Imperial College London

Business School

Essays in Life-Cycle Portfolio Choice

Yuxin Zhang

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy of Imperial College London

ABSTRACT

This dissertation presents three essays in life-cycle portfolio choice. Chapter 1 solves 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, it shows that, in the presence of stock market predictability, 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). Out-of-sample analysis supports these conclusions.

Chapter 2 studies the effect of observable predictors that imperfectly predict conditional expected stock returns on optimal life-cycle consumption and portfolio choice in the presence of undiversifiable labor income risk. Investors filter the unobservable expected stock returns from realized predictive variables and stock returns. Young stockholders hold more conservative

portfolios, better matching empirical observations, than models assuming a predictor perfectly delivering the conditional expected stock return or models assuming i.i.d. stock returns. Welfare losses from ignoring imperfect predictability can be substantial.

Chapter 3 uses different stock return predictors at quarterly frequency to solve for optimal consumption and portfolio choice in a life-cycle model with short-sales and borrowing constraints and undiversifiable labor income risk. Both wealth accumulation and asset allocation look similar qualitatively to their i.i.d. unconditional averages, but are quantitatively different and depend on predictors in different ways. Therefore, enhanced target-date funds (ETDFs) that condition saving and portfolio choice on predictor variables can lead to substantial welfare gains.

Contents

List of Figures

List of Tables

Declaration of Originality

'I herewith certify that this thesis constitutes my own work and that all material, which is not my own work, has been properly acknowledged'

Yuxin Zhang

Copyright

Yuxin Zhang

Acknowledgments

I am indebted to Alex Michaelides and Francisco Gomes for guidance and encouragement. I would also like to thank Franklin Allen, Harjoat Bhamra, Rustam Ibragimov, Alexandros Kostakis, David Miles, Lubos Pastor, Tarun Ramadorai and seminar participants at Imperial College London for many helpful comments. All remaining errors are my own.

Chapter 1

Stock Market Mean Reversion and Portfolio Choice over the Life Cycle

I. Introduction

In the absence of stock market predictability and in the presence of borrowing and short-sale constraints,[1] models with background labor income risk (e.g., Cocco, Gomes, and Maenhout (2005)) predict that households should invest a larger proportion of their savings in the stock market when young: future labor income acts like an implicit risk free asset crowding out riskless financial assets from household portfolios. This prediction resembles the advice given by financial planning consultants in recommending target-date funds (TDFs) that reduce exposure to stocks as retirement

[1] Deaton (1991) and Carroll (1997) propose models with background risks and potentially binding liquidity constraints ("buffer stock saving models") as the leading alternative to the classic Permanent Income-Life Cycle Hypothesis.

approaches.

How does the presence of stock market predictability affect these predictions? Various papers have analyzed the implications of stock market predictability 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 using Epstein–Zin (1989) and Weil (1990) preferences.[4] The results make the case for enhanced

[2]Kim and Omberg (1996), Brennan, Schwartz, and Lagnado (1997), Brandt (1999), Campbell and Viceira (1999), Balduzzi and Lynch (1999), Barberis (2000), Campbell, Cocco, Gomes, Maenhout, and Viceira (2001), Campbell, Chan, and Viceira (2003), and Wachter (2002) show that stock market exposure varies substantially as a response to the predictive factor(s).

[3]For instance, Heaton and Lucas (2000), Viceira (2001), Haliassos and Michaelides (2003), Gomes and Michaelides (2005), and Cocco, et al. (2005).

[4]Benzoni, Collin-Dufresne, and Goldstein (2007) investigate the implications of a cointegrating relation 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

target-date funds (ETDFs) that condition decisions on expected risk premia: introducing such funds generates improvements in investor welfare.

As in the buffer stock saving literature, optimal consumption is a concave function of liquid wealth, 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 independent and identically distributed (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

variation in interest rates and expected income growth to illustrate the effects on portfolio choice, while Koijen, Nijman, and Werker (2010) 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.

models with i.i.d. stock returns, namely that financial wealth tends to be the main predictor of life-cycle portfolio choice.

When calibrated to the observed dividend yield as a factor of predictability, the asset allocation profiles retain the 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.[5] Taken together, these findings make the case for enhanced TDFs (ETDFs) that condition on the market timing ability of the investor.

From all the underlying correlations studied, the main correlation that is found to quantitatively affect our conclusions is the correlation between permanent earnings shocks and the stock market innovation. Financial

[5] Aggressive market timing behavior is similar to the behavior predicted in Brennan et al. (1997) and Barberis (2000), models that do not feature undiversifiable labor income uncertainty.

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 persistence parameter and the correlation between the factor and stock market innovation and find that they do not substantially alter the asset allocation and wealth accumulation profiles.

To emphasize the results with regards to ETDFs, we next 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.

We also perform out-of-sample analysis adapting Lan's (2015) methodology to our life cycle setting and using the realized dividend yield and stock returns to simulate life cycle wealth and utility in 2 subperiods: 1985–2014 and 1999–2014. We find that welfare is statistically and economically significantly higher when comparing the baseline model to either the Vanguard or i.i.d. stock returns model, especially in the 1999–2014 period. The analysis demonstrates the value in devising ETDFs for individual investors and calls for further research in this area.

The paper is organized as follows: Section II describes the theoretical model, outlines the numerical solution algorithm and discusses the parameter choices for the calibration. Section III discusses the effects of stock market mean reversion by comparing the benchmark results to the i.i.d. stock returns model. Section IV discusses hedging demands and how different correlation changes also affect wealth accumulation, while Section V briefly discusses the implications of model parameter uncertainty. Section VI discusses the implications of the model for lifestyle funds and Section VII concludes.

II. The Model

Time is discrete, there is one nondurable 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 $\tilde{R}$. 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 nondurable good. Following Deaton (1991), cash on hand in period t is denoted by $X_t = S_{t-1}\tilde{R}_t + B_{t-1}R_f + Y_t$. 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:

A. 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

$$V_t = \left\{ (1-\beta) C_t^{1-1/\psi} + \beta \left(\mathrm{E}_t (p_{t+1} V_{t+1}^{1-\gamma} + b(1-p_{t+1}) X_{t+1}^{1-\gamma}) \right)^{\frac{1-1/\psi}{1-\gamma}} \right\}^{\frac{1}{1-1/\psi}}, \quad (1.1)$$

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}.

B. Labor Income Process

Following a relatively standard specification in the literature (Carroll (1997)), the labor income process before retirement is given by

$$Y_{it} = Y_{it}^{p} U_{it}, \tag{1.2}$$

$$Y_{it}^{p} = \exp(g(t, Z_{it})) Y_{it-1}^{p} N_{it}, \tag{1.3}$$

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 i.i.d. with mean $\{-0.5 \times \sigma_u^2, -0.5 \times \sigma_n^2\}$, and variances σ_u^2 and σ_n^2, respectively. The natural 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 $Y_{it} = \lambda Y^p_{iK}$, 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 the Panel Study Income Dynamics (PSID) data. We choose not to model explicitly the returns from housing following the empirical evidence (e.g., 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 (2015)). 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).

C. Liquidity Constraints

Borrowing and short sales of stocks are not allowed: $B_t \geqq 0$ and $S_t \geqq 0$ to avoid the counterfactual implication that households lever up to invest in the stock market. The share of wealth in stocks (α_t) is defined as $S_t/(S_t + B_t)$.

D. 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

$$r_{t+1} - r_f = f_t + z_{t+1}, \tag{1.4}$$

$$f_{t+1} = \mu + \phi(f_t - \mu) + \varepsilon_{t+1}, \tag{1.5}$$

where the two innovations $\{z_{t+1}, \varepsilon_{t+1}\}$ are i.i.d. Normal variables with mean equal to 0 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.[6]

E. Numerical Solution

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

$$x_{it+1} = \frac{Y_{it}^{p}}{Y_{it+1}^{p}}(r_{t+1}\alpha_{it} + r_f(1-\alpha_{it}))x_{it} + U_{it+1}. \tag{1.6}$$

An Internet Appendix (available at www.jfqa.org) 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

[6]The i.i.d. returns specification is the one found in recent papers with either the constant relative risk aversion preferences (Cocco et al. (2005)) or the Epstein–Zin–Weil preferences (Gomes and Michaelides (2005) or Cooper and Zhu (2016)).

because of the strong persistence in the factor f_t that requires a substantial number of grid points to replicate key moments: the persistence, the conditional variance and the conditional (1-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 first-order autoregression (AR(1)) case. 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. The Internet Appendix provides extensive details from different experiments to test numerical accuracy.

F. 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 for σ_u and 0.1 for σ_n. The deterministic component of labor income is identical to the one used by most life cycle papers in this literature (Cocco et al. (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% 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 modeling tax-deferred saving through retirement accounts.

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. We 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 mean reversion system (3.3) and (3.4) using three data sources:

the data from Pastor and Stambaugh (2012),[7] the publicly available data set on Robert Shiller's Web site,[8] and by 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 $\phi = 0.9$. 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 0. Most estimates in the literature are toward the higher negative

[7] We thank Lubos Pastor for kindly providing this data set.

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

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) $\sigma_{\varepsilon} = 0.007$[9] for the baseline model but we also find it useful to compare results with the case when the factor is more volatile ($\sigma_{\varepsilon} = 0.015$). 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 $\mu = 0.04$ (an unconditional equity premium at

[9] 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.

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 0. 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 et al. (2005)). We therefore set the correlation between transitory labor income shocks and stock returns equal to 0. 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. (2001), 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.

III. 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?

The Internet Appendix presents the consumption function that has the familiar shape from the buffer stock saving literature without risky asset choice (Deaton (1991), Carroll (1997)): below a cutoff point x^* no saving takes place, while the marginal propensity to consume falls quickly beyond x^*. Once the constraint stops binding, the saving reaction to the factor realization is not quantitatively important.

The policy functions for the share of wealth in stocks illustrate the large dependence of portfolios on the factor realization (age 25 (Graph A of Figure 1), age 55 (Graph B), and age 75 (Graph 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 upward. 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 (e.g., comparing age 25 with age 55 or age 75). Moreover, the presence of labor income prevents the portfolio from depicting the yo-yo type behavior found in Brennan et al.(1997), where the portfolios move from 0 to 1 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 volatility of the dividend yield.

Figure 2 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, Kinniry, Aliaga-Diaz, Patterson, and DiJoseph (2013)) for 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[10] is similar across all three models (Graph A of Figure 2). This is slightly surprising since Graph 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

[10]The model is simulated 2000 times for 500 individual life histories.

simulating based on different initial realizations of the factor and then averaging over them. For most of these factor states, 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 ($\sigma_\varepsilon = 0.015$). This is illustrated in Graph D of Figure 2: 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 3 plots what happens over the life cycle when starting from the lowest (state = 1), 6th, 10th 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 et al. (1997) that features no labor income risk since the 0 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 portfolio behavior between 0 to full allocation in the stock market arises once again.

We have also experimented with increasing the volatility of the stock return innovation (σ_z) by 10%, 20% and 50% as a way to capture the uncertainty over the coefficient relating the equity premium to the dividend yield. We find that the share of wealth in stocks is reduced throughout the life cycle relative to the baseline mean reversion case but portfolio behavior remains qualitatively unchanged and do not report these results.

IV. Hedging Demands

How do these results change when the correlations between the different innovations vary? We next use the model to quantitatively assess the magnitude of such hedging demands.

A. Variation in Correlations

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 0. In the Shiller dividend yield data set, we have found this correlation to be close to 0, contrary to our findings using CRSP data. Setting the correlation to 0 can therefore be useful in assessing the possible range of hedging demands that might be generated from this correlation.

We also evaluate hedging demands when changing the correlation between permanent earnings shocks and stock market innovations ($\rho_{z,n}$). 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. Bonaparte et al. (2014) find that this correlation can vary for different households from -0.6 to 0.6 and we therefore use 0.5 to investigate how the results change.

There is no known empirical estimate for $\rho_{n,\varepsilon}$ (the correlation between the factor innovation and the permanent labor income shock) in the literature. There are potentially some a priori reasons to expect it not to be statistically different from 0 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. 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 0. 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: setting $\rho_{n,\varepsilon}$ equal to 0.15 satisfies this constraint.

The wealth accumulation and mean shares of wealth in stocks over the life cycle are depicted in Graph A of Figure 4, and Graph B, respectively.[11] When the correlation between the factor and the stock return innovation is set to 0 ($\rho_{z,\varepsilon} = 0.0$) from $\rho_{z,\varepsilon} = -0.6$, the share of wealth in stocks is lower on average between 0% and 5% over different parts of the working life cycle (Graph B) and the effects are larger when the factor is more volatile (Graph D), leading to a discernibly lower average wealth accumulation (Graph C). It can be shown that the variance of longer term stock returns is higher when

[11] 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 since portfolios in the i.i.d. model do not exhibit any time variation in response to the factor realizations. For a specific correlation (ρ) we can compute these hedging demands by comparing the two simulated profiles and compute the percentage differences:
$hedg(\rho) = 100 * (\alpha_{\text{factor}}(\rho) - \alpha_{\text{I.I.D.}}(\rho)) \;/\; \alpha_{\text{I.I.D.}}(\rho)$. For space considerations we only report the simulated shares of wealth in stocks generated from different correlations of interest.

$\rho_{z,\varepsilon} = 0.0$ 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 ($\rho_{n,\varepsilon}$) does not materially affect the average share of wealth in stocks (Graphs B and D of Figure 4 for the two cases), and therefore also 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 5 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 (Graph A) and the higher factor variability case (Graph 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 0 for the highest factor realizations that signal high future expected stock returns.

V. Model Parameter Uncertainty

Model uncertainty is a feature in the stock market predictability literature as emphasized, for example, in Xia (2001), Brandt, Goyal, Santa-Clara, and Stroud (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 values: the baseline value (0.9) with a 50% probability, a lower persistence (0.7) with 25% probability and a higher persistence (0.93) with a 25% probability. This is necessarily asymmetric to respect factor stationarity. For $\rho_{z,\varepsilon}$ we assume the investor expects two values: the baseline (-0.6) with a 50% probability and an extreme equal to 0 with a 50% probability.

The recursive model is modified by simply assuming that the investor is aware of this uncertainty and takes expectations over these three possibilities, weighting each case with the assumed probability. In the case of $\rho_{z,\varepsilon}$, for example, the current value function is the result of an optimal saving-portfolio choice that expects next period returns to be generated with $\rho_{z,\varepsilon} = -0.6$ and with $\rho_{z,\varepsilon} = 0.0$ with a 50% probability each. The value function is not otherwise affected since investors are assumed not to learn the true value of this parameter. We then simulate stock returns assuming a 50-50 realization of $\rho_{z,\varepsilon}$. It turns out that results are not quantitatively affected if instead we simulate by fixing $\rho_{z,\varepsilon}$ at either -0.6 or 0.0.

Figure 6 shows the results from introducing model uncertainty in these particular ways. Average portfolios are slightly more balanced over the life cycle (Graphs 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 Graph C). The effect is larger for the case where factor variability is larger (compare Graphs D and B), but the quantitative average effects are not substantial relative to the no uncertainty cases.

VI. Are Lifestyle Funds Optimal?

A. 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 retiring in 2008 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 life cycle. 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 7 produces such a diagram by assuming different factor realizations starting from each of the 15 initial states that persist throughout the life cycle. Graph B produces the same information by comparing the highest and lowest asset allocation experiences relative to the Vanguard recommendation and Graphs 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.

B. 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 $v_n\left(x_{it}, f_t\right)$ be the value function for a new model. We compute a measure of welfare change for a particular age group (AGE) as:

$$\overline{\mu}_{\text{AGE}} = \text{average of } \left[\left(\frac{v_n(x_{it}, f_t)}{v_0(x_{it}, f_t)} \right)^{\frac{1}{1-\gamma}} - 1 \right], \text{ for all } i \in I_{\text{AGE}} \text{ and all factor states} \tag{1.7}$$

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 8 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. Graph A reports the results for the baseline model ($\sigma_\varepsilon = 0.007$) and Graph B for the higher factor volatility model ($\sigma_\varepsilon = 0.015$). Graph 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 (Graph B of Figure 2). With higher factor volatility (Graph B) these effects are magnified as the differences across the average portfolio shares further diverge (Graph D of Figure 2).

We make two observations based on the results in Figure 8. First, the welfare losses are economically significant: they can reach 2%–7% 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 3 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 3 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 9 for the baseline case ($\sigma_{\varepsilon} = 0.007$ for Graphs A and B) and for the higher factor volatility case (Graphs C and D).

The Vanguard recommendation generates substantially more volatility in consumption over the working part of the life cycle (Graphs B and D of Figure 9), despite generating mildly higher average consumption in the baseline case (Graph 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 8.

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 (e.g., compare Graph C of Figure 9 with Graph A). 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 (Graph D relative to Graph 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.

C. Out-of-Sample Analysis

We next compare how realized wealth and certainty equivalents evolve out of sample, partly adapting the calculations in Lan (2015) for our model. Specifically, for the 2 cases (high and low σ_ε) we start simulating for every age using the realized dividend yield and stock returns between 1985 and 2014 and between 1999 and 2014. For instance, given a simulated initial distribution of assets per age in 1985, from 1986 onward we use the realized dividend yield to pick the relevant factor state and the realized stock return to shock financial asset returns. We save financial wealth for every age group between 1986 and 2014, as well as the CEQs as defined in expression (3.6). For an investor at age 30, for example, we track the evolution of individual wealth and CEQs over the 1986 –2014 period and average in the cross section every year.

It might be helpful to start the discussion by plotting the wealth levels for different models. Figure 10 shows how the mean wealth evolves for different starting ages according to the different models. As can be seen, the baseline models typically generate higher mean wealth than either the i.i.d. or Vanguard rule would predict. Moreover, younger households display the mean wealth accumulation rising over time (and over the life cycle), while older households can display the decreasing wealth after retirement.

Tables 1 and 2 report the mean differences across CEQs for different models ($\sigma_\varepsilon = 0.007$ and $\sigma_\varepsilon = 0.015$, respectively). Specifically, as in Lan (2015), the mean differences across CEQs for each age are computed over the 1985–2014 (Panel A) and 1999–2014 (Panel B) periods. The t-statistics are computed using a Newey–West (1987) procedure to correct for serial correlation. One conclusion from Table 1 is that the increase in welfare is not always statistically significant (compare the baseline model with the i.i.d. stock returns model in the 1985–2014 period). Nevertheless, the baseline model statistically and economically outperforms the Vanguard model in both periods, while the baseline model statistically and economically outperforms both the i.i.d. and Vanguard models in the 1999–2014 period. These conclusions are strengthened when the factor is more volatile (Table 2) since in that instance the dividend yield as a signal generates more aggressive investor behavior. These results emphasize the importance of searching for good signals when devising long-term asset allocation strategies, but also how the same variable can generate different conclusions over different time periods.

D. Rules of Thumb

Can a rule of thumb be devised in the presence of factor predictability that dominates widely used recommendations like the Vanguard one? Based on our results we devise approximate asset allocation 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 following different rules of thumb are applied. If the factor realization is above 5%, the investor allocates all financial wealth to the stock market. If the factor realization is between 2% and 5%, then the Vanguard rule is applied. If the factor realization is below 2%, then the maximum of 0 and the Vanguard rule minus 50% is chosen. We also report results for changing the thresholds between 1% and 6%, rather than between 2% and 5%, and also by adjusting the allocation relative to the Vanguard recommendation by 30% instead of 50%. The consumption/saving rules are chosen optimally in both cases. We think a richer model of retirement uncertainties is needed to make similar recommendations for the retirement period and we leave this extension for future work.

Tables 3 and 4 report the mean differences across CEQs for the 2 models ($\sigma_\varepsilon = 0.007$ and $\sigma_\varepsilon = 0.015$, respectively). A positive percentage means that the rule of thumb dominates the Vanguard recommendation. We can see that over most time periods and experiments there is a statistically and economically significant welfare improvement relative to the Vanguard recommendations.

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 10% 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 but we do not report these welfare comparisons due to space considerations.

VII. 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 magnitude of 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. 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 learning mechanism about the true underlying model through either a Kalman filtering or Bayesian learning approach (Brandt et al. (2005), Pastor and Stambaugh (2012)), an explicit treatment of housing and introducing time-varying volatility and risk aversion (e.g., through a stochastic discount factor). All these extensions will require additional computational power to achieve the desired required solution accuracy but will further improve our understanding of life-cycle portfolio choice under uncertainty and offer scientific advice to billions of households increasingly making their own individual financial decisions.

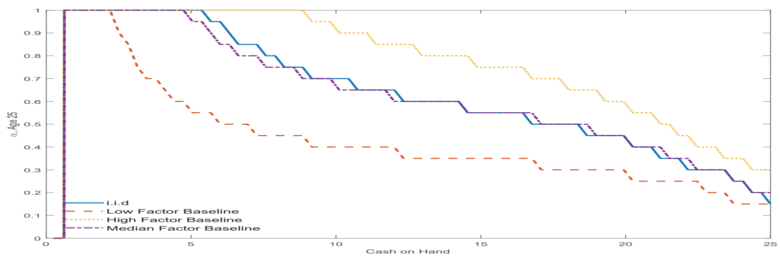
VIII. Appendix: Figures

FIGURE 1

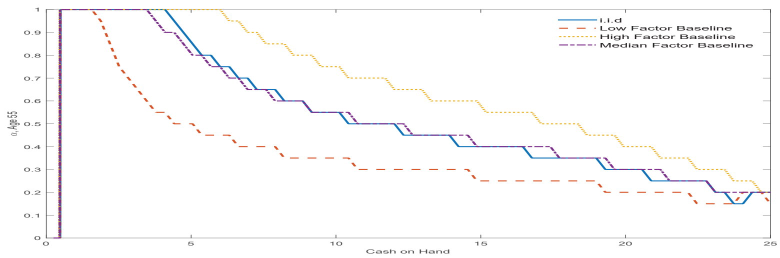
Policy Function Comparison

Figure 1 presents the policy function comparisons between the i.i.d. stock returns model (solid line) and three factor states from the baseline mean reversion model (the median factor (dash-dot line), the lowest factor (dashed line) and the highest factor (dotted line)).

Graph A. Share of Wealth in Stocks (α), Age 25



Graph B. Share of Wealth in Stocks (α), Age 55



Graph C. Share of Wealth in Stocks (α), Age 75

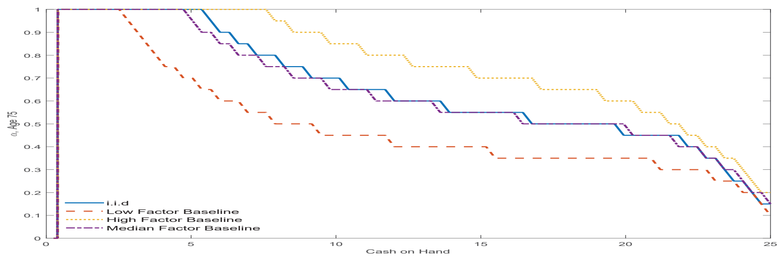
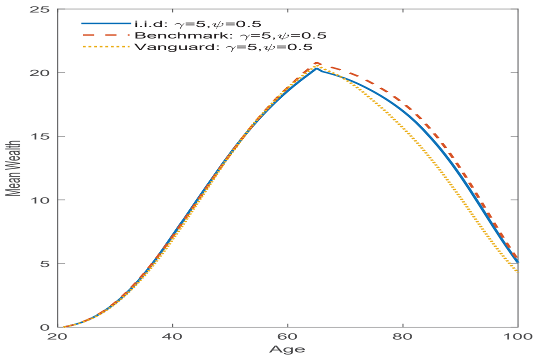


FIGURE 2

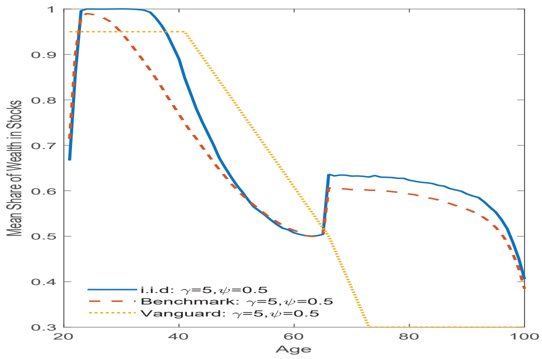
Life-Cycle Profile Comparison

Figure 2 presents the life-cycle profile comparison between the benchmark stock market mean reversion results (dashed line), the Vanguard recommendation (dotted line) and the i.i.d. stock returns model (solid line).

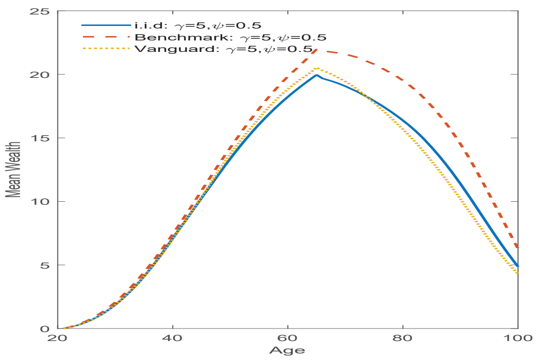
Graph A. Mean Wealth: i.i.d. versus Benchmark and Vanguard (σ_ϵ = 0.007)



Graph B. Mean Share of Wealth in Stocks: i.i.d. versus Benchmark and Vanguard (σ_ϵ = 0.007)



Graph C. Mean Wealth: i.i.d. versus Benchmark and Vanguard (σ_ϵ = 0.015)



Graph D. Mean Share of Wealth in Stocks: i.i.d. versus Benchmark and Vanguard (σ_ϵ = 0.015)

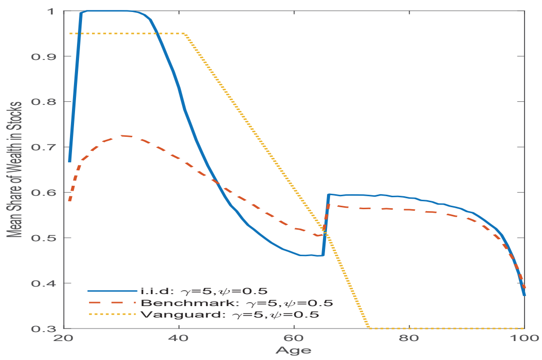
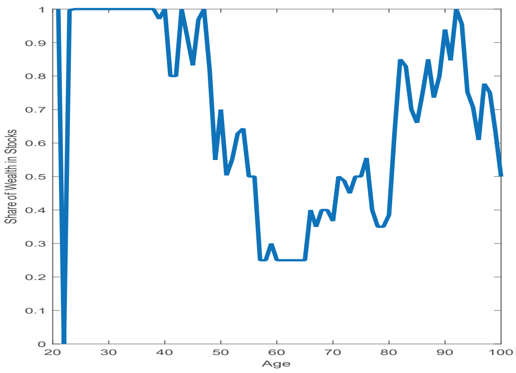


FIGURE 3

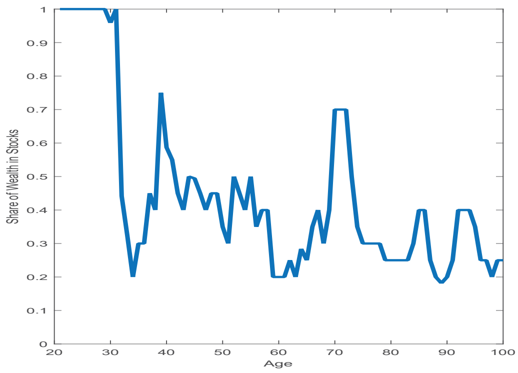
Life-Cycle Individual Share of Wealth in Stocks

Figure 3 presents the life-cycle individual share of wealth in stocks for different initial factors starting from age 21 for the benchmark stock market mean reversion model.

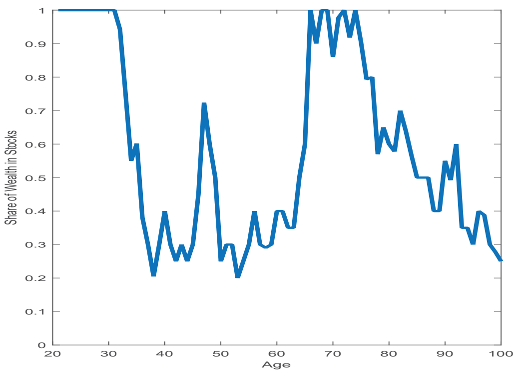
Graph A. Benchmark (factor state = 1)



Graph B. Benchmark (factor state = 6)



Graph C. Benchmark (factor state = 10)



Graph D. Benchmark (factor state = 15)

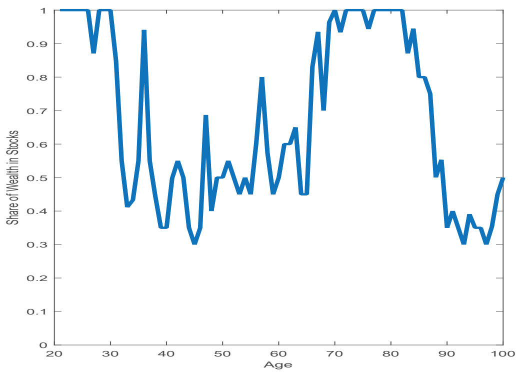
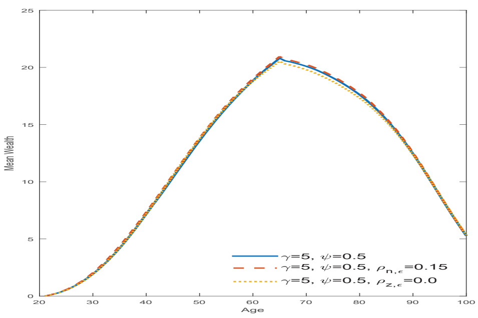


FIGURE 4

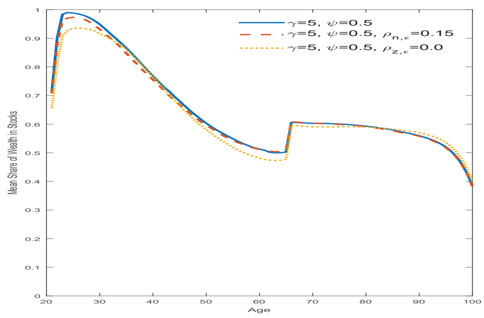
Life-Cycle Profiles for Average Wealth and Portfolio Shares for Different Parameters

Figure 4 presents 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.

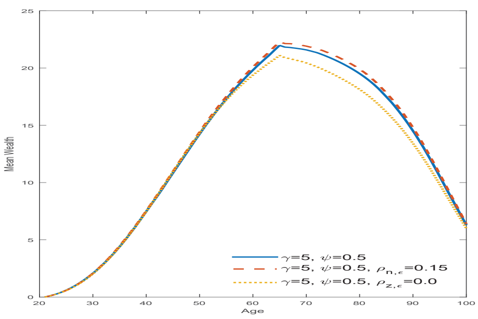
Graph A. Mean Wealth with Different Parameters (σ_ϵ = 0.007)



Graph B. Mean Share of Wealth in Stocks with Different Parameters (σ_ϵ = 0.007)



Graph C. Mean Wealth with Different Parameters (σ_ϵ = 0.015)



Graph D. Mean Share of Wealth in Stocks with Different Parameters (σ_ϵ = 0.015)

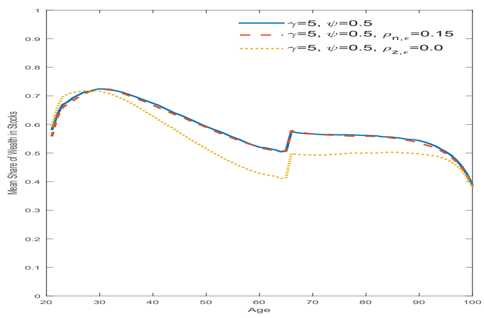
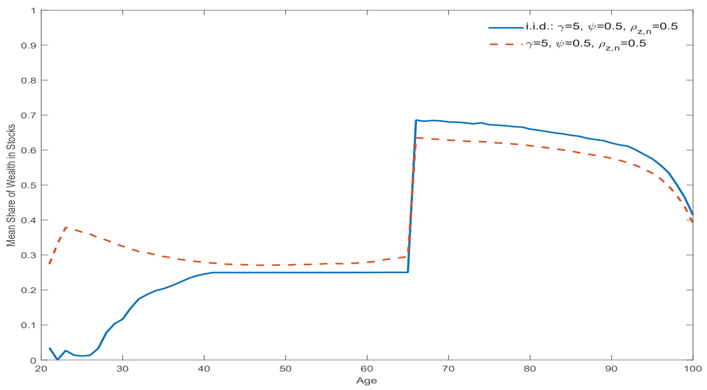


FIGURE 5

Life-Cycle Mean Portfolio Shares in Stocks

Figure 5 presents 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.

Graph A. Life-Cycle Mean Portfolio Shares for Different Parameters ($\sigma_\epsilon = 0.007$)



Graph B. Life-Cycle Mean Portfolio Shares for Different Parameters ($\sigma_\epsilon = 0.015$)

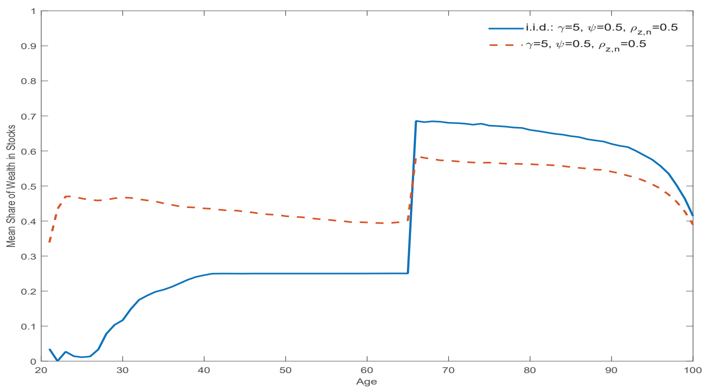
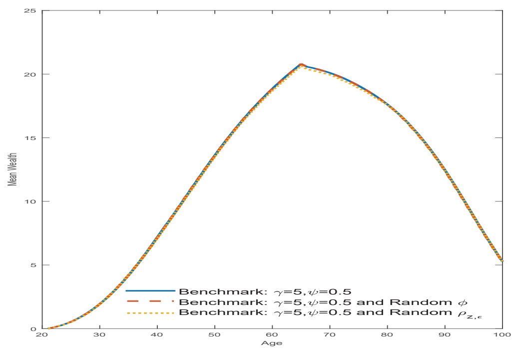


FIGURE 6

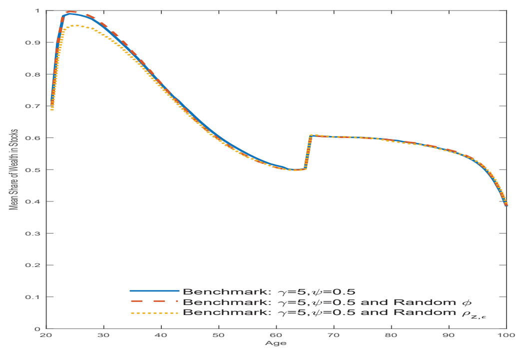
Life-Cycle Profiles for Average Wealth and Portfolio Shares for Different Parameters

Figure 6 presents the 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 $\rho_{z,\varepsilon}$.

