Dynamic Portfolio Allocation Under Market Incompleteness and Wealth Effects

Shen, Yiwen; Li, Chenxu; Scaillet, Olivier; Jiang, Yueting

How to cite

SHEN, Yiwen et al. Dynamic Portfolio Allocation Under Market Incompleteness and Wealth Effects. In: Operations research, 2025, p. 78. doi: 10.1287/opre.2024.0976

This publication URL: https://archive-ouverte.unige.ch/unige:187714

Publication DOI: 10.1287/opre.2024.0976

Dynamic Portfolio Allocation under Market Incompleteness and Wealth Effects

Yiwen Shen
School of Business and Management, HKUST, yiwenshen@ust.hk

Chenxu Li
Guanghua School of Management, Peking University, cxli@gsm.pku.edu.cn

Olivier Scaillet
University of Geneva and Swiss Finance Institute, olivier.scaillet@unige.ch

Yueting Jiang
CUHK Business School, CUHK, yueting.jiang@link.cuhk.edu.hk

This paper develops a novel decomposition of optimal dynamic portfolio choice under flexible incomplete market models and the wealth-dependent HARA utility. The decomposition reveals the fundamental impacts of market incompleteness and wealth effect in portfolio allocation. With hedgeable interest rate risk, we show that the optimal portfolio under HARA utility can be decomposed into a pure CRRA optimal portfolio and a financing bond portfolio that matches the investor future subsistence requirements. In this case, the wealth growth rate is always higher for HARA investors with more initial wealth, leading to increased wealth inequality regardless of the market scenario. As an application of our decomposition, we solve the HARA optimal policy in closed-form under an incomplete market model with both stochastic interest rate and volatility. Using parameters calibrated from U.S. market data, we find that the wealth effect generates a procyclical pattern in investor stock positions and time-varying risk aversion levels. Moreover, the wealth effect in investor utility and the increased risk premium in stressed market combined lead to a novel "buy-high-sell-low" channel that may hurt HARA investors with low initial wealth.

Key words: optimal portfolio choice, incomplete market, wealth-dependent utility, closed-form analysis, wealth inequality, heterogeneous investors.

1. Introduction

Optimal portfolio choice has been a central topic in modern financial economics, drawing long-standing interest in both finance industry and academic research. The static mean-variance framework of Markowitz (1952) laid a foundation for modern portfolio theory. Following the seminal work by Samuelson (1969) and Merton (1969, 1971), various studies have been developed for the optimal dynamic portfolio choice; see the surveys in, e.g., Brandt (2010), Wachter (2010), and Detemple (2014). As an optimal stochastic control problem in a continuous-time setting, the solution of optimal policies usually relies on two approaches. The first one is the well-known dynamic programming method, which characterizes the optimal policy via partial differential equations (PDEs). However, the resulting PDEs are usually very difficult to solve in high-dimensional problems. It hinders implementations of optimal portfolios for models with multiple state variables, non-linear dynamics, and wealth-dependent utilities (Detemple 2014). The second approach is the martingale method pioneered and developed by, e.g., Cox and Huang (1989), Ocone and Karatzas (1991), and Detemple et al. (2003). It first solves the optimal consumption and bequest. Then, the optimal portfolios are represented as conditional expectations of random variables with explicit dynamics. Accordingly, Monte Carlo simulation can be used to solve the optimal portfolios numerically (Detemple et al. 2003).

To reveal the economic nature of the optimal portfolio, the decomposition of optimal policies into mean variance (myopic) and hedge components is developed by Merton (1971) and has become a state-of-the-art approach (e.g., Liu 2007, Detemple and Rindisbacher 2010, and Moreira and Muir 2019). See also Basak and Chabakauri (2010) for decomposing the optimal portfolio under the mean-variance framework. In recent works, Capponi and Rubtsov (2022) study the portfolio allocation problem that accounts for losses under systemic tail events, and decompose the optimal portfolio into a mean-variance term and an adjustment term for systemic risk. He and Jiang (2020) show that the mean-variance efficiency of a fractional Kelly portfolio can be improved by adding a hedge component and a corresponding adjustment term.

For the purpose of implementing and analyzing the behavior of optimal portfolios, existing works largely focus on specific affine models (e.g., Duffie et al. 2000) and wealth-independent utilities, such as the basic constant relative risk aversion (CRRA) utility and the recursive utility that generalizes it.[1] While these specifications bring analytical convenience, e.g., closed-form optimal portfolio policies in specific cases,[2] they limit the model capacity to capture empirically flexible market dynamics and realistic investor preferences. For general diffusion models without closed-form policies, Detemple et al. (2003) develop an effective Monte Carlo simulation approach based on the decomposition of the optimal policy. However, this method is by far largely limited to the complete market setting.[3] Thus, the dynamic optimal portfolio choice problem under incomplete market models with wealth-dependent utilities is still an open challenge. This is exactly the focus of our work.

In this paper, we use the martingale approach to develop a decomposition for the optimal portfolio policy under a general class of incomplete market diffusion models. In an incomplete market, investors cannot fully hedge the risk by investing in the risky assets, making the optimal policy hard to solve. In addition, we focus on the wealth-dependent hyperbolic absolute risk aversion (HARA) utility with lower bounds for both intermediate consumption and terminal wealth. Compared with the CRRA utility commonly used in the literature, the HARA utility offers more flexibility in modelling the investor preference, and captures realistic features in the investment decisions such as portfolio insurance, investment goals, and subsistence level requirements. However, it is much less studied due to its mathematical inconvenience (see Kim and Omberg 1996 for a rare case with closed-form policy and Duffie et al. 1997 for solving the optimal policy in a model with constant coefficients for stock and income dynamics). To establish the decomposition, we apply the “least favorable completion” method in Karatzas et al. (1991) for general diffusion models. It completes the market by introducing suitable fictitious assets. Then, it establishes the equivalence between the optimal policy in the completed market and that in the original market by setting appropriate price of risk for these fictitious assets. Such price of risk is endogenously determined by the investor utility function, and is thus referred to as the investor-specific price of risk in incomplete market models. It is also known as the “shadow price” of market incompleteness (e.g., Detemple and Rindisbacher 2010).

We first develop the optimal portfolio decomposition under general incomplete market models with the CRRA and HARA utilities. The optimal policy is decomposed into a mean-variance component and a hedge component for the uncertainty in interest rate and price of risk. In our decomposition, each component is expressed as conditional expectation of random variables with explicit dynamics, and the suitable investor-specific price of risk is characterized by an integral-type equation system. Our decomposition reveals that the optimal policy is indeed wealth-independent under the CRRA utility, but not so under the HARA utility. We find that the mean-variance component of the HARA investor satisfies a ratio relationship with its CRRA counterpart. The HARA investor first sets aside the amount of wealth that equals the present value of her future subsistence requirements under the martingale measure determined by the investor-specific price of risk; then she constructs the mean-variance component based on the remaining wealth just like a CRRA investor. Moreover, we show that the hedge component under the HARA utility contains an additional term for hedging the uncertainty in the present value of investor future subsistence requirements. In general incomplete market models, the investor-specific price of risk may not coincide for the CRRA and HARA investors. It imposes a major challenge for establishing connections between their optimal policies.

We then proceed to a specific, but highly flexible class of incomplete market models. We consider the models in which the interest rate risk can be fully hedged by the bond assets in the market, although the full market is still incomplete. As we do not impose restriction on the dynamics of asset prices and state variables, the above set-up is general enough to cover a wide range of classical models considered in the dynamic portfolio allocation literature.[4] We show that, in this case, the investor-specific price of risk coincides under the HARA and CRRA utilities, and the optimal policies under the two utilities have a closed-form relationship with intuitive economic interpretations. First, when interest rate risk is hedgeable, we can explicitly calculate the present value of the HARA investor future subsistence requirements as the market value of a hypothetical bond holding scheme. Then, the optimal portfolio of the HARA investor can be constructed as follows. First, the HARA investor sets aside the amount of wealth equal to the market value of the hypothetical bond holding scheme, whose payments exactly finance her future subsistence requirements. Then, she invests as a CRRA investor for both the mean-variance and hedge components. Finally, she holds an additional portfolio of bond assets that replicates the dynamics of the hypothetical bond holding scheme. As such, we can decompose the HARA optimal portfolio into a CRRA optimal portfolio and a financing portfolio for future subsistence requirements. Such structure is only valid when the interest rate risk is hedgeable, thus future cash flows can be perfectly synthesized by investing in the bond assets.

When the interest rate is nonrandom, the additional financing portfolio vanishes, as it suffices for the HARA investor to hold the riskless asset to finance her subsistence requirements.

We apply our theoretical decomposition to analyze the optimal portfolio allocation of investors in different wealth groups. We consider two HARA investors with different initial wealth levels, but coincide in other aspects of their utility functions. We show that with hedgeable interest rate risk, the ratio of the two investor remaining wealth, after subtracting the hypothetical bond scheme for future subsistence requirements, stays constant over time. Essentially, it is because the two investors hold the same CRRA portfolio in addition to their financing portfolios. As an important consequence, the overall wealth growth rate, after all subsistence requirements are met, is always higher for the high-wealth investor than for the low-wealth one. This finding contributes to the studies on wealth inequality, which has drawn growing attention in recent decades (see a review in De Nardi and Fella 2017). Specifically, it provides support to the wealth return channel for explaining the growth of wealth inequality, and is consistent with the empirical evidence that wealth returns are increasing in investor wealth level (Fagereng et al. 2020 and Bach et al. 2020). Our study complements the literature in two novel aspects. First, we solve the dynamic portfolio optimization problem for the HARA investors, thus providing a theoretical foundation for their decisions. Second, we show that the wealth gap increases regardless of the underlying model dynamics and realized market scenarios (e.g, bull or bear markets).

Our decomposition greatly facilitates the implementation of HARA optimal policy, which is rare in the literature due to its analytical inconvenience. With hedgeable interest rate risk, we show that we can conveniently obtain the HARA optimal policy from its CRRA counterpart, which is usually much easier to compute via either closed-form solution or numerical approaches. After the CRRA policy is solved, a straightforward Monte Carlo simulation can be used to calculate the value of the hypothetical bond holding scheme and the additional hedging portfolio of the HARA investor, as the dynamics of the random variables involved are explicitly given. As a benchmark, we consider a high-dimensional incomplete market model with both stochastic interest rate and stochastic volatility. The model includes a bond asset and a stock asset. The interest rate follows a Cox-Ingersoll-Ross process as in Cox et al. (1985), and its uncertainty can be fully hedged by the bond asset. On the other hand, the stock price and its variance process follow the the classical stochastic volatility model of Heston (1993). The Cox-Ingersoll-Ross-Heston stochastic volatility and stochastic interest rate (CIRH-SVSIR) model includes two risky assets driven by three independent Brownian motions. We solve the optimal policy in closed-form for a HARA investor under this incomplete market model. It demonstrates the application potential of our theoretical decomposition results.

We conduct a comprehensive comparative study to reveal the wealth effect in optimal portfolio allocation. To get fresh empirical support, we calibrate the CIRH-SVSIR model by the maximum likelihood estimation approach in Aït-Sahalia and Kimmel (2007) and Ait-Sahalia and Kimmel (2010), which is widely used for estimating continuous-time models. We use the SPDR S&P 500 ETF as the stock asset, and extract the values of underlying interest rate and volatility from the US treasury yields and VIX index. With the estimated parameters, we show that the optimal stock (resp. bond) weight increases (resp. decreases) with the HARA investor wealth level. It is consistent with the empirical findings that the investment in risky assets increases concavely in investor financial wealth; see, e.g., Roussanov (2010), Wachter and Yogo (2010), and Calvet and Sodini (2014). Moreover, under the HARA utility, the optimal stock weight increases with the investment horizon via two channels. First, longer investment horizon increases the hedging demand of the investor. Second, with a longer investment horizon, the bond scheme for financing the HARA investor future subsistence requirements becomes cheaper, increasing the remaining wealth allocated on the stock. The second channel is absent under the CRRA utility.

In addition to the above static analysis, we reveal the wealth effect from a dynamic aspect by checking how the complex market dynamics affect the optimal allocation strategy and overall investment performance of HARA investors.[5] We show that under the CIRH-SVSIR model, the optimal stock weight of HARA investors depends on the entire paths of market dynamics and exhibits a procyclical behavior unseen under the CRRA utility. That is, the HARA investor increases (resp. decreases) her stock holding during bull (resp. bear) markets, a pattern consistent with the empirical observations (Amromin and Sharpe 2014). Such cycle-dependence is more significant for HARA investors with lower initial wealth levels, as their optimal policies are more sensitive to market scenarios. Moreover, the wealth-dependent HARA utility endogenously generates the time-varying risk aversion of investors: the HARA investors become more (resp. less) risk averse during bear (resp. bull) market regimes. It contributes to the fast growing literature on investor time-varying risk aversion and its implications in portfolio allocation. Using portfolio survey data, Guiso et al. (2018) find that investor risk aversion substantially increases after the Global Financial Crisis. Berrada et al. (2018) develop a model with regime-dependent risk preference and show that it can explain the excess equity premium and volatility observed in data. Li et al. (2022) study the dynamic portfolio allocation problem with both regime-dependent return and risk aversion. They find that investors with regime-dependent risk aversion achieve better investment performance than those with constant ones.

The wealth level of HARA investors introduces a risk-return trade-off that substantially impacts the overall investment performance: the HARA investor with higher initial wealth invests more in the risky asset, leading to a higher return but also more risk. We quantify such trade-off by simulating a large number of paths under our estimated model to account for possible market scenarios. With the simulated paths and the closed-form optimal policy, we compute ex-ante expectations of performance statistics including excess return mean, volatility, 99% Value-at-Risk, and maximum drawdown. We find that as HARA investor initial wealth increases by ten times from the subsistence level, the average annual excess return increases from 12.5% to 28.4%, while the volatility increases from 14.3% to 31.6%, and the maximum drawdown jumps from 23.3% to 42.9%. These monotonic increasing patterns are statistically significant. The huge differences in the investment performance highlight the practical relevance of understanding the wealth effects in delegated portfolio management.

Finally, we identify a novel market timing effect in the stock trading of HARA investors. We show that the high-wealth HARA investors can better "time" the market than low-wealth ones as they tend to have larger weight on the stock during the periods with higher risk premium, even after the average level of stock weight is controlled. This market timing effect contributes to a higher Sharpe ratio of high-wealth HARA investors. It can be explained by the interplay of the leverage effect in the CIRH-SVSIR model and the wealth effect of HARA utility. By the leverage effect, expected stock returns are higher during bear market regimes. However, by the wealth effect, HARA investors tend to reduce their stock position in bear markets. Such impact is more significant for low-wealth HARA investors, making them less capable to benefit from the higher expected returns in stressed markets. This generates a "buy-high-sell-low" channel for explaining the variation in wealth growth rates of different investors. Such channel is empirically observed in the recent work of Sakong (2022). It finds that poorer households consistently buy houses in booms and sell after a boost, leading to a 60 basis points difference in expected annual returns between the first and third quartiles of US households. We show that this channel can be potentially explained by the wealth effect in optimal portfolio allocation.

The rest of this paper is organized as follows. Section 2 gives the model set-up and describes the fictitious completion method in incomplete market models. In Section 3, we develop the decomposition for general incomplete market models under CRRA and HARA utilities. In Section 4, we apply the decomposition to incomplete market models with hedgeable interest rate risk, and establish a closed-form relation between optimal policies under the HARA and CRRA utilities. Section 5 conducts a comprehensive comparative study on the wealth effect using the CIRH-SVSIR model. Section 6 concludes and provides discussions. We collect auxiliary results and proofs in the Electronic Companion.

2. Model Set-up and Fictitious Completion Method

We begin by setting up the model, the utility function, and the optimal dynamic portfolio choice problem under a general incomplete market framework. We then briefly introduce the fictitious completion method used to develop our portfolio decomposition results.

2.1. Model Set-up

Assume that the market consists of m risky assets and one savings account (risk-free asset). The price S_{it} of risky asset $i = 1, 2, \ldots, m$, follows the generic stochastic differential equation (SDE):

$$\frac{dS_{it}}{S_{it}} = (\mu_i(t, Y_t) - \delta_i(t, Y_t))\, dt + \sigma_i(t, Y_t) dW_t, \tag{1}$$

where Y_t is an n-dimensional state variable driven by the following generic SDE:

$$dY_t = \alpha(t, Y_t)dt + \beta(t, Y_t)dW_t. \tag{2}$$

In (1), W_t is a standard $d-$dimensional Brownian motion; $\mu_i(t,y)$ and $\delta_i(t,y)$ are scalar functions for modeling the mean rate of return and the dividend rate respectively; $\sigma_i(t,y)$ is a $d-$dimensional vector-valued function for modeling the volatility. In (2), $\alpha(t,y)$ is an $n-$dimensional vector-valued function for modeling the drift of the state variable Y_t; $\beta(t,y)$ is an $n \times d$ dimensional matrix-valued function for modeling the diffusion of Y_t. We assume the existence and uniqueness of solutions to SDEs (1) and (2). The savings account appreciates at an instantaneous interest rate $r_t = r(t, Y_t)$ for some scalar-valued function $r(t,y)$. The state variable Y_t governs all the investment opportunities in the market through the rate of return, the dividend rate, the volatility, and the instantaneous interest rate.

We focus on the incomplete market case where the number of independent Brownian motions is strictly larger than the number of tradable risky assets, i.e., $d > m$. In this case, we cannot fully hedge the uncertainty stemming from the Brownian motion by investing in the risky assets. As we will show, due to market incompleteness, the decomposition and implementation for the optimal portfolio policy become a challenging issue. Denote the investor wealth process by X_t. Then, it satisfies the following wealth equation:

$$dX_t = (r(t,Y_t)X_t - c_t)dt + X_t\pi_t^\top \left[(\mu(t,Y_t) - r(t,Y_t)1_m)dt + \sigma(t,Y_t)dW_t\right]. \tag{3}$$

In (3), the functions $\mu(t,y) := (\mu_1(t,y), \mu_2(t,y), \cdots, \mu_m(t,y))^\top$ and $\sigma(t,y) := (\sigma_1(t,y), \sigma_2(t,y), \cdots, \sigma_m(t,y))^\top$ represent the mean rate of return and volatility of the risky assets. We assume the volatility function $\sigma(t,y)$ has rank m, i.e., its rows are linearly independent. Besides, the scalar c_t is the instantaneous consumption rate; π_t is an $m-$dimensional vector representing the weights of the risky assets in the portfolio; 1_m denotes an $m-$dimensional column vector with all elements equal to one.

The investor maximizes her expected utility over both intermediate consumption and terminal wealth by dynamically allocating her wealth among the risky assets and the risk-free asset, subject to the non-bankruptcy condition. We focus on a general class of wealth-dependent utility functions: the hyperbolic absolute risk aversion (HARA) utility. Following the convention (see, e.g., Carroll and Kimball 1996), we formulate the optimization problem as

$$\sup_{(\pi_t, c_t)} E\left[\int_0^T we^{-\rho t}\frac{(c_t - \bar{c}_t)^{1-\gamma}}{1-\gamma}dt + (1-w)e^{-\rho T}\frac{(X_T - \bar{x}_T)^{1-\gamma}}{1-\gamma}\right], \text{ with } X_t \geq 0 \text{ for all } t \in [0,T], \tag{4}$$

where $\gamma > 0$ is the risk aversion coefficient; $w \in [0,1]$ is the weight for the intermediate consumption part in the utility, and ρ is the discount rate. The parameters $\bar{c}_t$ for $t \in [0,T]$ and $\bar{x}_T$ represent the minimum allowable amounts, i.e., subsistence levels, of intermediate consumption c_t and terminal wealth X_T, all of which are scalars. They are assumed to be positive and exogenously given in the optimization problem.

We allow the subsistence level for consumption $\bar{c}_t$ to be time-varying, reflecting potential variation in the consumption requirements over the investment horizon.

The HARA utility function (4) is defined for $c_t > \bar{c}_t$ and $X_T > \bar{x}_T$. If it is not satisfied, we assume the utility takes value of $-\infty$ (see Detemple and Rindisbacher 2010).The HARA utility allows for imposing lower bound constraints on investor consumption and/or terminal wealth, which is suitable for incorporating realistic features such as portfolio insurance, investment goal constraints, and subsistence requirements. That said, closed-form optimal policies under the HARA utility are rare due to technical difficulties. In addition, potential numerical methods (e.g., the Monte Carlo simulation approach in Detemple et al. 2003) are largely deployed under complete market settings.

As a simpler and special case, the HARA utility reduces to the widely used CRRA utility when $\bar{c}_t$ and $\bar{x}_T$ are set to zero in (4). With a CRRA utility, the investor optimization problem is formulated as:

$$\sup_{(\pi_t, c_t)} E\left[\int_0^T w e^{-\rho t}\frac{c_t^{1-\gamma}}{1-\gamma}dt + (1-w)e^{-\rho T}\frac{X_T^{1-\gamma}}{1-\gamma}\right], \text{ with } X_t \geq 0 \text{ for all } t \in [0,T]. \tag{5}$$

The wealth independence nature of CRRA utility brings mathematical convenience that leads to closed-form solution of the optimal policy or significant simplifications of the optimization problem under specific models. However, it is unable to capture the wealth effect in optimal portfolio allocation, which can be important in realistic settings. In the following, we assume $\gamma > 1$ in both the HARA and CRRA utility functions (4) and (5), i.e., the investor is more risk averse than that with the log-utility (see Wachter 2002).

Although the subsistence levels $\bar{c}_t$ and $\bar{x}_T$ are exogenously given, the HARA optimal policy cannot be directly derived from the CRRA optimal policy via simple transformations. In particular, the optimal policy has very different structures under the two utilities: wealth-independent under the CRRA utility and wealth-dependent under the HARA utility. Technically, as we discuss in Section 3, the present values of $\bar{c}_t$ and $\bar{x}_T$ are still stochastic. Thus, HARA investors need to consider how to finance her future subsistence requirements when making portfolio decisions. This makes solving the HARA optimal policy challenging.

2.2. Fictitious Completion Method

As a foundation for solving the optimal policy in incomplete market models, we briefly introduce the fictitious completion method in Karatzas et al. (1991). Specifically, the investor "completes" the market by bringing in $d-m$ fictitious assets without dividend payment. Their prices F_{it}, for $i = 1, 2, \ldots, d-m$, satisfy the following SDE:

$$\frac{dF_{it}}{F_{it}} = \mu_{it}^f dt + \sigma_i^f(t, Y_t)dW_t, \tag{6}$$

where the mean rates of returns μ_{it}^f are stochastic processes adaptive to the filtration generated by the Brownian motion W_t. We can choose the volatility function $\sigma^f(t,y) := (\sigma_1^f(t,y), \cdots, \sigma_{d-m}^f(t,y))^\top$ of the fictitious assets arbitrarily, as long as it has rank $d-m$ and satisfies the following orthogonal condition with the volatility function $\sigma(t,y)$ of the real risky assets S_t:

$$\sigma(t,y)\sigma^f(t,y)^\top \equiv 0_{m\times(d-m)}. \tag{7}$$

It guarantees that the fictitious and real assets are driven by different Brownian shocks, and thus completes the market.

Combining the m real risky assets with prices S_t in (1) and the $d-m$ fictitious risky assets with prices F_t in (6), we construct a completed market consisting of d risky assets and driven by d independent Brownian motions. In this completed market, we represent the prices of the risky assets, including both the real and fictitious ones, by a d–dimensional column vector $C_t = (S_t^\top, F_t^\top)^\top$. Denote their mean return rate and volatility by $\mu_t^c = ((\mu(t,Y_t)-\delta(t,Y_t))^\top, (\mu_t^f)^\top)^\top$ and $\sigma^c(t,Y_t) = (\sigma(t,Y_t)^\top, \sigma^f(t,Y_t)^\top)^\top$. By linear algebra, the orthogonal condition (7) implies that $\sigma^c(t,y)$ must be nonsingular. Thus, we are now in a complete market, where we can fully hedge the uncertainty stemming from all Brownian motions. Similar to (4) or (5), we consider the utility maximization problem in this completed market, which allows for investing in both the real assets S_t and the fictitious assets F_t.

In the completed market, we define the total price of risk as $\theta_t^c := \sigma^c(t,Y_t)^{-1}(\mu_t^c - r(t,Y_t)1_d)$. By the orthogonal condition (7), we can decompose the total price of risk as:

$$\theta_t^c = \theta^h(t,Y_t) + \theta_t^u. \tag{8}$$

Here, $\theta^h(t,Y_t)$ and θ_t^u are the prices of risk associated with the real and fictitious assets, respectively. They are both d–dimensional column vectors, defined as:

$$\theta^h(t,Y_t) := \sigma(t,Y_t)^+(\mu(t,Y_t) - r(t,Y_t)1_m), \tag{9a}$$

and

$$\theta_t^u := \sigma^f(t,Y_t)^+(\mu_t^f - r(t,Y_t)1_{d-m}), \tag{9b}$$

where $A^+ := A^\top(AA^\top)^{-1}$ denotes the Moore–Penrose inverse (Penrose 1955) of a general matrix A with linearly independent rows. The term $\theta^h(t,Y_t)$ in (9a) is referred to as the market price of risk, as it is fully determined by the real assets shared by all investors in the market. The term θ_t^u in (9b), however, is purely associated with the fictitious assets, which are specifically introduced for solving the optimal portfolio choice problem (4) or (5) in the incomplete market. As we will show momentarily, θ_t^u is endogenously determined by the investor utility function and the investment horizon. Thus, in line with the literature (Detemple 2014), we refer to θ_t^u as the investor-specific price of risk. It plays a central role in solving the optimal portfolio allocation problem in incomplete market models.

With the total price of risk in (8), we introduce the state price density as

$$\xi_t := \exp\left(-\int_0^t r(v,Y_v)dv - \int_0^t (\theta_v^c)^\top dW_v - \frac{1}{2}\int_0^t (\theta_v^c)^\top \theta_v^c dv\right).^6 \tag{10}$$

For any $s \geq t \geq 0$, we define the relative state price density as $\xi_{t,s} = \xi_s/\xi_t$. By Ito's formula, it satisfies

$$d\xi_{t,s} = -\xi_{t,s}[r(s,Y_s)ds + (\theta_s^c)^\top dW_s] \tag{11}$$

with initial value $\xi_{t,t} = 1$. The dynamics of $\xi_{t,s}$ hinges on the unknown investor-specific price of risk θ_v^u.

In the completed market, we can solve the optimal policy (π_t, π_t^F) by the martingale approach pioneered by Karatzas et al. (1987) and Cox and Huang (1989). We briefly discuss the general steps below and include more details in Section EC.4.1. The martingale approach first formulates the dynamic problem (4) as a static optimization problem. Then, we can obtain the optimal intermediate consumption and terminal wealth by the standard method of Lagrangian multiplier. On the other hand, we express the optimal policy (π_t, π_t^F) for the completed market via the martingale representation theorem (see, e.g., Section 3.4 in Karatzas and Shreve 1991). With the Clark-Ocone formula (Ocone and Karatzas 1991), we can further represent the optimal policy in the form of conditional expectations of suitable random variables. See Detemple et al. 2003 for decomposition under general complete-market diffusion models and a Monte Carlo simulation method.

We denote by π_t and π_t^F the optimal weights of real and fictitious assets respectively, which are m and $(d-m)$–dimensional vectors. By the least favorable completion principle proposed in Karatzas et al. (1991), the optimal policy π_t for the real assets in the completed market coincides with its counterpart in the original incomplete market, as long as we properly choose the investor-specific price of risk θ_v^u such that the optimal weights for the fictitious assets are always identically zero, i.e.,

$$\pi_v^F \equiv 0_{d-m}, \text{ for any } 0 \leq v \leq T. \tag{12}$$

The least favorable[7] constraint (12) determines the proper investor-specific price of risk θ_v^u and thus the state price density ξ_v in (10) for $0 \leq v \leq T$. Then, the corresponding optimal policy π_t of the real assets for the completed market is also optimal for the original incomplete market. In particular, the desired θ_v^u satisfying (12) and the resulting optimal policy π_t are independent of the specific choice of $\sigma^f(v,y)$, as long as it satisfies the orthogonal condition (7).

3. Optimal Policy for General Incomplete Market Models

In this section, we decompose the optimal policy under general incomplete market models (1) – (2) with CRRA and HARA utilities. We express components in the decomposition as conditional expectations of suitable random variables. The decomposition not only reveals the structure of the optimal policy, but also serves as an indispensable foundation for our further analysis under more concrete cases.

We first introduce some building blocks for our decomposition. Define the scalar function $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ as

$$\tilde{\mathcal{G}}_{t,T}(\theta^u) := (1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} \left(\xi_{t,T}\right)^{1-\frac{1}{\gamma}} + w^{\frac{1}{\gamma}} \int_t^T e^{-\frac{\rho s}{\gamma}} \left(\xi_{t,s}\right)^{1-\frac{1}{\gamma}} ds, \tag{13}$$

and the d–dimensional vector-valued function $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ as:

$$\tilde{\mathcal{H}}_{t,T}(\theta^u) := (1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} \left(\xi_{t,T}\right)^{1-\frac{1}{\gamma}} H_{t,T} + w^{\frac{1}{\gamma}} \int_t^T e^{-\frac{\rho s}{\gamma}} \left(\xi_{t,s}\right)^{1-\frac{1}{\gamma}} H_{t,s} ds. \tag{14}$$

In above, w, γ, ρ, and T are the utility parameters given in (4) or (5); $\xi_{t,T}$ is defined by (11); the term $H_{t,s}$ is a d–dimensional column vector given by

$$H_{t,s} = \int_t^s \left[L^r_{t,v} + L^\theta_{t,v}\theta^c_v\right] dv + \int_t^s L^\theta_{t,v} dW_v, \tag{15}$$

where θ^c_v is the total price of risk in (8).

In (15), we have $L^r_{t,v} = \mathcal{D}_t r(v, Y_v)$ and $L^\theta_{t,v} = \mathcal{D}_t \theta^c_v$, which are a d-dimensional vector and a $d \times d$-dimensional matrix, respectively. Here $\mathcal{D}_t$ denotes the time–t Malliavin derivative with respect to the Brownian motion W_t, which is introduced in Section EC.1.1 of the Electronic Companion. As a natural analogue to a classical derivative, we can intuitively understand the Malliavin derivative as the sensitivity to the underlying Brownian motion. See Appendix D of Detemple et al. 2003 for an accessible survey of Malliavin calculus in finance). Thus, $L^r_{t,v}$ and $L^\theta_{t,v}$ measure the impact of a time–t perturbation in the Brownian motion W_t on the time–v value of the interest rate and total price of risk, respectively. By (15), the term $H_{t,s}$ captures the cumulative impact over the horizon t to s.

For our decomposition of the optimal policy, it suffices to view $L^r_{t,v}$ and $L^\theta_{t,v}$ as standard diffusion processes with dynamics given in Section EC.1.1. As shown in Section 4.4, they facilitate the implementation of optimal policy. By (13) and (14), $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ depend on the unknown process of investor-specific price of risk θ^u_v for $v \in [t, T]$ via the dynamics of $\xi_{t,v}$ in (11). We highlight such dependence by the form $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$. In Section EC.4, we characterize the investor-specific price of risk θ^u_v, which also appears in the SDEs of the Malliavin derivatives, by the least favorable completion in Karatzas et al. (1991) and the dual optimization problem in He and Pearson (1991).

The following proposition establishes the decomposition of the optimal policy for general incomplete market models under the CRRA utility.

Proposition 1 *Under the incomplete market model (1) – (2) and the CRRA utility function given in (5), the optimal policy is wealth-independent. It can be decomposed as*

$$\pi_C(t, Y_t) = \pi^{mv}_C(t, Y_t) + \pi^h_C(t, Y_t),$$

where the mean-variance component $\pi^{mv}(t, Y_t)$ is explicitly given by

$$\pi^{mv}_C(t, Y_t) = \frac{1}{\gamma}(\sigma(t, Y_t)\sigma(t, Y_t)^\top)^{-1}(\mu(t, Y_t) - r(t, Y_t)1_m); \tag{16}$$

the hedge component $\pi^h(t, Y_t)$ follows by

$$\pi_C^h(t, Y_t) = -\left(1 - \frac{1}{\gamma}\right)(\sigma(t, Y_t)^+)^\top \frac{E_t[\tilde{\mathcal{H}}_{t,T}(\theta^u)]}{E_t[\tilde{\mathcal{G}}_{t,T}(\theta^u)]}, \tag{17}$$

where, throughout the paper, E_t denotes the expectation condition on the information up to time t; $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}$ are defined in (13) and (14), respectively. With $w > 0$, the optimal consumption and wealth-consumption ratio are given by

$$c_t = \frac{w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}}}{E_t[\tilde{\mathcal{G}}_{t,T}(\theta^u)]} X_t \;\; \text{and} \;\; \phi_C(t, Y_t) = w^{-\frac{1}{\gamma}} e^{\frac{\rho t}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}(\theta^u)]. \tag{18}$$

The investor-specific price of risk θ_v^u under the least favorable completion is also wealth-independent. It is characterized by the following equation:

$$\theta_v^u = \left(\sigma(v, Y_v)^+ \sigma(v, Y_v) - I_d\right)(1 - \gamma)\frac{E_v[\tilde{\mathcal{H}}_{v,T}(\theta^u)]}{E_v[\tilde{\mathcal{G}}_{v,T}(\theta^u)]}, \tag{19}$$

where I_d is the d–dimensional identity matrix.

