

Intertemporal Asset Allocation Strategies under Inflationary Risk

Dissertation

Zur Erlangung des Grades
eines Doktors der Wirtschaftswissenschaften
der Fakultät für Wirtschaftswissenschaft
Universität Bielefeld

Dipl.-Math. Chih-Ying Hsiao

September 2006

1. Supervisor: *Prof. Willi Semmler*,
University of Bielefeld, Germany
2. Supervisor: *Prof. Carl Chiarella*,
University of Technology, Sydney, Australia

Contents

1	Introduction	13
2	An Initial Intertemporal Example	23
3	Intertemporal Asset Allocation under Inflation	31
3.1	The Model	31
3.2	Solution via Dynamic Programming	36
3.3	Representation of the Solution via the Feynman-Kac Formula	41
3.4	The Martingale Method: An Alternative Solution Strategy .	42
3.5	Summary and Remarks	45
4	Empirical Financial Markets	47
4.1	The General Duffie-Kan Affine Term Structure Model	49
4.2	Data	50
4.3	Model I; Unspecified Factors	52
4.3.1	The Model	53
4.3.2	Model Identification	53
4.3.3	Estimation of Underlying Factors	55
4.4	Model II; Inflation-Indexed Bonds	61
4.4.1	The Model	62
4.4.2	Model Estimation	66
4.4.3	A Modification of the Model	76
4.4.4	Estimation of the Modified Model	79
4.4.5	Estimation of Realized Inflation Dynamics	84
4.4.6	Estimation of Stock Return Dynamics	86
4.5	Summary and Comparison of Models I and II	88
5	Optimal Investment Strategies	91
5.1	How large should the Relative Risk Aversion be?	91
5.2	Optimal Portfolio Allocation for Model I	94

5.2.1	Explicit Forms for Investment Strategies	94
5.2.2	Examples: Intertemporal and Information Effects	98
5.2.3	Examples: Risk Aversion and Horizon Effects	104
5.3	Optimal Portfolio Allocation for Model II	106
5.3.1	Analytical Solution	107
5.3.2	Example: The Hedging Effect of Inflation-Index Bonds	112
5.4	Summary	118
6	Portfolio Strategies under Short-Sale Constraints	119
6.1	The Backward Markov Chain Approximation Method	120
6.2	Optimal Portfolio Application with Short-Sale Constraints	121
6.3	Summary	126
7	Conclusions	127
8	Appendix	131
8.1	The Discrete-Time Counterpart	131
8.1.1	The Discrete-Time Model	131
8.1.2	The Backward Iteration Formula	133
8.1.3	The Jacobi Iteration for the Infinite Time Horizon Problem	135
8.1.4	Remarks on the Convergence of the Algorithms	137
8.1.5	Self-Financing and Nominal Wealth Dynamics	137
8.1.6	Sequential Optimization	139
8.2	Some Basic Results	143
8.3	Proofs	146

List of Figures

4.1	US Nominal Bond Yields and Federal Funds Rate (FFR)	51
4.2	US Real Bond Yields and Federal Funds Rate (FFR)	51
4.3	Estimated Factors in the Three-Factor Model	58
4.4	Estimated Factors, all Three Models	59
4.5	Time Series of Real Yields and the Estimated Real Rate	67
4.6	Market and Estimated 1Y bond yield	74
4.7	Market and Estimated 10Y bond yield	74
4.8	Nominal Yields and Estimated Factors	75
4.9	Federal Fund Rate and the Estimated Rate	76
4.10	Comparing Estimated Real Rates	81
4.11	Estimated π_t in both models	83
4.12	Realized and Filtered Annualized Inflation	85
4.13	SP500 Index	87
4.14	Daily Excess Returns(S&P500)	88
4.15	S&P500 Excess Daily Returns Distribution	89
4.16	Comparison of the Estimated Factors in the Model I and Model II	89
5.1	Example: Final Wealth Distribution, $\gamma = 7$, Model I	101
5.2	Example: Final Wealth Distribution, $\gamma = 15$, Model I	102
5.3	Example: Final Wealth Distribution, $\gamma = 30$, Model I	103
5.4	Optimal Portfolio Weights for Model I, against γ	105
5.5	Optimal Investment Weights for Model I, Horizon Effect	106
5.6	Optimal Portfolio Weights for Model II with IIB	114
5.7	Optimal Portfolio Weights for Model II without IIB	115
5.8	Optimal Weights for Model II, Horizon Effect, with IIB	117
5.9	Optimal Weights for Model II, Horizon Effect, without IIB	117

List of Tables

4.1	Statistics for US Nominal/Real Bond Yields	52
4.2	Estimated Parameters: One-, Two-/Three-factor Models . . .	57
4.3	Comparison Estimation Behaviors	60
4.4	Upper Panel: estimated parameters for the real yield formula and Lower Panel: statistics, fitting errors, and price sensitivities	68
4.5	Ordinary Linear Regression on filtered r_t	69
4.6	Upper Panel:estimated parameters for nominal term struc- ture; Lower Panel: statistics, fitting errors, and yield sensitivity	73
4.7	OLS results for the filtered π_t	75
4.8	Upper Panel: estimated parameters for the real yield for- mula/ Lower Panel: real yield statistics and price Sensitivities	80
4.9	Comparing information criteria for the real yield models . . .	80
4.10	Upper Panel:estimated parameters for the modified nominal term structure; Lower Panel: fitting errors and yield sensitivity	82
4.11	Comparing Information Criteria	84
4.12	Estimation Results for the CPIU	85
5.1	Asset Returns and Risk Compensation for Model I	92
5.2	Asset Returns and Risk Compensation for Model II with the parameter separations	93
5.3	Time schedule of the bond maturities	98
5.4	Example, Statistics of Wealth Distributions, Model I	100
5.5	Optimal Weights against Risk Aversion	104
5.6	Limits of the Investment Weights when $\tau \rightarrow \infty$	106
5.7	Parameter summary for Model II	113
5.8	Decomposition of Portfolio for Model II with IIB	114
5.9	Portfolio Decomposition without IIB	115

6.1	Parameters Used for Simulation Examples with Short-Sale Constraints	122
6.2	Average Errors of the Value Function with Parameters given in Table 6.1	123
6.3	Average Relative Errors of the Portfolio Decisions with Parameters given in Table 6.1	123
6.4	Effect of Short-Sale Constraints on Portfolio Allocations . . .	124
6.5	Effect of Commission Fees on Portfolio Decisions	125
6.6	Risk Aversion and Short-Sale Exclusion (SSE).	125

List of Notation

Abbreviations

IIB: Inflation-Index Bond

i.i.d: independent and identically distributed

ODE: Ordinary Differential Equation

PDE: Partial Differential Equation

SDE: Stochastic Differential Equation

Notation (pages)

$A(\tau)$: the Duffie-Kan coefficient (level term) (p. 49)

$A_n(\tau)$: the Duffie-Kan coefficient for the nominal bond model (p. 62)

$A_r(\tau)$: the Duffie-Kan coefficient for the real bond mode (p. 63)

$B(\tau)$: the Duffie-Kan coefficient (linear term) (p. 49)

$B_{nr}(\tau), B_{rr}(\tau)$: the Duffie-Kan coefficient for the nominal bond model (p. 62)

$B_{rr}(\tau)$: the Duffie-Kan coefficient for the real bond mode (p. 63)

C_t : nominal consumption at t (p. 25)

c_t : real consumption at t (p.35)

$F(X_t)$: the drift coefficient for X_t (p. 32)

\mathcal{F}_t : the augmented natural filtration at t (p. 31)

$G(X_t)$: the diffusion coefficient for X_t (p. 32)

$g_r (g_\pi)$: the volatility parameter for $r_t (\pi_t)$ (p. 62)

I_t : the price index at t (p31)

$J_0(V_0, R_0)$: the value function in the initial example (p. 25)

$J_1(V_1, R_1)$: the value function in the initial example (p. 25)

$J(t, T, v_t, X_t)$: the value function for the intertemporal asset allocation problem (p. 27)

- $J_v = \partial J / \partial v$ (p. 37)
 $J_X = \partial J / \partial X$ (p. 37)
 m : the number of the assets (p. 33)
 $\min(a, b)$ gives the smaller value among $\{a, b\}$ (p. 36)
 \max : maximization operator (p. 37)
 n : the number of the factors (p. 32)
 P_{it} : price of the i -th asset at t (p. 33)
 $P_n(t, T, X_t)$: the nominal bond price at t (p. 49)
 $P_I(t, T)$: the price of IIB (p. 63)
 $P_r(t, T)$: the real bond price (p. 63)
 $P^S(t)$: the stock price (p. 87)
 \mathcal{P} : the real world measure (p. 23)
 R_t : the instantaneous nominal interest rate at t (pp. 24, 49)
 r_t : the instantaneous real interest rate at t (p. 62)
 \bar{r} : the mean for r_t (p. 62)
 \mathcal{R}_{AA} : the correlation matrix of W_t (p. 32)
 \mathcal{R}_{AI} : the correlation matrix between W_t and W_t^I (p. 33)
 \mathcal{R}_{IX} : the correlation matrix between W_t^I and W_t^X (p. 32)
 \mathcal{R}_{XA} : the correlation matrix between W_t^X and W_t (p. 33)
 \mathcal{R}_{XX} : the correlation matrix for W_t^X (p. 32)
 t : time (p. 23)
 $U(\cdot)$: the utility function (p. 34)
 V_t : nominal wealth at t (p. 25)
 v_t : real wealth at t (p. 35) $X_t (= (X_{1t}, \dots, X_{nt})^\top)$: the process of the factors (p. 32)
 $Y_n(t, T, X_t)$: the nominal bond yield (p. 49)
 $Y_r(t, T)$: the real bond yield (p. 64)
 W_t^I : the risk source of I_t (p. 31)
 W_t^O : the complement set of W_t^X in W_t
 W_t^r : the risk source for r_t (p. 62)
 W_t^π : the risk source for π_t (p. 62)
 W_t^S : the risk source for $dP^S(t)/P^S(t)$ (p. 87)
 $W_t^X = (W_{1t}^X, \dots, W_{nt}^X)^\top$: the process of risk sources for X_t (p. 32)
 $W_t = (W_{1t}, \dots, W_{mt})^\top$: the process of risk sources for the assets (p. 32)
 $\alpha_t = (\alpha_{1t}, \dots, \alpha_{mt})^\top$: the vector of investment weights/portfolio proportions at t (p. 35)
 Γ : the volatility matrix for X_t (p. 49)
 Γ_i : the i -row of Γ (p. 55)
 Γ_{ij} : the (i, j) -element in Γ (p. 56)

- γ : risk aversion parameter/absolute risk aversion (p. 23)
 δ : the subjective discount rate (p. 16)
 ϵ_1 : indicator for the consideration of intermediate consumption (p. 35)
 ϵ_t^e : measurement errors (p. 56)
 η : short-sale commission (p. 36)
 θ : the reverting mean for X_t (p. 49)
 $\lambda_i(X_t)$: the market price of risk for W_{it} at t (p. 33)
 $\lambda(X_t) = (\lambda_1(X_t), \dots, \lambda_m(X_t))^\top$: the vector of the market price of risk (p. 33)
 $\lambda_t = \lambda(X_t)$ (p. 34)
 λ_x ; $x = \{r, \pi, I, S\}$: the market price of risk of W_t^x (p. 65)
 λ : the vector of constant market prices of risk (p.50)
 \mathcal{K} : the mean-reversion matrix for X_t (p. 49)
 κ_i : the mean-reversion parameter for each component of X_t (p. 55)
 κ_r (κ_π): the mean-reversion parameter for r_t (π_t) (p. 62)
 $\mu_i(X_t, t)$: the drift coefficient for the return of the i -th asset (p. 33)
 $\mu(X_t, t) = (\mu_1(X_t, t), \dots, \mu_m(X_t, t))^\top$: the vector of the drift coefficients (p. 33)
 $\mu_t = \mu(X_t, t)$ (p. 34)
 π_t : the expected instantaneous inflation rate at t (p. 31, 62)
 $\bar{\pi}$: the mean for π_t (p. 62)
 ρ_{xy} ; $x, y = \{r, \pi, I, S\}$: the correlation between W_t^i and W_t^j
 $\Sigma_i(X_t, t)$: the diffusion coefficient of the return of the i -th asset (p. 33)
 $\Sigma(X_t, t) = (\Sigma_1(X_t, t), \dots, \Sigma_m(X_t, t))^\top$: the matrix of the diffusion coefficients (p. 34)
 $\Sigma_t = \Sigma(X_t, t)$ (p. 34)
 σ_ϵ : the volatility parameter for ϵ_t^e (p. 56)
 σ_I : the volatility parameter for dI_t/I_t (p. 31)
 σ_S : the volatility parameter for $dP^S(t)/P^S(t)$ (p. 87)
 τ : time to maturity
 $\Phi(t, T, X_t)$: the part of the value function relating to the factors (p. 38)
 $\Phi_X := \partial\Phi/\partial X$ (p. 39)
 $\xi_0 := A'(0)$ (p. 49)
 $\xi_1 := B'(0)$ (p. 49)
 ψ_t : consumption ratio at t (pp. 26, 35)

$$\underline{1} = (1, \dots, 1)^\top \text{ (p. 34)}$$

\top : matrix transpose (p. 32)

Acknowledgments

Here I would like to thank those people who have also contributed to the development of this thesis.

I would like to thank my first supervisor Professor Willi Semmler, who suggested the topic of this thesis, the intertemporal asset allocation problem, and advised the consideration of macroeconomic impacts on this problem, which has formed the guiding line of this thesis. Professor Semmler's comprehensive and balanced treatment of theoretical and empirical macroeconomics has had a remarkable influence on the structure of this thesis.

I also want to express my deep gratitude to my second supervisor Professor Carl Chiarella in the School of Finance and Economics, at the University of Technology, Sydney (UTS). Many crucial points in this thesis, such as, the solution approach under inflation risk in Chapter 3, the no-arbitrage principle under inflation risk in Chapter 4, the second term structure models for modelling inflation-indexed bonds in Chapter 4 and the consideration of short-sale constraints for the intertemporal asset allocation problems in Chapter 6 were developed based upon ideas obtained in discussions with him.

Some parts in the thesis have already appeared as working papers and have been submitted to journals for publication. I would also like to thank the critiques and comments from anonymous referees of *Computational Economics* on papers arising out of Section 4.3 and Chapter 6, which helped a lot to improve this thesis. Further thanks must be given to my colleagues at Bielefeld University and those at UTS, especially Dr. Wolfgang Lemke for many important comments on Section 4.3.3, Dr. Hing Hung for the discussion on technical details of the Kalman filter technique in Section 4.4.2, Evan Shell-shear for the English wording, also Dr. Chen Pu, Dr. Christoph Woester, Marten Hillebrand, Dennis Krichhoff, Dr. Wenlang Zhang, and participants at the "10th International Conference on Computing in Economics and Finance" in Amsterdam and in the "13th Symposium of Society of Non-linear Dynamics and Econometrics" at London for many useful discussions.

Chapter 1

Introduction

The object of this dissertation is the study of how one constructs optimal intertemporal asset allocation strategies. The construction of optimal asset allocation decisions is an important economic activity. From a micro-economic point of view, individuals need to consider how much they can consume now and how much they should save for future consumption. To achieve their goals they will often have to consider the possibility of investing their savings in financial markets. To achieve this they will have to know how to organize their investment portfolios. The portfolio decision problem is also important from a macro-economic point of view, as it affects the financial structure of the proportions of bonds and stocks observed in the economy.

For modern day investors there exists a multitude of assets in financial markets such as stocks, bonds, and numerous financial derivatives that can be utilized to form investment portfolios. In addition to the complexity and diversity of financial markets, the difficulty presented by the range of time periods inherent in many financial products makes the task of constructing intertemporal asset allocation strategies even more difficult. Over long time scales, the investment environment will vary considerably. Different interest rates for short-term and long-term investments come into play and these are always in a state of flux. To exacerbate the situation for long-term investments, purchasing power is no longer constant, so inflation risk should be taken into account. So the question is, how can the agents find their way through the maze of financial market considerations to rationally decide their long-term asset allocation? Recently, there has been a resurgence of research interest in this subject. The main thrust of this research is to solve the intertemporal asset allocation problem for practical applications

as well as to extend the theoretical analysis of such problems.

As a piece of research on the intertemporal asset allocation problem, this dissertation contains a normative and a positive part. The normative part studies *how* to construct optimal asset allocation strategies. The first part of the dissertation is devoted to studying solution techniques for the stochastic optimal control problem which is applied to the intertemporal asset allocation problem. The solution techniques include analytical solutions and computational solutions in the case where analytical solutions are not available. The second part – the positive part – models the real financial markets in which investors may wish to invest. Knowledge of the financial markets is indispensable if one wishes to offer useful financial advice. The second part of this dissertation provides an empirical study of financial markets, especially from the viewpoint of constructing asset allocation strategies. At the conclusion of these two parts, we will be able to provide concrete asset allocation suggestions based on current market situations.

The Beginning

The study of the portfolio selection problem of a rational investor facing *uncertainty* began with the works of Markowitz (1959) and Tobin (1958). It was further developed by Sharpe (1964) and Lintner (1965). These contributions have been established as the *Capital Asset Pricing Model* (CAPM) which is widely used in modern finance research as well as in practice. Albeit that there is a wide acceptance of the CAPM, it is still not general enough to deal with real asset allocation problems because the CAPM can only solve a *one-period* optimization problem. While in reality, asset allocation problems need to be solved over many time periods.

The Intertemporal Model

Samuelson (1969) extended the one-period model to a many-period model and also included consumption decisions. However, he found the multi-period portfolio decision was exactly the same as the one-period period decision because he assumed the asset returns are independent and identically distributed (i.i.d.). The first intertemporal model to include the key features that yield a multi-period decision¹ that is different from a one-period decision² was given by Merton (1973). He showed, when there are some

¹The terms an intertemporal asset allocation decision, a strategic asset allocation decision, or a dynamic asset allocation decision are also used.