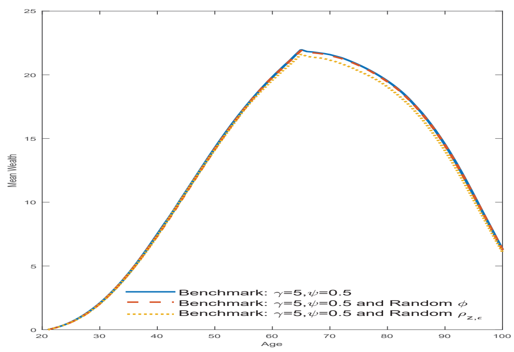
Graph A.
Mean Wealth with random ϕ and random $\rho_{z,\epsilon}$ ($\sigma_\epsilon = 0.007$)



Graph B. Mean Share of Wealth in Stocks with random ϕ and random $\rho_{z,\epsilon}$ ($\sigma_\epsilon = 0.007$)



Graph C.
Mean Wealth with random ϕ and random $\rho_{z,\epsilon}$ ($\sigma_\epsilon = 0.015$)



Graph D. Mean Share of Wealth in Stocks with random ϕ and random $\rho_{z,\epsilon}$ ($\sigma_\epsilon = 0.015$)

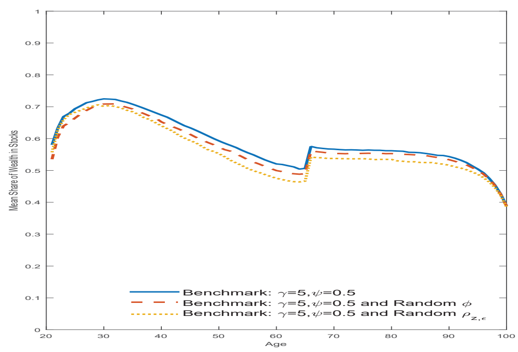
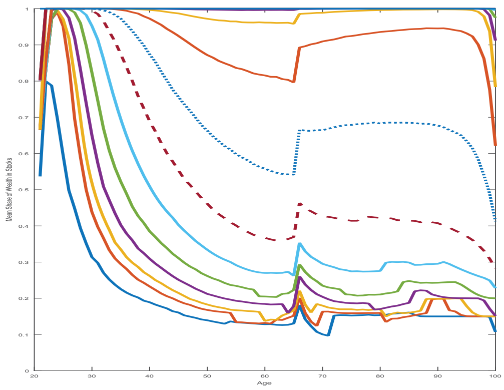


FIGURE 7

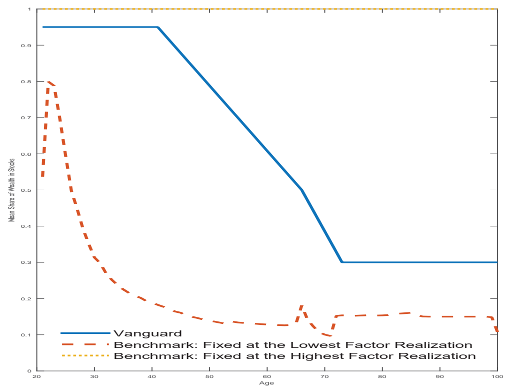
Portfolio Shares with Different Constant Factor Realization

Figure 7 presents average portfolio shares with different constant factor realizations for the stock market mean reversion model. Graphs B and D report the highest and lowest states and also report the Vanguard recommendation that lies in between the two extremes.

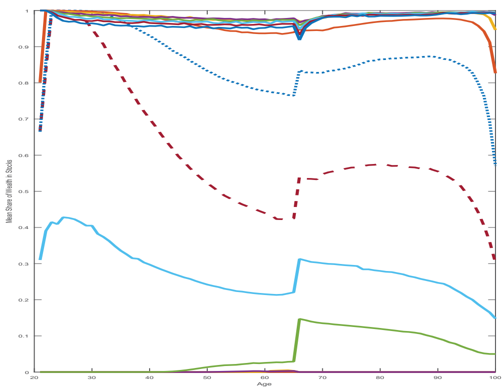
Graph A. Mean Share of Wealth in Stocks for different fixed factor realizations (σ_ϵ = 0.007)



Graph B. Mean Share of Wealth in Stocks: Vanguard versus Benchmark with fixed factor realization (σ_ϵ = 0.007)



Graph C. Mean Share of Wealth in Stocks for different fixed factor realizations (σ_ϵ = 0.015)



Graph D. Mean Share of Wealth in Stocks: Vanguard versus Benchmark with fixed factor realization (σ_ϵ = 0.015)

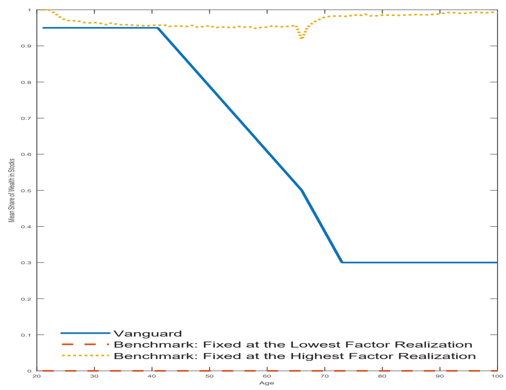
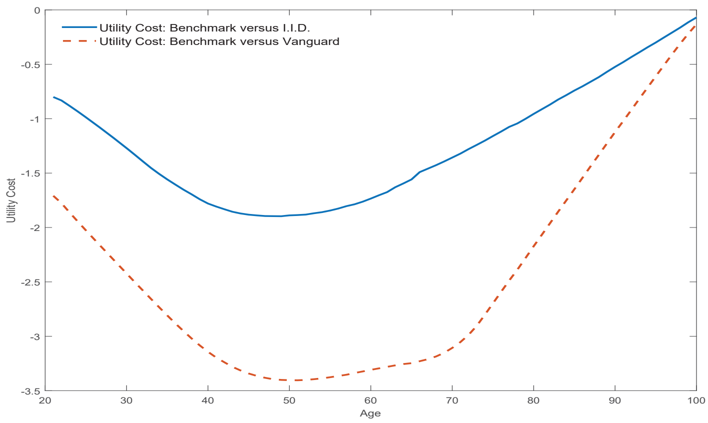


FIGURE 8

Welfare Evaluation

Figure 8 presents average consumption certainty equivalents 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.

Graph A. Consumption Certainty Equivalent Comparison (σ_ϵ = 0.007)



Graph B. Consumption Certainty Equivalent Comparison (σ_ϵ = 0.015)

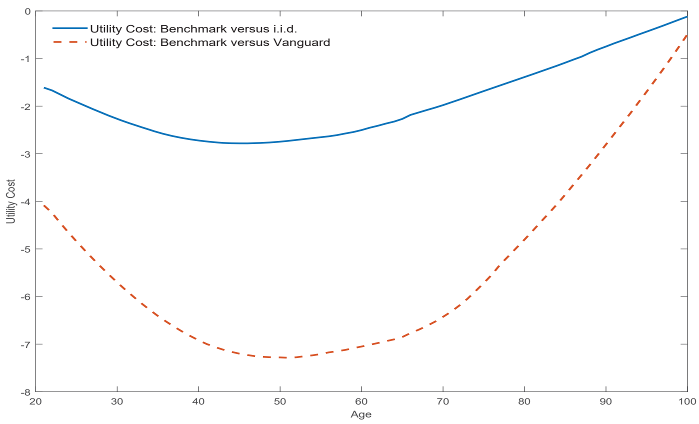
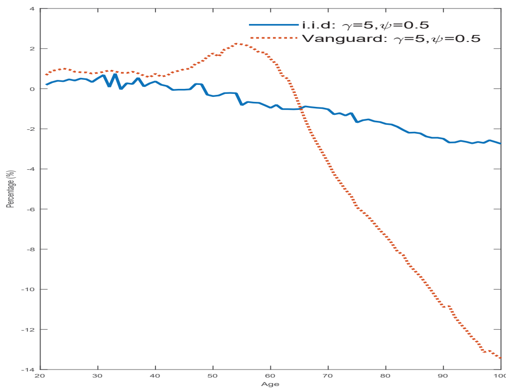


FIGURE 9

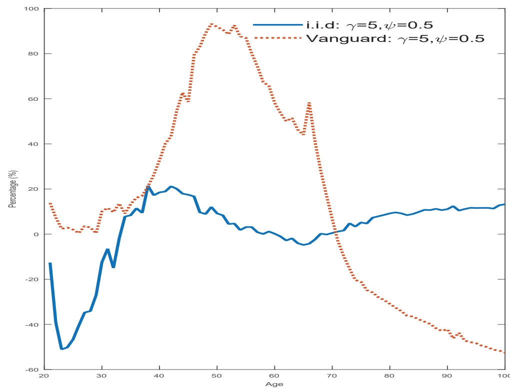
Consumption Evaluation

Figure 9 presents percentage changes 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.

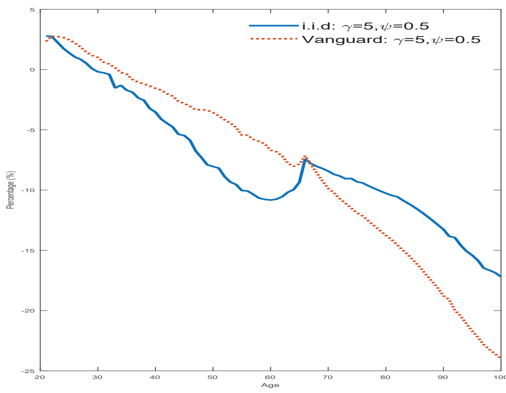
Graph A. Mean Consumption Change: I.I.D. versus Benchmark and Vanguard versus Benchmark (σ_ϵ = 0.007)



Graph B. S.D. Consumption Change: I.I.D. versus Benchmark and Vanguard versus Benchmark (σ_ϵ = 0.007)



Graph C. Mean Consumption Change: I.I.D. versus Benchmark and Vanguard versus Benchmark (σ_ϵ = 0.015)



Graph D. S.D. Consumption Change: I.I.D. versus Benchmark and Vanguard versus Benchmark (σ_ϵ = 0.015)

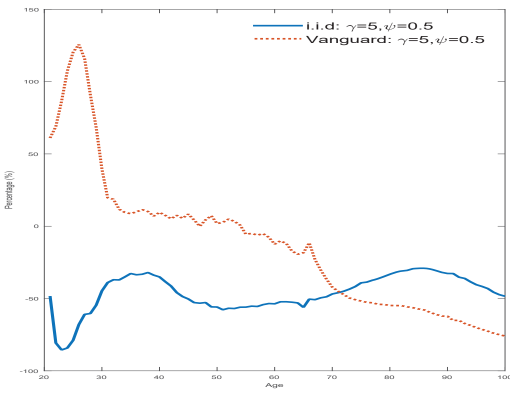
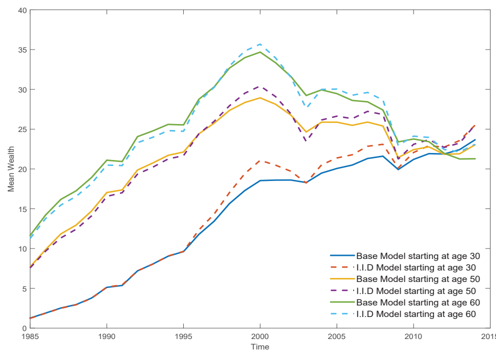


FIGURE 10

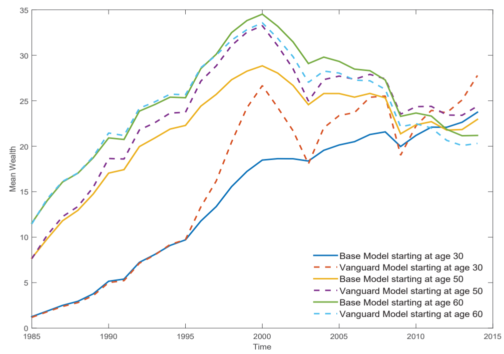
Mean Wealth Profiles for Different Starting Investor Ages in 1985

Figure 10 presents mean wealth profiles for different starting investor ages in 1985. The initial wealth distribution for all ages is generated through a simulation of the baseline model. From that same initial wealth distribution in 1985, investors start behaving according to the baseline, i.i.d. and Vanguard models, while realized dividend yields and stock returns are used from 1986 to 2014 to generate individual wealth profiles.

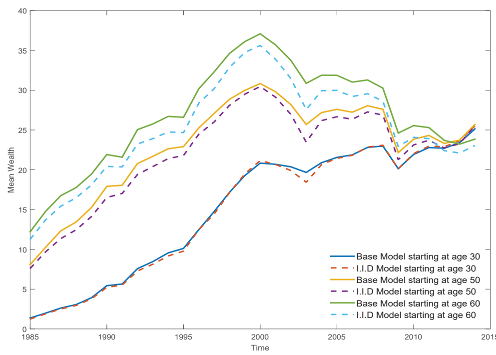
Graph A. Baseline and I.I.D. Models for Different Starting Ages (σ_ϵ = 0.007)



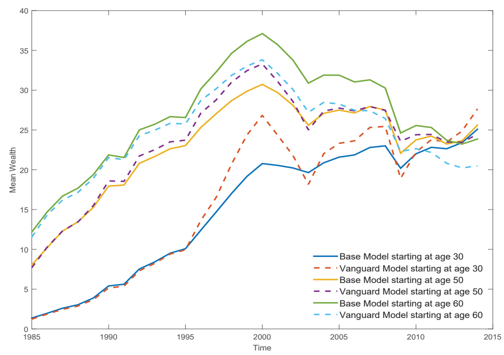
Graph B. Baseline and Vanguard Models for Different Starting Ages (σ_ϵ = 0.007)



Graph C. Baseline and I.I.D. Models for Different Starting Ages (σ_ϵ = 0.015)



Graph D. Baseline and Vanguard Models for Different Starting Ages (σ_ϵ = 0.015)



IX. Appendix: Tables

TABLE 1

Out of Sample Analysis with $\sigma_\varepsilon = 0.007$

Table 1 presents the mean differences in certainty equivalents (CEQs) between the benchmark model (with $\sigma_\varepsilon = 0.007$) and the i.i.d. stock returns and Vanguard models, respectively. The t-statistics, based on the asymptotic distribution, are in parentheses. In Panel A, for each starting age and initial wealth distribution based on an initial life-cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014.

Starting	*Panel A. 1985–2014*		*Panel B. 1999–2014*	
Age	I.I.D.	Vanguard	I.I.D.	Vanguard
30	0.3(0.3)	-1.5(-1.6)	-1.5(-5.1)	-4.9(-13.6)
40	-0.3(-0.4)	-0.5(-0.6)	-1.8(-4.7)	-6.3(-12.4)
45	-0.1(-0.1)	0.0(0.0)	-1.8(-4.6)	-6.0(-11.0)
50	0.3(0.2)	0.9(0.8)	-1.6(-4.3)	-5.7(-9.9)
55	0.7(0.5)	-0.1(-0.2)	-1.6(-4.1)	-5.5(-10.8)
60	0.0(-0.0)	-3.3(-11.6)	-1.8(-4.6)	-5.3(-13.9)
65	-0.9(-0.6)	-7.5(-9.5)	-3.0(-5.9)	-4.6(-11.0)
70	-0.6(-0.4)	-12.8(-11.9)	-3.3(-6.5)	-5.2(-64.0)

TABLE 2

Out of Sample Analysis with $\sigma_\varepsilon = 0.015$

Table 2 presents the mean differences in certainty equivalents (CEQs)between the benchmark model (with $\sigma_\varepsilon = 0.015$) and the i.i.d. stock returns and Vanguard models, respectively. The t-statistics, based on the asymptotic distribution, are in parentheses. In Panel A, for each starting age and initial wealth distribution based on an initial life cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014.

Starting	*Panel A. 1985–2014*		*Panel B. 1999–2014*	
Age	I.I.D.	Vanguard	I.I.D.	Vanguard
30	-5.8(-13.5)	-6.2(-16.0)	-6.5(-15.0)	-8.6(-19.6)
40	-6.4(-10.4)	-5.5(-15.7)	-6.7(-14.9)	-9.7(-20.4)
45	-6.6(-8.9)	-5.4(-13.9)	-6.6(-16.4)	-9.6(-18.0)
50	-6.8(-8.4)	-5.3(-10.6)	-6.5(-18.2)	-9.4(-16.9)
55	-7.3(-9.8)	-7.3(-15.7)	-6.5(-16.9)	-9.5(-26.5)
60	-8.0(-10.9)	-10.1(-13.5)	-7.1(-17.8)	-9.8(-29.0)
65	-8.2(-8.6)	-13.5(-16.7)	-9.1(-12.7)	-10.0(-15.9)
70	-9.4(-11.1)	-19.7(-12.6)	-0.1(-13.8)	-10.8(-21.0)

TABLE 3

Out of Sample Analysis with $\sigma_\varepsilon = 0.007$

Table 3 presents the mean differences in certainty equivalents (CEQs) between the Vanguard model (with $\sigma_\varepsilon = 0.007$) and the different rule-of-thumb, respectively. The t-statistics, based on the asymptotic distribution, are in parentheses. In Panel A, for each starting age and initial wealth distribution based on an initial life-cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014. U5L2R3: If the dividend yield is greater than 5%, then set the share of wealth in stocks to 100%. If the dividend yield is below 2%, then take the maximum of the Vanguard recommendation minus 30% and 0. Otherwise, follow the Vanguard recommendation. U5L2R5: Same as U5L2R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%. U6L1R3: Same as U5L2R3 but the thresholds of inaction are between 1% and 6% instead of 2% and 5%. U6L1R5: Same as U6L1R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%.

Starting	Panel A: 1985–2014			
Age	U5L2R3	U6L1R3	U5L2R5	U6L1R5
30	1.5(4.5)	0.8(21.9)	1.8(4.2)	0.8(20.6)
40	1.4(3.1)	0.7(61.3)	1.7(3.1)	0.7(69.6)
45	1.3(3.3)	0.7(9.4)	1.6(3.2)	0.7(9.6)
50	1.7(3.2)	0.6(6.3)	2.4(2.5)	0.6(6.1)
55	1.6(3.5)	0.5(11.2)	1.5(3.1)	0.5(11.1)
60	1.0(3.2)	0.5(3.2)	0.9(2.9)	0.5(3.24)
65	1.2(3.7)	0.8(4.2)	1.2(3.6)	0.9(4.30)
70	2.0(6.5)	1.4(7.6)	2.1(6.8)	1.4(8.1)

TABLE 3 (Continued)

Out of Sample Analysis with $\sigma_\varepsilon = 0.007$

Table 3 presents the mean differences in certainty equivalents (CEQs) between the Vanguard model (with $\sigma_\varepsilon = 0.007$) and the different rule-of-thumb, respectively. The t-statistics, based on the asymptotic distribution, are in parenthesis. In Panel A, for each starting age and initial wealth distribution based on an initial life cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014. U5L2R3: If the dividend yield is greater than 5%, then set the share of wealth in stocks to 100%. If the dividend yield is below 2%, then take the maximum of the Vanguard recommendation minus 30% and 0. Otherwise, follow the Vanguard recommendation. U5L2R5: Same as U5L2R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%. U6L1R3: Same as U5L2R3 but the thresholds of inaction are between 1% and 6% instead of 2% and 5%. U6L1R5: Same as U6L1R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%.

Starting	Panel B: 1999–2014			
Age	U5L2R3	U6L1R3	U5L2R5	U6L1R5
30	1.7(19.2)	0.8(95.6)	0.0(15.2)	0.9(109.1)
40	2.5(12.1)	0.8(48.5)	3.4(9.91)	0.9(32.2)
45	2.8(10.2)	0.8(32.5)	3.7(8.4)	0.8(49.8)
50	3.0(8.6)	0.8(44.2)	4.0(7.3)	0.8(71.7)
55	3.1(8.0)	0.7(172.4)	4.2(7.2)	0.7(38.4)
60	3.5(10.2)	0.7(20.5)	4.7(8.4)	0.7(18.7)
65	4.4(10.5)	0.9(39.7)	6.3(8.5)	0.9(38.4)
70	4.9(10.9)	1.2(66.1)	5.3(10.5)	1.3(57.0)

TABLE 4

Out of Sample Analysis with $\sigma_\varepsilon = 0.015$

Table 4 presents the mean differences in certainty equivalents (CEQs) between the Vanguard model (with $\sigma_\varepsilon = 0.015$) and the different rule-of-thumb, respectively. The t-statistics, based on the asymptotic distribution, are in parenthesis. In Panel A, for each starting age and initial wealth distribution based on an initial life-cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014. U5L2R3: If the dividend yield is greater than 5%, then set the share of wealth in stocks to 100%. If the dividend yield is below 2%, then take the maximum of the Vanguard recommendation minus 30% and 0. Otherwise, follow the Vanguard recommendation. U5L2R5: Same as U5L2R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%. U6L1R3: Same as U5L2R3 but the thresholds of inaction are between 1% and 6% instead of 2% and 5%. U6L1R5: Same as U6L1R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%.

Starting	Panel A: 1985–2014			
Age	U5L2R3	U6L1R3	U5L2R5	U6L1R5
30	6.1(15.2)	5.0(32.4)	6.7(10.9)	5.2(32.1)
40	5.7(18.2)	4.6(27.6)	6.2(11.1)	4.6(29.7)
45	5.9(18.3)	4.2(11.6)	6.7(8.4)	4.2(11.6)
50	7.0(9.3)	3.8(9.1)	8.1(5.6)	3.8(9.0)
55	7.3(10.0)	3.4(11.4)	7.2(8.7)	3.5(11.5)
60	8.7(12.8)	3.4(19.2)	8.5(12.2)	3.4(19.8)
65	12.4(14.4)	4.3(9.8)	12.1(13.8)	4.3(9.9)
70	18.7(11.6)	6.4(9.2)	18.6(11.4)	6.4(9.4)

TABLE 4 (Continued)

Out of Sample Analysis with $\sigma_\varepsilon = 0.015$

Table 4 presents the mean differences in certainty equivalents (CEQs) between the Vanguard model (with $\sigma_\varepsilon = 0.015$) and the different rule-of-thumb, respectively. The t-statistics, based on the asymptotic distribution, are in parenthesis. In Panel A, for each starting age and initial wealth distribution based on an initial life-cycle simulation, we do the out of sample analysis from 1985 to 2014, using the realized stock returns and the realized dividend yield to pick the relevant policy function. In Panel B, we do the same from 1999 to 2014. U5L2R3: If the dividend yield is greater than 5%, then set the share of wealth in stocks to 100%. If the dividend yield is below 2%, then take the maximum of the Vanguard recommendation minus 30% and 0. Otherwise, follow the Vanguard recommendation. U5L2R5: Same as U5L2R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%. U6L1R3: Same as U5L2R3 but the thresholds of inaction are between 1% and 6% instead of 2% and 5%. U6L1R5: Same as U6L1R3 but subtract 50% from Vanguard recommendation when the dividend yield is below 2%.

Starting	Panel B: 1999–2014			
Age	U5L2R3	U6L1R3	U5L2R5	U6L1R5
30	6.4(34.9)	5.1(202.8)	7.4(23.7)	5.3(181.5)
40	7.5(69.7)	5.2(26.8)	9.0(28.8)	5.4(24.7)
45	7.8(52.7)	5.1(24.4)	9.4(24.6)	5.2(21.8)
50	8.0(39.4)	5.1(28.0)	9.7(20.7)	5.1(23.9)
55	8.3(27.7)	5.0(27.7)	10.3(17.6)	5.0(25.9)
60	9.5(21.5)	5.0(22.5)	12.2(12.6)	4.9(20.3)
65	11.7(14.1)	5.6(121.6)	14.6(9.8)	5.4(164.0)
70	13.1(14.4)	6.9(121.1)	13.4(13.6)	6.8(142.4)

Chapter 2

Life-Cycle Consumption and Portfolio Choice with an Imperfect Predictor

I. Introduction

Optimal life-cycle portfolio choice is a classic problem in financial economics, encountered by every investor. Samuelson (1969) argues that the investment decision is independent of wealth and consumption-saving decisions. However, Samuelson's conclusion is confined to the assumption of independent and identically distributed (i.i.d.) stock returns and the absence of undiversifiable, risky labor income. Cocco, Gomes, and Maenhout (2005) solve for optimal portfolio choice, consumption and saving decisions numerically and show that the labor income stream is a key factor for optimal life-cycle portfolio choice with mortality risk, borrowing and short-sale constraints, and time-separable power utility preferences. Their findings provide rationale for age-varying investment advice such as recommending target-date funds (TDFs) that reduce exposure to stocks as retirement approaches.[1] These authors, however, assume that the stock returns are i.i.d., a classical view meaning that the expected return is constant over time.

Nevertheless, recent empirical studies provide evidence supporting the predictability of stock returns. Many papers find that a number of variables forecast stock returns. The main method is a simple predictive regression: if we can find $|b| > 0$ in $r_{t+1} = \alpha + bq_t + z_{t+1}$, then we know that $E_t\left(r_{t+1}\right) = bq_t$. This implies that the expected stock return can be perfectly predicted by the predictor. The popular predictors (q_t) provided by the

[1] Heaton and Lucas (2000), Viceira (2001), Haliassos and Michaelides (2003) and Gomes and Michaelides (2005) also study the effect of labor income risk on portfolio choice while ignoring the predictability of stock returns.

literature are the dividend/price ratio(D/P), earnings per share (EPS) or consumption-wealth ratio (CAY).[2] Since these predictors themselves follow a persistent auto-regressive process (AR model), the r_t essentially is a mean reversion process.[3]

In response to the evidence on the stock market predictability, various papers have studied its implications for optimal portfolio choice and

---

[2]See Lettau and Ludvigso (2001) and Lan (2015).

[3]For instance, Campbell (1987) and Fama and French (1988) show that dividend/price ratios predict stock returns. Campbell and Shiller (1988) also make this point by proposing the following regressions:

$$\begin{cases} r_{t+1} = r_f + b\mu_t + z_{t+1} \\ \mu_{t+1} = a + \beta\mu_t + \varepsilon_{t+1} \end{cases}, \quad \begin{bmatrix} z_{t+1} \\ \varepsilon_{t+1} \end{bmatrix} \sim Normal\left(0, \Omega\right),$$

where r_{t+1} denotes the real stock market return from time t to $t+1$, μ_t is the predictor such as the dividend/price ratio at time t, α and β are the regression's intercept and slop coefficients of the predictor, r_f is the real risk free interest rate and z_{t+1} and ε_{t+1} are the white noises following a bi-variate normal distribution with mean of zero and covariance structure of Ω. When $\beta = 0$, this regression becomes the i.i.d. stock return model. Fama and French really focus on the importance of the D/P on long-time horizon. These observations show that the predictability of stock return is economically and statistically significant phenomenon that can not be dismissed. Fama and French (1989) is an excellent summary and example that documents and illustrates the time variation of expected stock returns.

consumption.[4] Michaelides and Zhang (2016) build a model in which an investor chooses consumption and optimal asset allocation over the life cycle to maximize an Epstein-Zin-Weil preference function assuming that the dividend yield can perfectly predict the expected stock returns (hereafter, the **perfect predictor model**). This model, however, seems restrictive because it assumes that an observable predictor such as the dividend yield can perfectly predict expected stock returns. This assumption can be criticized for data mining, non-robustness of test statistics and incorrect inference in small samples. Goyal and Welch (2008) re-examine the performance of predictors such as the dividend yield and find that these predictors are both weak in-sample, and out-of-sample, indicating that the predictability of expected stock returns is quite uncertain. It seems more likely that the predictors are

[4]Kim and Omberg (1996), Brennan, Schwartzc, and Lagnad (1997), Brandt (1999), Campbell and Viceira (1999), Balduzzia and Lynch (1999), Campbell, Cocco, Gomes, Maenhout, and Viceira (1999), and Wacher (2002) show that stock market risk premiums change materially with respect to the predictive factor(s) and analyze the implications for optimal portfolio choice.

noisy proxies, in that they are correlated with the time-varying expected stock returns but can not predict them perfectly.[5]

More recently, the idea that the predictive relation between the predictor and expected stock returns is quite uncertain has gained more ground. For example, Xia (2001) assumes that the predictability parameter (b) in the predictive regression is ambiguous. This uncertainty in b is just one specific example that the expected risk premium is hard to precisely observe. Pastor and Stambaugh (2009) generalize Xia (2001) by assuming that the current expected stock return is unobservable and the predictor is imperfect so that the estimation of expected stock returns using the predictive regression omits some important features. In fact, the unexpected stock returns negatively correlate with the innovations in the unobservable expected stock returns, when the stock returns exhibit mean reversion (Pastor and

[5] Ang and Bekaert (2007) also examine the predictive power of the dividend yield for forecasting the excess stock returns. They find that the univariate dividend yield regression provides a rather poor proxy to the true expected stock return.

Stambaugh (2012)). Pastor and Stambaugh (2009) construct an imperfectly predictive system with noisy predictors to estimate the expected stock returns and find that this imperfection has a significant effect on the conditional expected stock returns.

How does the presence of such imperfect predictability affect optimal consumption and portfolio choice for a stockholder over the life cycle? In this paper, I solve a life-cycle portfolio choice model with an imperfect predictor, jointly modeling an imperfect predictive system, liquidity constraints and non-diversifiable background labor income risk to analyze the normative implications for life-cycle consumption and portfolio choice using Epstein-Zin (1989) preferences (hereafter, the **imperfect predictor model**). The key feature of this model is to include the imperfection in the predictive relation of stock returns model to understand how this type of uncertainty affects saving and portfolio choice over the life cycle.

When calibrated to the observed dividend yield and stock returns from

1946 to 2015, under the imperfect predictive system of stock returns, the portfolio allocation is more conservative than that in the perfect predictor model or in the i.i.d. stock returns model. This result substantially alters one of the main insights of models ignoring imperfect predictability. Specifically, such models predict that "stocks are for the young" and such advice has been popularized by Target Date Funds (TDFs) that advise a more aggressive asset allocation in stocks when young and a gradual reduction in this exposure as the investor gets older. With imperfect predictability, consistent with Pastor and Stambaugh (2012), stocks become more volatile in the long run, and therefore young households hold more conservative (balanced) portfolios.

Interestingly, this prediction of the imperfect predictor model is more consistent with empirical observation than either the i.i.d. stock returns or the perfect predictor models. When compared with the data from the U.S. Survey of Consumer Finances (hereafter, **SCF**), the imperfect predictor model matches the data better than either the perfect predictor model or the i.i.d.

stock returns model. Specifically, in the SCF data stockholder portfolios are balanced between bonds and stocks. Recently, Wachter and Yogo (2010) generate balanced portfolios through nonhomothetic utility over basic and luxury goods. In this paper, the balanced portfolio early in life arises due to the additional stock market uncertainty arising from imperfect predictability.

From all the underlying parameters studied, the main parameters that materially affect the optimal consumption and investment choice are the volatility of the unobservable expected stock return, the persistence of the unobservable expected stock returns and the correlation between the innovations to stock returns and shocks to unobserved expected stock returns. Therefore, we should pay more attention to these parameters when making investment decisions. I also experiment with respect to the correlation between permanent earnings shocks and stock market innovations, the correlation between innovations to stock returns and shocks to the dividend yield and the correlation between shocks to the dividend yield and innovations to the unobserved expected stock returns. I find that these correlations do not substantially change wealth accumulation and consumption, but they do significantly alter the portfolio allocation.

These findings influence the design of target date funds (TDFs) because market timing through the utilization of different information affects optimal portfolio choice. The presence of imperfect predictability affects tactical asset allocation and alters the prediction of models where investors expect either i.i.d. stock returns or use a model with a perfect predictor to compute expected stock returns. Therefore, the imperfection of the predictor significantly changes the asset allocation decision, with potentially significant implications for the design of optimal TDFs.

To illustrate the importance of taking imperfect predictability into account when designing TDFs, I compare the welfare across different models by computing the consumption certainty equivalent under different settings. Specifically, I simulate 10,000 individual life histories assuming that the data generating process of stock returns is an imperfect predictive system. In the imperfect predictor model, the investor chooses the investment policy according to the expected return filtered from the observed data. On the contrary, investors using the perfect predictor model or the i.i.d. stock returns model make investment decisions without caring about any observed stock returns. As to the investors using the Vanguard TDFs investment rules (hereafter, **Vanguard TDF model**), they adjusts their portfolio allocation only depending on age. I can then calculate the ratio of value functions from the imperfect predictor model to the ones from the other models and report the consumption certainty equivalent based on this ratio. In this way, I can compare the change in investor welfare between the imperfect predictor model and the other three models: the perfect predictor model, the i.i.d. stock returns model, and the Vanguard TDF model.

The perfect predictor model has the smallest welfare loss, and the i.i.d. stock returns model generates the largest welfare loss. The Vanguard TDF model obtains the second largest welfare loss. All of these welfare losses vary with the correlation between unexpected returns and shocks to the predictors, and increase as this correlation approaches 1. 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.

Where do these welfare rankings come from? I show that these substantial welfare losses relative to the baseline can be explained by the differences in the first two moments of household consumption. The imperfect predictor model has the highest mean consumption and volatility of consumption, and the i.i.d. stock returns model generates the lowest mean consumption and volatility of consumption over the working life. In the middle is the perfect predictor model.

The paper is organized as follows. Section 2 explains the theoretical model in the paper and a rough description of the numerical solution. Section

3 illustrates the estimation method and discusses the calibration. Section 4 builds a baseline model with the risky labor income and Epstein–Zin preferences to study the effect of the imperfect predictive system on the portfolio choice over the life cycle, Section 5 contains the welfare analysis across different models including the TDFs and Section 6 concludes.

II. The Model

A. Model Specification

Preference Model

I denote adult age by t ($t \in [20, 100]$). The investor chooses the portfolio and consumption policies to maximize the following Epstein-Zin preferences:

$$\begin{cases} V_t = \underset{(c_t,\alpha_t)}{max} \left\{ (1-\beta)\, C_t^{1-1/\psi} + \beta \left[\Re_t\left(V_{t+1}\right)\right]^{1-1/\psi} \right\}^{1/(1-1/\psi)} \\ \Re_t\left(V_{t+1}\right) = \left[E_t \left(p_{t+1} V_{t+1}^{1-\gamma} + b\left(1-p_{t+1}\right) X_{t+1}^{1-\gamma} \right) \right]^{1/(1-\gamma)} \end{cases} \tag{2.1}$$

where V_t is the continuation value at age t, $\Re_t$ is the uncertainty aggregator, X_{t+1} is the terminal wealth if the investor is dead at age t+1, β is the discount factor, ψ is the elasticity of inter-temporal substitution (hereafter, **EIS**), γ is the risk aversion parameter, b is the strength of the bequest motive and p_{t+1} is the conditional probability of surviving next period conditional on having survived until age t.

