Network for Studies on Pensions, Aging and Retirement

NETSPAR ACADEMIC SERIES

Stock Market Mean Reversion and Portfolio Choice over the Life Cycle

Alexander Michaelides and Yuxin Zhang

DP 06/2015-018

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

Alexander Michaelides†

Imperial College London, CEPR and NETSPAR

Yuxin Zhang

Imperial College London‡

June 11, 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 an anonymous referee, Nicholas Barberis, John Campbell, John Cochrane, George Constantinides, Steve Davis, Christian Gollier, Francisco Gomes, Michael Haliassos, James Kahn, Martin Lettau, Sydney Ludvigson, Lubos Pastor, Annette Vissing-Jorgensen, Paul Willen and seminar participants at the Athens University of Economics and Business, the NETSPAR international conference on pensions, the University of Chicago Graduate School of Business, the University of Cyprus, the University of Southampton, the University of Sussex, 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 and show that the investor pursues aggressive market timing strategies. Importantly, in the presence of stock market predictability, the model suggests that the conventional financial advice of reducing stock market exposure as retirement approaches is correct on average, but ignoring changing market information can lead to substantial welfare losses. Therefore, enhanced target-date funds (ETDFs) that condition on expected equity premia increase welfare relative to target-date funds (TDFs).

JEL Classification: E21, G11.

Key Words: Portfolio Choice over the Life Cycle, Stock Market Mean Reversion, Stock Market Predictability, Hedging Demands, Lifestyle funds, enhanced target-date 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,[2] whereas others focus on the effect of background labor income risk on portfolio choice while ignoring stock market predictability.[3] 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.

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 Sorensen (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.

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). 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, 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[4]. 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"[5] and "excess sensitivity" puzzles.[6]

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, but does not respond substantially to changes in the investment opportunity set. On the other hand, 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. This translates to large variations over the life cycle depending on the factor realization 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 shift up (for high expected risk premia) or down (for low expected risk premia) the share of wealth allocated in stocks.

When calibrated to the observed dividend yield as a factor of predictability, the asset allocation profiles retain the target date funds (TDF) feature of slowly decreasing stock market exposure as the household ages. Nevertheless, the level of asset allocation moves up or down depending on the factor realization: optimal portfolio choice shifts up or down depending on the expected risk premium. When experimenting with a more volatile process, the portfolio movements become a lot more aggressive to take advantage of factor predictability[7]. Taken together, these findings make the case for enhanced TDFs (ETDFs) that condition on the market timing ability of the investor.

To emphasize this insight we make welfare comparisons across the Vanguard TDF recommendations, an i.i.d. stock returns model and the baseline mean reversion model. We show substantial welfare losses relative to the baseline and we show how these can arise either from lower mean consumption or higher consumption volatility over the life cycle when the incorrect portfolio rule is being used. Moreover, these losses are maximized at around age 50 when the increase in average wealth accumulation slows down and the net saving rate (the difference between labor income and consumption) turns negative. We therefore experiment with rules of thumb that alter the Vanguard recommendation depending on the expected equity premium and show how these rules of thumb (a proxy for ETDFs) reduce welfare less than following either the i.i.d. model or Vanguard's recommendation.

From all the underlying correlations studied, the main correlation that can quantitatively affect our conclusions is the correlation between permanent earnings shocks and the stock market innovation. Specifically, 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), a substantial hedging demand is generated. Financial advisors should therefore pay special attention to this correlation when devising rules of thumb about life cycle portfolio allocations.

We also experiment with introducing model uncertainty with regards to the underlying parameters driving the factor model and compute hedging demands by changing different correlations in the model. Model uncertainty with regards to the persistence parameter and the correlation between the factor and stock market innovation do not substantially alter the asset allocation and wealth accumulation profiles, and thereby do not alter the implications of the model with regards to ETDFs.

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, while Section 6 briefly discusses the implications of model parameter uncertainty. Section 7 discusses the implications of the model for lifestyle funds and Section 8 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

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

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

where $g(t, Z_{it})$ 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, $g(t, Z_{it})$. 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 ($\lambda = 0.68$) of the last working period permanent component of labor income.

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), and this fraction includes both rental and mortgage expenditures. This empirical process is taken from Gomes and Michaelides (2005) and is based on Panel Study Income Dynamics (PSID) data. We choose not to model explicitly the returns from housing following the empirical evidence (for instance, Cocco and Lopes (2015) and references therein) that households tend not to decumulate housing as fast as life-cycle models predict. A prominent explanation tends to be a psychological benefit from continuing to own one's house, an explanation that is consistent with the low observed demand for home equity conversion mortgages (Davidoff (forthcoming)). For these reasons we do not explicitly model the potential effects of housing returns, and focus instead only on investments of liquid financial wealth for rich households (that empirically tend to be both stockholders and homeowners).

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. The share of wealth in stocks (α_t) is defined as $\frac{S_t}{S_t+B_t}$.

2.4 Mean Reversion

We follow Campbell and Viceira (1999) and Pastor and Stambaugh (2012) 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 ($\phi > 0$) and negative correlation between the excess stock market return innovation (z_{t+1}) and the innovation to the factor (ε_{t+1}). 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

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

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:

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

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. Appendix A details the numerical solution technique and the numerical accuracy in the implementation of the Tauchen (1986) and 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 and also solves the model using the Tauchen (1986) numerical approximation approach, following the recommendations in Floden (2008) for very persistent processes.

2.6 Parameter Choice

Even though empirical predictability studies are typically done at a monthly or quarterly frequency, we solve the model at an annual frequency to maintain comparability with the life-cycle literature. 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. 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 relatively large estimate for the replacement ratio during retirement (68 percent of last working period labor income) arises from using both social security and private pension accounts to estimate the benefits in the PSID data and is consistent with not explicitly modelling tax-deferred saving through retirement accounts. Bagliano et. al. (2014) further discuss the implications from varying replacement rates.

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 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 $\gamma = \{2\}$ and $\psi = \{0.2\}$. We also set the discount factor (β) equal to 0.96.

To calibrate the stock market predictability parameters we use one of the more popular predictors of stock returns, namely the dividend yield. We estimate the equation above using three different data sources: the data from Pastor and Stambaugh (2012)[8], the publicly available data set on Robert Shiller's website[9], and decomposing the total return CRSP value weighted index into the capital gain and dividend return part. All of these data series are annual, as is the model frequency, and start at 1802, 1871 and 1926 respectively. All three estimations of the model generate a very persistent factor predicting the real log return on the U.S. equity market. This is relatively stable across different subperiods and we therefore set this persistence parameter at $\phi = .9$. The unconditional stock market volatility is given by the unconditional standard deviation of stock returns and is set equal to 0.18. A key parameter turns out to be the correlation between the factor and the return innovation ($\rho_{z,\varepsilon}$). For the Siegel and CRSP data sets this parameter is estimated at -0.6, while for the Shiller data set, we estimate it at around zero. Most estimates in the literature are towards the higher negative number (Campbell and Viceira (1999) for a quarterly estimation and Pastor and Stambaugh (2012) for both an annual and quarterly estimation). We therefore use $\rho_{z,\varepsilon} = -0.6$ for the baseline model but also experiment with the $\rho_{z,\varepsilon} = 0.0$ estimate, that also motivates a case with uncertainty about this correlation in what follows. The factor innovation is very smooth and we estimate (and use) $\sigma_{\varepsilon} = .007$ for the baseline model but we also find it useful to compare results with the case when the factor is more volatile ($\sigma_\varepsilon = .015$). This second case is interesting because it allows us to admit the possibility of negative expected returns (the factor can be negative), which could be a possibility for a small investor taking prices as given (and using a factor other than the dividend yield). The other benchmark parameters for the generation of stock market returns are $\mu = .04$ (an unconditional equity premium at 4%) as in most of the life-cycle literature. Given these estimates, we can infer that the unconditional variance of the stock market return innovation equals $\sigma_z^2 = 0.18^2 - \sigma_f^2$.

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 and we therefore set this correlation equal to zero. 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 ($\rho_{n,z}$) is set equal to 0.15, consistent with the mean estimates in most empirical work (Campbell et. al. (2002) or Davis, Kubler and Willen (2006)). Nevertheless, this can vary and be higher across heterogeneous occupations (Angerer and Lam (2009)) and/or workers (Bonaparte, Korniotis and Kumar (2014)) and we therefore experiment with values up to 0.5.

It is important to point out an issue that arises when approximating numerically persistent processes like the dividend yield, an issue that we discuss extensively in the numerical appendix. In numerically approximating very persistent autoregressive processes, a large number of grid points is needed to replicate key moments: the persistence, the conditional variance and the conditional (one-period ahead) equity risk premium. Moreover, the conditional expected risk premium must be positive, especially when the factor is denoting the dividend yield, a theoretical restriction that is not typically imposed in most estimation methods (a recent counter example is Pettenuzzo, Timmermann and Valkanov (2014)). Given the low estimated volatility of the factor, we impose the theoretical restriction that the conditional expected equity premium always be positive in the baseline model. To do so, we follow Floden (2008) and use Tauchen's (1986) approximation for the very persistent AR(1) case in the baseline model[10]. This allows us to explicitly control the range of the factor (thereby ensuring it is always positive) irrespective of the number of grid points chosen to replicate the chosen persistence and cross correlations. For the case with higher volatility that allows negative factor states we revert to the Tauchen and Hussey (1991) approximation[11].

3 Optimal Consumption and Portfolio Choice

3.1 Consumption and Portfolio Choice in Baseline Model

The consumption policy functions for ages 25, 55 and 75 are plotted in figure 1 (left column). 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, once the constraint stops binding, the saving reaction to the factor realization is not quantitatively important.

The portfolio policy functions for ages 25, 55 and 75 are plotted in figure 1 (right column). For the higher (lower) factor states that predict higher (lower) one-period ahead stock market returns, the investor increases (decreases) the share of wealth in the stock market, as long as the borrowing constraint is not binding ($\alpha = 1$). The optimal portfolio choice policy rule depends on factor realizations in an intuitive way but the presence of labor income prevents the portfolio from depicting the yo-yo type behavior found in Brennan, Schwartz and Lagnado (1997) where the portfolios move from zero to one depending on the factor realization. This happens in the baseline case because the factor realization is always constrained to be positive, a feature that arises from the low estimated variability of the dividend yield as a predictability factor.

Figure 2 plots the results from simulating the model starting with an initial wealth distribution assuming all households start from initial financial wealth equal to zero.[12] The model is simulated 2000 times for 500 individual life histories and the results report the averages from these simulations. Panel A is the standard life cycle wealth accumulation profile 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 below one, and the average profile generates a movement away from stocks as the household approaches retirement.

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). Figure 3 performs the simulations associated with changing these preference parameters. For this configuration of preference parameters, 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 3, panel C) due to the standard intuition of the lower implicit value of human capital in this model when wealth accumulation is higher. Having human capital as an implicit riskless asset, a higher wealth accumulation tilts portfolios away from stocks to replenish the lower value of riskless assets in the ovarall portfolio. This is not completely monotonic here, however, since averaging over different factor realizations can generate some non-monotonicities (as in the early part of the life cycle). Higher risk aversion also increases wealth accumulation and also generates safer portfolios (panel D) except in the early part of the life cycle where faster wealth accumulation for higher risk aversion implies earlier stock market participation and higher share of wealth in stocks (because liquidity constrained households are defined to have a zero share of wealth in stocks).

4 The Effects 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 differences between the i.i.d. model and the factor realizations are not quantitatively substantial, and therefore we focus our attention on the share of wealth in stocks. The policy functions for the share of wealth in stocks illustrate the large dependence of portfolios on the factor realization (Figure 4 for age 25 (Panel A), age 55 (Panel B) and age 75 (Panel C)). Portfolios can shift from 40% to 100% and vice versa depending on the factor realization and the household's age, something that does not happen in the i.i.d. model. The median factor state resembles closely the i.i.d. case, while higher factor realizations (higher expected stock returns) shift the share of wealth in stocks upwards. These policy function graphs illustrate clearly that portfolio allocations should be a lot more volatile in the mean reversion than in the i.i.d. model, while the range of outcomes is wider when households have longer horizons (comparing age 25 with age 55 or age 75, for example).

Figure 5 illustrates the differences in simulated profiles between the i.i.d. and the mean reversion model, while also introducing the life style recommendations issued by Vanguard (Donaldson et. al. (2013)) for target date funds (TDFs). We introduce Vanguard in anticipation of the discussion that follows on the optimality of TDF advice. Vanguard's basic recommendation is to invest 90% of a household's financial wealth in equities until age 40, and start decreasing that share as retirement approaches to reach 50% at age 65. To simulate wealth profiles for this case, we take the portfolio rule as exogenous but the household still makes optimal consumption-saving decisions, taking this portfolio decision as given.

Average wealth accumulation is similar across all three models (Figure 5, Panel A). This is slightly surprising since Panel B illustrates that the average share of wealth in stocks can differ substantially across models, even though average behavior still follows the life style fund intuition: reduce exposure to the stock market as retirement approaches. The fact that the average share of wealth in stocks is never one might be surprising given the results in the i.i.d. version of the model. 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 deviates from the 100% asset allocation to stocks.

The deviation from the i.i.d. model is even larger when the factor is perceived to be more volatile ($\sigma_\varepsilon = .015$). This is illustrated in Figure 5, Panel D: the average share of wealth in stocks is now substantially below one over the largest part of the life cycle. This reflects the fact that the household now expects negative expected risk premia for the lowest factor realizations, generating a full allocation in cash over this range.

The effect of the factor on portfolios can be more clearly seen by tracking individual portfolio simulations starting from different initial factor realizations. Figure 6 plots what happens over the lifecycle when starting from the lowest (state=1), sixth, tenth and highest (state=15) factor realization for the baseline case. Because the factor is persistent, it takes a substantial amount of time for a change to happen: when it does happen, the portfolio moves relatively quickly. Nevertheless, the change is different from the bang-bang movement in the share of wealth in stocks found in Brennan, Schwartz and Lagnado (1997) that features no labor income risk since the zero allocation in cash is never visited in these simulations. This happens in the baseline model because of the low volatility of the factor and the fact that the factor is always constrained to be positive. In the model with higher factor variability, the bang-bang movement arises once again.

Figure 7 presents the same information in the sense that the initial factor is different, but we then simulate over many different realizations of the factor. The average share of wealth in stocks is lower over most of the working life cycle for the simulations starting from the lower initial factor state rather than the ones starting from the higher factor realizations. Initial factor realizations therefore provide substantial differences across asset allocations, especially in the early part of the life cycle where horizon effects are more important.

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), we set it 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, 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 ($\rho_{n,\varepsilon}$). Bonaparte, Korniotis and Kumar (2014) 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. 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). Given that most factors predicting stock returns have very low volatility, we expect the correlation between the factor innovation and the much more volatile idiosyncratic earnings shocks to be close to zero. Nevertheless, we consider the potential effects of this correlation by increasing it to a high enough value that can simultaneously maintain the positive definiteness of the variance covariance matrix of the different innovations given the other chosen parameters. We therefore set $\rho_{n,\varepsilon}$ equal to 0.15 in our comparative statics experiments.

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. The figures that follow allow the reader to compute the hedging demand component that arises when the benchmark mean reversion model is simulated for a specific correlation (ρ) relative to the i.i.d. model by comparing the two simulated profiles and compute the percentage differences:

$$hedg(\rho) = 100 * \{\frac{\alpha_{\text{factor}}(\rho) - \alpha_{I.I.D.}(\rho)}{\alpha_{I.I.D.}(\rho)}\} \tag{6}$$

For space considerations we only report the simulated shares of wealth in stocks but (6) can be readily performed using the reported results and can be used to evaluate the hedging demand arising from the intertemporal variation relative to the model with i.i.d. returns under different parameterizations of the correlations of interest.

The wealth accumulation and mean shares of wealth in stocks over the life cycle are depicted in Figure 8, Panel A and Figure 8, Panel B, respectively. When the stock return innovation is set to zero ($\rho_{z,\varepsilon} = 0$) from $\rho_{z,\varepsilon} = -0.6$, the share of wealth in stocks is lower on average between 0 and five percent over different parts of the working life cycle (Panel B) and the effects are larger when the factor is more volatile (panel D), leading to a discernibly lower wealth accumulation (panel C). It can be shown that the variance of longer term stock returns is higher when $\rho_{z,\varepsilon} = 0$ rather than when $\rho_{z,\varepsilon} = -0.6$, therefore explaining the slight decrease in the share of wealth in stocks.

Changing the correlation between the factor innovation and the permanent income shock does not materially affect the average share of wealth in stocks (Panels B and D for the two cases), and therefore does not substantially affect average wealth accumulation.

What happens when the correlation between the permanent earnings shock and the stock market innovation is raised from 0.15 to 0.5? Figure 9 plots the average life cycle portfolio share allocation for the mean reversion model versus the i.i.d. model when this correlation is 0.5. There is a substantial difference across the life-cycle profiles both for the baseline case (Panel A) and the higher factor variability case (Panel B). Interestingly, the average share of wealth in stocks does not drop as much in the mean reversion model as it does in the i.i.d. model. This happens because the higher correlation does not push the share of wealth in stocks to zero for the highest factor realizations that signal high future expected stock returns.

6 Model Parameter Uncertainty

Model uncertainty is a feature in in the stock market predictability literature as emphasized, for example, in Xia (2001), Brandt et. al. (2005) and Pastor and Stambaugh (2012). We do not introduce learning in the model because this can be a substantial extension that requires either Kalman filtering or Bayesian updating techniques, extensions that we view as an interesting and challenging future research project in the context of life cycle models. We therefore take a more simplistic approach and assume that the investor observes a noisy signal of two key structural parameters, whereas the truth is generated with the baseline parameters analyzed above. We focus on the persistence of the factor (ϕ) and the correlation between the stock return and factor innovations ($\rho_{z\varepsilon}$).

For ϕ, we assume the investor expects three different values: the baseline value (0.9) with a 50 percent probability, a lower persistence (0.7) with 25 percent probability and a higher persistence (0.93) with a 25 percent probability. This is necessarily asymmetric to respect factor stationarity. For $\rho_{z\varepsilon}$ we assume the investor expects two different values: the baseline (-0.6) with a 50 percent probability and an extreme equal to zero with a 50 percent probability.