²This is also called a myopic decision or a tactical asset allocation decision.

underlying *time-varying* “factors” affecting asset returns so that the asset returns are no longer independent and identically distributed (i.i.d.), there is an additional *intertemporal hedging term* which does not appear in the myopic (one-period) decision. The appearance of the intertemporal hedging terms arises because an intertemporal strategy considers not only the trade-off between return and risk but also the possible future development of asset returns. In order to obtain the optimal intertemporal strategies one needs to forecast future development of asset returns because agents have the possibility of hedging against undesired future developments of asset returns.

The intertemporal term can solve the *asset allocation puzzle* raised by Canner, Mankiw and Weil (1997): the optimal investment strategy should always have a *constant* bond to stock investment ratio according to the theoretical result based on the one-period asset model³, while in practice financial advisors usually suggest to more conservative investors to hold more bonds relative to stocks.

The underlying time-varying factors in Merton’s intertemporal model are, formally, exogenous continuous-time Markovian stochastic processes. In practical asset allocation applications one question arises quite naturally: which are those factors that can change the investment environment and affect the mechanism of the asset returns? In the literature we can find a number of factor specifications. In Cox, Ingersoll and Ross (CIR)(1985a), the factor was a technological index. Some other researchers take the view that excess stock returns are *predictable* and regard the predictable excess return as a factor. Authors adopting this approach including Kim and Omberg (1996), Wachter (2002), Campbell, Chacko, Rodriguez and Viceira (2004), as well as Munk, Sørensen and Vinther (2004). Another important time-varying factor is the *stochastic interest rate* which is taken into account in Cox, Ingersoll and Ross (1985b), Brennan, Schwartz and Lagnado (1997), Brennan and Xia (2002), and Munk, Sørensen and Vinther (2004). Other approaches include the model of Brennan, Schwartz and Lagnado (1997)

³We remark that this ratio is independent of agents’ risk aversion. In the standard CAPM framework, see for example, Chapter 5 in Campbell, Lo and MacKinlay (1997), agents with different risk preference will have different investment proportions in the risk-less asset and the market portfolio which consists of the risky assets on the market, including both bonds and stocks, and whose investment structure is independent of agents’ risk preference. So, a more risk averse agent will invest less in the market portfolio, that is, invest less both in bonds and stock while the bond to stock investment ratio remains constant.

where a predictable dividend stream was considered as a factor. In the sophisticated model of Brennan, Wang and Xia (2004) the maximal Sharpe Ratio was considered as a factor. Due to increasing transactions of financial derivatives, Liu and Pan (2003) include stochastic volatility as a factor. All the aforementioned papers adopt the continuous-time framework of Merton where the agents are allowed to decide and revise their decisions in every moment. Parallel to the continuous-time framework, there are also discrete-time models. For example, Campbell and Viceira (1999) take mean-reverting excess stock returns into account and in a second paper Campbell and Viceira (2001) discuss the impact of a stochastic interest rate and inflationary expectations on asset allocation.

For long-term investments, the consumption price index varies over the investment horizon, so purchasing power cannot remain constant. With this feature of long-term investments, one focus of this dissertation is on a consideration of inflation risk in the intertemporal asset allocation model, that is, we consider inflationary expectations as a factor. Furthermore, related to inflation risk, this dissertation extends Merton's intertemporal asset allocation model to accommodate a time-varying consumption price index. With this extension we have two different terms for different economic activities: agents' consumption is counted in real terms while their financial investments are arranged in nominal terms.

Solution Techniques

To solve the intertemporal asset allocation problem, the method of *dynamic programming* was used by Merton (1971). In his continuous-time framework, it turned out that one needs to solve the *Hamilton-Bellman-Jacobi* (HJB) equation which is a nonlinear second order partial differential equation. The reader will note that there is a gap between the time of Merton's initiation of the intertemporal asset allocation problem in 1971 and the recent resurgence of research activity in this subject from the second half of the 1990s. The reason for this delay is the difficulty in obtaining analytical solutions. Kim and Omberg (1996) provided an analytical solution in their model through the *verification theorem*, which refers to a procedure that consists of trying a possible solution candidate for the HJB equation and then verifying it. Through the verification theorem, Liu (2001, 2005) was able to solve dynamic programming problems analytically for a wide class of models by suggesting a fairly general solution structure.

An alternative method that may be used to solve the intertemporal as-

set allocation problem is the *martingale method* of Cox and Huang (1989), Karatzas, Lehoczky and Shreve (1987), and Pliska (1986). The application of this method can be found, for example, in Wachter (2002). The martingale method for the optimal strategies has a very close relation to martingale pricing in financial market theory, which was proved by Harrison and Kreps (1979), and Harrison and Pliska (1981) to be equivalent to the no-arbitrage principle in a perfect financial market. The no-arbitrage principle, suggested by Black and Scholes (1973), is now established as a fundamental tenet in financial market theory. The discrete-time counterpart of martingale pricing is the pricing scheme of *stochastic discount factors*.

How can one solve the intertemporal asset allocation problem in the extended framework with a stochastic consumption pricing index? Brennan and Xia (2002) employed the scheme based on the real pricing kernel. In the discrete-time model of Campbell and Viceira (2001) the stochastic discount factor approach was employed to guarantee the no-arbitrage condition. Munk, et al (2004) provided the solution of the intertemporal problem under inflation risk by using dynamic programming. This dissertation will study both of the main solution methods: the method of dynamic programming and the martingale method for the intertemporal asset allocation problem and will explore their application to the extended model with a stochastic consumption price index.

The Term Structure of Interest Rates

For long-term investments, bonds are regarded as “safe” assets and recommended to conservative investors. For a bond portfolio, agents will invest in various bonds with different maturities, so different interest rates (short-term and long-term interest rates) will come into play and the *term structure of interest rates* will provide essential information for constructing a bond portfolio. Apart from the applications in constructing bond portfolios, the term structure of interest rates incorporates important information relevant for an intertemporal decision. Long-term interest rates give a clue about market expectations and the future development of short-term interest rates. Given this, a rational agent must strive to acquire information about the term structure of interest rates for her/his long-term investment strategies. With the aforementioned motivation, one of the focuses of this dissertation is the study of the term structure of interest rates.

Two main approaches have been developed in the theory of the term structure of interest rates. The classical approach started with Fisher’s (1896)

unbiased expectations hypotheses. After that two new hypothesis followed: Hicks' (1939) liquidity preference hypothesis and the preferred habitat hypothesis of Modigliani and Sutch (1966). The second, modern, approach is based on the no-arbitrage principle of Black and Scholes (1973). Since the development of the no-arbitrage principle, there has been a rapid development in the modelling of financial derivatives. With regard to the term structure model, the no-arbitrage principle provides a unified perspective in which bonds may be considered as financial derivatives. Hand in hand with the rapid development of financial derivatives based on the no-arbitrage principle, the modern arbitrage approach to the study of the term structure still remains an active area of research.

In the early stages of the modern approach to the study of the term structure, a lot of effort was devoted to finding out the factors which determine the term structure. For example, in Vasicek (1977) and Cox, Ingersoll and Ross (1985b), where the only factor considered was the stochastic instantaneous interest rate⁴. Richard (1978) considered a two-factor model, where the two factors were the real instantaneous interest rate and the anticipated inflation rate. The two factors in Brennan and Schwartz (1979) were the instantaneous interest rate and the yield on a long term consol bond. Longstaff and Schwartz (1992) provided a general equilibrium approach and obtained the instantaneous interest rate and its variance as two common factors for pricing bonds. Hull and White's two-factor model (1994) took into account the instantaneous interest rate and a stochastic mean-reverting tendency. All the above term structure models can be resumed within the *affine yield-factor* family of Duffie and Kan (1996). The bond yields of all these models can be expressed as an affine combination of the underlying factors. Due to this finding there is an invertible relation between the bond yields and the underlying factors when the number of the yields is the same as that of the factors. Because of this invertible relation, the factors can be replaced by the bond yields. This explains the name of the Duffie-Kan model: the *yield-factor* model because the bond yields can serve as factors for bond pricing. The insight of Duffie and Kan opened up a new possibility in the study of the term structure of interest rates: we do not need to specify which factors affect the term structure. Data on bond yields already contain the information about the term structure.

It is necessary to note here that there is still another way to unify the mod-

⁴In CIR (1985b) the instantaneous interest rate is an equilibrium interest rate, which is derived from a stochastic technological index.

els of the term structure of interest rates other than the Duffie-Kan model. This is the approach proposed by Heath, Jarrow and Morton (HJM) (1992) based on forward rate information. The HJM model and its extensions by Brace, Gatarek and Musiela (1997) find wide applications in pricing interest rate derivatives in practice, see for example Brigo and Mercurio (2001). This dissertation adopts the Duffie-Kan approach instead of the latter one for the following two reasons. First, for the main task of this dissertation, the construction of intertemporal asset allocation strategies, the stochastic interest rate and inflation risk are two important factors. By employing the Duffie-Kan model, we can model them directly. Second, the Duffie-Kan approach has the advantage of being analytically more tractable.

After the breakthrough in term structure modelling proposed by Duffie and Kan and their insight that the bond yields can replace the underlying determining factors, the focus of term structure research has moved to analyze bond yield data empirically in order to find out the common determining factors. It can be observed very easily in reality that interest rates or forward rates of different maturities are highly correlated. Therefore, they should be affected by common factors. Using the principal component method, Rebonato (1998) finds that the first two principal components can already describe 99% of the entire variability of forward rates in UK. Similar results have been found in most major economies.

This dissertation employs the *Kalman filter* method⁵ to filter out the common factors “hidden” behind the bond yield data, an approach initiated by Jegadeesh and Pennacchi (1996), Babbs and Nowman (1999), and De Jong (2000) and many others. We use the Duffie and Kan framework to implement the Kalman filter method. Comparing it with the principal component method for detecting the common factors, the Duffie and Kan framework has an advantage in that it expresses the no-arbitrage principle in terms of the maturity-dependent relation between different bond yields, whereas the principal component method is unable to be easily matched with the no-arbitrage principle. The Kalman filter method provides a filtered likelihood function based on the Duffie-Kan model which facilitates further the implementation of maximum likelihood estimation. When undertaking parameter estimation we still need to be aware of the identification problem as pointed out in Dai and Singleton (2000).

⁵See Harvey (1990).

Inflation-Indexed Bonds

Inflation risk can be hedged by *inflation-indexed bonds*, whose principal and coupon payments are adjusted with respect to some price index. Through this they provide certain purchasing power and can hedge the inflation risk of long term investment plans. The US Treasury has been issuing Treasury Indexed-Protected Securities (TIPS) since January 1997, these are securities whose payments are adjusted to the Consumption Price Index. The outstanding amount of IIBs in 2004 was about \$200bn in the US and \$500bn worldwide.⁶ Jarrow and Yildirim (2003) model inflation-indexed bonds based on the HJM approach.

Computational Solution Methods

Let us come back to the issue of solving the intertemporal asset allocation problem. When the underlying asset pricing models must fit observed empirical data as described above, then the solution of the optimal asset allocation strategy is often difficult to obtain. This is also the case when certain market imperfections are considered, such as short-sale constraints or transaction costs. In such cases solving the intertemporal asset allocation problem requires the utilization of computational methods. Tapiero and Sulem (1994) summarize such computational methods into four categories: (i) A direct solution of the Hamiltonian-Jacobi-Bellman (HJB) equation; (ii) The *Markov chain approximation method* of Kushner (1977) which approximates the original controlled process by finite-state processes; (iii) Methods for such examples with well-known solutions, for example, the linear quadratic problem; (iv) Methods using simulation-based techniques, such as Monte-Carlo simulations. In the application of numerical methods to the asset allocation problem, Brennan et al. (1997) employ a solution approach of the first category using a finite difference approximation. Kushner and Dupuis (2000) give the convergence conditions for the Markov chain approximation method of the second category in a quite general setting that also allows for jump processes and stochastic stopping rules. Methods of this category are quite widely implemented because they apply two classical iteration methods: the policy space iteration (the Howard improvement algorithm) and the state space iteration (the Jacobi iteration). These methods are usually employed for the intertemporal asset allocation problem with *infinite* time

⁶Liquidity in the TIPS market is improving, with the daily trading volume having doubled during 2002-2004 and amounting to about \$5bn in 2004. For details see: <http://www.treas.gov/offices/domestic-finance/key-initiatives/tips.shtml>.

horizon. Camilli and Falcone (1995) provide a useful truncation technique for the Jacobi iteration and estimate its errors. Santos and Vigo-Aguiar (1998) consider the errors of both classical methods for a stochastic growth model. Munk (2003) employs the policy space iteration for the asset allocation problem and compares different Markov chain approximation schemes. For discrete-time models, Campbell, Chan and Viceira (2003) develop an approximation method (with Epstein-Zin utility). Grüne and Semmler (2003) apply the Jacobi iteration to an asset pricing model using the technique of adaptive grids, which is proposed by Grüne (1997).

The Contributions of this Dissertation

In the framework of the intertemporal asset allocation problem, this dissertation focuses on the modelling of inflation risk and its impact on intertemporal asset allocation strategies. The contributions of this dissertation to the current literature are as follows. First, in considering inflation risk, Merton's continuous-time framework of the intertemporal model is extended to accommodate a time-varying consumption price index. Recall that it is important to consider a time-varying consumption price index for the intertemporal asset allocation problem because: (i) a time-varying price level will affect consumption decisions since agents care about how many goods they can consume (real terms) instead of how much money they spend for consumption. (ii) A time-varying price level will affect the portfolio strategies because the implied inflation risk from the time-varying price index affects interest rates (bond yields) of different maturities. Second, this dissertation extends the solution method of the intertemporal asset allocation problem, that is, the method of dynamic programming, to a framework with a stochastic consumption price level. An analytical solution formula is provided for the optimal intertemporal investment strategy in this extended framework by using the Feymann-Kac formula. With regard to inflation modelling, the third contribution of this work is to develop a new interest rate model to include inflation-indexed bonds. Based on this new model, we can study the hedging performance of inflation-indexed bonds against the inflation risk within the intertemporal asset allocation problem. Fourthly, due to the difficulty in solving the intertemporal asset allocation problem for some extended cases, this dissertation develops a computational algorithm based on the Jacobi iteration method in the Markov Chain Approximation family of Kushner (1977) . This algorithm is then applied to our intertemporal asset allocation problem taking account of various kinds of short-sale constraints.

The Structure of this Dissertation

The remainder of this work is organized as follows. We begin with a simple discrete-time example in Chapter 2, which can already illustrate the essential features of the intertemporal asset allocation model. Chapter 3 introduces at first Merton's general continuous-time framework of the intertemporal asset allocation model and then extends Merton's framework to incorporate a stochastic consumption price index. An analytical solution for the optimal investment strategy is obtained by using the Feynman-Kac formula. In order to give concrete suggestions for intertemporal asset allocation strategies, some real market situations are examined and analyzed. Chapter 4 is devoted to empirical research of the current market situation on which asset allocation strategies will be based. We employ two interest rate models related to two different approaches: The first model employs a data-oriented approach and the second one is based on theoretical considerations. The second model allows for inflation-indexed bonds. Both models belong to the general Duffie and Kan term structure framework. Chapter 5 provides examples to illustrate the intertemporal effect, the information effect (which information is important when conducting intertemporal portfolio strategies), and the hedging performance of inflation-indexed bonds. Real markets contain a number of imperfections that need to be taken into account when discussing optimal portfolio strategies. In Chapter 6 we consider the impact of several short-sale constraints on the intertemporal portfolios for which it is necessary to develop a numerical algorithm, in order to obtain the optimal strategies. The numerical algorithm we employ is the Markov Chain Approximation Method. Chapter 7 concludes the whole dissertation and suggests some future research directions. Technical derivations and proofs of the results are provided in the Appendix, including the link between Merton's continuous-time model and its discrete-time counterpart.

Chapter 2

An Initial Intertemporal Example

In this introductory section we show that many properties related to intertemporal portfolio decisions found by Merton (1971) can be shown in a simple discrete-time two-period model. Merton's intertemporal hedging effect states that, if there are time-varying factors affecting asset returns, the portfolio decision of a two-period investor is not only to maximize her/his utility of the asset return (according to the trade-off between expected return and risk) for one period but she/he will also hedge against the future development of the time-varying factors.

In this simple example, we can also provide the conditions where the intertemporal effect does not appear, i.e. the two-period optimal portfolio (an intertemporal portfolio) is exactly the same as the one-period optimal portfolio (a myopic portfolio) for each period. The conditions are, (i) when the asset returns are i.i.d., (ii) when the factor shocks and the asset return risk are independently distributed, and (iii) when the utility function is logarithmic. Although our two-period discrete-time model is much simpler than the general intertemporal continuous-time model, the conditions for the disappearance of the intertemporal effect are identical for these two different kinds of models.

In our discrete-time example, there are representative agents having a utility function of the constant-relative-risk-aversion (CRRA) type

$$U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma}. \quad (2.1)$$

The agents are given an initial wealth $V_0 > 0$ and invest it in a financial market in order to maximize their expected utility over two-period

$$\mathbf{E}_0 \left[U(C_0) + e^{-\delta} U(C_1) + e^{-2\delta} U(C_2) \right], \quad (2.2)$$

where \mathbf{E}_0 is expectation operator based on information until $t = 0$. C_t represents agents' consumption at $t = 0, 1, 2$. The parameter $\gamma \in (0, 1) \cup (1, \infty)$ represent the risk aversion of the agents. For the case $\gamma = 1$ we take the log-utility function¹

$$U(C_t) = \ln C_t .$$

The parameter $\delta > 0$ represents the discount effect due to deferred consumption.