Labor Income Process

Following the same method as Cocco, et al. (2005) andCarroll (1997), I build the labor income process before retirement as follows:

$$Y_{it} = Y_{it}^{p} U_{it} \tag{2.2}$$

$$Y_{it}^{p} = exp\left[g\left(t, Z_{it}\right)\right] Y_{it-1}^{p} N_{it} \tag{2.3}$$

where $g\left(t, Z_{it}\right)$ is a deterministic function of age and household i's characteristics Z_{it}, Y_{it}^{p} is a permanent component with innovation N_{it} of household i's age t labor income, and U_{it} is a transitory component of household i's age t labor income.

In equations (2.2) - (2.3), I assume that $ln\left(U_{it}\right)$ and $ln\left(N_{it}\right)$ are independent and identically distributed with mean $\{-0.5\sigma_u^2, -0.5\sigma_n^2\}$, and variances σ_u^2 and σ_n^2, respectively. As to Y_{it}^{p}, $ln\left(Y_{it}^{p}\right)$ 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 (hereafter, **PSID**) data. We choose not to model explicitly the returns from housing following the empirical evidence (e.g., 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 (2015)). 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).

For convenience, I will take logarithms on both sides of (2.2) and (2.3) while solving the investor's problem. Hence,

$log\left(Y_{it}^{p}\right) = g(t, Z_{it}) + log\left(Y_{it-1}^{p}\right) + log\left(N_{it}\right)$ and

$log\,(Y_{it}) = log\,(Y_{it}^{p}) + log\,(U_{it}).$

Stock Return Predictability Model

I assume that there are two assets in which the investor can invest, a risk-free asset, such as T-bills, and a risky asset, such as stocks. The risk free asset has a constant gross real return of r_f, and the risky asset has a gross real return r_t. As to modeling the gross real return of risky asset, I follow the idea of Pastor and Stambaugh (2009) that the expected stock returns are unobservable and that investor must filter these expected stock returns from the other observable information. Denote $(r_t,\, q_t,\, \mu_t)$ as the stock return, the predictor and the unobservable expected stock return, respectively. Then, an imperfect predictive system can be defined as follows:

$$\mu_{t+1} = \alpha_\mu + \phi_\mu \mu_t + \varepsilon_{t+1} \tag{2.4}$$

$$r_{t+1} = r_f + \mu_t + z_{t+1} \tag{2.5}$$

$$q_{t+1} = \alpha_q + \phi_q q_t + v_{t+1} \tag{2.6}$$

where $\begin{bmatrix} \varepsilon_{t+1}, & z_{t+1} & v_{t+1} \end{bmatrix} \sim Normal\,(0, \Omega)$ and $\Omega = \begin{bmatrix} \sigma_\varepsilon^2 & \sigma_{z\varepsilon} & \sigma_{v\varepsilon} \\ & \sigma_z^2 & \sigma_{zv} \\ & & \sigma_v^2 \end{bmatrix}$.

This imperfect predictive system is a generalization of the classical predictive regression. The unobservable expected stock return (μ_t) follows a simple AR(1) process described by equation (2.4). Equation (2.5) defines the next period's stock return (r_{t+1}) as a sum of the risk free rate (r_f), the unobservable expected stock return (μ_t) and an innovation term (unexpected stock return, z_t). Equation (2.6) assumes that the predictor (q_t) evolves in a manner of a persistent AR(1) process, which is a standard assumption in the literature about the predictability of stock returns. This model is consistent with a variety of economic models in which the expected return not only varies over time but also exhibits mean reversion.

Based on this imperfect predictive system, the investor must filter out μ_t from the other observable variables (r_t, q_t). Applying the simplest filtering algorithm (see theorem A in the appendix, the conditional distribution of a multivariate normal distribution), the first two conditional expected moments of μ_t can be rewritten as

$$E\left(\mu_t|d_t\right) = E_r + \Sigma_{\mu d}\Sigma_d^{-1}\left(\begin{bmatrix} r_t - r_f \\ q_t \end{bmatrix} - \begin{bmatrix} E_r \\ E_q \end{bmatrix}\right) \tag{2.7}$$

$$Var\left(\mu_t|d_t\right) = \sigma_\mu^2 - \Sigma_{\mu d}\Sigma_d^{-1}\Sigma_{\mu d}' \tag{2.8}$$

where $d_t = [r_t,\, q_t]$, $\Sigma_{\mu d} = [\sigma_{\mu r}, \sigma_{\mu q}]$ and $\Sigma_d = \begin{bmatrix} \sigma_r^2 & \sigma_{rq} \\ \sigma_{rq} & \sigma_q^2 \end{bmatrix}$.

(7) and (8) can be further simplified as:

$$E(\mu_t|[r_t, q_t]) = \hat{\mu}_{t|t} = E_r + \kappa_r [r_t - r_f - E_r] + \kappa_q [q_t - E_q] \tag{2.9}$$

$$Var(\mu_t|[r_t, q_t]) = \sigma_\mu^2 - \kappa_r \sigma_{\mu r} - \kappa_q \sigma_{\mu q} \tag{2.10}$$

where $\kappa_r = \frac{\sigma_{\mu r}\sigma_q^2 - \sigma_{rq}\sigma_{\mu q}}{\sigma_r^2\sigma_q^2 - \sigma_{rq}^2} < 0$, $\kappa_q = \frac{\sigma_r^2\sigma_{\mu q} - \sigma_{rq}\sigma_{\mu r}}{\sigma_r^2\sigma_q^2 - \sigma_{rq}^2} > 0$, $E_r = \frac{\alpha_\mu}{1-\phi_\mu}$, $E_q = \frac{\alpha_q}{1-\phi_q}$,

$\sigma_r^2 = \sigma_\mu^2 + \sigma_z^2$, $\sigma_\mu^2 = \frac{\sigma_\varepsilon^2}{\left(1-\phi_\mu^2\right)}$, $\sigma_q^2 = \frac{\sigma_v^2}{\left(1-\phi_q^2\right)}$, $\sigma_{\mu r} = \rho_{z\varepsilon}\sigma_z\sigma_\varepsilon + \frac{\sigma_\varepsilon^2}{\left(1-\phi_\mu^2\right)}$,

$\sigma_{\mu q} = \frac{\rho_{v\varepsilon}\sigma_v\sigma_\varepsilon}{\left(1-\phi_\mu\phi_q\right)}$ and $\sigma_{rq} = \rho_{zv}\sigma_z\sigma_v + \frac{\phi_q\sigma_{v\varepsilon}}{\left(1-\phi_\mu\phi_q\right)}$.

(2.9) and (2.10) say that the conditional moments of μ_t consist of three information sources. The first source is the unconditional mean of risk premium (E_r). The second source is the current stock return (r_t), and the last one is the current dividend yield (q_t). Similarly, the conditional variance of μ_t can be decomposed into three parts: the variance of unobservable expected stock returns $\left(\sigma_\mu^2\right)$, the covariance between the unobservable expected stock returns and the realized stock returns ($\sigma_{\mu r}$) and the covariance between the unobservable expected stock returns and the dividend yield ($\sigma_{\mu q}$).

Several important conclusions can be drawn from (2.9) and (2.10). First, κ_r is negative, which implies that an unexpected increase in the stock return leads to the decrease in the next period's expected stock return. κ_r , therefore, measures the mean reversion effect. In contrast, the positive κ_q measures the predictability effect because a positive shock to the dividend yield predicts an increase in the next period's expected stock return and vice versa.

Second, when $\rho_{\mu q} = 1$, $E_r = E_q$, $\sigma_\mu = \sigma_q$ and $\rho_{\mu r} = \rho_{rq}$, $\kappa_r = 0$ and $\kappa_q = 1$. (2.9) and (2.10), therefore, become

$$\hat{\mu}_{t|t} = E\left(\mu_t | [r_t, q_t]\right) = q_t \tag{2.11}$$

$$Var\left(\mu_t | [r_t, q_t]\right) = 0 \tag{2.12}$$

(2.11) and (2.12) implies that $E_t(r_{t+1}) = q_t$, namely, the predictor perfectly predicts the expected stock return. The imperfect predictive system ((2.4) - (2.6)) degenerates into the classical predictive regression used in Campbell and Shiller (1988), Campbell and Viceira (1999), Michaelides and Zhang (2016) etc. Similarly, the i.i.d. stock returns model is also a special case of this imperfect predictive system. In contrast, if $|\rho_{v\varepsilon}| < 1$ and $\rho_{v\varepsilon} \neq 0$, the predictor (q_t), is not a perfect proxy of μ_t, and the information from r_t and q_t enters the conditional expected μ_t according to (2.9) - (2.10). Hence, the expected stock return of the next period is not completely determined by the observed predictor so that uniquely relying on the this predictor can deliver an inaccurate estimation.

Third, the conditional moments of the unobservable expected stock return depend on both the observed data (r_t, q_t) and the correlations among the unobservable expected stock return, the observed predictor and the current stock return $(\rho_{\mu r}, \rho_{rq}, \rho_{\mu q})$. This also explains why the correlation between the innovations to observable predictor and the shocks to current stock return does not play a key role in the perfect predictor model solved by Michaelides and Zhang (2016)[6]. The perfect predictor model rule out the effect of these correlations from calculating the conditional expected stock return of the next period $(E_t [r_{t+1}] = q_t)$ and conditional variance $\left(Var_t [r_{t+1}] = \sigma_z^2\right)$, which means that these correlations only have a small effect on the optimal investment and consumption decision.

B. The Investor's Optimization Problem

At the beginning of period t, investor i has a wealth $W_{i,t}$. Then, during this period, labor income $Y_{i,t}$ is realized. Following Deaton (1991), cash on hand $X_{i,t}$ can be defined as $X_{i,t} = W_{i,t} + Y_{i,t}$. Then, the investor must determine how much to consume, $C_{i,t}$ and how to invest the remaining savings

[6]Michaelides and Zhang (2016) use the perfect predictor model/classical predictive regression to solve the life-cycle portfolio choice problem and find that only the correlation between the innovations of stock returns and the permanent earning shocks of labor income (ρ_{zn}) materially affects the optimal portfolio choice.

in stocks $S_{i,t}$ and the risk free asset $B_{i,t}$. In the next period, before earning period $t+1$'s labor income, the wealth at $t+1$ is given by

$$W_{i,t+1} = S_{i,t}\left(1 + r_{t+1}\right) + B_{i,t}\left(1 + r_f\right) = \alpha_{it}\left(1 + r_{t+1}\right) + \left(1 - \alpha_{it}\right)\left(1 + r_f\right) \ ,$$

where $S_{i,t}$ is the investment in the stock market in the previous period, $B_{i,t}$ is the investment in risk-free asset in the previous period and $\alpha_{i,t}$ is the share of wealth in stocks in the previous period and defined as $\alpha_{i,t} = \frac{S_{i,t}}{B_{i,t}+S_{i,t}}$. The budget constraint of investor i at time t is $S_{i,t} + B_{i,t} = W_{i,t} + Y_{i,t} - C_{i,t}$.

The investor maximizes the household's utility subject to the budget constraint and the constraints (2.2) through (2.6) with the non-negativity restrictions on $C_{i,t}$, B_{it} and $S_{i,t}$. These non-negativity constraints on B_{it} and $S_{i,t}$ guarantee the investor not to borrow against his/her future labor income or retirement wealth.

In this optimization problem, μ_t is unobservable and the investor has to estimate it through (2.9) - (2.10) conditional on the observed information (r_t, q_t) available at time t. The state variables of the investor's problem are t,

$X_{i,t}$, $\hat{\mu}_{t|t}$ and Y_{it}^{p}, the control variables are $C_{i,t}$ and $\alpha_{i,t}$, and the policy functions are defined as $C_{i,t}\left(X_{i,t}, Y_{i,t}^{p}, \hat{\mu}_{t|t}\right)$ and $\alpha_{i,t}\left(X_{i,t}, Y_{i,t}^{p}, \hat{\mu}_{t|t}\right)$.

Since, the problem uses the Epstein-Zin utility, the value function is homogeneous with respect to the current permanent part of labor income. This property allows us to normalize the investor's cash on hand $(X_{i,t})$ by dividing $Y_{i,t}^{p}$, which means the number of state variables is reduced by one. The policy functions, therefore, become $c_{i,t}\left(x_{i,t}, \hat{\mu}_{t|t}\right)$ and $\alpha_{i,t}\left(x_{i,t}, \hat{\mu}_{t|t}\right)$, where $x_{i,t} = \frac{X_{i,t}}{Y_{i,t}^{p}}$.

C. Numerical Solution

The optimization problem faced by the investor can be rewritten as the following optimization model:

$$V_t\left(x_{i,t},\hat{\mu}_{t|t}\right)=\underset{(c_{i,t},\alpha_{i,t})}{Max}\left\{\begin{array}{c}(1-\beta)\,c_{i,t}^{1-\frac{1}{\psi}}+\beta\left[\left\{E_t\left(p_{t+1}V_{t+1}^{1-\gamma}\left(x_{i,t+1},\hat{\mu}_{t+1|t+1}\right)\right.\right.\right.\\ \left.\left.\left.+b\left(1-p_{t+1}\right)x_{i,t+1}^{1-\gamma}\right)\right\}^{\frac{1}{\gamma}}\right]^{1-\frac{1}{\psi}}\end{array}\right\}^{\frac{1}{1-\frac{1}{\psi}}}$$

$$s.t.\begin{cases}\mu_{t+1}=\alpha_\mu+\phi_\mu\mu_t+\varepsilon_{t+1}\\ r_{t+1}=r_f+\mu_t+z_{t+1}\\ q_{t+1}=\alpha_q+\phi_q q_t+v_{t+1}\\ ln\left(N_{t+1}\right)=\mu_n+n_{t+1}\\ ln\left(U_{t+1}\right)=\mu_u+u_{t+1}\\ x_{i,t+1}=\frac{Y_{i,t}^p}{Y_{i,t+1}^p}\left(r_{t+1}\alpha_{i,t}+r_f\left[1-\alpha_{i,t}\right]\right)\left(x_{i,t}-c_{i,t}\right)+U_{i,t+1}\end{cases} \tag{2.13}$$

where $\hat{\mu}_{t|t}$ is a linear function of (r_t, q_t) and updated through formula (2.9), c_{it} is the normalized consumption of household i at time t, x_{it} is the normalized cash on hand of household i at time t and $\alpha_{i,t}$ is the risky asset allocation of household i at time t.

This problem has no analytical solution. I, therefore, solve this problem numerically by using backward induction. In the last period (hereafter, T), the optimal policy functions are easy to solve because the investor does not invest any more and consumes all wealth except for the saving bequeathed to heirs. Then, I can now replace the value function in the Bellman equation (2.13) with the optimal policy function solved at time T and calculate the optimal policies for T-1. Repeating this procedure up to age 20, I can obtain the policy functions at each age.

In the backward induction algorithm, grid search is used to find the optimal policy functions of the problem (2.13) based on a fine discrete approximation of the following VAR model:

$$
\begin{cases}
\mu_{t+1} = \alpha_\mu + \phi_\mu \mu_t + \varepsilon_{t+1} \\
r_{t+1} = r_f + \mu_t + z_{t+1} \\
q_{t+1} = \alpha_q + \phi_q q_t + v_{t+1} \\
ln\left[N\right]_{t+1} = \mu_n + n_{t+1}
\end{cases}
\tag{2.14}
$$

I use Tauchen and Hussey (1991) method to discretize the state space of the VAR model (2.14) and calculate the transition probabilities among these grid points assuming that they follow a Markov Chain. Then, using the grid points from the discretization of (2.14)[7], I can construct the next period's return by:

$$
\begin{cases}
r_{t+1|t} = r_f + \hat{\mu}_{t|t} + z_{t+1} + w_{t+1} \\
\hat{\mu}_{t|t} = E_r + \kappa_r \left[r_t - r_f - E_r\right] + \kappa_q \left[q_t - E_q\right]
\end{cases}
\tag{2.15}
$$

where w_{t+1} is an independent innovation term introduced by the

[7]The temporary part of labor income ($ln\left(U_t\right)$) is not correlated with the other variables. Its grid points are, therefore, generated independently.

filtering algorithm and follows $N\left(0, Var\left\{\mu_t|[r_t, q_t]\right\}\right)$.

Finally, I iteratively apply the backward induction algorithm to solve the consumption and investment policy functions of the optimization problem (2.13) based on $r_{t+1|t}$ from age T to age 20. The details of numerical implements are the same as the Online Appendix of Michaelides and Zhang (2016).

I implement this numerical algorithm using Fortran 2003 on a Windows workstation[8]. For accelerating the time of computation, I parallelize this algorithm according to the state variables using OpenMP[9], which makes the problem can be solved in twenty four hours.

[8]Intel Xeon E5-2699 v3 2.3GHz RAM 256GB

[9]OpenMP is a set of compiler directives, library routines, and environment variables to enable programmers to develop parallel applications for shared memory multiprocessor computer.

III. Empirical Analysis

A. Data

The stock market data used in this paper comes from the Center for Research in Securities Prices (CRSP). I screen out the annual one year bond return, annual CPI growth rate, monthly value-weighted cumulative return of S&P 500 and monthly value-weighted ex-return of S&P 500 from Dec. 31st, 1946 to Dec. 31st, 2015. Next, I construct annual cumulative and ex-dividend S&P 500 price index based on the monthly data with the initial cumulative price of 1.00. Using the difference between annual cumulative and ex-dividend price index, I can easily obtain the annual cumulative return and annual ex-dividend return. The annual dividend is calculated by multiplying the lagged total annual price index by the difference between the annual cumulative return and ex-dividend return. Finally, I compute the real return as the difference between the annual cumulative return and annual CPI growth rate. Table 1 shows the summary of stock market data.

The empirical portfolio holding data are based on the SCF 2007. The empirical asset holding is defined as either $\alpha = equity/(equity + bond)$ or $\alpha = equity/(equity + bond + liquidity)$, where *liquidity* is the financial wealth with high liquidity such as the cash.

B. Parameter Estimation

The first step of solving the investor's optimization problem is to estimate the parameters of the equation (2.4) - (2.6). For estimating this VAR model through the observed data, I transform it into the following VARMA(1,1) model:

$$\begin{cases} r_{t+1} - r_f = \left(1 - \phi_\mu\right) E_r + \phi_\mu (r_t - r_f) + n v_t - \left(\phi_\mu - m\right) \omega_t + \omega_{t+1} \\ q_{t+1} = (1 - \phi_q) E_q + \phi_q q_t + v_{t+1} \end{cases} \tag{2.16}$$

where m and n are constant parameters derived based on the equations (2.4) - (2.6) and ω_t is forecast error $\left(\omega_t = r_t - \hat{r}_{t|t-1}\right)$ and serially uncorrelated.

The appendix describes how to derive this VARMA(1,1) model and estimate it using MLE.[10] Table 2 summarizes the results of the parameter estimation.

Some parameters in the covariance matrix of equations (2.4) - (2.6) remain unidentified because the covariance matrix consisting of three variables can not be exactly estimated through only two observed variables (see appendix C). I, therefore, describe the solution space of the covariance matrix (Ω) with respect to a specific variable. As $\sigma_{z\varepsilon}$ play a critical role in determining the conditional expected return, I solve the solution space of the covariance matrix (Ω) with respect to $\sigma_{z\varepsilon}$. The details about how to derive the solution space of the covariance matrix are explained in the appendix. In short, the solution space of the covariance matrix with respect to $\sigma_{z\varepsilon}$ is

[10] I thank Lubos Pastor for kindly providing matlab codes to perform this estimation.

simplified into the following linear system:

$$\begin{cases} \sigma_z^2 = \sigma_r^2 - \left(Cov\left(r_t, r_{t-1}\right) - \sigma_{z\varepsilon}\right)/\phi_\mu \\ \\ \sigma_\varepsilon^2 = \left(Cov\left(r_t, r_{t-1}\right) - \sigma_{z\varepsilon}\right)\left(1 - \phi_\mu^2\right)/\phi_\mu \end{cases} \tag{2.17}$$

$$s.t.\ |\rho_{zv}| < 1\, and\, |\rho_{v\varepsilon}| < 1$$

As $\sigma_{z\varepsilon} = \rho_{z\varepsilon}\sigma_\varepsilon\sigma_z$, I can only discuss the correlation between the shocks to unobservable expected stock returns and the innovation of stock returns ($\rho_{z\varepsilon}$) instead of the covariance, $\sigma_{z\varepsilon}$. Various studies provide empirical evidence that $\rho_{z\varepsilon} < 0$. Pastor and Stambaugh (2009) find that this correlation is negative if the stock returns exhibit mean reversion. Figure 1, Graph A, plots the solution space of $(\rho_{v\varepsilon}, \rho_{zv}, \rho_{z\varepsilon})$ while changes $\rho_{z\varepsilon}$ from -1 to 0. Graph B projects the solution space onto the plane consisting of ρ_{zv} and $\rho_{z\varepsilon}$, and Graph C describes the relationship between $\rho_{z\varepsilon}$ and $\rho_{v\varepsilon}$.

Several conclusions can be drawn from the Figure 1. First, based on the data, the ranges of $\rho_{z\varepsilon}$, ρ_{zv} and $\rho_{v\varepsilon}$ are approximately [-0.66,-0.99], [-0.65,

-0.99] and [0.37, 0.94], respectively. Second, Graph B shows that the correlation between the innovations of stock returns and the shocks to the dividend yield (ρ_{zv}) has approximately a negative relation with the correlation between the innovations of stock returns and the shocks to the unobservable expected stock returns ($\rho_{z\epsilon}$). When $\rho_{z\varepsilon}$ tends to be a perfect negative correlation, $|\rho_{zv}|$ decreases from 0.99 to 0.65. In contrast, the correlation between the shocks to the unobservable expected stock returns and the innovations of the dividend yield ($\rho_{v\varepsilon}$) positively relates with $\rho_{z\varepsilon}$. When $\rho_{z\varepsilon}$ tends to be a perfect negative correlation, $\rho_{v\varepsilon}$ is close to a perfect positive correlation.

For better understanding the effect of the imperfect predictive system of stock returns on the life-cycle consumption and portfolio choice, I set up a baseline model, where $\rho_{z\varepsilon} = -0.7$, $\sigma_\varepsilon = 0.09852$, $\sigma_z = 0.1646$, $\rho_{zv} = -0.723$, $\rho_{v\varepsilon} = 0.56$ and $\rho_{zn} = 0.15$.

IV. Optimal Consumption and Portfolio Choice

A. *The Baseline Model*

Parameter Choice

Even though empirical predictability studies are typically done on a monthly or quarterly frequency, I solve the model at an annual frequency to maintain comparability with the existing life-cycle portfolio literature.Carroll (1997) estimates the variances of the idiosyncratic shocks using data from the PSID, and the baseline simulations use values close to those: 0.1 for σ_u and 0.1 for σ_n. The deterministic component of labor income is identical to the values used by most life cycle papers, for example, Cocco, et al. (2005), and this setting also facilitates comparisons between this model and its counterparts such as perfect predictor model and i.i.d. stock returns model. The relatively large estimate for the replacement ratio during retirement (68% 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 modeling tax-deferred saving through retirement accounts.

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. The elasticity of inter-temporal substitution (ψ) is set to be 0.5. These choices are close to the empirical estimates for the EIS 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. As to the discount rate, much macroeconomic research estimates this rate to be 1% per quarter or approximate 4% per year. In order to emphasize that the results in this paper does not stem from extreme assumptions about discount factor, β in the baseline model is 0.96, which means the discount rate is assumed to be 4% per year.

The parameters used in the imperfect predictive system of the stock market are listed in Table 1 and 2. In addition, I set a trading cost of 2.9% to reflect transaction cost, tax and other implicit trading costs, which implies a risk premium of 4% the same as in the most of the life-cycle portfolio literature.

There is no estimate of the correlation between the innovations of the unobservable expected stock returns and the permanent, idiosyncratic earnings shocks to the labor income ($\rho_{n\varepsilon}$) in the literature. I therefore set this correlation equal to zero. Angerer and Lam (2009) note that the correlation between the innovations of stock returns and transitory part of labor income (ρ_{zu}) does not empirically affect portfolios and this is consistent with the simulation results in life cycle models (Cocco, et al. (2005)). I set this correlation at zero. Similarly, I also set ρ_{nv} to zero. The correlation between the permanent earning shocks to the labor income and the innovations of stock returns (ρ_{zn}) is set equal to 0.15 in the baseline model, which follows the same setting as Michaelides and Zhang (2016). Table 3 summarizes the parameter values used in the baseline model.

Consumption and Portfolio Choice in the Baseline Model

Figure 2 plots the life-cycle profiles of wealth accumulation, consumption, labor income and share of wealth in stocks by simulating 10,000 individual life histories and reports the average.

Graph A shows the mean wealth accumulation and consumption over the life cycle in the presence of a bequest motive and labor income. The wealth accumulation increases as the investor approaches retirement and reaches the peak at the retirement age. After the retirement, the wealth accumulation starts to decrease as agent ages.

Graph A also shows that the consumption tracks labor income very closely before retirement and the gap between consumption and labor income gets larger as the wealth deaccumulates, reflecting that the liquidity constraint becomes less binding. When the agent approaches death, the consumption path decreases. Graph B graphs the mean share of wealth in stocks over the life cycle. Early in life, a higher proportion of wealth is invested in the risky asset except for at the very beginning of life. As the agent approaches retirement, the share of wealth in stocks slopes down. After retirement, the mean stock allocation bounces up a little then keeps highly stable until the agent reaches the end of life. During the whole life cycle, the mean stock allocation is clearly less than 1 and fluctuates between 40% and 65%. These findings (Graph A and B) are consistent with Cocco, et al. (2005), Gomes and Michaelides (2005) and Michaelides and Zhang (2016).

Figure 3 compares the life-cycle profiles between the baseline model (imperfect predictor model), perfect predictor model, i.i.d. stock returns model and the Vanguard TDF model. The Vanguard TDF model's basic recommendation is to invest 90% of a household's financial wealth in stocks until age 40, and start decreasing that share as retirement approaches reach

50% at age 65. After retirement, the Vanguard TDF model recommends the investor to continuously reduce the stock market exposure to approximate 30% and keep this proportion until death. To simulate wealth profiles for this case, I take the portfolio rule as exogenous but the household still makes optimal consumption-saving decisions, taking this portfolio decision into account.

The mean wealth accumulation and consumption shows a notable difference between the baseline model and the other three models. Graph A shows that the wealth accumulation and the consumption in the baseline model are the highest in all of these models. This arises here because the imperfection of the predictive system leads the investor to increase the precautionary saving in the baseline model. Graph B describes the difference in simulated average consumption over the life cycle. The mean consumption of the baseline and perfect predictor model are the highest and the second highest respectively because the investor takes advantage of predictability. Graph C depicts the mean share of wealth in stocks over the life cycle. The i.i.d. stock returns model maintains the highest proportion of wealth in the stock market, except between ages 45 and 65, and the baseline model has the lowest mean share allocation. On the other hand, the perfect predictor model falls in between. The glide path of the Vanguard TDF model is exogenous as it is fixed at each age without considering any information.

The remarkable difference of mean portfolio allocation across these models can be explained by the investment policy functions. Figure 4 shows the share of wealth in stocks with respect to low, medium and high estimations of the expected stock return for age 25, 55 and 75 (Graph A, B and C show the share of wealth in stocks for age 25, Graph D, E and F for age 55, and Graph G, H and I for age 75). The investment policy functions of the i.i.d. stock returns model vary with age besides the cash on hand, and does not depend on the other factors. In the baseline and perfect predictor models, the portfolio allocation can drastically shift up or down depending on the estimation of the expected stock return besides age and cash on hand.

When focusing on the baseline model and the perfect predictor model, I find that the investment policy functions in the baseline model are always less than that of the perfect predictor model. This result arises because the imperfection of the predictive system increases the conditional variance of the next period's return given the same estimation of the expected return.

An empirical puzzle arises that the predictions of portfolio allocation from the i.i.d. and perfect predictor model have a large gap during the working age over the life cycle. Figure 5 compares the mean share of wealth in stocks from the perfect predictor model and the imperfect predictor model with the data of SCF 2007. Graph A compares the mean share of wealth in stocks with the empirical portfolio allocation without considering liquidity. Graph B, however, includes the asset with high liquidity in the calculation of empirical portfolio allocation. The smoothed empirical portfolio allocation is calculated by the linear regression method. From Figure 5, we can find that the prediction from the imperfect predictor model matches the SCF data better than the perfect predictor model, which shows that the imperfection of the predictive system possibly make an important contribution to explain the observed pattern of household portfolio choice.

The Analysis of Model Parameter Uncertainty

Even though the baseline model has considered the imperfection of predictability in the stock returns, the estimation of parameter such as the unconditional expected risk premium (E_r), persistence of the unobservable expected stock returns (ϕ_μ) and standard error of the unobservable expected stock returns (σ_ε) possibly still have an estimation error. These parameters materially affect the mean wealth accumulation, consumption and asset allocation when the preference parameters such as risk aversion(γ) and EIS(ψ) remain unchanged. Therefore, this section measures the sensitivity of the baseline model to these parameters.

Figure 6 shows the effect of a higher unconditional expected risk premium (E_r) on the mean wealth accumulation (Graph A), consumption

(Graph B) and share of wealth in stocks (Graph C) over the life cycle of the baseline model, perfect predictor model and i.i.d. stock returns model, respectively. For obtaining a higher unconditional expected risk premium ($E_r = 7\%$), I set up a 0% of the Trading Cost. Under the scenario in which the risk premium is perceived to be higher, the mean wealth accumulation, consumption and portfolio allocation all shift up. A higher unconditional expected risk premium makes investor lean to holding stocks, which leads to a higher wealth accumulation and, then, a higher consumption.

Figure 7 plots how a lower standard error of the unobserved expected stock return (σ_ε) affects the life-cycle profiles of the baseline model. When the volatility of the unobserved expected stock return (σ_ε) decreases to 0.005 from 0.0985, the mean wealth accumulation and consumption decrease, and portfolio allocation shifts up except for the 45 - 65 age group. A lower σ_ε leads to the unobservable expected stock returns fluctuating around the unconditional expectation of risk premium within a narrow band, which makes the imperfect predictive system act as a i.i.d. stock returns. The life-cycle profiles are, therefore, closer to that of the i.i.d. stock returns model.

The parameter ϕ_μ measures the persistence of the unobservable expected stock returns. This parameter is of our interest because the predictor used in the predictive regression is often a highly persistent process in the classical literature such as Campbell and Shiller (2009), Fama and French (1988), Xia (2001) and Cochrane (2005). Figure 8 depicts the life-cycle profile given a higher persistence of the unobservable expected stock returns (Graph A shows the mean wealth accumulation and consumption, and Graph B describes the mean share of wealth in stocks). From Graph A and B, a higher persistence makes the agent take advantage of predictability so that the mean share of wealth in stocks shifts up and seems close to that of the perfect predictor model. This is reasonable because the unobservable expected stock return is close to the high persistent predictor process when its persistence is high. On the other hand, the high persistence of the unobservable expected stock returns makes the investor more willing to consume in the earlier stage of life, attaining a lower mean wealth accumulation at retirement. The conclusions drawn from Figure 7 and 8 remind us that it is dangerous to depend entirely on an imperfect predictor such as dividend yield. The characteristics of high persistence and low volatility in the dividend yield process can lead to more aggressive investment polices and inappropriate consumption decisions.

Admittedly, the unconditional expected risk premium and standard error and persistence of the unobservable expected stock return are not the whole story. The variations due to correlations such as ρ_{zn} and $\rho_{z\varepsilon}$ are also crucial in the household financial decisions. I analyze these effects in the next subsection.

B. Hedging Demands

How does the correlations among the different innovations change the results of baseline model? In the i.i.d. stock returns model and perfect predictor model, the most important correlation generating quantitatively substantial hedging demands is the correlation between the permanent earnings shocks and the innovations to stock returns (ρ_{zn}), and the other correlations such as $\rho_{z\varepsilon}$ do not materially affect the results. Does this conclusion change when I introduce the imperfection to the predictive regression?

Correlation between the Shocks to the Unobservable Expected Stock Returns and the Innovations of Stock Returns

To investigate the importance of the correlation between the shocks to the unobservable expected stock returns and the innovations of stock returns ($\rho_{z\varepsilon}$), I vary $\rho_{z\varepsilon}$ from -0.9 to -0.5 and use the baseline model ($\rho_{z\varepsilon}$ = - 0.7) and perfect predictor model as benchmarks for comparison. Figure 9 plots the mean wealth accumulation, consumption (Graph A) and the mean share of wealth in stocks (Graph B) over the life cycle due to the variation of $\rho_{z\varepsilon}$. When $\rho_{z\varepsilon}$ tends to be 0 from a perfect negative correlation, the investor views the dividend yield as a better predictor of the unobserved expected stock return. From Table 4, we know that a smaller $|\rho_{z\varepsilon}|$ decreases the mean reversion effect and increases the predictability effect. This implies that results are close to that from the perfect predictor model. The investor, therefore, decreases the wealth accumulation and consumption (Graph A) and increases the stock holding. On the contrary, when this correlation is close to perfect negative, the mean asset allocation in risky stocks shifts down and the mean wealth accumulation and consumption move up.

Correlation between the Permanent Earnings Shocks and the Innovations of Stock Returns

I also measure the sensitivity of the correlation between the permanent earnings shocks and the innovations of stock returns (ρ_{zn}). In the baseline model, this correlation is 0.15, a value that reflects the substantial idiosyncratic risk that exists in labor income data. I vary this correlation from -0.15 to 0.3.

Figure 10 plots its effect of ρ_{zn} on the results from the baseline model.

From Graph A, when ρ_{zn} changes, I find that the mean wealth accumulation and consumption rarely change. However, in Graph B, I find that the investor is more willing to invest risky stocks when this correlation decreases. The labor income acts more as a risk less asset when ρ_{zn} is small, which leads investors to taking more risk exposure in the stock market. On the contrary, when this correlation increase, it crowds out the risky investment because the labor income acts more like a risky stock. Hence, the portfolio allocation negatively correlates with ρ_{zn}, which is consistent with the results found in Cocco, et al. (2005) and Michaelides and Zhang (2016).

Correlation between the Innovations of Stock Returns and the Shocks to Dividend Yield

Changing the correlation between the innovations of stock returns and the shocks to dividend yield (ρ_{zv}) does not materially affect the mean wealth accumulation and consumption (Figure 11, Graph A), but does significantly change the portfolio allocation (Figure 11, Graph B). According to Table 5, when $|\rho_{zv}|$ close to 0, the predictability effect from dividend yield becomes stronger, while the mean reversion effect from the current return does not change. This makes the investment behavior from the imperfect predictor model more like that from the perfect predictor model. This conclusion is different from the result of Michaelides and Zhang (2016) because ρ_{zv} has little effect on determining the conditional moments of the next period's return in the perfect predictor model.

Correlation between the Innovations of the Dividend Yield and the Shocks to the Unobservable Expected Stock Returns

What happens when the correlation between the innovations of the dividend yield and the shocks to the unobservable expected stock returns ($\rho_{v\varepsilon}$) varies? Figure 12, Graph A plots the mean wealth accumulation and consumption over the life cycle, and Figure 12, Graph B plots the mean share of wealth in stocks. When $\rho_{v\varepsilon}$ increases from 0.2 to 0.8, the mean wealth accumulation and consumption rarely change. The mean share of wealth in stocks, however, shows a positive correlation with $\rho_{v\varepsilon}$ before retirement. Table

6 shows that the predictability effect from the dividend yield becomes strong, while the mean reversion effect from the stock returns only slightly decreases. Therefore, as $\rho_{v\varepsilon}$ increases, the portfolio choices from the imperfect predictor model tend toward the predictions from the perfect predictor model.