Figure 10 shows the results from introducing model uncertainty in these particular ways. Average portfolios are slightly more balanced over the life cycle (Panels B (for $\sigma_\varepsilon = 0.007$) and D (for $\sigma_\varepsilon = 0.015$)), especially for the model with the correlation uncertainty, generating also slightly lower wealth accumulation as a result (more clearly seen in Panel C). The effect is larger for the case where factor variability is larger (compare Panels D and B), but the quantitative average effects are not substantial relative to the no uncertainty cases.

7 Are Lifestyle funds Optimal?

7.1 Deviations from the I.I.D. model

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 factor realization. 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 that the investor follows the asset allocation implied by the same factor state throughout the lifecycle. This could arise from heterogeneous expectations about the underlying factor among different investors. As we have seen before, this implies that households faced with the lower factor realization should have lower exposure in the stock market. Figure 11 produces such a diagram by assuming different factor realizations starting from each of the 15 different initial states that persist throughout the life cycle. Panel B produces the same information by comparing the highest and lowest asset allocation experiences relative to the Vanguard recommendation and panels C and D repeat the same exercise for the higher factor variability case.

We can observe from these figures that the Vanguard life style recommendation is between the upper and lowest factor realization asset allocation profiles and that the deviation is even more pronounced when the factor is more volatile. Thus, the Vanguard recommendation is roughly correct on average but for an investor with market timing ability or expectations about time-varying expected risk premia, this recommendation will deviate substantially from the optimal portfolio choice. We next evaluate the welfare losses from ignoring stock market mean reversion.

7.2 Welfare Evaluations

To calculate welfare changes we use the value functions across different experiments. Given that we have solved for saving, portfolio choices and value functions for all periods in the life cycle, we know that the value functions at a particular age are a sufficient statistic for welfare effects. Let $v_o(x_{it}, f_t)$ be the value function for the benchmark model and $v_n(x_{it}, f_t)$ be the value function for a new model. We compute a measure of welfare change for a particular age group (age) as:

$$\overline{\mu}_{age} = \text{average of } \left[\left(\frac{v_n(x_{it}, f_t)}{v_0(x_{it}, f_t)} \right)^{\frac{1}{1-\gamma}} - 1 \right] \text{ for all } i \in I_{age} \text{ and all factor states.} \tag{7}$$

This is the unconditional (across factor states) certainty consumption equivalent because we convert the change of the value into the dimension of expenditure before taking the average.

Figure 12 plots the life cycle certainty equivalents in percent when returns are simulated based on the mean reversion model and the comparison is between the mean reversion and the i.i.d. model and between the mean reversion and Vanguard recommendation. Panel A reports the results for the baseline model ($\sigma_\varepsilon = 0.007$) and Panel B for the higher factor volatility model ($\sigma_\varepsilon = 0.015$). Panel A illustrates substantial welfare losses from following the i.i.d. model relative to the optimal portfolio rule in the presence of the factor, and the welfare losses are even more substantial when following the Vanguard recommendation. This arises naturally given that the deviations of the average portfolio allocations are even larger between the Vanguard recommendation and the factor model than the ones between the i.i.d. and factor model (Figure 5, Panel B). With higher factor volatility (Panel B) these effects are magnified as the differences across the average portfolio shares further diverge (Figure

5, Panel D).

We make two observations based on the results in Figure 12. First, the welfare losses are economically significant: they can reach two to seven percent of consumption equivalents depending on the specification, and this represents a substantial welfare loss. Second, the losses tend to get maximized at around age 50, whereas average wealth accumulation is maximized at the exogenous retirement age (65). What can explain these findings?

To better understand these welfare shapes and magnitudes, it is helpful to recall that given the preference for consumption smoothing, welfare in the model is increasing in consumption and decreasing in the variability of consumption. We can therefore gain an insight on where the welfare differences are coming from by investigating how average consumption and consumption inequality evolve over the life cycle across the three models. To do so we compute the average consumption and standard deviation of consumption as a cohort ages. Mean consumption is very similar over the life cycle across the three models in a life cycle graph. To emphasize the differences we therefore report the percentage difference between the profiles generated by the baseline versus the i.i.d. model and the baseline versus the Vanguard recommendation. These differences are plotted in Figure 13 for the baseline case ($\sigma_\varepsilon = 0.007$ for Panel A and Panel B) and for the higher factor volatility case (Panel C and D).

The Vanguard recommendation generates substantially more volatility in consumption over the working part of the life cycle (Panel B and D), despite generating mildly higher average consumption in the baseline case (Panel A). Given the preferences for smoother consumption, this increased consumption inequality translates into a welfare loss that essentially gets maximized at mid life (around age 50), justifying the peak in welfare loss depicted in

Figure 12.

The i.i.d. model on the other hand generates lower mean consumption but lower volatility over the life cycle. The biggest welfare loss is therefore coming from lower mean consumption and this is even more pronounced in the case where the factor is more volatile (compare Panel C with Panel A, for example). Consumption variability is actually lower with the i.i.d. model since portfolio rules are a lot more stable, and this is especially so for the more volatile factor (Panel D relative to Panel B).

The question remains as to why welfare losses are maximized at around age 50. This happens because the saving rate (defined as the percent of labor income that is saved) turns negative at around that age. This is driven by the hump shape in the labor income process and the fact that the rate of average wealth accumulation begins to slow down after that age as average labor income begins to fall. Given the reduction in the saving rate, the welfare loss from following an imperfect portfolio rule is reduced.

7.3 Rules of Thumb

Can a rule of thumb be devised in the presence of factor predictability? Based on our results we devise approximate rules that are similar to Vanguard's but are conditional on the factor realization for the baseline case: they are more aggressive for higher factor realizations and less aggressive for lower ones. For the baseline case, and relative to the Vanguard recommendation, the rule of thumb is to decrease the portfolio allocation by around 50 percentage points (that is from a 90% to a 40% allocation in stocks) until retirement is reached, if the investor expects equity premia to be lower than 2%. On the other hand, the asset allocation should be increased to 100% when the factor realization is above 5%. We think a richer model of retirement uncertainties is needed to make similar recommendations for the retirement period. For the high variability case, the rule of thumb is taken to the extreme of having full allocation to stocks for the sufficiently high factor realizations and full allocation to cash for the sufficiently low factor realizations. This behavior can reduce the welfare losses relative to the ones shown in Figure 13 for the i.i.d. case and the Vanguard recommendation.

We can also provide a rule of thumb of the saving rate (the proportion of labor income that is saved). Depending on when the household begins saving for retirement, the model predicts a rule of thumb of around ten percent of labor income saved for thirty years of working life to achieve adequate consumption smoothing, regardless of the factor variability. Implementing these rules of thumb lowers the welfare loss relative to both the i.i.d. and Vanguard models, assuming stock returns are generated through the mean reversion process.

8 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. Optimal portfolio choice is shown to be heavily dependent on the realization of the factor predicting future returns. In the baseline case where the factor is always positive, the share of wealth in stocks is a parallel shift of the i.i.d. model, and the extent of the shift depends on the factor realization and the degree that short sale constraints bind. When the factor is more volatile and can take negative expected values, portfolio holdings will very often be either completely allocated in the stock market or in the riskless asset market, consistent with Barberis (2000) and Brennan, Schwartz and Lagnado (1997). The large welfare losses from failing to condition on market information support the case for enhanced TDFs (ETDFs) that condition asset allocation on expected risk premia.

Future directions of research include the explicit introduction of tax deferred retirement accounts (for the i.i.d. case, see Gomes, Michaelides and Polkovnicheno (2009)), an explicit treatment of capital gains taxation (in the spirit of Dammon, Spatt and Zhang (2004)), an explicit learning mechanism about the true underlying model through either a Kalman filtering or Bayesian learning approach (Brandt et. al. (2005) and Pastor and Stambaugh (2012)), an explicit introduction of housing (Cocco and Lopes (2015)) and introducing time-varying volatility and risk aversion (through a stochastic discount factor, for example). All these extensions will require additional computational power to achieve the desired required solution accuracy but will further improve our understanding of life cycle portolio choice under uncertainty and offer scientific advice to billions of households increasingly making their own individual financial decisions.

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) and Tauchen (1986) 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).

In numerically approximating very persistent autoregressive processes like the dividend yield, a large number of grid points is needed to replicate key moments: the persistence, the conditional variance and the conditional risk premium. Moreover, the conditional expected risk premium must be positive, a theoretical restriction that is not typically imposed in most estimation methods (a recent counter example is Pettenuzzo, Timmermann and Valkanov (2014)). To determine the number of quadrature points needed to accurately capture the dynamics of the estimated VAR we pay attention to satisfy two important constraints. First, we ensure that the numerical approximation method respects the theoretical restriction that the conditional expected equity premium be positive. Second, that the model can be numerically solved without causing a computational intractability arising from the curse of dimensionality.

Respecting these constraints using the Tauchen and Hussey (1991) method is not easy. When a large number of grid points is used to accurately replicate the persistence of the process, the quadrature points determined by the method become wider and can generate a negative expected, one-period ahead, equity premium. This typically happens at the edges of the grids where the approximation is poorer. For the cases that this problem arises, we revert to the older Tauchen (1986) procedure that allows the researcher to determine the range of quadrature points. This allows the researcher to ensure the conditional equity premium is positive for all factor states.

To determine the number of grid points that satisfy these requirements, we rely both on simulation and on a numerical integration method to compute the key conditional moments that are relevant for the household's optimization problem. A detailed appendix with various experiments is available upon request. We next report some of these results.

We report extensive experiments with the Tauchen (1986) 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 for the baseline case ($\sigma_\varepsilon = 0.007$):

$$\begin{bmatrix} f_{t+1} \\ r_{t+1} \\ \ln N_{t+1} \\ \ln U_{t+1} \end{bmatrix} = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix} + \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} f_t \\ r_t \\ \ln N_t \\ \ln U_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \\ n_{t+1} \\ u_{t+1} \end{bmatrix}$$

Denote the coefficient matrix as follows:

$$A1 = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix}, \; A2 = \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

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

$$\Omega = \begin{bmatrix} 0.000049 & -0.000753 & 0 & 0 \\ & 0.03214 & 0 & 0.0027 \\ & & 0.01 & 0 \\ & & & 0.01 \end{bmatrix} \text{ and } \rho = \begin{bmatrix} 1 & -0.6 & 0 & 0 \\ & 1 & 0 & 0.15 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}.$$

To test the accuracy of the approximation method, we use two different methods. The first method is based on simulation. We simulate based on a Tauchen (1986) discretization for a given number of grid points and then perform a Monte Carlo analysis to investigate how close the estimated parameters are to the actual parameters used to generate the discrete approximation. Specifically, we generate 100 simulation paths and each simulation path has 52,000 periods. In each simulation, the first 2000 periods are discarded. The VAR is then estimated in each of 100 times and the coefficients averaged and reported below. The numbers in parentheses are the standard deviations of the reported average coefficients. Since the means of can always be estimated correctly in all the experiments, we only report the VAR coefficient averages.

The second method uses the quadrature nodes and weights used in the numerical solution to compute the conditional expected equity risk premium (which should equal the factor state) and the conditional variance of the risk premium. The conditional correlation matrix is also computed and reported. These values should be close to the simulations reported in the first method, and should also be internally consistent. For example, the realization of the factor should equal the expected risk premium conditional on the current factor realization.

The internet appendix contains many different experiments, we report here the main experiment that led us to choose 15 grid points for f_t, 20 points for r_t, 5 points for $\ln(N_t)$ and 3 points for $\ln(U_t)$. The average VAR coefficients and variance covariance matrix are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0001)}{0.00465} \\ \underset{(0.005)}{0.0209} \\ \underset{(0.002)}{-0.007} \\ \underset{(0.003)}{-0.0048} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.004)}{0.884} & \underset{(0.00043)}{-0.00005} & \underset{(0.0006)}{0.00014} & \underset{(0.00069)}{0.00016} \\ \underset{(0.11)}{0.977} & \underset{(0.012)}{0.0003} & \underset{(0.018)}{-0.0021} & \underset{(0.016)}{-0.0042} \\ \underset{(0.06)}{-0.053} & \underset{(0.0006)}{0.0001} & \underset{(0.01)}{-0.0014} & \underset{(0.01)}{0.0018} \\ \underset{(0.07)}{0.002} & \underset{(0.0068)}{0.0003} & \underset{(0.011)}{-0.0023} & \underset{(0.01)}{-0.0008} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000006)}{0.0000489} & \underset{(0.000001)}{-0.0007} & \underset{(0.000007)}{-0.00000082} & \underset{(0.000007)}{-0.000001} \\ & \underset{(0.00047)}{0.00313} & \underset{(0.00014)}{0.0026} & \underset{(0.00015)}{0.00002} \\ & & \underset{(0.00012)}{0.01} & \underset{(0.0001)}{-0.000028} \\ & & & \underset{(0.0001)}{0.01} \end{bmatrix}$$

The conditional moments for the first, seventh and 15th factors are (the approximation is worst for the states at the edges)

f_t	$E_{f_t}(RP_{t+1})$	$Var_{f_t}(RP_{t+1})$
0.007882	0.00929	0.019
0.030823	0.0334	0.033
0.07212	0.0710	0.0193

The same conditional moments using the numerical integration approach (used in the value function solution) are:

f_t	$E_{f_t}(RP_{t+1})$	$Var_{f_t}(RP_{t+1})$
0.007882	0.0102	0.019
0.030823	0.0309	0.0327
0.07212	0.0698	0.019

.

When using the Tauchen-Hussey (1991) procedure instead of the Tauchen (1986) one for the same parameterization, the results are

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.00413} \\ \underset{(0.005)}{0.0174} \\ \underset{(0.003)}{-0.0056} \\ \underset{(0.002)}{-0.0051} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.005)}{0.897} & \underset{(0.0003)}{0.00004} & \underset{(0.0007)}{0.00006} & \underset{(0.00068)}{-0.0001} \\ \underset{(0.11)}{1.06} & \underset{(0.01)}{-0.0004} & \underset{(0.018)}{-0.003} & \underset{(0.018)}{0.0023} \\ \underset{(0.07)}{0.0155} & \underset{(0.007)}{0.00054} & \underset{(0.01)}{-0.0033} & \underset{(0.01)}{0.0002} \\ \underset{(0.06)}{-0.0004} & \underset{(0.006)}{-0.0005} & \underset{(0.011)}{-0.0019} & \underset{(0.01)}{0.0029} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.0000489} & \underset{(0.0000014)}{-0.000748} & \underset{(0.000008)}{0.0000005} & \underset{(0.000008)}{0.0000002} \\ & \underset{(0.00044)}{0.0032} & \underset{(0.00018)}{0.0027} & \underset{(0.00018)}{-0.00002} \\ & & \underset{(0.00013)}{0.01} & \underset{(0.0001)}{-0.000005} \\ & & & \underset{(0.00012)}{0.01} \end{bmatrix}$$

The conditional moments for the first, seventh and 15th factors are (the approximation is worst for the states at the edges, that is, the first and 15th state)

f_t	$E_{f_t}(RP_{t+1})$	$Var_{f_t}(RP_{t+1})$
-0.0045	-0.0212	0.02845
0.03441	0.03396	0.0326
0.08455	0.1	0.02812

.

The same conditional moments using the numerical integration approach (used in the value function solution) are:

f_t	$E_{f_t}(RP_{t+1})$	$Var_{f_t}(RP_{t+1})$
-0.0045	-0.0212	0.0287
0.03441	0.03441	0.03214
0.08455	0.1012	0.02867

.

It is because the Tauchen-Hussey (1991) method produces a negative factor when a large number of grid points is used that we use the Tauchen (1986) method for the baseline case. When the volatility of the factor increases ($\sigma_\varepsilon = 0.015$) the results are very similar across the two methods and we revert to using the Tauchen-Hussey (1991) approximation.

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

$$v_{it}(x_{it}, f_j) = \Bigg[\underset{\{c_{it},\alpha_{it}\}}{MAX} (1-\beta) c_{it}^{1-1/\psi} + \beta \left\{ \sum_{k,U,N,R} \pi_{j,kUNR} \left[\begin{array}{c} p_{t+1}(v_{it+1}(x_{it+1}, f_k))^{1-\gamma} (Y_{it+1}^p / Y_{it}^p)^{1-\gamma} + \\ b(1-p_{t+1})(x_{it+1})^{1-\gamma} (Y_{it+1}^p / Y_{it}^p)^{1-\gamma} \end{array} \right] \right\}^{\frac{1-1/\psi}{1-\gamma}} \Bigg]^{\frac{1}{1-1/\psi}}$$

while $\pi_{j,kUNR}$ is the transition probability matrix given that the current factor is f_j. We discretize the state variable x by dividing it into 151 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 ($E_t v^j_{t+1}(.)$), which we can compute since we have just obtained v^j_{t+1}. 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 2000 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] Brandt, M.W., A. Goyal, P. Santa-Clara, and J.R. Stroud, 2005, "A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning about Return Predictability," *Review of Financial Studies*, 18, 831-873.

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

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

[13] 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.

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

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

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

[17] 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.

[18] 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.

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

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

[21] —— 1998, "How important is precautionary saving?" *Review of Economics and Statistics,* 80 no. 3: 410-19.

[22] Cocco, Joao and Paula Lopes, 2015, "Reverse Mortgage Design," LBS/LSE working paper.

[23] 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.

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

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

[26] Dammon, Robert M., Chester S. Spatt, and Harold Zhang, 2004, Optimal asset location and allocation with taxable and tax-deferred investing, *Journal of Finance,* 59 (3), 999–1037.

[27] Davidoff, Thomas, “Can High Costs Justify Weak Demand for the Home Equity Conversion Mortgage?” Forthcoming, *Review of Financial Studies.*

[28] 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.

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

[30] Donaldson, Scott, Francis Kinniry, Roger Aliaga-Diaz, Andrew Patterson, Michael DiJoseph, 2013, “Vanguard’s Approach to Target-Date Funds.”

[31] 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.

[32] Floden, Martin, 2008, "A Note on the Accuracy of Markov-chain approximations to highly persistent AR(1) processes," *Economics Letters*, 99, 516-20.

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

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

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

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

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

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

[39] 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.

[40] —— 1997. "Market Frictions, Savings Behavior, and Portfolio Choice," *Macroeconomic Dynamics* 1: 76-101.

[41] —— 2000. "Portfolio Choice in the Presence of Background Risk," *The Economic Journal*, 110, 1-26.

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

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

[44] 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.

