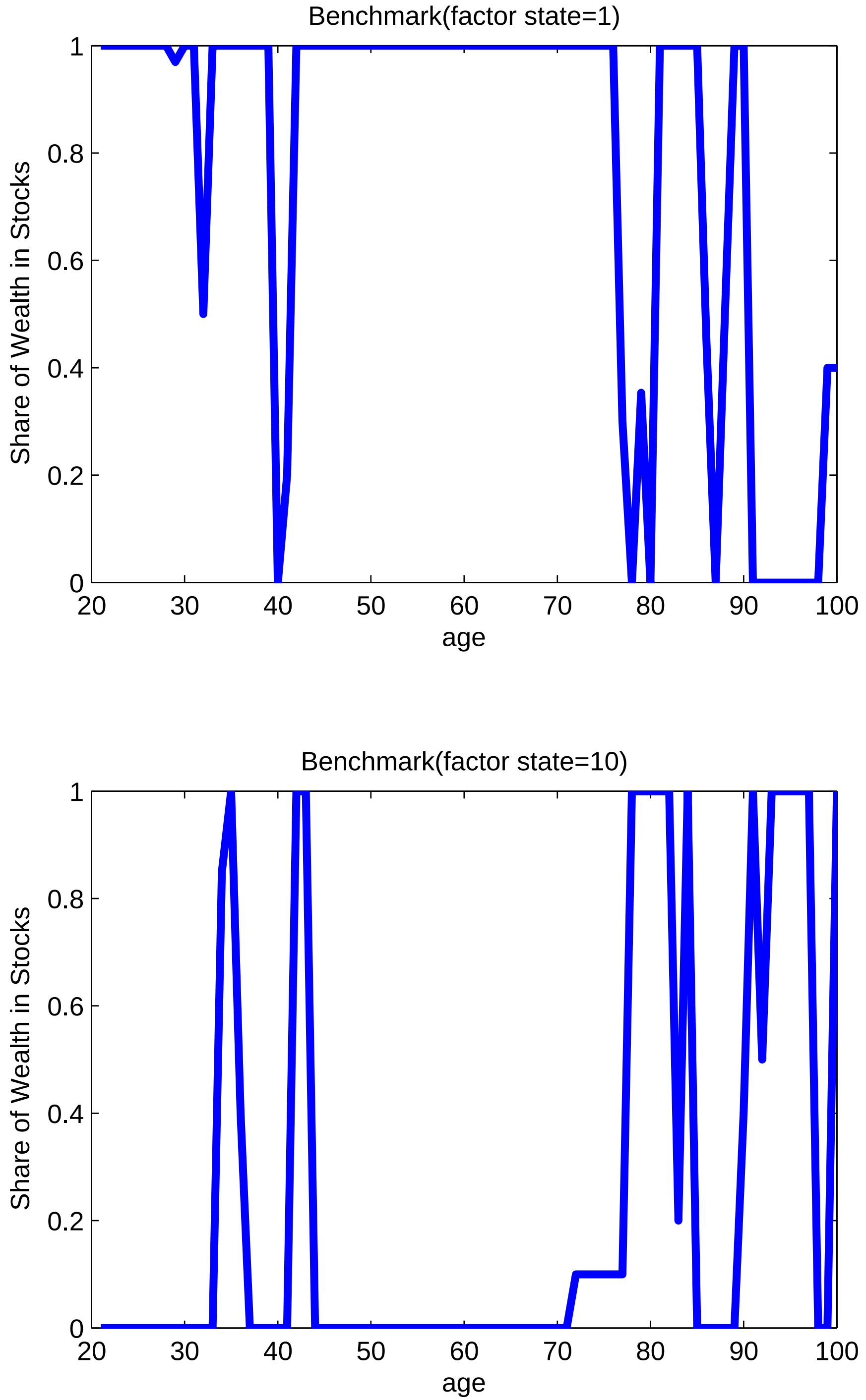