V. Optimal TDFs

Financial advisors argue that the share of wealth in stocks should decrease as the investor approaches retirement and also quantify this as what the i.i.d. stock returns model predicts. The target date fund (TDF) using the results from the i.i.d. stock returns model has therefore become quite a popular financial advice, commonly recommended by large financial advisors like Vanguard TDFs. When the stock returns are predictable, however, the share of wealth in stocks should change according to market timing. This is what the perfect predictor model predicts. Retrospecting the financial crisis in 1929, 1997, 2001 and 2008, blindly following the rules suggested by life style funds for households entering retirement would not have been sound investment advice. Hence, the enhanced TDF (eTDF) has been proposed to take advantage of changing market conditions and expectations. This paper, however, shows that it is not easy to take advantage of changing market conditions and expectations. When the predictor is imperfect, the investment decision from the perfect predictor model seems over optimistic. But, what is the quantitative magnitude of investor welfare from investment rules given by the different models? In this section I evaluate the welfare loss of the investor with respect to the perfect predictor model, the i.i.d. stock returns model and the Vanguard TDF model when the stock market is modeled as the imperfect predictive system.

A. Welfare Evaluation

To measure welfare changes I use the value functions across different models. Given that I have solved for saving, portfolio choices and value functions for all periods in the life cycle, I know that the value functions at a particular age are a sufficient statistic for welfare effects. Let $V_0\left(x_{i,t},\, \hat{\mu}_{t|t}\right)$ be the value function from the baseline model, and $V_n(x_{i,t},\ f_t)$ be the value function from the alternative model such as the perfect predictor model or the i.i.d. stock returns model or the Vanguard TDF model. In these notations, $\hat{\mu}_{t|t}$ is the conditional expectation of unobservable expected stock return and f_t is the observed state factor. The f_t in perfect predictor model or Vanguard TDF model is the dividend yield. In contrast, f_t in i.i.d. stock returns model is a null variable because the policy functions are all the same for the different dividend yield.

I compute consumption certainty equivalent for a particular age as the follows:

$$E\left\{\left[\frac{V_n(x_{i,t},\ f_t)}{V_0\left(x_{i,t},\ \hat{\mu}_{t|t}\right)}\right]^{1/(1-\gamma)}-1\right\} \tag{2.18}$$

where $i \in I_{age}$ and $x_{i,t}$ is the same in both V_0 and V_n. This definition is the consumption certainty equivalent because I convert the change of the value into the dimension of expenditure before taking the unconditional expectation. Moreover, this consumption certainty equivalent is computed when stock returns are simulated based on the imperfect predictive system.

Figure 13 plots the consumption certainty equivalent of the different models relative to the baseline mode over the life cycle when changing $\rho_{v\varepsilon}$ from 0.2 to 0.8. Graph A illustrates substantial welfare loss from following the strategy predicted by the perfect predictor model relative to using the optimal investment policy given by the imperfect predictor model. Graph B shows that the welfare losses are even more substantial from following the i.i.d. stock returns model, and Graph C reports that welfare loss from taking the investment rules from the Vanguard TDF model is in the middle.

Several observations can be drawn based on Figure 13. First, the welfare losses are economically significant: they vary from 2% to 4% of consumption equivalent when the investor follows the strategy from the perfect predictor model, from 5% to 11% when the investor follows the strategy from the i.i.d. stock returns model and from 2% to 6% when investor follows the strategy from the Vanguard TDF model. Second, these welfare losses positively correlated with $\rho_{v\varepsilon}$. When the predictor is close to perfect positive ($\rho_{v\varepsilon} \approx 1$), the welfare loss becomes even larger. Third, the welfare losses from the i.i.d. stock returns model and perfect predictor model tend to get maximized at around age 50, whereas average wealth accumulation is maximized at the exogenous retirement age of 65. On the contrary, the welfare losses from the Vanguard TDF model has a peak at around age 70.

To better understand these welfare shapes and magnitudes, it is helpful to recall that given the preference for consumption smoothing, welfare is increasing in average consumption and decreasing in the volatility of consumption. I can therefore obtain an insight on where the welfare differences are coming from by comparing the mean change of consumption and the change of standard deviation of consumption over the life cycle. To do so, I define the average change of consumption for a particular age as $E_t\left(\frac{C_1-C_2}{C_2}\right)$, where C_1 is the consumption stream from the first model and C_2

is from the second model, and the change of standard deviation of consumption as $\frac{SD(C_1)-SD(C_2)}{SD(C_2)}$.

Figure 14 plots the mean change of consumption and the change of standard deviation of consumption for the baseline model relative to the perfect predictor model (Graphs A and B), the i.i.d. stock returns model (Graphs C and D) and the Vanguard TDF model (Graphs E and F). The I.I.D model produces the largest change in consumption volatility over the working life. 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 losses depicted in Figure 14.

The perfect predictor model on the other hand generates a lower mean consumption change over the life cycle. Since the portfolio rules are more stable than i.i.d. stock returns model, consumption variability is actually lower with the perfect predictor model relative to the i.i.d. stock returns model but higher relative to the baseline model.

When compared to the Vanguard TDF model, the welfare loss approaches the peak at about age 70 (see Figure 14, Graph E) because the mean change of consumption reaches the summit at age 70 (Figure 14, Graph E and F).

VI. Conclusions

In this paper, I jointly analyze the implications of an imperfect predictive system, undiversifiable labor income risk and exogenously imposed liquidity constraints on optimal consumption and portfolio decisions over the life cycle. In the presence of an imperfect predictor of the unobservable expected stock returns, the optimal portfolio choice is shown to be more conservative than that predicted by an i.i.d. and perfect predictor model when calibrated to the observed data from 1946 to 2015. Different from Wachter and Yogo (2010) which use the nonhomothetic utility over basic and luxury goods to generate balanced portfolios, this paper generates the balanced portfolios through introducing the imperfection to the predictor of stock returns. Compared with the SCF 2007, the imperfect predictor model matches the data better than the i.i.d. stock returns model and the perfect predictor model. Moreover, when the imperfection is introduced into the perfect predictor model, almost all correlations ($\rho_{z\varepsilon}, \rho_{zv}, \rho_{zn}, \rho_{v\varepsilon}$) become important, which is different from one of the conclusion of Michaelides and Zhang (2016). Hence, a financial advisor should pay more attention to these correlations when giving investment advice.

To measure the benefits of taking the imperfect predictor into account, I compare the welfare loss of the perfect predictor, i.i.d. and Vanguard TDF model relative to baseline model. The largest welfare loss is obtained from following the rules predicted by the i.i.d. stock returns model. The perfect predictor model has the smallest welfare loss, and the Vanguard TDF model is in the middle. Hence, an investment strategy uniquely relying on a single information source or the unconditional expected stock return leads to an incorrect investment decision and substantial welfare loss.

Future directions of research include the explicit introduction of ambiguity aversion in preferences, ambiguity in the parameters such as the risk premium and persistence of the unobservable expected return process, Bayesian posterior distributions for the parameters (Barberis (2000), Xia (2001) and Pastor and Stambaugh (2009)), a stochastic volatility in stock returns and an explicit model of housing. All these extensions will require additional computational power to achieve the desired required solution accuracy, but will further improve our understanding of life cycle portfolio choice under uncertainty and provide reasonable advice to billions of households increasingly making their own financial decisions.

VII. Appendix

A. A Theorem of Multivariate Normal Distribution

Theorem 7.1 (Tsay (2010), Ch11): Suppose that x and y are two random vectors such that their joint distribution is multivariate normal. In addition, assume that the diagonal block covariance matrix Σ_{ww} is non-singular for w = x, y. Then,

1. $E(x|y) = \mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)$
2. $Var(x|y) = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma'_{yx}$

This theorem provides us with an algorithm of filtering the unobservable state, x, from the observable variables y.

B. *Definitions and Notations*

- μ_t is the unobservable expected stock return
- q_t is the observable dividend yield.
- E_r is the unconditional expectation of μ_t;
- E_q is the unconditional expectation of q_t;
- $d_t = \begin{pmatrix} r_t \\ q_t \end{pmatrix}$, $D_t = (d_0, d_1, ..., d_t)$, where D_t is the full history of d_t;

- $e_t = E\left(d_t|\mu_t,\, D_{t-1}\right)$ is the expectation of observable d_t conditional on the current unobservable μ_t and historical observed values of d_t up to time $t-1$;

- $\hat{\mu}_{t|t-1} = E\left(\mu_t|D_{t-1}\right)$ is the expectation of unobservable μ_t conditional on the historical observed values of d_t up to time $t-1$;

- $\hat{d}_{t|t-1} = E\left(d_t|D_{t-1}\right)$ is the expectation of observable d_t conditional on the historical observed values of d_t up to time $t-1$;

- $\hat{\mu}_{t|t} = E(\mu_t|D_t)$ is the estimation of unobservable μ_t conditional on the historical observed values of d_t up to time t;

- $G_t = Cov\left(d_t, \mu_t|D_{t-1}\right)$ is the covariance between the observable variables d_t and the unobservable μ_t conditional on the historical observed data up to time $t-1$.

- $P_t = Var\left(\mu_t|D_{t-1}\right)$ is the variance of unobservable μ_t conditional on the historical observed data up to $t-1$.

- $R_t = Var\left(d_t|\mu_t, D_{t-1}\right)$ is the variance of d_t conditional on the current unobservable expected stock return (μ_t) and the historical observed data up to time $t-1$.

- $Q_t = Var\left(\mu_t|D_t\right)$ is the variance of the unobservable μ_t conditional on the historical observed data up to t.

- $S_t = Var(d_t|D_{t-1})$ is the variance of d_t conditional on the historical observed data up to time $t-1$.

C. Maximum Likelihood Estimation

Since the expected stock return (μ_t) is unobservable, the classical maximum likelihood estimation needs a modification that rewrites the VAR (2.24) to make it only involve the observable variables, the stock returns (r_t) and the predictor (q_t). For doing this modification, I follow the same idea of Pastor and Stambaugh (2012) and use this method to estimate the parameters of the imperfect predictive system conditional on all the observed data. The first step is to set up the recursive formula to update conditional moments of the unobservable expected stock return using the observed data.

Starting the Recursion

The recursion begins with $\mu_0|D_0$, where D_0 is the empty set. At time 0, since there are no observations of d_t, I use the unconditional mean of μ_t as the estimation of unobservable expected stock returns, $\hat{\mu}_{0|0} = E(\mu_0|D_0) = E_r$. Similarly, the $Q_0 = Var\,(\mu_0|D_0) = V_{\mu\mu}$ is the unconditional variance of μ_t.

At time 0, I can predict the moments of d_1 conditional on D_0. Assuming that the process of μ_t is stationary, I obtain $E\,(\mu_1|D_0) = \alpha_q + \phi_\mu E\,(\mu_0|D_0)$. This implies:

$$\hat{\mu}_{1|0} = E\,(\mu_1|D_0) = E_r;$$

$$P_1 = Var\,(\mu_1|D_0) = \Sigma_{\mu\mu};$$

and

$$\hat{d}_{1|0} = E\,(d_1|D_0) = \begin{pmatrix} 0 \\ \\ \alpha_q \end{pmatrix} + E\left(\begin{bmatrix} \mu_0 \\ \\ q_0 \end{bmatrix} |D_0\right) \Longrightarrow$$

$$\hat{d}_{1|0} = \begin{pmatrix} E_r \\ \\ E_q \end{pmatrix} \text{and } S_1 = Var\left(d_1|D_0\right) = \begin{bmatrix} \Sigma_{rr} & \Sigma_{rq} \\ \\ \Sigma_{qr} & \Sigma_{qq} \end{bmatrix}.$$

As $(d_1|D_0)$ and $(\mu_1|D_0)$ follow bi-variate normal distribution, I can obtain the following through applying theorem A;

$$e_1 = E\left(d_1|\mu_1,\, D_0\right) = \begin{pmatrix} E_r \\ \\ E_q \end{pmatrix} + \begin{bmatrix} \Sigma_{r\mu} \\ \\ \Sigma_{q\mu} \end{bmatrix} \Sigma_{\mu\mu}^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right)$$

and

$$R_1 = Var\left(d_1|\mu_1, D_0\right) = \begin{bmatrix} \Sigma_{rr} & \Sigma_{rq} \\ \\ \Sigma_{qr} & \Sigma_{qq} \end{bmatrix} - \begin{bmatrix} \Sigma_{r\mu} \\ \\ \Sigma_{q\mu} \end{bmatrix} \Sigma_{\mu\mu}^{-1} \begin{bmatrix} \Sigma_{r\mu} & \Sigma_{q\mu} \end{bmatrix}$$

Hence, I find the formulae for e_1 and R_1 as follows:

$$e_1 = \hat{d}_{1|0} + G_1 P_1^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right) \tag{2.19}$$

where $G_1 = Cov\left(d_1, \mu_1|D_0\right) = \begin{bmatrix} \Sigma_{r\mu} \\ \\ \Sigma_{q\mu} \end{bmatrix}$

and

$$R_1 = S_1 - G_1 P_1^{-1} G_1' \tag{2.20}$$

where $S_1 = \begin{bmatrix} \Sigma_{rr} & \Sigma_{rq} \\ \Sigma_{qr} & \Sigma_{qq} \end{bmatrix}$ and $P_1^{-1} = \Sigma_{\mu\mu}^{-1}$

Updating via Bayes Theorem

Recalling the conditional probability formula, I can obtain the following equation, $f\left(\mu_1|d_1, D_0\right) = \frac{f(d_1|\mu_1,D_0)f(\mu_1|D_0)}{f(d_1|D_0)}$. Since $D_1 = (d_1,\ D_0)$, I have a probability density function as follows.

$$f\left(\mu_1|d_1, D_0\right) \propto f\left(d_1|\mu_1, D_0\right) f\left(\mu_1|D_0\right) \tag{2.21}$$

Now, I expand the right hand side of equation (2.21) as follows:

$$f\left(\begin{bmatrix} r_1 & q_1 \end{bmatrix} |\mu_1, D_0\right) f\left(\mu_1|D_0\right) =$$

$$\left(\frac{1}{\sqrt{2\pi}}\right)^2 |P_1|^{-1}|R_1|^{-1} exp\left\{-\frac{1}{2}\left[\left(d_1 - e_1\right)' R^{-1}\left(d_1 - e_1\right) + \left(\mu_1 - \hat{\mu}_{1|0}\right) P^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right)\right]\right\}$$

$$\propto exp\left\{-\frac{1}{2}\left(d_1 - \hat{d}_{1|0} + G_1 P_1^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right)\right)' R^{-1}\left(d_1 - \hat{d}_{1|0} + G_1 P_1^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right)\right)\right.$$

$$\left. + \left(\mu_1 - \hat{\mu}_{1|0}\right) P^{-1}\left(\mu_1 - \hat{\mu}_{1|0}\right)\right\}$$

After rearranging the terms and ignoring the irrelevant quadratic terms of d_1, I can obtain the follows:

$$f\left([\ r_1,\ \ q_1\ ]|\mu_1, D_0\right) f\left(\mu_1|D_0\right) \propto$$

$$exp\left\{-\tfrac{1}{2}\left[\left(d_1-\hat{d}_{1|0}\right)' G_1P_1^{-1}R^{-1}\left(\mu_1-\hat{\mu}_{1|0}\right)+\left(\mu_1-\hat{\mu}_{1|0}\right)'\left(P_1^{-1}\right)' G_1' R^{-1}\left(d_1-\hat{d}_{1|0}\right)\right.\right.$$

$$\left.\left.+\left(\mu_1-\hat{\mu}_{1|0}\right)'\left(P_1^{-1}\right)' G_1' R^{-1} G_1 P_1^{-1}\left(\mu_1-\hat{\mu}_{1|0}\right)+\left(\mu_1-\hat{\mu}_{1|0}\right) P^{-1}\left(\mu_1-\hat{\mu}_{1|0}\right)\right]\right\}.$$

Since μ_t, P_t and $\left(d_1-\hat{d}_{1|0}\right)' G_1P_1^{-1}R^{-1}$ are scalars, I can rewrite this formula as

$$f\left([\ r_1,\ \ q_1\ ]|\mu_1, D_0\right) f\left(\mu_1|D_0\right) \propto$$

$$exp\left\{-\tfrac{1}{2}\left[2P_1^{-1}G_1'R^{-1}\left(d_1-\hat{d}_{1|0}\right)\left(\mu_1-\hat{\mu}_{1|0}\right)+\right.\right.$$

$$\left.\left.\left(P_1^{-1}+\left(P_1^{-1}\right)' G_1' R^{-1} G_1 P_1^{-1}\right)\left(\mu_1-\hat{\mu}_{1|0}\right)^2\right]\right\}$$

$$=exp\left\{-\tfrac{1}{2}\left[2P_1^{-1}G_1'R^{-1}\left(d_1-\hat{d}_{1|0}\right)\left(\mu_1-\hat{\mu}_{1|0}\right)\right.\right.$$

$$\left.\left.+P_1^{-1}P_1\left(P_1^{-1}+\left(P_1^{-1}\right)' G_1' R^{-1} G_1 P_1^{-1}\right) P_1P_1^{-1}\left(\mu_1-\hat{\mu}_{1|0}\right)^2\right]\right\}$$

$$=exp\left\{-\tfrac{1}{2}\left[2P_1^{-1}G_1'R^{-1}\left(d_1-\hat{d}_{1|0}\right)\left(\mu_1-\hat{\mu}_{1|0}\right)+\right.\right.$$

$$\left.\left.P_1^{-1}\left(P_1+G_1'R^{-1}G_1\right)P_1^{-1}\left(\mu_1-\hat{\mu}_{1|0}\right)^2\right]\right\}$$

$$\propto exp\left\{-\frac{\left[\mu_1-\hat{\mu}_{1|0}-P_1\left(P_1+G_1'R^{-1}G_1\right)^{-1}G_1'R^{-1}\left(d_1-\hat{d}_{1|0}\right)\right]^2}{2P_1\left(P_1+G_1'R^{-1}G_1\right)^{-1}P_1}\right\}.$$

This is the kernel of the normal distribution again. Hence, after the

Bayesian update, conditional probability density of μ_t is still a normal distribution, and its conditional moments are:

$$\hat{\mu}_{1|1} = E\left(\mu_1 | D_1\right) = \hat{\mu}_{1|0} + P_1 \left(P_1 + G_1' R_1^{-1} G_1\right)^{-1} G_1' R_1^{-1} \left(d_1 - \hat{d}_{1|0}\right) \quad (2.22)$$

and

$$Var\left(\mu_1 | D_1\right) = Q_1 = P_1 \left(P_1 + G_1' R_1^{-1} G_1\right)^{-1} P_1 \quad (2.23)$$

For finding all conditional probability densities and moments of μ_t for t = 2, ... , T, I rewrite the equations (2.4) - (2.6) as follows:

$$\begin{bmatrix} r_t - r_f - E_r \\ q_t - E_q \\ \mu_t - E_r \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_q & 0 \\ 0 & 0 & \phi_\mu \end{bmatrix} \begin{bmatrix} r_{t-1} - r_f - E_r \\ q_{t-1} - E_q \\ \mu_{t-1} - E_r \end{bmatrix} + \begin{bmatrix} z_t \\ v_t \\ \varepsilon_t \end{bmatrix} \quad (2.24)$$

Assuming the VAR represented by (2.24) is stationary, I obtain the following:

$$E\left(\mu_t - E_r|D_{t-1}\right) = \phi_\mu E\left(\mu_{t-1} - E_r|D_{t-1}\right) \Longrightarrow$$

$$E\left(\mu_t|D_{t-1}\right) = \left(1-\phi_\mu\right)E_r + \phi_\mu E\left(\mu_{t-1}|D_{t-1}\right).$$

Since the μ_t is unobservable, I can not delete the $E(.)$ in $E\left(\mu_{t-1}|D_{t-1}\right)$.

Simplifying the formula above, I rewrite it as:

$$\hat{\mu}_{t|t-1} = \left(1-\phi_\mu\right)E_r + \phi_\mu \hat{\mu}_{t-1|t-1} \tag{2.25}$$

Similarly, I can rewrite $\hat{d}_{t|t-1}$ as the following:

$$\hat{d}_{t|t-1} = E\left(d_t|D_{t-1}\right) = \begin{bmatrix} \hat{\mu}_{t-1|t-1} \\ \left(1-\phi_q\right)E_q + \phi_q q_{t-1} \end{bmatrix} \tag{2.26}$$

and

$$P_t = Var\left(\mu_t|D_{t-1}\right) = \phi_\mu^2 Var\left(\mu_{t-1}|D_{t-1}\right) + \sigma_\varepsilon^2 = \phi_\mu^2 Q_{t-1} + \sigma_\varepsilon^2 \tag{2.27}$$

Next, for getting the other update formulae, I take the variance on both sides of the VAR (2.24).

$$Var\left(\begin{bmatrix} r_t \\ q_t \\ \mu_t \end{bmatrix} | D_{t-1}\right) =$$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_q & 0 \\ 0 & 0 & \phi_\mu \end{bmatrix}' Var\left(\begin{bmatrix} r_{t-1} - E_r \\ q_{t-1} - E_q \\ \mu_{t-1} - E_r \end{bmatrix} | D_{t-1}\right) \begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_q & 0 \\ 0 & 0 & \phi_\mu \end{bmatrix} + Var\left(\begin{bmatrix} z_t \\ v_t \\ \varepsilon_t \end{bmatrix} | D_{t-1}\right)$$

$$= \begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_x & 0 \\ 0 & 0 & \phi_\mu \end{bmatrix}' \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Q_{t-1} \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & \phi_q & 0 \\ 0 & 0 & \phi_\mu \end{bmatrix} + \begin{bmatrix} \sigma_z^2 & \rho_{zv}\sigma_z\sigma_v & \rho_{z\varepsilon}\sigma_z\sigma_\varepsilon \\ & \sigma_v^2 & \rho_{v\varepsilon}\sigma_v\sigma_\varepsilon \\ & & \sigma_\varepsilon^2 \end{bmatrix}$$

$$= \begin{bmatrix} S_t & G_t \\ G_t' & P_t \end{bmatrix}$$

Hence, I obtain

$$e_t = E\left(d_t | \mu_t, D_{t-1}\right) = \hat{d}_{t|t-1} + G_t P_t^{-1}\left(\mu_t - \hat{\mu}_{t|t-1}\right) \tag{2.28}$$

$$R_t = Var\left(d_t | \mu_t,\, D_{t-1}\right) = S_t - G_t P_t^{-1} G_t' \tag{2.29}$$

and, the conditional moments of μ_t

$$\hat{\mu}_{t|t} = E\left(\mu_t | D_t\right) = \hat{\mu}_{t|t-1} + P_t\left(P_t + G_t' R_t^{-1} G_t\right)^{-1} G_t' R_t^{-1}\left(d_t - \hat{d}_{t|t-1}\right) \tag{2.30}$$

$$Q_t = Var\left(\mu_t | D_t\right) = P_t\left(P_t + G_t' R_t^{-1} G_t\right)^{-1} P_t \tag{2.31}$$

Maximum Likelihood Estimation of Parameter

Denote $[m_t, n_t] = P_t\left(P_t + G_t' R_t^{-1} G_t\right)^{-1} G_t' R_t^{-1}$. Then, according to (2.30), I have

$$\hat{\mu}_{t|t} = \hat{\mu}_{t|t-1} + [m_t, n_t]\left[\begin{pmatrix} r_t - r_f - \hat{\mu}_{t-1|t-1} \\ \\ x_t - E_{t-1}(q_t) \end{pmatrix}\right]$$
$$= \left(1 - \phi_\mu\right) E_r + \phi_\mu \hat{\mu}_{t-1|t-1} + m_t\left(r_t - r_f - \hat{\mu}_{t-1|t-1}\right) + n_t v_t$$

The last equality hold because $\hat{\mu}_{t|t-1} = \left(1 - \phi_\mu\right) E_r + \phi_\mu \hat{\mu}_{t-1|t-1}$ (equation 2.25).

Then, I rewrite this formula as follows.

$$\hat{\mu}_{t|t} = \left(1 - \phi_{\mu}\right) E_r + \left(\phi_{\mu} - m_t\right) \hat{\mu}_{t-1|t-1} + m_t(r_t - r_f) + n_t v_t \tag{2.32}$$

Next, I define the forecast error of $r_{t+1} - r_f$ conditional on information at time t as

$$\omega_{t+1} = (r_t - r_f) - E_t\left(r_t - r_f\right) \tag{2.33}$$

Since $r_{t+1} - r_f = \mu_t + z_{t+1}$ (equation (2.5)), I have $E_t\left(r_{t+1} - r_f\right) = E_t\left(\mu_t\right) = \hat{\mu}_{t|t}$.

Now, replacing $E_t\left(r_t - r_f\right)$ with $\hat{\mu}_{t|t}$ in (2.33), I obtain $\omega_{t+1} = r_{t+1} - r_f - \hat{\mu}_{t|t}$.

Rearrange this equation, I obtain as the following:

$$r_{t+1} - r_f = \left(1 - \phi_{\mu}\right) E_r + \phi_{\mu}(r_t - r_f) + n_t v_t - \left(\phi_{\mu} - m_t\right) \omega_t + \omega_{t+1} \tag{2.34}$$

Combining (2.34) with (2.6), I obtain a new equation system consisting of only observable data as follows:

$$\begin{cases} r_{t+1} - r_f = \left(1-\phi_\mu\right) E_r + \phi_\mu (r_t - r_f) + n_t v_t - \left(\phi_\mu - m_t\right) \omega_t + \omega_{t+1} \\ \\ q_{t+1} = \alpha_q + \phi_q q_t + v_{t+1} \end{cases} \tag{2.35}$$

In steady state, I can delete the time index of m_t and n_t and rewrite the equation system (2.35) as the following VARMA(1,1) model.

$$\begin{bmatrix} r_t - r_f \\ \\ q_t \end{bmatrix} - \begin{bmatrix} E_r \\ \\ E_q \end{bmatrix} = \begin{bmatrix} \phi_\mu & 0 \\ \\ 0 & \phi_q \end{bmatrix} \left(\begin{bmatrix} r_{t-1} - r_f \\ \\ q_{t-1} \end{bmatrix} - \begin{bmatrix} E_r \\ \\ E_q \end{bmatrix} \right) + \tag{2.36}$$

$$\begin{bmatrix} -\left(\phi_\mu - m\right) & n \\ \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{t-1} \\ \\ v_{t-1} \end{bmatrix} + \begin{bmatrix} \omega_t \\ \\ v_t \end{bmatrix}$$

Now, define the state variable as $\xi_t = (\omega_t, v_t, \omega_{t-1}, v_{t-1})'$. This

VARMA(1,1) model can be rewritten as a state space model consisting of a observation equation and a state equation and estimated by the Kalman Filter algorithm.

$$\begin{cases} d_t^* = A d_{t-1}^* + H\xi_t \\ \xi_t = F\xi_{t-1} + \begin{bmatrix} e_t \\ 0 \end{bmatrix} \end{cases} \tag{2.37}$$

where $d_t^* = d_t - [E_r, E_q]'$, $A = \begin{bmatrix} \phi_\mu & 0 \\ 0 & \phi_q \end{bmatrix}$, $H_{12} = \begin{bmatrix} -\left(\phi_\mu - m\right) & n \\ 0 & 0 \end{bmatrix}$,

$F = \begin{bmatrix} 0_{2\times2} & 0_{2\times2} \\ I_{2\times2} & 0_{2\times2} \end{bmatrix}$, $H = \begin{bmatrix} I_{2\times2} & H_{12} \end{bmatrix}$, $e_t = [\omega_t, v_t]'$, and $Var(e_t) = \Sigma^*$.

Given a sample of d_t, the joint likelihood function of the state space model (2.37) is $L = \prod_{t=1}^{T} f\left(d_t^* | d_{t-1}^*\right)$, where $f\left(d_t^* | d_{t-1}^*\right)$ is the conditional probability density of d_t^*. I, therefore, can estimate the parameters by maximizing the following log-likelihood function[11]:

$$-2ln(L) = \sum_{t=1}^{T} \left(ln|V_{t|t-1}| + \left[d_t^* - \hat{d}_{t|t-1}^*\right]' V_{t|t-1}^{-1} \left[d_t^* - \hat{d}_{t|t-1}^*\right] \right) \tag{2.38}$$

The terms in equation (2.38) are defined as the following:

- $Var(r_t) = \sigma_r^2 = (1-\phi_\mu)^{-1} \left[n\sigma_v^2 n + (1-\phi_\mu^2 + m^2)\sigma_\omega^2 + 2m\sigma_{v\omega}n\right]$ [12];
- $Cov(r_t, q_t) = \sigma_{qr} = (1-\phi_\mu\phi_q)^{-1} \left[\phi_q\sigma_v^2 n + \left(1-(\phi_\mu - m)\phi_q\right)\sigma_{v\omega}\right]$ [13];
- $Var(q_t) = \sigma_q^2 = (1-\phi_q^2)^{-1}\sigma_v^2$;
- $\hat{d}_{t|t-1}^* = E_{t-1}(d_t^*) = Ad_{t-1}^* + H_{12}\Sigma^* V_{t-1|t-2}^{-1} \left[d_{t-1}^* - \hat{d}_{t-1|t-2}^*\right]$; and
- $V_{t|t-1} = V_{t-1}(d_t^*) = H_{12}\left(\Sigma^* - \Sigma^* V_{t-1|t-2}^{-1}\Sigma^*\right) H_{12}' + \Sigma^*$

[11] This log-likelihood function (2.38) is based on the logarithm of the equation (13.4.1) from Hamilton (1994)

[12] Taking variance on both sides of (2.34), I obtain $\sigma_r^2 = \phi_\mu^2\sigma_r^2 - 2\phi_\mu(\phi_\mu - m)\sigma_\omega^2 + (\phi_\mu - m)^2\sigma_\omega^2 + n^2\sigma_v^2 + \sigma_\omega^2 - 2(\phi_\mu - m)n\sigma_{v\omega} + 2\phi_\mu n\sigma_{v\omega}$. Solve this equation for σ_r^2

[13] Taking covariance between r_t and q_t based on (2.35), I obtain the equation $\sigma_{qr} = \phi_q\phi_\mu\sigma_{xr} + n\phi_q\sigma_v^2 - (\phi_\mu - m)\phi_q\sigma_{v\omega} + \sigma_{v\omega}$. Then, σ_{qr} is from solving this equation.

The formula of updating $\hat{d}^*_{t|t-1}$ is from taking $E_{t-1}(\cdot)$ on the both sides of the observation equation of (2.37).

$$\hat{d}^*_{t|t-1} = E_{t-1}(d^*_t) = Ad^*_{t-1} + H\hat{\xi}_{t|t-1}$$

$$= Ad^*_{t-1} + HF\hat{\xi}_{t-1|t-1}$$

$$= Ad^*_{t-1} + \begin{bmatrix} H_{12} & 0_{2\times 2} \end{bmatrix} \left(\hat{\xi}_{t-1|t-2} + \Sigma_{\xi d^*} V^{-1}_{t-1|t-2}\left[d^*_{t-1} - \hat{d}^*_{t-1|t-2}\right]\right)$$

$$= Ad^*_{t-1} + \begin{bmatrix} H_{12} & 0_{2\times 2} \end{bmatrix} \begin{pmatrix} \hat{\omega}_{t-1|t-2} \\ \hat{\upsilon}_{t-1|t-2} \\ \hat{\omega}_{t-2|t-2} \\ \hat{\upsilon}_{t-2|t-2} \end{pmatrix} +$$

$$\begin{bmatrix} H_{12} & 0_{2\times 2} \end{bmatrix} \Sigma_{\xi d^*} V^{-1}_{t-1|t-2}\left[d^*_{t-1} - \hat{d}^*_{t-1|t-2}\right].$$

As $\begin{bmatrix} \hat{\omega}_{t-1|t-2} \\ \hat{\upsilon}_{t-1|t-2} \end{bmatrix} = 0$ and $\begin{bmatrix} H_{12} & 0_{2\times 2} \end{bmatrix} \Sigma_{\xi d^*} = H_{12}\Sigma^*$, this equation can be simplified as:

$$\hat{d}^*_{t|t-1} = Ad^*_{t-1} + H_{12}\Sigma^* V^{-1}_{t-1|t-2}\left[d^*_{t-1} - \hat{d}^*_{t-1|t-2}\right].$$

As to $V_{t|t-1}$, the update formula is from taking $Var(\cdot)$ on both sides of state equation of (2.37).

$$V_{t|t-1} = V_{t-1}(d_t^*) = HFVar_{t-1}(\xi_t)F'H' + HFVar\begin{pmatrix} e_t \\ \\ 0 \end{pmatrix}F'H'$$

$$= \begin{bmatrix} I_{2\times2} & H_{12} \end{bmatrix}\begin{bmatrix} 0_{2\times2} & 0_{2\times2} \\ \\ I_{2\times2} & 0_{2\times2} \end{bmatrix}\left(\Sigma_{\xi\xi} - \Sigma_{\xi d}V_{t-1|t-2}^{-1}\Sigma'_{d\xi}\right)\begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ \\ 0_{2\times2} & 0_{2\times2} \end{bmatrix}\begin{bmatrix} I_{2\times2} \\ \\ H_{12} \end{bmatrix}$$

$$+\begin{bmatrix} I_{2\times2} & H_{12} \end{bmatrix}\begin{bmatrix} 0_{2\times2} & 0_{2\times2} \\ \\ I_{2\times2} & 0_{2\times2} \end{bmatrix}\Sigma^*\begin{bmatrix} 0_{2\times2} & I_{2\times2} \\ \\ 0_{2\times2} & 0_{2\times2} \end{bmatrix}\begin{bmatrix} I_{2\times2} \\ \\ H_{12} \end{bmatrix}.$$

Since $\begin{bmatrix} I_{2\times2} & H_{12} \end{bmatrix}\begin{bmatrix} 0_{2\times2} & 0_{2\times2} \\ \\ I_{2\times2} & 0_{2\times2} \end{bmatrix} = H_{12}$, I have:

$$V_{t|t-1} = H_{12}\left(\Sigma^* - \Sigma^*V_{t-1|t-2}^{-1}\Sigma^*\right)H'_{12} + \Sigma^*.$$

Last, the necessary initial values for solving the log-likelihood function (2.38) are:

- $\hat{d}_{1|0}^* = 0$; and
- $V_{1|0} = Var(d_t) = \begin{bmatrix} Var(r_t) & Cov(r_t, q_t) \\ \\ & Var(q_t) \end{bmatrix}.$

Through the MLE, I can obtain the estimation of parameters in VARMA model (2.16). Our goal, however, is to find the covariance matrix of equations (4) - (6), $\Omega = \begin{bmatrix} \sigma_z^2 & \sigma_{zv} & \sigma_{z\varepsilon} \\ & \sigma_v^2 & \sigma_{v\varepsilon} \\ & & \sigma_\varepsilon^2 \end{bmatrix}$. Unfortunately, in Ω, σ_{vv} is the only term that can be identified. As to estimate the other terms in Ω, obviously, a good starting point is the moments because the moments of VARMA model (2.16) should equal to that of the original VAR model (2.24). The moments we need in VARMA model can be computed as follows:

- $\sigma_r^2 = (1-\phi_\mu)^{-1}\left[n\sigma_v^2 n + (1-\phi_\mu^2+m^2)\sigma_\omega^2 + 2m\sigma_{v\omega}n\right]$
- $\sigma_{qr} = (1-\phi_\mu\phi_q)^{-1}\left[\phi_q\sigma_v^2 n + \left(1-(\phi_\mu - m)\phi_q\right)\sigma_{v\omega}\right]$
- $Cov\left(r_t, r_{t-1}\right)$

  $= Cov\left(\phi_\mu(r_{t-1}-r_f) + nv_{t-1} - \left(\phi_\mu - m\right)\omega_{t-1} + \omega_t, r_{t-1} - r_f\right)$

$$= \phi_\mu \sigma_r^2 + nCov\left(v_{t-1}, r_{t-1} - r_f\right) - \left(\phi_\mu - m\right) Cov\left(\omega_{t-1}, r_{t-1} - r_f\right)$$

- $Cov\left(r_{t-1}, q_t\right)$

$$= Cov\left(r_{t-1}, \phi_q q_{t-1} + v_t\right)$$

$$= \phi_q Cov\left(r_{t-1}, q_{t-1}\right)$$

- $Cov\left(r_t, q_{t-1}\right)$

$$= Cov\left(\phi_\mu r_{t-1} + n v_{t-1} - \left(\phi_\mu - m\right) \omega_{t-1} + \omega_t, q_{t-1}\right)$$

$$= \phi_\mu \sigma_{rq} + nCov\left(v_{t-1}, q_{t-1}\right) - \left(\phi_\mu - m\right) Cov\left(\omega_{t-1}, q_{t-1}\right)$$

On the other hand, based on the VAR model (2.24), I can write down the following linear equation system about the moments as follows:

$$\begin{bmatrix} \sigma_r^2 \\ \sigma_{qr} \\ Cov\left(r_t, r_{t-1}\right) \\ Cov\left(r_{t-1}, q_t\right) \\ Cov\left(r_t, q_{t-1}\right) \end{bmatrix} =$$

$$\begin{bmatrix} 1/\left(1-\phi_{\mu}^{2}\right) & 1 & 0 & 0 & 0 \\ 0 & 0 & \phi_{q}/\left(1-\phi_{\mu}\phi_{q}\right) & 1 & 0 \\ \phi_{\mu}/\left(1-\phi_{\mu}^{2}\right) & 0 & 0 & 0 & 1 \\ 0 & 0 & \phi_{q}^{2}/\left(1-\phi_{\mu}\phi_{q}\right) & \phi_{q} & 0 \\ 0 & 0 & 1/\left(1-\phi_{\mu}\phi_{q}\right) & 0 & 0 \end{bmatrix} \begin{bmatrix} \sigma_{\varepsilon}^{2} \\ \sigma_{z}^{2} \\ \sigma_{v\varepsilon} \\ \sigma_{zv} \\ \sigma_{z\varepsilon} \end{bmatrix}$$

Since both the VARMA (2.16) and VAR (2.24) models describe the same thing, their moments must be the same as each other. This linear equation system, therefore, becomes solvable, when I replace the right hand side of this equation with the corresponding calculation from the VARMA model (2.16).