Proof. See Section EC.4.2. □

Proposition 1 provides a decomposition of the optimal policy under CRRA utility (5) for general incomplete market models. It follows by applying the least favorable completion approach and then simplifying the results using the special structure of the CRRA utility. The optimal policy $\pi_C(t, Y_t)$ is decomposed into two components. The first component $\pi_C^{mv}(t, Y_t)$ is the mean-variance component. As reflected by the right-hand side of (16), it equals the product of the inverse covariance matrix $(\sigma(t, Y_t)\sigma(t, Y_t)^\top)^{-1}$ and the excess return $\mu(t, Y_t) - r(t, Y_t)1_m$, and further divided by the investor risk aversion level γ. The mean-variance component is "myopic" under the CRRA utility, as it is independent of the investor horizon and future market state. The second component $\pi_C^h(t, Y_t)$ is the hedge component for future investment opportunities. It can be further decomposed into two parts, which hedge the uncertainty in interest rate and price of risk respectively (see, e.g., Detemple et al. 2003). As our study focuses on the wealth effect in optimal portfolio allocation, here we do not separate the two hedge components in (17) to ease exposition.

The decomposition in Proposition 1 clearly reveals the wealth-independent property under the CRRA utility, which is widely noticed in the literature under both general set-up and specific models (e.g., Wachter 2002, Detemple et al. 2003, Liu 2007). The wealth-independent property of the CRRA utility is reflected by two aspects. First, the wealth level X_t does not appear in the components (16) – (17) as well as the wealth-consumption ratio (18). Second, as we show in Section EC.4.2, the functions $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ in (17) and (18) are also wealth-independent under the CRRA utility, ensuring the wealth level X_t does not affect the optimal policy implicitly via them. It is essentially because the investor-specific price of risk θ_v^u does not depend on investor wealth under the CRRA utility, which can be verified by (19).

Next, the following theorem establishes the decomposition of optimal policy under the HARA utility (4) for general incomplete market models (1) – (2).

Theorem 1 *Under the incomplete market model (1) – (2) and the HARA utility (4) with both terminal wealth and intermediate consumption (i.e., $w \in (0,1)$), the optimal policy is given by $\pi_H(t, X_t, Y_t) = \pi_H^{mv}(t, X_t, Y_t) + \pi_H^h(t, X_t, Y_t)$. The mean-variance component $\pi_H^{mv}(t, X_t, Y_t)$ satisfies*

$$\pi_H^{mv}(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t} \pi_C^{mv}(t, Y_t), \tag{20}$$

where $\pi_C^{mv}(t, Y_t)$ is the CRRA mean-variance component in (16); $\bar{X}_t$ is given by

$$\bar{X}_t = X_t - Z_{t,T} \tag{21}$$

with

$$Z_{t,T} := \bar{x}_T E_t \left[\xi_{t,T} \right] + \int_t^T \bar{c}_s E_t \left[\xi_{t,s} \right] ds. \tag{22}$$

The hedge component $\pi_H^h(t, X_t, Y_t)$ follows by:

$$\pi_H^h(t, X_t, Y_t) = -(\sigma(t, Y_t)^+)^\top \frac{\bar{X}_t}{X_t} \frac{E_t[\tilde{\mathcal{H}}_{t,T}(\theta^u)]}{E_t[\tilde{\mathcal{G}}_{t,T}(\theta^u)]} - (\sigma(t, Y_t)^+)^\top \frac{\Psi(t, X_t, Y_t)}{X_t}, \tag{23}$$

where $\Psi(t, X_t, Y_t)$ is a d–dimensional column vector, defined as:

$$\Psi(t, X_t, Y_t) := \bar{x}_T E_t \left[\xi_{t,T} H_{t,T} \right] + \int_t^T \bar{c}_s E_t \left[\xi_{t,s} H_{t,s} \right] ds. \tag{24}$$

The optimal consumption is given by

$$c_t = \bar{c}_t + \frac{w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}}}{E_t[\tilde{\mathcal{G}}_{t,T}(\theta^u)]} \bar{X}_t. \tag{25}$$

The functions $\tilde{\mathcal{G}}_{t,T}(\theta^u)$, $\tilde{\mathcal{H}}_{t,T}(\theta^u)$, and $H_{t,s}$ are still defined by (13), (14), and (15) respectively, except that we now plug in the investor-specific price of risk θ_v^u for the HARA investor. It is characterized by the following equation:

$$\theta_v^u = \left(\sigma(v, Y_v)^+ \sigma(v, Y_v) - I_d \right) \left[(1-\gamma) \frac{E_v[\tilde{\mathcal{H}}_{v,T}(\theta^u)]}{E_v[\tilde{\mathcal{G}}_{v,T}(\theta^u)]} - \frac{\gamma}{\bar{X}_v} \Psi(v, X_v, Y_v) \right], \tag{26}$$

which also depends on investor wealth X_v. In the case with only terminal wealth (resp. intermediate consumption) in (4), the above results still follow except for dropping the terms related to $\bar{c}_s$ (resp. $\bar{x}_T$) in (22) and (24).

Proof. See Section EC.4.3. □

Theorem 1 develops a decomposition of optimal policy under HARA utility for general incomplete market models, which is novel in the literature. Importantly, it reveals how the investor wealth affects the optimal policy when we move from the wealth-independent CRRA utility to the wealth-dependent HARA utility. In general incomplete market models, the investor wealth level X_t impacts the optimal policy under

HARA utility via two channels. First, by (20) and (23), the wealth level explicitly appears in both the mean-variance and hedge components via the multiplier $\bar{X}_t/X_t$. Second, under the HARA utility, the investor-specific price of risk θ_v^u becomes wealth-dependent. It can be seen by Equation (26), as a wealth-related term $\gamma\Psi(v, X_v, Y_v)/\bar{X}_v$ is now involved in the right-hand side. Economically, it means that the fictitious assets used by the HARA investor to complete the market is impacted by her wealth level. Then by (10), (13), and (14), the state price density $\xi_{t,s}$ and the functions $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ also depend on investor wealth under the HARA utility. Thus, the fictitious completion introduces an implicit channel for the wealth level to affect the HARA optimal policy. It strongly contrasts with the optimal policy under the CRRA utility, in which the investor-specific price of risk and the building blocks are all wealth-independent.

To interpret the structure of optimal policy under the HARA utility, we first focus on the term $Z_{t,T}$ in (22). By He and Pearson (1991), the conditional expectation $E_t\left[\xi_{t,s}\right]$ is the time–t present value of a unit payment at time s under the equivalent martingale measure characterized by the process θ_v^u. As such, the term $Z_{t,T}$ in (22) represents the present value at time t of all future subsistence requirements of the HARA investor, including the parts for both the terminal wealth, $\bar{x}_T E_t\left[\xi_{t,T}\right]$, and intermediate consumption $\int_t^T \bar{c}_s E_t\left[\xi_{t,s}\right] ds$. Then, $\bar{X}_t = X_t - Z_{t,T}$ is the HARA investor remaining wealth after she first sets aside $Z_{t.T}$ amount of her wealth to satisfy future subsistence requirements. When there is no subsistence requirement, i.e., with $\bar{x}_T = 0$ and $\bar{c}_s \equiv 0$, the term $Z_{t,T}$ vanishes under the CRRA utility.

We have the following observations on how the HARA optimal policy in Theorem 1 differs with its CRRA counterpart in Proposition 1. First, by (20) , the mean-variance components $\pi_H^{mv}(t, X_t, Y_t)$ and $\pi_C^{mv}(t, Y_t)$ satisfy a simple ratio relationship with a wealth-related multiplier $\bar{X}_t/X_t = 1 - Z_{t,T}/X_t$. It can be interpreted as follows. The HARA investor first sets aside $Z_{t.T}$ amount of her wealth to satisfy her future subsistence requirements. Then, she constructs the mean-variance component for her portfolio using the remaining wealth $\bar{X}_t$ just like a CRRA investor. Such a structure reflects the complete intolerance for violation of subsistence requirements under the HARA utility. In light of the relationship for mean-variance components in (20), we define an equivalent relative risk aversion level for HARA investors as

$$\gamma_H(X_t) := \frac{\gamma X_t}{\bar{X}_t} = \gamma\left(1 - \frac{Z_{t,T}}{X_t}\right)^{-1}. \tag{27}$$

Then by (16) and (20), we have

$$\pi_H^{mv}(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t}\pi_C^{mv}(t, Y_t) = \frac{1}{\gamma_H(X_t)}(\sigma(t, Y_t)\sigma(t, Y_t)^\top)^{-1}\left(\mu(t, Y_t) - r(t, Y_t)1_m\right).$$

That is, the mean-variance component of a HARA investor with wealth X_t and relative risk aversion γ coincides with that of a CRRA investor who has a relative risk aversion of $\gamma_H(X_t)$. Thus, $\gamma_H(X_t)$ captures how the wealth level affects the risk aversion of the HARA investor. A similar measure is also used in Detemple and Rindisbacher (2010) for analyzing the HARA optimal policy under complete market models.

Next, we compare the hedge components $\pi_H^h(t, X_t, Y_t)$ and $\pi_C^h(t, Y_t)$ under the two utilities. Unlike the mean-variance component, the hedge component is more complicated under the HARA utility. By (17) and (23), we see the HARA hedge component differs from its CRRA counterpart in three aspects. First, the first term in $\pi_H^h(t, X_t, Y_t)$ includes the wealth-related multiplier $\bar{X}_t/X_t$, which can be interpreted in the same way as that for the mean-variance component. Second, the hedge component $\pi_H^h(t, X_t, Y_t)$ includes an additional term $(\sigma(t, Y_t)^+)^\top \Psi(t, X_t, Y_t)/X_t$ under the HARA utility. As discussed in Section EC.4.3, this additional term essentially hedges the uncertainty in the present value of investor future subsistence requirements $Z_{t,T}$. Accordingly, it vanishes under the CRRA utility when $\bar{x}_T = 0$ and $\bar{c}_s = 0$. Finally, the functions $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ in HARA hedge component (23) can be wealth-dependent and thus may differ from their CRRA counterparts in (17). It is essentially because the investor-specific price of risk θ_v^u may differ for CRRA and HARA investors in general incomplete market models.

By (25), the optimal consumption under the HARA utility contains two parts. The first part is the subsistence requirement for intermediate consumption $\bar{c}_t$, which must be satisfied under the HARA utility. The second part is the "surplus" consumption based on the remaining wealth $\bar{X}_t$, after subtracting the present value of all future subsistence requirements. By (25), we can write the wealth-consumption ratio under HARA utility as $\phi_H(t, X_t, Y_t) = X_t/c_t$. It is also wealth-dependent, which contrasts with the CRRA utility.

By (19) and (26), the investor-specific price of risk satisfies an integral-type equation system, in which the whole process of θ_v^u is involved. As shown in He and Pearson (1991), it can also be characterized by a complex second-order quasilinear PDE, which includes multiple products and quotients of first order partial derivatives of the unknown wealth process (Theorem 7 therein). Note that if we can solve θ_v^u explicitly, then the existence of solution to (19) or (26) is proved by construction. This is the case for the CIRH-SVSIR model used in our comparative analysis (see Section 4.4). For general models, we discuss the existence results for (19) or (26) at the end of Section EC.4.3 of the Electronic Companion. We refer to Section 4 of He and Pearson (1991) for sufficient conditions on the existence of solution.

4. Incomplete Market Models with Hedgeable Interest Rate Risk

In the previous section, we have established the decompositions of the optimal policy under HARA and CRRA utilities for general incomplete market models. The decompositions clearly demonstrate the structural impact of the wealth-dependent HARA utility on the optimal portfolio choice. However, the complex structure of the optimal policy makes it difficult to conduct further analysis. In the following, we proceed to a specific, but highly flexible class of incomplete market models, in which the interest rate risk is fully hedgeable by the risky assets. We show that in this case, the optimal policies under the HARA and CRRA utilities have a closed-form relationship with intuitive economic interpretations.

We begin by setting up the models considered in this section. Suppose that the state variable Y_t can be separated to two parts $Y_t = ((Y_t^r)^\top, (Y_t^o)^\top)^\top$, where Y_t^r (resp. Y_t^o) is an n_r (resp. $n- n_r$) dimensional column vector. We assume the interest rate only depends on the state variable Y_t^r but not Y_t^o, i.e., $r_t = r(t, Y_t^r)$. Thus, Y_t^r is the state variable governing the interest rate, while Y_t^o controls the other aspects of investment opportunities. Accordingly, suppose we can decompose the d–dimensional Brownian motion W_t into two parts $W_t = ((W_t^r)^\top, (W_t^o)^\top)^\top$, with W_t^r (resp. W_t^o) being a d_r (resp. $d - d_r$) dimensional standard Brownian motion. We assume the interest rate-related state variable Y_t^r satisfies the generic SDE:

$$dY_t^r = \alpha^r(t, Y_t^r)dt + \beta^r(t, Y_t^r)dW_t^r, \tag{28}$$

for an n_r–dimensional vector-valued function $\alpha^r(t, y^r)$ and an $n_r \times d_r$–dimensional matrix-valued function $\beta^r(t, y^r)$. By (28), Y_t^r is Markovian in itself and driven only by the d_r–dimensional Brownian motion W_t^r. Thus, the innovation in the interest rate $r(t, Y_t^r)$ only depends on the Brownian motion W_t^r, but not W_t^o. On the other hand, the second state variable Y_t^o follows the generic SDE:

$$dY_t^o = \alpha^o(t, Y_t)dt + \beta^o(t, Y_t)dW_t, \tag{29}$$

for an $(n - n_r)$–dimensional vector-valued function $\alpha^o(t, y)$ and an $(n - n_r) \times d$–dimensional matrix-valued function $\beta^o(t, y)$. The dynamics of Y_t^o can depend on the full state variable Y_t and the d–dimensional Brownian motion W_t.

In addition, suppose the m risky assets in the market can be decomposed into two sets $S_t = ((S_t^{(1)})^\top, (S_t^{(2)})^\top)^\top$, with $S_t^{(1)}$ and $S_t^{(2)}$ including d_r and $m - d_r$ risky assets. For the assets in $S_t^{(1)}$, we assume their prices follow the SDE:

$$\frac{dS_{it}^{(1)}}{S_{it}^{(1)}} = \mu_i^{(1)}(t, Y_t^r)dt + \sigma_i^{(1)}(t, Y_t^r)dW_t^r, \text{ for } i = 1, 2, ..., d_r; \tag{30}$$

where $\mu_i^{(1)}(t, y^r)$ is a scalar function for modeling the mean rate of return; $\sigma_i^{(1)}(t, y^r)$ is a d_r-dimensional vector-valued function for modeling the return volatility. Thus, the dynamics of the asset prices in $S_t^{(1)}$ only hinge on the interest rate-related state variable Y_t^r and the Brownian motion W_t^r. For the remaining $m - d_r$ risky assets in $S_t^{(2)}$, their prices satisfy the dynamics:

$$\frac{dS_{it}^{(2)}}{S_{it}^{(2)}} = \mu_i^{(2)}(t, Y_t)dt + \sigma_i^{(2)}(t, Y_t)dW_t, \text{ for } i = 1, 2, ..., m - d_r, \tag{31}$$

with scalar mean return function $\mu_i^{(2)}(t, y)$ and d-dimensional vector-valued volatility function $\sigma_i^{(2)}(t, y)$. As such, their price dynamics can depend on the full state variable Y_t and Brownian motion W_t. To ease exposition, we call assets in $S_t^{(1)}$ as the bond assets since they only involve the uncertainty in the interest rate process, and assets in $S_t^{(2)}$ as the stock assets in the market.

Since the market has d_r bond assets with returns driven by d_r independent Brownian motions in W_t^r, the uncertainty in the interest rate $r(t, Y_t^r)$ can be fully hedged by investing in these bond assets. Thus, the market for interest rate risk is complete, although the full market is not. With the bond assets, the market price of interest rate risk can be uniquely determined as:

$$\theta^r(t,Y_t^r)=\sigma^{(1)}(t,Y_t^r)^{-1}\big(\mu^{(1)}(t,Y_t^r)-r(t,Y_t^r)1_{d_r}\big), \tag{32}$$

which is a d_r–dimensional column vector. On the other hand, the stock market is still incomplete, as the uncertainty in W_t^o cannot be fully hedged. Thus, the investor still has to "complete" the market with fictitious assets when solving the portfolio allocation problem. We describe the least favorable completion under this set-up in Section EC.5.1. The above setting directly includes the models with nonrandom interest rate, under which we do not need the bonds to hedge the interest rate risk, i.e., $n_r=d_r=0$.

Although we assume a complete market for interest rate risk, the class of model we consider is general and flexible since we do not impose any assumptions on the dynamics of state variables and asset prices. Besides, the stock market is allowed to be incomplete with unhedgeable risk. The set-up in (28) – (31) covers a wide range of classical models in the literature, such as the CIR model and the Heston stochastic volatility model in Liu (2007), the mean-reverting return models in Kim and Omberg (1996) and Wachter (2002), and the stochastic volatility model in Moreira and Muir (2019).

We now proceed to establish the decomposition of HARA optimal policy under the set-up in (28) – (31). With a complete market for interest rate risk, the value of a zero-coupon bond (ZCB hereafter) with a given maturity can be fully determined by the absence of arbitrage principle. It is because the payoff of a ZCB can be perfectly replicated by investing in the savings account and the bond assets in the market. Denote by $B_{t,s}$ the time–t price of a ZCB with unit face value that matures at time s. We first establish the following proposition, which plays an indispensable role for decomposing the HARA optimal portfolio.

Proposition 2 *Under the set-up in (28) – (31) where the market for interest rate risk is complete, we have*

$$E_t[\xi_{t,s}]=B_{t,s}=E_t[\eta_{t,s}^r], \tag{33}$$

where $\eta_{t,s}^r$ is defined as:

$$\eta_{t,s}^r=\exp\left(-\int_t^s r_v dv-\int_t^s(\theta_v^r)^\top dW_v^r-\frac{1}{2}\int_t^s(\theta_v^r)^\top\theta_v^r dv\right), \tag{34}$$

with $\theta_v^r=\theta^r(v,Y_v^r)$ in (32). The ZCB price $B_{t,s}$ satisfies the following dynamics:

$$\frac{dB_{t,s}}{B_{t,s}}=\mu_B(t,Y_t^r;s)dt+\sigma_B(t,Y_t^r;s)dW_t^r. \tag{35}$$

The expressions of the drift $\mu_B(t,Y_t^r;s)$ and volatility $\sigma_B(t,Y_t^r;s)$ are explicitly given in (EC.5.4) and (EC.5.5) of Section EC.5.2. They do not depend on the investor-specific price of risk.

Proof. See Section EC.5.2. □

Proposition 2 reveals the role of hedgeable interest rate risk in solving the HARA optimal policy. As discussed in Section 3, the dynamics of state price density $\xi_{t,s}$ in (11) hinges on the investor-specific price of risk θ_v^u in general incomplete market models. Thus, its conditional expectation $E_t[\xi_{t,s}]$ may vary for different investors. However, with a complete market model for interest rate risk, we show that $E_t[\xi_{t,s}]$ can be uniquely pinned down by the ZCB price $B_{t,s}$, regardless of the investor utility function. Its economic interpretation is as follows. Recall that $E_t[\xi_{t,s}]$ represents the time–t present value of one unit payment at time s under the equivalent martingale measure determined by the investor-specific price of risk. When there is a complete market for interest rate risk, we can perfectly replicate this payment by investing in the bond assets with an initial cost of $B_{t,s}$ at time t. Thus, the present value $E_t[\xi_{t,s}]$ equals the ZCB price $B_{t,s}$. Moreover, the dynamics of $B_{t,s}$ is explicitly given by (35), which is independent of the investor-specific price of risk.

4.1. Economic Structure of HARA Optimal Policy

With the above preparation, the following theorem presents our main results on the decomposition of the HARA optimal policy under a complete market for interest rate risk.

Theorem 2 *Under the set-ups in (28) – (31) where the market for interest rate risk is complete, the investor specific price of risk θ_v^u coincides for HARA and CRRA investors with the same utility parameters w, ρ, γ, and T in (4) and (5). The HARA optimal policy is given by $\pi_H(t, X_t, Y_t) = \pi_H^{mv}(t, X_t, Y_t) + \pi_H^h(t, X_t, Y_t)$, where*

$$\pi_H^{mv}(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t} \pi_C^{mv}(t, Y_t). \tag{36}$$

Here $\pi_C^{mv}(t, Y_t)$ is the optimal mean-variance component for the corresponding CRRA investor with $\bar{x}_T = \bar{c}_s = 0$ in (4). The remaining wealth $\bar{X}_t$ is given by $\bar{X}_t = X_t - Z_{t,T}$ with

$$Z_{t,T} = \bar{x}_T B_{t,T} + \int_t^T \bar{c}_s B_{t,s} ds, \ \ \textit{for } w \in (0,1). \tag{37}$$

The hedge component follows by:

$$\pi_H^h(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t} \pi_C^h(t, Y_t) + \frac{1}{X_t} \begin{pmatrix} \Pi_B(t, Y_t^r) \\ 0_{m-d_r} \end{pmatrix}, \tag{38}$$

where $\pi_C^h(t, Y_t)$ is the optimal hedge component for the CRRA investor; $\Pi_B(t, Y_t^r)$ is a portfolio consisting solely of the d_r bonds, given by

$$\Pi_B(t, Y_t^r) = \left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} \left(\bar{x}_T B_{t,T} \sigma_B(t, Y_t^r; T)^\top + \int_t^T \bar{c}_s B_{t,s} \sigma_B(t, Y_t^r; s)^\top ds\right), \tag{39}$$

where $\sigma_B(t, Y_t; s)$ is the instantaneous volatility of the bond return in (35). The optimal consumption for the HARA investor is given by

$$c_t = \bar{c}_t + \bar{X}_t/\phi_C(t, Y_t), \tag{40}$$

where $\phi_C(t, Y_t)$ is the wealth-consumption ratio of the CRRA investor. In the case with only terminal wealth (resp. intermediate consumption) in (4), the above results still follow except for dropping the terms related to $\bar{c}_s$ (resp. $\bar{x}_T$) in (37) and (39).

Proof. See Section EC.5.3. □

Theorem 2 establishes a novel decomposition of the HARA optimal policy for the models where the interest rate risk is fully hedgeable. We can make the following observations on its structure, which show how the general decomposition in Theorem 1 simplifies under hedgeable interest rate risk.

First, the amount of wealth $Z_{t,T}$ to satisfy the investor future subsistence requirement, as given by (22) for general models, can be explicitly determined by (37) as the value of a hypothetical bond holding scheme. The bond holding scheme consists of $\bar{x}_T$ shares ZCBs maturing at T and a continuum of $\bar{c}_s\, ds$ shares ZCB maturing at s for all $s \in [t, T]$. Notice that the payments from this bond holding scheme exactly finance the subsistence requirements of the HARA investor by (4). It is not surprising as with a complete market for interest rate risk, a ZCB that delivers a unit payment at any time $s > t$ can be perfectly synthesized by the savings account and the d_r bond assets in the market. Thus, the present value of a fixed amount of future payment can be uniquely determined by the no-arbitrage principle. It leads to the expression of $Z_{t,T}$ in (37).

Second, as a key component for solving the HARA optimal policy, we show that with a complete market for interest rate risk, the investor-specific price of risk θ_v^u under HARA utility coincides with its CRRA counterpart, thus is also independent of the investor wealth level. To prove this, we apply the dual problem method in He and Pearson (1991), which characterizes the investor-specific price of risk by an optimization problem (see Haugh et al. 2006 for using the duality approach to evaluate the performance of different policies). With the same investor-specific price of risk, the HARA investor completes the market using the same fictitious assets as the corresponding CRRA investor. It guarantees that the state price density $\xi_{t,s}$ as well as the functions $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ in (13) and (14) are the same under the two utilities. This simplification is indispensable for developing the relationship between the optimal policies in (36) – (38).

Third, the mean-variance components of the CRRA and HARA satisfy a ratio relationship by (36). With $Z_{t,T}$ given by (37), we can explicitly calculate the wealth-related multiplier $\bar{X}_t/X_t$ as $\frac{\bar{X}_t}{X_t} = 1 - \frac{1}{X_t}\left(\bar{x}_T B_{t,T} + \int_t^T \bar{c}_s B_{t,s} ds\right)$, which increases concavely in investor current wealth X_t and approaches one as X_t goes to infinity. For the mean-variance component, the HARA investor first holds the hypothetical bond holding scheme to finance her future subsistence requirements, then allocates the remaining wealth

$\bar{X}_t$ exactly as a CRRA investor. It corroborates the interpretation discussed after Theorem 1. In addition, the equivalent risk aversion $\gamma_H(X_t)$ in (27) specifies to

$$\gamma_H(X_t) = \gamma \left[1 - \frac{1}{X_t}\left(\bar{x}_T B_{t,T} + \int_t^T \bar{c}_s B_{t,s} ds\right)\right]^{-1}. \tag{41}$$

Clearly, $\gamma_H(X_t)$ is high (resp. low) when the HARA investor wealth level is low (resp. high). It is consistent with the empirical observations that investor risk aversion tends to be low (resp. high) in the bull (resp. bear) market (see, e.g., Berrada et al. 2018 and Li et al. 2022). Moreover, (41) suggests that under the wealth-dependent HARA utility, the fluctuation in investor wealth can generate time-varying risk aversion, which is fully absent under the CRRA utility.

We then look into the HARA hedge component $\pi_H^h(t, X_t, Y_t)$. By (38), we can further decompose $\pi_H^h(t, X_t, Y_t)$ into two parts. The first part $(\bar{X}_t/X_t)\pi_C^h(t, Y_t)$ scales the CRRA counterpart $\pi_C^h(t, Y_t)$ by the multiplier $\bar{X}_t/X_t$. Thus, it can be interpreted in the same way as for the mean-variance component. The second part, given by $-(\sigma(t, Y_t)^+)^\top \Psi(t, X_t, Y_t)$ in (23) for general incomplete market models, can now be explicitly calculated as $\Pi_B(t, Y_t^r)$ in (39) under the assumption of hedgeable interest rate risk. This term only depends on the market state variable Y_t^r but not investor current wealth X_t. It is a portfolio purely of the d_r bond assets. As discussed momentarily, its role is to replicate the dynamics of the hypothetical bond holding scheme that exactly finances the future subsistence requirements.

From an economic aspect, we can view relationship (36) and (38) as a decomposition of the HARA policy that separates the roles of the state variable Y_t and the investor wealth level X_t. The state variable Y_t impacts the optimal policy via the CRRA policies $\pi_C^{mv}(t, Y_t)$ and $\pi_C^h(t, Y_t)$, as well as the bond portfolio $\Pi_B(t, Y_t^r)$. On the other hand, the investor wealth level X_t impacts the optimal policy only via the ratio $\bar{X}_t/X_t$ and the denominator $1/X_t$ in (38). In the limit as wealth X_t goes to infinity, the HARA policies $\pi_H^{mv}(t, X_t, Y_t)$ and $\pi_H^h(t, X_t, Y_t)$ converge to their CRRA counterparts, with $\bar{X}_t/X_t$ increasing to one and $1/X_t$ dropping to zero.

In light of the bond-stock set-up in (30) and (31), we can also decompose the optimal portfolio as $\pi_H(t, X_t, Y_t) = (\pi_H^{(bond)}, \pi_H^{(stock)})^\top$, where $\pi_H^{(bond)}$ and $\pi_H^{(stock)}$ denote the optimal policy on the bond and stock assets, respectively. Then, it is easy to verify

$$\pi_H^{(stock)} = \frac{\bar{X}_t}{X_t}\pi_C^{(stock)} \quad \text{and} \quad \pi_H^{(bond)} = \frac{\bar{X}_t}{X_t}\pi_C^{(bond)} + \frac{\Pi_B(t, Y_t^r)}{X_t}, \tag{42}$$

where $\pi_C^{(stock)}$ and $\pi_C^{(bond)}$ denote the optimal policies of a CRRA investor. That is, the HARA investor invests in the stocks exactly as a CRRA investor based on the remaining wealth $\bar{X}_t$. However, her bond portfolio has an additional term $\Pi_B(t, Y_t^r)/X_t$, which is used to synthesize the hypothetical bond holding scheme for the subsistence requirement.

Finally, (40) decomposes the optimal consumption c_t under the HARA utility into two components. The first part is the subsistence requirement $\bar{c}_t$, which must be satisfied under the HARA utility. The second part $\bar{X}_t/\phi_C(t, Y_t)$ coincides with the optimal consumption level of a CRRA investor with wealth level $\bar{X}_t$. Thus, the HARA investor first consumes the subsistence requirement $\bar{c}_t$, then makes additional consumption based on her remaining wealth $\bar{X}_t$ just as a CRRA investor. We can verify that the wealth-consumption ratio $\phi_H(t, X_t, Y_t) = X_t/c_t$ under HARA utility is higher than its CRRA counterpart $\phi_C(t, Y_t)$ if and only if $Z_{t,T} > \bar{c}_t\phi_C(t, Y_t)$, i.e., the present value of future subsistence requirements $Z_{t,T}$ is high relative to the current consumption threshold $\bar{c}_t$.

Next, the following proposition shows that we can decompose the optimal portfolio of HARA investors into a financing sub-portfolio and a CRRA sub-portfolio.

Proposition 3 *Under the set-ups in (28) – (31) where the market for interest rate risk is complete, the HARA investor optimal portfolio can be decomposed into two parts. The first one is a financing portfolio that exactly delivers the future subsistence requirements. It starts with initial wealth*

$$X_0^{(finan)} = Z_{0,T}, \tag{43}$$

with $Z_{0,T}$ given by (37). For the financing portfolio, the investor consumes at rate $\bar{c}_t$ for $t \in [0, T]$ and invests in the d_r bonds according to the policy

$$\pi_t^{(finan)} = \Pi_B(t, Y_t^r)/X_t^{(finan)}, \tag{44}$$

where $\Pi_B(t, Y_t^r)$ follows by (39). The terminal wealth from the financing portfolio satisfies $X_T^{(finan)} = \bar{x}_T$. The second portfolio is a CRRA optimal portfolio that starts with initial wealth

$$X_0^{(crra)} = X_0 - X_0^{(finan)} = X_0 - Z_{0,T}.$$

For the CRRA portfolio, the investor consumes at rate $X_t^{(crra)}/\phi_C(t, Y_t)$ for $t \in [0, T]$ and invests in both the bonds and stocks following the CRRA optimal policy $\pi_C(t, Y_t)$.

Proof. See Section EC.5.4. □

Proposition 3 reveals that the HARA optimal portfolio can be decomposed into two separate sub-portfolios. The first one is a financing portfolio. It invests in the bond assets only and delivers the payments that exactly match the subsistence requirements on intermediate consumption and terminal wealth over the future investment horizon. The investment policy for the financing sub-portfolio is given by $\Pi_B(t, Y_t^r)/X_t^{(finan)}$. As shown in the proof, the value of the financing sub-portfolio always coincides with the hypothetical bond holding scheme, i.e., $X_t^{(finan)} = Z_{t,T}$ for all $t \in [0, T]$. Thus, the role of the additional term $\Pi_B(t, Y_t^r)$ in (38) is to replicate the dynamics of the hypothetical bond holding scheme, which is only possible when the market for interest rate risk is complete. The second sub-portfolio is a pure CRRA optimal portfolio based on the investor remaining wealth $\bar{X}_t$. The HARA investor consumes the minimum requirement $\bar{c}_t dt$ from the financing portfolio, and consumes like a CRRA investor from the CRRA portfolio. It further reveals the economic structure of the HARA optimal portfolio.

Finally, we show that the decomposition of HARA optimal policy can be further simplified under nonrandom, but possibly time-varying, interest rate. To save space, we discuss the results in Section EC.1.2 of the Electronic Companion. In this case, the additional term $\Pi_B(t, Y_t^r)$ in the HARA hedge component $\pi_H^h(t, X_t, Y_t)$ disappears in (38), as the HARA investor no longer needs to hedge the uncertainty in the interest rate. As shown in Proposition EC.1, the optimal HARA policy is parallel to its CRRA counterpart with:

$$\pi_H(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t} \pi_C(t, Y_t). \tag{45}$$

It implies that we can decompose the portfolio allocation problem for a HARA investor into two stages under nonrandom interest rate: the composition of the risky asset portfolio (i.e., $\pi_C(t, Y_t)$) and the wealth proportion allocated to this portfolio (i.e., $\bar{X}_t / X_t$). The HARA investor current wealth level and future subsistence requirements only affect the second stage, but not the first one. We can interpret such a decomposition as a "two-fund separation theorem" for HARA investors, which is originally proposed under the Markowitz mean-variance framework in Tobin (1958). In our case, it is driven by the wealth effect under the HARA utility.