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

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

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

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

[49] —— 1971, "Optimum Consumption and Portfolio Rules in a Continuous Time Model," *Journal of Economic Theory*, 3, 373-413.

[50] —— 1973, "An Intertemporal Capital Asset Pricing Model," *Econometrica*, 41, 867-87.

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

[52] Pastor, Lubos and Robert F. Stambaugh, 2012, "Are Stocks Really Less Volatile in the Long Run?" *The Journal of Finance*, Vol. LXVII, 2, 431-477.

[53] Pettenuzzo, Davide, Allan Timmermann and Rossen Valkanov, 2014, "Forecasting Stock Returns Under Economic Constraints," *Journal of Financial Economics*, 114 (3), 517-553.

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

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

[56] Siegel, Jeremy, 2008, *Stocks for the Long Run.* McGraw Hill, New York, NY.

[57] Tauchen George, 1986, "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," *Economics Letters*, 20, 177-181.

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

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

[60] 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.

[61] 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.

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

[63] 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] 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).

[3] Gollier and Pratt (1996), Heaton and Lucas (1996, 1997, 2000), Viceira (2001), Gomes and Michaelides (2005), Cocco, Gomes and Maenhout (2005) and Haliassos and Michaelides (2003).

[4] 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.

[5]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)).

[6]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).

[7]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.

[8]We thank Lubos Pastor for kindly providing this data set, which is based on Siegel (2008).

[9]At http://www.econ.yale.edu/~shiller/data.htm.

[10]When discretizing the factor with 15 states using the Tauchen (1986) method with $\sigma_\varepsilon = 0.007$, the actual factor states are 0.0079, 0.0125, 0.0171, 0.0216, 0.0262, 0.0308, 0.0354, 0.0400, 0.0446, 0.0492, 0.0538, 0.0584, 0.0629, 0.0675, 0.0721.

[11]When discretizing the factor with 15 states using the Tauchen and Hussey (1991) method with $\sigma_\varepsilon = 0.015$, the actual factor states are -0.0555, -0.0379, -0.0229, -0.0093, 0.0035, 0.0159, 0.0280, 0.0400, 0.0520, 0.0641, 0.0765, 0.0893, 0.1029, 0.1179, 0.1355.

[12]In previous versions of the paper we used the initial wealth distribution for stock market participants from the SCF in 2001 but given that this paper is more normative in nature we think it is instructive to understand what happens when starting all households from the same (zero) initial financial wealth.

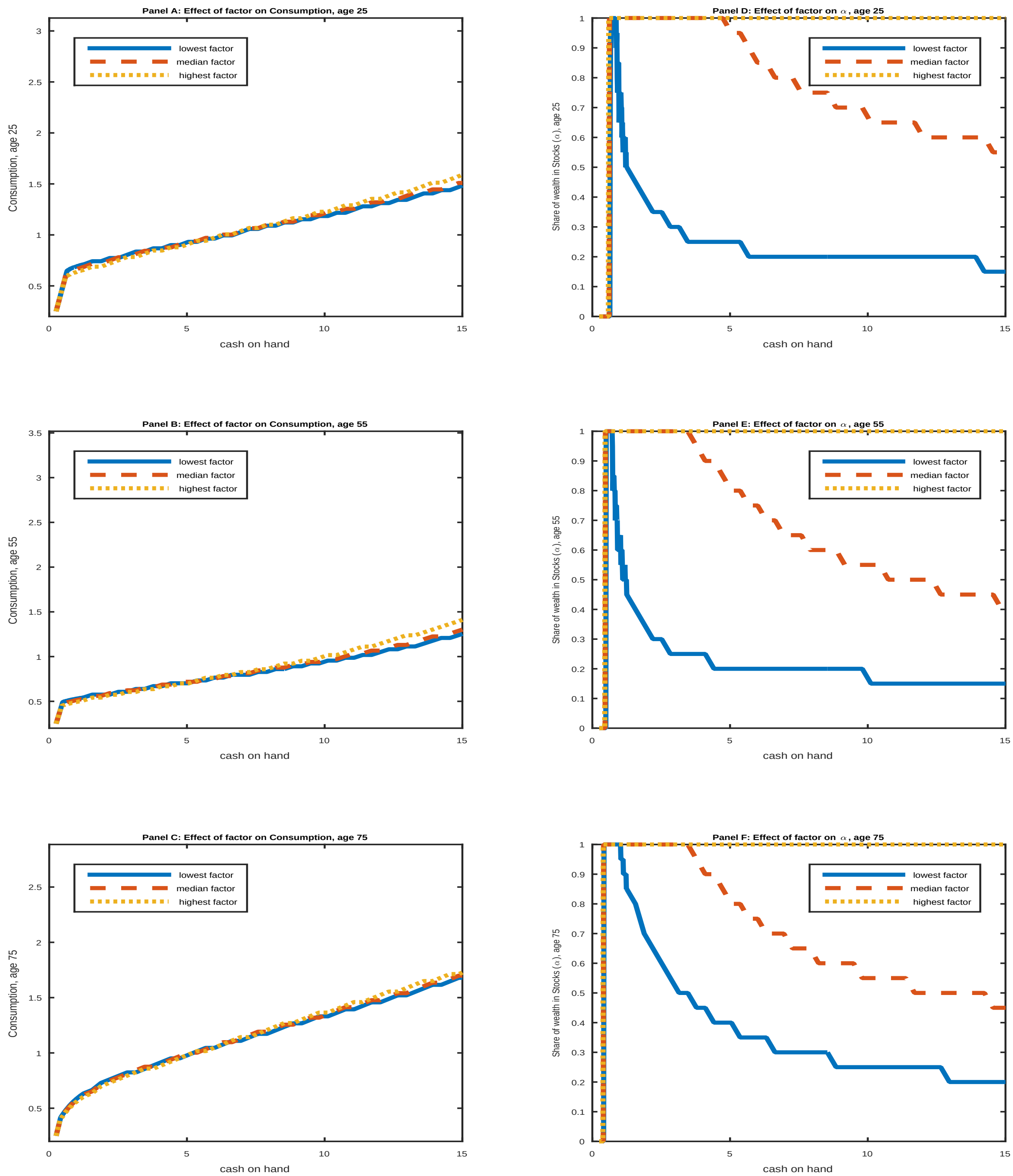


Figure 1: Policy functions for baseline case. Risk aversion (γ) is 5, the elasticity of intertemporal substitution (ψ) is 0.5, the discount factor (β) is 0.96 and the persistence of the factor is 0.9.

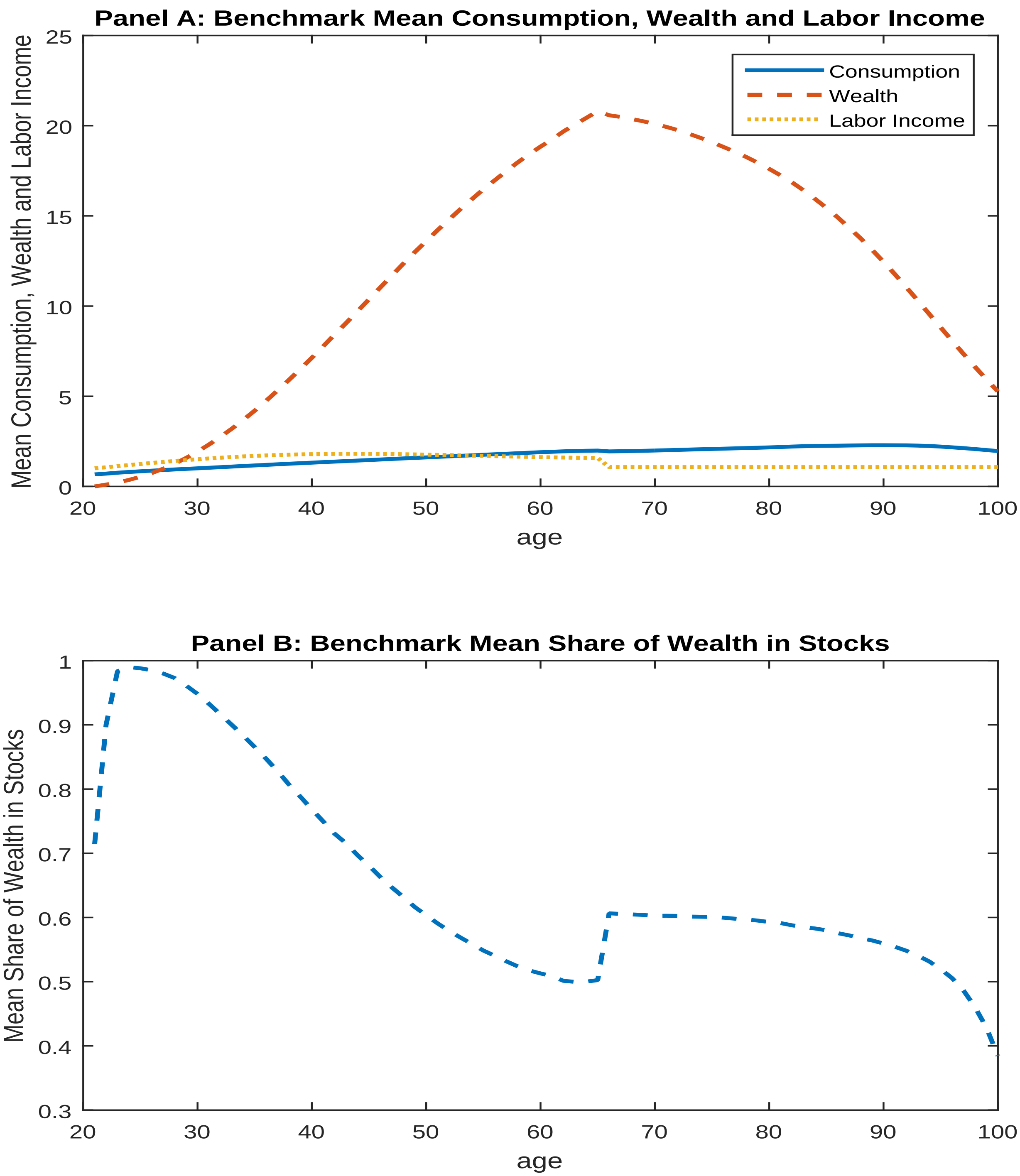


Figure 2: Life-cycle profiles for the baseline case. Risk aversion (γ) is 5, the elasticity of intertemporal substitution (ψ) is 0.5, the discount factor (β) is 0.96 and the persistence of the factor is 0.9.

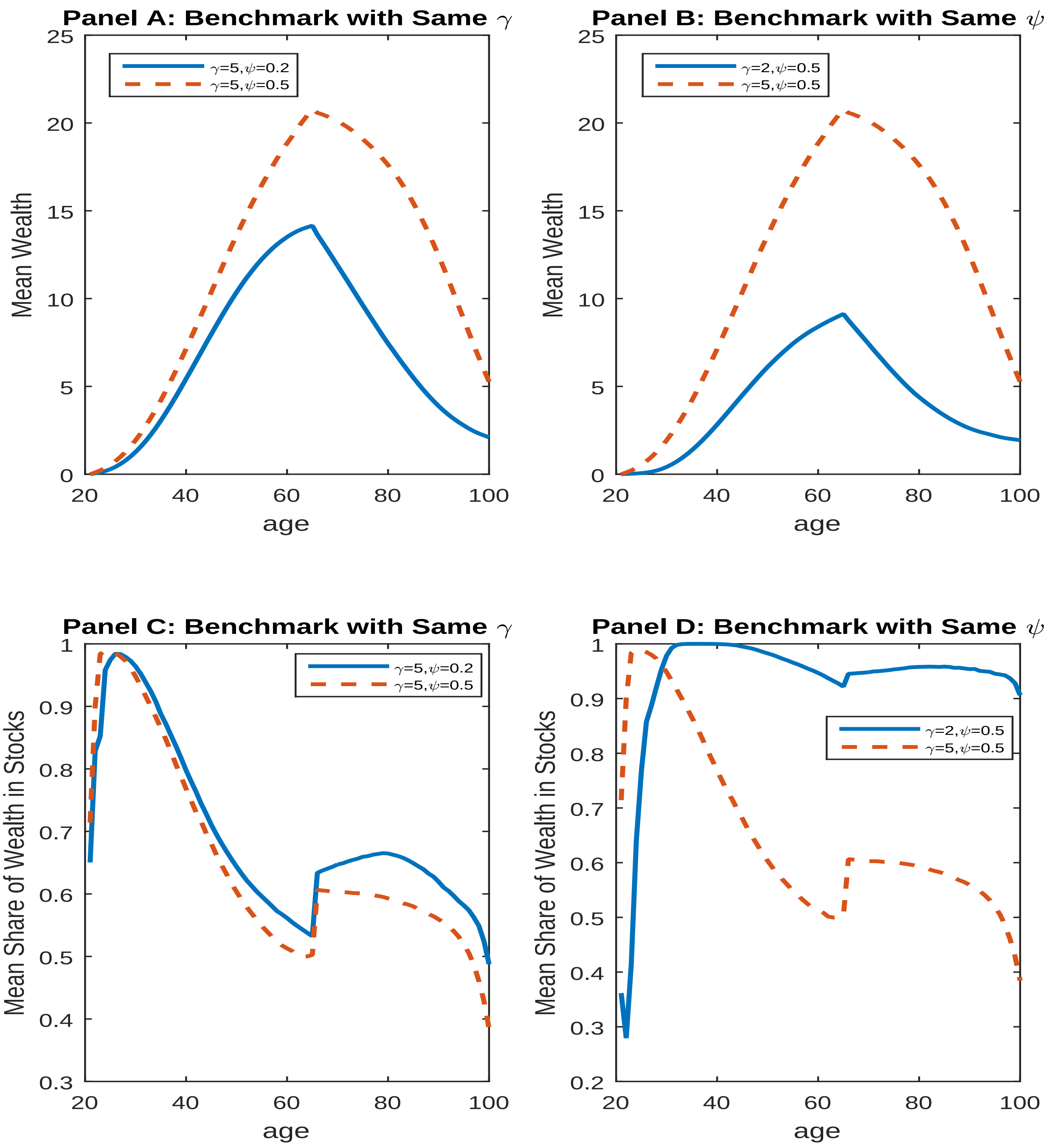


Figure 3: Life-cycle profiles for different preference parameters. Panels A and C hold risk aversion (γ) constant and vary the elasticity of intertemporal substitution (ψ). Panels B and D vary risk aversion (γ) and hold the elasticity of intertemporal substitution (ψ) constant.

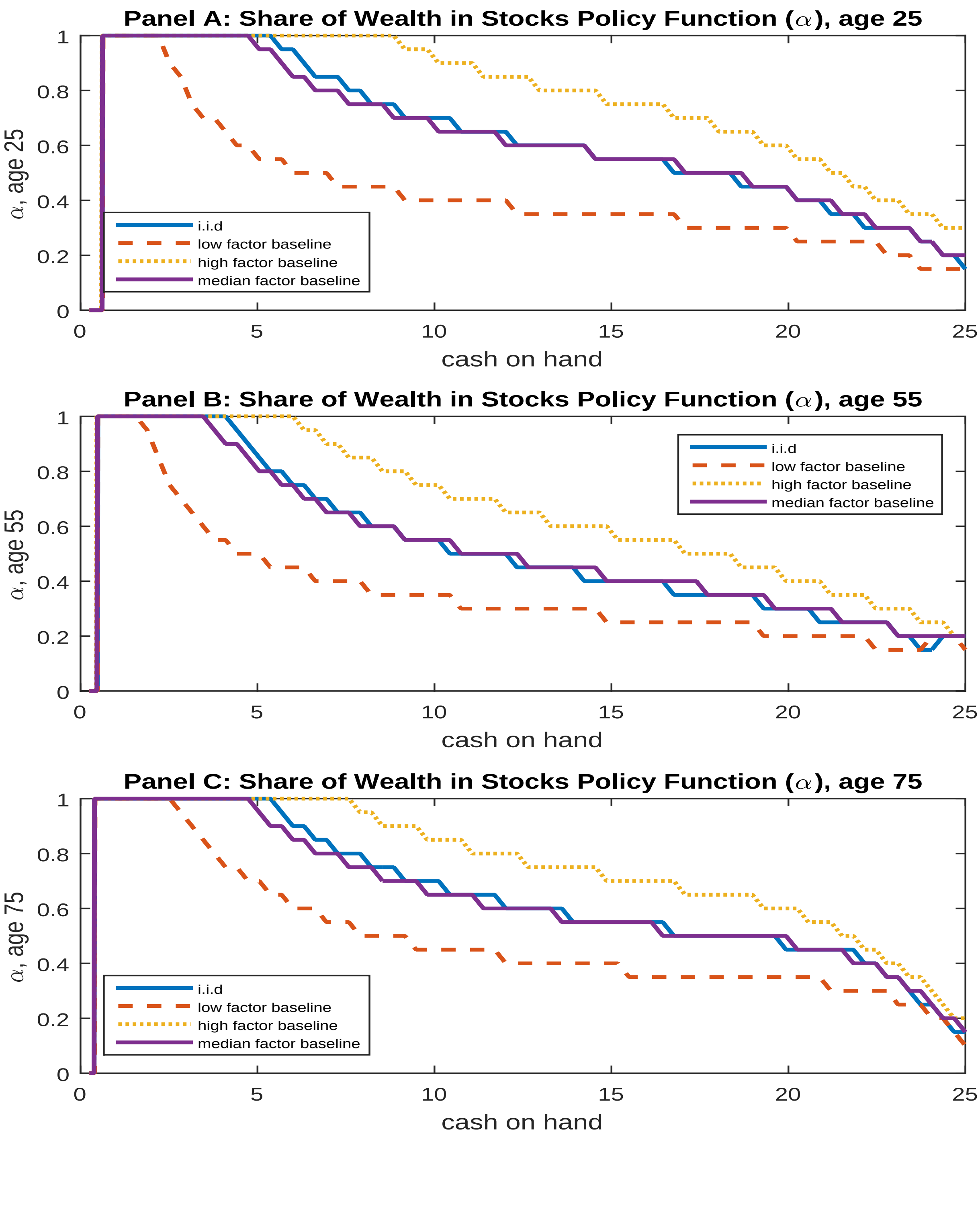


Figure 4: Policy function comparison between benchmark stock market mean reversion results and the I.I.D. stock returns case.