There is a financial market for borrowing/lending money and investment in risky assets. The (one-period) interest rate R_t for borrowing and lending is not constant but follows the exogenous dynamics

$$R_{t+1} = R_t + \kappa(\bar{R} - R_t) + g\Delta W_{t+1}^R . \quad (2.3)$$

The interest rate R_t follows a *mean-reverting* process around the “mean” \bar{R} .

The financial market consists of one risky asset with the return dynamics

$$\frac{P_{t+1} - P_t}{P_t} = \mu(R_t) + \sigma(R_t)\Delta W_{t+1} . \quad (2.4)$$

The return of the risky asset is characterized by an expected average return $\mu(R_t) > 0$, which is affected by the current interest rate R_t . A Wiener process W_t is used to represent the risk of the asset return where the Wiener increment ΔW_t is normally distributed $\sim \mathcal{N}(0, 1)$. The positive coefficient $\sigma(R_t)$ characterizes the size of the return risk, which also depends on the current interest rate R_t .

We note that the distribution of the risky asset return is not the same but always changes with time because the interest rate R_t is time-varying. This

¹For the case $\gamma = 1$ we consider the shifted utility function $U(C_t) = \frac{C_t^{1-\gamma}}{1-\gamma} - \frac{1}{1-\gamma}$ and note that $\lim_{\gamma \rightarrow 1} \frac{C_t^{1-\gamma} - 1}{1-\gamma} = \ln C_t$. We do not consider this shifted utility for the other case because we need the utility function (2.1) to be homothetic for later use.

²When $R_t < \bar{R}$, there is a positive “drift” $\kappa(\bar{R} - R_t)$ to increase R_t so that the interest rate can return towards \bar{R} , and vice versa.

feature of the variation in time of the underlying distribution is the key feature of an *intertemporal* asset allocation model.

Let V_t ($t = 0, 1, 2$) denote the wealth the agents possess at each time point, where the agents can decide on the amount of consumption C_t ($t = 0, 1, 2$), and invest the remainder in the financial market. For their investment they decide a proportion α_t at $t = 0, 1$ to invest in the risky asset. They can also keep some of their wealth for lending and earn a certain return R_t , or they can borrow money for their investment. So, given their consumption and investment decisions, the wealth evolves according to

$$V_{t+1} = (V_t - C_t)\Pi_{t+1}(\alpha_t, R_t, \Delta W_{t+1}) \quad (2.5)$$

for $t = 0, 1$, where

$$\Pi_{t+1}(\alpha_t, R_t, \Delta W_{t+1}) = 1 + \alpha_t(\mu(R_t) + \sigma(R_t)\Delta W_{t+1}) + (1 - \alpha_t)R_t \quad (2.6)$$

is the portfolio return during the period $[t, t + 1]$ for $t = 0, 1$. We note that the investment strategy is self-financing³.

We use $J_0(V_0, R_0)$ to denote the value function which is defined as the maximized objective function of the two-period decision (2.2)

$$J_0(V_0, R_0) := \max_{C_0, C_1, \alpha_0, \alpha_1} \mathbf{E}_0 \left[U(C_0) + e^{-\delta}U(C_1) + e^{-2\delta}U(C_2) \right]. \quad (2.7)$$

Due to the natural time structure and the law of iterated expectations, the two-period decision problem (2.7) can be solved backwards sequentially, that is, first for the second period $t \in [1, 2]$, then going back to the first period $t \in [0, 1]$

$$J_0(V_0, R_0) = \max_{C_0, \alpha_0} \mathbf{E}_0 \left[U(C_0) + \max_{C_1, \alpha_1} \mathbf{E}_1 [e^{-\delta}U(C_1) + e^{-2\delta}U(C_2)] \right]. \quad (2.8)$$

This solution method is called *backward solution scheme* in general.

We look at the partial optimization problem for the second period $t = [1, 2]$ at first and let $J_1(V_1, R_1)$ denote the value function

$$J_1(V_1, R_1) := \max_{C_1, \alpha_1} \mathbf{E}_1 [e^{-\delta}U(C_1) + e^{-2\delta}U(C_2)].$$

³For the details see Appendix 8.1.5.

The agents spend all wealth for consumption at final time due to the increasing utility function. So, following the wealth dynamics (2.5), we obtain

$$\begin{aligned} & \max_{C_1, \alpha_1} \mathbf{E}_1 [e^{-\delta} U(C_1) + e^{-2\delta} U(C_2)] \\ &= \max_{C_1, \alpha_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} E_1 \left[U \left((V_1 - C_1) \Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\} \end{aligned} \quad (2.9)$$

Based on Property 23 in Appendix 8.1.6 which claims that the agents can do the two asset allocation decisions sequentially: first choosing the optimal portfolio, then deciding their consumption spending⁴, we can rewrite our partial optimization problem further as

$$\begin{aligned} & J_1(V_1, R_1) \\ &= \max_{C_1, \alpha_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} E_1 \left[U \left((V_1 - C_1) \Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\} \\ &= \max_{C_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} U(V_1 - C_1) (1 - \gamma) \max_{\alpha_1} E_1 \left[U \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\} \\ &= e^{-\delta} U(V_1) (1 - \gamma) \max_{\psi_1} \left\{ U(\psi_1) + e^{-\delta} U(1 - \psi_1) (1 - \gamma) \max_{\alpha_1} E_1 \left[U \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\} \\ &=: e^{-\delta} U(V_1) \Theta_1(R_1), \end{aligned} \quad (2.10)$$

where $\psi_1 := C_1/V_1$ and $\Theta_1(R_1)$ is defined as

$$\Theta_1(R_1) := (1 - \gamma) \max_{\psi_1} \left\{ U(\psi_1) + e^{-\delta} U(1 - \psi_1) (1 - \gamma) \max_{\alpha_1} E_1 \left[U \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\}.$$

The above manipulations rely on the equality based on the feature of the CRRA utility function (2.1) that

$$U(V_1 \psi_1) = V_1^{1-\gamma} U(\psi_1) = (1 - \gamma) U(V_1) U(\psi_1).$$

The sequential decision making above implies that the agents choose their optimal portfolio independently of the consumption decision, where the portfolio decision is taken to maximize agents' expected utility of the one-period portfolio returns, that is,

$$\max_{\alpha_1} E_1 \left[U \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right]. \quad (2.11)$$

In the literature, the one-period portfolio decision is also called *myopic* portfolio decision.

⁴which depends on the expected returns of the optimal portfolio

The fourth line of the equation (2.10) illustrates that the value function $J_1(V_1, R_1)$ depends on the endowment V_1 from the last period and the current interest rate R_1 in a separable way. The dependence of the value function $J_1(V_1, R_1)$ on the interest rate R_1 arises from the fact that the portfolio return $\Pi_2(\alpha_1, R_1, \Delta W_2)$ depends on the interest rate R_1 .

After solving the partial optimization problem for the period $[1, 2]$ we go back to solve the two-period decision problem (2.8). Using the result (2.10), the value function (2.8) can be rewritten as

$$\begin{aligned} J_0(V_0, R_0) &= \max_{C_0, \alpha_0} \mathbf{E}_0 [U(C_0) + J_1(V_1, R_1)] \\ &= \max_{C_0, \alpha_0} \mathbf{E}_0 [U(C_0) + e^{-\delta} U(V_1) \Theta_1(R_1)] . \end{aligned}$$

The optimization problem above shares the same form as the optimization problem (2.9). So, using the same technique as was used to obtain equation (2.10), we obtain the equation

$$\begin{aligned} &\max_{C_0, \alpha_0} \mathbf{E}_0 [U(C_0) + e^{-\delta} U(V_1) \Theta_1(R_1)] \\ &= \max_{C_0} \left\{ U(C_0) + e^{-\delta} U(V_0 - C_0)(1 - \gamma) \max_{\alpha_0} E_0 [U(\Pi_1(\alpha_0, R_0, \Delta W_1)) \Theta_1(R_1)] \right\} , \end{aligned}$$

where the portfolio decision is to maximize expected *intertemporal* utility

$$\max_{\alpha_0} E_0 [U(\Pi_1(\alpha_0, R_0, \Delta W_1)) \Theta_1(R_1)] . \quad (2.12)$$

Different from the objective function of the one-period case (2.11), the agents not only consider the expected utility of future portfolio returns

$$E_0 [U(\Pi_1(\alpha_0, R_0, \Delta W_1))] .$$

The agents who make decisions at $t = 0$ need to predict the future interest rate R_1 in order to achieve the optimality stated in (2.12), where the utility of each future portfolio return $\Pi_1(\alpha_0, R_0, \Delta W_1)$ has different weight $\Theta_1(R_1)$. In other words, by constructing the optimal portfolio in the two-period model, the agents need to consider not only the distribution of the realized portfolio return $\Pi_1(\alpha_0, R_0, \Delta W_1)$ but also the distribution of the future interest rate R_1 . This is the intuition of the *intertemporal hedging term* in the intertemporal portfolio decision. The discussion above can be extended to an n -period model as shown in Appendix 8.1.2. The intertemporal hedging term in the continuous-time model of Merton (1971, 1973) can

be also explained based on the same intuition.

We now show that the three conditions stated at the beginning of this section, namely, (i) when the asset returns are i.i.d. distributed, (ii) when the factor shocks and the asset return risk are independently distributed, and (iii) when the utility function is logarithmic, lead to the disappearance of the intertemporal effect so that the two-period optimal investment plan becomes identical to two one-period optimal portfolio decisions for each period. Recall the asset return assumption (2.4), the first condition is satisfied if the interest rate is constant $R_0 = R_1 = \bar{R}$. Then, the constant term $\Theta_1(R_1)$ in the two-period portfolio decision rule (2.12) can be taken out of the expectation operator

$$\max_{\alpha_0} \mathbf{E}_0 \left[U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right) \Theta_1(R_1) \right] = \Theta_1(R_1) \max_{\alpha_0} \mathbf{E}_0 \left[U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right) \right],$$

so the two-period portfolio decision becomes identical to the myopic portfolio decision as shown on the RHS.

The second condition for the disappearance of the intertemporal effect can be verified easily because

$$\begin{aligned} & \max_{\alpha_0} \mathbf{E}_0 \left[U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right) \Theta_1(R_1) \right] \\ &= \mathbf{E}_0 \left[\Theta_1(R_1) \right] \max_{\alpha_0} \mathbf{E}_0 \left[U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right) \right] \end{aligned}$$

if ΔW_1 and R_1 are independently distributed.

Now we discuss the case of the log-utility function $U(C_t) = \ln C_t$. As mentioned before, we solve the partial optimization problem for the period $t = [1, 2]$ at first. The value function can be rewritten as

$$\begin{aligned} & \max_{C_1, \alpha_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} \mathbf{E}_1 [U(C_2)] \right\} \\ &= \max_{C_1, \alpha_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} \mathbf{E}_1 \left[\ln \left((V_1 - C_1) \Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \right\} \\ &= \max_{C_1} \left\{ e^{-\delta} U(C_1) + e^{-2\delta} \left[\ln(V_1 - C_1) \right] \right\} + e^{-2\delta} \max_{\alpha_1} \mathbf{E}_1 \left[\ln \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right] \\ &= e^{-\delta} U(V_1) + e^{-2\delta} U(V_1) + \Theta_1(R_1), \end{aligned} \tag{2.13}$$

where

$$\Theta_1(R_1) = e^{-\delta} \max_{\psi_1} \left\{ U(\psi_1) + e^{-\delta} U(1 - \psi_1) \right\} + e^{-2\delta} \max_{\alpha_1} \mathbf{E}_1 \left[\ln \left(\Pi_2(\alpha_1, R_1, \Delta W_2) \right) \right],$$

with $\psi_1 = C_1/V_1$.

According to the result (2.13), we can rewrite the two-period asset allocation problem (2.2) as

$$\begin{aligned}
& \max_{C_0, C_1, \alpha_0, \alpha_1} \mathbf{E}_0 \left[U(C_0) + e^{-\delta} U(C_1) + e^{-2\delta} U(C_2) \right] \\
&= \max_{C_0, \alpha_0} \left\{ U(C_0) + \mathbf{E}_0 \left[e^{-\delta} U(V_1) + e^{-2\delta} U(V_1) + \Theta_1(R_1) \right] \right\} \\
&= \max_{C_0, \alpha_0} \left\{ U(C_0) + (e^{-\delta} + e^{-2\delta}) \mathbf{E}_0 \left[U \left((V_0 - C_0) \Pi_1(\alpha_0, R_0, \Delta W_1) \right) \right] \right\} + \mathbf{E}_0 [\Theta_1(R_1)] \\
&= \max_{C_0} \left\{ U(C_0) + (e^{-\delta} + e^{-2\delta}) U(V_0 - C_0) \right\} \\
&\quad + (e^{-\delta} + e^{-2\delta}) \max_{\alpha_0} \mathbf{E}_0 \left[U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right) \right] + \mathbf{E}_0 [\Theta_1(R_1)].
\end{aligned}$$

In the last equation we can see that the two-period optimal portfolio α_0 maximizes the one-period expected utility $U \left(\Pi_1(\alpha_0, R_0, \Delta W_1) \right)$, independent of the interest rate effect $\Theta_1(R_1)$. The reason why the intertemporal hedging effect should vanish here in the case of the log-utility can be mathematically explained by comparing the value function (2.10) of the case for $\gamma \neq 1$ and the value function (2.13) for the log-utility. The effect of the factor development $\Theta_1(R_1)$ in the multiplicative form (2.10) becomes the additive form (2.13) in the case of the log-utility. So, the future development of the factor does not affect the current portfolio decision.

Although the two-period discrete-time model is simple, it accommodates the essential features of the intertemporal asset allocation problem. It is useful to gain some intuitive insight into Merton's (1971) general continuous-time framework introduced in the next chapter. In the appendix we provide the general framework of the discrete-time intertemporal optimization problem.

Chapter 3

Intertemporal Asset Allocation under Inflation

Agents who want to construct a long-term asset allocation strategy must consider the risks of stochastic interest rates and inflation. In this chapter, we extend Merton's (1971) continuous-time framework to an asset allocation model which can include such long-term risks. Also, we introduce two main solution methods: the method of dynamic programming and the martingale method and discuss their application in our extended framework with inflation.

3.1 The Model

Just as in Merton's (1971) general asset allocation model, there are agents in our model who maximize their life-time expected utility by constructing their consumption plans and elaborate strategies for investing in financial markets.

For our model we have a probability space with the augmented natural filtration¹ $\{\mathcal{F}_t, t \in [0, T]\}$ and the real world measure \mathcal{P} .

We extend Merton's model by introducing a *stochastic* price index I_t , which is modelled by the diffusion process

$$\frac{dI_t}{I_t} = \pi_t dt + \sigma_I dW_t^I, \quad (3.1)$$

¹See Karatzas and Shreve (1991).

where W_t^I is a one-dimensional Wiener process and π_t is the *anticipated* instantaneous inflation rate². We assume that the latter is a function of some underlying factors X_t and we write $\pi_t = \pi(X_t)$. We normalize the initial price level by setting $I_0 = 1$.

With the introduction of the stochastic price index, the main difference between our model and that of Merton is that we have to distinguish real terms and nominal terms. The utility of agents will depend on *real* consumption while the asset prices in financial markets are evaluated in *nominal* terms.

The key feature of the intertemporal model is that there are *time-varying underlying factors*, which cause the evolution of, and uncertainty in, the financial markets. The underlying factors $X_t = \{X_{1t}, \dots, X_{nt}\}$ are modelled by an n -dimensional exogenous diffusion process

$$dX_t = F(X_t)dt + G(X_t)dW_t^X, \quad (3.2)$$

where $X_t = (X_{1t}, \dots, X_{nt})^\top$, F is an $\mathbb{R}^{n \times 1} \rightarrow \mathbb{R}^{n \times 1}$ function, and G is an $\mathbb{R}^{n \times 1} \rightarrow \mathbb{R}^{n \times n}$ function. The factor uncertainty $W_t^X = (W_{1t}^X, \dots, W_{nt}^X)^\top$ is an n -dimensional Wiener process with the correlation matrix $\mathcal{R}_{XX}dt := dW_t^X dW_t^{X\top}$. The correlation matrix between the price level shock and the factor uncertainty source is denoted by $\mathcal{R}_{IX}dt := dW_t^I dW_t^{X\top}$.

Economically relevant examples for such factors would include: interest rates, inflationary expectations, stochastic trends and stochastic volatilities in asset returns, and the Sharpe ratio.

For the factor X_t we require

Assumption 1 *The weak solution of the stochastic differential equation (3.2) to exist.*

Assumption 2 *The stochastic process X_t has a stationary distribution.*

There is a financial market where money is borrowed and lent and assets are traded. Agents borrow or lend cash at the nominal instantaneous interest rate R_t . We assume that the interest rate is determined by the underlying factors $R_t = R(X_t)$.

There are m sources of financial market uncertainty which are modelled by an m -dimensional Wiener process $W_t = (W_{1t}, \dots, W_{mt})^\top$ with the correlation matrix $\mathcal{R}_{AA}dt := dW_t dW_t^\top$. These may be different than the factor

²See Richard (1978).

uncertainty W_t^X mentioned above. We assume that there are more than m risky assets in the financial market. The instantaneous return of each asset dP_{it}/P_{it} is stochastic and is modelled by the diffusion process

$$\frac{dP_{it}}{P_{it}} = \mu_i(X_t, t)dt + \Sigma_i(X_t, t)dW_t, \quad \text{for } i = 1, \dots, \quad (3.3)$$

where the drift coefficients μ_i are $\mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}$ functions and the diffusion coefficients Σ_i are $\mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^{1 \times m}$ functions. All of these coefficients are functions of the underlying factors X_t and time t . The risk sources W_t are allowed to be correlated with W_t^X and W_t^I and their correlation matrices are denoted by $\mathcal{R}_{IA}dt := dW_t^I dW_t^{\top}$ and $\mathcal{R}_{XA}dt := dW_t^X dW_t^{\top}$.