4.2. HARA Investors with Heterogeneous Initial Wealth

In this section, we apply our theoretical decomposition to analyze the portfolio allocation of heterogeneous investors. We consider a novel aspect of investor heterogeneity: HARA investors with different initial wealth levels. Our findings complement and contribute to the empirical studies on how wealth impacts investor investment behaviors (see, e.g., Wachter and Yogo 2010, Calvet and Sodini 2014). Importantly, we provide a rigorous micro-foundation for these empirical researches using the optimal decision from the investor utility maximization problem.

In the following, we assume the market for interest rate risk is complete, as in the set-ups (28) – (31). Consider two HARA investors with different initial wealth levels. The high-wealth HARA investor has an initial wealth of $X_0^{(h)}$, and the low-wealth HARA investor has an initial wealth of $X_0^{(l)}$ with $X_0^{(l)} < X_0^{(h)}$. The two investors have the same utility parameters in (4), including investment horizon T, weight on consumption w, utility discount level ρ, and risk aversion coefficient γ. Besides, we assume they have the same subsistence requirements on their consumption and terminal wealth, i.e., with same $\bar{x}_T$ and $\bar{c}_t$ for all t.[8] This set-up isolates the impact of initial wealth level on the optimal portfolio allocation of the two HARA investors. By Theorem 2, the two investors will have the same (CRRA) optimal policy when their wealth levels increase to infinity.

We investigate how the wealth processes $X_t^{(l)}$ and $X_t^{(h)}$ differ for the high- and low-wealth investors. As a benchmark, we first consider two CRRA investors with initial wealth $X_0^{(h)}$ and $X_0^{(l)}$. Under CRRA utility, the optimal policy $\pi_C(t, Y_t)$ and wealth-consumption ratio $\phi_C(t, Y_t)$ are the same for the two investors. By (3), we can verify they have the same instantaneous return of wealth dX_t/X_t. It leads to

$$X_t^{(h)}/X_t^{(l)} = X_0^{(h)}/X_0^{(l)} \tag{46}$$

for all t, i.e., the wealth ratio of the two CRRA investors stays constant over time.

For HARA investors, the relationship in (46) does not hold since their optimal policy $\pi_H(t, X_t, Y_t)$ and wealth-consumption ratio $\phi_H(t, X_t, Y_t)$ are wealth-dependent. Thus, the instantaneous wealth return rate may vary for the two HARA investors. However, we can show that a similar relationship holds for the remaining wealth of the two HARA investors, i.e., the amount of wealth after subtracting all future subsistence requirements. It is summarized in the following proposition.

Proposition 4 *Under the set-ups in (28) – (31) where the market for interest rate risk is complete, the ratio of the remaining wealth $\bar{X}_t^{(h)} = X_t^{(h)} - Z_{t,T}$ and $\bar{X}_t^{(l)} = X_t^{(l)} - Z_{t,T}$ of the high- and low-wealth HARA investors stays constant over time:*

$$\frac{\bar{X}_t^{(h)}}{\bar{X}_t^{(l)}} \equiv \frac{X_0^{(h)} - Z_{0,T}}{X_0^{(l)} - Z_{0,T}} > \frac{X_0^{(h)}}{X_0^{(l)}}, \ \forall t \in [0, T], \tag{47}$$

where $Z_{0,T}$ is defined by (37). Moreover, the difference in their wealth growth rates, after all subsistence requirements are subtracted, is explicitly given by

$$\frac{1}{T}\left[\ln\left(\frac{\bar{X}_T^{(h)}}{X_0^{(h)}}\right) - \ln\left(\frac{\bar{X}_T^{(l)}}{X_0^{(l)}}\right)\right] = \frac{1}{T}\left[\ln\left(1 - \frac{Z_{0,T}}{X_0^{(h)}}\right) - \ln\left(1 - \frac{Z_{0,T}}{X_0^{(l)}}\right)\right] > 0. \tag{48}$$

The difference is always positive and only depends on the initial wealth levels, investment horizon, and subsistence requirements.

Proof. See Section EC.5.5. □

Proposition 4 states that when the interest rate risk is fully hedgeable, the ratio between the remaining wealth of the two HARA investors, $\bar{X}_t^{(h)}/\bar{X}_t^{(l)}$, stays constant over time. We emphasize that this result holds regardless of the specific model dynamics (28) – (31) and the random realization of market scenarios (e.g., bull and bear regimes). Thus, it establishes a model-free relationship between the wealth processes of the two HARA investors. This relationship can be interpreted as follows based on our decomposition results. By Proposition 3, with hedgeable interest rate risk, the HARA optimal portfolio can be decomposed into a financing portfolio and a CRRA portfolio. For the two investors, their wealth allocated on the CRRA portfolio at time t is exactly given by $\bar{X}_t^{(h)}$ and $\bar{X}_t^{(l)}$. On the other hand, by the discussion for (46), the growth rates of their CRRA portfolios are always the same. Thus, the ratio $\bar{X}_t^{(h)}/\bar{X}_t^{(l)}$ stays constant over time. When there is no subsistence requirement, i.e., $\bar{x}_T = \bar{c}_s = 0$ in (4), we have $Z_{t,T} \equiv 0$ by (37) and the relation (47) reduces to (46) under the CRRA case.

The relationship in (47) has direct implication on the wealth gap of the low- and high-wealth HARA investors. To see this, we measure the HARA investor wealth growth rate using the terminal remaining wealth after all subsistence requirements are satisfied. Equation (48) then shows that the wealth growth rate is always higher for the high-wealth investor, i.e., $\ln(\bar{X}_T^{(h)}/X_0^{(h)}) > \ln(\bar{X}_T^{(l)}/X_0^{(l)})$. Moreover, we can explicitly calculate the difference by the right hand side of (48), which again does not depend on the underlying model dynamics and market scenarios. Consequently, the gap in the two investor terminal remaining wealth becomes larger than that of their initial wealth, i.e., $\bar{X}_T^{(h)}/\bar{X}_T^{(l)} > X_0^{(h)}/X_0^{(l)}$. The magnitude of such effect can be substantial. For example, consider two HARA investors with initial wealth $X_0^{(h)} = 10^6$ USD and $X_0^{(l)} = 5 \times 10^5$ USD. Suppose the total subsistence requirement is $Z_{0,T} = 10^5$ USD and the investment horizon is five years for both. By (48), the difference in their annualized wealth growth rates would be 2.4%, which is economically meaningful.

The above results suggest that the existence of subsistence requirements, which distinguishes HARA from CRRA utility, puts the low-wealth investor in a disadvantageous position and leads to a larger wealth gap at the end of investment horizon. That is, the high-wealth HARA investor always enjoys a higher wealth growth rate from optimal portfolio allocation. Such pattern is consistent with the empirical findings (e.g., Fagereng et al. 2020 and Bach et al. 2020). The wealth effect in optimal portfolio allocation contributes to the wealth return channel for explaining the wealth inequality. As a highlight of our study, the increase in wealth gap shown in Proposition 4 holds regardless of the underlying model dynamics or market scenarios.

In Section EC.1.2, we further investigate how the optimal portfolio differs for the two HARA investors under the special case of nonrandom interest rate. Denote the optimal policies of the high- and low-wealth investors by $\pi_t^{(h)}$ and $\pi_t^{(l)}$, respectively. We find that the ratio of their optimal portfolios, i.e., the amount of wealth allocated on the risky assets, stays constant over time, i.e.,

$$\frac{1_m^\top \pi_t^{(h)} X_t^{(h)}}{1_m^\top \pi_t^{(l)} X_t^{(l)}} \equiv \frac{X_0^{(h)} - Z_{0,T}}{X_0^{(l)} - Z_{0,T}}, \ \forall t \in [0, T]. \tag{49}$$

Again it holds regardless of the underlying model dynamics and market scenarios, as long as the interest rate is nonrandom. Furthermore, we show that the optimal policy ratio $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)})$ is larger when the interest rate is low or the investment horizon is short, i.e.,

$$\frac{\partial}{\partial r}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) < 0 \quad \text{and} \quad \frac{\partial}{\partial T}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) < 0.$$

Intuitively, a lower interest rate or shorter investment horizon makes the bond holding scheme more expensive, thus decreasing the weights on risky assets of both HARA investors. However, such effect is larger for the low-wealth investor, as her subsistence requirement is more binding. It enlarges the gap between the optimal policies from the two investors. To save space, we discuss the results in more details in Section EC.1.2 of the Electronic Companion.

4.3. Discussion on Complete Market Models

In this section, we briefly discuss the decomposition of HARA optimal policy under complete market models, i.e., the number of risky assets equals the number of Brownian motions. In this case, the innovations in both the interest rate and market price of risk can be fully hedged by the risky assets in the market. Thus, the investor do not need to complete the market with fictitious assets. The state price density $\xi_{t,s}$ in (11) simplifies to

$$d\xi_{t,s} = -\xi_{t,s}[r(s,Y_s)ds + \theta^h(s,Y_s)^\top dW_s], \tag{50}$$

which is explicitly determined by the risk-free rate $r(s,Y_s)$ and market price of risk $\theta^h(s,Y_s)$. Then, the building blocks $\mathcal{G}_{t,T}$, $\mathcal{H}_{t,T}$, and $H_{t,s}$ in (13) – (15) can be defined accordingly by plugging in $\xi_{t,s}$ with above dynamics and replacing the total price of risk θ_t^c by the market price of risk $\theta^h(t,Y_t)$ in (9a) for real assets. In complete market models, the investor-specific price of risk θ^u is not involved. Thus, the building blocks are the same for all investors regardless of their wealth and utility.

Technically, we can view the decomposition under complete market models as a special case of our incomplete market results with the above simplifications. Thus, our main decompositions in Proposition 1 and Theorems 1 and 2, the economic implications in Propositions 3 and 4, as well as the insights from the comparative analysis in Section 5, all hold under complete market models. However, there is a fundamental difference that highlights the challenges brought by the market incompleteness. In general incomplete market models, the decompositions hinge on the unknown investor-specific price of risk θ_v^u. It is characterized by a very complex forward-backward integral-type equation like the ones in (19) and (26), and may depend on investor's wealth and utility function. It makes solving the optimal policy under incomplete market models much more challenging than for complete market models (see, e.g., Detemple 2014).

The structure of HARA optimal policy is briefly analyzed in Detemple and Rindisbacher (2010) for complete market models. They develop a decomposition of optimal portfolio under the numeraire that uses discount bonds as account units. We emphasize that our decompositions and subsequent analysis are not simple extensions of the complete-market results. When the market is incomplete, the investors need to complete the market using fictitious assets. As we discussed following Theorem 1, the CRRA and HARA optimal policies cannot be directly related in general incomplete market models, as the unknown investor-specific price of risk θ_v^u for the fictitious assets may differ for the CRRA and HARA investors. Under the assumption that interest rate risk is fully hedgeable, we circumvent this obstacle by combining the dual problem method in He and Pearson (1991) and the least favorable completion method in Karatzas et al. (1991) (see the proof in Section EC.5.3). This technical contribution is essential in handling the market incompleteness and allows us to obtain the closed-form relationship between the CRRA and HARA optimal policy. Moreover, we apply our closed-form decomposition to reveal the wealth effects in the optimal portfolio allocation of HARA investors. These are fully absent in the work of Detemple and Rindisbacher (2010).

4.4. Implementation of HARA Optimal Policy

The decomposition developed so far not only reveals the economic structure of the HARA optimal portfolio, but also facilitates the implementation of the HARA optimal policy under specific models. In the literature, the study of optimal portfolio allocation under the wealth-dependent HARA utility and stochastic market environment is relatively rare, as the corresponding optimal policy is usually hard to solve. Our decomposition results provide a powerful remedy for this challenge under the models with a complete market for interest rate risk. By Theorem 2, we can conveniently obtain the HARA optimal policy from its CRRA counterpart, which is usually much easier to obtain. We discuss the general steps hereafter.

Suppose we can solve the optimal CRRA policy $\pi_C^{mv}(t, Y_t)$ and $\pi_C^h(t, Y_t)$ in closed-form or by numerical methods (e.g., Monte Carlo simulation). Then, we can solve the HARA optimal policy as follows. First, we can compute the value of the financing portfolio $Z_{t,T}$ by (37) using the bond price $B_{t,s}$. By Proposition 2, the bond price $B_{t,s}$ does not depend on the unknown investor-specific price of risk θ_v^u and is thus easy to compute. For example, standard Monte Carlo simulation can be used to evaluate the conditional expectation $E_t\left[\eta_{t,s}^r\right]$ using the SDE $d\eta_{t,v}^r = -\eta_{t,v}^r\left(r_v dv + (\theta_v^r)^\top dW_v^r\right)$, which is explicitly given. With the value of $Z_{t,T}$, the mean-variance component $\pi_H^{mv}(t, X_t, Y_t)$ under HARA utility directly follows by (36). Finally, we can solve the hedge component $\pi_H^h(t, X_t, Y_t)$ by (38). To obtain the additional term $\Pi_B(t, Y_t^r)$ in (39), the volatility $\sigma_B(t, Y_t^r; s)$ can be evaluated by a Monte Carlo simulation method based on (EC.5.5) in Section EC.5.2. The Malliavin derivatives therein can be viewed as random variables with explicit dynamics (see Detemple et al. 2003). In the above procedures, the unknown investor-specific price of risk is not involved. Thus, we do not need to solve the extremely complex forward-backward integral-type equation in (26). This greatly facilitates the implementation of the optimal policy under HARA utility.

In the following, we use our decomposition to solve the HARA optimal policy under a three-dimensional incomplete market model with both stochastic interest rate and volatility. It demonstrates the application potential of our theoretical results and serves as a foundation for our subsequent comparative analysis. We set up the model in the below.

Example 1 *The interest rate r_t follows a one-factor CIR process, given by*

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sigma_r\sqrt{r_t}dW_{1t}. \tag{51}$$

There is a zero-coupon bond asset in the market with maturity T_1. Under the CIR interest rate process, its price is given by $P_{t,T_1} = \exp\left(a(\tau_1) + b(\tau_1)r_t\right)$, which follows the dynamics:

$$dP_{t,T_1}/P_{t,T_1} = \left(r_t + b(\tau_1)\lambda_r\sigma_r^2 r_t\right)dt + b(\tau_1)\sigma_r\sqrt{r_t}dW_{1t}, \tag{52}$$

with $\tau_1 = T_1 - t$. The explicit forms of $a(\tau)$ and $b(\tau)$ are given in (EC.1.12a) and (EC.1.12b) of Section EC.1.3. In addition, the market has a stock asset with price S_t satisfying:

$$dS_t/S_t = (r_t + \lambda_v V_t)\,dt + \sqrt{V_t}dW_{2t}; \tag{53}$$

its variance V_t follows

$$dV_t = \kappa_v(\theta_v - V_t)dt + \sigma_v\sqrt{V_t}(\rho_v dW_{2t} + \sqrt{1-\rho_v^2}dW_{3t}). \tag{54}$$

In above, W_{1t}, W_{2t}, and W_{3t} are three independent standard one-dimensional Brownian motions. The positive parameters κ_r, θ_r, and σ_r (resp. κ_v, θ_v and σ_v) determine the rate of mean-reversion, the long-run mean, and the proportional volatility of the interest rate process r_t (resp. variance process V_t). The parameters λ_r and λ_v control the price of interest rate risk and volatility risk, respectively. The leverage effect parameter $\rho_v \in [-1, 1]$ measures the instantaneous correlation between the asset return and the change in its variance. We assume the Feller's condition holds: $2\kappa_r\theta_r > \sigma_r^2$ and $2\kappa_v\theta_v > \sigma_v^2$.

In the above Cox-Ingersoll-Ross-Heston stochastic volatility and stochastic interest rate (CIRH-SVSIR) model, there are two risky assets (a bond and a stock) driven by three independent Brownian motions. Thus the market is incomplete.[9] The interest rate r_t follows a CIR process. The shocks to interest rate can be fully hedged by the ZCB, which is driven by the same Brownian motion W_{1t}. The market price of risk of W_{1t} is uniquely determined by the bond as $\lambda_r\sigma_r\sqrt{r_t}$. The stock price S_t is driven by an independent Brownian motion W_{2t}. That is, the returns of stock and bond are instantaneously uncorrelated. The variance process V_t is modeled by (54) as in a classical Heston-SV model (Heston 1993). The instantaneous correlation between the stock return and the variance innovation is ρ_v. When $\rho_v < 0$, the model captures the "leverage effect" in stock volatility. Besides, the market price of risk (Sharpe ratio) of the risky asset, $\lambda_v\sqrt{V_t}$ increases in the volatility, consistent with the empirical evidence (Campbell and Cochrane 1999). In the model, the shocks to the interest rate r_t and the variance V_t are assumed to be uncorrelated.

The CIRH-SVSIR model is also studied in Liu (2007), which derives the closed-form optimal policy under the wealth-independent CRRA utility over terminal wealth. We now consider the optimal portfolio choice problem for an investor with the wealth-dependent HARA utility. Without loss of generality, we assume the investor investment horizon is shorter than the maturity of the bond in the market, i.e., $T < T_1$. The CIRH-SVSIR mode fits into our general set-up in (28) – (31) as follows. The full state variable includes the interest rate and variance, i.e., $Y_t = (r_t, V_t)$; and the interest rate-related state variable Y_t^r is simply r_t itself. Since the interest rate risk is fully hedgeable, we can apply our decomposition in Theorem 2 to solve the optimal HARA policy in closed-form, which is given in the following proposition.

Proposition 5 *Under the CIRH-SVSIR model in (51) – (54) and the HARA utility (4) over terminal wealth ($w = 0$), the optimal policy $\pi_H(t, X_t, r_t, V_t)$ can be solved in closed-form as follows. The optimal weight on bond P_{t,T_1} is given by*

$$\pi_H^{(bond)}(t, X_t, r_t, V_t) = \frac{\bar{X}_t}{X_t b(\tau_1)}\left(\frac{\lambda_r}{\gamma} + d_r(\tau)\right) + \frac{\bar{x}_T b(\tau)}{X_t b(\tau_1)}\exp\left(a(\tau) + b(\tau)r_t\right), \tag{55a}$$

where $\tau = T - t$ *and* $\tau_1 = T_1 - t$*; the remaining wealth* $\bar{X}_t = X_t - \bar{x}_T \exp\left(a(\tau) + b(\tau) r_t\right)$*. The optimal weight on stock* S_t *follows by*

$$\pi_H^{(stock)}(t, X_t, r_t, V_t) = \frac{\bar{X}_t}{X_t}\left(\frac{\lambda_v}{\gamma} + \rho_v \sigma_v d_v(\tau)\right). \tag{55b}$$

The functions $d_r(\tau)$ *and* $d_v(\tau)$*, as well as the CRRA optimal policies* $\pi_C^{(bond)}(t, r_t, V_t)$ *and* $\pi_C^{(stock)}(t, r_t, V_t)$*, are explicitly given in (EC.1.13a) – (EC.1.14b) of Section EC.1.3.*

Proof. See Section EC.5.6. □

We have the following observations on the HARA optimal policy under the CIRH-SVSIR model. First, the wealth level X_t impacts the optimal stock weight only via the multiplier $\bar{X}_t / X_t$ in (55b). Second, the optimal bond weight is decomposed to two parts in (55a), which are affected by the wealth level X_t via the multipliers $\bar{X}_t / X_t$ and $1/X_t$. These observations corroborate the results in Theorem 2. In addition, we see that the HARA optimal policy does not involve the current market state variables (r_t, V_t). As shown in Section EC.1.3, it is because the CRRA optimal policy is independent of the market state variables under the CIRH-SVSIR model, as also noted in Liu (2007). However, the path of r_t and V_t can affect the HARA optimal policy implicitly via the investor wealth X_t. The above comparisons reveal the wealth effect in the optimal portfolio allocation of HARA investors. We investigate such impacts using a comprehensive comparative study in the following section.

5. Comparative Analysis of Wealth-dependent Effects

5.1. Parameter Estimation

To get empirical validity in the subsequent analysis, we estimate the parameters of CIRH-SVSIR model based on the observed market data in recent years. The model has in total nine parameters to be estimated, including $(\kappa_r, \theta_r, \sigma_r)$ for the interest rate process, $(\kappa_v, \theta_v, \sigma_v)$ for the variance process, ρ_v for the correlation between the innovations to the stock and variance processes, as well as the two market price of risk λ_r and λ_v. The high-dimensional parameter space as well as the multi-asset nature of the model pose significant challenge in the estimation. We briefly discuss our estimation procedure in the following. The details of the estimation are documented in Section EC.2.1.

We employ the maximum likelihood estimation approach in Aït-Sahalia and Kimmel (2007) and Ait-Sahalia and Kimmel (2010), which is widely used for estimating continuous-time models with state variables. We use the SPDR S&P 500 ETF as the stock asset S_t in our model, as it is the most widely tracked index for US equity market. The instantaneous interest rate r_t and stock variance V_t are not directly observed in the market. Thus, we need to extract them from observable assets. As in Aït-Sahalia and Kimmel (2007) and Ait-Sahalia and Kimmel (2010), we use the VIX index and US treasury bonds to extract the underlying V_t and r_t, respectively. We detail the estimation procedure in Section EC.2.1. We estimate the model using

Table 1 Estimated parameters for the CIRH-SVSIR model.

κ_r	θ_r	σ_r	λ_r	κ_v	θ_v	σ_v	ρ_v	λ_v
0.3158	0.0144	0.0641	0.8312	6.5607	0.0317	0.6942	−0.7543	5.8785
(0.2238)	(0.0107)	(0.0011)	(56.891)	(1.4942)	(0.0079)	(0.0196)	(0.0104)	(2.5376)

the data from 2013/4/29 to 2019/12/31. The annualized parameter estimates and their standard errors are reported in Table 1.

We have the following observations. First, the large value of κ_v indicates that the variance process is highly mean-reverting, while the interest rate process is less so with a small κ_r. Second, the large negative value of ρ_v suggests a strong leverage effect, i.e., the changes in stock price and its variance are highly negatively correlated. In addition, the market is characterized by a mild long-run volatility $\sqrt{\theta_v} \approx 17.8\%$ and a relatively high risk premium for stock $\lambda_v = 5.8785$. These features are largely consistent with the bull U.S. financial market in the estimation horizon. Our main findings are robust to the parameter values.

5.2. Impact of Current Wealth on Optimal Policy

In this section, we apply the closed-form solution in Proposition 5 to investigate how the optimal policy is impacted by investor wealth under the CIRH-SVSIR model. We consider a HARA investor that maximizes her expected utility over terminal wealth. We set the investor risk aversion and investment horizon as $T-t=$ 10 years and $\gamma = 4$. For comparison, we also consider a CRRA investor with the same utility parameters but no subsistence requirement. The model parameters are provided in Table 1. Figure 1 shows the optimal stock (left) and bond (right) weights at different investor wealth levels. In each panel, the red dashed and blue circled curves show the optimal policies under the CRRA and HARA utilities.

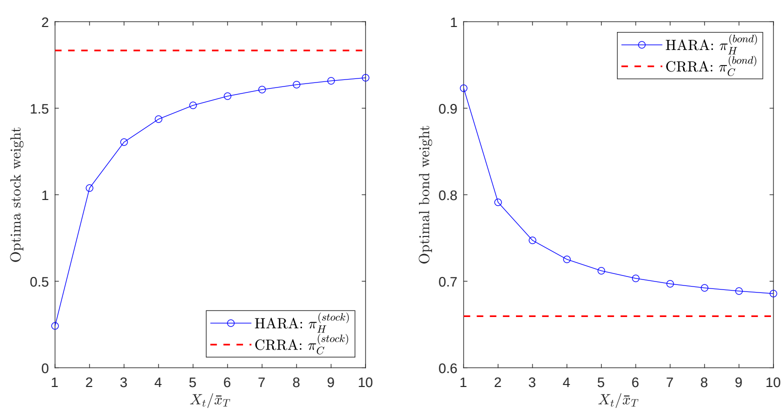

Figure 1 Optimal policy in the CIRH-SVSIR model by different investor wealth level: optimal stock weight (left panel) and optimal bond weight (right panel).

We first look into the optimal weight on the stock, which is explicitly given by (55b). It is impacted by the wealth level X_t only via the multiplier $\bar{X}_t/X_t = (1 - \bar{x}_T B_{t,T}/X_t)$, which increases concavely in X_t with $\lim_{X_t \to \infty} \bar{X}_t/X_t = 1$. This pattern is indeed observed in the left panel of Figure 1. The wealth effect is large: as X_t increases from $2\bar{x}_T$ to $10\bar{x}_T$, the optimal stock weight of the HARA investor increases from 103.9% to 167.6%. However, such effect diminishes as the wealth level becomes higher, i.e., as the subsistence constraint becomes less binding for the HARA investor. The concave increase in the allocation for risky asset with respect to investor wealth is empirically observed in the household finance literature; see, e.g., Roussanov (2010), Wachter and Yogo (2010), and Calvet and Sodini (2014). We show that this pattern can be explained by the optimal portfolio allocation under the wealth-dependent HARA utility.

Next, we check the optimal weight on the bond in the right panel, which is explicitly given by (55a). As discussed in Section 4.4, the optimal bond weight includes two parts, corresponding to the CRRA and financing sub-portfolios. We see that the optimal bond weight decreases in investor wealth level under the HARA utility. This decreasing pattern is driven by the second term in (55a), $\Pi_B(t, r_t)/X_t$, which is inversely proportional to the wealth level X_t. Intuitively, as the HARA investor becomes wealthier, the financing sub-portfolio for the subsistence requirement plays a less important role in her portfolio allocation. The magnitude of this wealth effect is also visible: the optimal bond weight decreases from 79.1% to 68.6% as the wealth level increases from $2\bar{x}_T$ to $10\bar{x}_T$.

The above analysis shows that the wealth effects in optimal portfolio allocation can be substantial. In Section EC.2.2, we further reveal how the optimal policy is affected by the investment horizon $T - t$ under the CRRA and HARA utilities. We find that under the HARA utility, the optimal stock weight increases in investment horizon due to two channels. First, a longer horizon increases the hedging demand of the investor. Second, a longer horizon decreases the ZCB price in the financing portfolio, leading to more remaining wealth allocated on the stock. The second channel is fully absent under the CRRA utility. We discuss the results in Section EC.2.2 of the Electronic Companion.

5.3. Cycle-dependence in Investment Decisions

In this section, we further reveal the wealth impact of HARA utility from a dynamic perspective. In particular, we investigate how the wealth effect interacts with the complex market dynamics over the investment horizon and affects the investor optimal portfolio allocation. Such dynamic analysis is rare in the literature on optimal portfolio choice (see the recent exception of Moreira and Muir 2019).

Consider a market under the CIRH-SVSIR model with parameters in Table 1. Without loss of generality, the initial levels are set as $r_0 = \theta_r = 0.0144$, $S_0 = 100$, $V_0 = \theta_v = 0.0317$. Assume there are two investors with HARA and CRRA utilities on the terminal wealth. Both investors have an investment horizon of $T = 10$ years. The HARA investor has an initial wealth of $X_0 = 4\bar{x}_T$. The risk aversion coefficient is set as $\gamma = 4$ for the HARA investor and $\gamma = 5.11$ for the CRRA investor. These choices lead to the same equivalent relative risk aversion $\gamma_H(X_0) = 5.11$ in (27) at the beginning of the horizon. Thus, the levels of optimal weights are comparable for the two investors. We simulate the market scenarios using a standard Euler's scheme on the dynamics (51) – (54), with a daily increment of $\Delta = 1/250$. Along the simulated path, we evaluate the optimal policies for the CRRA and HARA investors using the closed-form solutions in Proposition 5. The investor wealth levels then evolve according to Equation (3).

Figure 2 shows a representative market scenario. The upper panel plots the simulated path of the stock price S_t (black dotted with the left y-axis) and the corresponding optimal stock weights $\pi_H^{(stock)}$ and $\pi_C^{(stock)}$ of the HARA and CRRA investors (blue solid and red dashed with the right y-axis). The lower panel plots the path of the simulated stock price (black dotted with the left y-axis) and the equivalent risk aversion $\gamma_H(X_t)$ in (27) for the HARA investor (blue solid with the right y-axis). We classify the market regimes using the method in Lunde and Timmermann (2004): a transition from a bear to bull (resp. bull to bear) market is triggered when the stock price increases by 25% (resp. drops by 20%) from its lowest (resp. highest) level in the current regime. The bear markets are shown by the shaded areas of the figure. Not surprisingly, the price S_t plunges during bear markets.

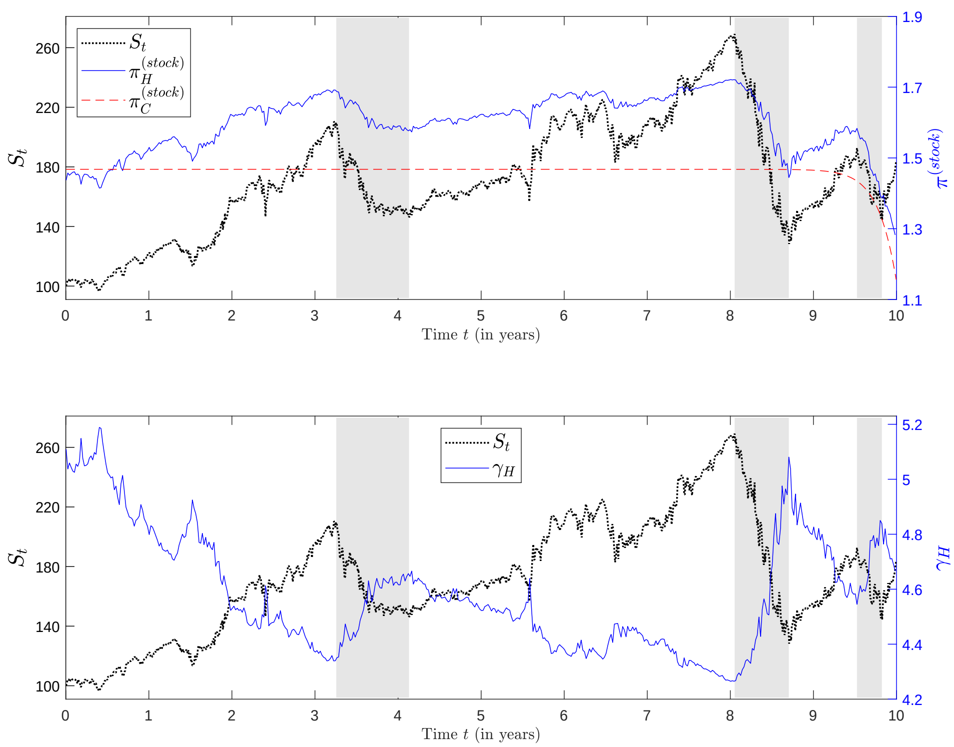

Figure 2 The upper panel plots a simulated path of stock price S_t, as well as the optimal stock weights $\pi_H^{(stock)}$ and $\pi_C^{(stock)}$ of HARA and CRRA investors. The lower panel plots the same simulated path of stock price S_t and the equivalent risk aversion $\gamma_H(X_t)$ in (27) for the HARA investor.

As discussed in Section EC.1.3, the CRRA optimal policy under the CIRH-SVSIR model is fully deterministic and independent of the market dynamics. It is seen in the upper panel: the optimal stock weight

$\pi_C^{(stock)}$ stays flat for most of the time, and only drops near the end of the investment horizon due to a smaller hedging demand. In contrast, the HARA investor stock weight $\pi_H^{(stock)}$ fluctuates significantly during the investment horizon. Moreover, it is positively correlated with the stock price S_t, implying the HARA investor tends to invest more (resp. less) on the stock during the bull (resp. bear) markets. Thus, the HARA investor is affected by the market cycles and invests in a procyclical way.[10] Such cycle-dependence essentially stems from the wealth-dependent property of the HARA utility, and is fully absent under the CRRA utility. As discussed in previous section, the HARA investor stock weight increases in her wealth level X_t, which hinges on the entire path of the market dynamics. During the bull markets, the investor wealth generally increases, leading to a larger position on the stock. Thus, the wealth effect of HARA utility provides a potential explanation for the procyclical investment behaviors of investors.

The cycle-dependence is also reflected by the HARA investor equivalent risk aversion level $\gamma_H(X_t)$, which is shown in the lower panel for the same simulated path. We see that the equivalent risk aversion is higher (resp. lower) during the bear (bull) market regimes. We interpret this as follows. In bear markets, the investor suffers a loss in her wealth, making her subsistence requirements more binding. This leads to a higher equivalent risk aversion by (27). Thus, the wealth-dependent property of HARA utility can endogenously generate the time-varying risk aversion of investors, which is fully absent under the CRRA utility. In particular, HARA investors become more (resp. less) risk averse in stressed (resp. bull) markets. It complements and contributes to the growing literature on time-varying risk preference (see, e.g., Guiso et al. 2018, Berrada et al. 2018, Li et al. 2022).