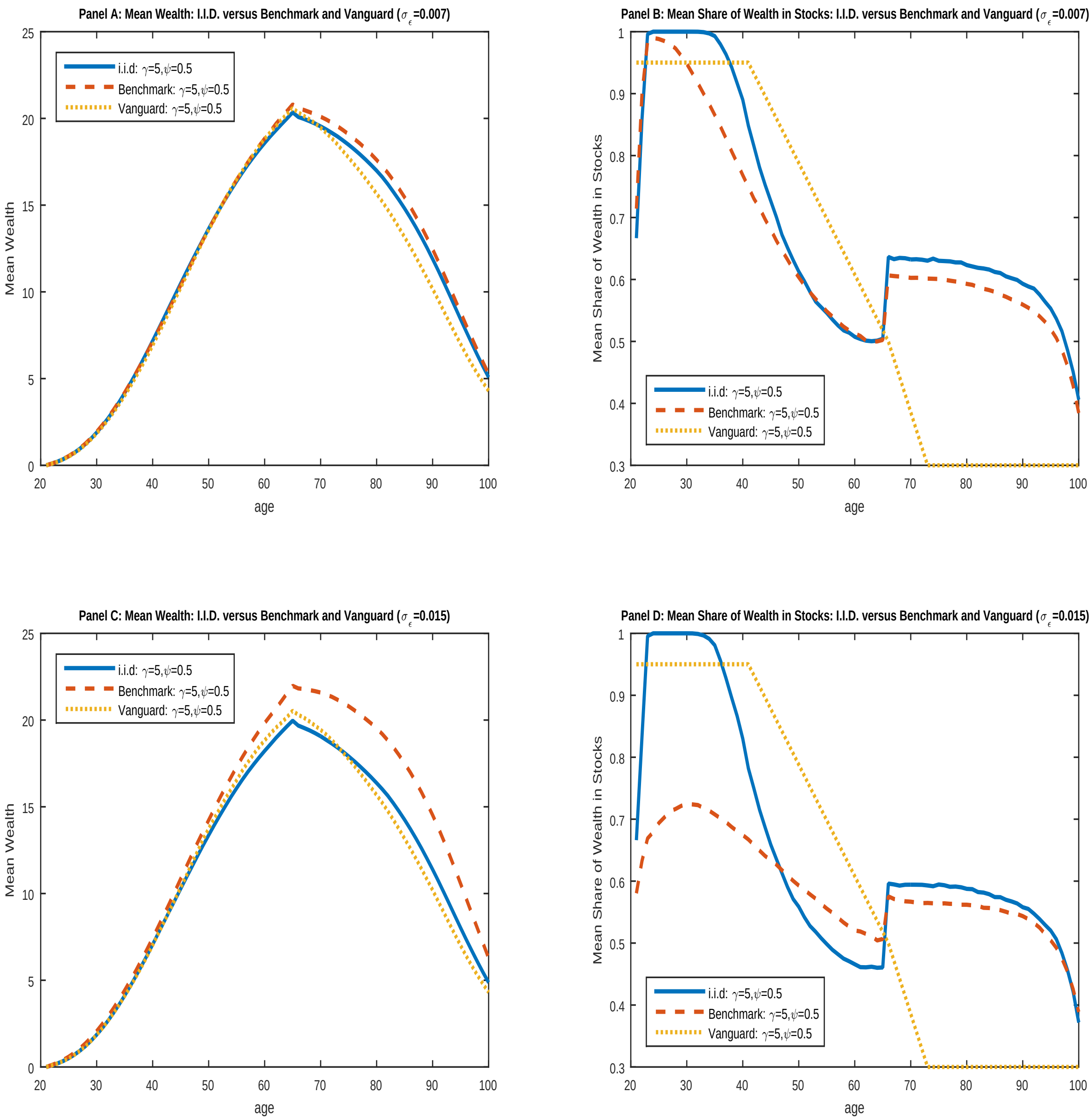


Figure 5: Life-cycle profile comparison between benchmark stock market mean reversion results with Vanguard recommendation and the I.I.D. stock returns model.

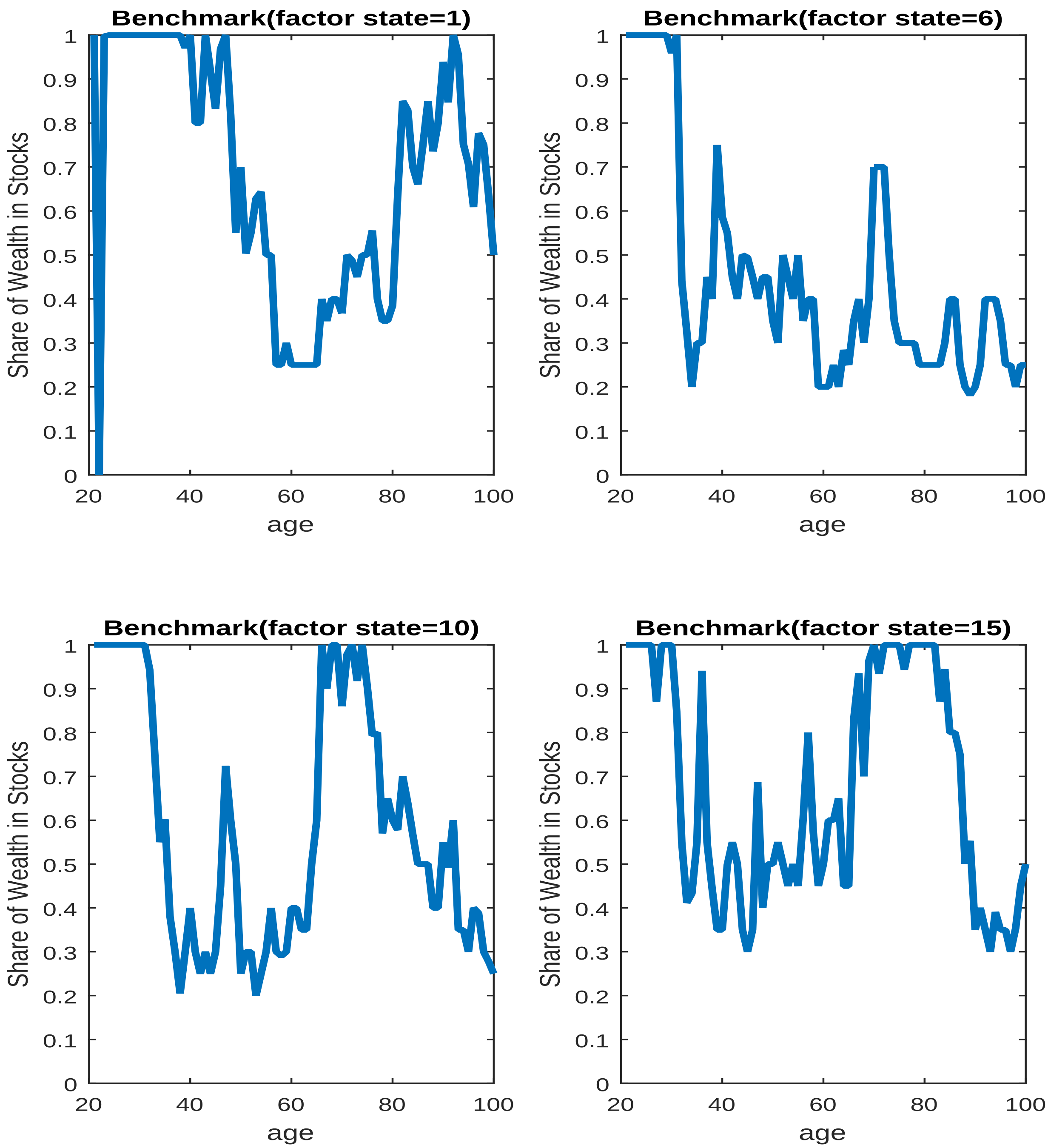


Figure 6: Life-cycle individual share of wealth in stocks for different initial factors starting from age 21 for the benchmark stock market mean reversion model.

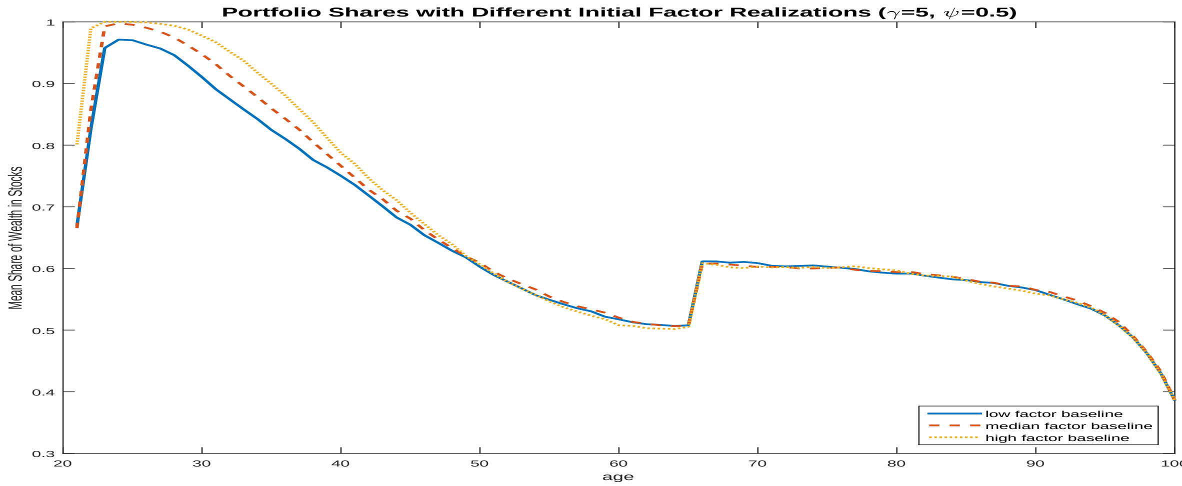


Figure 7: Life-cycle average share of wealth in stocks for different initial factors starting from age 21 for the benchmark stock market mean reversion model.

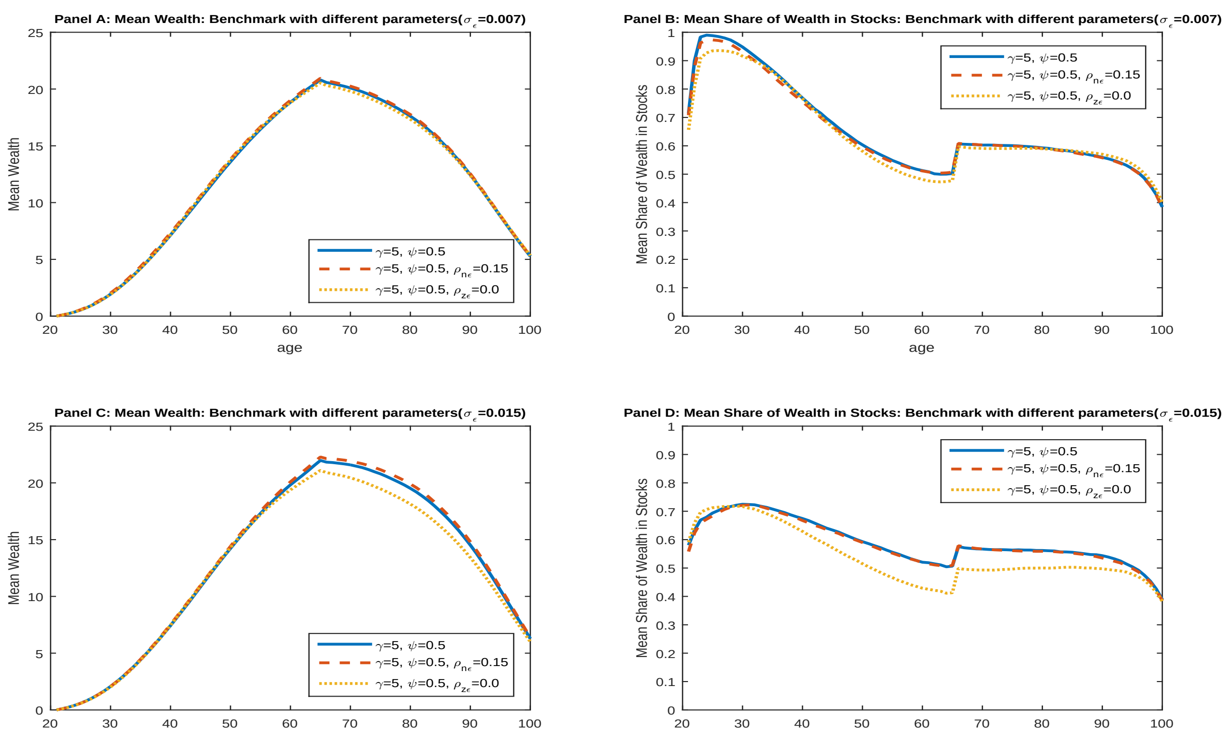
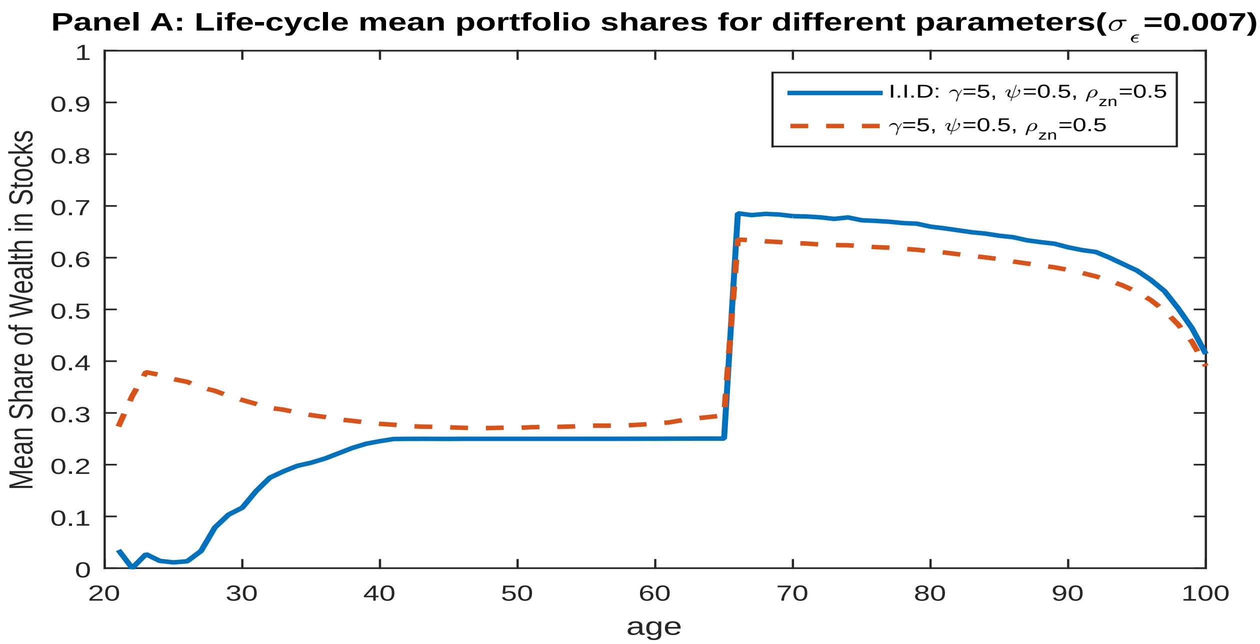


Figure 8: Life-cycle profiles for average wealth and portfolio shares for different parameters. The three cases are (i) the benchmark case ($\gamma = 5$ and $\psi = 0.5$) (ii) changing the correlation between permanent labor income shocks and the factor innovations ($\rho_{n\varepsilon}$) from 0.0 to 0.15 and (iii) changing the correlation between the stock return shocks and factor innovations ($\rho_{z\varepsilon}$) from -0.6 to 0.0



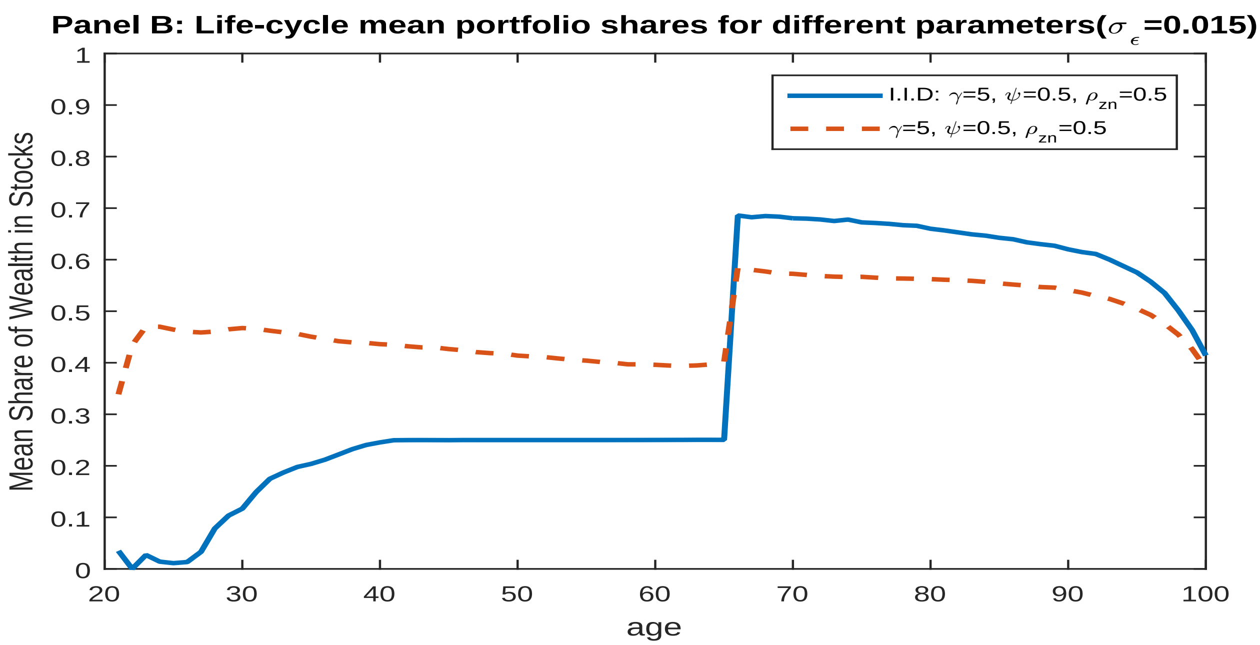


Figure 9: Life-Cycle mean portfolio shares in stocks. The average life-cycle portfolio shares in stocks for the stock market mean reversion model and the I.I.D. stock returns model when the correlation between the permanent earnings shocks and the stock market innovations is 0.5 instead of 0.15.