The rank of this linear equation system is, however, four because the second row can be eliminated by the fourth row, which makes this linear equation system be reduced into:

$$
\begin{cases}
\sigma_r^2 = \sigma_\varepsilon^2 / \left(1 - \phi_\mu^2\right) + \sigma_z^2 \\
\sigma_{qr} = \phi_q \sigma_{v\varepsilon} / \left(1 - \phi_\mu \phi_q\right) + \sigma_{zv} \\
Cov\left(r_t, r_{t-1}\right) = \sigma_{z\varepsilon} + \phi_\mu \sigma_\varepsilon^2 / \left(1 - \phi_\mu^2\right) \\
Cov\left(r_t, q_{t-1}\right) = \sigma_{v\varepsilon} / \left(1 - \phi_\mu \phi_q\right)
\end{cases}
\tag{2.39}
$$

(2.39) shows that this linear equation system has not unique solution but a solution space. For obtaining the solution space, I represent this linear equation system (2.39) with respect to $\sigma_{z\varepsilon}$, because this parameter is important in the estimation. Then, it becomes as follows:

$$
\begin{cases}
\sigma_z^2 = \sigma_r^2 - \left(Cov\left(r_t, r_{t-1}\right) - \sigma_{z\varepsilon}\right) / \phi_\mu \\
\sigma_\varepsilon^2 = \left(Cov\left(r_t, r_{t-1}\right) - \sigma_{z\varepsilon}\right) \left(1 - \phi_\mu^2\right) / \phi_\mu
\end{cases}
.
$$

$$
s.t.\ \rho_{zv} < 1\, and\, \rho_{v\varepsilon} < 1
$$

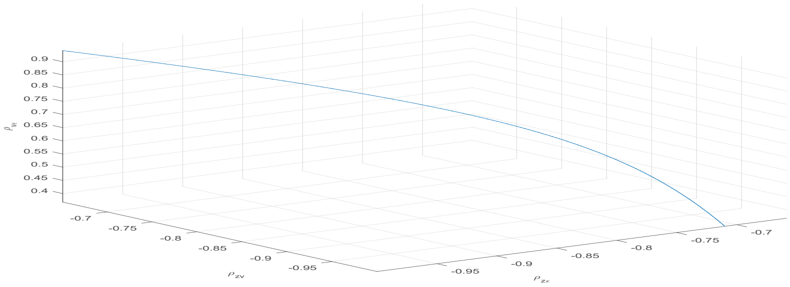
VIII. Appendix: Figures

FIGURE 1

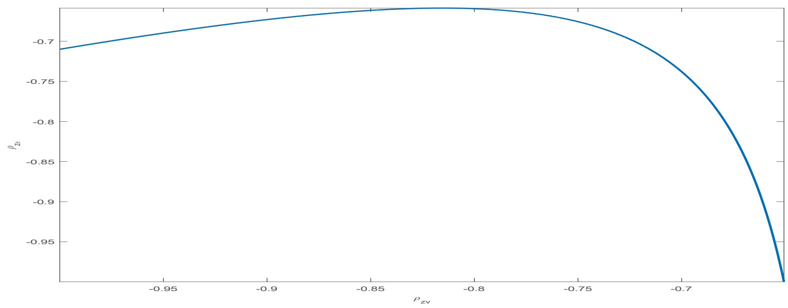
Solution Space of MLE

Figure 1 presents the solution space from MLE method. Graph A shows a 3D graph of the relationship between ρ_{zv}, $\rho_{z\varepsilon}$ and $\rho_{v\varepsilon}$, where ρ_{zv} is the correlation between shocks to stock returns and innovations of predictor, $\rho_{v\varepsilon}$ is the correlation between innovations of predictor and unobservable expected return, and $\rho_{z\varepsilon}$ is the correlation between innovations of stock returns and unobservable expected return. Graph B is the projection of the 3-D graph in Graph A on the plane of ρ_{zv} and $\rho_{z\varepsilon}$. Similarly, Graph C project the solution space on the plane of $\rho_{z\varepsilon}$ and $\rho_{v\varepsilon}$.

Graph A. Relationship among the Correlations



Graph B. Relationship between $\rho_{z\epsilon}$ and ρ_{zv}



Graph C. Relationship between $\rho_{z\epsilon}$ and $\rho_{v\epsilon}$

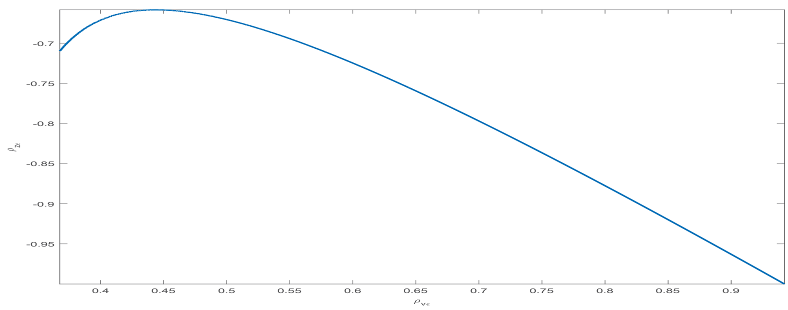
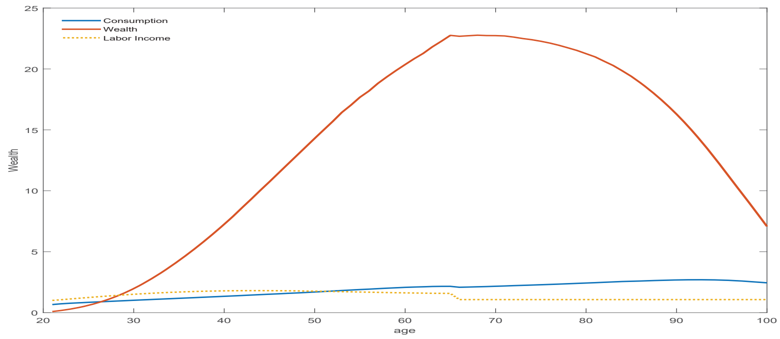


FIGURE 2

Mean Life-Cycle Profiles for Baseline Model

Figure 2, Graph A presents the mean wealth, consumption and labor income over the life cycle by simulating 6000 individual life histories. Graph B shows the mean share of wealth in stocks over the life cycle.

Graph A. Mean Consumption, Wealth and Labor Income for Baseline



Graph B. Mean Share of Wealth in Stocks for Baseline

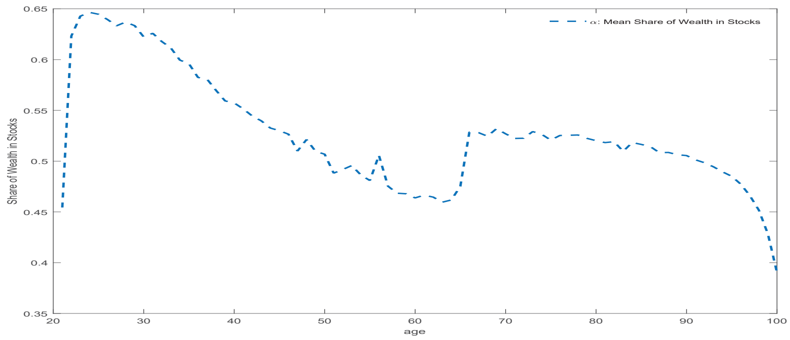
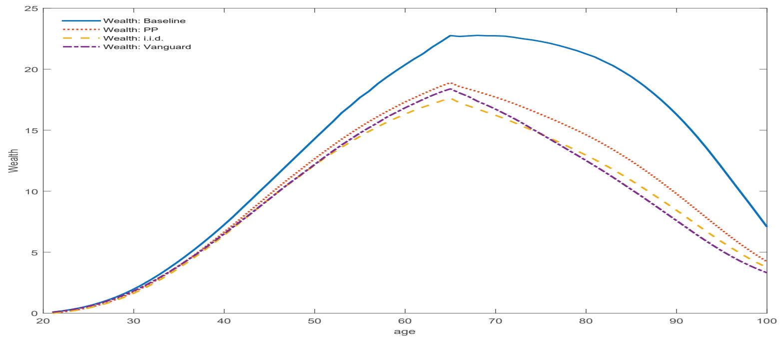


FIGURE 3

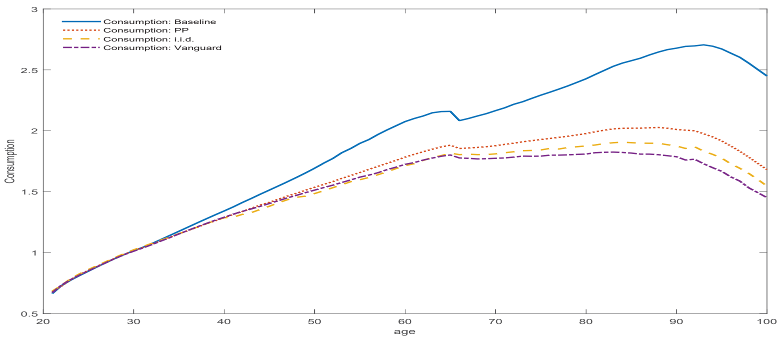
Comparison of Life-Cycle Profiles among Different Models

Figure 3, Graph A presents the mean wealth over the life cycle for four different models, where PP is the perfect predictor model, i.i.d. is the i.i.d. stock returns model and Vanguard represents the Vanguard TDF model. Graph B and C describe the mean consumption and share of wealth in stocks for the corresponding models.

Graph A. Comparison of Mean Wealth



Graph B. Comparison of Mean Consumption



Graph C. Comparison of Mean Share of Wealth in Stocks

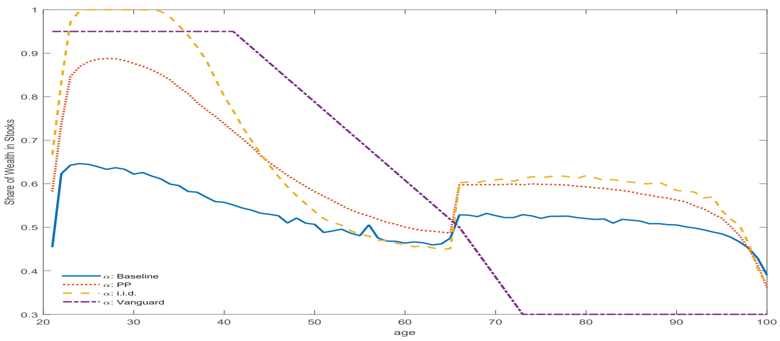
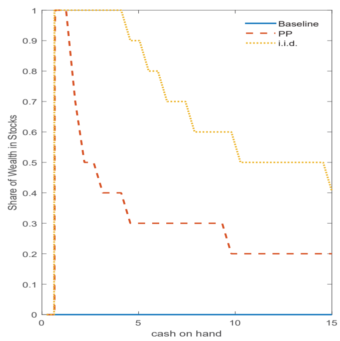


FIGURE 4

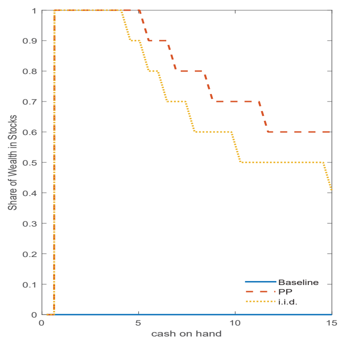
Investment Policy Function for Different Expected Stock Returns

Figure 4 presents the share allocation policy functions of different states, ages and models. The first row (Graph A, B and C) describes the share allocation policy functions between the baseline model (imperfect predictor model), perfect predictor model (PP model) and i.i.d. stock returns model for a 25-year-old investor when the estimations of the expected stock return is low, median and high respectively. The second row (Graph D, E and F) displays the share allocation policy functions for a 55-year-old agent, and the last row (Graph G, H and I) shows the share allocation policy functions for a 75-year-old agent.

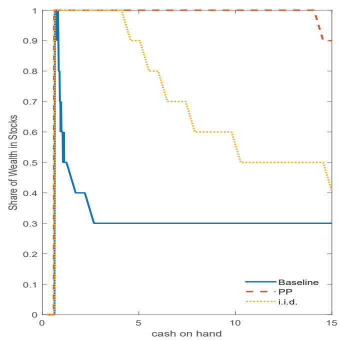
Graph A. Effect of μ_t on α When Estimate is Low, age 25



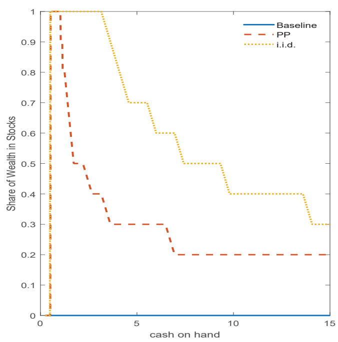
Graph B. Effect of μ_t on α When Estimate is Median, age 25



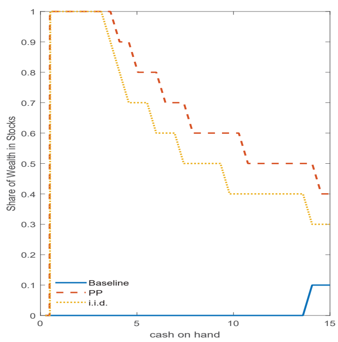
Graph C. Effect of μ_t on α When Estimate is High, age 25



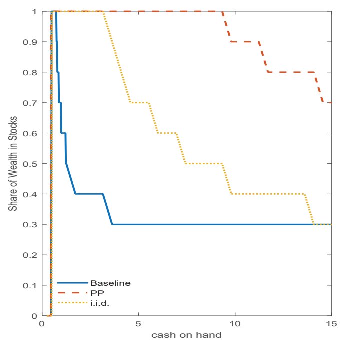
Graph D. Effect of μ_t on α When Estimate is Low, age 55



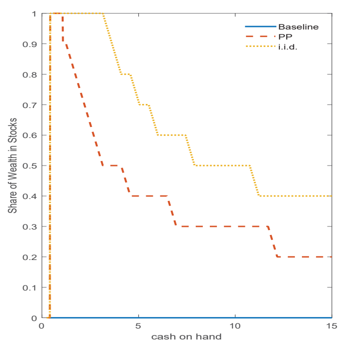
Graph E. Effect of μ_t on α When Estimate is Median, age 55



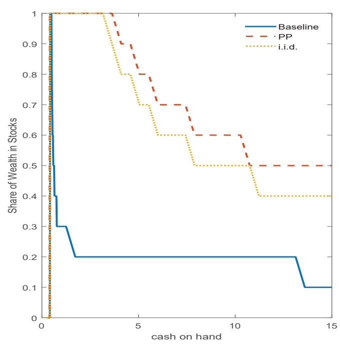
Graph F. Effect of μ_t on α When Estimate is High, age 55



Graph G. Effect of μ_t on α When Estimate is Low, age 75



Graph H. Effect of μ_t on α When Estimate is Median, age 75



Graph I. Effect of μ_t on α When Estimate is High, age 75

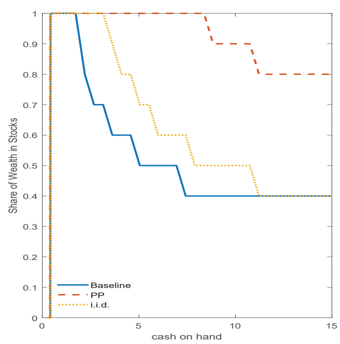
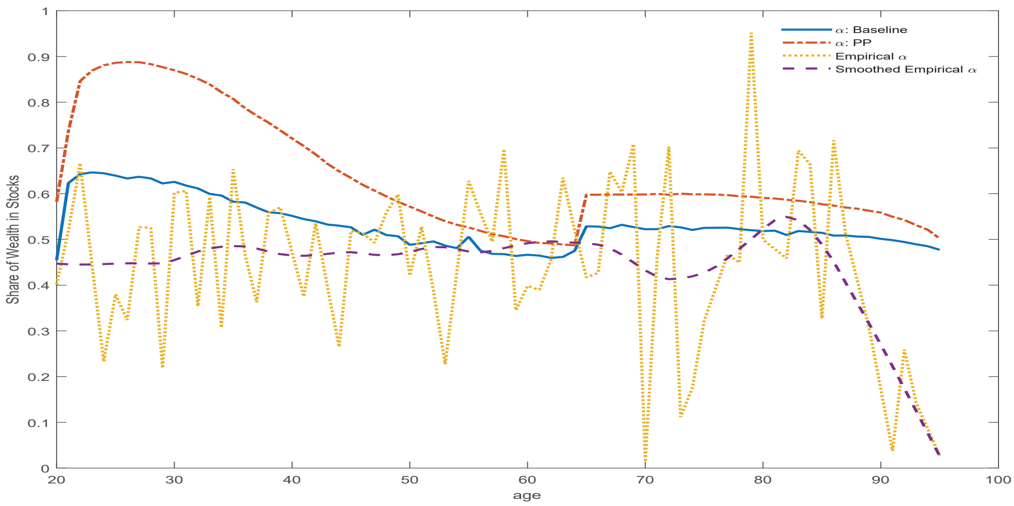


FIGURE 5

The Mean Share of Wealth in Stocks from the Baseline Model, Perfect Predictor Model and SCF Data

Figure 5, Graph A and B present the mean share of wealth in stocks (α) between the baseline model, perfect predictor model (PP model) and SCF data. The only difference between Graph A and B is the definition of the empirical α from SCF data. The empirical α in Graph A rules out the asset with high liquidity and is defined as equity/(equity+bonds). In contrast, in Graph B, the empirical α includes the asset with high liquidity and is calculated by equity/(equity+bonds+liquidity). I compute smoothed empirical α through running weighted linear least squares and a 2nd degree polynomial model with a span of 20% at each age.

Graph A. Mean Share of Wealth in Stocks of Baseline v.s. Empirical Portfolio Choice without Liquidity



Graph B. Mean Share of Wealth in Stocks of Baseline v.s. Empirical Portfolio Choice with Liquidity

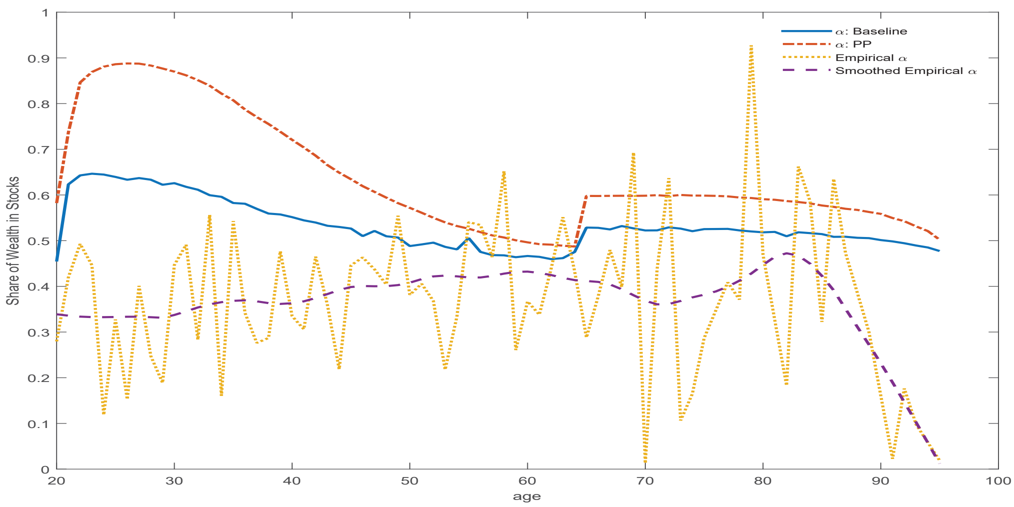
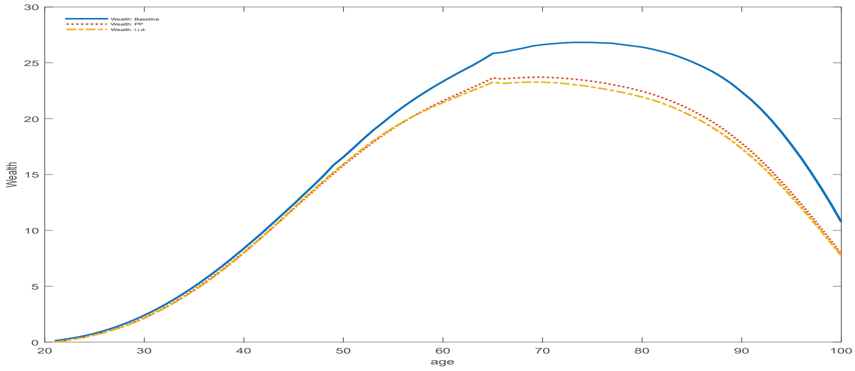


FIGURE 6

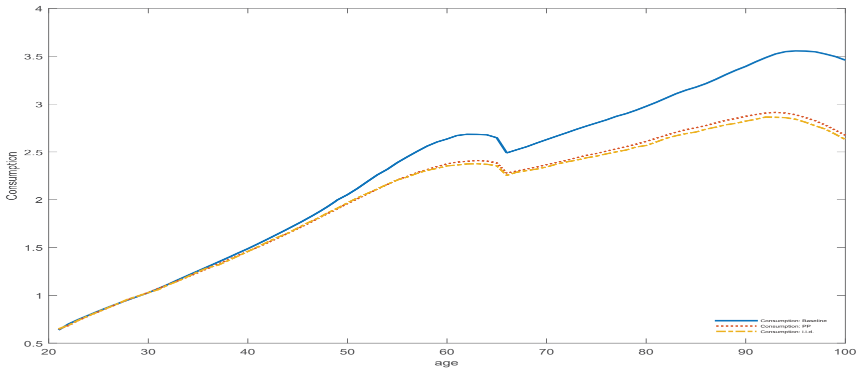
The Life-Cycle Profiles with High Risk Premium

Figure 6, Graph A, presents the comparison of mean wealth between the baseline model (imperfect predictor model), perfect predictor model (PP model) and i.i.d. stock returns model over the life cycle. Graph B reports the difference of mean consumptions between these models. Graph C compares mean share of wealth (α) in stocks.

Graph A. Comparison of Mean Wealth (Risk Premium=7%)



Graph B. Comparison of Mean Consumption (Risk Premium=7%)



Graph C. Comparison of Mean Share of Wealth in Stocks (Risk Premium=7%

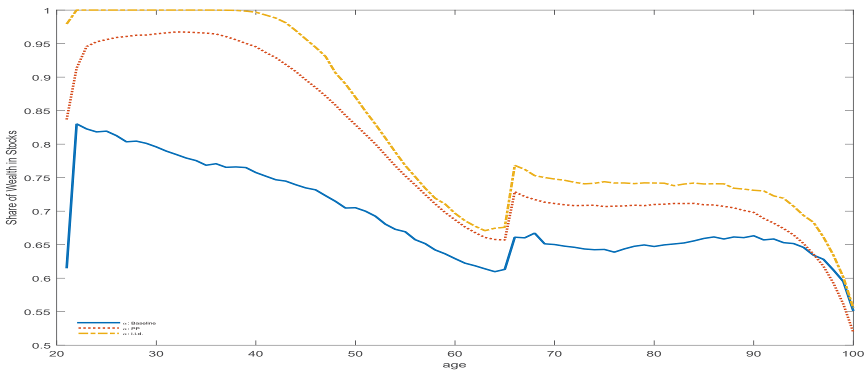
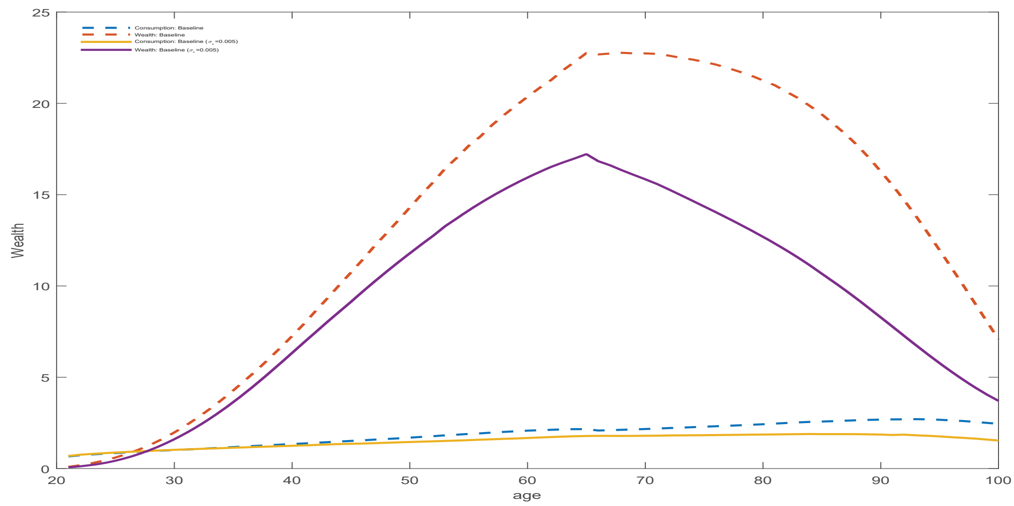


FIGURE 7

Life-Cycle Profile Comparison under the Low Volatility of Unobserved Expected Stock Returns

Figure 7, Graph A, presents how the low volatility of unobserved expected stock returns σ_ε affects the mean wealth accumulation and consumptions over the life cycle. Graph B shows the change of mean share of wealth in stocks from the baseline model when the volatility of unobserved expected stock returns is low.

Graph A. Comparison of Mean Consumption and Wealth



Graph B. Comparison of Mean Share of Wealth in Stocks)

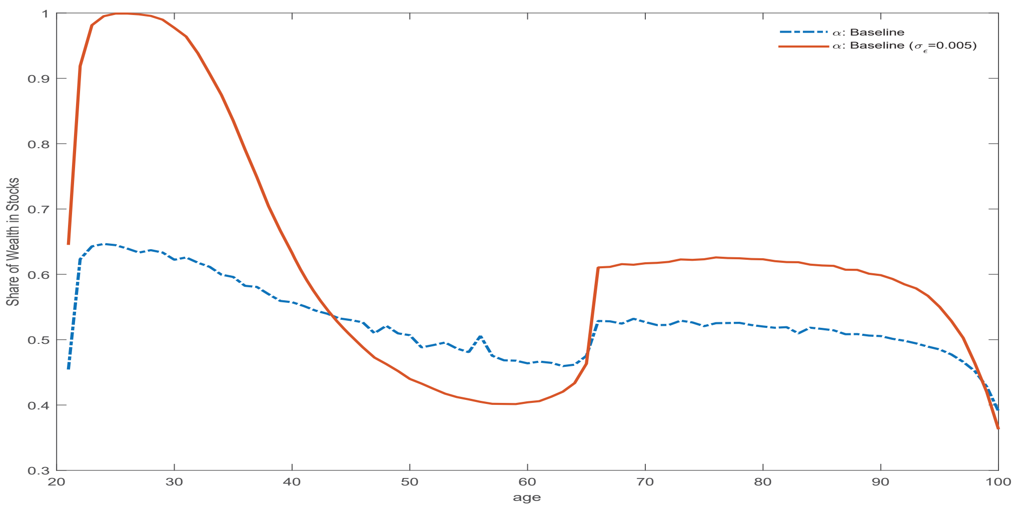
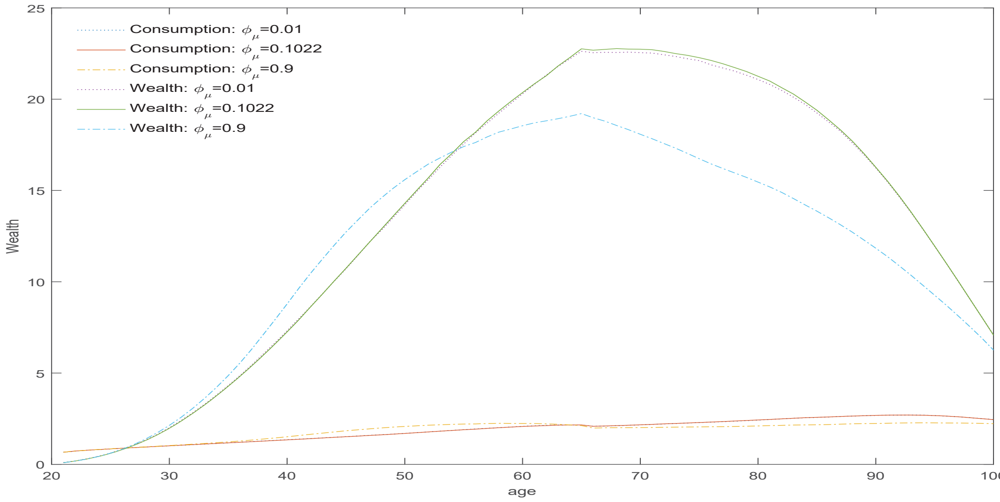


FIGURE 8

The Effect of High Persistence of Unobservable Expected Stock Return (ϕ_μ) on the Life Cycle Profiles

Figure 8 presents how the persistence of unobservable expected stock returns (ϕ_μ) affects mean wealth accumulation, consumption and portfolio choice. Graph A shows the mean wealth accumulation and consumption by varying ϕ_μ from 0.01 to 0.9. Graph B shows the mean share of wealth in stocks due to changing ϕ_μ.

Graph A. Effect of ϕ_μ on Mean Wealth and Consumption



Graph B. Effect of ϕ_μ on Mean Share of Wealth in Stocks)

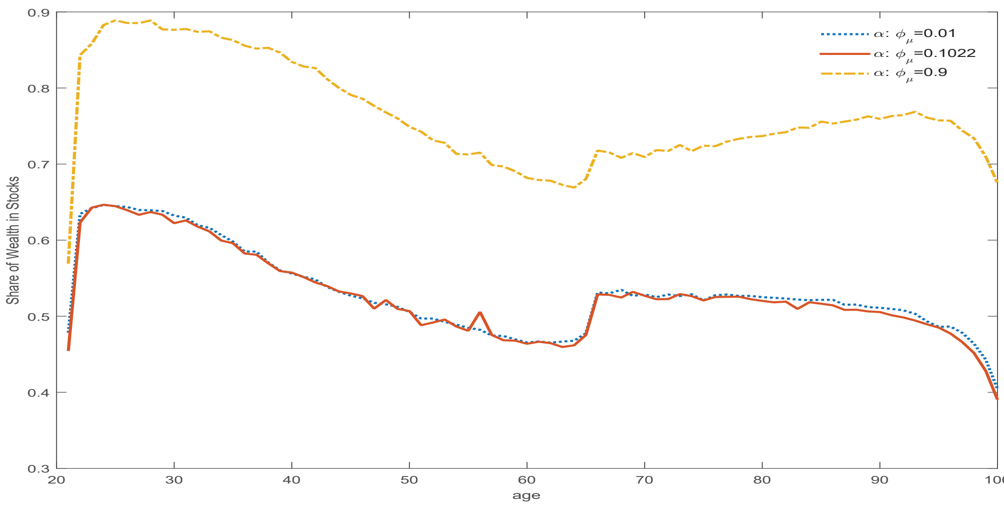
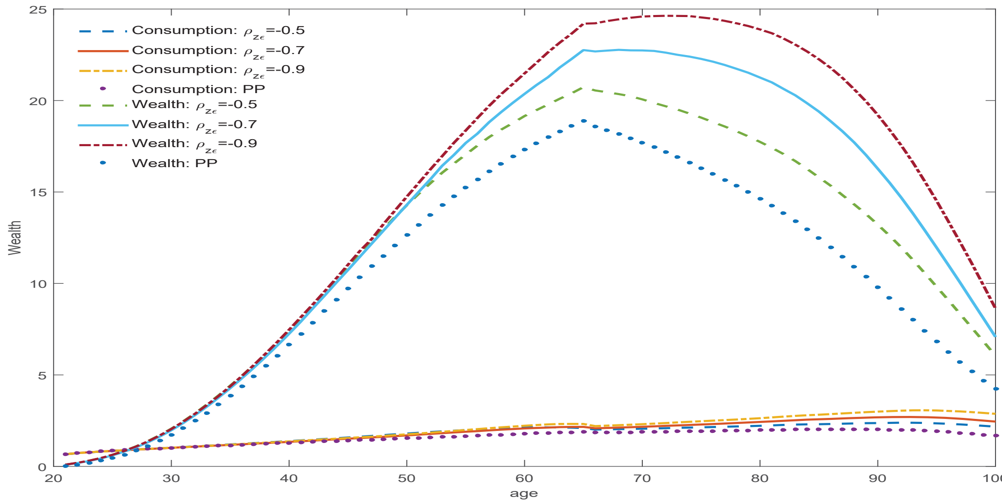


FIGURE 9

The Effect of Correlation between the Shocks to Stock Returns and Innovations of Unobservable Expected Stock Returns($\rho_{z\varepsilon}$)

Figure 9, Graph A presents the effect of correlation between the shocks to stock returns and the innovations of unobserved expected stock returns ($\rho_{z\varepsilon}$) on the mean wealth accumulation and consumption, and compares that with the perfect predictor model (PP model). Graph B shows its effect on mean share of wealth in stocks.