We use statistical tests to show that the patterns discussed above are not incidental results of a specific path. We run a large sample of $N_P = 10^4$ simulated path. Each path spans a horizon of ten years for the same HARA investor ($X_0/\bar{x}_T = 4$) as in Figure 2. First, we compute the correlation between daily stock price and the HARA optimal stock weight (resp. equivalent risk aversion level) for each path. From the 10^4 simulated paths, the average correlation between S_t and $\pi_t^{(stock)}$ (resp. $\gamma_H(X_t)$) is 0.71 (resp. -0.68); both statistically significant at the 0.1% level. We further do a regression analysis as follows. For the end of quarter q in path k, denote the stock price and HARA stock weight by $S_{k,q}$ and $\pi_{k,q}^{(stock)}$, respectively. We run the following regression using the observations from all quarters ($q = 1, 2, ..., 40$) and paths ($k = 1, 2, ..., 10^4$):

$$\frac{\pi_{k,q}^{(stock)}}{\tilde{\pi}_q^{(stock)}} = \beta_\pi \times \frac{S_{k,q}}{\tilde{S}_q} + \sum_{l=1}^{40} \lambda_l \mathbf{1}\{q = l\} + \varepsilon_{k,q}, \tag{56}$$

where $\tilde{\pi}_q^{(stock)} := \sum_{k=1}^{N_P} \pi_{k,q}^{(stock)}/N_P$ and $\tilde{S}_q := \sum_{k=1}^{N_P} S_{k,q}/N_P$ denote the average levels for quarter q across all paths. Thus, both $\pi_{k,q}^{(stock)}/\tilde{\pi}_q^{(stock)}$ and $S_{k,q}/\tilde{S}_q$ are normalized to have a mean of one for all q. The term $\sum_{l=1}^{40} \lambda_l \mathbf{1}\{q = l\}$ in (56) controls for the time fixed effects by assigning a coefficient λ_l for each quarter. Then, the coefficient β_π measures the (normalized) impact of stock price on optimal stock weight after their overall time trends are controlled. This addresses the concern that the positive correlation may appear because both stock price and weight are increasing over time. We also run regression (56) for the equivalent risk aversion γ_H in (27). We normalize γ_H for each quarter q in the same way as the optimal stock weight. That is, we plug in $\gamma_{H,k,q}/\tilde{\gamma}_{H,q}$ with $\tilde{\gamma}_{H,q} := \sum_{k=1}^{N_P} \gamma_{H,k,q}/N_P$ as the dependent variable.

For the HARA investor considered above, the estimated coefficient in (56) is $\beta_\pi = 0.094$ for optimal stock weight and $\beta_\gamma = -0.098$ for equivalent risk aversion level. Both coefficients are statistically significant at the 0.1% level. Thus, higher stock price is indeed associated with larger position on stock and lower risk aversion level of the HARA investor, even after the average time trend is controlled. This again validates the cycle-dependence in the portfolio allocation problem under the HARA utility. In contrast, both coefficients β_π and β_γ are zero in (56) for a CRRA investor. It is because the CRRA optimal policy is deterministic under the CIRH-SVSIR model, which can be fully explained by the time fixed effects.

We further investigate the cycle-dependence for HARA investors with different initial wealth levels. We find that the cycle-dependence is more pronounced for the low-wealth HARA investor than for the high-wealth one. That is, the fluctuations in the optimal stock weight and equivalent risk aversion are more dramatic for the low-wealth HARA investor (see a numerical example in Section EC.2.3). It is verified by the regression (56). In Figure 3, we plot the regression coefficients β_π and β_γ for HARA investors with different initial wealth level $X_0/\bar{x}_T$. All the coefficients are statistically significant. However, we see that their absolute values monotonically decrease in investor's initial wealth level. We interpret this as follows: the low-wealth HARA investor faces a more binding subsistence constraint, thus her optimal stock weight and equivalent risk aversion are more sensitive to the wealth level. Analytically, this is reflected by the concavity of the multiplier $\bar{X}_t/X_t$ in (55b). When the initial wealth level is very high, the HARA investor behaves like a CRRA investor, and the cycle-dependence vanishes in the portfolio allocation.

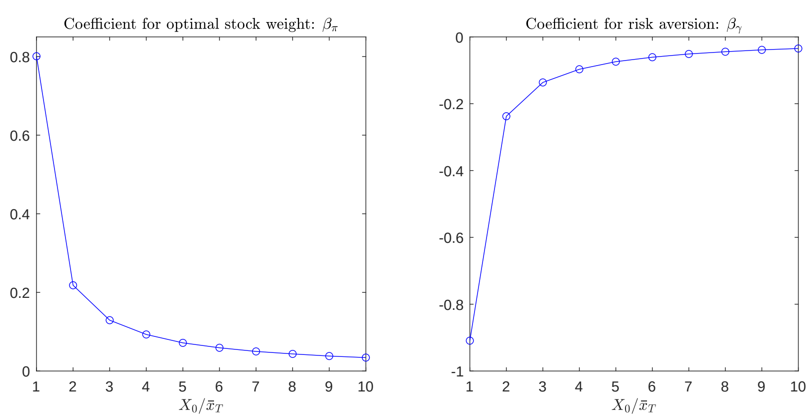

Figure 3 Regression coefficients β_π and β_γ in (56) for HARA investors with different initial wealth levels.

5.4. Wealth Effects on Investment Performance

In this section, we quantify the wealth effect on investor investment performance using the estimated CIRH-SVSIR model. Understanding such impact is practically important for setting return and risk targets in delegated portfolio management. For this question, analysis using one simulated path is insufficient, as it captures only one specific market scenario among possible others. This limitation is common in backtesting investment strategies. To overcome this challenge, we simulate a large number of paths for the investment problem of HARA investors with different initial wealth levels and evaluate their investment performance by averaging across the paths. Such large scale study takes possible market scenarios into account, and is facilitated by the closed-form solution in Proposition 5.

We compute the following performance metrics for each simulated path. Let $R_i = \ln(X_{i\Delta}/X_{(i-1)\Delta})/\Delta - r_{(i-1)\Delta}$ denote the annualized excess wealth return for day i. The excess return mean and return volatility are given by $\bar{R} = \sum_{i=1}^{N} R_i/N$ and $SD = \sqrt{\sum_{i=1}^{N}(R_i - \bar{R})^2/(N-1)}$ respectively, where $N = 2500$ denoting the total days in the horizon. Then, we check two risk measures of extreme losses in the investment horizon: the 99% level Value-at-Risk (VaR) of daily returns and the (percentage) maximum drawdown of investor wealth. The maximum drawdown is calculated as $MD = \max_{0\le n\le N}(1 - X_{n\Delta}/M_n^{\max})$, where $M_n^{\max} = \max_{0\le k\le n} X_{k\Delta}$ denotes the running maximum of investor wealth until day n (e.g., Zhang and Li 2023). The 99% VaR is obtained as the negative of the 25th lowest daily return in the path, given we have in total 2500 days. We also obtain the Sharpe ratio as $SR = \bar{R}/SD$, which represents the risk-adjusted return. After getting the performance metrics for each path, we then compute the average of these metrics across a large sample of 10^4 simulated paths. It allows us to estimate the ex-ante unconditional expectations over different realizations of market scenarios for each initial wealth level considered.

The results are displayed in Figure 4. The upper-left panel plots the average excess return mean $E\left[\bar{R}\right]$ (red squared) and volatility $E\left[SD\right]$ (blue circled). The upper-right panel presents the estimated maximum drawdown $E\left[MD\right]$ (blue circled with the right y-axis) and 99% level Value-at-Risk $E\left[VaR\right]$ (red squared with the left y-axis). The Sharpe ratio $E\left[SR\right]$ is shown in the lower-left panel. The lower-right panel plots the weighted stock return mean, which is discussed momentarily. Each point in the panel represents a given initial wealth level of the HARA investor, varying from $\bar{x}_T$ to $10\bar{x}_T$.

By the upper panels, we see that the investor initial wealth level substantially affects the investment performance in the expected direction. The HARA investor with higher initial wealth enjoys higher excess return, but also bears greater risk in terms of the volatility, maximum drawdown, and daily VaR. From this aspect, we can interpret the impact on investment performance as a risk-return trade-off triggered by the initial wealth of HARA investors. The magnitude of the trade-off is sizable: as the initial wealth X_0 increases from $\bar{x}_T$ to $10\bar{x}_T$, the average excess return mean increases from 12.5% to 28.4%, while the average return volatility increases from 14.3% to 31.6%. Similarly, the average maximum drawdown jumps

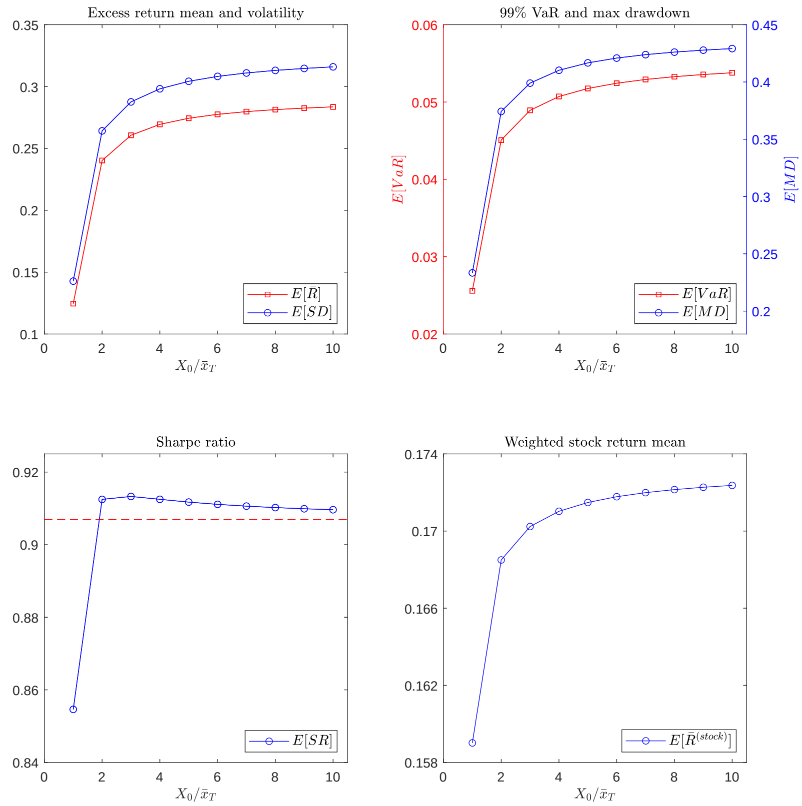

Figure 4 Performance statistics of HARA optimal dynamic portfolio with different initial wealth levels: Excess return mean and volatility (upper-left), 99% VaR and maximum drawdown (upper-right), Sharpe ratio (lower-left), and weighted stock return mean (lower-right).

from 23.3% to 42.9%, and the daily 99% VaR increases from 2.56% to 5.38%. We also find the 97.5% daily conditional VaR increases from 2.63% to 5.50%. The trade-off is due to the fact that high-wealth HARA investors allocate more wealth on the stock, which has much higher expected return and volatility than the bond.[11] The analysis highlights that it is crucial to understand the wealth-dependent feature of the investor utility in portfolio allocation.

Using 10^4 simulated paths, we find that the increasing patterns of excess return mean, volatility, maximum drawdown, and daily VaR are statistically significant. We run t-tests on the performance metrics of investors with different initial wealth levels, and show that the differences in means are all significant. The statistical test and results are described in Section EC.2.4 of the Electronic Companion. In addition, we see that the impact of initial wealth is more significant at low wealth levels but quickly decays as the initial wealth increases. It is because the optimal policy is more sensitive to wealth level for low-wealth HARA investors, as shown in Figure 1.

Finally, the lower-left panel plots the expected Sharpe ratio at different initial wealth levels, which measures the risk-adjusted return of the HARA investor. We see that the annualized Sharpe ratio jumps from 0.855 to 0.913 when the HARA investor's initial wealth increases from $X_0/\bar{x}_T = 1$ to $X_0/\bar{x}_T = 3$, and then experiences a very small drop afterwards (0.913 to 0.910). The dashed line in the panel denotes the Sharpe ratio of a CRRA investor, corresponding to the limiting case with $X_0/\bar{x}_T \to \infty$. Unlike the expected return and volatility, the Sharpe ratio is a normalized measure that adjusts for the investor's risk exposure. Thus, its change cannot be directly explained by the difference in the average stock positions of different HARA investors. In Section EC.2.5, we show that the increase in the Sharpe ratio can be attributed to two channels. First, high-wealth HARA investors benefit from a novel market timing effect that allows them to enjoy the higher stock risk premium during bear market periods. We explore this effect explicitly in the next section. Second, the optimal weights of high-wealth HARA investors are more stable over the investment horizon, as they are less sensitive to investor wealth level and market cycles. It reduces the time variation in the daily returns of high-wealth HARA investors. Both channels contribute to a higher Sharpe ratio. They are valid even after we adjust for the difference in the average stock weights of investors. We provide more numerical evidence and discussion on the two channels affecting the Sharpe ratio in Section EC.2.5.

5.5. Wealth-driven Market Timing Effect

In this section, we revel that the wealth-dependent HARA utility introduces a novel market timing effect in optimal portfolio allocation, which partly explains the increase in the Sharpe ratio. We compute the weighted stock (excess) return mean for each simulated path as

$$\bar{R}^{(stock)} = \frac{1}{N\Delta} \sum_{i=0}^{N-1} \frac{\pi_{i\Delta}^{(stock)}}{\bar{\pi}^{(stock)}} \times \left[\ln\left(\frac{S_{(i+1)\Delta}}{S_{i\Delta}} \right) - r_{i\Delta}\Delta \right],$$

which represents the average excess stock return weighted by the investor's stock position $\pi_{i\Delta}^{(stock)}/\bar{\pi}^{(stock)}$ on each day. Here we normalize the stock return by investor average stock weight (for a given path) $\bar{\pi}^{(stock)} = \sum_{i=1}^{N} \pi_{i\Delta}^{(stock)}/N$ in the denominator to account for the heterogeneity in the stock position of different investors. Thus, the weighted stock return mean primarily reveals the time variation effect in investor stock position, i.e., how well the investor "times" the market. In particular, we would have a larger $\bar{R}^{(stock)}$ if the investor tends to have larger stock weights during the periods with higher stock returns. We obtain the ex-ante unconditional expectation $E[\bar{R}^{(stock)}]$ by averaging $\bar{R}^{(stock)}$ across the 10^4 simulated paths.

The lower-right panel of Figure 4 plots $E[\bar{R}^{(stock)}]$ for investors with different initial wealth levels. Since the investors in our model do not intentionally time the market (e.g., buy low sell high), we may expect the weighted stock return $E[\bar{R}^{(stock)}]$ to be similar for different investors. However, we see that $E[\bar{R}^{(stock)}]$ increases monotonically in investor initial wealth, and the pattern is statistically significant (see Section EC.2.4). It suggests that HARA investors with higher initial wealth can better time the market: they are able to have larger stock position (relative to the path average) during the periods with high expected returns. As the initial wealth X_0 increases from $\bar{x}_T$ to $10\bar{x}_T$, the annualized weighted stock return mean changes from 15.9% to 17.2%, which is a 1.3% absolute annual increase. Such effect is economically significant, given we already adjust for the average stock position of investor.

The market timing effect essentially stems from the wealth-dependent property of the HARA utility. In the CIRH-SVSIR model, the stock risk premium is given by $\lambda_v V_t$ in (53), which is higher during volatile markets (when V_t is large). On the other hand, due to the leverage effect ($\rho_v < 0$), the stock price S_t tends to drop during high variance periods. It leads to potential loss in investor wealth, thus decreases their stock allocation according to the analysis in Section 5.2. Combining the two factors, we see that the procyclical behaviors make HARA investors less capable of benefiting from the high expected return under stressed markets. Such undesired effect is more prominent for investors with lower initial wealth, as their optimal policy is more sensitive to the wealth level. That is, the low-wealth HARA investors need to further decrease their stock positions during bear markets, which is the time with high risk premium of stocks. It explains the lower weighted stock return mean for them.

The market timing effect provides a novel channel for explaining the variation in wealth growth rates of investors with different wealth levels. In particular, high-wealth investors benefit from their ability to hold risky assets during stressed periods with high expected returns. It enlarges the wealth inequality among investors. Such timing effect and its implication on wealth inequality are empirically revealed in the recent work of Sakong (2022) using U.S. housing transaction and census data. It shows that poorer households consistently buy housing in booms and sell after a bust. This "buy-high-sell-low" channel leads to higher expected returns for wealthier households and enlarges the wealth inequality. We show that this channel can be generated by investor optimal portfolio allocation under the wealth-dependent HARA utility.

6. Conclusion

This paper establishes and implements a novel decomposition of the optimal policy under general incomplete-market diffusion models with the wealth-dependent HARA utility. The decomposition clearly reveals how the wealth level affects the HARA optimal portfolio under incomplete market models. We show that when the interest rate risk is hedgeable, we can connect the HARA optimal policy to its CRRA counterpart in closed-form. In particular, the HARA investors holds a CRRA optimal portfolio and a financing portfolio that exactly delivers the future subsistence requirements. We apply our decomposition to study the behaviors of HARA investors with heterogeneous wealth levels. We find that the high-wealth HARA investor always enjoys a higher wealth growth rate than the low-wealth one, regardless of the underlying model dynamics and realized market scenarios. It provides a potential channel for the increase in wealth inequality. Our decomposition facilitates the implementation of optimal policy via closed-form solution or numerical methods.

As an application, we explicitly solve the optimal policy for a HARA investor under an incomplete market model with both stochastic interest rate and volatility. We then conduct a comprehensive comparative study using parameters calibrated from the U.S. market data. Our study reveals the wealth effect in optimal portfolio allocation from various aspects. In particular, the wealth-dependent HARA utility introduces sophisticated cycle-dependence in the investor optimal strategies. It leads to a procyclical pattern in investor stock positions and lower (resp. higher) risk aversion in the bull (resp. bear) markets. We show that the initial wealth level of HARA investors can substantially impact their investment performance, leading to a risk-return tradeoff that stems from the wealth-dependent utility. We further identify a novel market timing effect: HARA investors with higher wealth can better time the market and benefit more from the high risk premiums during the volatile market periods. These findings are consistent with the empirical evidence on wealth inequality and investment behaviors of different investors.

We can adapt or generalize our optimal portfolio decomposition to other settings, e.g., the forward measure based representation considered in Detemple and Rindisbacher (2010). Moreover, it is interesting, among other possible extensions, to consider other (exotic) types of market incompleteness, e.g., the short-selling constraint or the "rectangular" constraint in Cvitanic and Karatzas (1992) and Detemple and Rindisbacher (2005), as well as the presence of jumps in, e.g., Jin and Zhang (2012). In addition, market incompleteness can arise due to uncertainty in labor income and wage (Mostovyi and Sîrbu 2020), which may have important implications on portfolio allocation. For the implementation of optimal policies, our theoretical decomposition might facilitate new numerical methods for incomplete market models and wealth-dependent utilities. From the economic perspective, a potential direction is to use empirical data to examine the implications on wealth-inequality and investor behaviors revealed in this paper. We defer these investigations, among others, to future research.

Endnotes

[1]The recursive utility generalizes the CRRA utility by separating risk aversion from elasticity of intertemporal substitution (Duffie and Epstein 1992). However, most applications with the recursive utility inherit the wealth-independent property of the CRRA utility (see, e.g., Chacko and Viceira 2005 and Moreira and Muir 2019).

[2]See, e.g., Kim and Omberg (1996) and Wachter (2002) for modeling stochastic market price of risk of the asset by using an Ornstein-Uhlenbeck model, Lioui and Poncet (2001) for considering stochastic interest rates by employing a constant-parameter instantaneous forward rate model, Liu et al. (2003) for studying impacts of event risk via affine stochastic volatility models with jumps, Liu (2007) for taking various stochastic environments (e.g., stochastic volatility) into account by modeling the asset return via quadratic affine processes, and Burraschi et al. (2010) for characterizing hedging components against both stochastic volatility and correlation risks under Wishart processes.

[3] We refer to the recent book of Dumas and Luciano (2017) for a survey of different numerical methods available for optimal portfolio choice.

[4]See, e.g., Kim and Omberg (1996), Wachter (2002), Liu (2007), and Moreira and Muir (2019) among others.

[5]This type of dynamic analysis is rare in the literature. A recent example is Moreira and Muir (2019), which show that, under a stochastic volatility model, ignoring the hedge component in the optimal policy leads to a substantial utility loss. However, the optimal policy in Moreira and Muir (2019) is independent of investor wealth level, contrasting with our focus on the wealth effect in portfolio allocation.

[6]To guarantee the martingale property of $\xi_t \exp(\int_0^t r_v dv)$, we assume that the total price of risk θ_v^c satisfies the Novikov condition: $E\left[\exp\left(\frac{1}{2}\int_0^T (\theta_v^c)^\top \theta_v^c dv\right)\right] < \infty$.

[7]The fictitious completion satisfying (12), i.e., zero weight for fictitious assets, must be the "least favorable" one among all possible completions as it is admissible in any other fictitious completions. See the discussion in Karatzas et al. (1991).

[8]Our results can be extended to the case where the two investors also differ in their subsistence requirements.

[9]In some cases, the volatility risk can be completed by financial derivatives. However, this may be unpractical for many assets, such as small cap stocks, stocks in emerging markets, crypto-currencies, or mutual funds.

[10]In contrast, we show in Section EC.2.3 that the HARA optimal bond weight is less sensitive to market cycles.

[11]In our investment horizon, the average annualized return of the stock and bond is 18.7% and 1.39%, respectively. The annualized volatility is 17.9% for the stock and 2.31% for the bond.

References

Aït-Sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics 83 (2), 413–452.

Ait-Sahalia, Y., Kimmel, R. L., 2010. Estimating affine multifactor term structure models using closed-form likelihood expansions. Journal of Financial Economics 98 (1), 113–144.

Amromin, G., Sharpe, S. A., 2014. From the horse's mouth: Economic conditions and investor expectations of risk and return. Management Science 60 (4), 845–866.

Bach, L., Calvet, L. E., Sodini, P., 2020. Rich pickings? risk, return, and skill in household wealth. American Economic Review 110 (9), 2703–2747.

Basak, S., Chabakauri, G., 2010. Dynamic mean-variance asset allocation. Review of Financial Studies 23 (8), 2970–3016.

Berrada, T., Detemple, J., Rindisbacher, M., 2018. Asset pricing with beliefs-dependent risk aversion and learning. Journal of Financial Economics 128 (3), 504–534.

Brandt, M. W., 2010. Portfolio choice problems. In: Aït-Sahalia, Y., Hansen, L. P. (Eds.), Handbook of Financial Econometrics, 1st Edition. Vol. 1. North-Holland, San Diego, pp. 269 – 336.

Burraschi, A., Porchia, P., Trojani, F., 2010. Correlation risk and optimal portfolio choice. Journal of Finance 65 (1), 393–420.

Calvet, L. E., Sodini, P., 2014. Twin picks: Disentangling the determinants of risk-taking in household portfolios. Journal of Finance 69 (2), 867–906.

Campbell, J. Y., Cochrane, J. H., 1999. By force of habit: A consumption-based explanation of aggregate stock market behavior. Journal of Political Economy 107 (2), 205–251.

Capponi, A., Rubtsov, A., 2022. Systemic risk-driven portfolio selection. Operations Research 70 (3), 1598–1612.

Carroll, C. D., Kimball, M. S., 1996. On the concavity of the consumption function. Econometrica 64 (4), 981–992.

Chacko, G., Viceira, L. M., 2005. Dynamic consumption and portfolio choice with stochastic volatility in incomplete markets. Review of Financial Studies 18 (4), 1369–1402.

Cox, J. C., Huang, C.-f., 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. Journal of Economic Theory 49 (1), 33–83.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

Cvitanic, J., Karatzas, I., 1992. Convex duality in constrained portfolio optimization. Annals of Applied Probability 2 (4), 767–818.

De Nardi, M., Fella, G., 2017. Saving and wealth inequality. Review of Economic Dynamics 26, 280–300.

Detemple, J., 2014. Portfolio selection: A review. Journal of Optimization Theory and Applications 161 (1), 1–21.

Detemple, J., Garcia, R., Rindisbacher, M., 2003. A Monte Carlo method for optimal portfolios. Journal of Finance 58 (1), 401–446.

Detemple, J., Rindisbacher, M., 2005. Closed-form solutions for optimal portfolio selection with stochastic interest rate and investment constraints. Mathematical Finance 15 (4), 539–568.

Detemple, J., Rindisbacher, M., 2010. Dynamic asset allocation: Portfolio decomposition formula and applications. Review of Financial Studies 23, 25–100.

Duffie, D., Epstein, L. G., 1992. Stochastic differential utility. Econometrica , 353–394.

Duffie, D., Fleming, W., Soner, H. M., Zariphopoulou, T., 1997. Hedging in incomplete markets with HARA utility. Journal of Economic Dynamics and Control 21 (4-5), 753–782.

Duffie, D., Pan, J., Singleton, K. J., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.

Dumas, B., Luciano, E., 2017. The Economics of Continuous-Time Finance. MIT Press, Cambridge.

Fagereng, A., Guiso, L., Malacrino, D., Pistaferri, L., 2020. Heterogeneity and persistence in returns to wealth. Econometrica 88 (1), 115–170.

Guiso, L., Sapienza, P., Zingales, L., 2018. Time varying risk aversion. Journal of Financial Economics 128 (3), 403–421.

Haugh, M. B., Kogan, L., Wang, J., 2006. Evaluating portfolio policies: A duality approach. Operations Research 54 (3), 405–418.

He, H., Pearson, N. D., 1991. Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. Journal of Economic Theory 54, 259–304.

He, X. D., Jiang, Z., 2020. Dynamic mean-variance efficient fractional kelly portfolios in a stochastic volatility model. Available at SSRN 3670621 .

Heston, S., 1993. A closed-form solution for options with stochastic volatility with applications to bonds and currency options. Review of Financial Studies 6, 327–343.

Jin, X., Zhang, A. X., 2012. Decomposition of optimal portfolio weight in a jump-diffusion model and its applications. Review of Financial Studies 25, 2877–2919.

Karatzas, I., Lehoczky, J. P., Shreve, S. E., 1987. Optimal portfolio and consumption decisions for a small investor on a finite horizon. SIAM Journal on Control and Optimization 25, 1557–1586.

Karatzas, I., Lehoczky, J. P., Shreve, S. E., Xu, G.-L., 1991. Martingale and duality methods for utility maximization in an incomplete market. SIAM Journal on Control and Optimization 29 (3), 702–730.

Karatzas, I., Shreve, S. E., 1991. Brownian Motion and Stochastic Calculus, 2nd Edition. Vol. 113 of Graduate Texts in Mathematics. Springer-Verlag, New York.

Kim, T., Omberg, E., 1996. Dynamic nonmyopic portfolio behavior. Review of Financial Studies 9, 141–161.

Li, H., Wu, C., Zhou, C., 2022. Time-varying risk aversion and dynamic portfolio allocation. Operations Research 70 (1), 23–37.

Lioui, A., Poncet, P., 2001. On optimal portfolio choice under stochastic interest rates. Journal of Economic Dynamics and Control 25, 1141–1865.

Liu, J., 2007. Portfolio selection in stochastic environments. Review of Financial Studies 20 (1), 1–39.

Liu, J., Longstaff, F., Pan, J., 2003. Dynamic asset allocation with event risk. Journal of Finance 58, 231–259.

Lunde, A., Timmermann, A., 2004. Duration dependence in stock prices: An analysis of bull and bear markets. Journal of Business & Economic Statistics 22 (3), 253–273.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7 (1), 77–91.

Merton, R. C., 1969. Lifetime portfolio selection under uncertainty: The continuous-time case. Review of Economics and Statistics 51, 247–257.

Merton, R. C., 1971. Optimum consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3, 373–413.

Moreira, A., Muir, T., 2019. Should long-term investors time volatility? Journal of Financial Economics 131 (3), 507–527.

Mostovyi, O., Sîrbu, M., 2020. Optimal investment and consumption with labor income in incomplete markets. Annals of Applied Probability 30 (2), 747–787.

Ocone, D., Karatzas, I., 1991. A generalized Clark representation formula, with application to optimal portfolios. Stochastics and Stochastics Reports 34, 187–220.

Penrose, R., 1955. A generalized inverse for matrices. Mathematical Proceedings of the Cambridge Philosophical Society 51, 406–413.

Roussanov, N., 2010. Diversification and its discontents: Idiosyncratic and entrepreneurial risk in the quest for social status. Journal of Finance 65 (5), 1755–1788.

Sakong, J., 2022. Cyclical transactions and wealth inequality. FRB of Chicago Working Paper (2022-05).

Samuelson, P., 1969. Lifetime portfolio selection by dynamic stochastic programming. Review of Economics and Statistics 51, 239–246.

Tobin, J., 1958. Liquidity preference as behavior towards risk. Review of Economic Studies 25 (2), 65–86.

Wachter, J. A., 2002. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. Journal of Financial and Quantitative Analysis 37, 63–91.

Wachter, J. A., 2010. Asset allocation. Annual Review of Financial Economics 2 (1), 175–206.

Wachter, J. A., Yogo, M., 2010. Why do household portfolio shares rise in wealth? Review of Financial Studies 23 (11), 3929–3965.

Zhang, G., Li, L., 2023. A general method for analysis and valuation of drawdown risk. Journal of Economic Dynamics and Control 152, 104669.

Dynamic Portfolio Allocation under Market Incompleteness and Wealth Effects — Electronic Companion

This Electronic Companion for "*Dynamic Portfolio Allocation under Market Incompleteness and Wealth Effects*" is organized as follows. Section EC.1 documents the auxiliary analytical results to our main text, including dynamics of the Malliavin-related terms in our portfolio decomposition (Section EC.1.1), HARA optimal policies under nonrandom interest rate (Section EC.1.2), and formulas under the CIRH-SVSIR model (Section EC.1.3). Section EC.2 provides the auxiliary numerical results, including estimation of the CIRH-SVSIR model (Section EC.2.1), impact of investment horizon on HARA policy (Section EC.2.2), analysis on cycle-dependence of HARA policy (Section EC.2.3), statistical tests for performance metrics (Section EC.2.4), and wealth effects on Sharpe ratio (Section EC.2.5). Section EC.3 documents the decomposition of the optimal policy under general incomplete market models and its proof. Sections EC.4 and EC.5 include the proof for the results in the main text.

Key words: optimal portfolio choice, incomplete market, wealth-dependent utility, closed-form analysis, wealth inequality, heterogeneous investors.

EC.1. Auxiliary Analytical Results

EC.1.1. Dynamics of the Malliavin Derivatives

In this section, we provide the dynamics of the Malliavin derivatives used in our optimal portfolio decompositions. As a natural analogue to a classical derivative, the Malliavin derivative measures the sensitivity with respect to the underlying Brownian motion. See the book Nualart (2009) and Appendix D of Detemple et al. (2003) for accessible surveys of Malliavin calculus in finance.