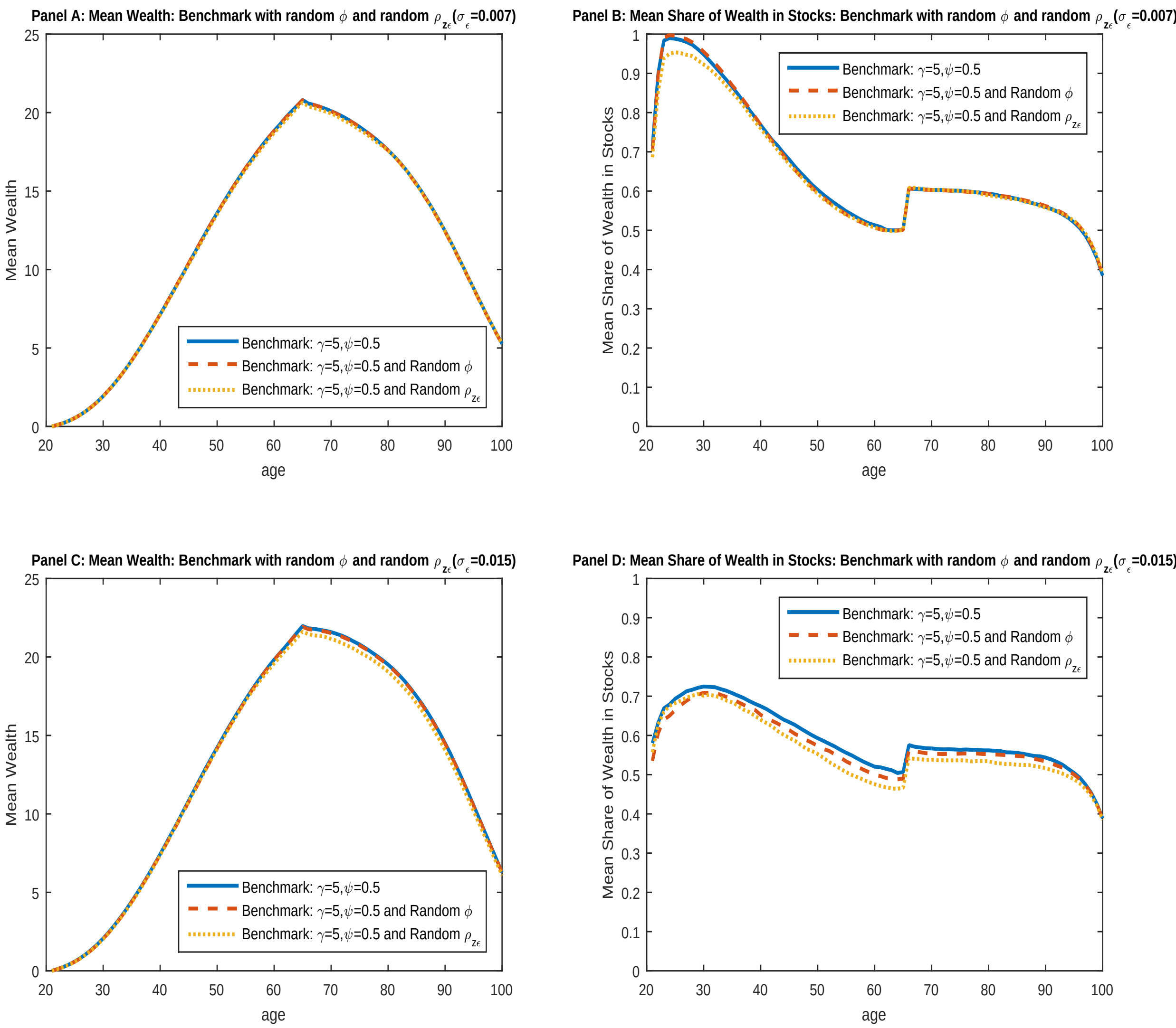


Figure 10: Life-cycle profiles for average wealth and portfolio shares for different parameters. The three cases are (i) the benchmark case ($\gamma = 5$ and $\psi = 0.5$) (ii) the benchmark case with uncertain ϕ and (iii) the benchmark case with uncertain $\rho_{z\varepsilon}$.

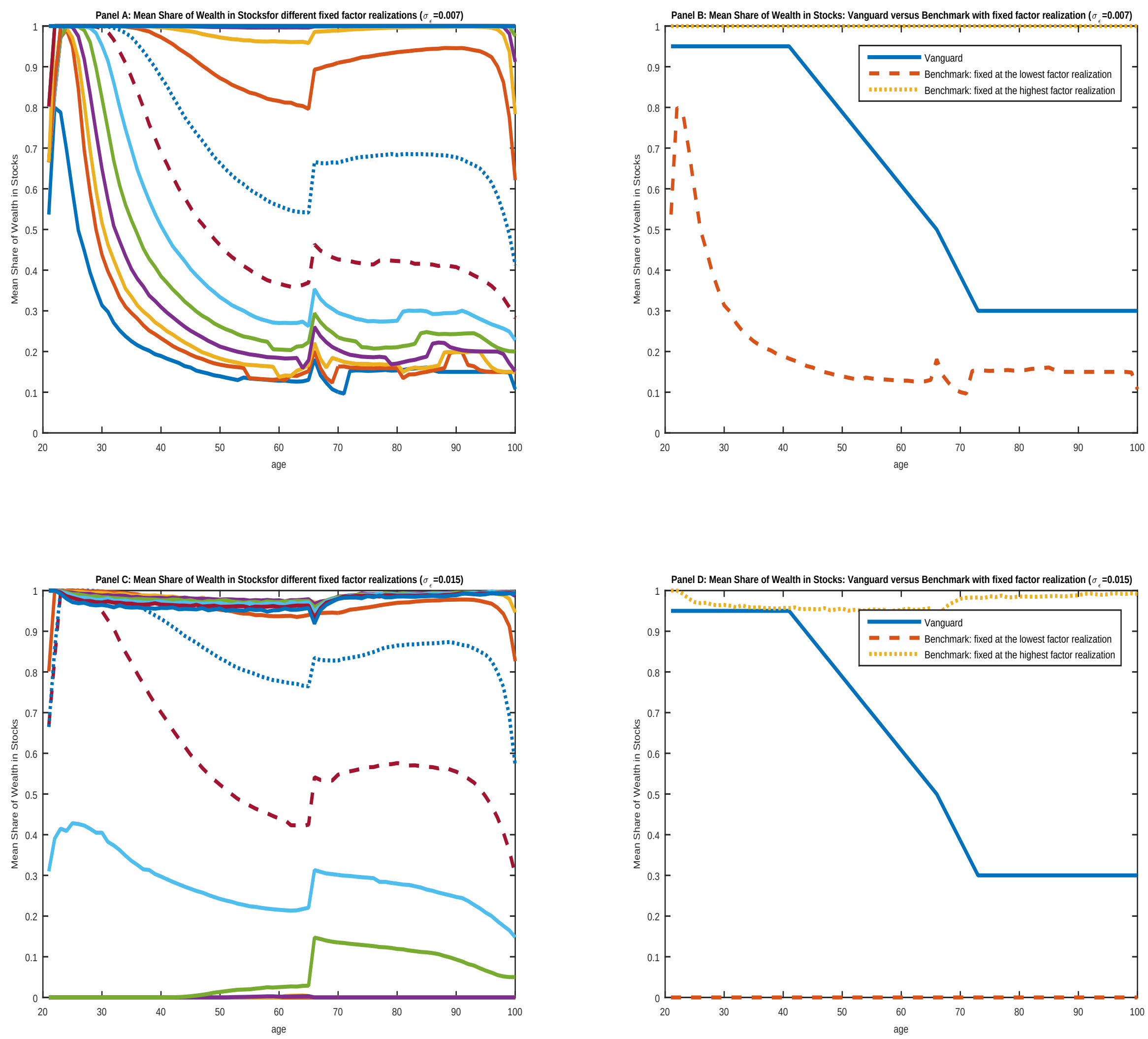


Figure 11: Portfolio shares with different constant factor realizations for the stock market mean reversion model. Panels B and D report the higest and lowest states and also report the Vanguard recommendation that lies in between the two extremes.

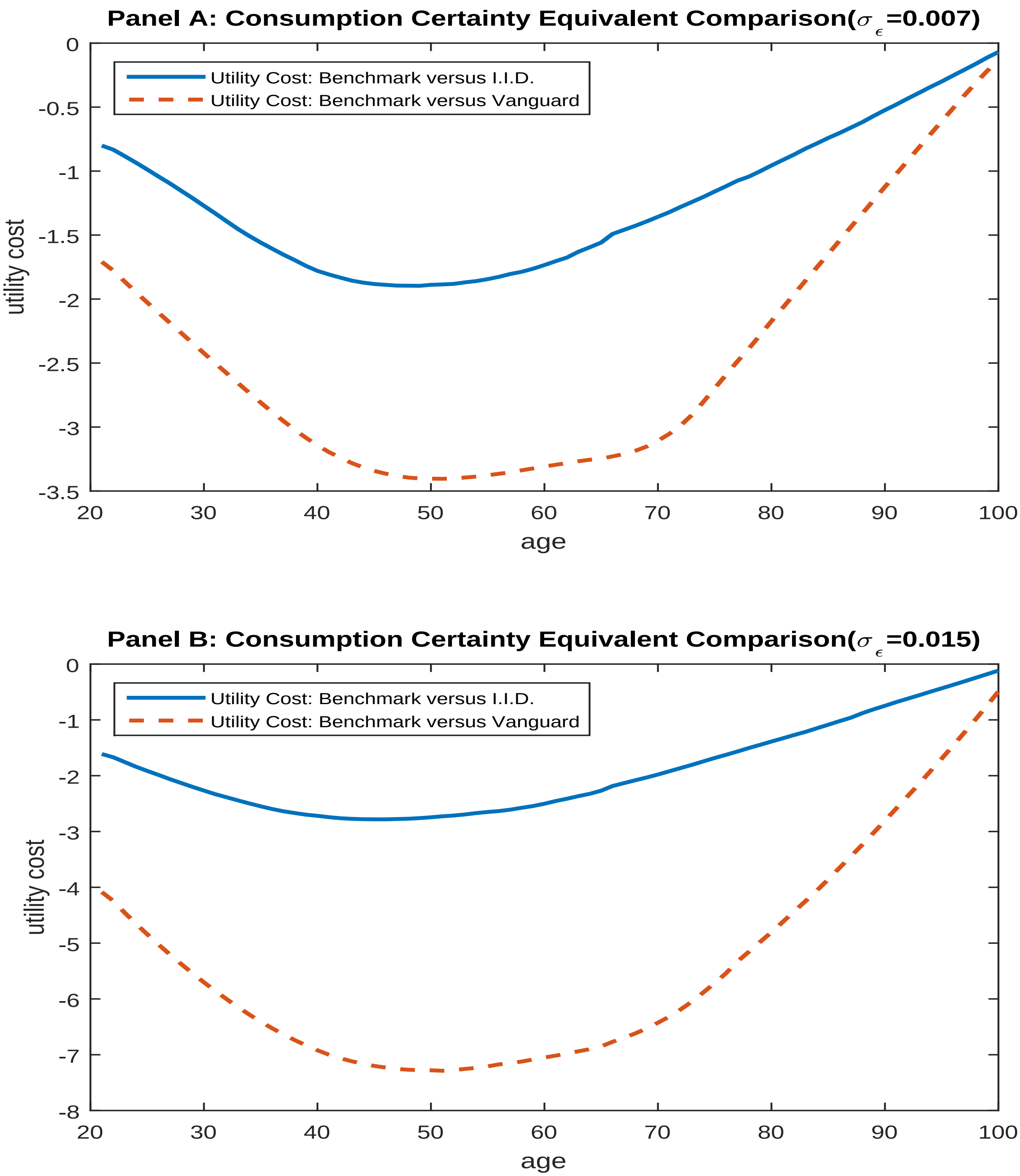


Figure 12: Welfare evaluation: average certainty equivalent in percent. We assume returns are generated based on the mean reversion model but households are either using the I.I.D. model or the Vanguard policy rule. Welfare loss is relative to the baseline case.

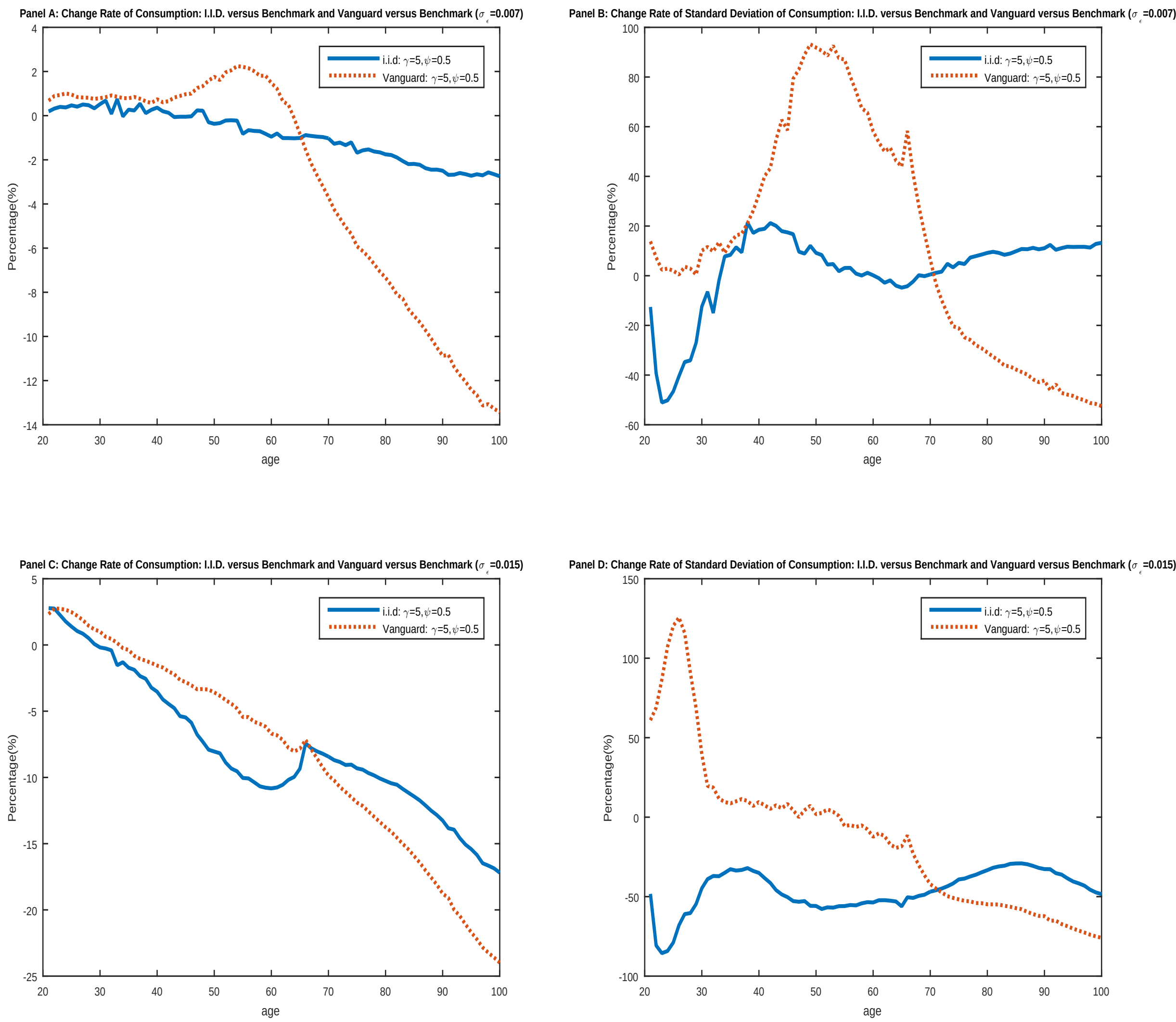


Figure 13: Consumption Evaluation: percentage change in average consumption and standard deviation of consumption over the life cycle. The two cases are (i) the I.I.D. case ($\gamma = 5$ and $\psi = 0.5$) relative to the baseline model and (ii) the Vanguard case($\gamma = 5$ and $\psi = 0.5$) relative to the baseline model.

Online Appendix for "Stock Market Mean Reversion and Portfolio Choice over the Life Cycle"

June 10, 2015

1 Appendix A

1.1 Tauchen-Hussey(1991) Parameterization Experiments

We now provide extensive experimentation with the Tauchen-Hussey (1991) parameterization when there are four different variables (as will be required for solving the model). 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 return (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 investigate how changing correlations between these variables affect hedging demands. Adding them as part of the VAR can help us determine the number of grid points required for the numerical accuracy 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:

$$\begin{bmatrix} f_{t+1} \\ r_{t+1} \\ lnN_{t+1} \\ lnU_{t+1} \end{bmatrix} = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix} + \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} f_t \\ r_t \\ lnN_t \\ lnU_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \\ n_{t+1} \\ u_{t+1} \end{bmatrix}$$

Denote the coefficient matrix and co-variance matrix as follows:

$$A1 = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix}, A2 = \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} 0.000049 & -0.000753 & 0 & 0 \\ & 0.03214 & 0.002689 & 0 \\ & & 0.01 & 0 \\ & & & 0.01 \end{bmatrix} \text{ and } \rho = \begin{bmatrix} 1 & -0.6 & 0 & 0 \\ & 1 & 0.15 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}$$

To test the accuracy of the approximation method, we use two different methods. The first method is based on simulation. We simulate based on a Tauchen(1986) discretization for a given number of grid points and then perform a Monte Carlo analysis to investigate how close the estimated parameters are to the actual parameters used to generate the discrete approximation. Specifically, we generate 100 simulation paths and each simulation path has 52,000 periods. In each simulation, the first 2000 periods are discarded. The VAR is then estimated in each of 100 times and the coefficients averaged and reported below. The numbers in parentheses are the standard deviations of the reported average coefficients. Since the means of f_t, r_t, $ln(N_t)$ and $ln(U_t)$ can always be estimated correctly in all the experiments, we only report the VAR coefficient averages.