We assume the financial market is well-functioning so that it satisfies the following *no-arbitrage* condition:

Assumption 3 *There are m functions of the underlying factors $\lambda_j : \mathbb{R}^n \rightarrow \mathbb{R}$ for $j = 1, \dots, m$ which satisfy the no-arbitrage relation*

$$\mu_i(X_t, t) - R_t = \Sigma_i(X_t, t)\lambda(X_t), \quad (3.4)$$

$\lambda(X_t) = (\lambda_1(X_t), \dots, \lambda_m(X_t))^{\top}$. The no-arbitrage relation (3.4) holds for all $t \in [0, T]$ and for any asset i in the financial market. Each function $\lambda_j(X_t)$ may be interpreted as the market price of each risk source W_{jt} , for $j = 1, \dots, m$.

The no-arbitrage condition is a central assumption in modern financial modelling. Quantitatively, the market price of risk is the excess return (price) per one unit volatility. So, the relation (3.4) has the interpretation that the excess return over the riskless return is equal to the sum of the risk premia required by the rational investors for bearing the risk associated with each risk source.

Furthermore, we assume that among the risky assets in the financial market there are m assets whose diffusion coefficients $\Sigma_i(X_t, t), i = 1, \dots, m$ are (almost surely) linearly independent. Under the no-arbitrage condition in Assumption A3 the other assets in the financial market can be replicated by a portfolio consisting of these m assets³. So, we need only to consider the m assets. We summarize the drift and diffusion coefficients of the m asset

³See Heath et al. (1992) for a fully rigorous discussion or Chiarella (2004) for a similar but more in intuitive discussion.

returns in the vector form

$$\mu_t = \begin{pmatrix} \mu_1(X_t, t) \\ \vdots \\ \mu_m(X_t, t) \end{pmatrix} \quad \text{and} \quad \Sigma_t = \begin{pmatrix} \Sigma_1(X_t, t) \\ \vdots \\ \Sigma_m(X_t, t) \end{pmatrix}$$

and note that Σ_t is (almost surely) full-rank so it is (almost surely) invertible. Then, the market prices of risk $\lambda(X_t)$ can be fully determined by these m assets. Mathematically, we state

Assumption 4

$$\lambda_t := \lambda(X_t) = \Sigma_t^{-1}(\mu_t - R_t \mathbf{1}), \quad (3.5)$$

where $\mathbf{1} = (1, \dots, 1)^\top$.

In addition, we still require

Assumption 5 *The economy described above is characterized by informational efficiency.*

By informational efficiency we mean that all sources of the factor uncertainty are included as a subset of the sources of the asset return uncertainty. Mathematically, we have $W_t^X \subseteq W_t$. Without loss of generality we specify the first n sources of the asset return uncertainty as the the factor uncertainty, that is $W_{it} = W_{it}^X$ for $i = 1, \dots, n$. The remaining Wiener processes for the asset return uncertainty are then denoted by $W_t^O = (W_{(n+1)t}, \dots, W_{mt})^\top$. So, all together we have $W_t = (W_t^X, W_t^O)^\top$.

The decision makers in our model are identical agents who are given (nominal) wealth endowment $V_0 > 0$ at initial time $t = 0$ and maximize their life-time expected utility

$$\max_{\alpha_t, c_t, t \in [0, T]} \mathbf{E}_0 \left[\epsilon_1 \int_0^T e^{-\delta t} U(c_t) dt + e^{-\delta T} U(c_T) \right] \quad (3.6)$$

by deciding their *real consumption* c_t and investment proportions α_t over the time horizon $[0, T]$. The utility at time t is a function of the real consumption c_t and is discounted by the factor $e^{-\delta t}$.

The utility function U is time-invariant and is of the constant relative risk aversion (CRRA) type so that

$$U(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}, \quad \text{with } \gamma > 0. \quad (3.7)$$

For the objective function (3.6) we can choose ϵ_1 to be 0 or 1. For $\epsilon_1 = 1$ all intermediate consumption is taken into consideration while for $\epsilon_1 = 0$ only final expected utility is considered. For the case $\epsilon_1 = 1$ we also allow $T \rightarrow \infty$.

The investment decision is denoted by $\alpha_t := (\alpha_{1t}, \dots, \alpha_{mt})^\top$ where $\alpha_{it}, i = 1, \dots, m$ is the investment in the i -th risky asset as a fraction of the wealth. The total position invested in the risky assets, $\sum_{i=1}^m \alpha_{it}$, may be greater than one (in the case of borrowing) or less than one (in the case of lending). So, the position α_{0t} defined as $\alpha_{0t} := 1 - \sum_{i=1}^m \alpha_{it}$ denotes the portfolio proportion of money holding. The (instantaneous) return on lending, or equivalently, the cost of borrowing is the instantaneous interest rate R_t .

Let V_t denote agents' nominal wealth at time t . Given decisions concerning investment proportions $\alpha_t = (\alpha_{1t}, \dots, \alpha_{mt})^\top$ and *nominal* consumption C_t , agents' nominal wealth changes at the rate

$$\begin{aligned} \frac{dV_t}{V_t} &= -\psi_t dt + \sum_{i=0}^m \alpha_{it} \frac{dP_{it}}{P_{it}} \\ &= (R_t - \psi_t) dt + \alpha_t^\top \left((\mu_t - R_t \mathbf{1}) dt + \Sigma_t dW_t \right), \end{aligned} \quad (3.8)$$

where $\psi_t := \frac{C_t}{V_t}$ is the nominal consumption ratio. These wealth dynamics are derived from their discrete-time counterpart and satisfy the self-financing budget constraint⁴.

Recall that the agents in our model are concerned with the utility of the *real* consumption c_t instead of the nominal consumption C_t . So, we need to transform the nominal wealth dynamics (3.8) into dynamics in real terms. As is well-known the relations between the real and nominal terms are given by $c_t = \frac{C_t}{I_t}$ and $v_t := \frac{V_t}{I_t}$ where v_t represents real wealth. Furthermore, the nominal consumption ratio is equal to the real consumption ratio, $\psi_t = \frac{C_t}{V_t} = \frac{c_t}{v_t}$.

Applying Itô's Lemma and employing the price dynamics (3.1) and the nominal wealth dynamics (3.8), we can obtain the evolution of the dynamics of

⁴This derivation, which is based on the discussion in Merton (1971), is shown in Appendix 8.1.5.

the real wealth, namely,

$$\begin{aligned} \frac{dv_t}{v_t} &= (R_t - \psi_t - \pi_t + \sigma_I^2)dt \\ &+ \alpha_t^\top (\mu_t - R_t \mathbf{1} - \sigma_I \Sigma_t \mathcal{R}_{AI})dt + \alpha_t^\top \Sigma_t dW_t - \sigma_I dW_t^I. \end{aligned} \quad (3.9)$$

Later, we will also consider short-sale constraints, for example, an additional short-sale commission η for each unit of short position. Taking into account the short sale commission η , the real wealth dynamics become

$$\begin{aligned} \frac{dv_t}{v_t} &= (R_t - \psi_t - \pi_t + \sigma_I^2)dt + \sum_{i=0}^m \min(0, \alpha_{it}) \eta dt \\ &+ \alpha_t^\top (\mu_t - R_t \mathbf{1} - \sigma_I \Sigma_t AI)dt + \alpha_t^\top \Sigma_t dW_t - \sigma_I dW_t^I. \end{aligned} \quad (3.10)$$

In summary, our asset allocation problem is to choose the real consumption ratios ψ_t and the portfolio strategies α_t for all $t \in [0, T]$ so that the life-time expected utility (3.6) of the real consumption $c_t = \psi_t v_t$ will be maximized. The financial market in which the agents invest is affected by the time-varying underlying factors following the dynamics (3.2). The dynamics (3.9) in the case without short-sale commissions and (3.10) in the case with the short-sale commissions, govern the evolutions of agents' real wealth given the asset allocation decisions $\{\psi_t, \alpha_t\}$ for $t \in [0, T]$.

We note that our control problem is time-dependent because the dynamics of the risky asset returns (3.3) are time-dependent. This time dependency is due to inclusion of bonds as assets, the expected return and the volatility of which change with time to maturity.

In the following sections two solution methods will be introduced and their application to the intertemporal asset allocation problem will be compared. The first one is the method of *dynamic programming* originally used by Merton (1971) and the second one is the *martingale method* introduced by Cox and Huang (1989).

3.2 Solution via Dynamic Programming

In this section we review the method of dynamic programming and derive the *Hamilton-Jacobi-Bellman* (HJB) equation. Here we only consider the case without short-sale commissions where the development of the real wealth follows the dynamics (3.9). The case with short-sale commissions and the other short-sale constraints will be considered in Chapter 6.

Let $J(t, T, v_t, X_t)$ denote the value function (the optimized objective function) for a sub-period $[t, T]$ with the given initial real wealth v_t and the given state of the factor X_t , so that

$$J(t, T, v_t, X_t) = \max_{\psi_s, \alpha_s; t \leq s \leq T} \left\{ \mathbf{E}_t \left[\epsilon_1 \int_t^T e^{-\delta s} U(\psi_s v_s) ds + e^{-\delta T} U(v_T) \right] \right\}. \quad (3.11)$$

The *Hamilton-Jacobi-Bellman*(HJB) equation⁵ characterizes the first order condition for the value function and is given by

$$\begin{aligned} 0 = \max_{\psi_t, \alpha_t} \left\{ \right. & \epsilon_1 e^{-\delta t} U(\psi_t v_t) + \frac{\partial J}{\partial t} \\ & + (R_t - \psi_t - \pi_t + \sigma_I^2 + \alpha_t^\top (\mu_t - R_t \mathbf{1} - \Sigma_t \mathcal{R}_{AI} \sigma_I)) J_v v_t \\ & + \frac{1}{2} (\alpha_t^\top \Sigma_t \mathcal{R}_{AA} \Sigma_t^\top \alpha_t - 2 \sigma_I \alpha_t^\top \Sigma_t \mathcal{R}_{AI} + \sigma_I^2) J_{vv} v_t^2 \\ & + (\alpha_t^\top \Sigma_t \mathcal{R}_{AX} G_t^\top - \sigma_I \mathcal{R}_I G_t^\top) J_{vX} v_t \\ & \left. + F_t^\top J_X + \frac{1}{2} \sum_{i,j=1}^n G_{it} \mathcal{R}_{XX} G_{jt}^\top J_{X_i X_j} \right\}, \quad (3.12) \end{aligned}$$

where $F_t := F(X_t)$, $G_t := G(X_t)$, and G_{it} is the i -th row of G_t .

We observe that during the time period $s \in [t, T]$, the factor dynamics dX_s defined by (3.2) are independent of the wealth level v_s . Since the utility function (3.7) is homothetic, the optimal consumption ratio ψ_s^* is independent of the wealth level v_s for all $s \in [t, T]$. Furthermore, the percentage change of the real wealth $\frac{dv_s}{v_s}$ defined by (3.9) is also independent of the wealth level v_s . Given the just-stated independence of v_t , it turns out that the initial wealth level v_t does not affect the optimal decisions ψ_s^* and α_s^* , and it can be treated as a scale multiplier of the intertemporal optimization problem (3.11). So, we can rewrite the value function as

$$\begin{aligned} & J(t, T, v_t, X_t) \\ = & v_t^{1-\gamma} \max_{\psi_s, \alpha_s; t \leq s \leq T} \left\{ \epsilon_1 \mathbf{E}_t \left[\int_t^T e^{-\delta s} U\left(\psi_s \frac{v_s}{v_t}\right) ds \right] + e^{-\delta T} \mathbf{E}_t \left[U\left(\frac{v_T}{v_t}\right) \right] \right\}. \end{aligned}$$

⁵The intuition concerning the HJB equation lies in the infinitesimal decomposition

$$J(t, T, v_t, X_t) = \max_{\psi_t, \alpha_t} \left\{ e^{-\delta t} U(\psi_t v_t) dt + J(t + dt, T, v_{t+dt}, X_{t+dt}) \right\}.$$

See pp.264-271 in Kamien and Schwartz (1991) for a heuristic discussion and Chapter 11 in Øksendal(2000) for a rigorous derivation.

Note that the re-scaled intertemporal optimization problem on the RHS is exactly the same as the intertemporal optimization problem with one unit of initial wealth, namely,

$$J(t, T, 1, X_t) = \max_{\psi_s, \alpha_s; t \leq s \leq T} \left\{ \epsilon_1 \mathbf{E}_t \left[\int_t^T e^{-\delta s} U(\psi_s \frac{v_s}{v_t}) ds \right] + e^{-\delta T} \mathbf{E}_t \left[U\left(\frac{v_T}{v_t}\right) \right] \right\}.$$

Using this result to transform the value function, we obtain the multiplicative expression

$$J(t, T, v_t, X_t) = v_t^{1-\gamma} J(t, T, 1, X_t) =: e^{-\delta t} U(v_t) \Phi(t, T, X_t)^\gamma. \quad (3.13)$$

where we set

$$\Phi(t, T, X_t)^\gamma := e^{\delta t} (1 - \gamma) J(t, T, 1, X_t) \quad (3.14)$$

and note that Φ is not a function of the initial wealth level v_t .

Based on the definition (3.11) of the value function $J(t, T, v_t, X_t)$, we obtain the boundary condition

$$J(T, T, v_T, X_T) = e^{-\delta T} U(v_T).$$

Considering the last equation in conjunction with the multiplicative formula (3.13), we obtain the boundary condition for the function $\Phi(t, T, X_t)$, namely,

$$\Phi(T, T, X_T) = 1. \quad (3.15)$$

From the first order conditions for ψ_t and α_t applied to the HJB equation (3.12), we have

$$\psi_t^* = \frac{1}{\Phi(t, T, X_t)}, \quad (3.16)$$

in the case with intermediate consumption, $\epsilon_1 = 1$. Using the result (3.16) to reduce the terms including ψ_t in the HJB equation (3.12), we obtain

$$e^{-\delta t} U(\psi_t^* v_t) - \psi_t^* J_v v = \frac{\gamma J}{\Phi}.$$

For the case without intermediate consumption ($\epsilon_1 = 0$) the optimal consumption decision is $\psi_t^* = 0$.

From the FOC for α_t we obtain the expression of the optimal α_t , which is

given by

$$\begin{aligned}
\alpha_t^* &= (\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} \left(-\frac{J_v v_t}{J_{vv} v_t^2} (\mu_t - R_t \mathbf{1}) - \frac{1}{J_{vv} v_t^2} \Sigma_t \mathcal{R}_{AX} G_t^\top J_{vX} v_t \right. \\
&\quad \left. + \frac{J_v v_t + J_{vv} v_t^2}{J_{vv} v_t^2} \sigma_I \Sigma_t \mathcal{R}_{AI} \right) \\
&= (\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} \left(\underbrace{\frac{1}{\gamma} (\mu_t - R_t \mathbf{1})}_{\text{(I)}} + \underbrace{\Sigma_t \mathcal{R}_{AX} G_t^\top \frac{\Phi_X}{\Phi}}_{\text{(II)}} + \underbrace{\left(1 - \frac{1}{\gamma}\right) \sigma_I \Sigma_t \mathcal{R}_{AI}}_{\text{(III)}} \right)
\end{aligned} \tag{3.17}$$

$$\begin{aligned}
&= (\Sigma_t^\top)^{-1} \left(\frac{1}{\gamma} \mathcal{R}_{AA}^{-1} \Sigma_t^{-1} (\mu_t - R_t \mathbf{1}) + \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} G_t^\top \frac{\Phi_X}{\Phi} - \frac{1-\gamma}{\gamma} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I \right) \\
&= (\Sigma_t^\top)^{-1} \left(\frac{1}{\gamma} \mathcal{R}_{AA}^{-1} \lambda_t + \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} G_t^\top \frac{\Phi_X}{\Phi} - \frac{1-\gamma}{\gamma} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I \right).
\end{aligned} \tag{3.18}$$

We can interpret the optimal portfolio allocation as being determined through the trade-off between the asset risks $\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top$ and the three "benefits", denoted as I, II, and III in the parentheses in equation (3.17). The first benefit I is the expected excess return, so the corresponding portfolio $(\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} \text{I}$ is called the *mean-variance efficient* portfolio. It is also known as the *myopic portfolio*. The second benefit II arises from the correlation between the asset return uncertainty and the factor shocks $\Sigma_t \mathcal{R}_{AX} G_t^\top$, and the factor effect on the objective function Φ_X/Φ , which appear only in an intertemporal model. It is called a benefit because a sophisticated portfolio decision can increase her/his utility by making use of these relations between the asset returns uncertainty and the factor shocks. We call the corresponding portfolio $(\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} \text{II}$ the *intertemporal hedging portfolio*. (Merton termed it as the intertemporal hedging term.) How to increase the utility using this term? We give an example where we consider the instantaneous (real) interest rate r_t is one of the underlying factors. Assume that a higher interest rate is more favored by the agents, that is, $J_r > 0$. For the case $\gamma > 1$,⁶ the positive effect of the interest rate $J_r > 0$ leads to the negative effect $\Phi_r < 0$.⁷ So, one can increase utility by investing more in an asset, for example, a bond, whose return is negatively correlated with interest rate

⁶This is the case we shall consider, see Section 5.1 for a discussion on the role of the risk aversion coefficient γ on the solution.

⁷This relationship is obtained by looking at equation (3.13) with r is one component in X . We can see J_r and Φ_r must have different signs because $U(v_t)$ is negative in the case $\gamma > 1$.

shocks. For the intuition of the structure of the intertemporal portfolio allocation α_t^* we refer to the initial example in Chapter 2 where we have seen the same logic for the myopic and the intertemporal hedging portfolio.