For an n–dimensional vector-valued function F_s, its Malliavin derivative $\mathcal{D}_t F_s$ with respect to the d–dimensional Brownian motion W_t is a $d \times n$ matrix with $\mathcal{D}_t F_s = ((\mathcal{D}_{1t} F_s)^\top, (\mathcal{D}_{2t} F_s)^\top, \cdots, (\mathcal{D}_{dt} F_s)^\top)^\top$, where each $\mathcal{D}_{it} F_s$ is an n–dimensional column vector representing the Malliavin derivative with respect to the ith Brownian motion W_{it}. Let $\mathcal{D}_t Y_s = ((\mathcal{D}_{1t} Y_s)^\top, (\mathcal{D}_{2t} Y_s)^\top, \cdots, (\mathcal{D}_{dt} Y_s)^\top)^\top$ denote the time–t Malliavin derivative of the time–s state variable Y_s. By Nunno et al. (2008), the dynamics of $\mathcal{D}_{it} Y_s$ can be derived from the SDE of Y_s in (2), which is given by:

$$d\mathcal{D}_{it} Y_s = (\nabla \alpha(s, Y_s))^\top \mathcal{D}_{it} Y_s ds + \sum_{j=1}^{d} \left(\nabla \beta_j(s, Y_s)\right)^\top \mathcal{D}_{it} Y_s dW_{js}, \ \lim_{s \to t} \mathcal{D}_{it} Y_s = \beta_i(t, Y_t). \quad \text{(EC.1.1)}$$

In above, $\beta_j(s, y)$ represents the jth column of matrix $\beta(s, y)$; W_{js} is the jth dimension of Brownian motion W_s; ∇ denotes the gradient of functions with respect to the arguments in the place of Y_s. For an

$m-$dimensional vector-valued function $f(t,y) = (f_1(t,y), f_2(t,y), \cdots, f_m(t,y))$, its gradient is an $n \times m$ matrix with each element given by $[\nabla f(t,y)]_{ij} = \partial f_j / \partial y_i(t,y)$, for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

For the $H_{t,s}$ term in (15), it satisfies the SDE:

$$dH_{t,s} = \left[L^r_{t,s} + L^\theta_{t,s}\theta^c_s\right] ds + L^\theta_{t,s} dW_s,$$

with initial value $H_{t,t} = 0$. As shown momentarily in Theorem EC.1, here the terms $L^r_{t,s}$ and $L^\theta_{t,s}$ are given by

$$L^r_{t,s} = \mathcal{D}_t r(s, Y_s) \text{ and } L^\theta_{t,s} = \mathcal{D}_t \theta^c_s. \tag{EC.1.2}$$

Thus, $L^r_{t,s}$ (resp. $L^\theta_{t,s}$) denotes the time–t Malliavin derivative of the interest rate (resp. total market price of risk). Using chain rule of Malliavin derivative, we further have

$$\mathcal{D}_t r(s, Y_s) = (\mathcal{D}_t Y_s) \nabla r(s, Y_s), \tag{EC.1.3a}$$

and

$$\mathcal{D}_t \theta^c_s = \mathcal{D}_t \theta^h_s + \mathcal{D}_t \theta^u_s = (\mathcal{D}_t Y_s) \nabla \theta^h(s, Y_s) + \mathcal{D}_t \theta^u_s. \tag{EC.1.3b}$$

In (EC.1.3b), $\theta^h(s, Y_s)$ is the market price of risk defined in (9a); $\mathcal{D}_t \theta^u_s$ denotes the Malliavin derivative of the investor-specific price of risk θ^u_s. It depends on the explicit form of θ^u_s, as characterized by the equation system (19) or (26).

Finally, in the models with hedgeable interest rate risk developed in Section 4, the volatility of the ZCB price $B_{t,s}$ is given by (EC.5.5) in Section EC.5.2 as

$$\sigma_B(t, Y^r_t; s) = -\frac{1}{B_{t,s}} E_t\left[\eta^r_{t,s}\left(\int_t^s M^r_{t,v} dv + \int_t^s M^\theta_{t,v}\left(dW^r_v + \theta^r_v dv\right)\right)^\top\right].$$

Here we have

$$M^r_{t,v} = \mathcal{D}^r_t r(v, Y^r_v) \text{ and } M^\theta_{t,v} = \mathcal{D}^r_t \theta^r\left(v, Y^r_v\right), \tag{EC.1.4}$$

where $\mathcal{D}^r_t$ denotes the time–t Malliavin derivative with respect to the d_r–dimensional Brownian motion W^r_t. Using the chain rule, they can be explicitly expressed as

$$\mathcal{D}^r_t r(v, Y^r_v) = (\mathcal{D}^r_t Y^r_v) \nabla r(v, Y^r_v) \text{ and } \mathcal{D}^r_t \theta^r\left(v, Y^r_v\right) = (\mathcal{D}^r_t Y^r_v) \nabla \theta^r(v, Y^r_v),$$

where $\mathcal{D}^r_t Y^r_v$ denotes the time–t Malliavin derivative of the state variable Y^r_v that governs the interest rate. As shown in (28), the dynamics of Y^r_v only depends on the d_r–dimensional Brownian motion W^r_v. We have $\mathcal{D}^r_t Y^r_v = ((\mathcal{D}^r_{1t} Y^r_v)^\top, (\mathcal{D}^r_{2t} Y^r_v)^\top, \cdots, (\mathcal{D}^r_{d_r t} Y^r_v)^\top)^\top$. Each $\mathcal{D}^r_{it} Y^r_v$ is an n_r–dimensional column vector and satisfies the following SDE:

$$d\mathcal{D}^r_{it} Y^r_v = \left(\nabla \alpha^r(v, Y^r_v)\right)^\top \mathcal{D}^r_{it} Y^r_v dv + \sum_{j=1}^{d_r} \left(\nabla \beta^r_j(v, Y^r_v)\right)^\top \mathcal{D}^r_{it} Y^r_v dW^r_{jv}, \ \lim_{v \to t} \mathcal{D}^r_{it} Y^r_v = \beta^r_i(t, Y_t). \tag{EC.1.5}$$

EC.1.2. HARA Optimal Policies Under Nonrandom Interest Rate

In this section, we decompose the HARA optimal policy under the models with nonrandom, but possibly time-varying, interest rate. In this case, we can explicitly calculate the ZCB price $B_{t,s}$ as

$$B_{t,s} = \exp\left(-\int_t^s r_v dv\right), \tag{EC.1.6}$$

which is just the discount factor from time t to s. The following proposition develops the connection between HARA and CRRA optimal policies.

Proposition EC.1 *With nonrandom, but possibly time-varying, interest rate r_t, the optimal policies under HARA and CRRA utilities satisfy the following simple ratio relationship:*

$$\pi_H^{mv}(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t}\pi_C^{mv}(t, Y_t) \text{ and } \pi_H^h(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t}\pi_C^h(t, Y_t). \tag{EC.1.7}$$

Thus, the optimal HARA policy is parallel to its CRRA counterpart:

$$\pi_H(t, X_t, Y_t) = \frac{\bar{X}_t}{X_t}\pi_C(t, Y_t). \tag{EC.1.8}$$

Proof. See Section EC.5.7. □

By Proposition EC.1, we see how the optimal HARA policy further simplifies under nonrandom interest rate. In particular, the additional term $\Pi_B(t, Y_t^r)$ in the hedge component $\pi_H^h(t, X_t, Y_t)$ in (38) vanishes in the optimal policy. It is because with nonrandom interest rate, the HARA investor does not need to hedge the uncertainty in the interest rate, and her future subsistence requirement can be perfectly matched by investing in the riskless asset. Consequently, the HARA optimal policy satisfies a simple ratio relationship with its CRRA counterpart, as shown in (EC.1.8).

In the following, we reveal how the optimal portfolio differs for heterogeneous HARA investors under the case of nonrandom interest rate. As in Section 4.2, we consider two HARA investors with initial wealth $X_0^{(h)} > X_0^{(l)}$. The utility parameters in (4) are the same for the two investors. Denote their optimal policies by $\pi_t^{(h)}$ and $\pi_t^{(l)}$. Define the optimal policy ratio of the two investors as $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)})$. Besides, define their optimal portfolio ratio as $1_m^\top \pi_t^{(h)} X_t^{(h)}/(1_m^\top \pi_t^{(l)} X_t^{(l)})$ using the amount of wealth allocated on the risky asset. We have the following proposition.

Proposition EC.2 *Under nonrandom interest rate, the optimal policy ratio of the high-wealth and low-wealth HARA investors satisfies*

$$\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}} > 1, \ \forall t \in [0, T]. \tag{EC.1.9}$$

The optimal portfolio ratio of the two HARA investors stays constant over time:

$$\frac{1_m^\top \pi_t^{(h)} X_t^{(h)}}{1_m^\top \pi_t^{(l)} X_t^{(l)}} \equiv \frac{X_0^{(h)} - Z_{0,T}}{X_0^{(l)} - Z_{0,T}}, \ \forall t \in [0, T]. \tag{EC.1.10}$$

Further assuming the interest rate is constant $r_t \equiv r$ and the investor utility only includes terminal wealth ($w = 0$ in eq. 4), we have

$$\frac{\partial}{\partial r}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) < 0 \text{ and } \frac{\partial}{\partial T}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) < 0, \tag{EC.1.11}$$

i.e., the optimal policy ratio decreases in the interest rate r and investment horizon T. The above results also hold for the optimal policy and portfolio ratio of each individual risky asset.

Proof. See Section EC.5.8. □

By (EC.1.9), the optimal policy ratio between the high- and low-wealth HARA investors is always larger than one, suggesting the high-wealth HARA investor invests more of her wealth on the risky assets. This introduces a risk-return trade-off investigated in Section 5. On the other hand, the optimal portfolio ratio of the two investors stays unchanged over time by (EC.1.10). It shows that the changes in the optimal policy ratio $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)})$ and investor wealth ratio $X_t^{(h)}/X_t^{(l)}$ exactly offset each other under the HARA utility with nonrandom interest rate.

Finally, Equation (EC.1.11) shows that the optimal policy ratio $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)})$ of the two HARA investors is larger when the interest rate is lower and/or investment horizon is shorter. With a lower interest rate and shorter horizon, the bond holding scheme for supporting the investor subsistence requirement, $\bar{x}_T \exp(-r(T-t))$, becomes more expensive. It decreases the remaining wealth of both HARA investors and reduces their weights on stocks. However, such effect is larger for the low-wealth investor due to the concavity of the wealth multiplier $\bar{X}_t/X_t$ in (EC.1.7). It leads to a larger optimal policy ratio $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)})$ of the two investors.

EC.1.3. Formulas for the CIRH-SVSIR Model

We first provide the function expressions used in the CIRH-SVSIR model. The functions $a(\tau)$ and $b(\tau)$ in the bond price $P_{t,T_1} = \exp(a(\tau_1) + b(\tau_1) r_t)$ and its dynamics (52) are given by

$$a(\tau) = \frac{2\kappa_r \theta_r}{\sigma_r^2} \ln\left(\frac{2\psi_r \exp\left(\tilde{\psi}_r \tau/2\right)}{\tilde{\psi}_r \left[\exp(\psi_r \tau) - 1\right] + 2\psi_r}\right) \tag{EC.1.12a}$$

and

$$b(\tau) = -\frac{2\left[\exp(\psi_r \tau) - 1\right]}{\tilde{\psi}_r \left[\exp(\psi_r \tau) - 1\right] + 2\psi_r}, \tag{EC.1.12b}$$

where $\psi_r = \sqrt{\left(\kappa_r + \lambda_r \sigma_r^2\right)^2 + 2\sigma_r^2}$ and $\tilde{\psi}_r = \psi_r + \kappa_r + \lambda_r \sigma_r^2$. The function $d_r(\tau)$ in (55a) is given by

$$d_r(\tau) = -\frac{2\left[\exp(\varphi_r \tau) - 1\right]\delta_r}{(\tilde{\kappa}_r + \varphi_r)\left[\exp(\varphi_r \tau) - 1\right] + 2\varphi_r}, \tag{EC.1.13a}$$

where $\delta_r = -(1-\gamma)\left[\lambda_r^2 \sigma_r^2/(2\gamma^2) + 1/\gamma\right]$, $\tilde{\kappa}_r = \kappa_r - (1-\gamma)\lambda_r \sigma_r^2/\gamma$, and $\varphi_r = \sqrt{\tilde{\kappa}_r^2 + 2\delta_r \sigma_r^2}$. The function $d_v(\tau)$ in (55b) follows by

$$d_v(\tau) = -\frac{2\left[\exp(\varphi_v \tau) - 1\right]\delta_v}{(\tilde{\kappa}_v + \varphi_v)\left[\exp(\varphi_v \tau) - 1\right] + 2\varphi_v}, \tag{EC.1.13b}$$

where $\delta_v = -(1-\gamma)\lambda_v^2/2\gamma^2$, $\tilde{\kappa}_v = \kappa_v - (1-\gamma)\lambda_v\sigma_v\rho_v/\gamma$, and $\varphi_v = \sqrt{\tilde{\kappa}_v^2 + 2\delta_v\left[\rho_v^2 + \gamma(1-\rho_v^2)\right]\sigma_v^2}$.

As a comparison, the CRRA optimal policy under the CIRH-SVSIR model is given as follows. For a CRRA investor, her optimal bond weight is:

$$\pi_C^{(bond)}(t, r_t, V_t) = \frac{1}{b(\tau_1)}\left(\frac{\lambda_r}{\gamma} + d_r(\tau)\right), \tag{EC.1.14a}$$

where $d_r(\tau)$ is defined in (EC.1.13a). The optimal stock weight of the CRRA investor follows by

$$\pi_C^{(stock)}(t, r_t, V_t) = \frac{\lambda_v}{\gamma} + \rho_v\sigma_v d_v(\tau), \tag{EC.1.14b}$$

with $d_v(\tau)$ defined in (EC.1.13b). The CRRA optimal policies (EC.1.14a) and (EC.1.14b) are indeed independent of the investor wealth X_t. In addition, we find the CRRA optimal policy is fully independent of the market state variables, and only depends on the remaining investment horizon $\tau = T - t$. This property is also noted in Liu (2007).

EC.2. Auxiliary Numerical Results

EC.2.1. Estimation of the CIRH-SVSIR Model

We describe how we estimate the CIRH-SVSIR model used in our comparative study. As mentioned in Section 5.1, we employ the maximum likelihood estimation method in Aït-Sahalia and Kimmel (2007) and Ait-Sahalia and Kimmel (2010). We use the SPDR S&P 500 ETF as the stock asset S_t, and extract the underlying variance V_t and interest rate r_t from VIX and US treasury bonds, respectively. We obtain the daily time series of S&P 500 index and the VIX index from the Center for Research in Security Prices (CRSP), and the U.S. treasury yields with maturities of one, two, and five years from the Federal Reserve Economic Data. We assume the one-year yield is observed without error, and the yields for two-year and five-year bonds contain normally distributed observation errors (see, e.g., Ait-Sahalia and Kimmel 2010).

First, following Section 5.1 of Aït-Sahalia and Kimmel (2007), we construct the integrated volatility proxy for the underlying variance V_t as

$$V_t \approx \frac{b_v V_{imp}(\tau_v) + a_v\tau_v}{\exp(b_v\tau_v) - 1} - \frac{a_v}{b_v}, \tag{EC.2.1}$$

where $V_{imp}(\tau_v)$ denotes the Black-Scholes implied variance calculated from the option price with maturity τ_v. The constants a_v and b_v are chosen such that the drift of the variance process has the linear form $a_v + b_v V_t$ under the risk-neutral measure. Under the CIRH-SVSIR model, they can be derived as

$$a_v = \kappa_v\theta_v \text{ and } b_v = -(\kappa_v + \lambda_v\rho_v\sigma_v), \tag{EC.2.2}$$

which are functions of the model parameters. The idea of the integrated proxy is to adjust the Black-Scholes implied volatility for the mean reversion effect in the volatility process. As in Aït-Sahalia and

Kimmel (2007) (see Section 7 therein), we use the VIX index to measure the implied volatility $V_{imp}(\tau_v)$, i.e., $VIX_t^2 = V_{imp}(\tau_v)$. The maturity τ_v is set as 22 trading days, in line with the calculation of VIX. As shown in Aït-Sahalia and Kimmel (2007), the integrated volatility proxy in (EC.2.1) is more accurate than the unadjusted Black-Scholes proxy with $V_t \approx VIX_t^2$.

Next, we calculate the instantaneous interest rate r_t from the treasury bond yields. We use the bond yields with maturities of 1, 2, and 5 years, denoted by $y_t^{(1)}$, $y_t^{(2)}$, and $y_t^{(5)}$. Following Ait-Sahalia and Kimmel (2010), we assume the one-year yield $y_t^{(1)}$ is accurately observed while the other two yields $y_t^{(2)}$ and $y_t^{(5)}$ contain observation errors. By the bond price expression $P_{t,T} = \exp\left(a(\tau) + b(\tau) r_t\right),$, we can calculate the interest rate as

$$r_t = \frac{1}{b(\tau_1)}(y_t^{(1)} - a(\tau_1)), \tag{EC.2.3}$$

with maturity $\tau_1 = 1$ year. The functions $a(\tau)$ and $b(\tau)$ are given by (EC.1.12a) and (EC.1.12b), which depend on the model parameters κ_r, θ_r, σ_r^2, and λ_r. With r_t solved from above, we can calculate the model implied two-year and five-year yields as

$$\tilde{y}_t^{(2)} = r_t b(\tau_2) + a(\tau_2) \;\text{ and }\; \tilde{y}_t^{(5)} = r_t b(\tau_5) + a(\tau_5),$$

with maturities $\tau_2 = 2$ years and $\tau_5 = 5$ years. Then, the observation errors are given by

$$\varepsilon_t^{(2)} = \tilde{y}_t^{(2)} - y_t^{(2)} \;\text{ and }\; \varepsilon_t^{(5)} = \tilde{y}_t^{(5)} - y_t^{(5)},$$

which are the difference between the implied and observed yields. As in Ait-Sahalia and Kimmel (2010), we assume the observation errors are normally distributed with constant mean and variance. In addition, they are independent across time and maturity, as well as the state variable processes. Thus, the joint likelihood of the observation errors $\{\varepsilon_t^{(2)}\}$ and $\{\varepsilon_t^{(5)}\}$ can be calculated using normal probability density function.

We obtain the likelihood of the observed data as follows for each parameter vector considered. We start from the daily times series of S&P 500 index, square of VIX, and one-year treasury yield $(S_t, VIX_t^2, y_t^{(1)})$. We first extract the value of the state variable (S_t, V_t, r_t) from the observed data by (EC.2.1) and (EC.2.3). Second, we evaluate the joint likelihood of the state variable process (S_t, V_t, r_t), using the Euler's approximation to their SDEs in (53), (54), and (51). Third, we multiply this joint likelihood by the Jacobian determinant of the mapping from (S_t, V_t, r_t) to $(S_t, VIX_t^2, y_t^{(1)})$. It yields the joint likelihood of the observed panel $(S_t, VIX_t^2, y_t^{(1)})$. By (EC.2.1) and (EC.2.3), the Jacobian determinant is given by

$$J = \det\left|\begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{\partial V_t}{\partial (VIX_t^2)} & 0 \\ 0 & 0 & \frac{\partial r_t}{\partial y_t^{(1)}} \end{pmatrix}\right| = \frac{b_v}{\left[\exp(b_v \tau_v) - 1\right] b(\tau_1)},$$

where b_v and $b(\tau_1)$ are given in (EC.2.2) and (EC.1.12b), respectively. Finally, we calculate the likelihood of the observation errors for the two-year and five-year yields, and multiply this likelihood by the joint likelihood of $(S_t, VIX_t, y_t^{(1)})$ obtained in previous step. We repeat this procedure for each candidate parameter vector to find the one that leads to the maximum likelihood.

EC.2.2. Impact of Investment Horizon on HARA Policy

In this section, we show how the HARA optimal policy under the CIRH-SVSIR model is affected by the investment horizon $T-t$. The optimal policy is solved in closed-form in Proposition 5. As in Section 5.2, we consider a HARA investor that maximizes her expected utility over terminal wealth with $X_0/\bar{x}_T = 4$ and $\gamma = 4$. For comparison, we also consider a CRRA investor with the same utility parameters but no subsistence requirement. We use the parameters in Table 1. Figure EC.1 plots the optimal stock (left panel) and bond (right panel) of the investor with different investment horizon $T-t$. In each panel, the red dashed and blue circled curves represent the optimal CRRA and HARA policies, respectively.

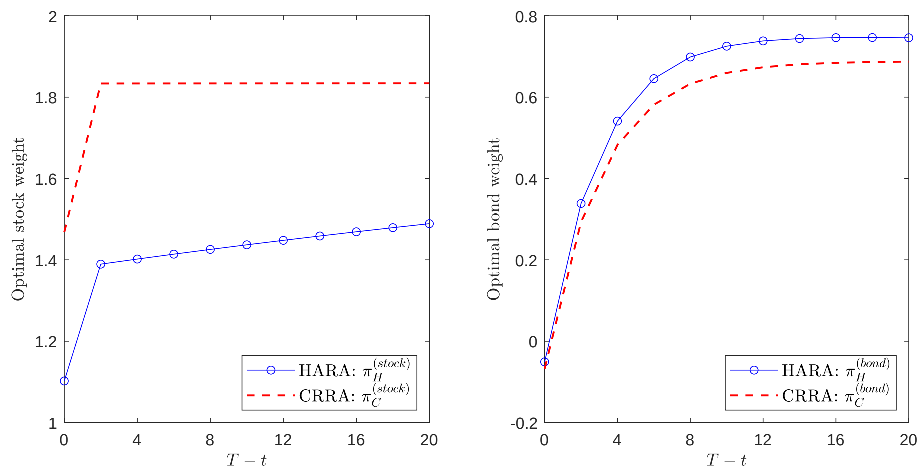

Figure EC.1 Optimal policy in the CIRH-SVSIR model by different investor horizon: optimal stock weight (left panel) and optimal bond weight (right panel).

By the left panel, we see that both the optimal stock weights $\pi_C^{(stock)}$ and $\pi_H^{(stock)}$ increase with the remaining investment horizon $T-t$. For the CRRA case, a sharp increase occurs when the investment horizon is short. It is due to the hedge component $\rho_v \sigma_v d_v\,(T-t)$ in (EC.1.14b). When $\rho_v < 0$, it increases monotonically in investment horizon $T-t$ due to more uncertainty. As the investment horizon becomes long, the CRRA optimal policy becomes almost insensitive to $T-t$. For the HARA case, however, in addition to the similar sharp increase for short horizons, $\pi_H^{(stock)}$ keeps increasing in $T-t$ even for longer investment horizons. As mentioned in Section 5.2, the increase in $\pi_H^{(stock)}$ under the HARA utility is generated by a combination of two effects. The first is the increase of the hedging demand as discussed above. The second lies in the multiplier $\bar{X}_t/X_t$, which is given by

$$\frac{\bar{X}_t}{X_t} = 1 - \frac{\bar{x}_T B_{t,T}}{X_t} = 1 - \frac{\bar{x}_T}{X_t}\exp\left(a(T-t) + b(T-t)r_t\right). \tag{EC.2.4}$$

The multiplier $\bar{X}_t/X_t$ also increases in $T-t$, as the ZCB price $B_{t,T}$ becomes cheaper when the investment horizon is longer. It increases the optimal weight on stock. Such wealth effect is significant even for long investment horizons, generating a more lasting impact of investment horizon under the HARA utility.

The right panel shows that the optimal bond weight increases concavely with remaining investment horizon under both HARA and CRRA utilities. Interestingly, the optimal bond weight changes its sign from negative to positive as the investment horizon increases. In the CIRH-SVSIR model, the bond asset has a negative risk premium as we always have $b(\tau_1) < 0$ in (52). Thus, the mean-variance component of the optimal bond weight is negative. On the other hand, the hedge component of the optimal bond weight is positive. It is because the stock return increases in risk-free rate r_t by (53), while the bond price decreases in r_t. Thus, investors will be long the bond to hedge potential decrease in the risk-free rate. When the investment horizon is short, the hedging demand is small. In this case, the negative mean-variance component dominates and leads to a short position in the bond. However, with a longer investment horizon, the hedging demand becomes larger and eventually generates a positive optimal bond weight. Moreover, the marginal impact of investment horizon becomes smaller as $T - t$ increases, leading to a concave increasing pattern. Finally, we notice that the optimal bond weight is always larger for the HARA investor than that for the CRRA investor. It is because the HARA investors will additionally hold the bond to finance their terminal subsistence requirements.

EC.2.3. Numerical Examples for Cycle-dependence under HARA Utility

In this section, we provide more numerical examples regarding the cycle-dependence of HARA optimal policy, which is discussed in Section 5.3. First, Figure EC.2 shows the bond price B_t (black dotted with the left y-axis) and the optimal bond weights of the CRRA and HARA investors (blue solid and red dashed with the right y-axis) under the same simulated path in Figure 2. Unlike the optimal stock weight, the optimal bond weight $\pi_H^{(bond)}$ of the HARA investor is relatively insensitive to market scenarios and tracks its CRRA counterpart closely. That is, we do not observe the cycle-dependence in the HARA investor bond position as in her stock position. It can be potentially explained as follows. By (55a), the bond position of the HARA investors can be decomposed into two parts, corresponding to the CRRA and financing portfolios respectively. The first part in the CRRA portfolio, $(\bar{X}_t/X_t)\pi_C^{(bond)}(t, r_t, V_t)$, increases in wealth level X_t via the multiplier $\bar{X}_t/X_t$, which is same as the stock weight $\pi_H^{(stock)}$. However, the second part from the financing portfolio, $\Pi_B(t, r_t)/X_t$, clearly decreases in investor wealth X_t. The two effects counteract each other and mitigate the cycle-dependence in the HARA bond weight. Thus, the bond position of the HARA investor does not fluctuate much with the market scenarios.

We then provide an example for how the cycle-dependence under HARA utility varies for investors with different wealth levels. We consider two HARA investors who have the same risk aversion level $\gamma = 4$ and an investment horizon $T = 10$ years. The two HARA investors only differ in their initial wealth levels, with the high-wealth and low-wealth investors having an initial wealth of $X_0^{(h)} = 5\bar{x}_T$ and $X_0^{(l)} = 2\bar{x}_T$, respectively. The optimal policies for the two HARA investors are solved by Proposition 5.

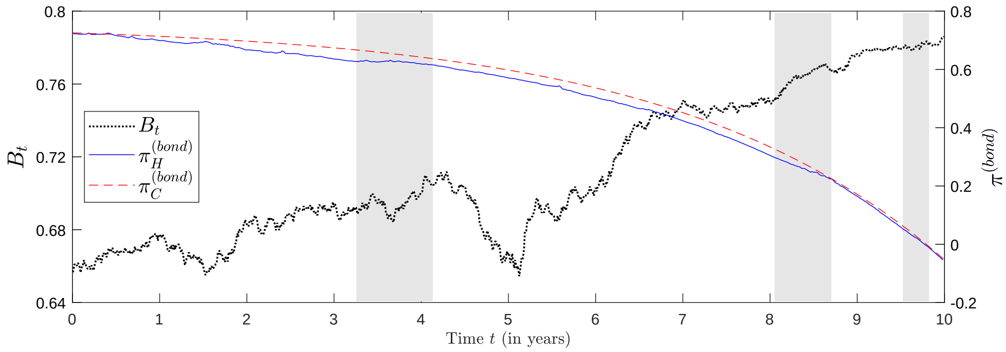

Figure EC.2 The figure plots the simulated bond price B_t, optimal bond weights $\pi_H^{(bond)}$ and $\pi_C^{(bond)}$ of HARA and CRRA investors. Here we use the same simulated path as in Figure 2.

The upper panel of Figure EC.3 plots the path of stock price S_t (black dotted with the left y-axis) and the corresponding optimal stock weights $\pi_t^{(h,s)}$ and $\pi_t^{(l,s)}$ of the high-wealth and low-wealth HARA investors (blue solid and red dashed with the right y-axis). We still use the same market path as in Figure 2. Unsurprisingly, the optimal stock weights of both investors are positively correlated with the stock price, reflecting the cycle-dependence seen in Figure 2. By comparing the optimal stock weights of the two investors, we see that the magnitude of variation is much larger for the low-wealth HARA investor. That is, the stock position of the low-wealth investor is more sensitive to market cycles. Such greater cycle-dependence can be explained as follows: the low-wealth HARA investor faces a more binding subsistence constraint, thus her optimal stock weight is more sensitive to the wealth level. While not reported here, we find that the optimal policy ratio for stock, $\pi_t^{(h,s)}/\pi_t^{(l,s)}$, becomes larger (resp. smaller) in bear (resp. bull) markets, suggesting the optimal portfolios of the two investors further diverge in stressed markets.

The lower panel of Figure EC.3 plots the equivalent risk aversion level $\gamma_H(X_t)$ in (27) for the two HARA investors along the simulated path. We see that both the equivalent risk aversion levels are higher during the bear market regimes. Moreover, the equivalent risk aversion of the low-wealth investor fluctuates more dramatically with the market than that of the high-wealth investor. It again reflects the wealth effect of the HARA utility. In bear markets, both the investors suffer a loss in their wealth, leading to a higher equivalent risk aversion by (27). Moreover, such effect is more significant for the low-wealth investor with more binding subsistence constraints. The greater cycle-dependence for low-wealth HARA investors is validated by the regression analysis discussed in Section 5.3 (see Figure 3).

EC.2.4. Statistical Tests for Investment Performance Metrics

In this section, we examine the statistical significance of the patterns of investment performance metrics for HARA investors with different initial wealth levels, which are discussed in Section 5.4. This is done as follows. For a given performance metrics M (e.g., excess return mean, volatility), we evaluate its ex-ante unconditional expectation using 10^4 simulated paths for HARA investor with initial wealth $X_0/\bar{x}_T =$

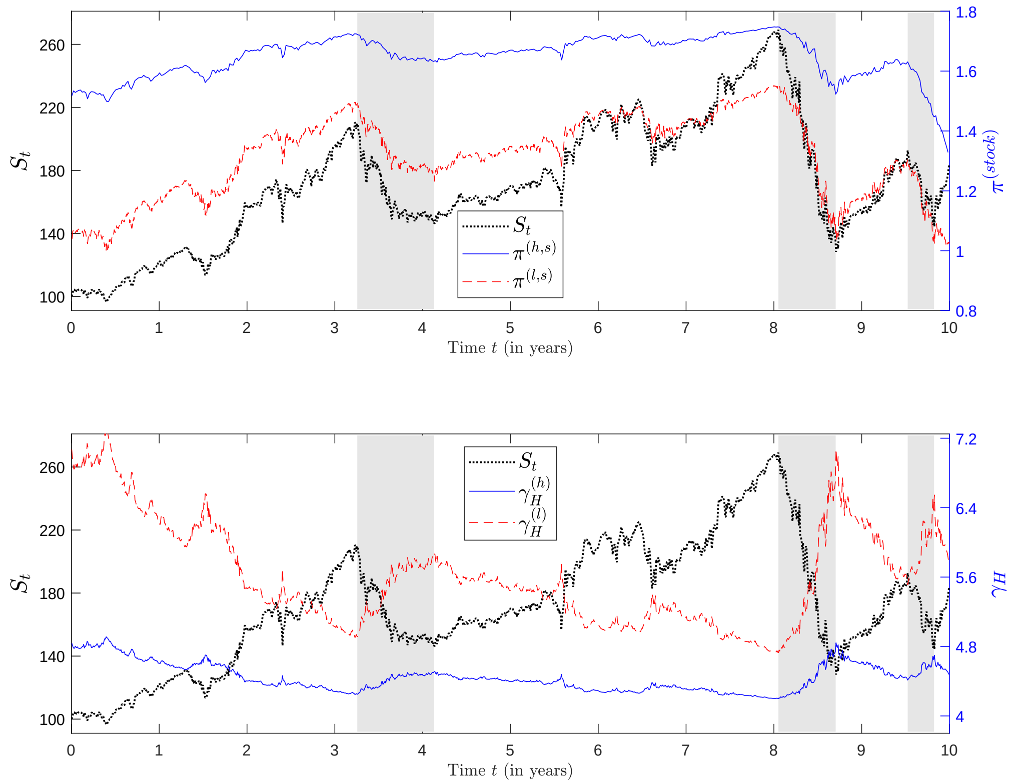

Figure EC.3 The upper panel plots the simulated path of stock price S_t, optimal stock weights $\pi^{(h,s)}$ and $\pi^{(l,s)}$ of high-wealth and low-wealth HARA investors. The lower panel plots the equivalent risk aversion $\gamma_H^{(h)}$ and $\gamma_H^{(l)}$ of the two investors.

l. Denote the corresponding estimate by $E[M_l]$. We choose a set of representative initial wealth levels $l \in \{1, 2, 4, 6, 8, 10\}$, which covers the x-axis range in Figure 4. We test the statistical significance of the difference in $E[M_l]$ for adjacent levels of l, e.g., $E[M_2] - E[M_1]$. We simulate the optimal strategies and wealth process of different investors using the same simulated paths. Thus, we employ a paired t-test for the mean difference, with each pair being the performance metrics of two HARA investors from a given path.

The statistical test results are reported in Table EC.1. The investment performance metrics are included in the columns, including average excess return mean ($\bar{R}$), return volatility (SD), maximum drawdown (MD), 99% Value-at-Risk (99% VaR), Sharpe ratio (SR), and the weighted stock return mean ($\bar{R}^{(stock)}$). All quantities except the VaR are annualized. The average differences are reported, and their standard errors are given in the parenthesis. If the average differences are all positive in a column, it implies a monotonically increasing pattern with respect to the initial wealth level.

With a large number (10^4) of sample paths, the standard errors of average differences are relatively small. Thus, we see that all the average differences are statistically significant at the 0.1% level, except for the difference in Sharpe ratio between $X_0/\bar{x}_T = 4$ and $X_0/\bar{x}_T = 2$.[1] It validates the our discussion in Section

[1] This occurs around the kink point in the pattern of Sharpe ratio, as shown in the lower-left panel of Figure 4.