The second method uses the nodes and weights used in the numerical solution to compute the conditional expected equity risk premium (which should equal the factor state) and the conditional variance of the risk premium. The conditional correlation matrix is also reported. These values should be close to the simulations. Moreover the current factor state should equal the expected risk premium conditional on the current factor realization.

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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A1} = \begin{bmatrix} \underset{(0.0002)}{0.00446} \\ \underset{(0.0045)}{0.0131} \\ \underset{(0.0024)}{-0.0054} \\ \underset{(0.0032)}{-0.0054} \end{bmatrix} = \begin{bmatrix} A1_1 \\ A1_2 \\ A1_3 \\ A1_4 \end{bmatrix}$$

$$\hat{A2} = \begin{bmatrix} \underset{(0.004)}{0.888} & \underset{(0.0004)}{0.000236} & \underset{(0.0006)}{-0.00017} & \underset{(0.0007)}{0.00006} \\ \underset{(0.1)}{1.17} & \underset{(0.01)}{0.000822} & \underset{(0.017)}{0.00086} & \underset{(0.015)}{-0.0017} \\ \underset{(0.054)}{0.008} & \underset{(0.0056)}{0.00067} & \underset{(0.01)}{-0.002} & \underset{(0.01)}{-0.0002} \\ \underset{(0.074)}{0.0136} & \underset{(0.0057)}{0.00037} & \underset{(0.012)}{-0.00022} & \underset{(0.01)}{0.0021} \end{bmatrix} = \begin{bmatrix} A2_{11} & A2_{12} & A2_{13} & A2_{14} \\ A2_{21} & A2_{22} & A2_{23} & A2_{24} \\ A2_{31} & A2_{32} & A2_{33} & A2_{34} \\ A2_{41} & A2_{42} & A2_{43} & A2_{44} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.00005} & \underset{(0.000015)}{-0.00074} & \underset{(0.0000054)}{-0.00000089} & \underset{(0.000006)}{-0.0000016} \\ & \underset{(0.00051)}{0.0319} & \underset{(0.000176)}{0.00268} & \underset{(0.000174)}{0.000014} \\ & & \underset{(0.00012)}{0.01} & \underset{(0.00011)}{-0.00000049} \\ & & & \underset{(0.00012)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000004 \\ 0.00002 \\ 0.0000058 \\ 0.0000105 \end{bmatrix}$. The average is 0.0000091.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.000018 & 0.00000015 & 0.0000003 & 0.00000047 \\ 0.011 & 0.000095 & 0.00028 & 0.000245 \\ 0.003 & 0.000032 & 0.0001 & 0.0001 \\ 0.0055 & 0.000033 & 0.00014 & 0.000083 \end{bmatrix}$$

The average is 0.0013.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1} \mid f_t]$	$Var[RP_{t+1} \mid f_t]$	$\rho_{RP_{t+1} \mid f_t}$

0.005984	-0.0166	0.0286	$\begin{bmatrix} 1 & -0.524 & -0.0102 \\ & 1 & 0.158 \\ & & 1 \end{bmatrix}$
0.01493	0.014566	0.03155	$\begin{bmatrix} 1 & -0.592 & -0.0013 \\ & 1 & 0.165 \\ & & 1 \end{bmatrix}$
0.02261	0.02149	0.03207	$\begin{bmatrix} 1 & -0.602 & -0.008 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.02974	0.02845	0.03153	$\begin{bmatrix} 1 & -0.596 & -0.002 \\ & 1 & 0.134 \\ & & 1 \end{bmatrix}$
0.036605	0.03885	0.03261	$\begin{bmatrix} 1 & -0.612 & 0.003 \\ & 1 & 0.155 \\ & & 1 \end{bmatrix}$
0.0434	0.04556	0.03256	$\begin{bmatrix} 1 & -0.59 & 0.007 \\ & 1 & 0.153 \\ & & 1 \end{bmatrix}$
0.05026	0.04817	0.03174	$\begin{bmatrix} 1 & -0.599 & -0.013 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.05739	0.05666	0.03136	$\begin{bmatrix} 1 & -0.59 & 0.036 \\ & 1 & 0.152 \\ & & 1 \end{bmatrix}$
0.06507	0.06284	0.0317	$\begin{bmatrix} 1 & -0.598 & -0.013 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$
0.07402	0.09232	0.02776	$\begin{bmatrix} 1 & -0.52 & -0.013 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.005984	-0.013522	0.028117	$\begin{bmatrix} 1 & -0.518 & 0.0 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.01493	0.0142	0.03183	$\begin{bmatrix} 1 & -0.59 & 0.003 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.02261	0.02261	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02974	0.02974	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.036605	0.036605	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.0434	0.0434	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05026	0.05026	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05739	0.05739	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06507	0.0658	0.03182	$\begin{bmatrix} 1 & -0.595 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07402	0.093522	0.02812	$\begin{bmatrix} 1 & -0.52 & 0.0 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

From these tables, we can see that the Tauchen-Hussey (1991) method can produce accurate conditional risk premia moments. Nevertheless, the accuracy of the conditional moments is not good at the lowest and highest factors. Moreover, we observe that the estimations of the VAR coefficients $A1_1$,$A1_2$and $A2_{21}$ are not as accurate as desired. To improve accuracy, we first experiment by increasing the number of grid points of r_t. This leads to experiment 2.

Experiment 2: The number of grid points for f_t is 10, 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 grid points for $ln(U_t)$ is 3. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A1} = \begin{bmatrix} \underset{(0.0002)}{0.0045} \\ \underset{(0.005)}{0.013} \\ \underset{(0.003)}{-0.0048} \\ \underset{(0.003)}{-0.005} \end{bmatrix}, \hat{A2} = \begin{bmatrix} \underset{(0.004)}{0.888} & \underset{(0.0004)}{-0.00011} & \underset{(0.0006)}{0.00001} & \underset{(0.0007)}{-0.00015} \\ \underset{(0.11)}{1.176} & \underset{(0.01)}{0.00172} & \underset{(0.02)}{0.0031} & \underset{(0.017)}{0.003} \\ \underset{(0.067)}{-0.0058} & \underset{(0.005)}{0.00021} & \underset{(0.01)}{0.0013} & \underset{(0.01)}{0.0007} \\ \underset{(0.067)}{-0.0028} & \underset{(0.006)}{0.0011} & \underset{(0.0093)}{-0.0012} & \underset{(0.01)}{-0.0017} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.000048} & \underset{(0.000013)}{-0.00074} & \underset{(0.000007)}{-0.0000014} & \underset{(0.000007)}{0.000000} \\ & \underset{(0.0004)}{0.0319} & \underset{(0.0002)}{0.0027} & \underset{(0.00019)}{-0.000029} \\ & & \underset{(0.00015)}{0.01} & \underset{(0.0001)}{0.0000002} \\ & & & \underset{(0.00014)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.000000004 \\ 0.000027 \\ 0.0000096 \\ 0.000009 \end{bmatrix}$. The average MSE is 0.0000114.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.000017 & 0.0000002 & 0.00000034 & 0.0000004 \\ 0.013 & 0.0001 & 0.0004 & 0.0003 \\ 0.0045 & 0.00003 & 0.00011 & 0.000085 \\ 0.0044 & 0.00004 & 0.000086 & 0.000095 \end{bmatrix}$$

The average is 0.00114.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1} \mid f_t]$	$Var[RP_{t+1} \mid f_t]$	$\rho_{RP_{t+1} \mid f_t}$
0.005984	-0.01373	0.02817	$\begin{bmatrix} 1 & -0.51 & 0.013 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.014927	0.01391	0.03161	$\begin{bmatrix} 1 & -0.605 & -0.005 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.022610	0.024357	0.032303	$\begin{bmatrix} 1 & -0.605 & -0.005 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02974	0.0282	0.03166	$\begin{bmatrix} 1 & -0.6 & 0.004 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.03661	0.035	0.03175	$\begin{bmatrix} 1 & -0.61 & 0.004 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.0434	0.04454	0.0322	$\begin{bmatrix} 1 & -0.61 & -0.013 \\ & 1 & 0.17 \\ & & 1 \end{bmatrix}$
0.050262	0.05123	0.0324	$\begin{bmatrix} 1 & -0.596 & 0.027 \\ & 1 & 0.128 \\ & & 1 \end{bmatrix}$
0.05739	0.05674	0.0325	$\begin{bmatrix} 1 & -0.6 & -0.03 \\ & 1 & 0.174 \\ & & 1 \end{bmatrix}$
0.06507	0.06501	0.03184	$\begin{bmatrix} 1 & -0.59 & -0.075 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$
0.07402	0.09424	0.02828	$\begin{bmatrix} 1 & -0.52 & 0.0076 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\vert f_t\right]$	$Var\left[RP_{t+1}\vert f_t\right]$	$\rho_{RP_{t+1}\vert f_t}$
0.005984	-0.013522	0.028117	$\begin{bmatrix} 1 & -0.52 & 0.0 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.014927	0.0142	0.03183	$\begin{bmatrix} 1 & -0.59 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02261	0.02261	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02974	0.02974	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.03661	0.03661	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.0434	0.0434	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05026	0.05026	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05739	0.05739	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06507	0.0658	0.031828	$\begin{bmatrix} 1 & -0.595 & 0.00 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07402	0.093522	0.02812	$\begin{bmatrix} 1 & -0.518 & 0.00 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

This does not improve accuracy too much and we therefore then increase the number of grid points of f_t instead of r_t.

Experiment 3: The number of grid points for f_t is 15, 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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.00413} \\ \underset{(0.005)}{0.0174} \\ \underset{(0.003)}{-0.0056} \\ \underset{(0.002)}{-0.0051} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.005)}{0.897} & \underset{(0.0003)}{0.00004} & \underset{(0.0007)}{0.00006} & \underset{(0.00068)}{-0.0001} \\ \underset{(0.11)}{1.06} & \underset{(0.01)}{-0.0004} & \underset{(0.018)}{-0.003} & \underset{(0.018)}{0.0023} \\ \underset{(0.07)}{0.0155} & \underset{(0.007)}{0.00054} & \underset{(0.01)}{-0.0033} & \underset{(0.01)}{0.0002} \\ \underset{(0.06)}{-0.0004} & \underset{(0.006)}{-0.0005} & \underset{(0.011)}{-0.0019} & \underset{(0.01)}{0.0029} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.0000489} & \underset{(0.0000014)}{-0.000748} & \underset{(0.000008)}{0.0000005} & \underset{(0.000008)}{0.0000002} \\ & \underset{(0.00044)}{0.0032} & \underset{(0.00018)}{0.0027} & \underset{(0.00018)}{-0.00002} \\ & & \underset{(0.00013)}{0.01} & \underset{(0.0001)}{-0.000005} \\ & & & \underset{(0.00012)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000004 \\ 0.000027 \\ 0.00001 \\ 0.000005 \end{bmatrix}$. The average is 0.00000106.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.00002 & 0.0000001 & 0.0000005 & 0.0000005 \\ 0.012 & 0.00086 & 0.00031 & 0.00031 \\ 0.005 & 0.000047 & 0.0001 & 0.000085 \\ 0.0034 & 0.000039 & 0.00013 & 0.0001 \end{bmatrix}$$

The average is 0.00139.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1}\|f_t]$	$Var[RP_{t+1}\|f_t]$	$\rho_{RP_{t+1}\|f_t}$
-0.004548	-0.021244	0.02845	$\begin{bmatrix} 1 & -0.53 & -0.002 \\ & 1 & 0.17 \\ & & 1 \end{bmatrix}$

0.00367	0.003517	0.03184	$\begin{bmatrix} 1 & -0.59 & 0.014 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.01063	0.009214	0.03323	$\begin{bmatrix} 1 & -0.598 & -0.004 \\ & 1 & 0.157 \\ & & 1 \end{bmatrix}$
0.01698	0.02011	0.0318	$\begin{bmatrix} 1 & -0.596 & 0.009 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.02297	0.02025	0.0338	$\begin{bmatrix} 1 & -0.62 & -0.003 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.02875	0.02761	0.0317	$\begin{bmatrix} 1 & -0.6 & -0.006 \\ & 1 & 0.156 \\ & & 1 \end{bmatrix}$
0.03441	0.03396	0.0326	$\begin{bmatrix} 1 & -0.61 & -0.012 \\ & 1 & 0.157 \\ & & 1 \end{bmatrix}$
0.04	0.04214	0.0319	$\begin{bmatrix} 1 & -0.59 & 0.0004 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.04559	0.04437	0.0328	$\begin{bmatrix} 1 & -0.61 & 0.017 \\ & 1 & 0.135 \\ & & 1 \end{bmatrix}$
0.05125	0.05262	0.0332	$\begin{bmatrix} 1 & -0.61 & 0.02 \\ & 1 & 0.128 \\ & & 1 \end{bmatrix}$
0.05703	0.05724	0.03162	$\begin{bmatrix} 1 & -0.59 & -0.01 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.06302	0.06169	0.03259	$\begin{bmatrix} 1 & -0.6 & -0.01 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

0.06937	0.06756	0.033	$\begin{bmatrix} 1 & -0.61 & -0.009 \\ & 1 & 0.158 \\ & & 1 \end{bmatrix}$
0.07633	0.07597	0.03236	$\begin{bmatrix} 1 & -0.6 & -0.01 \\ & 1 & 0.134 \\ & & 1 \end{bmatrix}$
0.08455	0.1	0.02812	$\begin{bmatrix} 1 & -0.51 & -0.04 \\ & 1 & 0.159 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
-0.004548	-0.02121	0.02866	$\begin{bmatrix} 1 & -0.53 & 0.0 \\ & 1 & 0.159 \\ & & 1 \end{bmatrix}$
0.00367	0.00271	0.03177	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.01063	0.01061	0.03213	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.01698	0.01698	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02297	0.02297	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02875	0.02875	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$

0.03441	0.03441	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04	0.04	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04559	0.04559	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05125	0.05125	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05703	0.05703	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06302	0.06302	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06937	0.06939	0.03213	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07633	0.07729	0.03177	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.08455	0.1012	0.02867	$\begin{bmatrix} 1 & -0.53 & 0.0 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

From these results, we conclude that the higher number of grid points for f_t does increase the accuracy of the estimation of VAR coefficients and conditional moments of the risk premium. Now, an interesting question is whether we can further improve the accuracy of estimation by increasing the number of grid points of f_t and r_t. The experiments 4 and 5 answer this question.

Experiment 4: The number of grid points for f_t is 20, 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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.004} \\ \underset{(0.0044)}{0.02} \\ \underset{(0.003)}{-0.0055} \\ \underset{(0.003)}{-0.0042} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.004)}{0.899} & \underset{(0.0004)}{-0.00002} & \underset{(0.0008)}{-0.000002} & \underset{(0.00076)}{0.00007} \\ \underset{(0.11)}{1.01} & \underset{(0.01)}{-0.001} & \underset{(0.02)}{-0.0035} & \underset{(0.019)}{-0.001} \\ \underset{(0.07)}{0.009} & \underset{(0.0056)}{-0.0003} & \underset{(0.013)}{-0.00097} & \underset{(0.01)}{0.003} \\ \underset{(0.06)}{-0.014} & \underset{(0.0053)}{-0.0006} & \underset{(0.011)}{0.0025} & \underset{(0.01)}{0.001} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000008)}{0.000049} & \underset{(0.00001)}{-0.000754} & \underset{(0.000006)}{-0.0000007} & \underset{(0.000007)}{-0.0000003} \\ & \underset{(0.0004)}{0.00322} & \underset{(0.00017)}{0.0027} & \underset{(0.00016)}{-0.00004} \\ & & \underset{(0.00014)}{0.01} & \underset{(0.0001)}{-0.000001} \\ & & & \underset{(0.00011)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000003 \\ 0.00002 \\ 0.00001 \\ 0.00001 \end{bmatrix}$. The average is 0.000009.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.00002 & 0.00000015 & 0.0000006 & 0.0000006 \\ 0.012 & 0.0001 & 0.0004 & 0.00037 \\ 0.0052 & 0.00003 & 0.00016 & 0.0001 \\ 0.0038 & 0.00003 & 0.00013 & 0.0001 \end{bmatrix}$$

The average is 0.0014.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1}\|f_t]$	$Var[RP_{t+1}\|f_t]$	$\rho_{RP_{t+1}\|f_t}$
-0.01333	0.02857	0.02967	$\begin{bmatrix} 1 & -0.54 & -0.007 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$

-0.005574	-0.008782	0.03197	$\begin{bmatrix} 1 & -0.59 & 0.006 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.000949	0.001329	0.0313	$\begin{bmatrix} 1 & -0.6 & 0.01 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.006858	0.007505	0.03264	$\begin{bmatrix} 1 & -0.606 & 0.01 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.01239	0.01041	0.03268	$\begin{bmatrix} 1 & -0.6 & -0.005 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$
0.01768	0.01876	0.0317	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.02279	0.01992	0.0325	$\begin{bmatrix} 1 & -0.6 & 0.007 \\ & 1 & 0.143 \\ & & 1 \end{bmatrix}$
0.02778	0.0251	0.033	$\begin{bmatrix} 1 & -0.6 & -0.004 \\ & 1 & 0.158 \\ & & 1 \end{bmatrix}$
0.0327	0.0339	0.0321	$\begin{bmatrix} 1 & -0.6 & 0.004 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.03757	0.03527	0.031	$\begin{bmatrix} 1 & -0.6 & 0.006 \\ & 1 & 0.139 \\ & & 1 \end{bmatrix}$
0.04243	0.046	0.0322	$\begin{bmatrix} 1 & -0.6 & 0.013 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.04730	0.0509	0.0307	$\begin{bmatrix} 1 & -0.6 & 0.005 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$

0.05222	0.05186	0.03235	$\begin{bmatrix} 1 & -0.6 & -0.02 \\ & 1 & 0.158 \\ & & 1 \end{bmatrix}$
0.05721	0.05702	0.0314	$\begin{bmatrix} 1 & -0.6 & 0.004 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.06232	0.06465	0.0319	$\begin{bmatrix} 1 & -0.59 & 0.04 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06761	0.06754	0.0321	$\begin{bmatrix} 1 & -0.6 & 0.003 \\ & 1 & 0.156 \\ & & 1 \end{bmatrix}$
0.07314	0.07425	0.032	$\begin{bmatrix} 1 & -0.59 & 0.01 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.07905	0.07842	0.0325	$\begin{bmatrix} 1 & -0.6 & -0.01 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.08557	0.08602	0.0311	$\begin{bmatrix} 1 & -0.58 & -0.014 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.09333	0.109	0.02927	$\begin{bmatrix} 1 & -0.54 & 0.014 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
-0.01333	-0.0275	0.029	$\begin{bmatrix} 1 & -0.54 & 0.0 \\ & 1 & 0.158 \\ & & 1 \end{bmatrix}$

-0.005574	-0.006574	0.0318	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.000949	0.000919	0.03213	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.006858	0.006858	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.01239	0.01239	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.01768	0.01768	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02279	0.02279	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02778	0.02778	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.0327	0.0327	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.03757	0.03757	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04243	0.04243	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04730	0.04730	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$

0.05222	0.05222	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05721	0.05721	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06232	0.06232	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06761	0.06761	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07314	0.07314	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07905	0.07908	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.08557	0.08657	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.09333	0.10753	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$

We conclude from these results that accuracy does not increase too much compared to the last experiment and therefore start using 15 grid points for factor f.