Graph A. Effect of ϕ_μ on Mean Wealth and Consumption



Graph B. Effect of ϕ_μ on Mean Share of Wealth in Stocks)

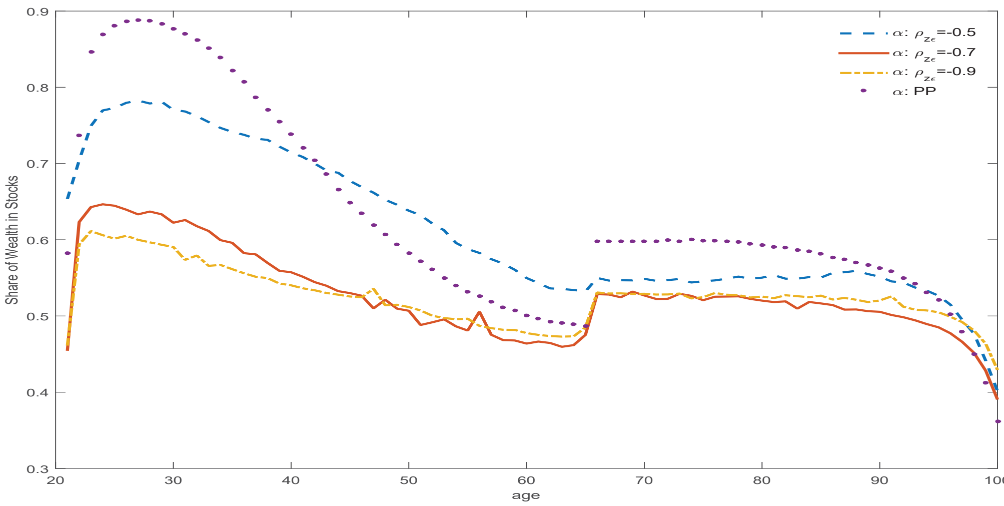
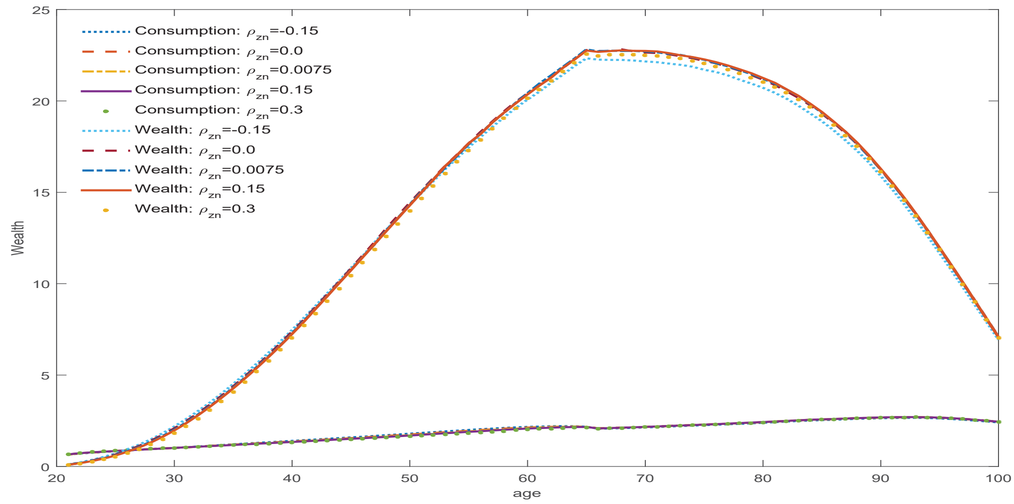


FIGURE 10

The Effect of Correlation between the Shocks to Stock Returns and the Permanent Part of Labor Income (ρ_{zn})

Figure 10, Graph A presents the effect of correlation between the shocks to stock returns and the permanent part of labor income (ρ_{zn}) on the mean wealth and consumption. Graph B depicts the change in the mean share of wealth in stocks due to the variation of ρ_{zn}.

Graph A. Effect of ρ_{zn} on Mean Consumption and Wealth



Graph B. Effect of ρ_{zn} on Mean Share of Wealth in Stocks)

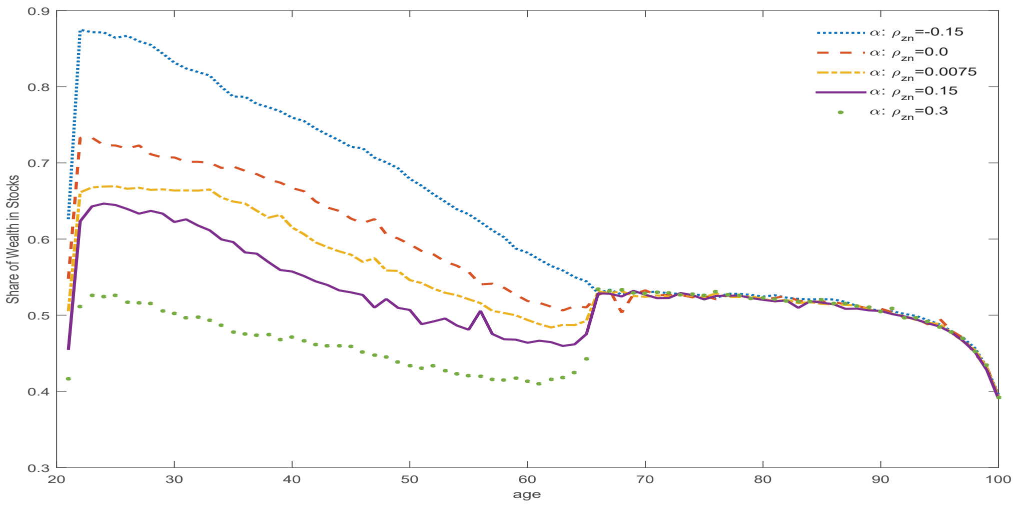
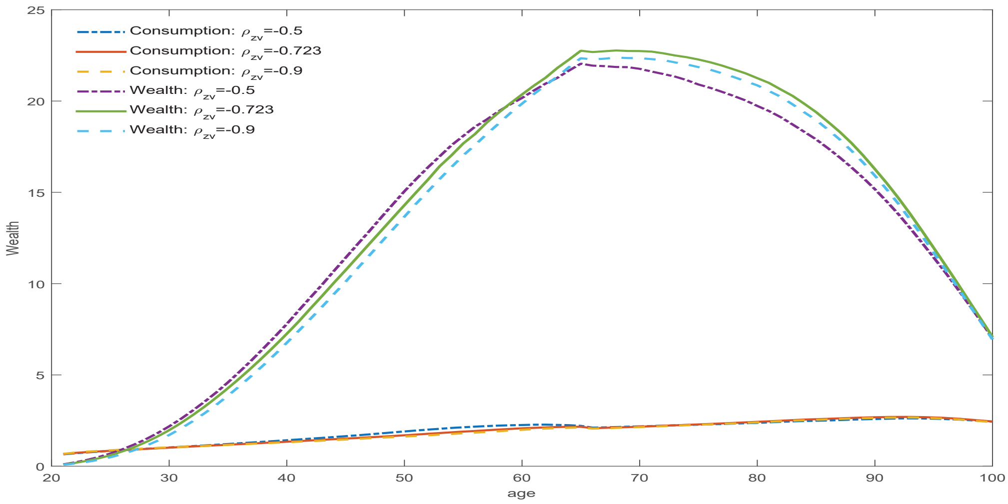


FIGURE 11

The Effect of Correlation between the Shocks to Stock Returns and the Innovations of Predictor (ρ_{zv})

Figure 11, Graph A presents the effect of correlation between the shocks to stock returns and the innovations of predictor (ρ_{zv}) on the mean wealth accumulation and consumption. Graph B exhibits its effect on the mean share of wealth in stocks.

Graph A. Effect of ρ_{zv} on Mean Wealth and Consumption



Graph B. Effect of ρ_{zv} on Mean Share of Wealth in Stocks)

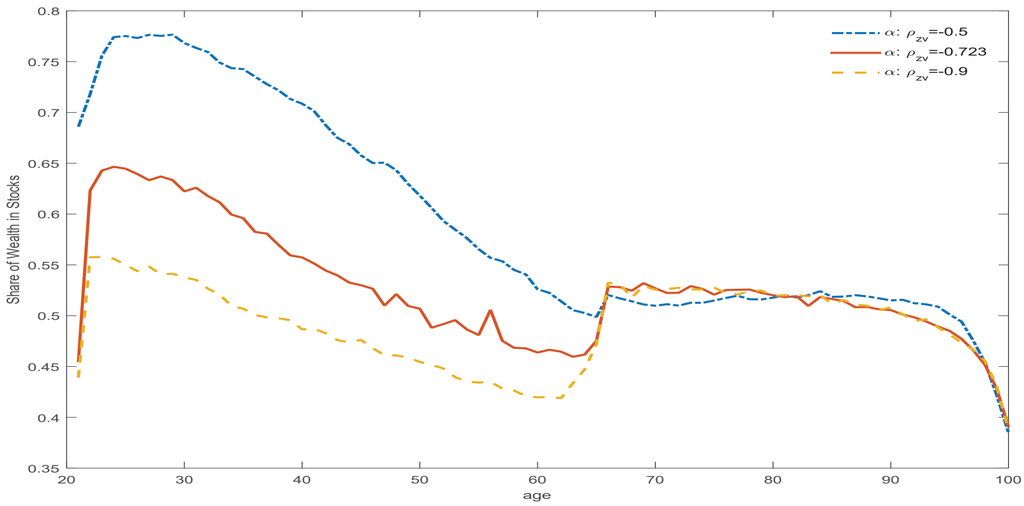
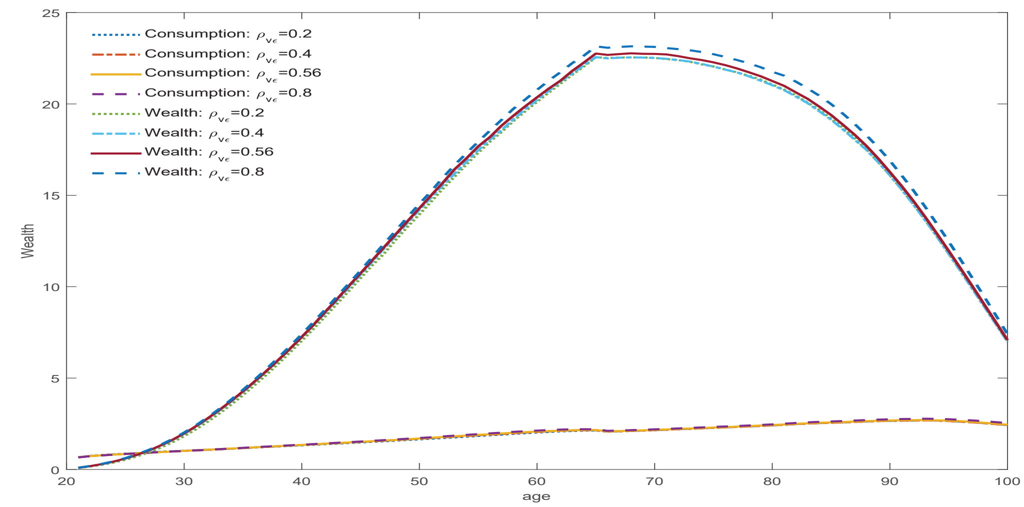


FIGURE 12

The Effect of Correlation between the Innovations of Predictor and the Shocks to the Unobservable Expected Stock Returns ($\rho_{v\varepsilon}$)

Figure 12, Graph A presents the effect of correlation between the innovations of predictor and the shocks to the unobservable expected stock returns ($\rho_{v\varepsilon}$) on the mean wealth accumulation and consumption. Graph B describes the change of the mean share of wealth in stocks by varying $\rho_{v\varepsilon}$.

Graph A. Effect of $\rho_{v\epsilon}$ on Mean Wealth and Consumption



Graph B. Effect of $\rho_{v\epsilon}$ on Mean Share of Wealth in Stocks)

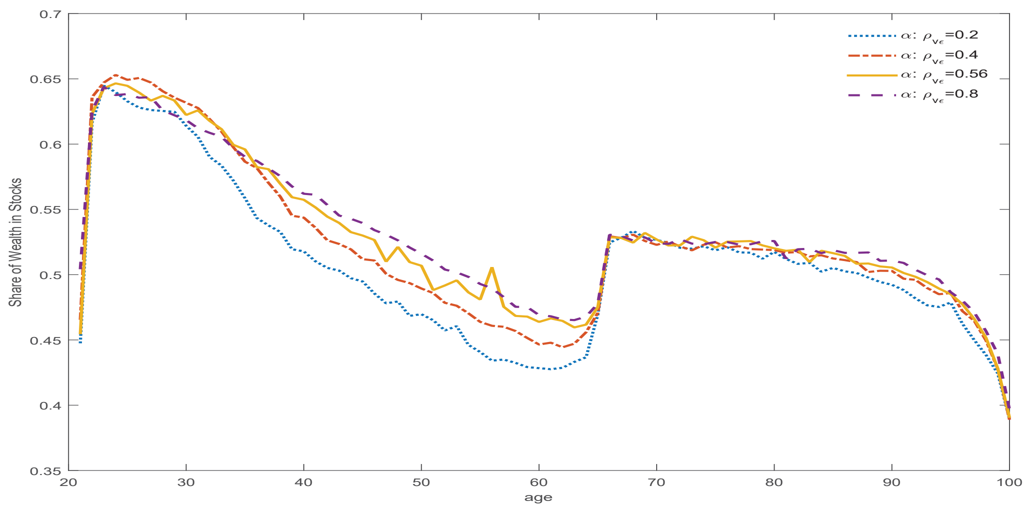
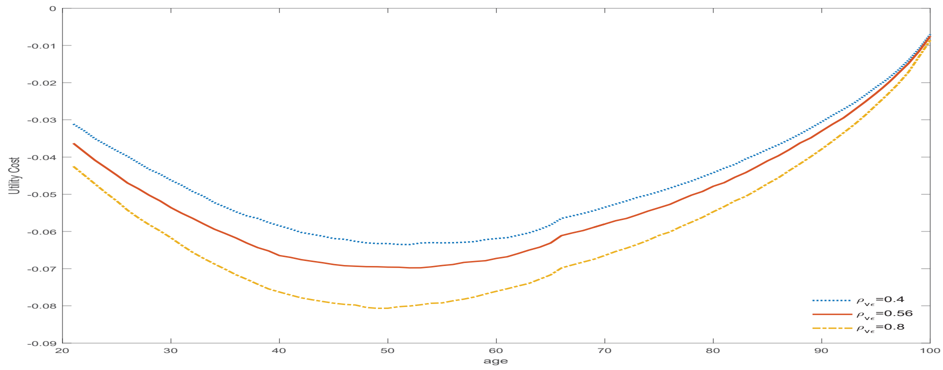


FIGURE 1

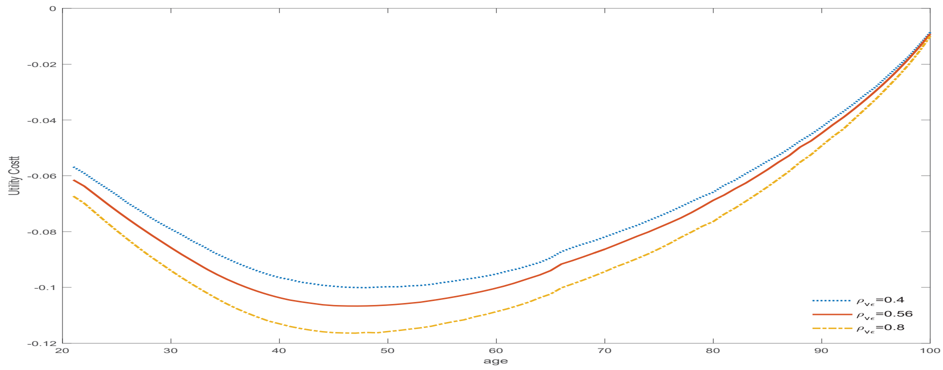
Welfare Evaluation

Figure 13 presents the change of consumption certainty equivalent of the baseline model with respect to the perfect predictor model (PP model), i.i.d. stock returns model and the Vanguard TDF model (Vanguard model) when changing the correlation between the shocks to the unobservable expected stock returns and the innovations of predictor ($\rho_{v\varepsilon}$). Graph A shows the welfare loss from the perfect predictor model when $\rho_{v\epsilon}$ varies. Graph B plots the welfare loss from the i.i.d. stock returns model when changing $\rho_{v\epsilon}$. Graph C gives the welfare loss from the Vanguard TDF model for different $\rho_{v\epsilon}$.

Graph A. Mean Consumption Certainty Equivalent of Baseline w.r.t. PP



Graph B. Mean Consumption Certainty Equivalent of Baseline w.r.t. i.i.d.



Graph C. Mean Consumption Certainty Equivalent of Baseline w.r.t. Vanguard

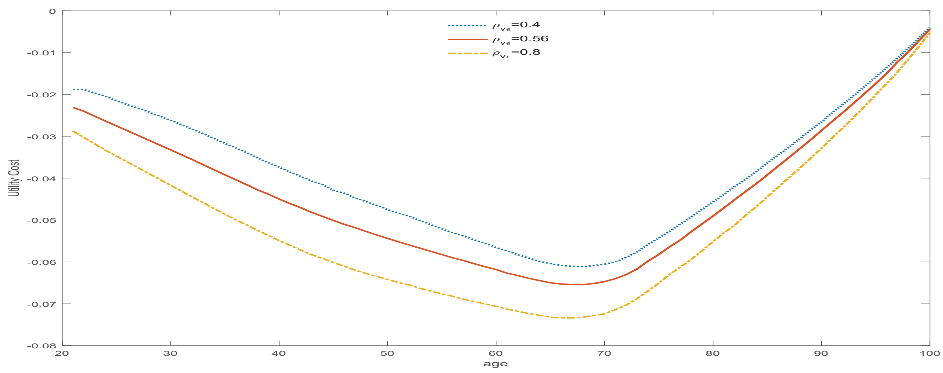
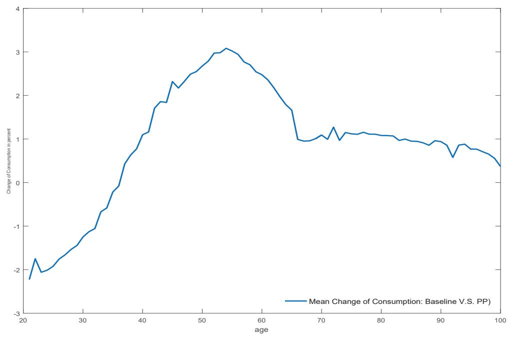


FIGURE 14

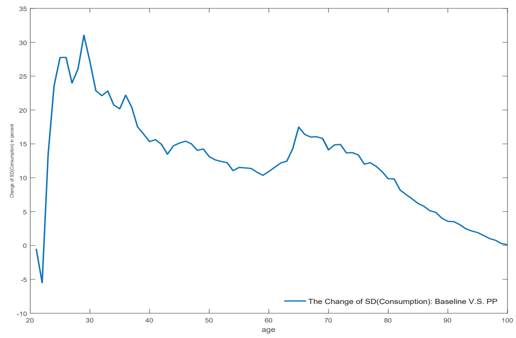
Consumption Evaluation

Figure 14 presents the average change of consumption and standard deviation of consumption over the life cycle. Graph A, C and E describe the mean change of consumption in the baseline model compared to the perfect predictor model (Graph A), the i.i.d. stock returns model (Graph C) and the Vanguard TDF model (Graph E). The mean change of consumption is defined as $E_t\left(\frac{C_1 - C_2}{C_2}\right)$. Graph B, D and F show the change of standard deviation of consumption for the corresponding models. The change of standard deviation is defined as $\left(\frac{sd[C_1] - sd[C_2]}{sd[C_2]}\right)$.

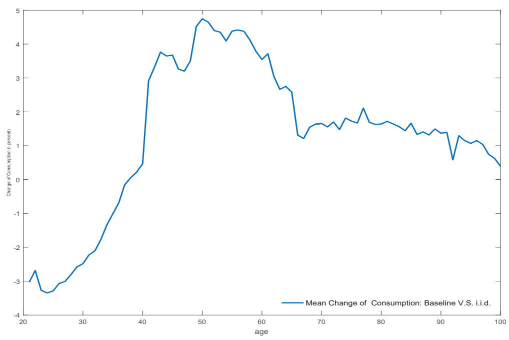
Graph A. Mean Change of Consumption between Baseline and PP



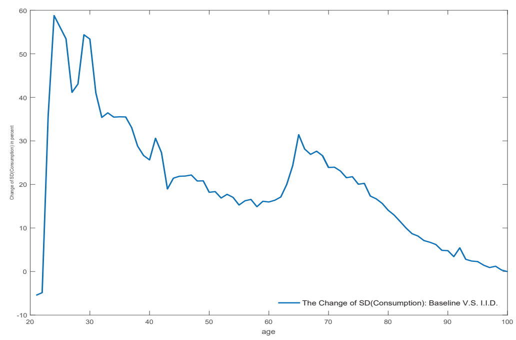
Graph B. The Change of SD (Consumption) of Baseline w.r.t. PP



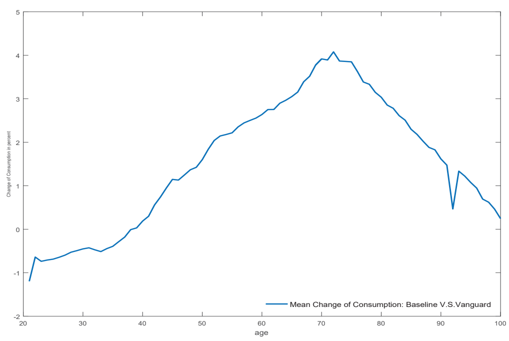
Graph C. Mean Change of Consumption between Baseline and i.i.d.



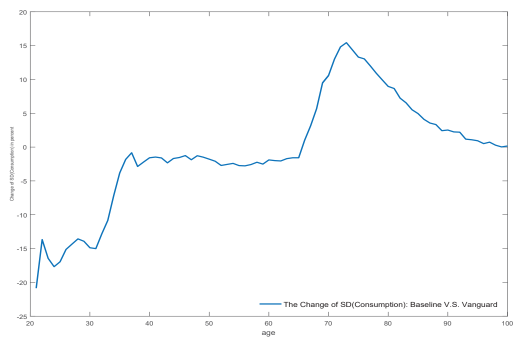
Graph D. The Change of SD(Consumption) between Baseline and i.i.d.



Graph E. Mean Change of Consumption between Baseline and Vanguard



Graph F. The Change of SD(Consumption) between Baseline and Vanguard



IX. Appendix: Tables

TABLE 1

Descriptive Statistics

Table 1 presents descriptive statistics of the annual data from CRSP. The real risk free is defined as the mean of the difference between the 1-Year bond return and annual CPI growth rate. Real adjusted return (r_t) is defined as the difference between the annual value weighted adjusted returns and annual CPI growth rate. SD is the standard deviation.

1946/12/31 –2015/12/31	Mean(%)	SD(%)	Skewness	Kurtosis
1-Year Bond Return	5.07	3.9	1.0	4.23
Annual CPI Growth Rate	3.79	3.4	2.81	7.14
Value Weighted Adjusted Returns	12	17	-0.4	3.02
Value Weighted Ex-Returns	8.28	17	-0.41	2.93
Dividend/Price	3.39	1.5	0.47	3.05
Real Adjusted Return	8.22	18	-0.43	3.04
Real Risk Premium ($r_t - r_f$)	6.93	18	-0.43	3.04
Real Risk Free Rate	1.29	–	–	–

TABLE 2

The Results of Parameter Estimation

Table 2 shows the parameter estimation of the equations (4) - (6). E_r is the unconditional expectation of the risk premium, E_q is the unconditional expectation of the predictor, ϕ_q is the persistence parameter of the predictor, ϕ_μ is the persistence parameter of the unobserved expected stock return process, σ_v is the standard deviation of the predictor's innovations, σ_ω is the standard deviation of the forecast error specified in (16), m and n are the parameters in (16) which are derived from equations (4) - (6), $\rho_{\omega v}$ is the correlation between the innovations of the predictor process and the forecast errors, σ_r is the standard deviation of stock returns, ρ_{rq} is the correlation between the stock returns and the predictors and σ_q is the standard deviation of the predictor.

Parameters	Estimation	Parameters	Estimation
E_q	0.0326	ϕ_q	0.9553
m	-0.2242	σ_r	0.1921
n	5.7292	E_r	0.069
σ_v	0.004968	$\rho_{r,q}$	-0.0932
ϕ_μ	0.1022	σ_ω	0.1813
$\rho_{\omega v}$	-0.6562	σ_q	0.01681

TABLE 3

Summary of Parameter Choice

Table 3 presents the parameter choice used in the baseline model. The σ_z is the standard error of the stock returns, σ_ε is the standard deviation of the shocks to the unobservable expected stock return, σ_v is the standard error of the predictor, σ_n is the standard error of the permanent part of labor income, σ_u is the standard deviation of the transitory component of labor income, E_r is the unconditional expected risk premium, E_q is the unconditional expected dividend yield, γ is the risk aversion, ϕ_q is the persistence parameter of the dividend yield process, ϕ_u is the persistence of the unobservable expected stock returns, r_f is the real risk free rate, $\rho_{z\varepsilon}$ is the correlation between the innovations of stock returns and the shocks to the dividend yield, ρ_{vz} is the correlation between the shocks to the dividend yield and the innovations of stock returns, $\rho_{v\varepsilon}$ is the correlation between the innovations of the dividend yield and the shocks to the unobservable stock returns, ψ is the elasticity of inter-temporal substitution, b is the bequest motive, ρ_{zu} is the correlation between the innovations of stock returns and the transitory component of labor income, ρ_{vn} is the correlation between the innovations of dividend yield process and the shocks to the permanent part of labor income, ρ_{vu} is the correlation between the innovations of dividend yield process and the transitory component of labor income, $\rho_{\varepsilon n}$ is the correlation between the innovations of unobservable expected stock returns and the shocks to the permanent part of labor income. $\rho_{\varepsilon u}$ is the correlation between the innovations of unobservable expected stock returns and the transitory component of labor income, $E\left[ln\left(N_t\right)\right]$ is the expectation of logarithm of the permanent earning shocks to the labor income, $E\left[ln\left(U_t\right)\right]$ is the expectation of logarithm of the transitory earning shocks to the labor income, and β is the discount factor of the utility function.

Parameters	Value	Parameters	Value
σ_z	0.1646	ϕ_q	0.955
ρ_{zu}	0.0	σ_r	0.1921
σ_ε	0.0985	E_r	0.069
σ_v	0.00497	r_f	0.0129
ϕ_μ	0.1022	ρ_{vu}	0.0
σ_n	0.1	σ_u	0.1
$\rho_{z\varepsilon}$	-0.7	$\rho_{\varepsilon n}$	0.0
ρ_{vz}	-0.723	$\rho_{\varepsilon u}$	0.0
$\rho_{v\varepsilon}$	0.56	ρ_{zn}	0.15
$E\left[ln\left(N_t\right)\right]$	-0.005	E_q	0.0326
$E\left[ln\left(U_t\right)\right]$	-0.005	γ	5
ψ	0.5	β	0.96
Trading Cost	0.029	b	2.5

TABLE 4

Conditional Expectation and Standard Deviation due to Variation of $\rho_{z\varepsilon}$

Table 4 presents the conditional moments of the unobservable expected stock return (μ_t) and the next period's stock return (r_{t+1}) with different correlations between the shocks to the stock market and the innovations of unobservable expected stock returns ($\rho_{z\varepsilon}$). $E(\mu_t | r_t, q_t)$ is the conditional expectation of the current unobserved expected stock return based on the current observed return (r_t) and dividend yield (q_t), $\sigma_{\mu_t | r_t, q_t}$ is the conditional standard deviation of the current unobserved expected stock return based on the current observed return (r_t) and dividend yield (q_t), and $\sigma_{r_{t+1} | r_t, q_t}$ is the conditional standard deviation of next period's stock return based on the current observed return (r_t) and dividend yield (q_t).

$\rho_{z\varepsilon}$	$E(\mu_t \| r_t, q_t)$	$\sigma_{\mu_t \| r_t, q_t}$	$\sigma_{r_{t+1} \| r_t, q_t}$
-0.5	$E_r - 0.185\,[r_t - r_f - E_r] + 0.88\,[q_t - E_q]$	0.091	0.188
-0.7(baseline)	$E_r - 0.274\,[r_t - r_f - E_r] + 0.79\,[q_t - E_q]$	0.082	0.184
-0.9	$E_r - 0.363\,[r_t - r_f - E_r] + 0.69\,[q_t - E_q]$	0.068	0.178

TABLE 5

Conditional Expectation and Standard Deviation due to Variation of ρ_{zv}

Table 5 presents the conditional moments of unobservable expected stock return (μ_t) and the next period's stock return (r_{t+1}) with different correlations between the shocks to the stock market and the innovations of dividend yield (ρ_{zv}). $E(\mu_t|r_t, q_t)$ is the conditional expectation of the current unobservable expected stock return based on the current observed return (r_t) and the dividend yield (q_t), $\sigma_{\mu_t|r_t,q_t}$ is the conditional standard deviation of the current unobservable expected stock return based on the current observed return (r_t) and dividend yield (q_t), and $\sigma_{r_{t+1}|r_t,q_t}$ is the conditional standard deviation of the following period's stock return based on the current observed return (r_t) and dividend yield (q_t).

ρ_{zv}	$E(\mu_t|r_t, q_t)$	$\sigma_{\mu_t|r_t,q_t}$	$\sigma_{r_{t+1}|r_t,q_t}$
-0.5	$E_r - 0.277\,[r_t - r_f - E_r] + 0.96\,[q_t - E_q]$	0.082	0.184
-0.723(baseline)	$E_r - 0.274\,[r_t - r_f - E_r] + 0.79\,[q_t - E_q]$	0.082	0.184
-0.9	$E_r - 0.273\,[r_t - r_f - E_r] + 0.65\,[q_t - E_q]$	0.082	0.184

TABLE 6

Conditional Expectation and Standard Deviation due to Variation of $\rho_{v\varepsilon}$

Table 6 presents the conditional moments of the unobservable expected stock return (μ_t) and the next period's stock return (r_{t+1}) for different correlations between the shocks to the unobservable expected stock returns and the innovations of dividend yield ($\rho_{v\varepsilon}$). $E\left(\mu_t|r_t,q_t\right)$ is the conditional expectation of the current unobservable expected stock return based on the current observed return (r_t) and the dividend yield (q_t), $\sigma_{\mu_t|r_t,q_t}$ is the conditional standard deviation of the current unobservable expected stock return based on the current observed stock return (r_t) and the dividend yield (q_t), and $\sigma_{r_{t+1}|r_t,q_t}$ is the conditional standard deviation of next period's return based on the current observed return (r_t) and the dividend yield (q_t).

$\rho_{v\varepsilon}$	$E\left(\mu_t\|r_t,q_t\right)$	$\sigma_{\mu_t\|r_t,q_t}$	$\sigma_{r_{t+1}\|r_t,q_t}$
0.2	$E_r - 0.28\left[r_t - r_f - E_r\right] - 0.10\left[q_t - E_q\right]$	0.0831	0.1844
0.4	$E_r - 0.28\left[r_t - r_f - E_r\right] + 0.39\left[q_t - E_q\right]$	0.0828	0.1843
0.56(baseline)	$E_r - 0.27\left[r_t - r_f - E_r\right] + 0.79\left[q_t - E_q\right]$	0.0820	0.1839
0.8	$E_r - 0.27\left[r_t - r_f - E_r\right] + 1.37\left[q_t - E_q\right]$	0.0798	0.1830

Chapter 3

How Should Stock Market Predictability Affect Target-date Fund Design?

I. Introduction

In analyzing the effects of stock market predictability on life cycle portfolio choice, the choice of the factor predicting stock returns and the frequency at which decisions are made could become important determinants of life cycle saving and portfolio choice. Either of these choices can substantially alter the conventional advice that households should invest a larger proportion of their financial wealth in the stock market when young and gradually reduce the exposure to the stock market as they grow older. This is the standard advice given by financial planning consultants (for instance, Vanguard) who recommend target-date funds (TDFs) that reduce exposure to the stock market as retirement approaches and is also a typical result in the academic literature in the presence of undiversifiable labor income risk (for example, Cocco, Gomes, and Maenhout (2005), Gomes and Michaelides (2005), Polkovnichenko (2007) and Fagereng, Gottlieb and Guiso (2016)).

Nevertheless, the conventional TDF advice rests on a key assumption about stock returns, namely that they are independently and identically distributed (i.i.d.). Substantial evidence exists, however, that certain factors can inform the investor about the future evolution of the stock market. As a result, various papers have analyzed the implications of stock market predictability for consumption and/or portfolio choice while ignoring labor income risk,[1] whereas others focus on the effect of background labor income

[1]Kim and Omberg (1996), Brennan, Schwartz and Lagnado (1997), Brandt (1999), Campbell and Viceira (1999), Balduzzi and Lynch (1999), Barberis (2000), Campbell et. al. (2001), (2003), and Wachter (2002) show that stock market exposure varies substantially as a response to the predictive factor(s).

risk on portfolio choice while ignoring stock market predictability.[2] Michaelides and Zhang (2016) jointly model stock market predictability and non-diversifiable background labor income risk and analyze the normative implications for optimal consumption and portfolio choice over the life cycle using Epstein-Zin (1989) and Weil (1990) preferences.[3] Michaelides and Zhang (2016) make the case for enhanced target-date funds (ETDFs) that condition decisions on expected risk premia: introducing such funds generates improvements in investor welfare.

In both the i.i.d. stock returns or stock market mean reversion literature analyzed above, the investor allocates financial assets strategically (from a

[2]For instance, Heaton and Lucas (2000), Viceira (2001), Haliassos and Michaelides (2003), Gomes and Michaelides (2005), and Cocco, Gomes, and Maenhout (2005).

[3]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 (2010) 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.

longer run perspective, see Campbell and Viceira, 2002). In particular, this means that most models are solved at an annual frequency: the investor is assumed to make decisions once a year and stick with them for the rest of the year.