Table EC.1 Statistical Tests for Differences in Investment Performance of HARA Investors with Varying Initial Wealth Level

	$\bar{R}$	SD	MD	$99\%\ VaR$	SR	$\bar{R}^{(stock)}$
$E[M_2]-E[M_1]$	1.16×10^{-1} (1.82×10^{-4})	1.21×10^{-1} (2.33×10^{-4})	1.41×10^{-1} (6.92×10^{-4})	1.95×10^{-2} (5.07×10^{-5})	5.79×10^{-2} (8.01×10^{-4})	9.49×10^{-3} (1.20×10^{-4})
$E[M_4]-E[M_2]$	2.92×10^{-2} (2.68×10^{-5})	3.40×10^{-2} (1.10×10^{-4})	3.59×10^{-2} (2.86×10^{-4})	5.63×10^{-3} (2.47×10^{-5})	1.60×10^{-5} (2.40×10^{-4})	2.52×10^{-3} (2.87×10^{-5})
$E[M_6]-E[M_4]$	8.10×10^{-3} (6.51×10^{-6})	1.00×10^{-2} (3.78×10^{-5})	1.07×10^{-2} (9.41×10^{-5})	1.72×10^{-3} (9.58×10^{-6})	-1.39×10^{-3} (7.37×10^{-5})	7.56×10^{-4} (7.87×10^{-6})
$E[M_8]-E[M_6]$	3.81×10^{-3} (3.20×10^{-6})	4.85×10^{-3} (1.96×10^{-5})	5.20×10^{-3} (4.70×10^{-5})	8.46×10^{-4} (5.23×10^{-6})	-8.89×10^{-4} (3.57×10^{-5})	3.69×10^{-4} (3.67×10^{-6})
$E[M_{10}]-E[M_8]$	2.22×10^{-3} (1.82×10^{-6})	2.86×10^{-3} (1.19×10^{-5})	3.07×10^{-3} (2.81×10^{-5})	5.03×10^{-4} (3.30×10^{-6})	-5.89×10^{-4} (2.12×10^{-5})	2.19×10^{-4} (2.13×10^{-6})

5.4 on how the initial wealth level impacts the investment performance of HARA investors. In particular, the HARA investor with more initial wealth has higher average excess return, volatility, maximum drawdown, and Value-at-Risk. For Sharpe ratio, there is an upward jump when the initial wealth level is low, followed by a very small drop under high initial wealth levels. The weighted stock return mean also increases in investor's initial wealth level.

EC.2.5. Wealth Effects on Expected Sharpe Ratio

In this section, we investigate the wealth effects on the Sharpe ratio of HARA investors. By the lower-left panel of Figure 4, the Sharpe ratio increases from 0.855 to 0.913 when the HARA investor's initial wealth increases from $X_0/\bar{x}_T = 1$ to $X_0/\bar{x}_T = 3$. The Sharpe ratio is a risk-adjusted measure. Thus, its increase cannot be simply explained by the fact that high-wealth HARA investors put more wealth on the stock asset, as it leads to both higher expected return and volatility.

We interpret the increase in Sharpe ratio as follows. Under the CIRH-SVSIR model, the pattern of the Sharpe ratio is largely driven by the investor's position of the stock asset (the average return from the bond asset is very low). Thus, we focus on the stock returns of the HARA investors. Define the investor's weighted excess stock return for day i as

$$R_i^{(stock)} = \frac{\pi_{i\Delta}^{(stock)}}{\bar{\pi}^{(stock)}} \times \left[\ln\left(\frac{S_{(i+1)\Delta}}{S_{i\Delta}}\right)/\Delta - r_{i\Delta}\right], \tag{EC.2.5}$$

where $\bar{\pi}^{(stock)} = \sum_{i=1}^{N} \pi_{i\Delta}^{(stock)}/N$ is the investor's average stock weight for the entire horizon. We consider the average and standard deviation of the weighted returns, i.e., $\bar{R}^{(stock)} = \sum_{i=1}^{N} R_i^{(stock)}/N$ and $SD^{(stock)} = \sqrt{\sum_{i=1}^{N}(R_i^{(stock)} - \bar{R}^{(stock)})^2/(N-1)}$. Both $\bar{R}^{(stock)}$ and $SD^{(stock)}$ are adjusted for the average stock position of the investor. Thus, they are comparable for HARA investors with different initial wealth levels. In the special case where the investor has a constant stock weight, $\bar{R}^{(stock)}$ and $SD^{(stock)}$ reduce to the average and standard deviation of the stock excess returns as $\pi_{i\Delta}^{(stock)} \equiv \bar{\pi}^{(stock)}$. Note that the ratio between the mean and standard deviation of excess stock returns is not affected by the normalization level $\bar{\pi}^{(stock)}$.

We expect the investor's risk-adjusted return to be higher if the average weighted stock return $\bar{R}^{(stock)}$ is higher or the standard deviation $SD^{(stock)}$ is lower. Using 10^4 simulated paths, we estimate their ex-ante expectations for HARA investors with different initial wealth levels. The estimated $E[\bar{R}^{(stock)}]$ and $E[SD^{(stock)}]$ are plotted in the left and right panels of Figure EC.4, respectively. Note that $E[\bar{R}^{(stock)}]$ is also shown in the lower-right panel of Figure 4 in the main text. We see that the average weighted stock return $E[\bar{R}^{(stock)}]$ increases in investor initial wealth, while the standard deviation $E[SD^{(stock)}]$ exhibits a substantial drop as the initial wealth first increases. Both the two channels contribute to a higher risk-adjusted return of stock for HARA investors with higher initial wealth.

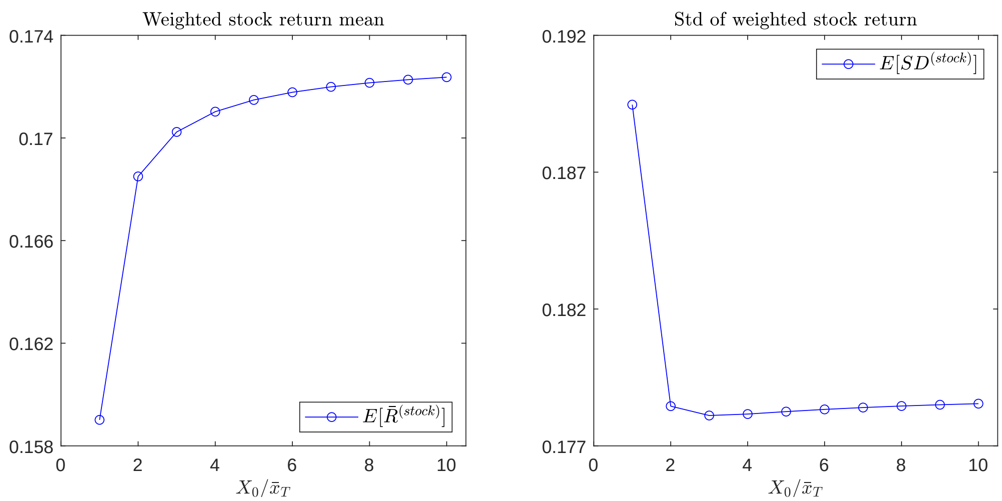

Figure EC.4 Average and standard deviation of the weighted stock return of HARA investors: $E[\bar{R}^{(stock)}]$ (left) and $E[SD^{(stock)}]$ (right).

As discussed in Section 5.5 of the main text, the increase in $E[\bar{R}^{(stock)}]$ can be explained by the market timing effect due to the procyclical investment behaviors of HARA investors: high-wealth HARA investors are more capable to hold risky assets during stressed periods with high expected returns. On the other hand, the decrease in the standard deviation $E[SD^{(stock)}]$ can be interpreted by the time variation in investor stock position, even after we normalize it with the average level $\bar{\pi}^{(stock)}$. First. the investor's wealth generally increases over the investment horizon as the stock asset has a positive risk premium. Thus, HARA investors will gradually increase their stock weight by the static analysis in Section 5.2. Second, due to the cycle-dependence analyzed in Section 5.3, the optimal policy under HARA utility depends on the historical path of market performance. Both factors introduce time variation to the optimal stock position of HARA investors. Moreover, due to the wealth constraint, such effects are more pronounced for low-wealth investors. This introduces additional uncertainty in the optimal policy for HARA investors with low initial wealth, even after we adjust for their average stock weight $\bar{\pi}^{(stock)}$.

We use simulations to illustrate the additional uncertainty in the optimal stock weight for low-wealth HARA investors. Define the normalized policy $\pi'_{i\Delta}$ as $\pi'_{i\Delta} = \pi^{(stock)}_{i\Delta}/\bar{\pi}^{(stock)}$ for each path. Figure EC.5 exhibits the representative quantiles of $\pi'_{i\Delta}$ for HARA investors with initial wealth $X_0/\bar{x} = 1$ (top) and $X_0/\bar{x} = 4$ (bottom). For ease of comparison, we use the same vertical axis in the two panels. We show the quantiles at the beginning of each quarter in the investment horizon, which are computed based on 10^4 simulation trials. The drop near the end is due to the decrease in the hedge component when investment horizon shrinks to zero, as analyzed in Section EC.2.2. Comparisons between the two panels clearly support our interpretations. The distributions of $\pi'_{i\Delta}$ of the low-wealth investor exhibit a more significant increasing pattern over the investment horizon, and spread out over much wider ranges than that of the high-wealth investor. These effects introduce more time variation in the normalized policy $\pi'_{i\Delta}$, leading to a higher $E[SD^{(stock)}]$.

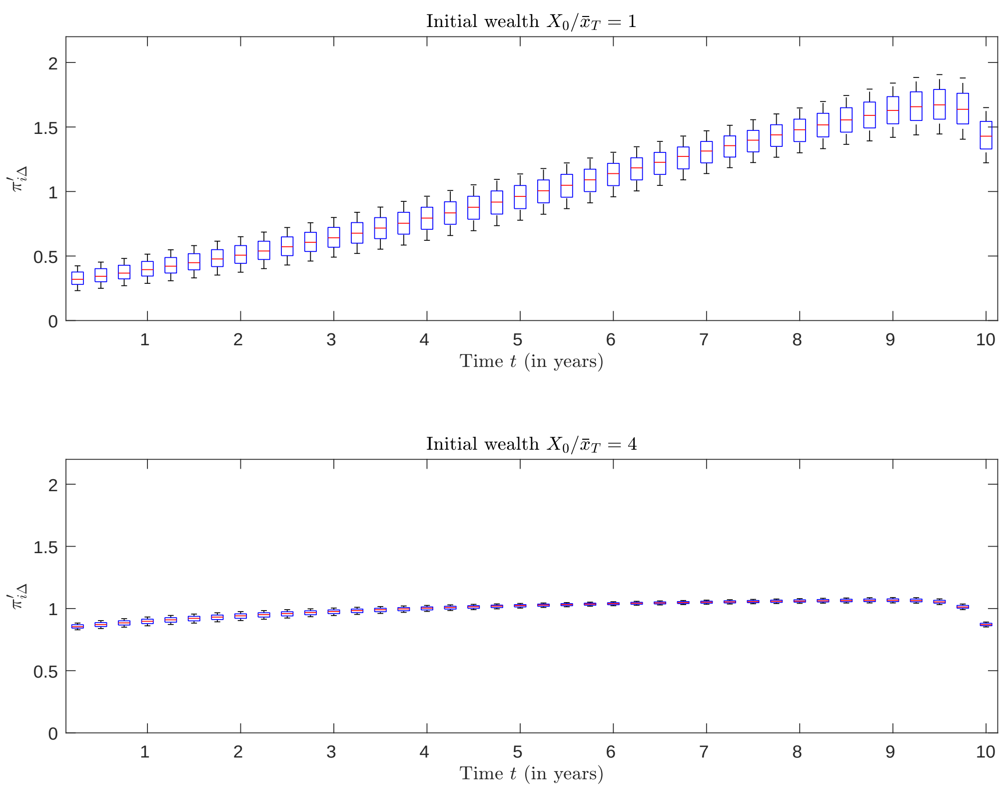

Figure EC.5 Representative quantiles of the normalized optimal stock weight $\pi'_{i\Delta}$ for HARA investors: $X_0/\bar{x}_T = 1$ (top) and $X_0/\bar{x}_T = 4$ (bottom).

By above discussion, high-wealth HARA investors benefit from the market timing effect in stock investment and have less time variation in their optimal stock weights. Both factors lead to higher risk-adjusted returns, which explain the upward jump in Sharpe ratio shown in Figure 4.

We also observe a drop in Sharpe ratio when the initial wealth further increases, although the economic magnitude is negligible (from 0.913 to 0.910). This slight drop can be potentially explained as follows. As the initial wealth level increases, the HARA investor will allocate more of her wealth on the stock. This can reduce the risk-adjusted return due to the quadratic variation of the Brownian motion. Consider a simplified example in which a single stock follows a Geometric Brownian motion with return rate μ and volatility σ. The interest rate is a constant r. Assume an investor always allocates a proportion π of her wealth on the stock. Then the investor wealth is given by

$$X_t = X_0 \exp\left\{\left(r + \pi(\mu - r) - \frac{\pi^2\sigma^2}{2}\right)t + \pi\sigma W_t\right\}.$$

This translates to a Sharpe ratio of $(\mu - r)/\sigma - \pi\sigma/2$, which decreases in the weight π due to the second term. It also explains the small upward tick in $E[SD^{(stock)}]$ as shown in the right panel of Figure EC.4. However, this effect tends to be small as the investor's optimal stock weight becomes insensitive to wealth level when her initial wealth is already high (see Figure 1).

EC.3. Decomposition of Optimal Policy under General Incomplete Market Models

In this section, we establish a decomposition of optimal policy under the general incomplete market model (1) – (2) with flexible utilities. In particular, we consider the following general optimization problem:

$$\sup_{(\pi_t, c_t)} E\left[\int_0^T u(t, c_t)dt + U(T, X_T)\right], \text{ with } X_t \geq 0 \text{ for all } t \in [0, T], \tag{EC.3.1}$$

where $u(t,\cdot)$ and $U(T,\cdot)$ are time-additive utility functions of the intermediate consumption and the terminal wealth; they are allowed to be time-varying to reflect the time value in utility, e.g., the discount effect. We assume they are strictly increasing and concave with $\lim_{x\to\infty} \partial u(t,x)/\partial x = \lim_{x\to\infty} \partial U(T,x)/\partial x = 0$. These assumptions are satisfied by the CRRA and HARA utilities in (4) and (5).

We solve the optimal portfolio allocation problem (EC.3.1) in the incomplete market by the fictitious completion approach proposed in Karatzas et al. (1991). We introduce $d - m$ fictitious assets to complete the market under the orthogonal condition (7). Then, we are in a complete market with m real assets and $d - m$ fictitious assets. We solve the optimal policy (π_t, π_t^F) in this completed market by the martingale approach, which is pioneered by Karatzas et al. (1987) and Cox and Huang (1989). It starts by formulating the dynamic problem (EC.3.1) with information up to time t as the following equivalent static optimization problem:

$$\sup_{(c_t, X_T)} E_t\left[\int_t^T u(s, c_s)ds + U(T, X_T)\right] \text{ subject to } E_t\left[\int_t^T \xi_{t,s} c_s ds + \xi_{t,T} X_T\right] \leq X_t, \tag{EC.3.2}$$

where E_t denotes the expectation condition on the information up to time t and X_t is the wealth level assuming that the investor always follows the optimal policy. Following the standard method of Lagrangian multiplier, we can represent the optimal intermediate consumption and terminal wealth as $c_t = I^u(t, \lambda_t^*)$ and $X_T = I^U(T, \lambda_T^*)$, with $I^u(t,\cdot)$ and $I^U(t,\cdot)$ being the inverse marginal utility functions of $u(t,\cdot)$ and $U(t,\cdot)$, i.e., the functions satisfying $\partial u/\partial x(t, I^u(t,y)) = y$ and $\partial U/\partial x(t, I^U(t,y)) = y$. The quantity λ_t^* denotes the Lagrangian multiplier for the wealth constraint in (EC.3.2). It is uniquely characterized by

$$X_t = E_t[\mathcal{G}_{t,T}(\lambda_t^*)], \tag{EC.3.3}$$

where $\mathcal{G}_{t,T}(\lambda_t^*)$ is defined as

$$\mathcal{G}_{t,T}(\lambda_t^*) := \xi_{t,T} I^U\left(T, \lambda_t^* \xi_{t,T}\right) + \int_t^T \xi_{t,s} I^u\left(s, \lambda_t^* \xi_{t,s}\right) ds. \tag{EC.3.4}$$

Here, $\xi_{t,s} = \xi_s/\xi_t$ is the relative state price density defined by (10). By (EC.3.3), we can determine the multiplier λ_t^* with information up to time t.

By the least favorable completion principle proposed in Karatzas et al. (1991), the optimal policy π_t for the real assets in the completed market coincides with its counterpart in the original incomplete market, as long as we properly choose the investor-specific price of risk θ_v^u such that the optimal weights for the fictitious assets are always identically zero, i.e.,

$$\pi_v^F \equiv 0_{d-m}, \text{ for any } 0 \le v \le T. \tag{EC.3.5}$$

Given an arbitrary choice of the volatility function $\sigma^f(v,y)$, the least favorable constraint (EC.3.5) and the orthogonal condition (7) determine the proper θ_v^u for $0 \le v \le T$. Then, the corresponding optimal policy π_t of the real assets for the completed market is also optimal for the original incomplete market. In particular, the desired θ_v^u satisfying (EC.3.5) and the resulting optimal policy π_t are independent of the specific choice of $\sigma^f(v,y)$.

With these preparations, we establish the decomposition of the optimal policy for general incomplete market models with flexible utilities in the following theorem.

Theorem EC.1 *Under the incomplete market model (1) – (2), the optimal policy π_t for the real assets with prices S_t admits the following decomposition:*

$$\pi_t = \pi^{mv}(t, X_t, Y_t) + \pi^h(t, X_t, Y_t). \tag{EC.3.6}$$

The terms $\pi^{mv}(t, X_t, Y_t)$ and $\pi^h(t, X_t, Y_t)$ denote the mean-variance component and the hedge component. The components can be expressed as conditional expectations on random variables with explicit dynamics:

$$\pi^{mv}(t, X_t, Y_t) = -(\sigma(t,Y_t)^+)^\top \theta^h(t,Y_t) E_t[\mathcal{Q}_{t,T}(\lambda_t^*)]/X_t, \tag{EC.3.7a}$$

and

$$\pi^h(t, X_t, Y_t) = -(\sigma(t,Y_t)^+)^\top E_t[\mathcal{H}_{t,T}(\lambda_t^*)]/X_t,. \tag{EC.3.7b}$$

Hereof, λ_t^* *is the multiplier uniquely determined by (EC.3.3), i.e.,* $X_t = E_t[\mathcal{G}_{t,T}]$*. It depends on* X_t *and satisfies the relation* $\lambda_t^* = \lambda_0^* \xi_t$*. The expressions for* $\mathcal{Q}_{t,T}(\lambda_t^*)$ *and* $\mathcal{H}_{t,T}(\lambda_t^*)$ *are explicitly given by*

$$\mathcal{Q}_{t,T}(\lambda_t^*) = \lambda_t^* (\xi_{t,T})^2 \frac{\partial I^U}{\partial y}(T, \lambda_t^* \xi_{t,T}) + \lambda_t^* \int_t^T (\xi_{t,s})^2 \frac{\partial I^u}{\partial y}(s, \lambda_t^* \xi_{t,s})\, ds, \tag{EC.3.8a}$$

and

$$\mathcal{H}_{t,T}(\lambda_t^*) = \Upsilon^U(T, \lambda_t^* \xi_{t,T}) \xi_{t,T} H_{t,T} + \int_t^T \Upsilon^u(s, \lambda_t^* \xi_{t,s}) \xi_{t,s} H_{t,s} ds, \tag{EC.3.8b}$$

where $\Upsilon^U(t,y) = I^U(t,y) + y\partial I^U(t,y)/\partial y$ *and* $\Upsilon^u(t,y) = I^u(t,y) + y\partial I^u(t,y)/\partial y$*. The term* $H_{t,s}$ *satisfies following SDE*

$$dH_{t,s} = [\mathcal{D}_t r_s + (\mathcal{D}_t \theta_s^c)\theta_s^c]\, ds + (\mathcal{D}_t \theta_s^c)\, dW_s, \tag{EC.3.9}$$

with initial values $H_{t,t} = 0_d$*. Here* $\mathcal{D}_t r_s$ *and* $\mathcal{D}_t \theta_s^c$ *denote the Malliavin derivatives of the interest rate* $r(s, Y_s)$ *and total price of risk* θ_s^c*, respectively. The optimal intermediate consumption* c_t *and terminal wealth* X_T *are given by* $c_t = I^u(t, \lambda_t^*)$ *and* $X_T = I^U(T, \lambda_T^*)$*.*

Proof. *See Section EC.3.1.* □

Next, we provide the following proposition for characterizing the proper investor-specific price of risk θ_v^u for the least favorable completion (EC.3.5).

Proposition EC.3 *Under the incomplete market model (1) – (2) and the utility (EC.3.2), the investor-specific price of risk* θ_v^u *at time* v *for the least favorable completion can be solved from the following optimization problem:*

$$\inf_{\theta^u \in Ker(\sigma)} E_v \left[\int_0^T \tilde{u}(s, \lambda_s^*) ds + \tilde{U}(T, \lambda_T^*) \right], \tag{EC.3.10}$$

Here we use $\theta^u \in Ker(\sigma)$ *to abbreviate for* $\theta_s^u \in Ker(\sigma(s, Y_s))$ *for any* $v \le s \le T$*, with* $Ker(\sigma(s, Y_s)) := \{w \in R^d : \sigma(s, Y_s) w \equiv 0_m\}$ *denoting the kernel of* $\sigma(s, Y_s)$*;* λ_s^* *is the time–s multiplier. It satisfies the following SDE*

$$d\lambda_s^* = -\lambda_s^* [r(s, Y_s) ds + (\theta_s^h(s, Y_s) + \theta_s^u)^\top dW_s],$$

in which θ_s^u *for* $v \le s \le T$ *serves as a control process. Besides,* $\tilde{u}(t,y)$ *and* $\tilde{U}(t,y)$ *are the conjugates of utility functions* $u(t,x)$ *and* $U(t,x)$*, defined by* $\tilde{u}(t,y) := \sup_{x\ge 0}(u(t,x) - yx)$ *and* $\tilde{U}(t,y) := \sup_{x \ge 0}(U(t,x) - yx)$*, respectively. In particular, the investor-specific price of risk* θ_v^u *satisfies the following* $d-$*dimensional equation:*

$$\theta_v^u = \frac{\sigma(v, Y_v)^+ \sigma(v, Y_v) - I_d}{E_v[\mathcal{Q}_{v,T}(\lambda_v^*)]} \times E_v[\mathcal{H}_{v,T}(\lambda_v^*)], \tag{EC.3.11}$$

where I_d *denotes the d–dimensional identity matrix;* $\sigma(v, Y_v)^+$ *is given by*

$$\sigma(v, Y_v)^+ = \sigma(v, Y_v)^\top (\sigma(v, Y_v) \sigma(v, Y_v)^\top)^{-1}.$$

Proof. *See Section EC.3.2.* □

EC.3.1. Proof of Theorem EC.1

We first provide the following lemma that represents the optimal policy under the completed market with both real and fictitious assets, assuming the investor-specific price of risk process θ_s^u were known. Recall that in the completed market, the assets $C_t = (S_t^\top, F_t^\top)^\top$ follow the dynamics:

$$dC_t = \text{diag}(C_t)\left[\mu_t^c dt + \sigma^c(t, Y_t) dW_t\right], \tag{EC.3.12}$$

with $\mu_t^c = ((\mu(t,Y_t) - \delta(t,Y_t))^\top, (\mu_t^f)^\top)^\top$ and $\sigma^c(t,Y_t) = (\sigma(t,Y_t)^\top, \sigma^f(t,Y_t)^\top)^\top$.

Lemma EC.1 *In the completed market with dynamics (EC.3.12) and (2), the optimal policy* $(\pi_t, \pi_t^F)^\top$ *for both the real and fictitious assets admits the following representation*

$$(\pi_t, \pi_t^F)^\top = -\frac{1}{X_t}(\sigma^c(t,Y_t)^\top)^{-1}\left(\theta_t^c E_t[\mathcal{Q}_{t,T}(\lambda_0^*\xi_t)] + E_t[\mathcal{H}_{t,T}(\lambda_0^*\xi_t)]\right), \tag{EC.3.13}$$

where θ_t^c *is the total price of risk defined in (8);* E_t *denotes the expectation conditional on the information up to time* t*;* ξ_t *is the state price density defined in (10);* λ_0^* *is the multiplier uniquely determined by the wealth equation*

$$X_0 = E[\mathcal{G}_{0,T}(\lambda_0^*\xi_t)], \tag{EC.3.14}$$

where X_0 *is the initial wealth and the function* $\mathcal{G}_{t,T}(\cdot)$ *is defined by (EC.3.4). The components* $\mathcal{Q}_{t,T}(\lambda_0^*\xi_t)$ *and* $\mathcal{H}_{t,T}(\lambda_0^*\xi_t)$ *are given by (EC.3.8a) and (EC.3.8b).*

PROOF: The statement follows from the martingale approach arguments that lead to Theorem 1 in Detemple et al. (2003) (see also, e.g., Karatzas et al. 1987 and Cox and Huang 1989). The hedge component here includes both the interest rate and price of risk hedge components. □

In what follows, we prove Theorem EC.1 by two parts. First, we prove the relationship $\lambda_t^* = \lambda_0^*\xi_t$. Second, we establish the decompositions of the optimal policy (EC.3.6) with components (EC.3.7a) – (EC.3.7b).

We first briefly prove the relationship

$$\lambda_t^* = \lambda_0^*\xi_t. \tag{EC.3.15}$$

As a foundation, the existence and uniqueness of λ_t^*, as the solution to equation (EC.3.3), follow from standard calculus: the utilities $u(t,\cdot)$ and $U(t,\cdot)$ are strictly increasing and concave with $\lim_{x\to\infty}\partial u(t,x)/\partial x = 0$ and $\lim_{x\to\infty}\partial U(T,x)/\partial x = 0$ (see similar discussions in Cox and Huang 1989). We now proceed to show the relationship (EC.3.15). Assuming the investor follows the optimal policy in the completed market, we follow Karatzas et al. (1987) and Cox and Huang (1989) to derive that the time–t optimal wealth satisfies $\xi_t X_t = E_t\left[\xi_T I^U(T,\lambda_0^*\xi_T) + \int_t^T \xi_s I^u(s,\lambda_0^*\xi_s)ds\right]$, where λ_0^* is characterized by (EC.3.14). By dividing ξ_t on both sides of the above equation and using the relation $\xi_s = \xi_t\xi_{t,s}$ for any $s \geq t$, we obtain

$X_t = E_t\big[\xi_{t,T} I^U(T, \lambda_0^* \xi_t \xi_{t,T}) + \int_t^T \xi_{t,s} I^u(s, \lambda_0^* \xi_t \xi_{t,s}) ds\big]$. By the definition of $\mathcal{G}_{t,T}(\cdot)$ in (EC.3.4), the previous equation is equivalent to $X_t = E_t[\mathcal{G}_{t,T}(\lambda_0^* \xi_t)]$. By the uniqueness of solution to equation (EC.3.3), we establish the relationship (EC.3.15), i.e., $\lambda_t^* = \lambda_0^* \xi_t$.

We now proceed to prove the decomposition of the optimal policy given in (EC.3.7a) – (EC.3.7b). Since we apply Lemma EC.1 to the completed market (EC.3.12) with the total price of risk θ_s^c given by $\theta_s^c = \theta^h(s, Y_s) + \theta_s^u$ in (8), the components $\mathcal{Q}_{t,T}(\lambda_0^* \xi_t)$ and $\mathcal{H}_{t,T}(\lambda_0^* \xi_t)$ in (EC.3.13) of Lemma EC.1 exactly coincide with the components $\mathcal{Q}_{t,T}(\lambda_t^*)$ and $\mathcal{H}_{t,T}(\lambda_t^*)$ in (EC.3.8a) – (EC.3.8b) of Theorem EC.1. To see this, by the relationship $\lambda_t^* = \lambda_0^* \xi_t$ in (EC.3.15), we can substitute $\lambda_0^* \xi_t$ in $\mathcal{Q}_{t,T}(\lambda_0^* \xi_t)$ and $\mathcal{H}_{t,T}(\lambda_0^* \xi_t)$ by the time–t multiplier λ_t^*. Following the above discussions, we can represent the optimal policy (π_t, π_t^F) in (EC.3.13) for the completed market as

$$(\pi_t, \pi_t^F)^\top = -\frac{1}{X_t} (\sigma^c(t, Y_t)^\top)^{-1} \left(\theta_t^c E_t[\mathcal{Q}_{t,T}(\lambda_t^*)] + E_t[\mathcal{H}_{t,T}(\lambda_t^*)]\right), \tag{EC.3.16}$$

where $\theta_t^c = \theta^h(t, Y_t) + \theta_t^u$. Here, in (EC.3.16), the components $\mathcal{Q}_{t,T}(\lambda_t^*)$ and $\mathcal{H}_{t,T}(\lambda_t^*)$ are given by (EC.3.8a) and (EC.3.8b), respectively. Next, we combine (EC.3.16) with the following algebraic fact:

$$(\sigma^c(t, Y_t)^\top)^{-1} = (\sigma^c(t, Y_t)^{-1})^\top = ((\sigma(t, Y_t)^+)^\top, (\sigma^f(t, Y_t)^+)^\top)^\top; \tag{EC.3.17}$$

the second equality follows from $\sigma^c(t, Y_t)^{-1} = (\sigma(t, Y_t)^+, \sigma^f(t, Y_t)^+)$, which can be obtained by the orthogonal condition (7). We explicitly represent the optimal policy for real assets as

$$\pi_t = -\frac{1}{X_t} (\sigma(t, Y_t)^+)^\top \left(\theta_t^c E_t[\mathcal{Q}_{t,T}(\lambda_t^*)] + E_t[\mathcal{H}_{t,T}(\lambda_t^*)]\right). \tag{EC.3.18}$$

We can further simplify this expression using the following algebraic fact

$$(\sigma(t, Y_t)^+)^\top \theta_t^c = (\sigma(t, Y_t)^+)^\top \theta^h(t, Y_t), \tag{EC.3.19}$$

with $\theta^h(t, Y_t)$ defined in (9a). To verify this, we use definition of Moore-Penrose inverse for $\sigma(t, Y_t)^+$, the orthogonal condition in (7), as well as the definition of θ_t^u in (9b) to deduce that

$$(\sigma(t, Y_t)^+)^\top \theta_t^u = (\sigma(t, Y_t) \sigma(t, Y_t)^\top)^{-1} \sigma(t, Y_t) \theta_t^u = 0_m.$$

Then by (8), we can compute the term $(\sigma(t, Y_t)^+)^\top \theta_t^c$ in (EC.3.18) as

$$(\sigma(t, Y_t)^+)^\top \theta_t^c = (\sigma(t, Y_t)^+)^\top (\theta^h(t, Y_t) + \theta_t^u) = (\sigma(t, Y_t)^+)^\top \theta^h(t, Y_t).$$

Hence, by (EC.3.19), we can further simplify the representations (EC.3.18) as

$$\pi_t = -\frac{1}{X_t} (\sigma(t, Y_t)^+)^\top \left(\theta^h(t, Y_t) E_t[\mathcal{Q}_{t,T}(\lambda_t^*)] + E_t[\mathcal{H}_{t,T}(\lambda_t^*)]\right). \tag{EC.3.20}$$

The decomposition (EC.3.6) given by (EC.3.7a) and (EC.3.7b) of the optimal policy π_t for real assets directly follows the representation (EC.3.20).