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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.0041} \\ \underset{(0.006)}{0.0187} \\ \underset{(0.0024)}{-0.005} \\ \underset{(0.0026)}{-0.0042} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.004)}{0.897} & \underset{(0.00044)}{-0.0000015} & \underset{(0.0006)}{0.00007} & \underset{(0.0007)}{-0.00006} \\ \underset{(0.13)}{1.03} & \underset{(0.01)}{0.00057} & \underset{(0.016)}{-0.00017} & \underset{(0.017)}{0.0019} \\ \underset{(0.057)}{0.0014} & \underset{(0.0046)}{-0.000004} & \underset{0.011}{-0.0003} & \underset{(0.01)}{0.00056} \\ \underset{(0.06)}{-0.008} & \underset{(0.0066)}{-0.0009} & \underset{(0.012)}{0.00145} & \underset{(0.01)}{0.0018} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.0000488} & \underset{(0.000015)}{-0.000747} & \underset{(0.000006)}{0.0000006} & \underset{(0.0000067)}{0.0000004} \\ & \underset{(0.0005)}{0.0321} & \underset{(0.0002)}{0.00262} & \underset{(0.00014)}{0.000001} \\ & & \underset{(0.00014)}{0.01} & \underset{(0.0001)}{-0.000004} \\ & & & \underset{(0.00016)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000004 \\ 0.000034 \\ 0.00000057 \\ 0.000007 \end{bmatrix}$. The average is 0.0000012.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.00002 & 0.00000002 & 0.00000041 & 0.00000045 \\ 0.018 & 0.0001 & 0.00027 & 0.0003 \\ 0.0032 & 0.000021 & 0.00013 & 0.00012 \\ 0.0039 & 0.000043 & 0.00014 & 0.00011 \end{bmatrix}$$

The average is 0.0016.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1}\|f_t]$	$Var[RP_{t+1}\|f_t]$	$\rho_{RP_{t+1}\|f_t}$
-0.004548	-0.023133	0.02833	$\begin{bmatrix} 1 & -0.595 & -0.012 \\ & 1 & 0.162 \\ & & 1 \end{bmatrix}$
0.00367	0.001543	0.03296	$\begin{bmatrix} 1 & -0.6 & 0.0004 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

0.01063	0.012227	0.03116	$\begin{bmatrix} 1 & -0.6 & 0.002 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.01698	0.020389	0.03184	$\begin{bmatrix} 1 & -0.603 & -0.02 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02297	0.023539	0.032034	$\begin{bmatrix} 1 & -0.61 & -0.002 \\ & 1 & 0.141 \\ & & 1 \end{bmatrix}$
0.02875	0.03113	0.03217	$\begin{bmatrix} 1 & -0.59 & 0.005 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.03441	0.03072	0.03221	$\begin{bmatrix} 1 & -0.59 & 0.005 \\ & 1 & 0.135 \\ & & 1 \end{bmatrix}$
0.04	0.04123	0.03217	$\begin{bmatrix} 1 & -0.6 & 0.015 \\ & 1 & 0.153 \\ & & 1 \end{bmatrix}$
0.04559	0.04338	0.032382	$\begin{bmatrix} 1 & -0.6 & -0.001 \\ & 1 & 0.153 \\ & & 1 \end{bmatrix}$
0.05125	0.05167	0.03134	$\begin{bmatrix} 1 & -0.58 & 0.075 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05703	0.05274	0.03217	$\begin{bmatrix} 1 & -0.62 & 0.0005 \\ & 1 & 0.139 \\ & & 1 \end{bmatrix}$
0.06302	0.06356	0.03186	$\begin{bmatrix} 1 & -0.6 & 0.005 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06937	0.065961	0.03235	$\begin{bmatrix} 1 & -0.6 & -0.012 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$

0.07633	0.07574	0.031	$\begin{bmatrix} 1 & -0.6 & 0.008 \\ & 1 & 0.156 \\ & & 1 \end{bmatrix}$
0.08455	0.1006	0.02813	$\begin{bmatrix} 1 & -0.54 & -0.02 \\ & 1 & 0.166 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1} \vert f_t\right]$	$Var\left[RP_{t+1} \vert f_t\right]$	$\rho_{RP_{t+1} \vert f_t}$
-0.004548	-0.02121	0.02867	$\begin{bmatrix} 1 & -0.53 & 0.0 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.00367	0.002709	0.03177	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.01063	0.01061	0.03213	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.01698	0.01698	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02297	0.02297	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.02875	0.02875	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.03441	0.03441	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$

0.04	0.04	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04559	0.04559	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05125	0.05125	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05703	0.05703	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06302	0.06302	0.03214	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.06937	0.06939	0.03213	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.07633	0.07729	0.03177	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.08455	0.1012	0.02867	$\begin{bmatrix} 1 & -0.53 & 0.0 \\ & 1 & 0.159 \\ & & 1 \end{bmatrix}$

In this experiment, accuracy is not better than the one in experiment 3.

Based on the results above, we find that the results are insensitive to the number of grid points when the numbers for f_t and r_t are beyond 15 and 20 respectively, and the numbers for $ln(U_t)$ and $ln(N_t)$ are above 3 and 5, respectively.

2 Appendix B

2.1 Tauchen(1986) Parameterization Experiments

We now provide extensive experimentation with the Tauchen (1986) parameterization when there are four different variables (as will be required for solving the model). 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 return (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 investigate how changing correlations between these variables affect hedging demands. Adding them as part of the VAR can help us determine the number of grid points required for the numerical accuracy 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:

$$\begin{bmatrix} f_{t+1} \\ r_{t+1} \\ lnN_{t+1} \\ lnU_{t+1} \end{bmatrix} = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix} + \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} f_t \\ r_t \\ lnN_t \\ lnU_t \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \\ n_{t+1} \\ u_{t+1} \end{bmatrix}$$

Denote the coefficient matrix and co-variance matrix as follows:

$$A1 = \begin{bmatrix} 0.004 \\ 0.02 \\ -0.005 \\ -0.005 \end{bmatrix}, A2 = \begin{bmatrix} 0.9 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\Omega = \begin{bmatrix} 0.000049 & -0.000753 & 0 & 0 \\ & 0.03214 & 0.0027 & 0 \\ & & 0.01 & 0 \\ & & & 0.01 \end{bmatrix} \text{ and } \rho = \begin{bmatrix} 1 & -0.6 & 0 & 0 \\ & 1 & 0.15 & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix}$$

To test the accuracy of the approximation method, we use two different methods. The first method is based on simulation. We simulate based on a Tauchen(1986) discretization for a given number of grid points and then perform a Monte Carlo analysis to investigate how close the estimated parameters are to the actual parameters used to generate the discrete approximation. Specifically, we generate 100 simulation paths and each simulation path has 52,000 periods. In each simulation, the first 2000 periods are discarded. The VAR is then estimated in each of 100 times and the coefficients averaged and reported below. The numbers in parentheses are the standard deviations of the reported average coefficients. Since the means of f_t, r_t, $ln(N_t)$ and $ln(U_t)$ can always be estimated correctly in all the experiments, we only report the VAR coefficient averages.

The second method uses the nodes and weights used in the numerical solution to compute the conditional expected equity risk premium (which should equal the factor state) and the conditional variance of the risk premium. The conditional correlation matrix is also reported. These values should be close to the simulations. Moreover the current factor state should equal the expected risk premium conditional on the current factor realization.

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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A1} = \begin{bmatrix} \underset{(0.00018)}{0.0046} \\ \underset{(0.0045)}{-0.0226} \\ \underset{(0.003)}{-0.0064} \\ \underset{(0.0025)}{-0.0051} \end{bmatrix} = \begin{bmatrix} A1_1 \\ A1_2 \\ A1_3 \\ A1_4 \end{bmatrix}$$

$$\hat{A2} = \begin{bmatrix} \underset{(0.004)}{0.884} & \underset{(0.0004)}{0.000136} & \underset{(0.0008)}{-0.000198} & \underset{(0.0008)}{0.00011} \\ \underset{(0.1)}{0.93} & \underset{(0.0099)}{-0.00327} & \underset{(0.019)}{0.00678} & \underset{(0.2)}{0.00038} \\ \underset{(0.07)}{0.037} & \underset{(0.0062)}{0.00058} & \underset{(0.01)}{-0.00074} & \underset{(0.01)}{0.00083} \\ \underset{(0.06)}{0.0022} & \underset{(0.0048)}{0.0002} & \underset{(0.0093)}{0.00118} & \underset{(0.009)}{0.0011} \end{bmatrix} = \begin{bmatrix} A2_{11} & A2_{12} & A2_{13} & A2_{14} \\ A2_{21} & A2_{22} & A2_{23} & A2_{24} \\ A2_{31} & A2_{32} & A2_{33} & A2_{34} \\ A2_{41} & A2_{42} & A2_{43} & A2_{44} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.00005} & \underset{(0.000017)}{-0.00078} & \underset{(0.0000089)}{-0.00000057} & \underset{(0.000008)}{0.0000011} \\ & \underset{(0.00057)}{0.0316} & \underset{(0.000218)}{0.00258} & \underset{(0.0002)}{-0.000009} \\ & & \underset{(0.00012)}{0.01} & \underset{(0.00013)}{-0.000003} \\ & & & \underset{(0.00014)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.000000034 \\ 0.000021 \\ 0.00000954 \\ 0.0000064 \end{bmatrix}$. The average is 0.0000091.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.0000183 & 0.00000018 & 0.00000063 & 0.00000052 \\ 0.011 & 0.000097 & 0.000366 & 0.00037 \\ 0.00048 & 0.000039 & 0.0001132 & 0.000105 \\ 0.00425 & 0.000023 & 0.0000864 & 0.000083 \end{bmatrix}$$

The average is 0.00132.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.007882	0.0111	0.0196	$\begin{bmatrix} 1 & -0.601 & -0.0082 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01502	0.01732	0.02858	$\begin{bmatrix} 1 & -0.619 & 0.003 \\ & 1 & 0.128 \\ & & 1 \end{bmatrix}$
0.02216	0.02323	0.03161	$\begin{bmatrix} 1 & -0.618 & 0.005 \\ & 1 & 0.126 \\ & & 1 \end{bmatrix}$
0.02929	0.03119	0.03302	$\begin{bmatrix} 1 & -0.625 & 0.004 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$

0.03643	0.0347	0.03404	$\begin{bmatrix} 1 & -0.619 & -0.003 \\ & 1 & 0.128 \\ & & 1 \end{bmatrix}$
0.043569	0.042696	0.033408	$\begin{bmatrix} 1 & -0.614 & 0.011 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.05071	0.0498	0.03285	$\begin{bmatrix} 1 & -0.603 & 0.016 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.05784	0.05777	0.03195	$\begin{bmatrix} 1 & -0.62 & -0.012 \\ & 1 & 0.155 \\ & & 1 \end{bmatrix}$
0.06498	0.05908	0.02866	$\begin{bmatrix} 1 & -0.609 & -0.007 \\ & 1 & 0.135 \\ & & 1 \end{bmatrix}$
0.07212	0.06624	0.01962	$\begin{bmatrix} 1 & -0.6 & -0.009 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\vert f_t\right]$	$Var\left[RP_{t+1}\vert f_t\right]$	$\rho_{RP_{t+1}\vert f_t}$
0.007882	0.0112	0.0196	$\begin{bmatrix} 1 & -0.604 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01502	0.01825	0.02861	$\begin{bmatrix} 1 & -0.623 & 0.003 \\ & 1 & 0.135 \\ & & 1 \end{bmatrix}$
0.02216	0.02315	0.03237	$\begin{bmatrix} 1 & -0.619 & 0.005 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$

0.02929	0.02945	0.03325	$\begin{bmatrix} 1 & -0.615 & 0.004 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.03643	0.03645	0.03339	$\begin{bmatrix} 1 & -0.614 & -0.003 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.043569	0.04355	0.033387	$\begin{bmatrix} 1 & -0.614 & 0.011 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.05071	0.0498	0.03285	$\begin{bmatrix} 1 & -0.614 & 0.016 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.05784	0.05685	0.03237	$\begin{bmatrix} 1 & -0.619 & -0.012 \\ & 1 & 0.1436 \\ & & 1 \end{bmatrix}$
0.06498	0.06175	0.0286	$\begin{bmatrix} 1 & -0.623 & -0.007 \\ & 1 & 0.1351 \\ & & 1 \end{bmatrix}$
0.07212	0.06878	0.01956	$\begin{bmatrix} 1 & -0.604 & -0.009 \\ & 1 & 0.118 \\ & & 1 \end{bmatrix}$

From these tables, we can see that the Tauchen (1986) method can produce accurate conditional risk premia moments. Nevertheless, the accuracy of the conditional moments is not good at the lowest and highest factors. This pattern is similar to results from using the Tauchen-Hussey (1991) method. Moreover, we observe that the estimations of the VAR coefficients $A1_1$,$A1_2$,$A1_3$and $A2_{21}$ are not as accurate as desired. To improve accuracy, we first experiment by increasing the number of grid points of r_t. This leads to experiment 2.

Experiment 2: The number of grid points for f_t is 10, 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 grid points for $ln(U_t)$ is 3. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.0047} \\ \underset{(0.005)}{0.0232} \\ \underset{(0.0025)}{-0.0048} \\ \underset{(0.0028)}{-0.0051} \end{bmatrix}, \hat{A}2 = \begin{bmatrix} \underset{(0.004)}{0.884} & \underset{(0.0005)}{0.000011} & \underset{(0.0007)}{0.00012} & \underset{(0.0006)}{-0.00002} \\ \underset{(0.1)}{0.918} & \underset{(0.012)}{0.00172} & \underset{(0.017)}{-0.00413} & \underset{(0.014)}{0.0007} \\ \underset{(0.05)}{0.0017} & \underset{(0.007)}{-0.00153} & \underset{(0.0125)}{0.00048} & \underset{(0.01)}{-0.00003} \\ \underset{(0.06)}{0.0071} & \underset{(0.0053)}{0.00074} & \underset{(0.0094)}{-0.0004} & \underset{(0.011)}{-0.00023} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.000015} & \underset{(0.000016)}{-0.00078} & \underset{(0.0000028)}{0.0000007} & \underset{(0.000008)}{0.000000} \\ & \underset{(0.0005)}{0.0315} & \underset{(0.00018)}{-0.0025} & \underset{(0.00018)}{0.00013} \\ & & \underset{(0.00014)}{0.01} & \underset{(0.00011)}{-0.000007} \\ & & & \underset{(0.00013)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.000000004 \\ 0.0000235 \\ 0.0000063 \\ 0.000007 \end{bmatrix}$. The average MSE is 0.0000099.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.000015 & 0.00000026 & 0.00000056 & 0.0000004 \\ 0.011 & 0.00016 & 0.00029 & 0.00019 \\ 0.003 & 0.000045 & 0.00016 & 0.00011 \\ 0.0036 & 0.000028 & 0.000089 & 0.00013 \end{bmatrix}$$

The average is 0.001145.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.007882	0.01304	0.01925	$\begin{bmatrix} 1 & -0.603 & -0.01 \\ & 1 & 0.1248 \\ & & 1 \end{bmatrix}$
0.01502	0.01948	0.02917	$\begin{bmatrix} 1 & -0.626 & -0.007 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$

0.02216	0.02191	0.03256	$\begin{bmatrix} 1 & -0.622 & 0.001 \\ & 1 & 0.155 \\ & & 1 \end{bmatrix}$
0.02929	0.0316	0.03341	$\begin{bmatrix} 1 & -0.621 & -0.007 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.03643	0.0404	0.0329	$\begin{bmatrix} 1 & -0.618 & -0.006 \\ & 1 & 0.141 \\ & & 1 \end{bmatrix}$
0.043569	0.041	0.03405	$\begin{bmatrix} 1 & -0.615 & 0.006 \\ & 1 & 0.138 \\ & & 1 \end{bmatrix}$
0.05071	0.05043	0.0333	$\begin{bmatrix} 1 & -0.622 & 0.008 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.05784	0.05534	0.0325	$\begin{bmatrix} 1 & -0.62 & -0.009 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$
0.06498	0.0608	0.0284	$\begin{bmatrix} 1 & -0.621 & -0.026 \\ & 1 & 0.153 \\ & & 1 \end{bmatrix}$
0.07212	0.06624	0.01962	$\begin{bmatrix} 1 & -0.6 & -0.004 \\ & 1 & 0.128 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.007882	0.0114	0.0194	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$

0.01502	0.01834	0.02844	$\begin{bmatrix} 1 & -0.625 & 0.003 \\ & 1 & 0.134 \\ & & 1 \end{bmatrix}$
0.02216	0.02318	0.0322	$\begin{bmatrix} 1 & -0.62 & 0.005 \\ & 1 & 0.143 \\ & & 1 \end{bmatrix}$
0.02929	0.02946	0.033	$\begin{bmatrix} 1 & -0.616 & 0.004 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.03643	0.03644	0.03324	$\begin{bmatrix} 1 & -0.615 & -0.003 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.043569	0.04355	0.03324	$\begin{bmatrix} 1 & -0.614 & 0.011 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.05071	0.05055	0.0331	$\begin{bmatrix} 1 & -0.616 & 0.016 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.05784	0.05682	0.0322	$\begin{bmatrix} 1 & -0.62 & -0.012 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.06498	0.0616	0.0284	$\begin{bmatrix} 1 & -0.625 & -0.007 \\ & 1 & 0.134 \\ & & 1 \end{bmatrix}$
0.07212	0.0686	0.0194	$\begin{bmatrix} 1 & -0.606 & -0.009 \\ & 1 & 0.116 \\ & & 1 \end{bmatrix}$

This does not improve accuracy too much and we therefore then increase the number of grid points of f_t instead of r_t.