The third benefit III comes from the correlation between the asset return uncertainty and the price index shock. A sophisticated investment decision should consider this correlation because it affects the evolution of the real wealth. We call the corresponding portfolio $(\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} \text{III}$ the *index hedging portfolio* or *inflation hedging portfolio*. In Brennan and Xia (2002) and Munk et al. (2004) we can also find the same decomposition of the optimal portfolio.

We now elaborate further on the structure of the HJB equation (3.12) with the notations we have adopted. Using the results of the product form (3.13)⁸ and of the optimal decisions (3.16), (3.17), the HJB equation (3.12) is transformed into the form

$$\begin{aligned}
0 = & \epsilon_1 + \frac{\partial}{\partial t} \Phi + F_t^\top \Phi_X \\
& + \left(\frac{1-\gamma}{\gamma} G_t \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \Sigma_t^{-1} (\mu_t - R_t \mathbf{1}) - \frac{(1-\gamma)^2}{\gamma} G_t \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I - (1-\gamma) G_t \mathcal{R}_{XI} \sigma_I \right)^\top \Phi_X \\
& + \frac{1}{2} \sum_{i,j=1}^n \Phi_{X_i X_j} G_{it} \mathcal{R}_{XX} G_{jt}^\top \\
& + \frac{1-\gamma}{2\Phi} \sum_{i,j=1}^n \Phi_{X_i} \Phi_{X_j} (G_{it} \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} G_{jt}^\top - G_{it} \mathcal{R}_{XX} G_{jt}^\top) \\
& + \Phi \left(-\frac{\delta}{\gamma} + \frac{1-\gamma}{\gamma} (R_t - \pi_t + \sigma_I^2) + \frac{1-\gamma}{2\gamma^2} (\mu_t - R_t \mathbf{1})^\top (\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} (\mu_t - R_t \mathbf{1}) \right. \\
& \quad \left. + \frac{(1-\gamma)^3}{2\gamma^2} \sigma_I^2 \mathcal{R}_{IA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} - \frac{(1-\gamma)^2}{\gamma^2} (\mu_t - R_t \mathbf{1})^\top \Sigma_t^{-1} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I - \frac{1-\gamma}{2} \sigma_I^2 \right).
\end{aligned} \tag{3.19}$$

The main task now in obtaining the solution to the intertemporal asset allocation problem is to solve the HJB equation (3.19), a non-linear second

⁸From which we readily calculate:

$$\begin{aligned}
\frac{\partial}{\partial t} J &= -\delta J + \gamma \frac{\Phi_t}{\Phi} J, & J_v v &= (1-\gamma) J, \\
J_{vv} v^2 &= (1-\gamma)(-\gamma) J, & J_X &= \gamma \frac{\Phi_X}{\Phi} J, \\
J_{vX} v &= (1-\gamma) \gamma \frac{\Phi_X}{\Phi} J, & J_{X_i X_j} &= \left(\gamma(\gamma-1) \frac{\Phi_{X_i} \Phi_{X_j}}{\Phi} + \gamma \frac{\Phi_{X_i X_j}}{\Phi} \right) J.
\end{aligned}$$

order partial differential equation for the value function $\Phi(t, T, X_t)$.

Based on the Assumption 5, the informational efficiency, we can obtain the relation using standard matrix operations

$$\mathcal{R}_{XA}\mathcal{R}_{AA}^{-1}\mathcal{R}_{AX} = \mathcal{R}_{XA} \begin{pmatrix} \mathcal{I}_X \\ 0_{(m-n)\times n} \end{pmatrix} = \mathcal{R}_{XX}, \quad (3.20)$$

where \mathcal{I}_X is an n -dimensional unit matrix. Then the fourth line in the HJB equation (3.19) turns out to be zero and the HJB equation reduces to a *linear* second order partial differential equation (PDE).

In the next section we employ another approach, using the Feynman-Kac formula, to solve the HJB equation (3.19).

We remark that the structure of the HJB equation (3.19) is the same as that in Liu (2005) although in Liu's model a *constant* price index is considered. Due to the same structure of the two HJB equations, Liu's approach using the verification theorem⁹ may be extended to the intertemporal model with the time-varying price index given by (3.1).

3.3 Representation of the Solution via the Feynman-Kac Formula

To implement the Feynman-Kac formula, we first simplify the notation in the HJB equation (3.19) by setting

$$\begin{aligned} z(X_t) &:= \frac{1-\gamma}{\gamma}\mathcal{R}_{XA}\mathcal{R}_{AA}^{-1}\lambda(X_t) - \frac{(1-\gamma)^2}{\gamma}\mathcal{R}_{XA}\mathcal{R}_{AA}^{-1}\mathcal{R}_{AI}\sigma_I - (1-\gamma)\mathcal{R}_{XI}\sigma_I \\ &=: z_t \end{aligned} \quad (3.21)$$

$$\begin{aligned} h(X_t) &:= -\frac{\delta}{\gamma} + \frac{1-\gamma}{\gamma}(R(X_t) - \pi(X_t) + \sigma_I^2) + \frac{1-\gamma}{2\gamma^2}\lambda(X_t)^\top\mathcal{R}_{AA}^{-1}\lambda(X_t) \\ &\quad + \frac{(1-\gamma)^3}{2\gamma^2}\sigma_I^2\mathcal{R}_{IA}\mathcal{R}_{AA}^{-1}\mathcal{R}_{AI} - \frac{1-\gamma}{2}\sigma_I^2 - \frac{(1-\gamma)^2}{\gamma^2}\lambda(X_t)^\top\mathcal{R}_{AA}^{-1}\mathcal{R}_{AI}\sigma_I \\ &=: h_t. \end{aligned} \quad (3.22)$$

Then the HJB equation (3.19) becomes

$$0 = \frac{\partial}{\partial t}\Phi + (F_t + G_t z_t)^\top \Phi_X + \frac{1}{2} \sum_{i,j=1}^n \Phi_{X_i X_j} G_{it} G_{jt}^\top + \Phi h_t + \epsilon_1. \quad (3.23)$$

⁹Recall that this involves guessing a solution and then verifying it.

We apply *Feynman-Kac* formula to solve the HJB equation (3.19), or its simplified form (3.23). The details of applying the Feynman-Kac formula to solve a linear second order PDE can be found in Theorem 2 in Appendix 8.2.

Property 1 *Let $(X_s)_{s \in [0, T]}$ be the solution of the the SDE (3.2). Let $(z_s)_{s \in [0, T]}$ and $(h_s)_{s \in [0, T]}$ be the processes defined in (3.21) and (3.22) respectively. Further we assume that the process $(z_s)_{s \in [0, T]}$ satisfies the Novikov condition*

$$\mathbf{E} \left[\exp \left(\int_0^T z_s^\top \mathcal{R}_{XX}^{-1} z_s ds \right) \right] < \infty. \quad (3.24)$$

Then the function $\Phi(t, T, x)$ satisfying the PDE (3.23) and the boundary condition (3.15) is given by

$$\Phi(t, T, x) = \mathbf{E}_{t,x} \left[e^{\int_t^T h_s ds} \Lambda_T + \epsilon_1 \int_t^T e^{\int_t^s h_u du} \Lambda_s ds \right], \quad (3.25)$$

where

$$\Lambda_s := \exp \left(\int_0^s z_u^\top \mathcal{R}_{XX}^{-1} dW_u^X - \frac{1}{2} \int_0^s z_u^\top \mathcal{R}_{XX}^{-1} z_u du \right), \quad (3.26)$$

for $s \in [0, T]$. The expectation operator $\mathbf{E}_{t,x}$ takes the expectation with respect to the process $(X_s)_{s \in [t, T]}$ with given initial value $X_t = x$.

3.4 The Martingale Method: An Alternative Solution Strategy

Cox and Huang (1989), and Karatzas, Lechoczky and Shreve (1987) provided an elegant way, the martingale method, to solve the asset allocation problem when the financial market is complete and free of arbitrage. Later Karatzas et. al (1991) extended this method to the case of incomplete markets.

The advantage of the martingale method is to provide some insight into the general structure of asset allocation problems. However, the martingale method has not developed as a practical solution tool, so only a few papers, for example, Wachter (2002), have provided analytical solutions for the asset allocation problem based on this method. For the extended intertemporal framework with the time-varying price index, a direct application of the

martingale method is not available yet and requires more research¹⁰. Nevertheless, due to the general aspect of the martingale method we still study the martingale method and show the equivalence between the solution via the application of the Feynman-Kac formula given in Property 1 and the martingale method for the case of a *constant* price index: $I_t \equiv 1$ for all $t \in [0, T]$. This approach of showing the equivalence is different to that of Cox and Huang (1989)¹¹. Further research is needed to extend the application of the martingale method to the intertemporal model with a time-varying price index.

We review in brief the martingale method based on Chapters 2 and 5 in Korn and Korn (2000). The martingale method can be applied in the case where there exists a uniquely determined martingale measure (or the risk neutral measure) for asset pricing. The relation of the martingale measure P^* to the original measure \mathcal{P} is given by $\frac{dP^*}{d\mathcal{P}} = H(0, T)$ where the measure transformation $H(s, t)$ over the time $[s, t]$ is given by

$$H(t, s) = \exp\left(-\int_t^s R_u du\right) \exp\left(-\int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} dW_u - \frac{1}{2} \int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} \lambda_u du\right). \quad (3.27)$$

Under the new measure \mathcal{P}^* , all asset prices in the financial market are martingale processes, see Theorem 3.14 on p. 100 in Korn and Korn (2000).

The basic idea of the martingale to solve the intertemporal asset allocation problem is to replace the wealth dynamics (3.8) by the martingale inequality

$$V_0 \geq \max_{C_t, \alpha_t} \mathbf{E}^* \left[\int_0^T C_t dt + U(V_T) \right], \quad (3.28)$$

where the expectation \mathbf{E}^* is calculated with respect to the martingale measure P^* and $(C_t, \alpha_t)_{t \in [0, T]}$ is a self-financing strategy, see Theorem 2.63 on p. 65 in Korn and Korn (2000).

¹⁰The main difficulty of a direct application of the martingale method to the intertemporal framework with a time-varying price index is the two-tier structure: consumption evaluated in real terms while trading activities evaluated in nominal terms. The view taken in this dissertation is that for the the long-term asset allocation decision problem, consumption evaluated in *real* terms is more relevant than in nominal terms in agents' objective function. However, the martingale inequality given below in equation (3.28) is expressed in *nominal* terms. More research is needed to answer the question: should the martingale measure be based on nominal or real terms?

¹¹Cox and Huang showed the both methods satisfy the same first order conditions.

Theorem 1 *Let the price index be constant $I_s = 1, \forall s \in [0, T]$. The intertemporal asset allocation problem (3.11) becomes*

$$J(t, T, \bar{V}, x) = \max_{C_s, \alpha_s; s \in [t, T]} \left\{ \mathbf{E}_t \left[\epsilon_1 \int_t^T e^{-\delta s} U(C_s) ds + e^{-\delta T} U(V_T) \right] \right\}, \quad (3.29)$$

where the dynamics of the nominal wealth is given by equation (3.8). The initial conditions are given by $V_t = \bar{V}$ and $X_t = x$. Then the optimal consumption for the intertemporal problem (3.29) with the given initial conditions is given by

$$C_s^* = \epsilon_1 e^{-\frac{\delta(s-t)}{\gamma}} y^{-\frac{1}{\gamma}} H(t, s)^{-\frac{1}{\gamma}}, \quad \forall s \in [t, T) \quad (3.30)$$

and the final wealth under the optimal consumption plan is given by

$$V_T^* = e^{-\frac{\delta(T-t)}{\gamma}} y^{-\frac{1}{\gamma}} H(t, T)^{-\frac{1}{\gamma}}. \quad (3.31)$$

Here y is the Lagrangian constant satisfying the martingale constraint

$$\bar{V} = \mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T H(t, s) C_s^* ds + H(t, T) V_T^* \right].$$

Based on this martingale constraint, the Lagrangian constant y satisfies

$$y^{-\frac{1}{\gamma}} = \bar{V} \left(\mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T H(t, s)^{1-\frac{1}{\gamma}} e^{-\frac{\delta(s-t)}{\gamma}} ds + H(t, T)^{1-\frac{1}{\gamma}} e^{-\frac{\delta(T-t)}{\gamma}} \right] \right)^{-1}.$$

The solution for the value function (3.29) is obtained by employing the optimal strategies given by (3.30) and (3.31), so that

$$\begin{aligned} J(t, T, \bar{V}, x) &= \mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T e^{-\delta s} U(C_s^*) ds + e^{-\delta T} U(V_T^*) \right] \\ &= e^{-\delta t} U(\bar{V}) \left(\mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T H(t, s)^{1-\frac{1}{\gamma}} e^{-\frac{\delta(s-t)}{\gamma}} ds + H(t, T)^{1-\frac{1}{\gamma}} e^{-\frac{\delta(T-t)}{\gamma}} \right] \right)^\gamma. \end{aligned}$$

Furthermore, the term $\Phi(t, T, x)$ in the multiplicative form given (3.14) is solved by

$$\Phi(t, T, x) = \mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T e^{-\frac{\delta(s-t)}{\gamma}} H(t, s)^{1-\frac{1}{\gamma}} ds + e^{-\frac{\delta(T-t)}{\gamma}} H(t, T)^{1-\frac{1}{\gamma}} \right]. \quad (3.32)$$

Proof see Theorem 5.8 on p. 210 in Korn and Korn (2000).

Now we show the relation between the solution via the martingale method and that obtained by the Feynman-Kac formula introduced in the previous subsection.

Property 2 *If the price index is always equal to one $I_s = 1, \forall s \in [0, T]$ in the investment environment and if the risk sources of the asset returns $(W_s)_{s \in [0, T]}$ can be spanned by the risk sources of the factor $(W_s^X)_{s \in [0, T]}$, then the optimal asset allocation strategy obtained by the Feynman-Kac formula (3.25) is identical to the optimal strategy obtained by the martingale method (3.32).*

3.5 Summary and Remarks

In this chapter we have extended Merton's asset allocation model to accommodate a stochastic price level. The main innovation of our model is that the consumption and portfolio decisions are evaluated in different terms: the objective of agents is to maximize *real* consumption while the investment activities are evaluated in *nominal* terms. We were able to solve the asset allocation model with the stochastic price level. The two main solution methods for the intertemporal asset allocation problems: the method of dynamic programming and the martingale method were introduced, compared, and discussed. The application of the martingale method in an environment with the stochastic price level requires more new research before it can be considered to be a useful alternative to dynamic programming.

This chapter has only set up the mathematical "skeleton" for the solution of the intertemporal asset allocation problem. In order to be able to give actual asset allocation recommendations, we need further elaboration of the basic skeleton, such as, what are these underlying factors X_t ? How do the underlying factors affect the interest rate $R(X_t)$ and the inflation $\pi(X_t)$? How do the asset expected return and risks $\mu(X_t, t)$ and $\Sigma(X_t, t)$ depend on the factors X_t and time t ? How does one determine the market price of risk $\lambda(X_t)$?

The task of the next chapter is to obtain answers to these questions via an empirical study of specific financial markets.

Chapter 4

Empirical Financial Markets

Why should we consider intertemporal asset allocation strategies? Are there really time-varying underlying factors affecting the asset returns on the real market that render inadequate myopic investment strategies? This chapter provides empirical support for the existence of such time-varying factors so that it is indeed necessary to consider intertemporal strategies.

The previous chapter showed how to construct the optimal strategies for the intertemporal asset allocation problem. The actual implementation of the optimal asset allocation strategy requires information about the investment environment. So, this chapter is devoted to an empirical study of financial markets so as to obtain the required information for the optimal strategies.

When considering long-term investment plans, bond assets are usually considered as important financial instruments. They are considered as safer assets due to their regular and fixed payments although they have a very complicated structure as a whole. Purchasing bonds with different time to maturity provides different rates of return. So, the question of the term structure of interest rates naturally come into play in a consideration of the long-term asset allocation problem. For this reason, the major part of this chapter is devoted to modelling bond markets and the term structure of interest rates.

Two factors, interest rate fluctuations and inflationary expectations, are important for the modelling of the term structure and long-term investment strategies. Future inflationary expectations will affect the expected bond returns. Regarding the long-term investment decisions, interest rates will

definitely change their values for over a long horizon. Also, the expectation of future purchasing power will have some influence on long-term consumption plans.

We will consider two term structure models in this chapter. The first model is a data-oriented multi-factor interest rate model. Following the yield-factor framework of Duffie and Kan (1996), we assume in the first model that all interest rates, long-term and short-term, are affected by some common and unobservable factors. This model is a data-oriented model because we do not specify the common factors a priori as some economic variables but we “learn” them from observed market data using the Kalman filter.

The second model, in contrast to the data-oriented approach of the first model, is based on an exact specification of the underlying factors. The two factors chosen in the model are the instantaneous real interest rate and the expected instantaneous inflation rate due to their importance as mentioned above. In addition it is necessary to model *inflation-indexed bonds* (IIB), whose payout is adjusted to some price index. In the US, Treasury Indexed-Protected Securities (TIPS) issued by the US Treasury are adjusted to the Consumer Price Index of all Urban (CPI-U). The IIBs provide the possibility to determine *real* interest rates on the markets. The empirical investigation is carried out by applying the Kalman filter method to estimate the instantaneous real interest rate and inflationary expectations.

Both models will be estimated using U.S. interest rate data from 2003-2005. The IIBs in U.S. have been issued since January 1997. In the period of observation, the market real interest rate, based on market data of the IIBs, are available. The estimation results of the two models with the different approaches, the data-oriented and the theoretical, will be compared at the end of this chapter.

An empirical study of stock prices is also provided where a simple model for stock dynamics, geometric Brownian motion, is adopted.