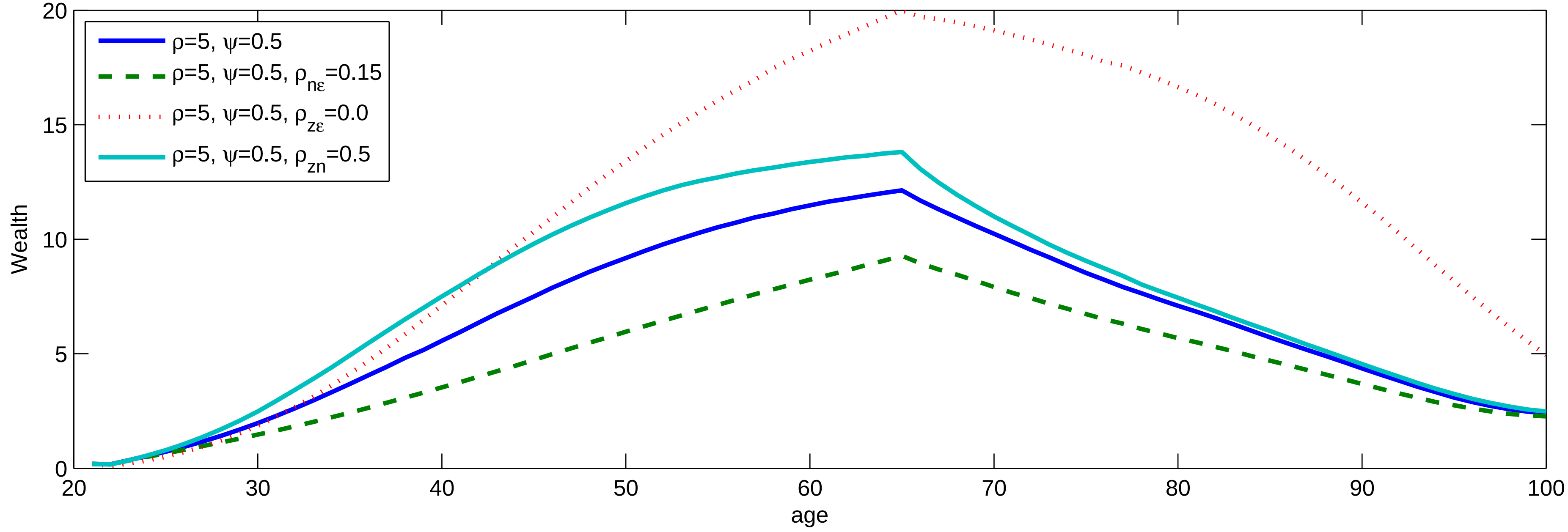