Experiment 3: The number of grid points for f_t is 15, 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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A1} = \begin{bmatrix} \underset{(0.0001)}{0.00465} \\ \underset{(0.005)}{0.0209} \\ \underset{(0.002)}{-0.007} \\ \underset{(0.003)}{-0.0048} \end{bmatrix} \quad \hat{A2} = \begin{bmatrix} \underset{(0.004)}{0.884} & \underset{(0.00043)}{-0.00005} & \underset{(0.0006)}{0.00014} & \underset{(0.00069)}{0.00016} \\ \underset{(0.11)}{0.977} & \underset{(0.012)}{0.0003} & \underset{(0.018)}{-0.0021} & \underset{(0.016)}{-0.0042} \\ \underset{(0.06)}{-0.053} & \underset{(0.0006)}{0.0001} & \underset{(0.01)}{-0.0014} & \underset{(0.01)}{0.0018} \\ \underset{(0.07)}{0.002} & \underset{(0.0068)}{0.0003} & \underset{(0.011)}{-0.0023} & \underset{(0.01)}{-0.0008} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000006)}{0.0000489} & \underset{(0.000001)}{-0.0007} & \underset{(0.000007)}{-0.00000082} & \underset{(0.000007)}{-0.000001} \\ & \underset{(0.00047)}{0.00313} & \underset{(0.00014)}{0.0026} & \underset{(0.00015)}{0.00002} \\ & & \underset{(0.00012)}{0.01} & \underset{(0.0001)}{-0.000028} \\ & & & \underset{(0.0001)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000004 \\ 0.000028 \\ 0.00007 \\ 0.000097 \end{bmatrix}$. The average is 0.000001.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.00002 & 0.0000002 & 0.0000004 & 0.0000005 \\ 0.012 & 0.00015 & 0.0003 & 0.00027 \\ 0.0032 & 0.00004 & 0.0001 & 0.000096 \\ 0.0043 & 0.000046 & 0.00013 & 0.00009 \end{bmatrix}$$

The average is 0.00128.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1}\|f_t]$	$Var[RP_{t+1}\|f_t]$	$\rho_{RP_{t+1}\|f_t}$
0.007882	0.00929	0.019	$\begin{bmatrix} 1 & -0.585 & 0.0008 \\ & 1 & 0.117 \\ & & 1 \end{bmatrix}$
0.01247	0.0167	0.02512	$\begin{bmatrix} 1 & -0.599 & -0.011 \\ & 1 & 0.113 \\ & & 1 \end{bmatrix}$

0.01706	0.0174	0.0298	$\begin{bmatrix} 1 & -0.614 & 0.001 \\ & 1 & 0.143 \\ & & 1 \end{bmatrix}$
0.02165	0.02281	0.0318	$\begin{bmatrix} 1 & -0.605 & -0.004 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.02624	0.0255	0.0338	$\begin{bmatrix} 1 & -0.62 & -0.003 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.030823	0.0334	0.033	$\begin{bmatrix} 1 & -0.61 & 0.006 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.0354	0.04	0.0336	$\begin{bmatrix} 1 & -0.61 & -0.001 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.04	0.0399	0.0327	$\begin{bmatrix} 1 & -0.6 & -0.032 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.04459	0.0452	0.0334	$\begin{bmatrix} 1 & -0.6 & -0.026 \\ & 1 & 0.154 \\ & & 1 \end{bmatrix}$
0.04918	0.051	0.0317	$\begin{bmatrix} 1 & -0.6 & 0.005 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.053765	0.05455	0.03287	$\begin{bmatrix} 1 & -0.6 & 0.005 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.05835	0.05441	0.03138	$\begin{bmatrix} 1 & -0.61 & 0.015 \\ & 1 & 0.137 \\ & & 1 \end{bmatrix}$
0.06294	0.06048	0.0299	$\begin{bmatrix} 1 & -0.62 & -0.009 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$

0.06753	0.06547	0.02545	$\begin{bmatrix} 1 & -0.6 & -0.01 \\ & 1 & 0.134 \\ & & 1 \end{bmatrix}$
0.07212	0.0710	0.01926	$\begin{bmatrix} 1 & -0.59 & -0.011 \\ & 1 & 0.11 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.007882	0.0102	0.019	$\begin{bmatrix} 1 & -0.588 & 0.0008 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01247	0.0158	0.0253	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$
0.01706	0.0193	0.0295	$\begin{bmatrix} 1 & -0.612 & 0.0 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.02165	0.02268	0.0316	$\begin{bmatrix} 1 & -0.605 & 0.0 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.02624	0.0266	0.0324	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.03082	0.0309	0.0327	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.0354	0.0354	0.0328	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$

0.04	0.04	0.0328	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.04459	0.04457	0.0328	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.04918	0.04908	0.0327	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.053765	0.05341	0.0324	$\begin{bmatrix} 1 & -0.61 & 0.00 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.05835	0.05732	0.0316	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.06294	0.0607	0.0295	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.06753	0.0642	0.02543	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.132 \\ & & 1 \end{bmatrix}$
0.07212	0.0698	0.019	$\begin{bmatrix} 1 & -0.588 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$

From these results, we conclude that the higher number of grid points for f_t does increase the accuracy of the estimation of VAR coefficients and conditional moments of the risk premium. Now, an interesting question is whether we can further improve the accuracy of estimation by increasing the number of grid points of f_t and r_t. The experiments 4 and 5 answer this question.

Experiment 4: The number of grid points for f_t is 20, 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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0001)}{0.00463} \\ \underset{(0.0067)}{0.022} \\ \underset{(0.003)}{-0.0065} \\ \underset{(0.0026)}{-0.0048} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.004)}{0.885} & \underset{(0.0004)}{0.00004} & \underset{(0.0006)}{0.00004} & \underset{(0.00068)}{0.0002} \\ \underset{(0.13)}{0.95} & \underset{(0.009)}{0.0005} & \underset{(0.018)}{-0.0017} & \underset{(0.019)}{-0.002} \\ \underset{(0.06)}{0.04} & \underset{(0.0056)}{-0.0001} & \underset{(0.01)}{0.0008} & \underset{(0.01)}{-0.002} \\ \underset{(0.06)}{0.002} & \underset{(0.006)}{-0.00006} & \underset{(0.011)}{-0.0012} & \underset{(0.01)}{0.0002} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000007)}{0.000048} & \underset{(0.000001)}{-0.0007} & \underset{(0.000008)}{-0.0000015} & \underset{(0.000007)}{-0.000001} \\ & \underset{(0.00004)}{0.00311} & \underset{(0.00019)}{0.0026} & \underset{(0.00017)}{0.00005} \\ & & \underset{(0.00014)}{0.01} & \underset{(0.0001)}{-0.00001} \\ & & & \underset{(0.00016)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.00000004 \\ 0.000044 \\ 0.0001 \\ 0.000066 \end{bmatrix}$. The average is 0.0000015.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.00002 & 0.0000001 & 0.0000003 & 0.0000005 \\ 0.018 & 0.00008 & 0.0003 & 0.00037 \\ 0.0039 & 0.00003 & 0.00007 & 0.00008 \\ 0.0033 & 0.00003 & 0.00012 & 0.0001 \end{bmatrix}$$

The average is 0.00168.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E\left[RP_{t+1}\mid f_t\right]$	$Var\left[RP_{t+1}\mid f_t\right]$	$\rho_{RP_{t+1}\mid f_t}$
0.00788	0.010	0.0194	$\begin{bmatrix} 1 & -0.584 & -0.0007 \\ & 1 & 0.124 \\ & & 1 \end{bmatrix}$
0.01126	0.01476	0.02435	$\begin{bmatrix} 1 & -0.61 & -0.022 \\ & 1 & 0.138 \\ & & 1 \end{bmatrix}$

0.01464	0.01833	0.0273	$\begin{bmatrix} 1 & -0.61 & -0.01 \\ & 1 & 0.125 \\ & & 1 \end{bmatrix}$
0.018	0.0227	0.0298	$\begin{bmatrix} 1 & -0.605 & 0.003 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.0214	0.0208	0.031	$\begin{bmatrix} 1 & -0.6 & -0.013 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.0248	0.027	0.0319	$\begin{bmatrix} 1 & -0.61 & 0.012 \\ & 1 & 0.133 \\ & & 1 \end{bmatrix}$
0.02817	0.0303	0.0325	$\begin{bmatrix} 1 & -0.61 & -0.006 \\ & 1 & 0.156 \\ & & 1 \end{bmatrix}$
0.0315	0.0322	0.03177	$\begin{bmatrix} 1 & -0.6 & 0.004 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.0349	0.0362	0.0316	$\begin{bmatrix} 1 & -0.59 & 0.013 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.0383	0.0395	0.0328	$\begin{bmatrix} 1 & -0.6 & 0.004 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.0417	0.0391	0.0328	$\begin{bmatrix} 1 & -0.61 & -0.012 \\ & 1 & 0.16 \\ & & 1 \end{bmatrix}$
0.04507	0.04498	0.0324	$\begin{bmatrix} 1 & -0.6 & -0.013 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.04845	0.04696	0.0324	$\begin{bmatrix} 1 & -0.62 & 0.013 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$

0.05183	0.0493	0.0335	$\begin{bmatrix} 1 & -0.61 & -0.02 \\ & 1 & 0.17 \\ & & 1 \end{bmatrix}$
0.05521	0.05378	.0319	$\begin{bmatrix} 1 & -0.6 & -0.04 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.0586	0.0564	0.03138	$\begin{bmatrix} 1 & -0.61 & 0.003 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.06198	0.06075	0.0301	$\begin{bmatrix} 1 & -0.61 & -0.001 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.06536	0.06225	0.02725	$\begin{bmatrix} 1 & -0.61 & 0.022 \\ & 1 & 0.1245 \\ & & 1 \end{bmatrix}$
0.06874	0.06497	0.02307	$\begin{bmatrix} 1 & -0.59 & 0.013 \\ & 1 & 0.114 \\ & & 1 \end{bmatrix}$
0.07212	0.06849	0.01872	$\begin{bmatrix} 1 & -0.57 & 0.02 \\ & 1 & 0.11 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\|f_t\right]$	$Var\left[RP_{t+1}\|f_t\right]$	$\rho_{RP_{t+1}\|f_t}$
0.00788	0.00988	0.01883	$\begin{bmatrix} 1 & -0.58 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01126	0.01437	0.0236	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$

0.01464	0.01833	0.0273	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.125 \\ & & 1 \end{bmatrix}$
0.018	0.0199	0.0299	$\begin{bmatrix} 1 & -0.608 & 0.0 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.0214	0.0225	0.0313	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.0248	0.02529	0.032	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.02817	0.02837	0.0324	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.0315	0.0316	0.0325	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.03492	0.03495	0.0326	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.03831	0.038315	0.0326	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.0417	0.04168	0.0326	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.04507	0.04505	0.0326	$\begin{bmatrix} 1 & -0.6 & -0.013 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.04845	0.04838	0.0325	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$

0.05183	0.0516	0.0324	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.05521	0.05471	0.0321	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.0586	0.0575	0.0313	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.145 \\ & & 1 \end{bmatrix}$
0.06198	0.06	0.02989	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.142 \\ & & 1 \end{bmatrix}$
0.06536	0.06258	0.0274	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.137 \\ & & 1 \end{bmatrix}$
0.06874	0.0656	0.0236	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.129 \\ & & 1 \end{bmatrix}$
0.07212	0.07012	0.0188	$\begin{bmatrix} 1 & -0.58 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$

We conclude from these results that accuracy does not increase too much compared to the last experiment and therefore start using 15 grid points for factor f.

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. The average VAR coefficients, variance covariance matrix and mean squared errors (MSE) are:

$$\hat{A}1 = \begin{bmatrix} \underset{(0.0002)}{0.0045} \\ \underset{(0.005)}{0.023} \\ \underset{(0.003)}{-0.0064} \\ \underset{(0.003)}{-0.005} \end{bmatrix} \quad \hat{A}2 = \begin{bmatrix} \underset{(0.005)}{0.89} & \underset{(0.0005)}{0.00007} & \underset{(0.0005)}{0.0002} & \underset{(0.0006)}{-0.0001} \\ \underset{(0.12)}{0.921} & \underset{(0.01)}{-0.00071} & \underset{(0.02)}{-0.0027} & \underset{(0.016)}{-0.0024} \\ \underset{(0.06)}{-0.035} & \underset{(0.005)}{-0.0024} & \underset{(0.01)}{0.001} & \underset{(0.009)}{-0.002} \\ \underset{(0.08)}{0.011} & \underset{(0.005)}{-0.0008} & \underset{(0.009)}{-0.002} & \underset{(0.01)}{-0.0016} \end{bmatrix}$$

$$\hat{\Omega} = \begin{bmatrix} \underset{(0.0000008)}{0.00004} & \underset{(0.000014)}{-0.00075} & \underset{(0.000007)}{0.00000017} & \underset{(0.000006)}{0.00000086} \\ & \underset{(0.0004)}{0.0312} & \underset{(0.0002)}{0.0026} & \underset{(0.00016)}{0.00005} \\ & & \underset{(0.00011)}{0.01} & \underset{(0.00009)}{0.000007} \\ & & & \underset{(0.00012)}{0.01} \end{bmatrix}$$

The MSE of elements of $A1$: $\begin{bmatrix} 0.000000042 \\ 0.000027 \\ 0.000007 \\ 0.000012 \end{bmatrix}$. The average is 0.0000012.

The MSE of elements of $A2$:

$$\begin{bmatrix} 0.000023 & 0.00000002 & 0.00000032 & 0.00000036 \\ 0.015 & 0.000112 & 0.0004 & 0.00026 \\ 0.0036 & 0.000034 & 0.00011 & 0.000083 \\ 0.0069 & 0.000027 & 0.000095 & 0.000123 \end{bmatrix}$$

The average is 0.0017.

In what follows, the column labeled f_t is the discretized univariate factor state. The conditional moments of the risk premium estimated by simulation are:

f_t	$E[RP_{t+1}\vert f_t]$	$Var[RP_{t+1}\vert f_t]$	$\rho_{RP_{t+1}\vert f_t}$
0.007882	0.009	0.0192	$\begin{bmatrix} 1 & -0.588 & 0.006 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01247	0.0149	0.0255	$\begin{bmatrix} 1 & -0.6 & -0.002 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$
0.01706	0.01717	0.0298	$\begin{bmatrix} 1 & -0.61 & 0.01 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$
0.02165	0.0242	0.0302	$\begin{bmatrix} 1 & -0.603 & 0.013 \\ & 1 & 0.137 \\ & & 1 \end{bmatrix}$

0.02624	0.0271	0.0326	$\begin{bmatrix} 1 & -0.61 & -0.007 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.030823	0.0309	0.0326	$\begin{bmatrix} 1 & -0.61 & 0.007 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.0354	0.0346	0.0339	$\begin{bmatrix} 1 & -0.61 & -0.001 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04	0.0392	0.033	$\begin{bmatrix} 1 & -0.61 & -0.012 \\ & 1 & 0.151 \\ & & 1 \end{bmatrix}$
0.04459	0.0425	0.0336	$\begin{bmatrix} 1 & -0.616 & -0.012 \\ & 1 & 0.15 \\ & & 1 \end{bmatrix}$
0.04918	0.04936	0.0325	$\begin{bmatrix} 1 & -0.61 & 0.015 \\ & 1 & 0.141 \\ & & 1 \end{bmatrix}$
0.053765	0.0538	0.032	$\begin{bmatrix} 1 & -0.6 & 0.006 \\ & 1 & 0.149 \\ & & 1 \end{bmatrix}$
0.05835	0.05617	0.0311	$\begin{bmatrix} 1 & -0.61 & -0.01 \\ & 1 & 0.157 \\ & & 1 \end{bmatrix}$
0.06294	0.06139	0.0293	$\begin{bmatrix} 1 & -0.62 & -0.0003 \\ & 1 & 0.129 \\ & & 1 \end{bmatrix}$
0.06753	0.06389	0.0252	$\begin{bmatrix} 1 & -0.62 & -0.014 \\ & 1 & 0.139 \\ & & 1 \end{bmatrix}$
0.07212	0.0687	0.0189	$\begin{bmatrix} 1 & -0.59 & -0.06 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$

The conditional moments computed using the numerical integration method (that will be used in the value function solution) are:

f_t	$E\left[RP_{t+1}\vert f_t\right]$	$Var\left[RP_{t+1}\vert f_t\right]$	$\rho_{RP_{t+1}\vert f_t}$
0.007882	0.0103	0.019	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$
0.01247	0.0159	0.0252	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$
0.01706	0.0193	0.0294	$\begin{bmatrix} 1 & -0.613 & 0.0 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.02165	0.0227	0.0315	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.02624	0.0266	0.0323	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.03082	0.0309	0.0326	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.0354	0.0354	0.0328	$\begin{bmatrix} 1 & -0.6 & 0.0 \\ & 1 & 0.148 \\ & & 1 \end{bmatrix}$
0.04	0.04	0.0327	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.04459	0.04457	0.0327	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$
0.04918	0.04908	0.0326	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.147 \\ & & 1 \end{bmatrix}$

0.053765	0.0534	0.03235	$\begin{bmatrix} 1 & -0.61 & 0.00 \\ & 1 & 0.146 \\ & & 1 \end{bmatrix}$
0.05835	0.0573	0.03156	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.144 \\ & & 1 \end{bmatrix}$
0.06294	0.0606	0.0294	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.14 \\ & & 1 \end{bmatrix}$
0.06753	0.06414	0.0252	$\begin{bmatrix} 1 & -0.61 & 0.0 \\ & 1 & 0.13 \\ & & 1 \end{bmatrix}$
0.07212	0.0697	0.019	$\begin{bmatrix} 1 & -0.59 & 0.0 \\ & 1 & 0.12 \\ & & 1 \end{bmatrix}$

In this experiment, accuracy is not better than the one in experiment 3.

Based on the results above, we find that the results are insensitive to the number of grid points when the numbers for f_t and r_t are beyond 15 and 20 respectively, and the numbers for $ln(U_t)$ and $ln(N_t)$ are above 3 and 5, respectively.