This chapter is organized as follows. The Duffie-Kan model of Duffie and Kan (1996), to which the two models considered in this chapter belong, is briefly reviewed in Section 4.1. Section 4.2 introduces the market data we will use. The data-oriented model is introduced in Section 4.3. In 4.4 we develop a new model, which can model both nominal bonds and inflation-indexed bonds. The last section compares these two models and draws some conclusions.

4.1 The General Duffie-Kan Affine Term Structure Model

In this section we give a brief review of the Duffie and Kan (1996) affine term structure model.

Let $P_n(t, T, X_t)$ denote the zero-coupon *nominal* bond at t with maturity date T . The payout of the nominal bond is normalized as one money unit $P_n(T, T, X_T) = 1$. The bond price of the affine term structure model is assumed to be of the form

$$P_n(t, T, X_t) = e^{-A(T-t) - B(T-t)^\top X_t}, \quad (4.1)$$

where $B(\tau)^\top = (B_1(\tau), \dots, B_n(\tau))$. The coefficients $A(\tau), B_1(\tau), \dots, B_n(\tau)$ are assumed to be differentiable. From the normalization $P_n(T, T) = 1$ we have $A(0) = B_1(0) = \dots = B_n(0) = 0$. Duffie and Kan (1996) show that this exponential affine structure can be supported by processes X with linear drift and square-linear diffusion coefficients

$$dX_t = \mathcal{K}(\theta - X_t)dt + \Gamma \sqrt{S_t} dW_t^X, \quad (4.2)$$

where $\theta \in \mathbb{R}^{n \times 1}, \mathcal{K} \in \mathbb{R}^{n \times n}$ and $\Gamma \in \mathbb{R}^{n \times n}$ and

$$\sqrt{S_t} = \begin{pmatrix} \sqrt{S_1(X_t)} & 0 & \cdots & 0 \\ 0 & \sqrt{S_2(X_t)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{S_n(X_t)} \end{pmatrix}$$

with $S_i(X_t) = \alpha_i + \beta_i^\top X_t$.

The *yield* on the nominal bond which is defined as the average return between t and T , can be denoted by

$$Y_n(t, T, X_t) := \frac{\ln P_n(T, T, X_t) - \ln P_n(t, T, X_t)}{T - t} = \frac{A(T-t)}{T-t} + \frac{B(T-t)^\top}{T-t} X_t. \quad (4.3)$$

The instantaneous nominal interest rate R_t is set equal to the “instantaneous” yield, given by

$$R_t := \lim_{T \downarrow t} Y_n(t, T) = \xi_0 + \xi_1^\top X_t, \quad (4.4)$$

where $\xi_0 = A'(0)$, $\xi_1 = (\xi_{11}, \dots, \xi_{1n})^\top$ and $\xi_{1i} = B'_i(0)$. The nominal money account is defined as the accumulation account

$$P_{0t} = \exp\left(\int_0^t R_s ds\right). \quad (4.5)$$

We assume that both of our bond models satisfy the no-arbitrage condition A4 in Section 2.1. In this chapter, we assume further that

Assumption 6 *The market prices of risk $\lambda(X_t) \equiv \lambda$ are constants*

and

Assumption 7 *All bonds are default-free.*

4.2 Data

The data used for model estimation are US bond data including nominal bond yields and real bond yields. They are daily data over the period from Jan. 02, 2003 until May 31, 2005 containing 603 observations¹. The nominal bond yields are calculated based on market returns of Treasury nominal bond securities using the cubic spline method² with time to maturity of 1 month, 3 months, 6 months, 1 year, 2, 3, 5, 7, 10, and 20 years. These daily nominal yields are shown in the dashed blue line in Figure 4.1 and compared with the effective Federal Fund Rate³ in the solid black line. Basic statistics of the nominal bond yields are given in Table 4.1 where we can see that the nominal bond yields increase with time to maturity.

The real yields are calculated based on market returns of Treasury Indexed-Protected Securities (TIPS) using the same cubic spline method as for the nominal yields. The daily real yields with time to maturity of 5, 7, and 10 years are also shown in the dash blue line in Figure 4.2 together with the effective Federal Fund Rate. The basic statistics for the real bond yields are also given in Table 4.1.

¹The data are provided by the US Treasury at <http://www.ustreas.gov/offices/domestic-finance/debt-management/interest-rate/>

²Go to the page <http://www.ustreas.gov/offices/domestic-finance/debt-management/interest-rate/> and click on “Treasury Yield Curve Methodology”.

³The data are provided by Federal Reserve Bank New York at <http://www.newyorkfed.org/markets/omo/dmm/fedfundsdata.cfm>.

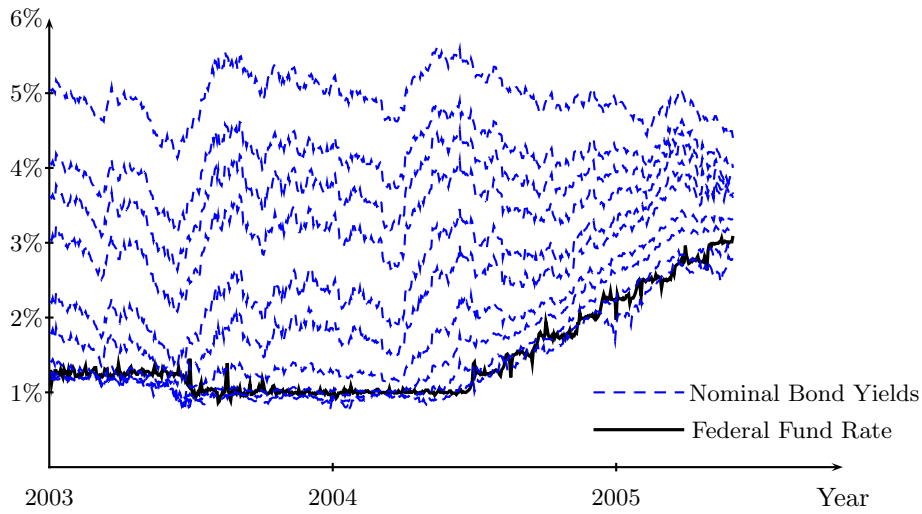


Figure 4.1: US Nominal Bond Yields and Federal Funds Rate (FFR)

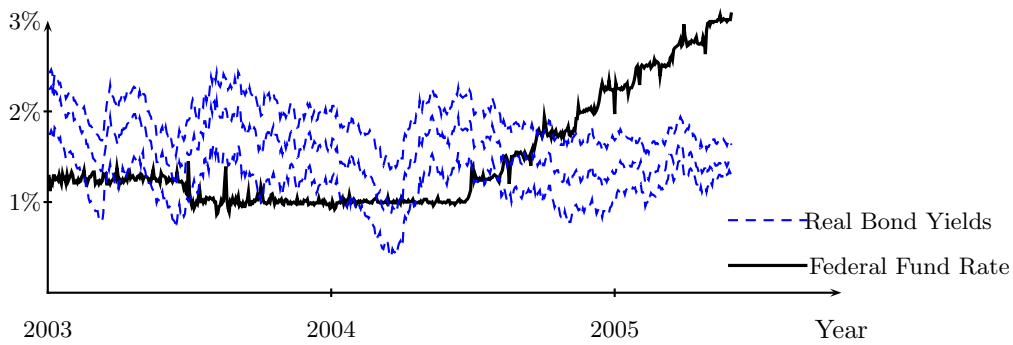


Figure 4.2: US Real Bond Yields and Federal Funds Rate (FFR)

During the observation period we have the macroeconomic scenario that the Federal Open Market Committee (FOCM) increased the target Federal Funds Rate continuously since the 2nd Quarter of 2004. From Figures 4.1 and 4.2 we make two observations. First, the short term nominal bond yields follow the increasing effective Federal Funds Rate while the long term nominal bond yields remain at the same level. So, the term premia reduced

Nominal Yields

Maturity	1M	3M	6M	1Y	2Y
Mean	1.37%	1.47%	1.63%	1.84%	2.28%
Std. dev.	0.59%	0.66%	0.75%	0.75%	0.75%
	3Y	5Y	7Y	10Y	20Y
	2.65 %	3.32%	3.76%	4.17%	4.95%
	0.65%	0.47%	0.37%	0.32%	0.29%

Real Yields

Maturity	5Y	7Y	10Y
Mean	1.16%	1.56%	1.90%
Std. dev.	0.25%	0.26%	0.23%

Table 4.1: Statistics for US Nominal/Real Bond Yields

over this period. Second, over the same period the real bond yields still remain stationary. We will discuss these empirical findings later based on the theoretical framework of the second model in Section 4.4.

For the estimation task we set one time unit equal to one year. The time interval for daily data is thus $1/250$ and for monthly data $1/12$.

4.3 Model I; Unspecified Factors

The first term structure model is a data-oriented model. We adopt the bond yield formula (4.3) where the nominal yields are assumed to be affected by some common unobservable factor X_t . The data used for the estimation are (only) nominal bond yields. Technically, we do not specify the factors X_t as specific economic variables but will “learn” them from the yield data.

This section contains the following subsections. Subsection 4.3.1 reviews briefly the bond pricing formula based on a Gaussian factor X_t . Before calibrating the model (4.3) we discuss the *model identification problem* in Subsection 4.3.2, in order to rule out a multiple parameter representations of the model. Finally, Subsection 4.3.3 provides the empirical study based on the US nominal yield data.

4.3.1 The Model

For the first model we assume the factors follow an n -dimensional Gaussian process

$$dX_t = \mathcal{K}(\theta - X_t)dt + \Gamma dW_t^X, \quad (4.6)$$

where W_t^X is a *standard* (orthogonal) n -dimensional Wiener process. We require \mathcal{K} to be positive definite so that the process X_t is stationary. We also require that \mathcal{K} has distinct eigenvalues.

Applying Itô's Lemma to the bond price formula (4.1), we can write the instantaneous return of the nominal bond as

$$\frac{dP_n(t, T, X_t)}{P_n(t, T, X_t)} = \mu_P(T - t, X_t)dt - B(T - t)^\top \Gamma dW_t^X, \quad (4.7)$$

where

$$\mu_P(\tau, X_t) = A'(\tau) + B'(\tau)^\top X_t - B(\tau)^\top \mathcal{K}(\theta - X_t) + \frac{1}{2} \sum_{i,j=1}^n B_i(\tau) B_j(\tau) \Gamma_i \Gamma_j^\top, \quad (4.8)$$

and Γ_i denotes the i -th row in Γ .

Our Assumption 4 in Section 3.1 and Assumption 6 in Section 4.1 lead to the no-arbitrage condition in the form

$$\mu_P(\tau, X_t) - R_t = -B(\tau)^\top \Gamma \lambda, \quad \text{for all } \tau > 0. \quad (4.9)$$

The arbitrage constraint (4.9) requires that the coefficients $A(\tau)$ and $B(\tau)$ satisfy the following ordinary differential equations.

$$\frac{d}{d\tau} B(\tau) = -\mathcal{K}^\top B(\tau) + \xi_1, \quad (4.10)$$

$$\frac{d}{d\tau} A(\tau) = (\mathcal{K}\theta - \Gamma\lambda)^\top B(\tau) - \frac{1}{2} \sum_{i,j=1}^n B_i(\tau) B_j(\tau) \Gamma_i \Gamma_j^\top + \xi_0. \quad (4.11)$$

4.3.2 Model Identification

The first estimation task is determine the parameters $(\theta, \mathcal{K}, \Gamma, \lambda, \xi_0, \xi_1)$ based on the empirically observed data. So, before the model estimation we have to discuss the *identification problem*, which arises due to the fact that one single data generating process may have many different parameter representations.

For our case where the model is based on unspecified factors, the identification problem may be understood by looking at the following simple example. For one data generating process of the bond yield $Y_n(t, t + \tau, X_t)$,

$$Y_n(t, t + \tau, X_t) = \frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X_t,$$

we can easily get an equivalent data generating process of the bond yield by applying a full-rank linear transformation \mathcal{L} on X_t and adjusting $B(\tau)$ correspondingly as shown in the equations

$$Y_n(t, t + \tau, X_t) = \frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X_t = \frac{A(\tau)}{\tau} + \frac{(\mathcal{L}^{-1\top} B(\tau))^\top}{\tau} \mathcal{L} X_t. \quad (4.12)$$

Now the *same* bond yield is based on a different set of factors $X_t^\mathcal{L} = \mathcal{L} X_t$ with a new coefficient $\mathcal{L}^{-1\top} B(\tau)$ and the dynamics of the transformed factor $X_t^\mathcal{L}$ satisfy the stochastic differential equation

$$dX_t^\mathcal{L} = \mathcal{L} dX_t = \mathcal{L} \mathcal{K} \mathcal{L}^{-1} (\mathcal{L} \theta - X_t^\mathcal{L}) dt + \mathcal{L} \Gamma dW_t^X,$$

which is different to the original factor dynamics (4.6). Due to the fact that the one bond yield model of $Y_n(t, T, X_t)$ can have different formulas related to different parameter representations, different parameter sets might give the same likelihood value so that the parameters cannot be determined uniquely through the maximum likelihood estimation. When running the numerical algorithm for the maximum likelihood estimation, multiple maxima might cause non-convergence of maximization process.

In order to solve this problem, we need to *normalize* the parameter space, that is, to restrict the parameter space so that for every data generating process there exists *only one* point in the restricted parameter space that corresponds to this data generating process.

Property 3 *Assume the following normalization conditions for the parameters $(\theta, \mathcal{K}, \Gamma, \lambda, \xi_0, \xi_1)$ appearing in the factor dynamics (4.6) and in the no-arbitrage conditions (4.10) and (4.11):*

- (i) \mathcal{K} in (4.6) is diagonal,
- (ii) θ in (4.6) is equal to $(0, \dots, 0)^\top$,
- (iii) ξ_1 in (4.10) is equal to $(1, \dots, 1)^\top$,
- (iv) Γ in (4.6) is lower triangular.

Then there is only one parameter representation corresponding to one given data generation process (4.12) for $Y_n(t, t + \tau, X_t)$ (up to permutations of the factors X_t).

We remark that we choose a different normalization conditions for the parameter space to those of Dai and Singleton (2000) and de Jong (2000). Instead of a *lower-triangular* \mathcal{K} and a *diagonal* Γ in the canonical representation in Dai and Singleton (2000) we have here a *diagonal* \mathcal{K} and a *lower-triangular* Γ . In fact, the two parameter representations are equivalent in the sense that for any given data generating process whose parameters follow the Dai and Singleton representation, we can find the corresponding parameter values following the representation given in Property 3 that lead to the same data generating process. Dai and Singleton (2000) characterize factors with orthogonal risk sources while we distinguish factors by their mean-reverting speeds. When solving the coefficients $A(\tau)$ and $B(\tau)$ and the intertemporal problem later, it turns out that calculations based on a diagonal \mathcal{K} are more convenient than those based on a lower-triangular \mathcal{K} . So, we employ the parameter representation characterized by Property 3.

Based on our parameter representation, we solve for the coefficients $B(\tau)$ and $A(\tau)$.

Property 4 Assume that all the normalization conditions in Property 3 are all satisfied. Let $\kappa_1, \dots, \kappa_n$ be the elements on the diagonal of \mathcal{K} . Then the coefficients $B(\tau) = (B_1(\tau), \dots, B_n(\tau))^\top$ and $A(\tau)$ satisfying the no-arbitrage conditions (4.10) and (4.11) are given by

$$B_i(\tau) = \frac{1}{\kappa_i}(1 - e^{-\kappa_i\tau}), \quad \forall i = 1, \dots, n \quad (4.13)$$

$$\begin{aligned} \frac{A(\tau)}{\tau} &= \sum_{i=1}^n \frac{\Gamma_i \lambda}{\kappa_i} \left(-1 + \frac{1 - e^{-\kappa_i\tau}}{\kappa_i\tau} \right) + \xi_0 \\ &\quad - \frac{1}{2} \sum_{i,j=1}^n \frac{\Gamma_i \Gamma_j^\top}{\kappa_i \kappa_j} \left(1 - \frac{1 - e^{-\kappa_i\tau}}{\kappa_i\tau} - \frac{1 - e^{-\kappa_j\tau}}{\kappa_j\tau} + \frac{1 - e^{-(\kappa_i + \kappa_j)\tau}}{(\kappa_i + \kappa_j)\tau} \right). \end{aligned} \quad (4.14)$$

4.3.3 Estimation of Underlying Factors

As an empirical model, the yield formula (4.3) cannot be satisfied exactly but only with measurement errors ϵ_t . The existence of the measurement errors might due to the nonobservability of the factor X_t , due to market frictions, or due to the imperfection of the model itself.

Using the model identification conditions given in Property 3, the identified empirical bond yield model is then given by

$$Y_n(t, t + \tau, X_t) = \frac{A(\tau)}{\tau} + \frac{1}{\tau} \sum_{i=1}^n B_i(\tau) X_{it} + \bar{\epsilon}_t, \quad (4.15)$$

where the measurement error $\bar{\epsilon}_t$ is assumed to be independent and $\mathcal{N}(0, \sigma_{\bar{\epsilon}})$ -distributed for all t and for all τ (times to maturity) of the observed nominal yields. Also, the measurement error $\bar{\epsilon}_t$ is assumed to be independently distributed of the factor process X_t .

We employ the Kalman filter⁴ to estimate the unobservable common factor X_t based on the nominal yield data. The *observation equation* in the Kalman filter is the estimation yield formula (4.15), where the coefficients $A(\tau)$ and $B(\tau)$ are given by (4.13) and (4.14). The *state equation* is the factor dynamics of X_t given by (4.6). In our implementation of the model estimation, we discretize the continuous-time dynamics (4.6) using the Euler-Maruyama method.

The task of the model calibration is to determine the parameters $(\mathcal{K}, \Gamma, \lambda, \xi_0)$ in order to fit the bond yield data as introduced in Section 4.2.