Figure 9: Life−cycle wealth and portfolio shares

Panel A: Life−cycle wealth with different parameters



Panel B: Life−cycle portfolio shares with different parameters

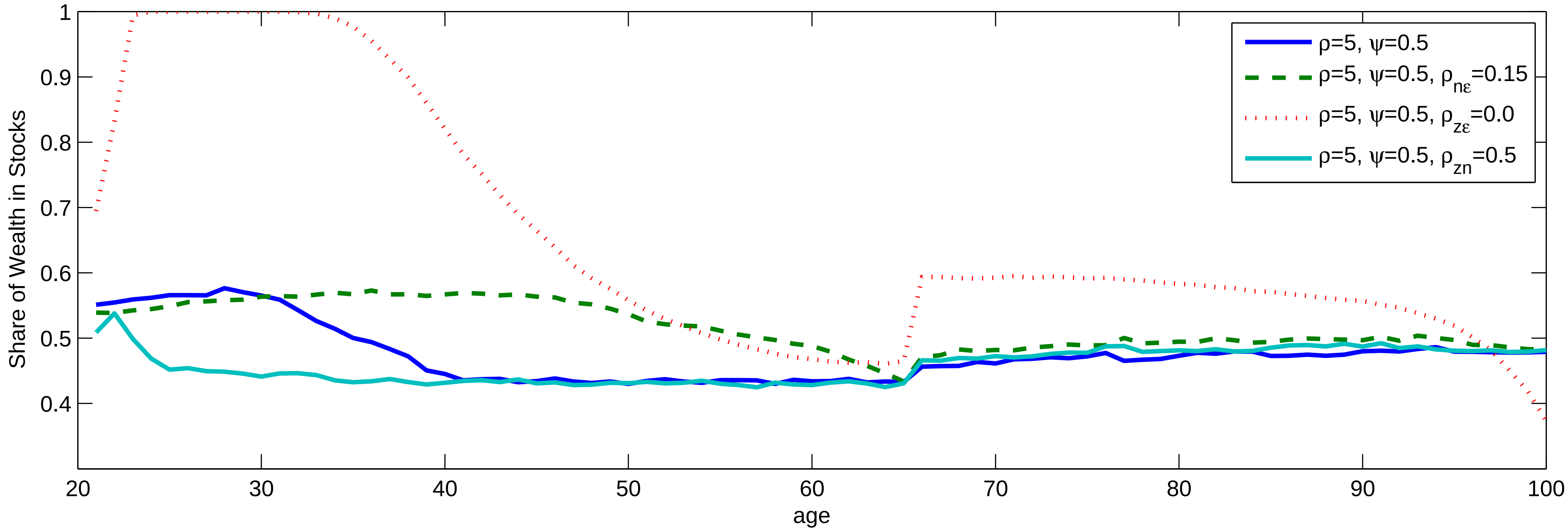