EC.3.2. Proof of Proposition EC.3

First, we verify that the proper investor-specific price of risk θ_v^u can be characterized as the solution to the optimization problem (EC.3.10). This verification follows by linking the least favorable completion approach of Karatzas et al. (1991) and the minimax local martingale approach of He and Pearson (1991), two independently developed martingale approaches for solving optimal portfolios under incomplete market settings. By Theorem 9.3 of Karatzas et al. (1991), the investor-specific price of risk θ_v^u satisfying (12) must lead to the smallest utility among all possible completions, i.e., the least favorable completion. More precisely, the desired θ_v^u satisfying (12) serves as the optimizer for the following dual problem

$$\inf_{\theta^u \in \text{Ker}(\sigma)} \left\{ \sup_{(c_t, X_T) \in \mathcal{A}_{\theta^u}} E\left[\int_0^T u(t, c_t) dt + U(T, X_T) \right] \right\}, \tag{EC.3.21}$$

where $\mathcal{A}_{\theta^u} = \{(c_t, X_T) : E[\int_0^T \xi_t c_t dt + \xi_T X_T] \leq X_0$ and $X_t \geq 0$ for all $t \in [0,T]\}$. The constraint $\theta^u \in$Ker(σ) corresponds to the orthogonal condition in (7). Problem (EC.3.21) is also discussed in He and Pearson (1991) for the same goal of characterizing the optimal portfolio in the incomplete market case, though the language of He and Pearson (1991) hinges on the class of arbitrage-free state prices, which indeed correspond to the state price density ξ_t of the completed market defined by (10). According to Theorem 2 and the discussion prior to Theorem 7 of He and Pearson (1991), the solution of problem (EC.3.21) also solves the following optimization problem (EC.3.22):

$$\inf_{\theta^u \in \text{Ker}(\sigma)} E\left[\int_0^T \tilde{u}(v, \lambda_v^*) dv + \tilde{U}(T, \lambda_T^*) \right], \tag{EC.3.22}$$

where $\tilde{u}(t,y)$ and $\tilde{U}(t,y)$ denote the conjugates of utility functions $u(t,x)$ and $U(t,x)$. We can verify that the conjugates $\tilde{u}(t,y)$ and $\tilde{U}(t,y)$ take their maximum at $x = I^u(t,y)$ and $x = I^U(t,y)$ respectively. To summarize, by linking the problems (EC.3.21) and (EC.3.22), we verify that the desired investor-specific price of risk θ_v^u for the least favorable completion (12) is also the solution of the optimization problem (EC.3.22). Then, according to the principle of dynamic programming, we can solve the optimal θ^u at arbitrary time v from the time-v version of problem (EC.3.22), which leads to the characterization in (EC.3.10)

Next, to develop equation (EC.3.11) that governs the investor-specific price risk, we decompose the optimal policy for the fictitious assets and then invoke the least favorable completion condition (EC.3.5). Similar to Theorem EC.1, we can decompose the optimal portfolio policy for the fictitious assets as $\pi^F(t, X_t, Y_t) = \pi^{mv,F}(t, X_t, Y_t) + \pi^{h,F}(t, X_t, Y_t)$, where the mean-variance component and hedge component can be expressed as:

$$\pi^{mv,F}(t, X_t, Y_t) = -\frac{1}{X_t}(\sigma^f(t, Y_t)^+)^\top \theta_t^u E_t[\mathcal{Q}_{t,T}(\lambda_t^*)], \tag{EC.3.23a}$$

and

$$\pi^{h,F}(t, X_t, Y_t) = -\frac{1}{X_t}(\sigma^f(t, Y_t)^+)^\top E_t[\mathcal{H}_{t,T}(\lambda_t^*)]; \tag{EC.3.23b}$$

the terms $\mathcal{Q}_{t,T}(\lambda_t^*)$ and $\mathcal{H}_{t,T}(\lambda_t^*)$ are defined in (EC.3.8a) and (EC.3.8b), respectively. By the least favorable completion (EC.3.5), the optimal policy for the fictitious assets should be equal to zero, i.e., $\pi^{mv,F}(t, X_t, Y_t) + \pi^{h,F}(t, X_t, Y_t) = 0_{d-m}$. Plugging in the components in (EC.3.23a) and (EC.3.23b), we can characterize the investor-specific price of risk θ_t^u by

$$(\sigma^f(t, Y_t)^+)^\top \theta_t^u E_t[\mathcal{Q}_{t,T}(\lambda_t^*)] = -(\sigma^f(t, Y_t)^+)^\top E_t[\mathcal{H}_{t,T}(\lambda_t^*)]. \tag{EC.3.24}$$

As it holds for any $t \in [0, T]$, we can replace t by v and derive that

$$(\sigma^f(v, Y_v)^+)^\top \theta_v^u = -(\sigma^f(v, Y_v)^+)^\top \frac{E_v[\mathcal{H}_{v,T}(\lambda_v^*)]}{E_v[\mathcal{Q}_{v,T}(\lambda_v^*)]}. \tag{EC.3.25}$$

Since $(\sigma^f(v, Y_v)^+)^\top$ is a $(d-m) \times d$ matrix, (EC.3.25) provides $(d-m)$ equations governing the $d-$dimensional column vector θ_v^u. We get the other m equations governing θ_v^u out of the orthogonal condition (7). By (9b) and (7), we can obtain

$$\left(\sigma(v, Y_v)^+\right)^\top \theta_v^u \equiv 0_m. \tag{EC.3.26}$$

By combining (EC.3.25) and (EC.3.26), the function θ_v^u solves

$$\theta_v^u = -\begin{pmatrix} (\sigma(v, Y_v)^+)^\top \\ (\sigma^f(v, Y_v)^+)^\top \end{pmatrix}^{-1} \begin{pmatrix} 0_{m\times d} \\ (\sigma^f(v, Y_v)^+)^\top \end{pmatrix} \cdot \frac{E_v[\mathcal{H}_{v,T}(\lambda_v^*)]}{E_v[\mathcal{Q}_{v,T}(\lambda_v^*)]}. \tag{EC.3.27}$$

We now further simplify the above equation. By (EC.3.17), we have

$$\begin{pmatrix} (\sigma(v, Y_v)^+)^\top \\ (\sigma^f(v, Y_v)^+)^\top \end{pmatrix}^{-1} = \sigma^c(v, Y_v)^\top = (\sigma(v, Y_v)^\top \;\; \sigma^f(v, Y_v)^\top). \tag{EC.3.28}$$

Plugging this into equation (EC.3.27), we can further simplify it as

$$\theta_v^u = -\sigma^f(v, Y_v)^\top (\sigma^f(v, Y_v)^+)^\top \frac{E_v[\mathcal{H}_{v,T}(\lambda_v^*)]}{E_v[\mathcal{Q}_{v,T}(\lambda_v^*)]}. \tag{EC.3.29}$$

By the definition of Moore-Penrose inverse, we can simplify the coefficient in the above equation as

$$\sigma^f(v, Y_v)^\top (\sigma^f(v, Y_v)^+)^\top = \sigma^f(v, Y_v)^\top (\sigma^f(v, Y_v)\sigma^f(v, Y_v)^\top)^{-1} \sigma^f(v, Y_v) = \sigma^f(v, Y_v)^+ \sigma^f(v, Y_v). \tag{EC.3.30}$$

Besides, by (EC.3.28), we note that

$$I_d = (\sigma(v, Y_v)^+ \;\; \sigma^f(v, Y_v)^+) \begin{pmatrix} \sigma(v, Y_v) \\ \sigma^f(v, Y_v) \end{pmatrix} \equiv \sigma(v, Y_v)^+ \sigma(v, Y_v) + \sigma^f(v, Y_v)^+ \sigma^f(v, Y_v). \tag{EC.3.31}$$

Combining (EC.3.31) with (EC.3.30), we get

$$\sigma^f(v, Y_v)^\top (\sigma^f(v, Y_v)^+)^\top = \sigma^f(v, Y_v)^+ \sigma^f(v, Y_v) = I_d - \sigma(v, Y_v)^+ \sigma(v, Y_v). \tag{EC.3.32}$$

Then, (EC.3.11) follows by plugging (EC.3.32) into (EC.3.29).

EC.4. Proof for Section 3

EC.4.1. Martingale Method under the HARA and CRRA Utility

We first describe the martingale approach for solving our dynamic portfolio allocation problem under the HARA and CRRA utilities in the completed market. It starts by formulating the dynamic problem (4) with information up to time t as the following equivalent static optimization problem:

$$\sup_{(c_t, X_T)} E_t\left[\int_t^T we^{-\rho t}\frac{(c_s-\bar{c}_s)^{1-\gamma}}{1-\gamma}dt+(1-w)e^{-\rho T}\frac{(X_T-\bar{x}_T)^{1-\gamma}}{1-\gamma}\right] \text{ s.t. } E_t\left[\int_t^T \xi_s c_s ds+\xi_T X_T\right]\leq X_t, \tag{EC.4.1}$$

where X_t is the wealth level assuming that the investor always follows the optimal policy. Under the HARA utility, the intermediate consumption and terminal wealth must satisfy the subsistence constraints, i.e., $c_t > \bar{c}_t$ and $X_T > \bar{x}_T$. We assume it is feasible for the investor by imposing the following condition on the investor initial wealth level: $\int_0^T \bar{c}_s E\left[\xi_s\right] ds + \bar{x}_T E\left[\xi_T\right] < X_0$.

With the static formulation (EC.4.1), we can represent the optimal intermediate consumption and terminal wealth following the standard method of Lagrangian multiplier (see Cox and Huang 1989): $c_t = \bar{c}_t + w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}}\left(\lambda_t^*\right)^{-\frac{1}{\gamma}}$ and $X_T = \bar{x}_T + (1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}}\left(\lambda_T^*\right)^{-\frac{1}{\gamma}}$, where the quantity λ_t^* denotes the Lagrangian multiplier for the wealth constraint in (EC.4.1). It is uniquely characterized by

$$X_t - E_t\left[\int_t^T \bar{c}_s \xi_{t,s} ds + \bar{x}_T \xi_{t,T}\right] = \left(\lambda_t^*\right)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}], \tag{EC.4.2}$$

where the scalar function $\tilde{\mathcal{G}}_{t,T}$ is defined in (13). By (EC.4.2), we can determine the multiplier λ_t^* with information up to time t. Similar formulation goes for the CRRA utility problem (5) by setting $\bar{c}_t \equiv 0$ and $\bar{x}_T = 0$ in (EC.4.1) and (EC.4.2).

EC.4.2. Proof of Proposition 1

To develop the optimal policy under the CRRA utility, we follow the general decomposition established in Theorem EC.1 for general utilities, and then substantially simplify the results based on the special structures of the CRRA utility. Under the CRRA utility in (5), we set $u(t,c) = we^{-\rho t}\frac{c^{1-\gamma}}{1-\gamma}$ and $U(T,x) = (1-w)e^{-\rho T}\frac{x^{1-\gamma}}{1-\gamma}$. in (EC.3.1). Then, the inverse marginal utility functions becomes

$$I^u(t,y) = w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}} y^{-\frac{1}{\gamma}} \text{ and } I^U(T,y) = (1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} y^{-\frac{1}{\gamma}}.$$

First, we note the following algebraic fact: with the specification of the CRRA utility in (5), the functions $\mathcal{Q}_{t,T}\left(\lambda_t^*\right)$, $\mathcal{H}_{t,T}\left(\lambda_t^*\right)$, and $\mathcal{G}_{t,T}\left(\lambda_t^*\right)$ defined in (EC.3.8a), (EC.3.8b), and (EC.3.4) are simplified to the following separable forms:

$$\mathcal{Q}_{t,T}\left(\lambda_t^*\right) = \left(\lambda_t^*\right)^{-\frac{1}{\gamma}}\tilde{\mathcal{Q}}_{t,T},\ \mathcal{H}_{t,T}\left(\lambda_t^*\right) = \left(1-\frac{1}{\gamma}\right)\left(\lambda_t^*\right)^{-\frac{1}{\gamma}}\tilde{\mathcal{H}}_{t,T}, \text{ and } \mathcal{G}_{t,T}(\lambda_t^*) = \left(\lambda_t^*\right)^{-\frac{1}{\gamma}}\tilde{\mathcal{G}}_{t,T}, \tag{EC.4.3}$$

with $\tilde{\mathcal{G}}_{t,T}$ and $\tilde{\mathcal{H}}_{t,T}$ defined in (13) and (14), respectively; the function $\tilde{\mathcal{Q}}_{t,T}$ is given by

$$\tilde{\mathcal{Q}}_{t,T} = -\frac{1}{\gamma}\tilde{\mathcal{G}}_{t,T}. \tag{EC.4.4}$$

Throughout the proof, we drop the dependence of $\tilde{\mathcal{G}}_{t,T}(\theta^u)$ and $\tilde{\mathcal{H}}_{t,T}(\theta^u)$ on the process θ^u to ease notation. With the separable forms, the wealth equation in (EC.3.3), i.e., $X_t = E_t[\mathcal{G}_{t,T}]$, is equivalent to

$$X_t = (\lambda_t^*)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}]. \tag{EC.4.5}$$

Combining (EC.4.3) and (EC.4.4), we have

$$E_t[\mathcal{Q}_{t,T}] = (\lambda_t^*)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{Q}}_{t,T}] = -(\lambda_t^*)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}]/\gamma. \tag{EC.4.6}$$

For the mean-variance component, which is given by $\pi_C^{mv}(t, X_t, Y_t) = -(\sigma(t,Y_t)^+)^\top \theta^h(t,Y_t) E_t[\mathcal{Q}_{t,T}]/X_t$ for general models in (EC.3.7a), we plug in (EC.4.6) under CRRA utility to get

$$\pi_C^{mv}(t, X_t, Y_t) = (\lambda_t^*)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}](\sigma(t,Y_t)^+)^\top \theta^h(t,Y_t)/(\gamma X_t) = (\sigma(t,Y_t)^+)^\top \theta^h(t,Y_t)/\gamma,$$

where the second equality follows by (EC.4.5). Note that the wealth level X_t vanishes in the mean-variance component. By the definition of $\theta^h(t,Y_t)$ in (9a), we obtain the representation in (16) as

$$\pi_C^{mv}(t, Y_t) = (\sigma(t,Y_t)^+)^\top \theta^h(t,Y_t)/\gamma = (\sigma(t,Y_t)\sigma(t,Y_t)^\top)^{-1}\left(\mu(t,Y_t) - r(t,Y_t)1_m\right)/\gamma.$$

Similarly, the optimal hedge component, which is given by $\pi^h(t, X_t, Y_t) = -(\sigma(t,Y_t)^+)^\top E_t[\mathcal{H}_{t,T}]/X_t$ for general models in (EC.3.7b), we plug in (EC.4.3) to get

$$\pi^h(t, X_t, Y_t) = -\left(1 - \frac{1}{\gamma}\right)(\lambda_t^*)^{-\frac{1}{\gamma}}(\sigma(t,Y_t)^+)^\top E_t[\tilde{\mathcal{H}}_{t,T}]/X_t.$$

Then (17) follows by plugging in $(\lambda_t^*)^{-\frac{1}{\gamma}} = X_t/E_t[\tilde{\mathcal{G}}_{t,T}]$ by (EC.4.5). For the optimal consumption, we have $c_t = w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}} (\lambda_t^*)^{-\frac{1}{\gamma}}$ by Theorem EC.1. By (EC.4.5), we derive the optimal consumption level c_t in (18) and wealth-consumption ratio $\phi(t, Y_t)$ in (18). Finally, Equation (19) for governing the investor-specific price of risk θ_v^u can be derived by specifying its general-model counterpart (EC.3.11) under the CRRA utility, and invoking the separable forms in (EC.4.3), wealth constraint (EC.4.5), and the relationship (EC.4.6).

We verify the optimal policy under CRRA utility is independent of investor wealth level. It is straightforward that the mean-variance component $\pi_C^{mv}(t, Y_t)$ in (16) is independent of wealth level. We next prove for the hedge component $\pi_C^h(t, Y_t)$. With the CRRA utility in (5), we can specify the dual problem (EC.3.22) as:

$$\inf_{\theta^u \in \mathrm{Ker}(\sigma)} E_v\left[(1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} (\lambda_T^*)^{1-\frac{1}{\gamma}} + w^{\frac{1}{\gamma}} \int_v^T e^{-\frac{\rho s}{\gamma}} (\lambda_s^*)^{1-\frac{1}{\gamma}} ds\right]. \tag{EC.4.7}$$

Using the relationship $\lambda_s^* = \lambda_0^* \xi_s = \lambda_v^* \xi_{v,s}$ as well as the fact that the multiplier λ_v^* is known with information available up to time v, we can extract the factor $(\lambda_v^*)^{1-\frac{1}{\gamma}}$ from the conditional expectation in (EC.4.7) to get $\inf_{\theta^u\in\mathrm{Ker}(\sigma)}(\lambda_v^*)^{1-\frac{1}{\gamma}}E_v\left[(1-w)^{\frac{1}{\gamma}}e^{-\frac{\rho T}{\gamma}}(\xi_{v,T})^{1-\frac{1}{\gamma}}+w^{\frac{1}{\gamma}}\int_v^T e^{-\frac{\rho s}{\gamma}}(\xi_{v,s})^{1-\frac{1}{\gamma}}ds\right]$. According to He and Pearson (1991), as the Lagrangian multiplier of the static optimization problem (EC.3.2), λ_0^* must be positive. Thus, $\lambda_v^*=\lambda_0^*\xi_v$ is also positive. Then, we can drop the factor $(\lambda_v^*)^{1-\frac{1}{\gamma}}$ in this optimization problem. Besides, the process $(Y_s,\xi_{v,s})$ for $v\le s\le T$ is Markovian with the initial value $(Y_v,1)$. By the feedback control law, we find that θ_v^u only depends on the time v and state variable Y_v, and is independent of the multiplier λ_v^*. It implies θ_v^u does not depend on investor wealth level under the CRRA utility. By (10) and (15), we conclude the state price density $\xi_{t,s}$ and the term $H_{t,s}$ are also independent of investor wealth. Then, by (13) and (14), the building blocks $\tilde{\mathcal{G}}_{t,T}$ and $\tilde{\mathcal{H}}_{t,T}$, thus the optimal hedge component $\pi_C^h(t,Y_t)$ in (17), are all independent of investor wealth level.

EC.4.3. Proof of Theorem 1

Under the HARA utility in (4), we set $u(t,c)=we^{-\rho t}\frac{(c-\bar{c}_t)^{1-\gamma}}{1-\gamma}$ and $U(T,x)=(1-w)e^{-\rho T}\frac{(x-\bar{x}_T)^{1-\gamma}}{1-\gamma}$. in (EC.3.1). Then, the inverse marginal utility functions becomes

$$I^u(t,y)=w^{\frac{1}{\gamma}}e^{-\frac{\rho t}{\gamma}}y^{-\frac{1}{\gamma}}+\bar{c}_t \text{ and } I^U(T,y)=(1-w)^{\frac{1}{\gamma}}e^{-\frac{\rho T}{\gamma}}y^{-\frac{1}{\gamma}}+\bar{x}_T.$$

We use the following algebraic fact: under the HARA utility (4), the functions $\mathcal{Q}_{t,T}(\lambda_t^*)$, $\mathcal{H}_{t,T}(\lambda_t^*)$, and $\mathcal{G}_{t,T}(\lambda_t^*)$ defined in (EC.3.8a), (EC.3.8b), and (EC.3.4) for general models can be expressed as follows:

$$\mathcal{G}_{t,T}(\lambda_t^*)=(\lambda_t^*)^{-\frac{1}{\gamma}}\tilde{\mathcal{G}}_{t,T}+\bar{x}_T\xi_{t,T}+\int_t^T\bar{c}_s\xi_{t,s}ds,\ \mathcal{Q}_{t,T}(\lambda_t^*)=(\lambda_t^*)^{-\frac{1}{\gamma}}\tilde{\mathcal{Q}}_{t,T}\ , \tag{EC.4.8a}$$

and

$$\mathcal{H}_{t,T}(\lambda_t^*)=\left(1-\frac{1}{\gamma}\right)(\lambda_t^*)^{-\frac{1}{\gamma}}\tilde{\mathcal{H}}_{t,T}+\bar{x}_T\xi_{t,T}H_{t,T}+\int_t^T\bar{c}_s\xi_{t,s}H_{t,s}ds, \tag{EC.4.8b}$$

where $\tilde{\mathcal{G}}_{t,T}$ and $\tilde{\mathcal{H}}_{t,T}$ are defined in (13) and (14); $\tilde{\mathcal{Q}}_{t,T}=-\tilde{\mathcal{G}}_{t,T}/\gamma$ as in (EC.4.4). In the above equations, we plug in the investor-specific price of risk θ_v^u under the HARA utility, which can be different from its CRRA counterpart. Then, we can specify the decomposition of optimal policy in Theorem EC.1 for general models as

$$\pi_H^{mv}(t,X_t,Y_t)=-\frac{1}{X_t}(\sigma(t,Y_t)^+)^\top\theta^h(t,Y_t)(\lambda_t^*)^{-\frac{1}{\gamma}}E_t[\tilde{\mathcal{Q}}_{t,T}] \tag{EC.4.9a}$$

for the mean-variance component and

$$\pi_H^h(t,X_t,Y_t)=-\frac{1}{X_t}(\sigma(t,Y_t)^+)^\top\left(\left(1-\frac{1}{\gamma}\right)(\lambda_t^*)^{-\frac{1}{\gamma}}E_t[\tilde{\mathcal{H}}_{t,T}]+\Psi(t,X_t,Y_t)\right) \tag{EC.4.9b}$$

for the hedge component. Here, $\Psi(t,X_t,Y_t)$ is defined in (24). In particular, we can show $\Psi(t,X_t,Y_t)=\mathcal{D}_tZ_{t,T}$, where $\mathcal{D}_t$ denotes the Malliavin derivative with respect to the Brownian motion W_t; Z_t is defined in (22). Thus, $\Psi(t,X_t,Y_t)$ essentially measures the sensitivity of $Z_{t,T}$ to the underlying Brownian motion. In addition, the wealth constraint (EC.3.3) for solving the multiplier λ_t^* specifies to

$$(\lambda_t^*)^{-\frac{1}{\gamma}}E_t[\tilde{\mathcal{G}}_{t,T}]+\bar{x}_TE_t\left[\xi_{t,T}\right]+E_t\left[\int_t^T\bar{c}_s\xi_{t,s}ds\right]=X_t. \tag{EC.4.10}$$

By Theorem EC.1, the optimal consumption level c_t under HARA utility can be expressed as

$$c_t = \bar{c}_t + w^{\frac{1}{\gamma}} e^{-\frac{\rho t}{\gamma}} (\lambda_t^*)^{-\frac{1}{\gamma}}. \tag{EC.4.11}$$

By the definition of $\bar{X}_t$ in (21), we can simplify (EC.4.10) as

$$(\lambda_t^*)^{-\frac{1}{\gamma}} E_t[\tilde{\mathcal{G}}_{t,T}] = \bar{X}_t, \tag{EC.4.12}$$

which can be viewed as the counterpart to the wealth constraint (EC.4.5) under CRRA utility. Plugging this into (EC.4.9a) and using the relationship $\tilde{\mathcal{Q}}_{t,T} = -\tilde{\mathcal{G}}_{t,T}/\gamma$, we have

$$\pi_H^{mv}(t, X_t, Y_t) = \frac{\bar{X}_t}{\gamma X_t} (\sigma(t, Y_t)^+)^\top \theta^h(t, Y_t).$$

Then, (20) can be obtained by plugging in the definition of $\theta^h(t, Y_t)$ in (9a) and comparing to the mean-variance component under CRRA utility in (16). Similarly, the optimal hedge component $\pi_H^h(t, X_t, Y_t)$ in (23) and consumption level c_t in (25) follow by plugging $(\lambda_t^*)^{-\frac{1}{\gamma}} = \bar{X}_t / E_t[\tilde{\mathcal{G}}_{t,T}]$ from (EC.4.12) into (EC.4.9b) and (EC.4.11), respectively.

Finally, we derive Equation (26) for characterizing the investor-specific price of risk θ_v^u under the HARA utility. Plugging (EC.4.8a) and (EC.4.8b) into Equation (EC.3.11) for general models in Proposition EC.3, we have

$$\theta_v^u = \frac{\sigma(v, Y_v)^+ \sigma(v, Y_v) - I_d}{(\lambda_v^*)^{-\frac{1}{\gamma}} E_v[\tilde{\mathcal{Q}}_{v,T}]} \times \left(\left(1 - \frac{1}{\gamma}\right) (\lambda_v^*)^{-\frac{1}{\gamma}} E_v[\tilde{\mathcal{H}}_{v,T}] + \Psi(v, X_v, Y_v) \right).$$

Then, Equation (26) follows by plugging in $\tilde{\mathcal{Q}}_{v,T} = -\tilde{\mathcal{G}}_{v,T}/\gamma$ and $(\lambda_v^*)^{-\frac{1}{\gamma}} = \bar{X}_v / E_v[\tilde{\mathcal{G}}_{v,T}]$ from (EC.4.12).

We briefly discuss the existence of solution to Equations (19) and (26). First, we notice that if we can find a solution to the equation system, then the existence is proved by construction. This is the case for the CIRH-SVSIR model considered in our comparative analysis (see Section 4.4). For general cases, we can show that the dual problem (EC.3.22) is differentiable to the multiplier λ_t^* under the CRRA and HARA utilities (see its expressions in (EC.4.7) and (EC.5.14)). Thus, Assumption T in He and Pearson (1991) is satisfied. Then, Lemma 2 and Theorem 2 in He and Pearson (1991) state that if problem (EC.3.22) has a solution, the optimization problems (EC.3.21) and (EC.3.1) also have a solution. In this case, the corresponding θ^u, which determines the minimax local martingale measure, satisfies the equation (19) or (26) by least favorable completion principle in Karatzas et al. (1991). Section 4 of He and Pearson (1991) discusses the conditions for the existence of solution to the dual problem (EC.3.22), although they are usually abstract for general models. In the case with only terminal wealth in investor's utility (i.e., $w = 0$ in (4) and (5)), we can apply Theorem 4 in He and Pearson (1991) to show the existence of solution under the HARA and CRRA utilities when we assume the relative risk aversion coefficient $\gamma > 1$.

EC.5. Proof for Sections 4 and EC.1.2

EC.5.1. Fictitious Completion under Hedgeable Interest Rate

We first introduce the least favorable completion under the set-up (28) – (31), i.e., with a complete market for interest rate risk. For the $m-d_r$ stock assets, we decompose their volatility matrix $\sigma^{(2)}(t,y) = \big(\sigma_1^{(2)}(t,y)^\top, \sigma_2^{(2)}(t,y)^\top, ..., \sigma_{m-d_r}^{(2)}(t,y)^\top\big)^\top$ as $\sigma^{(2)}(t,y) = \big(\sigma_r^{(2)}(t,y),\ \sigma_o^{(2)}(t,y)\big)$, where $\sigma_r^{(2)}(t,y)$ (resp. $\sigma_o^{(2)}(t,y)$) is a $(m-d_r)\times d_r$ (resp. $(m-d_r)\times(d-d_r)$) dimensional matrix, denoting the volatility on the Brownian motion W_t^r (resp. W_t^o).

Combining (30) and (31), the drift and volatility for all the m risky assets, including both bonds and stocks, are given by

$$\mu(t,Y_t) = \begin{pmatrix} \mu^{(1)}(t,Y_t^r) \\ \mu^{(2)}(t,Y_t) \end{pmatrix} \text{ and } \sigma(t,Y_t) = \begin{pmatrix} \sigma^{(1)}(t,Y_t^r) & 0_{d_r\times(d-d_r)} \\ \sigma_r^{(2)}(t,Y_t) & \sigma_o^{(2)}(t,Y_t) \end{pmatrix}. \tag{EC.5.1}$$

Since we have d_r bonds with returns driven by d_r Brownian motions, the market for interest rate risk is complete as the uncertainty in W_t^r can be fully hedged. On the other hand, the stock market is incomplete since we have more Brownian motions than the number of stocks. By simple algrebraic calculation, the market price of risk (9a) of the full Brownian motion $W_t = (W_t^r, W_t^o)$ can be expressed as:

$$\theta^h(t,Y_t) = \begin{pmatrix} \theta^r(t,Y_t^r) \\ \theta^o(t,Y_t) \end{pmatrix}, \tag{EC.5.2}$$

where the market price of interest rate risk $\theta^r(t,Y_t^r)$ follows by (32); $\theta^o(t,Y_t)$ is the market price of risk associated with the Brownian motion W_t^o, given by

$$\theta^o(t,Y_t) = \sigma_o^{(2)}(t,Y_t)^+\big(\mu^{(2)}(t,Y_t) - r(t,Y_t^r)1_{m-d_r} - \sigma_r^{(2)}(t,Y_t)\theta^r(t,Y_t^r)\big).$$

The premium for interest rate risk, $\sigma_r^{(2)}(t,Y_t)\theta^r(t,Y_t^r)$, is subtracted from the stock return $\mu^{(2)}(t,Y_t)$ in the above.

To solve the investor optimization problem, we introduce $d-m$ fictitious assets to complete the market. As the market for interest rate risk W_t^r is already complete with the bonds, we only need to use the fictitious assets to hedge the Brownian motion W_t^o related with stocks. Thus, we can assume the prices of fictitious assets follow the dynamics

$$\frac{dF_t}{F_t} = \mu_t^f dt + \sigma^f(t,Y_t)dW_t^o,$$

where $\sigma^f(t,Y_t)$ is a $(d-m)\times(d-d_r)$ dimensional matrix satisfying $\sigma^f(t,Y_t)\sigma_o^{(2)}(t,Y_t)^\top = 0_{(d-m)\times(m-d_r)}$. Then, the investor-specific price of risk associated with W_t^r is given by

$$\theta_t^u = \sigma^f(t,Y_t)^+\big(\mu_t^f - r(t,Y_t^r)1_{d-m}\big),$$

which is a $(d-d_r)$–dimensional vector. By (8) and (EC.5.2), the total price of risk in the completed market specifies to

$$\theta_t^c = \begin{pmatrix} \theta^r(t,Y_t^r) \\ \theta^o(t,Y_t) + \theta_t^u \end{pmatrix}.$$

That is, the investor-specific price of risk θ_t^u only appears in the price of stock-related risk for W_t^o. Then by (10), the relative state price density $\xi_{t,s}$ can be decomposed as

$$\xi_{t,s} = \eta_{t,s}^r \eta_{t,s}^o,$$

where $\eta_{t,s}^r$ is given in (34) as

$$\eta_{t,s}^r = \exp\left(-\int_t^s r_v dv - \int_t^s (\theta_v^r)^\top dW_v^r - \frac{1}{2}\int_t^s (\theta_v^r)^\top \theta_v^r dv\right) \tag{EC.5.3a}$$

and

$$\eta_{t,s}^o = \exp\left(-\int_t^s (\theta_v^o + \theta_v^u)^\top dW_v^o - \frac{1}{2}\int_t^s \left[(\theta_v^o)^\top \theta_v^o + (\theta_v^u)^\top \theta_v^u\right] dv\right). \tag{EC.5.3b}$$

EC.5.2. Proof of Proposition 2

We first provide the expressions of $\mu_B(t, Y_t^r; s)$ and $\sigma_B(t, Y_t^r; s)$ in (35), which denote the drift and volatility of the instantaneous return of ZCB $B_{t,s}$. They are given by

$$\mu_B(t, Y_t^r; s) = r_t + \sigma_B(t, Y_t^r; s)\theta^r(t, Y_t^r) \tag{EC.5.4}$$

and

$$\sigma_B(t, Y_t^r; s) = -\frac{1}{B_{t,s}} E_t\left[\eta_{t,s}^r \left(\int_t^s M_{t,v}^r dv + \int_t^s M_{t,v}^\theta \left(dW_v^r + \theta_v^r dv\right)\right)^\top\right], \tag{EC.5.5}$$

where $M_{t,v}^r$ and $M_{t,v}^\theta$ are two random variables with dynamics explicitly given in (EC.1.4). They are introduced using the Malliavin derivative with respect to the Brownian motion W_t^r.

We prove Proposition 2 in the below. First, we show that

$$E_t[\xi_{t,s}] = E_t[\eta_{t,s}^r], \tag{EC.5.6}$$

with $\eta_{t,s}^r$ defined in (EC.5.3a). Let $\{\mathcal{F}_t^r\}$ and $\{\mathcal{F}_t^o\}$ be the filtration generated by Brownian motions W_t^r and W_t^o up to time t respectively and denote the σ-algebra $\mathcal{F}_t = \mathcal{F}_t^r \cup \mathcal{F}_t^o$. By tower property, we have

$$E_t[\xi_{t,s}] = E[\eta_{t,s}^r \eta_{t,s}^o | \mathcal{F}_t] = E[E[\eta_{t,s}^r \eta_{t,s}^o | \mathcal{F}_t \cup \mathcal{F}_s^r] | \mathcal{F}_t].$$

Here, the σ-algebra $\mathcal{F}_t \cup \mathcal{F}_s^r$ includes the information of path of W_t^o up to time t and the path of W_t^r up to time $s > t$. By (32) and (EC.5.3a), the value of $\eta_{t,s}^r$ is fully determined by the path of the Brownian motion W_t^r up to time s. Thus, we have $E[\eta_{t,s}^r \eta_{t,s}^o | \mathcal{F}_t \cup \mathcal{F}_s^r] = \eta_{t,s}^r E[\eta_{t,s}^o | \mathcal{F}_t \cup \mathcal{F}_s^r]$. It leads to

$$E_t[\xi_{t,s}] = E\left[\eta_{t,s}^r E[\eta_{t,s}^o | \mathcal{F}_t \cup \mathcal{F}_s^r]\right| \mathcal{F}_t]. \tag{EC.5.7}$$

We next look at the inner conditional expectation $E\left[\eta_{t,s}^o | \mathcal{F}_t \cup \mathcal{F}_s^r\right]$. By (EC.5.3b), $\eta_{t,s}^o$ satisfies the SDE:

$$d\eta_{t,s}^o = -\eta_{t,s}^o \left(\theta_s^o + \theta_s^u\right) dW_s^o.$$

with initial value $\eta^o_{t,t}=1$. Then, we have

$$\eta^o_{t,s}=1-\int_t^s \eta^s_{t,v}\left(\theta^o_v+\theta^u_v\right)dW^o_v.$$

Taking conditional expectation, we get

$$E\left[\eta^o_{t,s}|\mathcal{F}_t\cup\mathcal{F}^r_s\right]=1-E\left[\int_t^s \eta^o_{t,v}\left(\theta^o_v+\theta^u_v\right)dW^o_v|\mathcal{F}_t\cup\mathcal{F}^r_s\right].$$

Since the Brownian motions W^r_v and W^o_v are independent, the σ-algebra $\mathcal{F}_t\cup\mathcal{F}^r_s$ contains no information of W^o_v for $v\in[t,s]$. Then, by the martingale property of Ito integral, we have

$$E\left[\int_t^s \eta^o_{t,v}\left(\theta^o_v+\theta^u_v\right)dW^o_v|\mathcal{F}_t\cup\mathcal{F}^r_s\right]=0,$$

leading to $E[\eta^o_{t,s}|\mathcal{F}_t\cup\mathcal{F}^r_s]=1$. Plugging this back to (EC.5.7), we prove (EC.5.6).