The inference of the unobservable factor X_t from observed bond yields $Y_n(t, t + \tau, X_t)$ is done by employing the *Kalman filter*. The parameter values are chosen by Maximum Likelihood Estimation. Our programs for the model estimation are based on the software packages "TSM" (Time Series Modelling) and "Optmum" (Opimization) supplied Programming Language "GAUSS"⁵.

Estimation

The model estimation is implemented for one-factor, two-factor and three-factor underlying dynamics (4.6). The estimation results are given in Table 4.2. The maximum likelihood estimation for the one- and two-factor models converge at the tolerance level of 10^{-5} for the gradient of the log-likelihood function while the convergence for the three-factor estimation cannot be achieved until we reduce the accuracy to 10^{-3} . We also note that for the three-factor estimation, the t-statistics for κ_i and Γ_{ij} become very large while the those for the λ become so small that all the three market prices

⁴See the Appendix on page 166.

⁵See the Homepage www.aptech.com.

	1-factor			
	Estimates	t-stat.		
κ_1	0.2792	64.45		
Γ_{11}	0.0089	58.63		
λ_1	-1.32	-5.73		
ξ_0	0.0153	2.22		
	2-factor		3-factor	
	Estimates	t-stat.	Estimates	t-stat.
κ_1	0.0307	9.99	0.0170	626.66
κ_2	0.2663	61.87	0.8122	441.15
κ_3			0.8605	809.44
Γ_{11}	0.0035	71.27	0.0079	285.47
Γ_{21}	-0.0019	-10.15	0.2152	743.00
Γ_{22}	0.0069	34.69	0.4524	1242.29
Γ_{31}			-0.2272	-666.50
Γ_{32}			-0.4509	-718.87
Γ_{33}			0.0038	231.29
λ_1	-0.4486	-4.15	-0.2159	-0.77
λ_2	-2.0456	-3.87	-0.6032	-0.53
λ_3			0.8900	0.10
ξ_0	0.0105	0.77	0.0279	0.20

Table 4.2: Estimated Parameters: One-, Two-/Three-factor Models

of risk are insignificant. We will come to discuss this finding later.

In Table 4.2 we see the mean-reversion parameter $\kappa_1 = 0.2792$ in the one-factor model, which corresponds to a half-life about two and half years. It is similar to the second mean-reversion parameter $\kappa_2 = 0.2663$ in the two-factor model. In the two-factor model the other trend is more persistent $\kappa_1 = 0.0307$, which corresponds to a half-life of 22.58 years. The market price of risk λ_1 in the one-factor model is -1.32 , meaning that, using the expression (4.9), the excess return $\mu_P - R_t$ is equal to $1.32(\Gamma_{11}B_1(\tau))$, which is 1.32 times the volatility $\Gamma_{11}B_1(\tau)$. For positive Γ_{ij} , the sign of the market price of risk λ_i is usually negative, corresponding to a positive excess return and in general, it depend on the signs of estimation results of Γ_{ij} on a case by case bearing.

In the estimation results for the three factor models, the volatility parameters $\Gamma_{21}, \Gamma_{22}, \Gamma_{31}$, and Γ_{32} are abnormally large when comparing with the bond yield volatility (standard deviation) given in Table 4.1. This leads to the fact that the corresponding t-statistics are also extremely large.

Recall that $\Gamma_j \Delta W_t^X$ denotes the factor innovation for the j -th factor, $j = 1, 2, 3$, where $\Gamma_j = (\Gamma_{j1}, \Gamma_{j2}, \Gamma_{j3})$. From Table 4.2 for the results of the three-factor model estimation, we can observe $\Gamma_2 = (0.2152, 0.4524, 0)$ is almost equal to $-\Gamma_3 = (0.2272, 0.4509, -0.0038)$. It turns out that the corresponding factor innovations $\Gamma_2 \Delta W_t^X$ and $\Gamma_3 \Delta W_t^X$ are almost (negatively) perfectly correlated (with correlation coefficient -0.9998). Also, these two factors have very similar mean-reversion parameters $\kappa_2 = 0.8122 \sim \kappa_3 = 0.8605$. It turns out that one factor X_{3t} is almost a mirror image of the other factor X_{2t} through the line $X = 0$. This result can also be observed in Figure 4.3, where these two factors fluctuate on an abnormally large scale so that the first factor X_{1t} on this scale appears to be like a horizontal line. We



Figure 4.3: Estimated Factors in the Three-Factor Model

mentioned earlier in Section 4.3.2 that we distinguish factors by their speeds mean-reversion. Since these two factors have similar mean-reversion parameters, we combine these two factors to one single factor. Mathematically, because of $\kappa_2 \sim \kappa_3$, we have $B_2(\tau) \sim B_3(\tau)$. Using this to rewrite the affine expression (4.15), the term $B_2(\tau)X_{2t} + B_3(\tau)X_{3t}$ can be approximated by

$$B_2(\tau)X_{2t} + B_3(\tau)X_{3t} \sim B_2(\tau)(X_{2t} + X_{3t}) .$$

This suggests considering $X_{2t} + X_{3t}$ as a new factor with the speed mean-reversion κ_2 .

In Figure 4.4 we plot the sum of these two factors ($X_{2t}^{(3)} + X_{3t}^{(3)}$) and the other estimated factors. In order to distinguish them we use a superscript to indicate the number of the factors in the model, for example, $X_{2t}^{(3)}$ denotes the second factor in the three-factor model. We observe that now all estimated factors are in the same range, instead of the wildly differing scales shown in Fig. 4.3. It is interesting to point out that the pair $(X_{1t}^{(3)}, X_{2t}^{(3)} + X_{3t}^{(3)})$ have similar trajectories to the two estimated factors of the two-factor model $(X_{1t}^{(2)}, X_{2t}^{(2)})$; The trajectory $X_{1t}^{(3)}$ in the three-factor model is similar to that of $X_{1t}^{(2)}$ in the two-factor model. Both correspond to the smaller mean-reversion parameters in each model respectively. Also, the evolutions of the second factor $X_{2t}^{(2)}$ in the two-factor model and the sum $X_{2t}^{(3)} + X_{3t}^{(3)}$ in the three-factor model are close to each other. On examining the estimated factor of the single-factor model, we can see this factor represents a kind of mixing of the two factors of the two-factor model. This is perhaps not surprising as this single factor has to somehow reflect the evolution of the two factors.

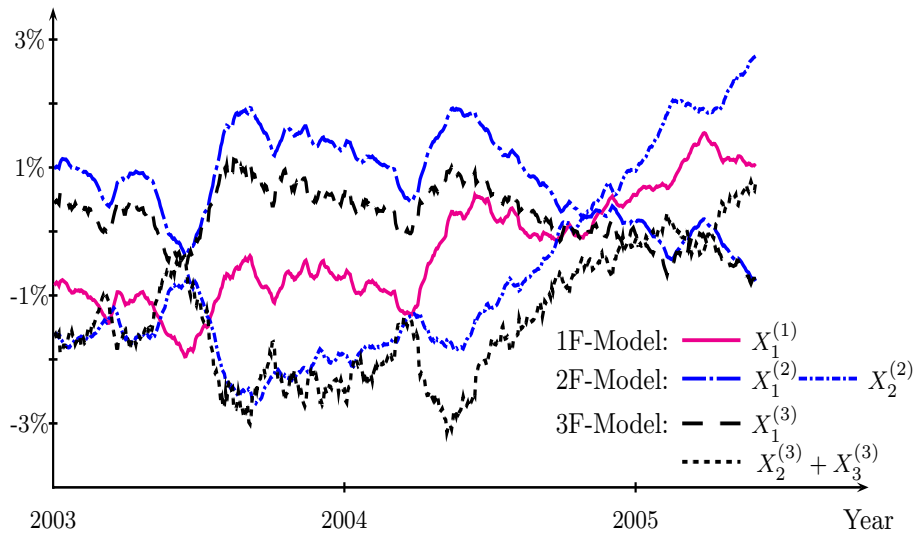


Figure 4.4: Estimated Factors, all Three Models

Model Selection

Here we compare the statistical performance of these three factor models. In

rel. fitting error	τ	1yr	3yr	5yr	10yr
1-Factor		31.4%	15.7%	33.4%	80.9%
2-Factor		17.4%	16.7%	15.8%	19.8%
3-Factor		2.2%	3.6%	5.9%	9.5%

Information criteria	AIC	BIC	HQIC
1-Factor	-9.370	-9.362	-9.367
2-Factor	-10.887	-10.873	-10.882
3-Factor	-12.853	-12.832	-12.845

Table 4.3: Comparison Estimation Behaviors

Table 4.3, the term “rel. fitting error” in the upper panel refers to the ratio of the standard deviations of the fitting errors with respect to the data. The lower panel provides the results of the three information criteria: the Akaike information criterion (AIC), the Bayesian information criterion (BIC), and the Hann-Quinn information criterion (HQIC)⁶. Intuitively, the value of an information criterion expresses an adjusted goodness of fit with a penalty for utilization of more degrees of freedom. The smaller the value, the better the model is.

Comparison of the results shows that the three-factor model is the best statistical model among the three factors. It has significantly the smallest fitting errors and the smallest values for all information criteria.

⁶Recall the definitions of the information criteria are given by

$$\begin{aligned} \text{AIC} &:= -\frac{2}{n} \ln(\text{Likelihood}) + 2\frac{k}{n}, \\ \text{BIC} &:= -\frac{2}{n} \ln(\text{Likelihood}) + 2\frac{k \ln(\ln n)}{n}, \\ \text{HQIC} &:= -\frac{2}{n} \ln(\text{Likelihood}) + k\frac{\ln n}{n}, \end{aligned}$$

where k is the number of the parameters and n is the number of observation data. The smaller the value, the better the model is. For the reference, see Akaike (1974), Schwarz (1978), and Hannan and Quinn (1979).

Concluding Remarks for the Estimation

As a conclusion to the estimation task one should ask the question: which model is the best model among the three models, one-, two-, or three-factor models?

The answer depends on the perspective that one chooses. On the one hand, the three-factor model performs best statistically in the task of modelling the nominal bond yields. On the other hand, however, the three-factor model has many undesirable aspects in providing a reasonable economic interpretation. Two of the filtered factors fluctuate on an abnormally large scale and they are almost perfectly negatively correlated. Furthermore, all estimated market prices of risk of the three-factor model are insignificant.

The two-factor model stands out when considering both the statistical performance and the resulting economic interpretations. Statistically, it also gives a reasonable fit. With regard to economic interpretation, the results show that it is indeed necessary to consider two factors which have different speeds of mean reversion. We argue further that the estimation results of the three-factor model does not necessarily suggest that we should consider a third factor and the two-factor model can actually embody the major part of the common movements of the bonds yields. Recall that the two estimated factors $X_{2t}^{(3)}$ and $X_{3t}^{(3)}$ in the three-factor model are almost perfectly correlated and they relate to very similar speeds of mean reversion. From our adopted perspective that the factors are distinguished with respect to the speeds of mean reversion we see that one of the two estimated factors is redundant. Furthermore, the sum $X_{2t}^{(3)} + X_{3t}^{(3)}$ gave a very similar factor trajectory to the one in the two-factor model.

Based on the discussions above, we have decided to choose the *two-factor model* as the best model for the modelling task.

4.4 Model II; Inflation-Indexed Bonds

In this section we construct a term structure model of inflation-indexed bonds (IIB) where there are specific factors: the instantaneous real interest rate and the expected inflation rate. The model is based on the no-arbitrage constraint of Jarrow and Yildirim (2003), from which we develop a two-factor term structure model that can also model inflation-indexed bonds.

The main idea of the construction of Jarrow and Yildirim (2003) is that they consider the “nominal world” and the “real world” as two countries and the price index as the “exchange rate” based on the no-arbitrage principle for the two-country model proposed by Amin and Jarrow (1991). Invoking an argument analogous to that in the two-country model that the no-arbitrage principle be satisfied on each national financial market, Jarrow and Yildirim obtain the no-arbitrage condition for the “nominal world”.

In this thesis we adopt the idea of Jarrow and Yildirim, however, we do not adopt the whole model directly. Their nominal term structure is based on a one-factor model, as in Munk et al. (2004). The shortcoming of such a one-factor model is two-fold; First, theoretically, if we consider the instantaneous nominal interest rate as the one factor, the model does not have the capacity to consider the inflation risk as the second factor, which we have argued is an important factor. Second, empirically, it is well known that a one-factor bond model does not provide satisfactory fitting to market data.

We extend the one-factor nominal bond model framework to that of a two-factor model of the type proposed by Richard (1978), where both the instantaneous real interest rate and the instantaneous expected inflation rate are factors for the nominal term structure.

4.4.1 The Model

Following Richard (1978) we assume that the instantaneous real interest rate r_t and the anticipated instantaneous inflation rate π_t are the two factors driving the nominal bond price. The two factors are assumed to follow the Gaussian mean-reverting process

$$dr_t = \kappa_r(\bar{r} - r_t)dt + g_r dW_t^r, \quad (4.16)$$

and

$$d\pi_t = \kappa_\pi(\bar{\pi} - \pi_t)dt + g_\pi dW_t^\pi, \quad (4.17)$$

where W_t^r and W_t^π are correlated Wiener processes with the instantaneous variance $dW_t^r dW_t^\pi = \rho_\pi dt$.

Let $P_n(t, T, r_t, \pi_t)$ be the nominal bond price at t , maturing at T , and depending on the current factor states r_t and π_t . With the assumptions (4.16) and (4.17), the nominal bond price can be modelled by the Duffie-Kan exponential affine formula, that is,)

$$P_n(t, T, r_t, \pi_t) = \exp(-A_n(T-t) - B_{nr}(T-t)r_t - B_{n\pi}(T-t)\pi_t). \quad (4.18)$$

The Duffie-Kan coefficients $A_n(\tau)$, $B_{nr}(\tau)$ and $B_{n\pi}(\tau)$ will be determined later by the no-arbitrage condition (4.31).

Applying Itô's Lemma to (4.18), we can write the return of the nominal bond as

$$\frac{dP_n(t, T, r_t, \pi_t)}{P_n(t, T, r_t, \pi_t)} = \mu_n(t, T - t)dt - B_{nr}(T - t)g_r dW_t^r - B_{n\pi}(T - t)g_\pi dW_t^\pi, \quad (4.19)$$

where

$$\begin{aligned} \mu_n(t, \tau) = & \frac{d}{d\tau}A_n(\tau) + \frac{d}{d\tau}B_{nr}(\tau)r_t + \frac{d}{d\tau}B_{n\pi}(\tau)\pi_t \\ & - B_{nr}(\tau)\kappa_r(\bar{r} - r_t) - B_{n\pi}(\tau)\kappa_\pi(\bar{\pi} - \pi_t) \\ & + \frac{1}{2} \left(B_{nr}(\tau)^2 g_r^2 + 2B_{nr}(\tau)B_{n\pi}(\tau)\sigma_r\sigma_\pi\rho_{r\pi} + B_{n\pi}(\tau)^2 g_\pi^2 \right). \end{aligned} \quad (4.20)$$

The nominal yield is defined by

$$Y_n(t, T, r_t, \pi_t) := \frac{-\ln P_n(t, T)}{T - t} = \frac{A_n(T - t)}{T - t} + \frac{B_{nr}(T - t)}{T - t}r_t + \frac{B_{n\pi}(T - t)}{T - t}\pi_t. \quad (4.21)$$

Let $P_I(t, T)$ denote the price of the (zero-coupon) **inflation-indexed bond** (IIB) that is issued at time 0^7 and matures at time T . The payout at the maturity date will be adjusted by the price index I_T so that

$$P_I(T, T) = I_T. \quad (4.22)$$

Define the *real bond* $P_r(t, T) := P_I(t, T)/I_t$ as the normalized IIB with respect to the corresponding price index. According to (4.22), we have $P_r(T, T) = 1$. In other words, the real bond has a payout of one unit of consumption good at T . We assume that the real bond is affected only by one factor, the instantaneous real interest rate r_t , which also follows the Duffie-Kan type dynamics so that

$$P_r(t, T) = \exp \left(-A_r(T - t) - B_{rr}(T - t)r_t \right), \quad (4.23)$$

where the Duffie-Kan coefficients $A_r(\tau)$ and $B_{rr}(\tau)$ will be determined later by the no-arbitrage conditions (4.32) and (4.33). The assumption (4.23) concerning the real bond implies the dynamics of the IIB $P_I(t, T)$ that will be showed later.

⁷We fix $I_0 = 1$

The real yield is defined as

$$Y_r(t, T) := \frac{-\ln P_r(t, T)}{T-t} = \frac{A_r(T-t)}{T-t} + \frac{B_{rr}(T-t)}{T-t} r_t. \quad (4.24)$$

We denote a *consumption good account* $M_r(t)$ as

$$M_r(t) := \exp\left(\int_0^t r_s ds\right),$$

and $M_I(t)$ as the *real money account*, which gives the nominal value of the consumption good account, that is,

$$M_I(t) := M_r(t) I_t. \quad (4.25)$$

To calculate return of the IIB, we apply Itô's Lemma at first to the real bond price (4.23) and obtain

$$\frac{dP_r(t, T, r_t)}{P_r(t, T, r_t)} = \mu_r(t, T-t) dt - B_{rr}(T-t) g_r dW_t^r, \quad (4.26)$$

where

$$\mu_r(t, \tau) := \frac{d}{d\tau} A_r(\tau) + \frac{d}{d\tau} B_{rr}(\tau) r_t - B_{rr}(\tau) \kappa_r (\bar{r} - r_t) + \frac{1}{2} B_{rr}(\tau)^2 g_r^2. \quad (4.27)$$

Next we apply Itô's Lemma to the expression for the IIB,

$$P_I(t, T, r_t, I_t) = P_r(t, T, r_t) I_t,$$

and recall the price index I_t follows the dynamics (3.1), so that we then have the return process of the IIB, namely

$$\frac{dP_I(t, T, r_t, I_t)}{P_I(t, T, r_t, I_t)} = \mu_I(t, T-t) dt - B_{rr}(T-t) g_r dW_t^r + \sigma_I dW_t^I, \quad (4.28)$$

where

$$\mu_I(t, T-t) := \mu_r(t, T-t) + \pi_t - B_{rr}(T-t) g_r \sigma_I \rho_{Ir}, \quad (4.29)$$

with $\rho_{Ir} dt = dW_t^r dW_t^I$.