There is no a priori evidence that this modeling choice is correct. Rather, it seems to be dictated by computational constraints and the fact that the empirical evidence on most inputs in the model (the process for labor income risk, for example) has been determined at the annual frequency. Nevertheless, many factors that have been used to predict the stock market are available at the quarterly frequency, implying that a quarterly life cycle model is needed to analyze the implications for predictability for portfolio choice at the quarterly frequency. Moreover, investigating differences between the quarterly and annual versions of the same model could be interesting in its own right (that is, even with i.i.d. stock returns). This will be our first innovation in this paper relative to the rest of the literature: solve a life cycle model at a quarterly rather than annual frequency.

Once this can be achieved, then we can use recently discovered factors in predicting stock returns that have not been previously used in the life cycle portfolio choice literature, despite being robust predictors of expected stock returns. We focus on two particular factors. The first is the variance risk premium (VRP) to predict expected stock returns, a factor that has been found to be a robust predictor both using US data (Bollerslev, Tauchen and Zhou (2009)) and international data (Bollerslev, Marrone, Xu, and Zhou (2014)). The second is CAY, Lettau and Ludvigson (2001), found to be the most robust predictor in a recent paper in RFS (2015, Magdalinos et al). A drawback of the predictive model is that quarterly monitoring the predictor has a higher cost than i.i.d. model or Vanguard TDFs so that we have to consider the effect of transaction cost. But, the risky asset in our study is the market index. The transaction cost is therefore only several bps. In this paper, we report the results with the extra trading cost of 10bp or 50bp.

We find that tactical asset allocation (using a model with stock market predictability than a model without one) can increase life cycle wealth accumulation and the effect is larger for models that have a more reliable predictability. Such market timing also affects the total share of wealth in stocks. Large variations over the life cycle depend 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.

It is perhaps important to note that VRP is not a persistent factor predicting stock returns thereby differentiating its importance from previously used factors (like the dividend yield, for example). Using this factor to predict stock returns we find that the asset allocation profiles retain the target date funds (TDF) feature of slowly decreasing stock market exposure as the household ages, on average. 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.[4] Similar conclusions qualitatively are generated when CAY is used to predict the stock market. Taken together, these findings make the case for enhanced TDFs (ETDFs) that condition on the market timing ability of the investor.

We also present different comparative statics adjusting both the bequest parameter and also other parameters determining hedging demands (different correlations between the different innovations in the stock market and labor income generating processes). The bequest motive affects primarily the speed with which wealth is decumulated during retirement and that affects the average allocation to the stock market during retirement, as found in the i.i.d. stock returns literature. The main correlation that affects hedging demands is

[4]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.

the correlation between the permanent labor income shock and the stock return innovation, as in the i.i.d. stock returns literature. Nevertheless, the correlation between the stock market innovation and the shock to the factor predicting stock returns affects the average share of wealth in stocks when CAY is used as a predictor. A more negative correlation reduces the average share of wealth in stocks during working life.

What is the quantitative impact on investor welfare if the data generating process for stock returns follows any of the predictors but instead investors are making decisions based on the i.i.d. stock returns model or based on the Vanguard TDF recommendations? We show that the VRP model generates improved certainty equivalent welfare relative to either the CAY or the i.i.d. stock returns model. This improvement is still significant even though the trading cost is included. We therefore conclude that basing decisions on the factor realizations can improve investor welfare.

The paper is organized as follows. Section II describes the empirical analysis and reproduces two popular predictors in the literature (CAY and VRP). Section III outlines the theoretical life-cycle model, outlines the numerical solution algorithm and discusses the parameter choices for the calibration. Section IV discusses the effects of stock market mean reversion by comparing the benchmark results to the i.i.d. stock returns model. Section IV discusses hedging demands and how different correlation changes also affect wealth accumulation, while Section V briefly discusses the implications of model parameter uncertainty. Section VI discusses the implications of the model for lifestyle funds and Section VII concludes.

II. Empirical Analysis

A. Data

The stock market data come from the Center for Research in Securities Prices (CRSP). We use the quarterly bond return, CPI growth rate, monthly cumulative and ex-dividend returns of the US S&P 500 index from January 1st, 1990 to September 30th, 2016. The quarterly cumulative and ex-dividend return are constructed from the monthly return of the US S&P 500 index.

The quarterly volatility index is from Federal Reserve Bank of St. Louis, and the quarterly consumption wealth ratio (CAY) as defined in Lettau and Ludvigson (2001) is downloaded from Lettau and Ludvigson's Web site. We construct the quarterly realized volatility based on the monthly realized volatility defined in Bollerslev et al. (2009):

$$RV_t \equiv \sum_{j=1}^{n} \left[p_{t-1+\frac{j}{n}} - p_{t-1+\frac{j-1}{n}(\Delta)} \right]^2,$$

where RV_t is the realized variation over the discrete $t-1$ to t time interval and p_t is the natural logarithmic price of the asset.

Table 1 contains the descriptive statistics from the data set. The risk free rate has a 0.16% real quarterly return with a very low standard deviation and we will therefore be assuming it to be constant in what follows. The real stock quarterly return has a mean of 1.98% with a standard deviation equal to 7.84%. CAY is very smooth (s.d. equal to 1.98%) and has a persistent (AR(1)) coefficient equal to 0.88. IV and RV are both individually persistent but when combined have very weak serial correlation.

B. Empirical Results

Table 2 contains the predictive regression results that use the predictors of CAY (Lettau and Ludvigson, 2001) and VRP (variance risk premium) as in Bollerslev et al. (2009). The results in Table 2 reveal a fact that VRP has a higher degree of predictability (adjusted R^2 of 15%) comparing with CAY which has an adjusted R^2 of 5.5%.

To illustrate the data, Figure 1 reproduces the Figure 2 in Bollerslev et al. (2009). Our plots are consistent with the results in Bollerslev et al. (2009) and Lettau and Ludvigson (2001). Figure 1 Graphs A and B show how implied and realized variance, respectively, rise in a recession and Graph C shows their difference, which can be used as a factor to predict the stock market. Graph D in Figure 1 shows how the consumption to wealth ratio rises in recessions, as discussed in Lettau and Ludvigson (2001).

C. Conclusion

There are good predictors of future returns and can use as input in a quarterly frequency model. Some predictors more persistent than others and correlations between factor and stock market innovations can be different.

III. The Model

Time is discrete, but contrary to previous literature we solve the model at a quarterly than annual frequency. This is an important deviation from the previous literature that solves similar models at an annual frequency. There is no a priori reason to solve the model at an annual frequency other than the availability of annual labor income income (typically from the *Panel Study of Income Dynamics* (PSID hereafter)). Given that predictability is posited at a quarterly frequency we solve both the i.i.d. and model with predictability at a quarterly frequency.

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 $\tilde{R}$. 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 $X_t = S_{t-1}\tilde{R}_t + B_{t-1}R_{\mathrm{f}} + Y_t$. 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:

A. 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

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

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}.

B. Labor Income Process

Following a relatively standard specification in the literature (eg Cocco et al. (2005)), the labor income process before retirement is given by

$$Y_{it} = Y_{it}^p U_{it}, \tag{3.1}$$

$$Y_{it}^p = \exp(g(t, Z_{it}))Y_{it-1}^p N_{it}, \tag{3.2}$$

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 $\{-0.5 \times \sigma_u^2, -0.5 \times \sigma_n^2\}$, 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. We will assume that the quarterly data generating process for labor income is the same as the one at the annual frequency. This will allow us to pick the variances of the labor income shocks in a specific way to make them consistent with their estimated annual equivalents: essentially, permanent variances will cumulate from quarterly to annual while transitory variances will stay the same. The calibration section discusses this issue in greater detail.

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 the PSID annual data. When moving to quarterly data we linearly interpolate across years to find the relevant quarterly expenditure. We choose not to model explicitly the returns from housing following the empirical evidence (e.g., 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 (2015)). 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).

C. *Liquidity Constraints*

In the presence of a positive expected equity premium households in the model will counterfactually lever up to invest in the stock market in the absence of any borrowing constraints and vice versa when the VRP predicts lower expected stock returns in the future. To prevent such counterfactually levered and volatile household portfolios (that are rarely advised by financial consultants), it is common to assume no borrowing and no short sales of stocks: $B_t \geqq 0$ and $S_t \geqq 0$. Equivalently, the share of wealth in stocks (α_t) is defined as $S_t/(S_t + B_t)$ and constrained between zero and one.

D. *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 construct a VAR:

$$r_{t+1} - r_f = \alpha + \beta f_t + z_{t+1}, \tag{3.3}$$

$$f_{t+1} = \mu + \phi(f_t - \mu) + \varepsilon_{t+1}, \tag{3.4}$$

where the two innovations $\{z_{t+1}, \varepsilon_{t+1}\}$ are i.i.d. Normal 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.

E. Numerical Solution

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

$$x_{it+1} = \frac{Y_{it}^p}{Y_{it+1}^p}(r_{t+1}\alpha_{it} + r_f(1 - \alpha_{it}))x_{it} + U_{it+1}. \tag{3.5}$$

We use Tauchen and Hussey (1991) approximation procedure for a vector autoregression model. 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 replicate key moments: the persistence, the conditional variance and the conditional (one-period ahead) equity risk premium.

F. Parameter Choice

Consumers/investors start working life at age 21. Typically, in annual frequency models, investors die for sure by age 100, therefore living for a maximum of 81 years. In the quarterly analog we construct, there are therefore 324 periods.

Using an annual frequency model, Cocco et al. (2005) estimate the variances of the idiosyncratic shocks using data from the PSID, and the literature uses values close to those: 0.1 for σ_u and 0.1 for σ_n. We assume that the quarterly frequency model is identical to the annual frequency model. It can then be shown that the transitory variance (σ_u^2) remains the same as in the annual model and the permanent variance (σ_n^2) should be divided by four. The deterministic component of labor income is also identical to the one used by most life cycle papers in this literature (Cocco, Gomes, and Maenhout (2005)): the only additional complication is that we linearly interpolate in between years to derive the quarterly counterpart. The replacement ratio during retirement (68% 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 not affected by the move from annual to quarterly frequency.

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 inter-temporal substitution (ψ) equal to 0.5. These choices are consistent with the empirical estimates for the elasticity of inter-temporal substitution in Vissing-Jorgensen (2002) and the empirical preference parameter estimates in Gomes, Michaelides, and Polkovnichenko (2009) and Cooper and Zhu (2016). The bequest parameter is set to 0.0. We set the discount factor (β) equal to 0.99.

The net constant real interest rate, r_f, equals 0.16%. The estimates from Table 1 show that the mean unconditional stock return in this period is around 2% quarterly. In most of the life cycle literature, the equity premium used is around 4% at an annual frequency to reflect transaction costs. To make comparisons across models more meaningful, we subtract a constant transaction cost in the excess return each period to make the mean unconditional return equal to 1% per quarter.

To calibrate the stock market predictability parameters we use two of the most popular predictors of stock returns, namely VRP and CAY, both in quarterly frequency. CAY generates a very persistent factor predicting the real log return on the U.S. equity market but the VRP predictor is not very persistent. The unconditional stock market volatility is given by the unconditional standard deviation of stock returns and is set equal to 0.08.

A key parameter turns out to be the correlation between the factor and the return innovation ($\rho_{z,\varepsilon}$). 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 estimate it at around therefore use $\rho_{z,\varepsilon} = -0.5$ for the CAY model and -0.04 for the VRP model. We provide experiments about this parameter in what follows.

The factor innovations of CAY and VRP model are very smooth and we estimate (and use) $\sigma_\varepsilon = 0.008$ for the CAY baseline model and $\sigma_\varepsilon = 0.0074$ for the VRP baseline model. Given these estimates, we can infer that the unconditional variance of the stock market return innovation equals $\sigma_z^2 = 0.08^2 - \sigma_f^2$.

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). 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. (2001) 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.

Finally, we discuss the transaction cost. Quarterly model has a higher re-balancing frequency so that the cost of monitoring factors is much higher than the annual model. Since the risky asset considered in this study is the market index, the trading cost is really low. For example, the expense of

Vanguard 500(VFINX), Fidelity 500 Index fund(FUSEX) and Schwab S&P 500 Index(SWPPX) are only 12bp, 9bp and 3bp respectively. Therefore, the transaction cost of 10bp or 50bp is enough.

IV. Comparing CAY and VRP with i.i.d. Model

How does the presence of a factor predicting returns affect saving and portfolio choice behavior relative to the i.i.d. and Vanguard models?

Figure 2 illustrates the differences in simulated profiles between the i.i.d., Vanguard and the mean reversion models (CAY and VRP).

Average wealth accumulation[5] varies across all three models (Figure 2, Graph B). Higher wealth accumulation generates higher mean consumption: Figure 2, Graph A. This allows "better" tactical asset allocation decisions given the assumption of no model misspecification during the simulations.

The average share of wealth in stocks is plotted in Figure 2, Graph C. The fact that the average share of wealth in stocks is never one or close to one

[5] The model is simulated 2000 times for 500 individual life histories.

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 realizations of the factor and then averaging over them. For most of these factor states, the investor deviates from the 100% asset allocation to stocks and when taking the average we have the optimal share of wealth in stocks substantially below one.

Figure 3 repeats this experiment but for an extra transaction cost specification. The transaction cost is 10bp and 50bp. The qualitative nature of the comparative statics discussed earlier does not change (higher wealth accumulation and consumption for the VRP model relative to either the Vanguard or i.i.d. models and more balanced portfolios for the VRP models). Even though the higher transaction cost the lower wealth accumulation, consumption and stock allocation, the extra transaction cost of 10bp or 50bp is not able to change the main results.

V. Hedging Demands

How do these results change when the correlations between the different innovations vary? We next use the model to quantitatively assess the magnitude of such hedging demands.

A. Variation in Correlations

To investigate the importance of hedging demand due to $\rho_{z,\varepsilon}$ (the correlation between the factor and the stock market innovation), we change it to -0.1 and -0.3 from the VRP baseline of -0.05 and -0.1 and -0.8 from the CAY baseline of -0.5.

We also evaluate hedging demands when changing the correlation between permanent earnings shocks and stock market innovations ($\rho_{z,n}$). In our baseline models (VRP and CAY) 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. Bonaparte, Korniotis, and Kumar (2014) find that this correlation can vary for different households from -0.6 to 0.6 and we therefore use 0.3 to investigate how the results change.

The life profiles (consumption, wealth accumulation and mean shares of wealth in stocks) over the life cycle for different $\rho_{z,\varepsilon}$ are depicted in Figure 4 (VRP) and 6 (CAY), respectively.[6] In the VRP model, when the correlation between the factor and stock return innovations decreases to -0.8 ($\rho_{z,\varepsilon} = -0.8$) from $\rho_{z,\varepsilon} = -0.05$, the effect on the consumption, wealth accumulation and share of wealth in stocks is not obvious. When using CAY as the predictor, we find that the effects on the share of wealth in stocks become larger than that of VRP case.

[6] 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 since portfolios in the i.i.d. model do not exhibit any time variation in response to the factor realizations. For a specific correlation (ρ) we can compute these hedging demands by comparing the two simulated profiles and compute the percentage differences: $hedg(\rho) = 100 * (\alpha_{\text{factor}}(\rho) - \alpha_{\text{I.I.D.}}(\rho)) \;/\; \alpha_{\text{I.I.D.}}(\rho)$. For space considerations we only report the simulated shares of wealth in stocks generated from different correlations of interest.

What happens when the correlation between the permanent earnings shock and the stock market innovation is raised from 0.15 to 0.3? Figure 5 and 7 plot the average life profiles over the life cycle for the mean reversion models versus the i.i.d. model when this correlation increases to 0.3. We find that this correlation has a significant effect on the asset allocation only (Graph C). An increasing in $\rho_{z,n}$ leads to an decrease in the share of wealth in stocks over the working life. This is because the labor income looks more like the risky asset rather than the riskless asset when $\rho_{z,n}$ increases, which squeezes the share of wealth in stocks. This conclusion is consistent with the findings in Cocco et. al. (2005) and Michaelides and Zhang (2016).

Changing the correlation between the factor innovation and the permanent income shock ($\rho_{n,\varepsilon}$) does not materially affect the average share of wealth in stocks (Figure 6 and 9 for the two cases with different bequest motives), and therefore also does not substantially affect average wealth accumulation.

VI. Are Lifestyle funds Optimal?

A. 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 retiring in 2008 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 and assume that the stock market can be predicted by the VRP model throughout the life cycle.

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 $v_n\left(x_{it}, f_t\right)$ be the value function for a new model. We compute a measure of welfare change for a particular age group (age) as:

$$\overline{\mu}_{\text{age}} = \text{average of } \left[\left(\frac{v_n(x_{it}, f_t)}{v_0(x_{it}, f_t)}\right)^{\frac{1}{1-\gamma}} - 1\right] \text{ for all } i \in I_{\text{age}} \text{ and all factor states} \tag{3.6}$$

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.

When comparing the VRP with the CAY or i.i.d. models, we simulate wealth profiles based on the optimal solution when using the VRP factor. This produces a series for cash on hand, and we use this same cash on hand to evaluate how the investor would have behaved under the CAY or i.i.d. models. This allows us to compare welfare keeping cash on hand constant and we can track these differences in welfare throughout the life cycle.

Figure 8 plots the life cycle certainty equivalents when returns are simulated based on the VRP model and the comparison is between the VRP and the two other models: the CAY and the i.i.d. model. Graph A illustrates substantial welfare losses from following the CAY model relative to the optimal portfolio rule in the presence of the VRP factor, and the welfare losses are even more substantial when following the i.i.d. recommendation. This arises naturally given that the deviations of the average portfolio allocations are even larger between the i.i.d. recommendation and the VRP model than the ones between the VRP and CAY models. The transaction cost of 10bp and 50bp does decrease the welfare loss, but it is still significant.

Graph B recalculates the life cycle certainty equivalents when returns are simulated based on the CAY model. There is a maximum of 2% welfare loss comparing with the VRP model and a maximum of 2.5% welfare gain relative to i.i.d model. Similarly, Graph C figures out the life cycle certainty equivalents when returns are simulated based on the i.i.d. model. The i.i.d. model generates a lower welfare loss at the same cash on hand than that using the VRP model as a base. The transaction cost of 10bp or 50bp can decrease the welfare loss but can not change the qualitative pattern.

We make three observations based on the results in Figure 8. First, the welfare losses are economically significant: they can reach 5% of consumption equivalents depending on the specification, and this represents a substantial welfare loss at a quarterly frequency. Second, the VRP model has the largest welfare gain relative to the other models (CAY and i.i.d.). The i.i.d model has the largest welfare loss comparing to VRP and CAY model. The CAY model is in an intermediate position. Third, the transaction cost can decrease the welfare loss, but welfare gain of VRP is still significant after deducting the transaction cost of 10bp or 50bp.

C. Out-of-Sample Analysis

We next compare how realized wealth evolve out of sample, partly adapting the calculations in Lan (2015) for our model. Specifically, we start simulating for every age using the realized VRP, CAY and stock returns between 1996Q1 and 2016Q1. For instance, given a simulated initial distribution of assets per age in 1996, from 1996 onward we use the realized VRP and CAY to pick the relevant factor state and the realized stock return to shock financial asset returns. We save financial wealth for every age group between 1996Q1 and 2016Q1. For an investor at age 30, for example, we track the evolution of individual wealth over the 1996Q1 –2016Q1 period and average in the cross section every quarter.

It might be helpful to start the discussion by plotting the wealth levels for different models. Figure 9 shows how the mean wealth evolves for different starting ages according to the different models. As can be seen, the VRP models typically generate higher mean wealth than either the i.i.d. or CAY model would predict, while the CAY model can not statistically and economically outperforms the i.i.d. model in every age groups. These results emphasize the importance of searching for good predictors when devising long-term asset allocation strategies. Moreover, younger households display the mean wealth accumulation rising over time (and over the life cycle), while older households can display the decreasing wealth after retirement.

VII. 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, and a more volatile factor generates more volatile movements in the asset allocation policy rule. On average, the asset allocation based on the predictive model is much lower than that from the i.i.d. case. Ignoring the stock market predictability can lead to a severe welfare loss. Moreover, the investor should pay more attention to the predictor. A reliable predictor has a positive effect on reducing the welfare loss and better out of sample performance. A unreliable predictor can lead to weak out-of-sample performance comparing with the i.i.d model, even though it can generate a better theoretical results.

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 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 treatment of housing 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.

VIII. Appendix: Tables

TABLE 1

Descriptive Statistics

Table 1 presents descriptive statistics of the quarterly data from the first quarter of 1990 to the third quarter of 2016. The r denotes the accumulate return on the S&P 500 in excess of quarterly CPI rate. IV denotes the quarterly "model free" implied variance or VIX index. RV is the quarterly "model free" realized variance. The other predictor variable is quarterly consumption wealth ratio (CAY). CPI and r_f refers to the consumer price index and the 90-day T-bill rate respectively, and are downloaded from CRSP as well.

Summary Statistics

1990/01/01 –2016/09/30	r	IV	RV	$IV-RV$	CAY	CPI	r_f
Mean(%)	1.98	1.11	0.62	0.49	0.04	0.06	0.16
SD(%)	7.84	0.94	0.98	0.75	1.98	0.08	0.09
Kurtosis	3.24	8.16	54.23	31.83	2.68	9.64	5
Skewness	-0.4	2.25	6.45	-3.24	-0.45	-1.39	0.32
AR(1)	0.0	0.41	0.47	-0.17	0.88	0.0011	0.09

Correlation Matrix

1990/01/01 –2016/09/30	r	IV	RV	$IV-RV$	CAY	CPI	r_f
r	1.00	-0.52	-0.42	-0.096	-0.092	-0.11	0.096
IV	–	1.00	0.7	0.34	0.26	-0.18	0.12
RV	–	–	1.00	-0.43	0.095	-0.46	0.3
$IV-RV$	–	–	–	1.00	0.2	0.38	-0.24
CAY	–	–	–	–	1.00	0.058	0.3
CPI	–	–	–	–	–	1.00	-0.76
r_f	–	–	–	–	–	–	1.00

TABLE 2

Predictive Regressions

Table 2 presents predictive regressions based on the quarterly data from the first quarter of 1990 to the third quarter of 2016. The parameters related to the predictive regression using VRP as a predictor are estimated from the following restricted VAR:

$$\begin{bmatrix} VRP_{t+1} \\ r_{t+1} - r_f \end{bmatrix} = \begin{bmatrix} Const \\ \alpha \end{bmatrix} + \begin{bmatrix} \phi & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} VRP_t \\ r_t - r_f \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \end{bmatrix}$$

The parameters related to the predictive regression using CAY as a predictor are estimated from the following restricted VAR:

$$\begin{bmatrix} CAY_{t+1} - E(CAY) \\ r_{t+1} - r_f \end{bmatrix} = \begin{bmatrix} 0 \\ \alpha' \end{bmatrix} + \begin{bmatrix} \phi & 0 \\ \beta & 0 \end{bmatrix} \begin{bmatrix} CAY_t - E(CAY) \\ r_t - r_f \end{bmatrix} + \begin{bmatrix} \varepsilon_{t+1} \\ z_{t+1} \end{bmatrix}$$

Hence, the constant term in the CAY AR(1) model is $E(CAY) * (1 - \phi)$ and α in the predictive regression is $\alpha' - \beta * E(CAY)$.

1990/01/01 –2016/09/30	VRP	CAY
Constant	0.0058 (6.72)	0.00003 (0.23)
α	0.0	0.018 (2.45)
β	3.6 (4.48)	0.6 (1.4)
ϕ	-0.18 (-1.84)	0.93 (22.66)
$\rho_{z,\varepsilon}$	-0.04	-0.51
σ_ε	0.0074	0.008
σ_z	0.0746	0.078
σ_r	0.079	0.079
Adj. R^2 (%)	15	5.5

IX. Appendix: Figures

FIGURE 1

Implied and Realized Variance, Variance Risk Premium and Consumption Wealth Ratio

Figure 1 plots the implied variance (Graph A), the realized variance (Graph B) and variance risk premium (Graph C) and consumption wealth ration (Graph D) for the S&P 500 index from the first quarter of 1990 to the third quarter of 2016. The shaded areas represent NBER recessions.

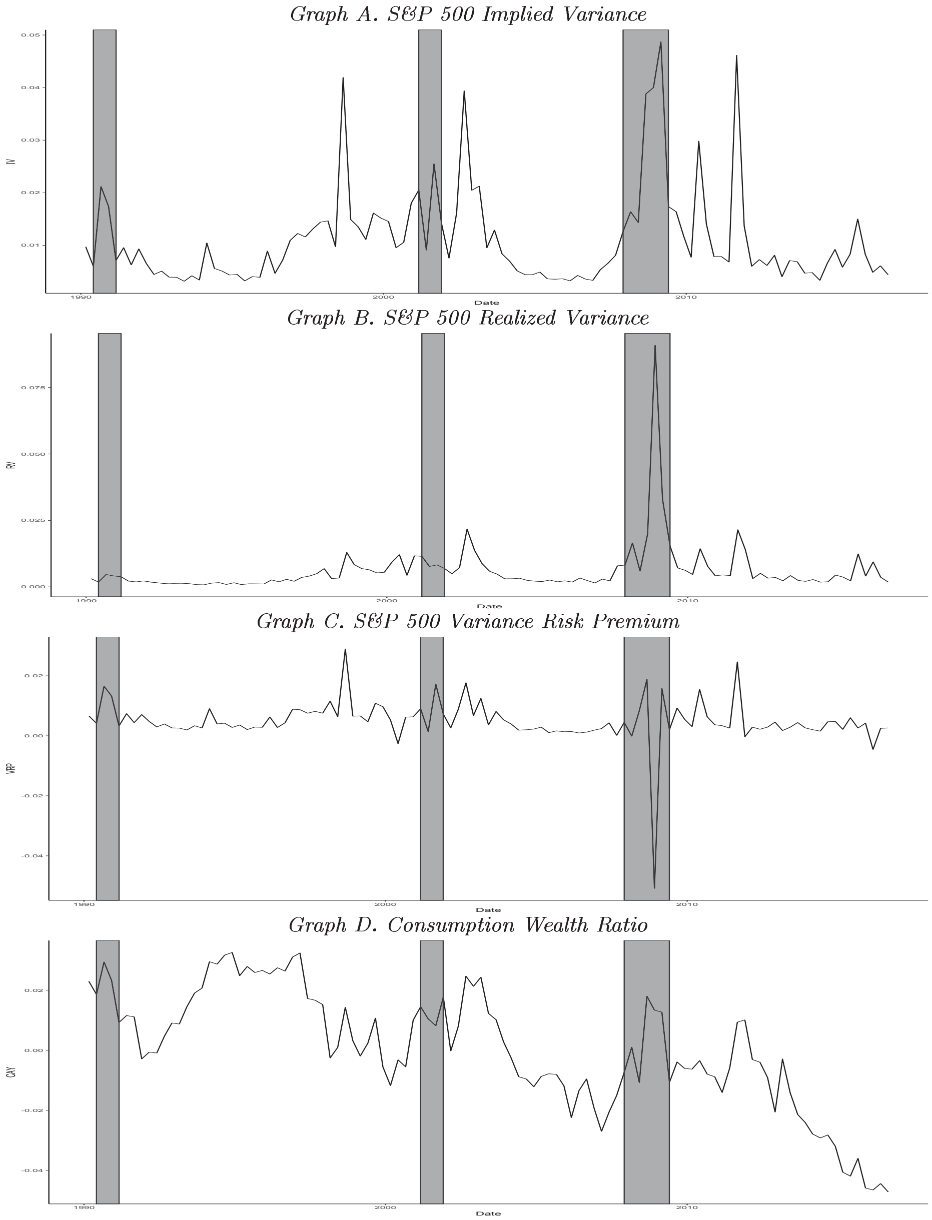
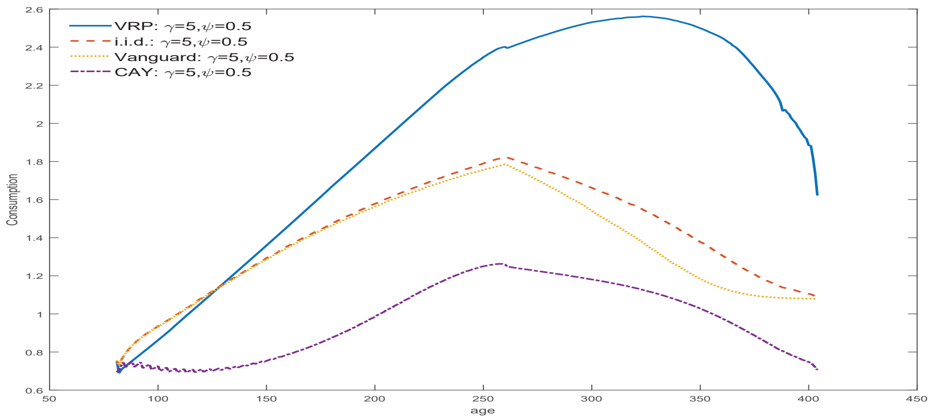


FIGURE 2

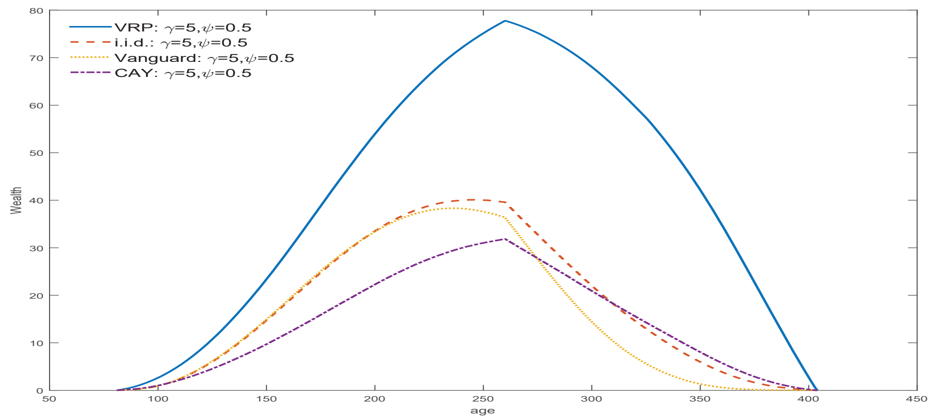
The Life-Cycle Profiles (Bequest = 0.0)

Figure 2 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stocks (Graph C) between the mean reversion models using VRP and CAY as predictor and i.i.d. case while keeping bequest motive at 0.0.

Graph A. Consumption



Graph B. Wealth



Graph C. Share of Wealth in Stock

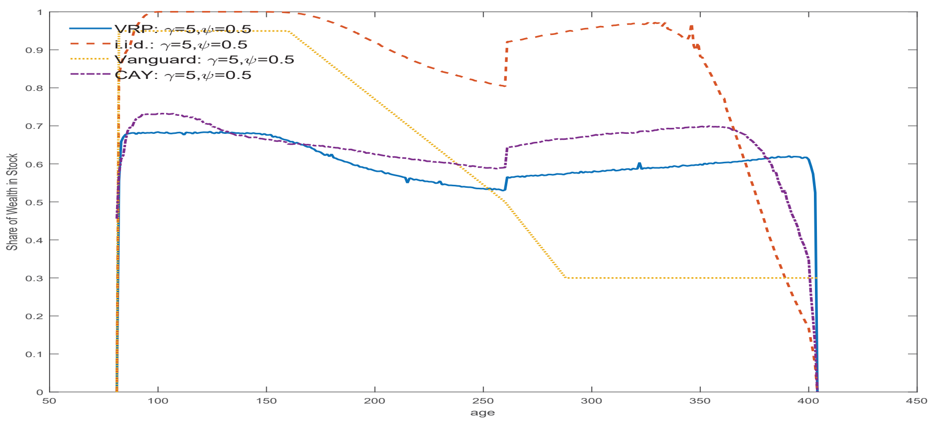
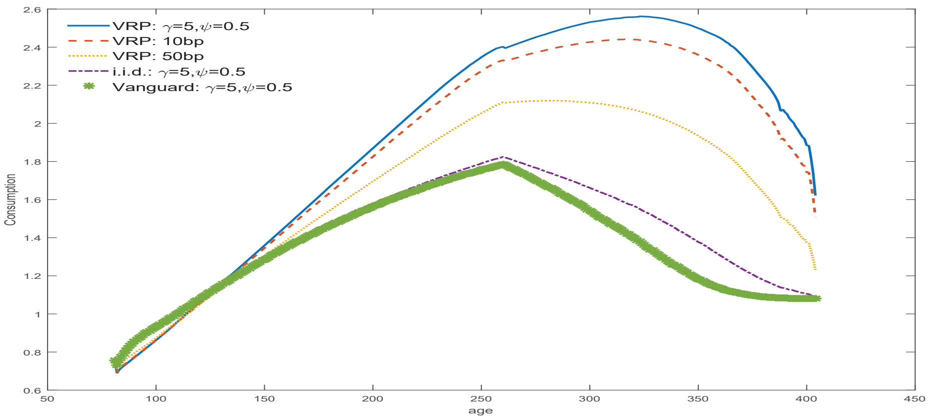


FIGURE 3

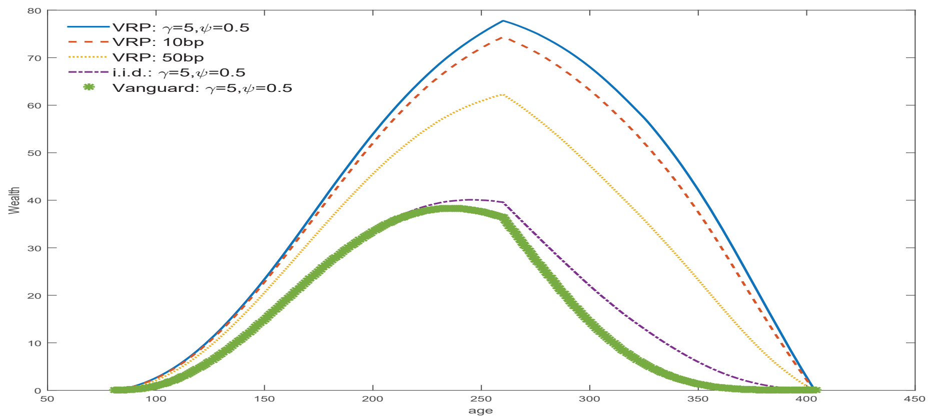
The Life-Cycle Profiles (Extra Transaction Cost)

Figure 3 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stock (Graph C) between the mean reversion models using different transaction cost and i.i.d. case while keeping bequest motive at 0.0.