In the next step, we prove

$$B_{t,s}=E_t[\eta^r_{t,s}]. \tag{EC.5.8}$$

Under the set-up in (28) – (31), the uncertainty in the interest rate can be fully hedged by investing in the d_r bond assets. Thus, a unit payment at time s can be perfectly replicated by holding a portfolio of the savings account and the bond assets. Then, by no-arbitrage principle, the ZCB price $B_{t,s}$ allows the following representation:

$$B_{t,s}=E_t\left[\exp\left(-\int_t^s r_v dv\right)\phi^r_{t,s}\right], \tag{EC.5.9}$$

which is the conditional expectation of the pricing kernel for the Brownian motion W^r_t. With a complete market for interest rate risk, the term $\phi^r_{t,s}$ in (EC.5.9) is uniquely determined by the d_r bond assets (see Karatzas and Shreve 1991) as

$$\phi^r_{t,s}=\exp\left(-\int_t^s(\theta^r_v)^\top dW^r_v-\frac{1}{2}\int_t^s(\theta^r_v)^\top\theta^r_v dv\right).$$

Plugging this into (EC.5.9), we have

$$\begin{aligned}B_{t,s}&=E_t\left[\exp\left(-\int_t^s r_v dv\right)\phi^r_{t,s}\right]\\&=E_t\left[\exp\left(-\int_t^s r_v dv-\int_t^s(\theta^r_v)^\top dW^r_v-\frac{1}{2}\int_t^s(\theta^r_v)^\top\theta^r_v dv\right)\right]=E\left[\eta^r_{t,s}\right].\end{aligned}$$

The last equality invokes the definition of $\eta^r_{t,s}$ in (EC.5.3a). Combining (EC.5.6) and (EC.5.8), we prove the relationship (33).

Finally, we derive the dynamics for the bond price $B_{t,s}$ in (35). By the definition of $\eta^r_{t,s}$ in (EC.5.3a), we can write $B_{t,s}$ as

$$B_{t,s}=\ E_t\left[\eta^r_s\right]/\eta^r_t, \tag{EC.5.10}$$

where

$$\eta_t^r = \exp\left(-\int_0^t r_v dv - \int_0^t (\theta_v^r)^\top dW_v^r - \frac{1}{2}\int_0^t (\theta_v^r)^\top \theta_v^r dv\right).$$

By Ito's lemma, we have

$$d\left(\frac{1}{\eta_t^r}\right) = \frac{1}{\eta_t^r}\left[\left(r_t + (\theta_t^r)^\top \theta_t^r\right) dt + \theta_t^r dW_t^r\right]. \tag{EC.5.11}$$

On the other hand, notice that $E_t[\eta_s^r]$ is a martingale with respect to filtration $\{\mathcal{F}_t\}$. Using Clark-Ocone formula (see Appendix D in Detemple et al. 2003), we can get

$$dE_t[\eta_s^r] = E_t[\mathcal{D}_t^r \eta_s^r]^\top dW_t^r, \tag{EC.5.12}$$

where $\mathcal{D}_t^r \eta_s^r$ is a d_r–dimensional column vector, representing the Malliavin derivative of η_s^r with respect to the Brownian motion W_t^r.

Using Malliavin calculus, the Malliavin derivative $\mathcal{D}_t^r \eta_s^r$ can be calculated as

$$\mathcal{D}_t^r \eta_s^r = -\eta_s^r \left(\theta_t^r + \int_t^s \mathcal{D}_t^r r_v dv + \int_t^s \mathcal{D}_t^r \theta_v^r (dW_v^r + \theta_v^r dv)\right).$$

Taking conditional expectation on both sides leads to:

$$\begin{aligned} E_t[\mathcal{D}_t^r \eta_s^r] &= -E_t\left[\eta_s^r \left(\theta_t^r + \int_t^s \mathcal{D}_t^r r_v dv + \int_t^s \mathcal{D}_t^r \theta_v^r (dW_v^r + \theta_v^r dv)\right)\right] \\ &= -\eta_t^r \theta_t^r E_t\left[\eta_{t,s}^r\right] - \eta_t^r E_t\left[\eta_{t,s}^r \left(\int_t^s \mathcal{D}_t^r r_v dv + \int_t^s \mathcal{D}_t^r \theta_v^r (dW_v^r + \theta_v^r dv)\right)\right]. \end{aligned}$$

Here, we use the fact that $E_t[\eta_s^r] = \eta_t^r E_t\left[\eta_{t,s}^r\right]$. Using $B_{t,s} = E_t\left[\eta_{t,s}^r\right]$ and the definition of $\sigma_B(t, Y_t^r; s)$ in (EC.5.5), we can further express $E_t[\mathcal{D}_t^r \eta_s^r]$ as

$$E_t[\mathcal{D}_t^r \eta_s^r] = \eta_t^r B_{t,s}\left(-\theta_t^r + \sigma_B(t, Y_t^r; s)^\top\right).$$

Recall that the terms $M_{t,s}^r$ and $M_{t,s}^\theta$ in (EC.5.5) are given by $M_{t,v}^r = \mathcal{D}_t^r r(v, Y_v^r)$ and $M_{t,v}^\theta = \mathcal{D}_t^r \theta^r(v, Y_v^r)$. Plugging this into (EC.5.12), we have

$$dE_t[\eta_s^r] = \eta_t^r B_{t,s}\left(-(\theta_t^r)^\top + \sigma_B(t, Y_t^r; s)\right) dW_t^r. \tag{EC.5.13}$$

Then, the dynamics of $B_{t,s}$ follows by applying Ito's formula on (EC.5.10) based on the SDEs in (EC.5.11) and (EC.5.13). After some algebraic simplification, we obtain

$$dB_{t,s} = B_{t,s}\left[(r_t + \sigma_B(t, Y_t^r; s)\theta_t^r) dt + \sigma_B(t, Y_t^r; s) dW_t^r\right].$$

It proves the dynamics of the bond price in (35).

EC.5.3. Proof for Theorem 2

We first prove that with a complete market for interest rate risk, the investor-specific price of risk coincides for the HARA and CRRA utility investors. We use the dual problem introduced in He and Pearson (1991). To begin with, using the specific function forms under HARA utility (4), we can explicitly specify the dual problem under HARA utility as

$$\inf_{\theta^u \in \mathrm{Ker}(\sigma)} E_v \left[(1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} \left(\lambda_T^*\right)^{1-\frac{1}{\gamma}} + w^{\frac{1}{\gamma}} \int_v^T e^{-\frac{\rho s}{\gamma}} \left(\lambda_s^*\right)^{1-\frac{1}{\gamma}} ds + \frac{\gamma-1}{\gamma} A_{v,T} \right], \tag{EC.5.14}$$

where $A_{v,T} = \bar{x}_T \lambda_T^* + \int_v^T \bar{c}_s \lambda_s^* ds$. On the other hand, the dual problem under CRRA utility specifies to

$$\inf_{\theta^u \in \mathrm{Ker}(\sigma)} E_v \left[(1-w)^{\frac{1}{\gamma}} e^{-\frac{\rho T}{\gamma}} \left(\lambda_T^*\right)^{1-\frac{1}{\gamma}} + w^{\frac{1}{\gamma}} \int_v^T e^{-\frac{\rho s}{\gamma}} \left(\lambda_s^*\right)^{1-\frac{1}{\gamma}} ds \right]. \tag{EC.5.15}$$

Comparing (EC.5.14) and (EC.5.15), we see that the term $A_{v,T}$ in (EC.5.14) distinguishes the dual problem under HARA utility from that under CRRA utility. When there is a bond market, we then verify that $E_v\left[A_{v,T}\right]$ does not depend on the control process θ_v^u for $v \in [v,T]$ and thus can be dropped from the dual problem (EC.5.14) to simplify it as the CRRA counterpart (EC.5.15). To see this, we use the relationship $\lambda_s^* = \lambda_0^* \xi_s = \lambda_v^* \xi_{v,s}$ to derive that

$$E_v\left[A_{v,T}\right] = \bar{x}_T E_v\left[\lambda_T^*\right] + \int_v^T \bar{c}_s E_v\left[\lambda_s^*\right] ds = \lambda_v^* \left[\bar{x}_T E_v[\xi_{v,T}] + \int_v^T \bar{c}_s E_v[\xi_{v,s}] ds \right].$$

By Proposition 2, we have $E_v[\xi_{v,T}] = B_{v,T}$ and $E_v[\xi_{v,s}] = B_{v,s}$ under the set-ups (28) – (31). It leads to

$$E_v\left[A_{v,T}\right] = \lambda_v^* [\bar{x}_T B_{v,T} + \int_v^T \bar{c}_s B_{v,s} ds],$$

which obviously does not depend on the control process θ_v^u. Thus, we can drop the term $A_{v,T}$ from (EC.5.14). By the above arguments, we show that with a complete market for interest rate risk, the investor-specific price of risk θ_v^u under HARA utility is uniquely characterized as the control process for the dual problem (EC.5.15), with the underlying Markov process $\left(Y_s, \xi_{v,s}\right)$ for $v \le s \le T$. Thus, we verify that the dual problems, as well as the underlying Markov process, coincide under the HARA and CRRA utilities. They directly lead to the same optimal control process θ_v^u. It proves that with a complete market for interest rate risk, the investor-specific price of risk θ_v^u of HARA investor is the same as that for the CRRA investor, and is thus independent of investor wealth level.

Next, we derive the optimal policy under the HARA utility. The mean-variance component (36) simply follows from (20). For the hedge component $\pi_H^h(t, X_t, Y_t)$, comparing (23) and (38), we only need to prove

$$-(\sigma(t,Y_t)^+)^\top \Psi(t, X_t, Y_t) = \begin{pmatrix} \Pi_B(t, Y_t^r) \\ 0_{m-d_r} \end{pmatrix}, \tag{EC.5.16}$$

where

$$\Psi(t, X_t, Y_t) = \bar{x}_T E_t\left[\xi_{t,T} H_{t,T}\right] + \int_t^T \bar{c}_s E_t\left[\xi_{t,s} H_{t,s}\right] ds \tag{EC.5.17}$$

following (24) and

$$\Pi_B(t, Y_t^r) = \left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} \left(\bar{x}_T B_{t,T} \sigma_B(t, Y_t^r; T)^\top + \int_t^T \bar{c}_s B_{t,s} \sigma_B(t, Y_t^r; s)^\top ds\right) \tag{EC.5.18}$$

by (39). By the property of Malliavin calculus, we can establish

$$E_t\left[\xi_{t,s} H_{t,s}\right] = -E_t\left[\mathcal{D}_t \xi_{t,s}\right] = -\mathcal{D}_t E_t\left[\xi_{t,s}\right].$$

The first equality can be verified by calculating the Malliavin derivative of $\xi_{t,s}$ (see Appendix A in Detemple et al. 2003). The second equality follows from Proposition 3.12 in Nunno et al. (2008). Further combining the above equation with (EC.5.6), we get

$$\mathcal{D}_t E_t\left[\xi_{t,s}\right] = \mathcal{D}_t E_t\left[\eta_{t,s}^r\right] = E_t\left[\mathcal{D}_t \eta_{t,s}^r\right],$$

where $\eta_{t,s}^r$ is given by (34). Note that $\eta_{t,s}^r$ only depends on the Brownian motion W_t^r. So we only need to consider the Malliavin derivative with respect to the Brownian motion W_t^r, denoted by $\mathcal{D}_t^r$. It leads to

$$E_t\left[\xi_{t,s} H_{t,s}\right] = -E_t\left[\mathcal{D}_t \eta_{t,s}^r\right] = \begin{pmatrix} -E_t\left[\mathcal{D}_t^r \eta_{t,s}^r\right] \\ 0_{d-d_r} \end{pmatrix}, \tag{EC.5.19}$$

where $\mathcal{D}_t^r \eta_{t,s}^r$ is a d_r–dimensional column vector. Using Malliavin calculus, we can obtain $\mathcal{D}_t^r \eta_{t,s}^r$ as

$$E_t\left[\mathcal{D}_t^r \eta_{t,s}^r\right] = -E_t\left[\eta_{t,s}^r \left(\int_t^s \mathcal{D}_t^r r_v dv + \int_t^s \mathcal{D}_t^r \theta_v^r \left(dW_v^r + \theta_v^r dv\right)\right)\right] = B_{t,s} \sigma_B(t, Y_t^r; s). \tag{EC.5.20}$$

The last equality uses the definition of $\sigma_B(t, Y_t; s)$ in (EC.5.5). On the other hand, given the volatility matrix $\sigma(t, Y_t)$ in (EC.5.1), we can show $(\sigma(t, Y_t)^+)^\top$ allows the following form:

$$(\sigma(t, Y_t)^+)^\top = \begin{pmatrix} \left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} & -\left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} \sigma^r(t, Y_t)^\top (\sigma^{(2)}(t, Y_t)^+)^\top \\ 0_{(d-d_r)\times d_r} & (\sigma^{(2)}(t, Y_t)^+)^\top \end{pmatrix}. \tag{EC.5.21}$$

Combining (EC.5.19), (EC.5.20), and (EC.5.21), we get

$$(\sigma(t, Y_t)^+)^\top E_t\left[\xi_{t,s} H_{t,s}\right] = -\left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} \begin{pmatrix} B_{t,s} \sigma_B(t, Y_t^r; s) \\ 0_{m-d_r} \end{pmatrix}.$$

Then, by the definition of $\Psi(t, X_t, Y_t)$ in (EC.5.17), we can obtain

$$-(\sigma(t, Y_t)^+)^\top \Psi(t, X_t, Y_t) = \left(\sigma^{(1)}(t, Y_t^r)^\top\right)^{-1} \begin{pmatrix} \bar{x}_T B_{t,T} \sigma_B(t, Y_t^r; T) + \int_t^T \bar{c}_s B_{t,s} \sigma_B(t, Y_t^r; s) ds \\ 0_{m-d_r} \end{pmatrix}.$$

Comparing with (EC.5.18), we can see the right-hand side of the above equation is exactly $\Pi_B(t, Y_t^r)$. It proves the relationship (EC.5.16) and thus Theorem 2.

EC.5.4. Proof for Proposition 3

We first prove that the value of the financing portfolio always equals that of the hypothetical bond holding scheme, i.e., $X_t^{(finan)} \equiv Z_{t,T}$ for all t. To show this, we first write out the dynamics of $Z_{t,T}$ in (37), which is the wealth of the HARA investor allocated to the hypothetical bond holding scheme. Using the SDE of the bond price $B_{t,s}$ in (35), market price of interest risk $\theta^r(t,Y_t^r)$ defined in (32), and the bond portfolio $\Pi_B(t,Y_t^r)$ in (39), we can derive

$$dZ_{t,T} = \left(-\bar{c}_t + r_t Z_{t,T} + \Pi_B(t,Y_t^r)^\top(\mu_t^{(1)} - r_t 1_{d_r})\right) dt + \Pi_B(t,Y_t^r)^\top \sigma_t^{(1)} dW_t^r, \tag{EC.5.22}$$

where r_t, $\mu_t^{(1)}$, and $\sigma_t^{(1)}$ abbreviate for $r\left(t,Y_t^r\right)$, $\mu^{(1)}\left(t,Y_t^r\right)$, and $\sigma^{(1)}\left(t,Y_t^r\right)$, respectively. Next, we get the dynamics of the wealth process of the financing portfolio $X_t^{(finan)}$. With the optimal policy in (44) and consumption rate $\bar{c}_t$, $X_t^{(finan)}$ satisfies

$$dX_t^{(finan)} = \left(-\bar{c}_t + r_t X_t^{(finan)} + \Pi_B(t,Y_t^r)^\top(\mu_t^{(1)} - r_t 1_{d_r})\right) dt + \Pi_B(t,Y_t^r)^\top \sigma_t^{(1)} dW_t^r. \tag{EC.5.23}$$

Taking the difference of the two equations, we can obtain

$$d(X_t^{(finan)} - Z_{t,T}) = r_t(X_t^{(finan)} - Z_{t,T})dt.$$

By (43), we have $X_0^{(finan)} = Z_{0,T}$ at $t=0$, i.e., the initial value $X_0^{(finan)} - Z_{0,T} = 0$. Then, we can conclude $X_t^{(finan)} - Z_{t,T} \equiv 0$, i.e., $X_t^{(finan)} = Z_{t,T}$ holds for all t.

Given $X_t^{(finan)} = Z_{t,T}$ always holds, we have the wealth on CRRA sub-portfolio satisfies $X_t^{(crra)} = X_t - Z_{t,T} = \bar{X}_t$. Then the optimal portfolio (amount of wealth allocated on each asset) is given by

$$\pi_t^{(finan)} X_t^{(finan)} + \pi_C(t,Y_t) X_t^{(crra)} = \begin{pmatrix} \Pi_B(t,Y_t^r) \\ 0_{m-d_r} \end{pmatrix} + \pi_C(t,Y_t)\bar{X}_t.$$

This coincides with the HARA optimal policy in Theorem 2 and concludes our proof.

EC.5.5. Proof of Propositions 4

We first obtain the dynamics of the wealth process X_t of a HARA investor. By (3), X_t satisfies

$$dX_t = X_t\left[\pi_H(t,X_t,Y_t)^\top(\mu_t - r_t 1_m) + r_t\right] dt - c_t dt + X_t \pi_H(t,X_t,Y_t)^\top \sigma_t dW_t. \tag{EC.5.24}$$

In above, μ_t, r_t, and σ_t abbreviate for $\mu(t,Y_t)$, $r(t,Y_t^r)$, and $\sigma(t,Y_t)$, respectively; $\pi_H(t,X_t,Y_t)$ and c_t denote the optimal policy and consumption under the HARA utility. By (36) and (38), we have

$$X_t \pi_H(t,X_t,Y_t)^\top = \bar{X}_t \pi_C(t,Y_t)^\top + \left(\Pi_B(t,Y_t^r)^\top, 0_{m-d_r}^\top\right).$$

Using this as well as the expression (40) for consumption c_t, we can further express dX_t in (EC.5.24) as

$$\begin{aligned} dX_t = &\left\{X_t r_t + \bar{X}_t\left[\pi_C(t,Y_t)^\top(\mu_t - r_t 1_m) - 1/\phi_C(t,Y_t)\right] - \bar{c}_t + \Pi_B(t,Y_t^r)^\top(\mu_t^{(1)} - r_t 1_{d_r})\right\} dt \\ &+ \bar{X}_t \pi_C(t,Y_t)^\top \sigma_t dW_t + \Pi_B(t,Y_t^r)^\top \sigma_t^{(1)} dW_t^r, \end{aligned}$$

where $\pi_C(t,Y_t)$ and $\phi_C(t,Y_t)$ are the optimal policy and optimal wealth-consumption ratio under the CRRA utility; $\mu_t^{(1)}$ and $\sigma_t^{(1)}$ abbreviate for $\mu^{(1)}(t,Y_t^r)$ and $\sigma^{(1)}(t,Y_t^r)$ of the bond assets. Subtracting the dynamics of $Z_{t,T}$ in (EC.5.22) from above leads to

$$d\bar{X}_t = \bar{X}_t\left[\pi_C(t,Y_t)^\top(\mu_t - r_t) - 1/\phi_C(t,Y_t) + r_t\right]dt + \bar{X}_t\pi_C(t,Y_t)^\top\sigma_t dW_t.$$

That is,

$$d\bar{X}_t/\bar{X}_t = \left[\pi_C(t,Y_t)^\top(\mu_t - r_t) - 1/\phi_C(t,Y_t) + r_t\right]dt + \pi_C(t,Y_t)^\top\sigma_t dW_t.$$

Note that the right-hand side is exactly the wealth return rate for a CRRA investor, and is independent of wealth level X_t. Thus, the remaining wealth $\bar{X}_t$ of a HARA investor evolves exactly as that for a CRRA investor.

By Ito's formula, we can write $\bar{X}_t$ as

$$\bar{X}_t = \bar{X}_0 \exp\left(\int_0^t \mu_C(s,Y_s)ds + \int_0^t \pi_C(s,Y_s)\sigma_s dW_s\right), \tag{EC.5.25}$$

where

$$\mu_C(s,Y_s) := r_s + \pi_C(s,Y_s)(\mu_s - r_s) - 1/\phi_C(s,Y_s) - \frac{1}{2}\pi_C(s,Y_s)^\top\sigma_s\sigma_s^\top\pi_C(s,Y_s).$$

Now consider two HARA investors with different initial wealth $X_0^{(l)} < X_0^{(h)}$. By (EC.5.25), we can verify the ratio of their remaining wealth stays constant over time, i.e.,

$$\frac{\bar{X}_t^{(h)}}{\bar{X}_t^{(l)}} \equiv \frac{\bar{X}_0^{(h)}}{\bar{X}_0^{(l)}}. \tag{EC.5.26}$$

It proves (47).

Finally, we prove (48). Setting $t = T$ in (EC.5.26) leads to

$$\frac{\bar{X}_T^{(h)}}{\bar{X}_T^{(l)}} = \frac{\bar{X}_0^{(h)}}{\bar{X}_0^{(l)}} = \left(\frac{X_0^{(h)}}{X_0^{(l)}}\right)\left(\frac{1 - Z_{0,T}/X_0^{(h)}}{1 - Z_{0,T}/X_0^{(l)}}\right), \tag{EC.5.27}$$

where the second equality uses (21). Note that with $X_0^{(l)} < X_0^{(h)}$, the multiplier on the right-hand side is greater than one, i.e.,

$$1 - \frac{Z_{0,T}}{X_0^{(h)}} > 1 - \frac{Z_{0,T}}{X_0^{(l)}}. \tag{EC.5.28}$$

Taking logarithm on both sides of (EC.5.27), we get

$$\ln(\bar{X}_T^{(h)}) - \ln(\bar{X}_T^{(l)}) = \ln(X_0^{(h)}) - \ln(X_0^{(l)}) + \ln\left(1 - \frac{Z_{0,T}}{X_0^{(h)}}\right) - \ln\left(1 - \frac{Z_{0,T}}{X_0^{(l)}}\right).$$

Rearranging the terms and dividing both sides by T leads to (48). It concludes our proof.

EC.5.6. Proof of Proposition 5

We derive the optimal HARA policy under the CIRH-SVSIR model. In this model, the market for interest rate risk is complete. Following our set-ups in (28) – (31), we can specify the state variable for interest rate risk Y_t^r as the interest rate r_t itself, which is driven by W_{1t} by (51). The market has only one bond maturing at T_1. By (52) the volatility of the bond is given by $\sigma^{(1)}(t,r_t) = b(\tau_1)\sigma_r\sqrt{r_t}$. In addition, the market price of interest rate risk can be uniquely determined by the bond as

$$\theta^r(t,r_t) = \frac{b(\tau_1)\lambda_r\sigma_r^2 r_t}{b(\tau_1)\sigma_r\sqrt{r_t}} = \lambda_r\sigma_r\sqrt{r_t}. \tag{EC.5.29}$$

According to Cox et al. (1985), when the interest rate is driven by a single-factor CIR process and the market price of interest rate risk is given by (EC.5.29), the time–t price of a ZCB with unit face value maturing at time T can be determined as

$$B_{t,T} = \exp\left(a(\tau) + b(\tau)r_t\right). \tag{EC.5.30}$$

It follows the dynamics

$$dB_{t,T}/B_{t,T} = \left(r_t + b(\tau)\lambda_r\sigma_r^2 r_t\right)dt + b(\tau)\sigma_r\sqrt{r_t}dW_{1t}, \tag{EC.5.31}$$

with $\tau = T - t$. When $T = T_1$, $B_{t,T}$ coincides with the bond price P_t in the market.

We now solve the optimal policy for investors with HARA utility on terminal wealth. First, the corresponding optimal CRRA policy in (EC.1.14a) and (EC.1.14b) are derived in Liu (2007) using a separation theorem (see Proposition 3 therein). We next apply Theorem 2 to derive the optimal HARA policy, as the market for interest rate risk is complete in the CIRH-SVSIR model. Plugging (EC.5.30) into (37), the remaining wealth $\bar{X}_t$ is given by

$$\bar{X}_t = X_t - \bar{x}_T \exp\left(a(\tau) + b(\tau)r_t\right).$$

We then calculate the additional term $\Pi_B(t,r_t)$ in the HARA hedge component (38). By (EC.5.31), the volatility of $B_{t,T}$ is $\sigma_B(t,r_t;T) = b(\tau)\sigma_r\sqrt{r_t}$. Plugging this and $\sigma^{(1)}(t,r_t) = b(\tau_1)\sigma_r\sqrt{r_t}$ into (39), we have

$$\Pi_B(t,r_t) = \sigma^{(1)}(t,r_t)^{-1}\bar{x}_T B_{t,T}\sigma_B(t,r_t;T) = \frac{b(\tau)\bar{x}_T}{b(\tau_1)}\exp\left(a(\tau) + b(\tau)r_t\right).$$

Combining the above with the optimal CRRA policy, we obtain the closed-form solution for the optimal HARA policy as in Proposition 5.

EC.5.7. Proof of Proposition EC.1

We apply our general decomposition results in Theorem 2 to the case with nonrandom interest rate. First, when the interest rate is nonrandom, the time–t price for a ZCB maturing at time s with unit face value is directly given by $B_{t,s} = \exp\left(-\int_t^s r_v dv\right)$ according to non-arbitrage principle, as shown in (EC.1.6). Plugging the bond price to (37), the amount for financing portfolio follows by

$$Z_{t,T} = \bar{x}_T \exp\left(-\int_t^T r_v dv\right) + \int_t^T \bar{c}_s \exp\left(-\int_t^s r_v dv\right) ds, \text{ for } w \in (0,1). \tag{EC.5.32}$$

The remaining wealth is given by $\bar{X}_t = X_t - Z_{t,T}$.

Next, we prove the optimal HARA policy under nonrandom interest rate in (EC.1.7). Comparing (EC.1.7) with the general results in (36) and (38), we only need to prove the additional term $\Pi_B(t, Y_t^r) = 0$. Under nonrandom interest rate, the bond price $B_{t,s}$ follows

$$dB_{t,s} = -r_t B_{t,s} dt.$$

Thus, it has zero instantaneous volatility $\sigma_B(t, Y_t^r; s) = 0$. It implies that there is no uncertainty in the bond price when the interest rate is nonrandom. Then by (39), we have $\Pi_B(t, Y_t^r) = 0$. It proves the optimal HARA policy (EC.1.7) under the nonrandom interest rate case.

EC.5.8. Proof of Proposition EC.2

First, we prove the results for the optimal policy ratio in (EC.1.9). By the ratio relationship (EC.1.7) for the optimal HARA policy in Proposition EC.1, we can derive

$$\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}} = \frac{\bar{X}_t^{(h)}}{X_t^{(h)}} \frac{X_t^{(l)}}{\bar{X}_t^{(l)}} = \frac{1 - Z_{t,T}/X_t^{(h)}}{1 - Z_{t,T}/X_t^{(l)}}, \tag{EC.5.33}$$

where the first equality follows by canceling out the optimal CRRA policy in the nominator and denominator; the second equality follows by plugging in $\bar{X}_t^{(h)} = X_t^{(h)} - Z_{t,T}$ and $\bar{X}_t^{(l)} = X_t^{(l)} - Z_{t,T}$. We then show the optimal policy ratio is always greater than one. By Proposition 4, we have

$$\frac{\bar{X}_t^{(h)}}{\bar{X}_t^{(l)}} = \frac{X_t^{(h)} - Z_{t,T}}{X_t^{(l)} - Z_{t,T}} \equiv \frac{X_0^{(h)} - Z_{0,T}}{X_0^{(l)} - Z_{0,T}} > 1.$$

Thus, we have $X_t^{(h)}/X_t^{(l)} > 1$, for all $t \in [0,T]$. It leads to $1_m^\top \pi_t^{(h)}/(1_m^\top \pi_t^{(l)}) > 1$ by (EC.5.33). For the optimal portfolio ratio, we multiply both sides of (EC.5.33) by $X_t^{(h)}/X_t^{(l)}$ to get

$$\frac{1_m^\top \pi_t^{(h)} X_t^{(h)}}{1_m^\top \pi_t^{(l)} X_t^{(l)}} = \frac{\bar{X}_t^{(h)}}{\bar{X}_t^{(l)}} \equiv \frac{X_0^{(h)} - Z_{0,T}}{X_0^{(l)} - Z_{0,T}},$$

where the second equality follows from Proposition 4. It proves the relation in (EC.1.10).

Finally, we prove relationship (EC.1.11) under constant interest rate $r(t, Y_t) = r$ and HARA utility on terminal wealth only ($w = 0$ in (4)). Under constant interest rate, the term $Z_{t,T}$ in (EC.5.32) can be simplified as $Z_{t,T} = \bar{x}_T \exp(-r(T-t))$. Plugging this into (EC.5.33) and calculating the partial derivatives, we get

$$\frac{\partial}{\partial r}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) = \frac{X_t^{(l)} \bar{x}_T (T-t)\exp(-r(T-t))}{X_t^{(h)}\left(X_t^{(l)} - \bar{x}_T \exp(-r(T-t))\right)^2} \times \left(X_t^{(l)} - X_t^{(h)}\right) < 0$$

and

$$\frac{\partial}{\partial T}\left(\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}}\right) = \frac{X_t^{(l)} \bar{x}_T r \exp(-r(T-t))}{X_t^{(h)}\left(X_t^{(l)} - \bar{x}_T \exp(-r(T-t))\right)^2} \times \left(X_t^{(l)} - X_t^{(h)}\right) < 0.$$

The inequalities follow as $X_t^{(h)}/X_t^{(l)} > 1$ for all $t \in [0, T]$. It proves (EC.1.11).

Note that by (EC.1.8), the vectors $\pi_t^{(h)}$ and $\pi_t^{(l)}$ are parallel to each other, as both of them are parallel to the common CRRA optimal policy. Then, we have

$$\frac{1_m^\top \pi_t^{(h)}}{1_m^\top \pi_t^{(l)}} = \frac{\pi_{t,i}^{(h)}}{\pi_{t,i}^{(l)}},$$

for $i = 1, 2, \ldots, m$ denoting each asset. Thus, the above proof also applies to the optimal policy and portfolio ratio of each individual asset.

References

Aït-Sahalia, Y., Kimmel, R., 2007. Maximum likelihood estimation of stochastic volatility models. Journal of Financial Economics 83 (2), 413–452.

Ait-Sahalia, Y., Kimmel, R. L., 2010. Estimating affine multifactor term structure models using closed-form likelihood expansions. Journal of Financial Economics 98 (1), 113–144.

Cox, J. C., Huang, C.-f., 1989. Optimal consumption and portfolio policies when asset prices follow a diffusion process. Journal of Economic Theory 49 (1), 33–83.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985. A theory of the term structure of interest rates. Econometrica 53 (2), 385–407.

Detemple, J., Garcia, R., Rindisbacher, M., 2003. A Monte Carlo method for optimal portfolios. Journal of Finance 58 (1), 401–446.

He, H., Pearson, N. D., 1991. Consumption and portfolio policies with incomplete markets and short-sale constraints: The infinite dimensional case. Journal of Economic Theory 54, 259–304.

Karatzas, I., Lehoczky, J. P., Shreve, S. E., 1987. Optimal portfolio and consumption decisions for a small investor on a finite horizon. SIAM Journal on Control and Optimization 25, 1557–1586.

Karatzas, I., Lehoczky, J. P., Shreve, S. E., Xu, G.-L., 1991. Martingale and duality methods for utility maximization in an incomplete market. SIAM Journal on Control and Optimization 29 (3), 702–730.

Karatzas, I., Shreve, S. E., 1991. Brownian Motion and Stochastic Calculus, 2nd Edition. Vol. 113 of Graduate Texts in Mathematics. Springer-Verlag, New York.

Liu, J., 2007. Portfolio selection in stochastic environments. Review of Financial Studies 20 (1), 1–39.

Nualart, D., 2009. Malliavin calculus and its applications. No. 110. American Mathematical Soc.

Nunno, G. D., Øksendal, B., Proske, F., 2008. Malliavin calculus for Lévy processes with applications to finance. Springer.