Figure 10: Life−Cycle Hedging Demands

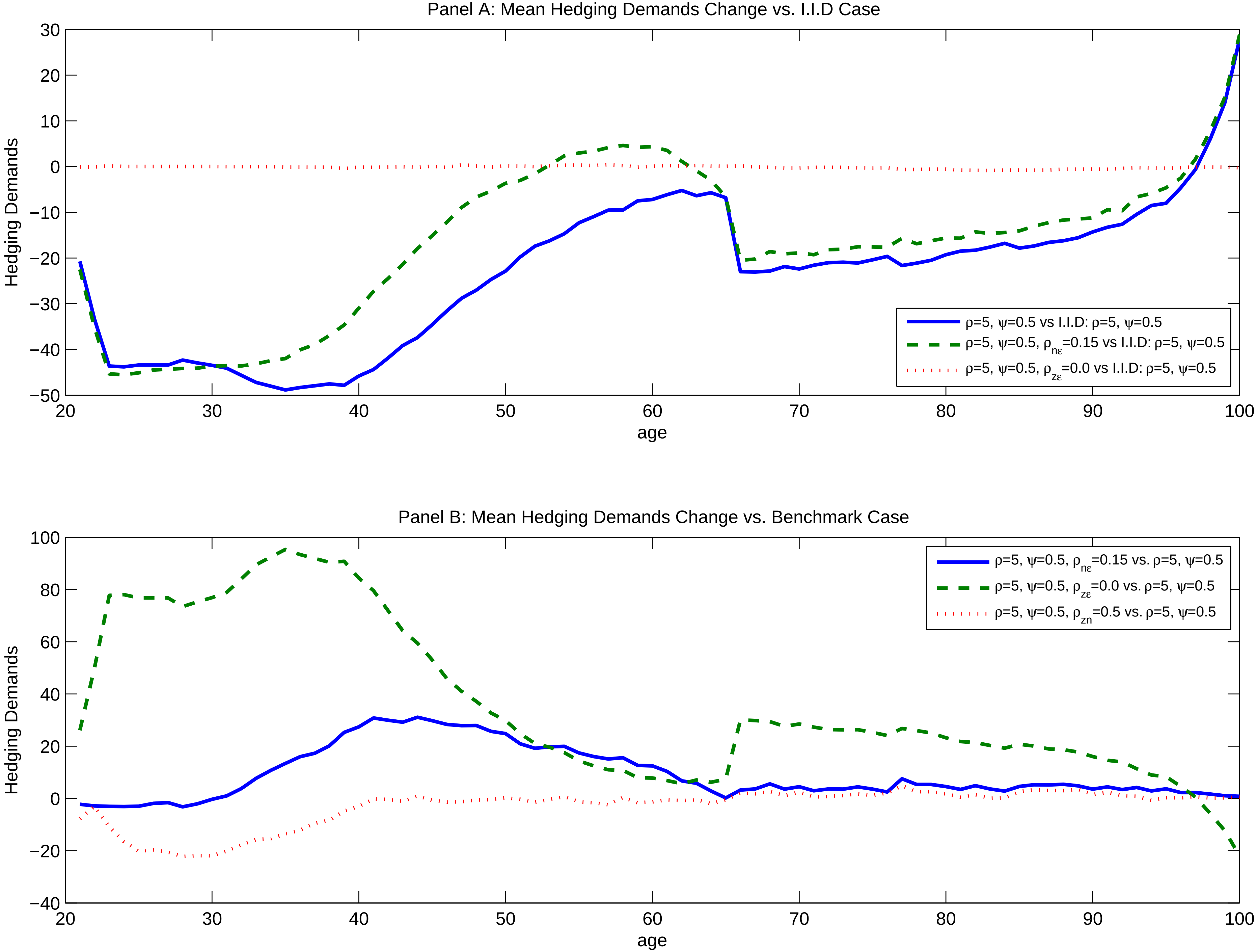


Figure 11: Life−Cycle Hedging Demands

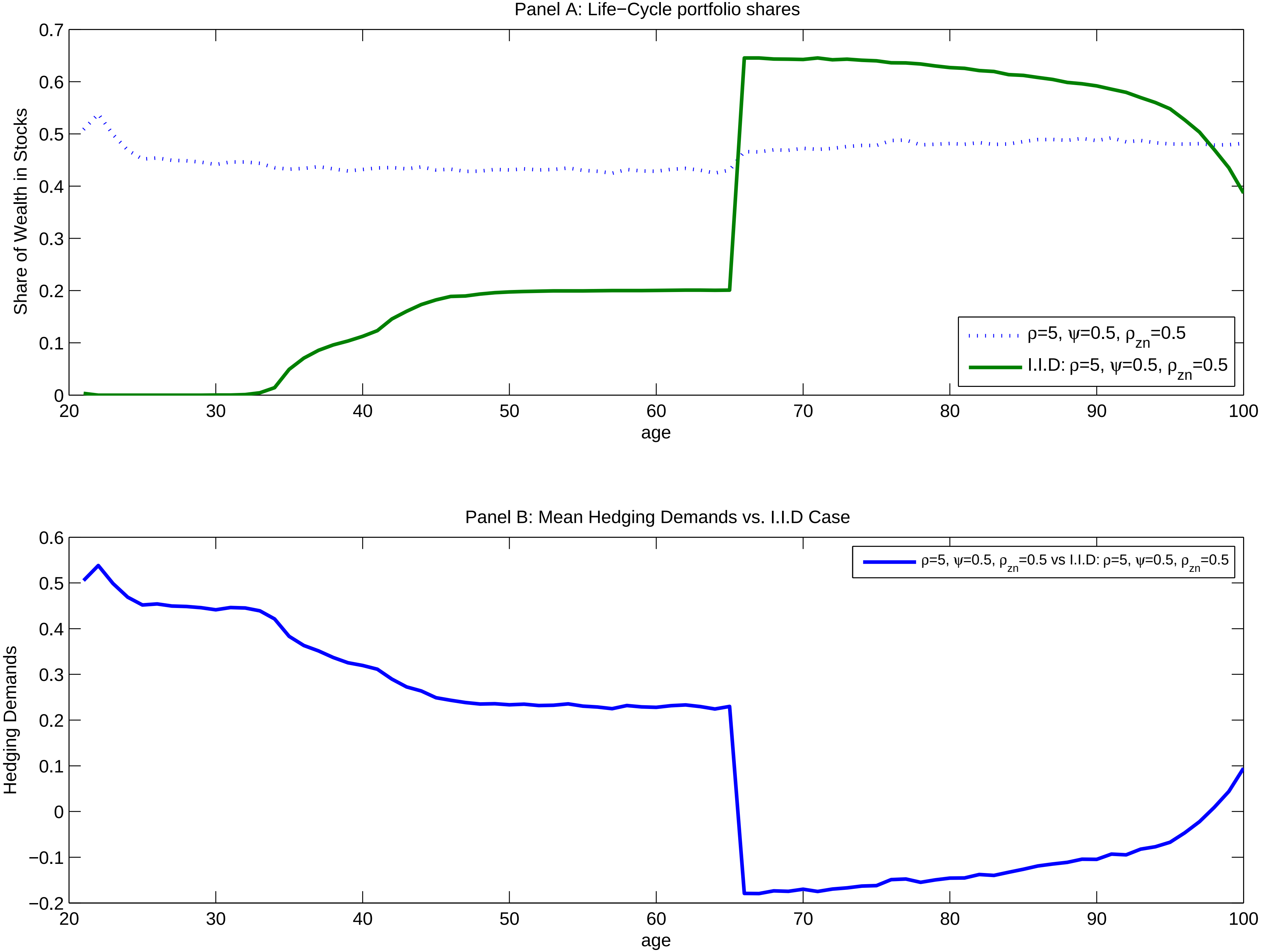


Figure 12: Life−Cycle portfolio shares and Hedging Demands