Graph A. Consumption



Graph B. Wealth



Graph C. Share of Wealth in Stock

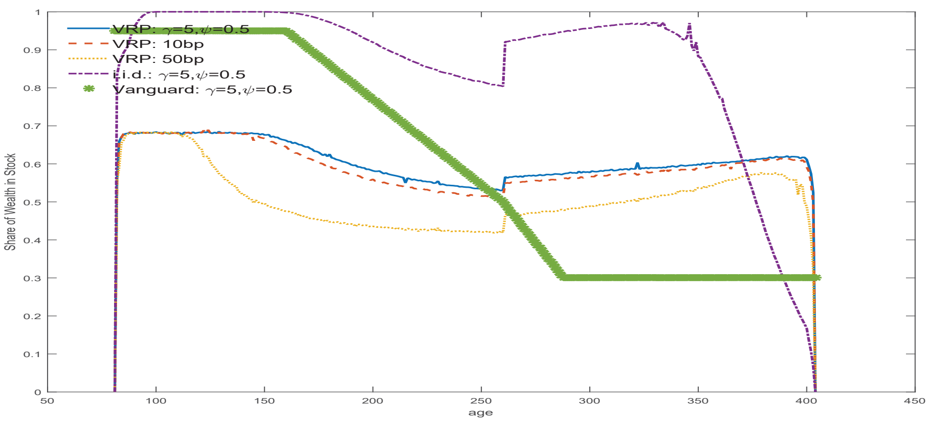
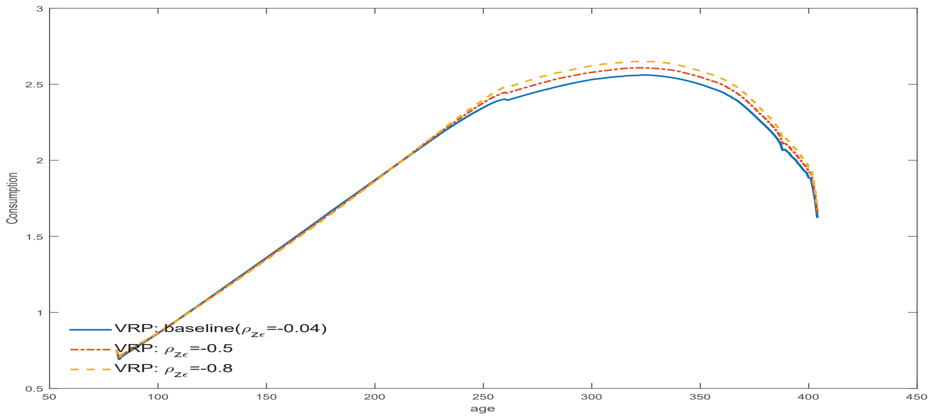


FIGURE 4

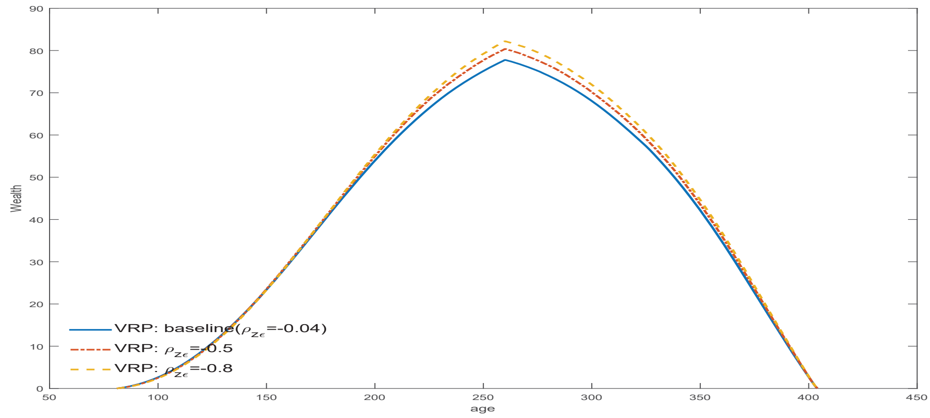
The Life-Cycle Profiles due to Variations in $\rho_{z\varepsilon}$

Figure 4 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stock (Graph C) between the mean reversion models using VRP as predictor and i.i.d. case due to variations in $\rho_{z\varepsilon}$.

Graph A. Consumption (Bequest = 0.0)



Graph B. Wealth (Bequest = 0.0)



Graph C. Share of Wealth in Stock (Bequest = 0.0)

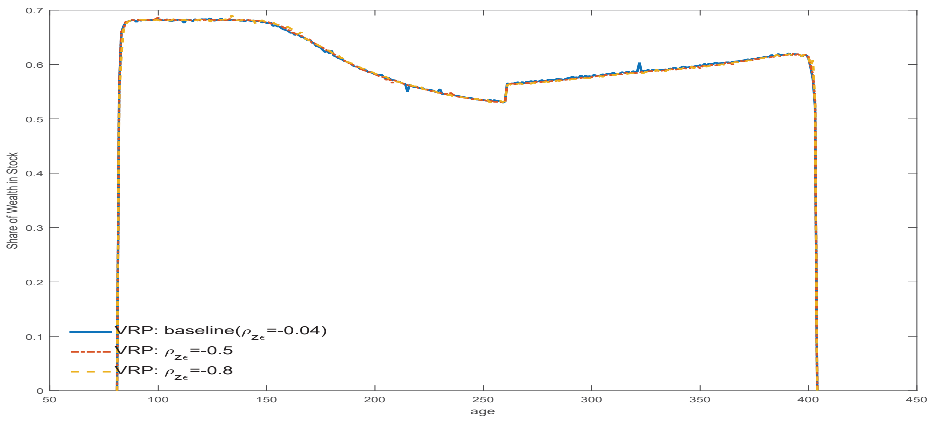
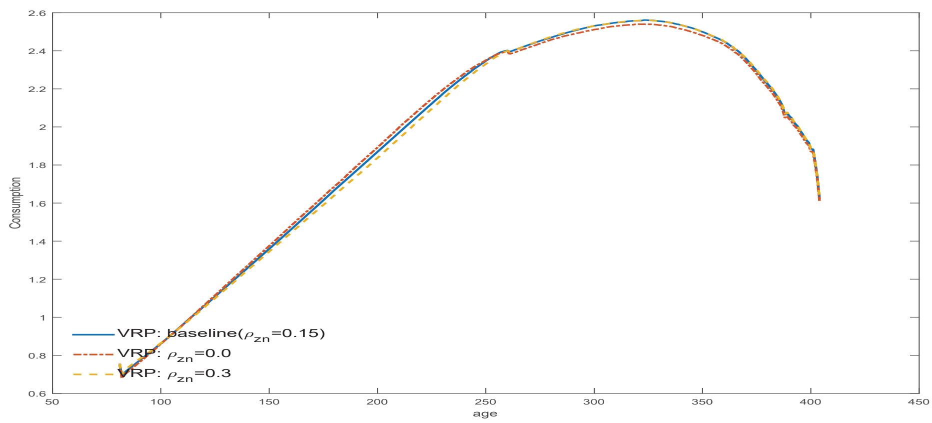


FIGURE 5

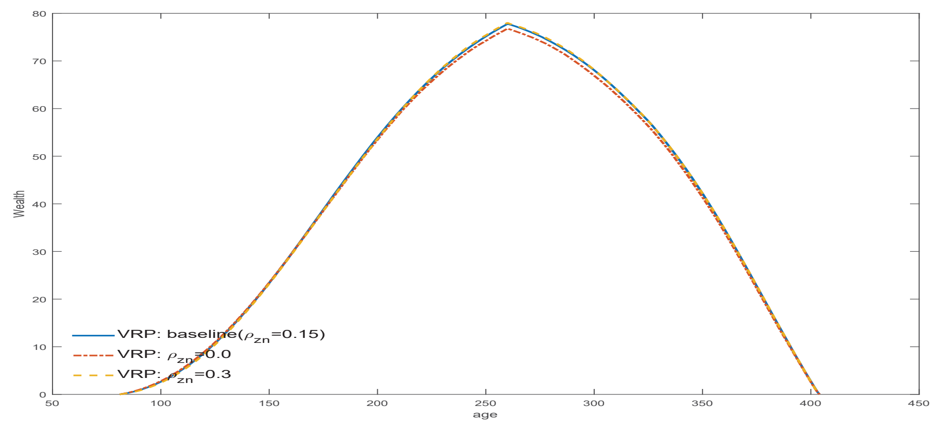
The Life-Cycle Profiles due to ρ_{zn}

Figure 5 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stock (Graph C) between the mean reversion models using VRP as predictor and i.i.d. case due to variations in ρ_{zn}.

Graph A. Consumption (Bequest = 0.0)



Graph B. Wealth (Bequest = 0.0)



Graph C. Share of Wealth in Stock (Bequest = 0.0)

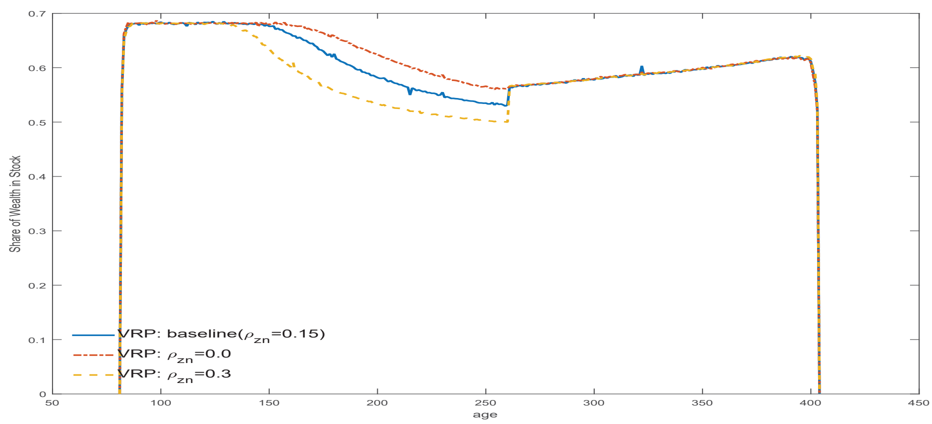
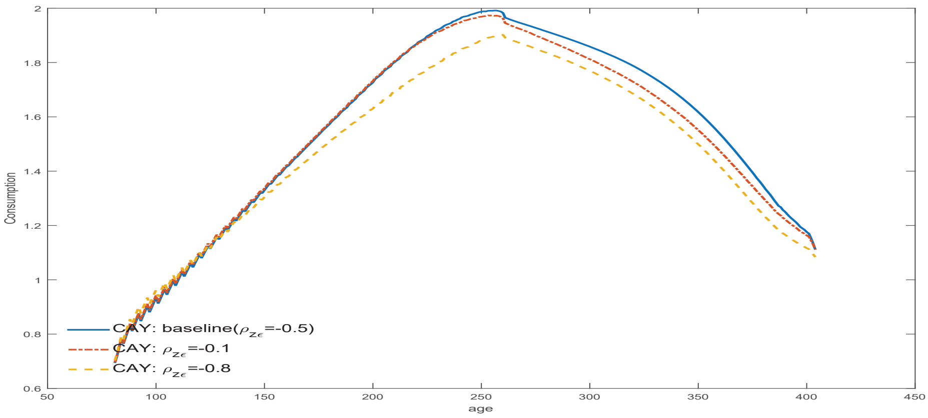


FIGURE 6

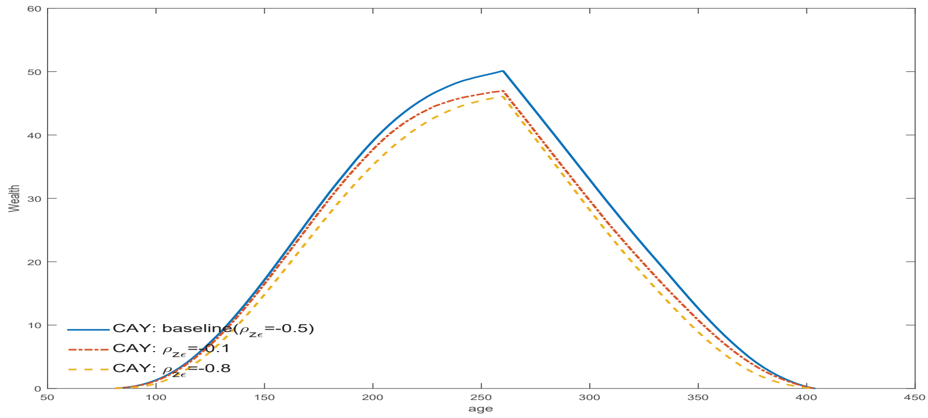
The Life-Cycle Profiles due to Variations in $\rho_{z\varepsilon}$

Figure 6 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stock (Graph C) between the mean reversion models using CAY as predictor and i.i.d. case due to variations in $\rho_{z\varepsilon}$.

Graph A. Consumption (Bequest = 0.0)



Graph B. Wealth (Bequest = 0.0)



Graph C. Share of Wealth in Stock (Bequest = 0.0)

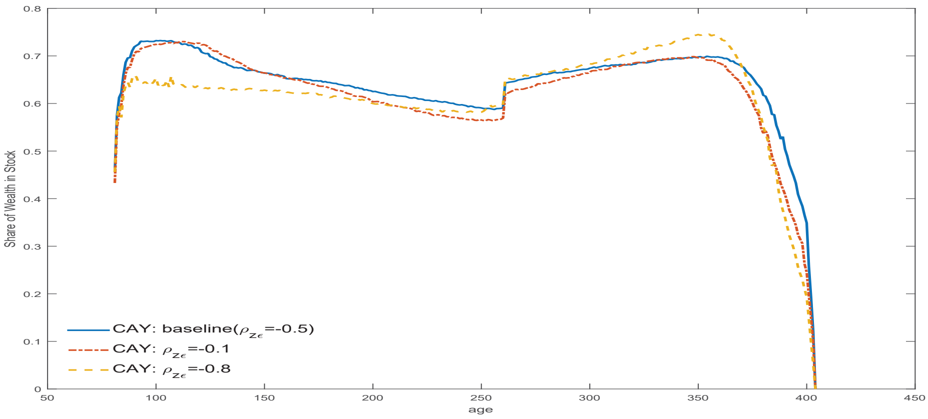
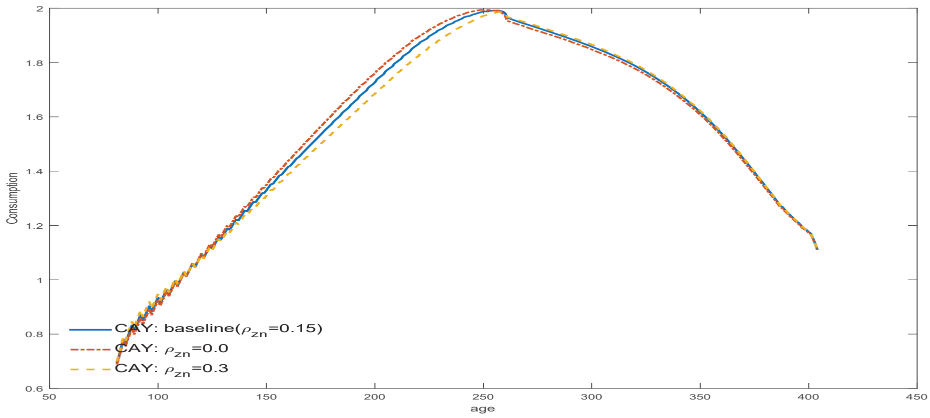


FIGURE 7

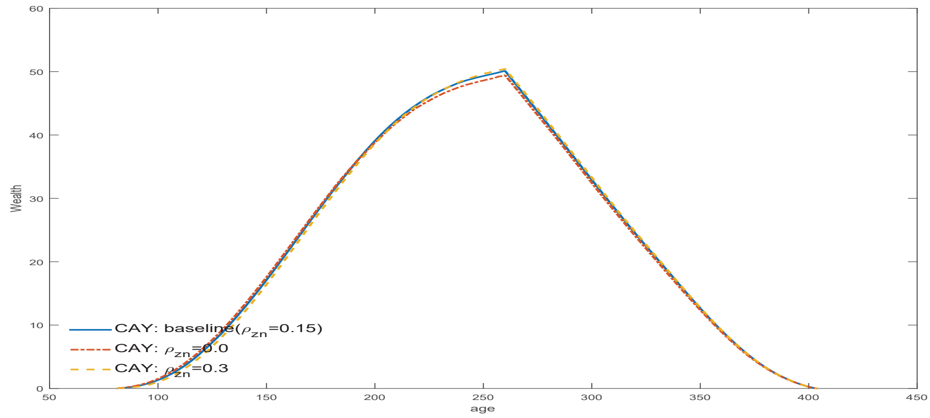
The Life-Cycle Profiles due to Variations in ρ_{zn}

Figure 7 shows the difference of consumption (Graph A), wealth accumulation (Graph B) and share of wealth in stock (Graph C) between the mean reversion models using CAY as predictor and i.i.d. case due to variations in ρ_{zn}.

Graph A. Consumption (Bequest = 0.0)



Graph B. Wealth (Bequest = 0.0)



Graph C. Share of Wealth in Stock (Bequest = 0.0)

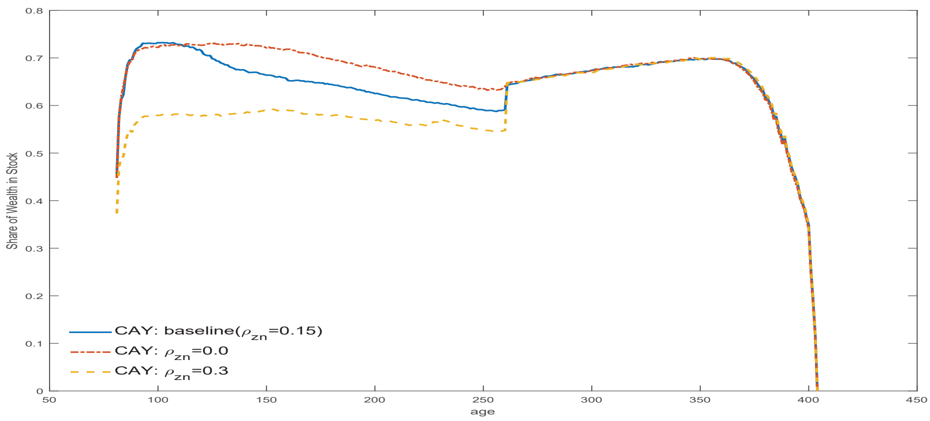
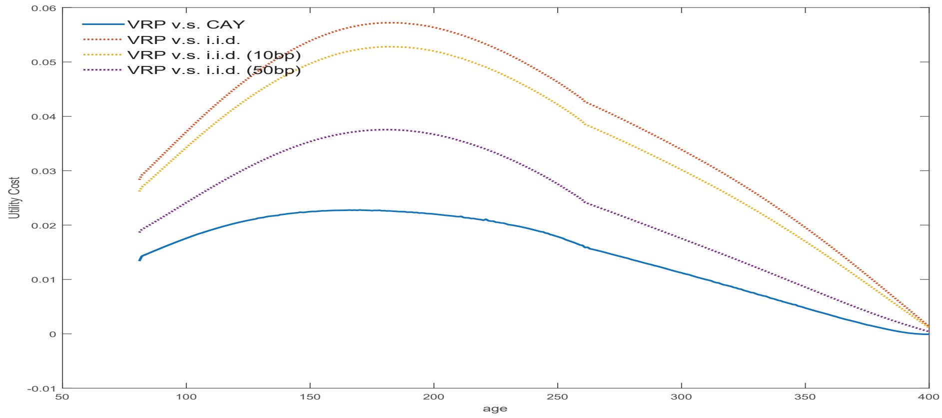


FIGURE 8

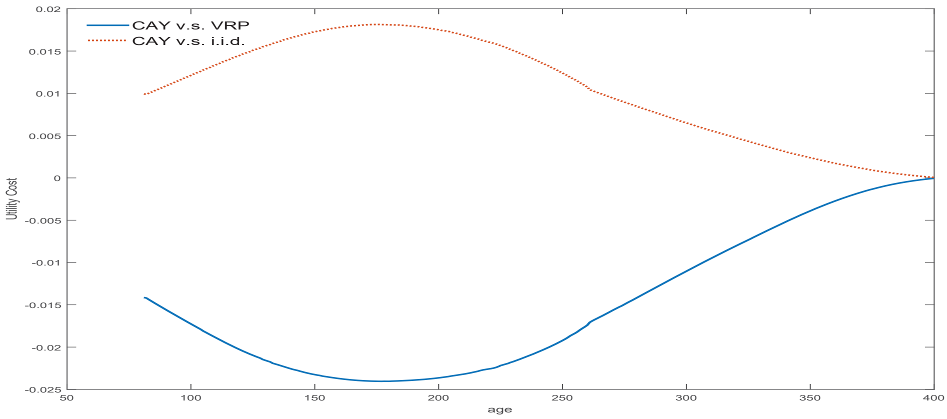
Welfare Evaluation

Figure 8 presents the welfare evaluation across different models.

Graph A. Welfare Evaluation using VRP as base (Bequest Motive = 0.0)



Graph B. Welfare Evaluation using CAY as base (Bequest Motive = 0.0)



Graph C. Welfare Evaluation using i.i.d. as base (Bequest Motive = 0.0)

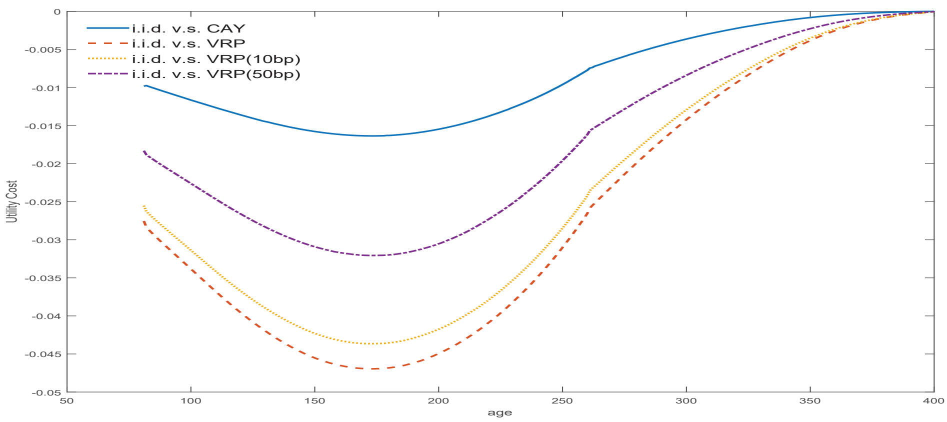
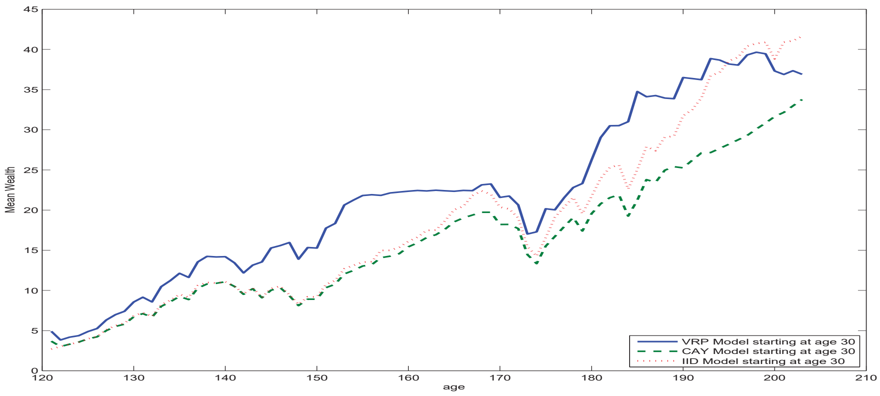


FIGURE 9

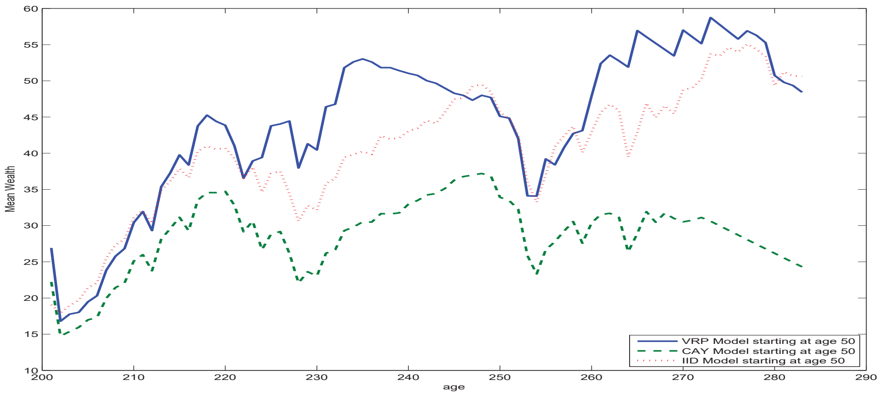
Out of Sample Testing

Figure 9 presents the out of sample testing across different models.

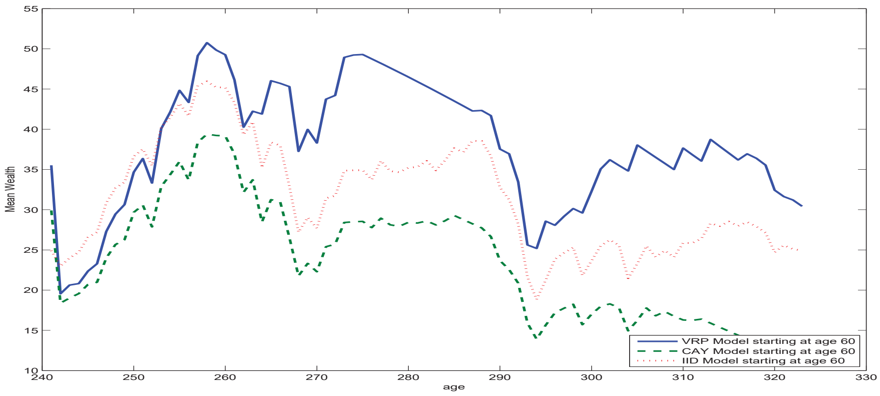
Graph A. Out of Sample Testing for Age Group 30 (Bequest Motive = 0.0)



Graph B. Out of Sample Testing for Age Group 50 (Bequest Motive = 0.0)



Graph C. Out of Sample Testing for Age Group 60 (Bequest Motive = 0.0)



Bibliography

[1] Ang, A., and G. Bekaert "Stock return predictability: Is it there? " *Review of Financial Studies*, 20 (2007), 651–707.

[2] Angerer, X., and P. Lam "Income Risk and Portfolio Choice: An Empirical Study." *Journal of Finance*, 64 (2009), 1037–1055.

[3] Balduzzi, P., and A. W. Lynch "Transaction Costs and Predictability: Some Utility Cost Calculations." *Journal of Financial Economics,* 52 (1999), 47–78.

[4] Barberis, N. "Investing for the Long Run when Returns Are Predictable." *Journal of Finance*, 55 (2000), 225–264.

[5] Benzoni, L.; P. Collin-Dufresne; and R. Goldstein "Portfolio Choice over the Life-Cycle when the Stock and Labor Markets Are Cointegrated." *Journal of Finance*, 62 (2007), 2123–2167.

[6] Bollerslev, T.; G. Tauchen; and H. Zhou "Expected Stock Returns and Variance Risk Premia." *Review of Financial Studies*, 22 (2009),

4463–4492.

[7] Bonaparte, Y.; G. Korniotis; and A. Kumar."Income Hedging and Portfolio Decisions." *Journal of Financial Economics*, 113 (2014), 300–324.

[8] Brandt, M. "Estimating Portfolio and Consumption Choice: A Conditional Euler Equations Approach." *Journal of Finance*, 54 (1999), 1609–1646.

[9] Brandt, M. W.; A. Goyal; P. Santa-Clara; and J. R. Stroud "A Simulation Approach to Dynamic Portfolio Choice with an Application to Learning about Return Predictability." *Review of Financial Studies*, 18 (2005), 831–873.

[10] Brennan, M.; E. Schwartz; and R. Lagnado "Strategic Asset Allocation." *Journal of Economic Dynamics and Control*, 21 (1997), 1377–1403.

[11] Brennan, M., and Y. Xia "Dynamic Asset Allocation under Inflation." *Journal of Finance,* 57 (2002), 1201–1238.

[12] Campbell, J. Y. "Stock returns and the term structure." *Journal of Financial Economics*, 18(2) (1987), 373–399.

[13] Campbell, J. Y., and R. J. Shiller "The dividend-price ratio and expectations of future dividends and discount factors." *Review of Financial Studies*, 1(3) (1988a), 195–228.

[14] Campbell, J. Y., and R. J. Shiller "Stock-prices, earnings, and expected dividends." *Journal of Finance*, 43(3) (1988b), 661–676.

[15] Campbell, J. Y., and L. Viceira "Consumption and Portfolio Decisions When Expected Returns Are Time Varying." *Quarterly Journal of Economics*, 114 (1999), 433–495.

[16] Campbell, J.; Y. L. Chan; and L. Viceira "A Multivariate Model of Strategic Asset Allocation." *Journal of Financial Economics*, 67 (2003), 41–80.

[17] Campbell, J. Y.; J. Cocco; F. Gomes; P. Maenhout; and L. Viceira "Stock Market Mean Reversion and the Optimal Allocation of a Long-Lived Investor." *European Finance Review*, 5 (2001), 269–292.

[18] Carroll, C. "Buffer-Stock Saving and the Life Cycle/Permanent Income Hypothesis." *Quarterly Journal of Economics,* 112 (1997), 3–55.

[19] Cocco, J., and P. Lopes "Reverse Mortgage Design." Working Paper, LBS/LSE (2015).

[20] Cocco, J.; F. Gomes; and P. Maenhout "Consumption and Portfolio Choice over the Life–Cycle." *Review of Financial Studies*, 18 (2005), 491–533.

[21] Cochrane, J. H. "Asset Pricing (Revised Edition)." *Princeton University Press*, (2005).

[22] Cooper, R., and G. Zhu "Household Finance over the Life-Cycle: What does Education Contribute?" *Review of Economic Dynamics*, 20 (2016), 63–89.

[23] Davidoff, T. "Can High Costs Justify Weak Demand for the Home Equity Conversion Mortgage?" *Review of Financial Studies*, 28 (2015), 2364–2398.

[24] Davis, S.; F. Kubler; and P. Willen "Borrowing Costs and the Demand for Equity over the Life Cycle." *The Review of Economics and Statistics* 88 (2006), 348–362.

[25] Deaton, A. “Saving and Liquidity Constraints.” *Econometrica* 59 (1991), 1221–1248.

[26] Donaldson, S.; F. Kinniry; R. Aliaga-Diaz; A. Patterson; and M. DiJoseph “Vanguard’s Approach to Target-Date Funds.” Valley Forge, PA: Vanguard Research (2013).

[27] Epstein, L., and S. Zin “Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework.” *Econometrica*, 57 (1989), 937–969.

[28] Fagereng, Andreas, Charles Gottlieb and Luigi Guiso, 2016, “Asset Market Participation and Portfolio Choice over the Life-Cycle,” *Journal of Finance*.

[29] Fama, E. F. and K. R. French “Dividend yields and expected stock returns.” *Journal of Financial Economics* 22(1) (1988), 3–25.

[30] Fama, E. F. and K. R. French “Business conditions and expected returns on stocks and bonds.” *Journal of Financial Economics* 25(1) (1989), 23–49.

[31] Floden, M. “A Note on the Accuracy of Markov-Chain Approximations to Highly Persistent AR(1) Processes.” *Economics Letters*, 99 (2008), 516–520.

[32] Gomes, F., and A. Michaelides “Optimal Life-Cycle Asset Allocation: Understanding the Empirical Evidence.” *Journal of Finance*, 60(2005), 869–904.

[33] Gomes, F., A. Michaelides, and V. Polkovnichenko “Optimal Savings with Taxable and Tax-Deferred Accounts.” *Review of Economic Dynamics*, 12 (2009), 718–735.

[34] Goyal, A. and I. Welch “A comprehensive look at the empirical performance of equity premium prediction.” *Review of Financial Studies*, 21 (2008), 1455–1508.

[35] Haliassos, M., and A. Michaelides “Portfolio Choice and Liquidity Constraints.” *International Economic Review*, 44 (2003), 144–177.

[36] Hamilton, J. D. “Time Series Analysis. ” *Princeton University Press*, (1994).

[37] Heaton, J., and D. Lucas "Portfolio Choice in the Presence of Background Risk." *The Economic Journal*, 110 (2000), 1–26.

[38] Kim, T. S., and E. Omberg "Dynamic Nonmyopic Portfolio Behavior." *Review of Financial Studies*, 9 (1996), 141–161.

[39] Koijen, R.; T. Nijman; and B. Werker "When Can Life Cycle Investors Benefit from Time-Varying Bond Risk Premia?" *Review of Financial Studies*, 23 (2010), 741–780.

[40] Lan, C. "An Out-of-Sample Evaluation of Dynamic Portfolio Strategies." *Review of Finance*, 19 (2015), 2359–2399.

[41] Lettau, M. and S. Ludvigson "Consumption, aggregate wealth, and expected stock returns. " *Journal of Finance*, 56 (2001), 815–849.

[42] Lynch, A., and S. Tan "Labor Income Dynamics at Business-Cycle Frequencies: Implications for Portfolio Choice." *Journal of Financial Economics*, 101 (2011), 333–359.

[43] Michaelides, A. and Y. Zhang "Stock market mean reversion and portfolio choice over the life cycle. " *Journal of Financial and Quantitative Analysis*, Forthcoming.

[44] Munk, C., and C. Sorensen. “Dynamic Asset Allocation with Stochastic Income and Interest Rates.” *Journal of Financial Economics*, 96 (2010), 433–462.

[45] Newey, W. K., and K. D. West. “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. ” *Econometrica*, 55 (1987), 703–708.

[46] Pastor, L., and R. F. Stambaugh. “Predictive systems: Living with imperfect predictors. ” *Journal of Finance*, 64 (2009), 1583–1628.

[47] Pastor, L., and R. F. Stambaugh. “Are Stocks Really Less Volatile in the Long Run?” *Journal of Finance*, 67 (2012), 431–477.

[48] Pettenuzzo, D.; A. Timmermann; and R. Valkanov. “Forecasting Stock Returns under Economic Constraints.” *Journal of Financial Economics*, 114 (2014), 517–553.

[49] Samuelson, P. “Lifetime portfolio selection by dynamic stochastic programming.” *Review of Economics and Statistics*, 51 (1969), 239–246.

[50] Tauchen, G. “Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions.” *Economics Letters*, 20 (1986), 177–181.

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

[52] Tsay, R. S. "Analysis of Financial Time Series, 3rd Edition." *Wiley*, 2010.

[53] Viceira, L. "Optimal Portfolio Choice for Long-Horizon Investors with Nontradable Labor Income." *Journal of Finance*, 55 (2001), 1163–1198.

[54] Vissing-Jorgensen, A. "Limited Asset Market Participation and the Elasticity of Intertemporal Substitution." *Journal of Political Economy*, 110 (2002), 825–853.

[55] Wachter, J. "Optimal Consumption and Portfolio Allocation with Mean-Reverting Returns: An Exact Solution for Complete Markets." *Journal of Financial and Quantitative Analysis*, 37 (2002), 63–91.

[56] Weil, P. "Nonexpected Utility in Macroeconomics." *Quarterly Journal of Economics*, 2 (1990), 29–42.

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

[58] Xia, Y. "Learning about Predictability: The Effects of Parameter Uncertainty on Dynamic Asset Allocation." *Journal of Finance*, 56 (2001), 205–246.