Applying Itô's Lemma to (4.25), we find that the return on the real money account $M_I(t)$ is given by

$$\frac{dM_I(t)}{M_I(t)} = (r_t + \pi_t) dt + \sigma_I dW_t^I. \quad (4.30)$$

We employ the assumption of the no-arbitrage principle Assumption 3 represented by equation (3.4) in Section 3.1. It requires that the nominal bond return (4.19), the IIB return (4.28) and the return of the real money account (4.30) satisfy the no-arbitrage equalities

$$\mu_n(t, \tau) - R_t = -B_{nr}(\tau)g_r\lambda_r - B_{n\pi}(\tau)g_\pi\lambda_\pi, \quad \forall \tau > 0 \quad (4.31)$$

$$\mu_I(t, \tau) - R_t = -B_{rr}(\tau)g_r\lambda_r + \lambda_I\sigma_I, \quad \forall \tau > 0 \quad (4.32)$$

$$\pi_t + r_t - R_t = \lambda_I\sigma_I, \quad (4.33)$$

where the market prices of risk λ_r , λ_π , and λ_I are constants. Since the three conditions above belong to special cases of the no-arbitrage condition (3.4), they can be obtained by the standard hedging argument used to obtain the general condition (3.4); see Chiarella (2004).

Property 5 *From the relations (4.32) and (4.33), the excess return for the real bond can be derived as*

$$\mu_r(t, \tau) - r_t = -B_{rr}(\tau)g_r(\lambda_r - \sigma_I\rho_{Ir}). \quad (4.34)$$

The relation (4.34) might be interpreted as the excess return of the real bond can be explained by the risk premium on the RHS with the adjusted market price of risk $\lambda_r - \sigma_I\rho_{Ir}$. There is an adjustment term $-\sigma_I\rho_{Ir}$ in the market price of risk because the real bonds are not directly tradable and their returns are calculated through the returns of the inflation-indexed bonds.

Property 6 *If the no-arbitrage equalities (4.31) – (4.33) are satisfied, then (i) the coefficients $A_n(\tau)$, $B_{nr}(\tau)$, $B_{n\pi}(\tau)$ in the expression (4.18) for the nominal bond price have the form*

$$B_{nr}(\tau) = \frac{1}{\kappa_r}(1 - e^{-\kappa_r\tau}), \quad (4.35)$$

$$B_{n\pi}(\tau) = \frac{1}{\kappa_\pi}(1 - e^{-\kappa_\pi\tau}), \quad (4.36)$$

$$\begin{aligned} \frac{A_n(\tau)}{\tau} &= \left(1 - \frac{1}{\tau\kappa_r} + \frac{e^{-\tau\kappa_r}}{\tau\kappa_r}\right)\left(\bar{r} - \frac{g_r\lambda_r}{\kappa_r}\right) + \left(1 - \frac{1}{\tau\kappa_\pi} + \frac{e^{-\tau\kappa_\pi}}{\tau\kappa_\pi}\right)\left(\bar{\pi} - \frac{g_\pi\lambda_\pi}{\kappa_\pi}\right) \\ &\quad - \frac{g_r^2}{2\kappa_r^2}\left(1 - 2\frac{1 - e^{-\kappa_r\tau}}{\kappa_r\tau} + \frac{1 - e^{-2\kappa_r\tau}}{2\kappa_r\tau}\right) - \frac{g_\pi^2}{2\kappa_\pi^2}\left(1 - 2\frac{1 - e^{-\kappa_\pi\tau}}{\kappa_\pi\tau} + \frac{1 - e^{-2\kappa_\pi\tau}}{2\kappa_\pi\tau}\right) \\ &\quad - \frac{g_r g_\pi \rho_{r\pi}}{\kappa_r \kappa_\pi} \left(1 - \frac{1 - e^{-\kappa_r\tau}}{\kappa_r\tau} - \frac{1 - e^{-\kappa_\pi\tau}}{\kappa_\pi\tau} + \frac{1 - e^{-(\kappa_r + \kappa_\pi)\tau}}{(\kappa_r + \kappa_\pi)\tau}\right) + \xi_0. \end{aligned} \quad (4.37)$$

(ii) The coefficients $A_r(\tau), B_r(\tau)$ in the expression (4.23) for the real yield have the form

$$B_{rr}(\tau) = \frac{1}{\kappa_r}(1 - e^{-\kappa_r\tau}) \quad (4.38)$$

$$\begin{aligned} \frac{A_r(\tau)}{\tau} = & \left(1 - \frac{1}{\tau\kappa_r} + \frac{e^{-\tau\kappa_r}}{\tau\kappa_r}\right)(\bar{r} - g_r \frac{\lambda_r - \sigma_I \rho I_r}{\kappa_r}) \\ & - \frac{g_r^2}{2\kappa_r^2} \left(1 - 2\frac{1 - e^{-\kappa_r\tau}}{\kappa_r\tau} + \frac{1 - e^{-2\kappa_r\tau}}{2\kappa_r\tau}\right). \end{aligned} \quad (4.39)$$

Property 7 If the no-arbitrage equalities (4.31) – (4.33) are satisfied, then (i) the instantaneous nominal interest rate is given by

$$R_t = \xi_0 + r_t + \pi_t. \quad (4.40)$$

(ii) When the IIBs are included in the investment set, then we have

$$\xi_0 = -\lambda_I \sigma_I. \quad (4.41)$$

4.4.2 Model Estimation

The model estimation in this subsection has three tasks. The first one is to estimate the parameters that are required to implement the optimal intertemporal portfolio rules described in Chapter 3. The second task is to use the Kalman filter to estimate the instantaneous real interest rate and the instantaneous expected inflation rate that are not directly observed, but are reflected implicitly in the evolution of the real and nominal term structures. The third task is to provide a validation check of the estimation results where the fitting errors of the market data should be small and the estimation results should be economically reasonable.

The U.S. Treasury provides daily data of real bond yields from 2003. These data allow us to estimate the term structure in a new way. We can estimate the instantaneous real interest rate directly from the market real yield data, whereas the conventional way of estimating the real interest rate would require us to first estimate the expected rate of inflation. Once the real interest rate has been estimated, we can utilize nominal bond yield data, which are considered to bear inflation risk, to estimate the expected rate of inflation. This new estimation procedure has the advantage that although our nominal term structure has two unobservable state variable, r_t and π_t , we can still identify them and estimate them through the market data.

The Term Structure of Real Yields

The data used for estimating the term structure of real yields are the real yields data described in Section 4.2. We employ the Kalman filter to filter out the unobservable factor X_t from the US data for the real yields. In the implementation of the Kalman filter⁸, the *observation equation* is the real yield formula (4.24) with measurement errors. Recall that the Duffie-Kan coefficients $A_r(\tau)$ and $B_{rr}(\tau)$ are given by (4.39) and (4.38), Thus, the observation equation here is given by

$$Y_r(t, t + \tau, r_t) = \frac{A_r(\tau)}{\tau} + \frac{B_{rr}(\tau)}{\tau} r_t + \epsilon_t^\tau. \quad (4.42)$$

The state equation here is the discretized factor dynamics of r_t (4.16) obtained by using the Euler-Maruyama scheme. The discretized process should very be close to the continuous-time process because the discretization interval is 0.004, corresponding to one day.

The results of the parameter estimation are given in Table 4.4 and the filtered time series for r_t is plotted in Fig. 4.5.

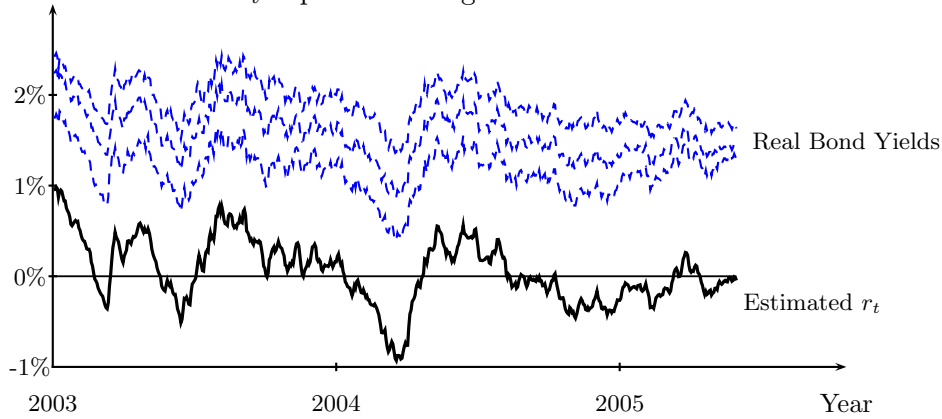


Figure 4.5: Time Series of Real Yields and the Estimated Real Rate

The average measurement errors σ_{ϵ^τ} for each real yields are given in the last row of Table 4.4. Compared with the standard deviations of the real yields above, the model can explain around 70% variation of the real yields⁹.

⁸See Appendix on page 166.

⁹The values for $\frac{\sigma_{\epsilon^\tau}}{\text{SD}}$ in Table 4.4 represents the unexplained fraction in the total variation of the real yields.

Log Likelihood = 10056.45

	Estimate	t-stat.
κ_r	0.1248	7.31
\bar{r}	0.0040	0.02
g_r	0.0101	27.51
λ_r^*	-0.5161	-0.22
$\sigma_{\mathcal{E}^r}$	0.0008	49.84

$$\lambda_r^* = \lambda_r - \sigma_I \rho_{I r}$$

τ	5Y	7Y	10Y
Mean	1.16%	1.56%	1.90%
SD	0.25%	0.26%	0.23%
$\frac{A_r(\tau)}{\tau}$	1.14%	1.48%	1.89%
$\frac{B_{rr}(\tau)}{\tau}$ (Sensitivity)	74%	67%	57%
$\hat{\sigma}_{\mathcal{E}^r}$	8.63e-4	5.83e-4	7.76e-4
$\hat{\sigma}_{\mathcal{E}^r}/SD$	35.01%	22.52%	33.23%

Table 4.4: Upper Panel: estimated parameters for the real yield formula and Lower Panel: statistics, fitting errors, and price sensitivities

We find that these fitting results are quite satisfactory. The horizontal line indicates the estimated mean of \bar{r} . We take the value 0 for this parameter because the result of the t -statistic in Table 4.4 does not suggest that the estimate should differ from zero.

With regard to the estimation result for the market price of the real interest risk λ_r , we note that from the real yield data we cannot determine the market price of risk λ_r directly, but only the adjusted market price $\lambda^* = \lambda - \sigma_I \rho_{I r}$. The reason for this has been given in the discussion after Property 5.

An Investigation on the double roles of κ_r

The parameter κ_r is related to two features in the real bond model. The first feature relates to the fact that this parameter can represent the speed of the mean-reversion of the factor r_t as represented in the dynamics (4.16). The higher this parameter is the faster the factor r_t comes back to its mean \bar{r} and also the more frequently the factor crosses the mean.

The other feature is the real yield sensitivity with respect to the change of the factor r_t as formulated in the real yield formula (4.42) where you can see one unit change of r_t leads to a $\frac{B_{rr}(\tau)}{\tau} (= \frac{1-e^{-\kappa_r\tau}}{\tau})$ unit change of the bond yield $Y_r(t, t + \tau, r_t)$. Based on our estimation result, one unit change of r_t leads to a change of the 5-year real yield by 74% of a unit as shown in in the lower panel in Table 4.4.

With regard to the first feature of the mean-reversion of the parameter κ_r , the half-life of a mean-reversion level κ_r is $(\ln 2)/\kappa_r$. Taking our estimation result of $\kappa_r = 0.1248$ in Table 4.4 as example, the half-life is around 5.55 years. Comparing this fact with the impression given in Fig. 4.5, where the trajectory of the estimated r_t crosses the mean already over ten times during the observation period of two and half year, we find that the estimated value $\kappa_r = 0.1248$ in Table 4.4 is too small for the mean-reversion behavior. In order to check this we implement the ordinary linear regression on the estimated r_t obtained by the Kalman filtering with the discretized relation

$$r_{t+\Delta} - r_t = \kappa_r(\bar{r} - r_t) + g_r\Delta W_t, \quad (4.43)$$

where $\Delta = 1/250$. The estimation results of this ordinary linear regression are given in Table 4.5. There we see that the estimate for the mean-reversion parameter $\kappa_r = 3.9625$ is much higher than the estimate 0.1248 in Table 4.4. The half-life corresponding to the estimated parameter by the linear regression is 0.1749 year, which is about two months. This result is more consistent with the mean-reversion behavior shown in Figure 4.5. Later in Section 4.4.3 we will try to modify the real bond model at this point in order to accommodate the two features of κ_r more properly.

	Estimate	t-stat.
κ_r	3.9625	2.68
\bar{r}	-0.0005	-0.36

Table 4.5: Ordinary Linear Regression on filtered r_t

The Term Structure of Nominal Yields

Here we will estimate the nominal bond yield model for $Y_n(t, t + \tau, r_t, \pi_t)$ given in equation (4.21). The data we use for the estimation are the US nominal and real yields data described in Section 4.2.

The observation equation is based on the yield formula (4.21) and with the measurement error $\tilde{\epsilon}_t$

$$Y_n(t, t + \tau, r_t, \pi_t) = \frac{A_n(\tau)}{\tau} + \frac{B_{nr}(\tau)}{\tau} r_t + \frac{B_{n\pi}(\tau)}{\tau} \pi_t + \tilde{\epsilon}_t^\tau, \quad (4.44)$$

where $A_n(\tau)$ is given by (4.37), $B_{nr}(\tau)$ and $B_{n\pi}(\tau)$ are given by (4.35) and (4.36), and the measurement errors $\tilde{\epsilon}_t^\tau$ are identically and independently distributed for all t and τ .

For the real interest rate r_t in equation (4.44) we use the previous estimated results presented in Fig. 4.5 where we assume that the investors both on the nominal bond market and on the IIB market share the same belief in the instantaneous real rate. The instantaneous inflation expectation π_t , however, is treated as unknown and will be estimated by using the Kalman filter. So, the state equation for the implementation the Kalman filter is the discretized dynamics of the expected inflation rate π_t given in (4.17) by the Euler-Maruyama Scheme. The parameters are determined by using the maximum likelihood method.¹⁰

Before the model estimation is undertaken it is necessary to point out that we still encounter the model identification problem. The model identification problem arises due to the fact that there is a specific relation between the two parameters, $\bar{\pi}$ in the expression (4.17) and the ξ_0 in the coefficient $A_n(\tau)$ given by (4.37). The discussion for the model identification problem is one special case of the general model identification conditions given in Property 3. The setting of $\bar{\pi} = 0$ is just one possible parameter representation. In fact, we can set $\bar{\pi}$ to any arbitrary level and can still identify the parameters by the data. For any given level $\bar{\pi}$, we denote the parameter ξ_0 in equation (4.37) by $\xi_0^{\bar{\pi}}$ in order to stress its relation to $\bar{\pi}$. Any parameter pair $(\bar{\pi}, \xi_0^{\bar{\pi}})$ satisfying the relation (4.45) will generate equivalent bond yield models (4.21) for any arbitrary $\bar{\pi}$.

Property 8 *Let $\xi_0^{\bar{\pi}}$ denote the parameter in equation (4.37) for a given level of $\bar{\pi}$ in the dynamics (4.17).*

If the relation

$$\xi_0^{\bar{\pi}} = \xi_0^0 - \bar{\pi} \quad (4.45)$$

¹⁰An alternative estimation procedure will be estimate the two factors simultaneously which is left to future research.

is satisfied, then the nominal yield model with the parameter pairs $(\xi_0, 0)$ and $(\xi_0^{\bar{\pi}}, \bar{\pi})$ will generate the same bond yield dynamics (4.21) for any arbitrary $\bar{\pi}$.

The flexibility to change $\bar{\pi}$ to an arbitrary level will lead to a problem of determining the market price of risk λ_f as we will see later in Section 4.4.5.

Another related remark is that in contrast to the flexibility of varying $\bar{\pi}$, the mean \bar{r} for the factor r_t appearing in the estimation of the real yield model in the first part of this subsection cannot be shifted arbitrarily. The reason is that in the real yield formula (4.42), the change of the level \bar{r} cannot be absorbed into the coefficient $A_r(\tau)$ given in (4.39) because there is no equivalent to the free parameter as ξ_0 in (4.39) to adjust the yield level.

The estimation for the correlation coefficient $\rho_{r\pi}$ between the real interest rate shock W_t^r and the expected inflation shock W_t^π requires an iterative estimation scheme due to the following fact. In equation (4.37) $\rho_{r\pi}$ is a parameter to be determined through the maximum likelihood estimation method. However, after $\rho_{r\pi}$ and all the other parameters have been estimated, we can calculate the *sample correlation coefficient* based on the estimated residuals of (4.16) and (4.17), that is

$$\begin{aligned}\Delta \hat{W}_t^r &= \frac{1}{g_r} (\Delta r_t - \kappa_r (\bar{r} - r_{t-\Delta}) \Delta), \\ \Delta \hat{W}_t^\pi &= \frac{1}{g_\pi} (\Delta \pi_t - \kappa_\pi (\bar{\pi} - \pi_{t-\Delta}) \Delta),\end{aligned}$$

and

$$\hat{\rho}_{r\pi} := \mathbf{E}[\Delta \hat{W}_t^r \Delta \hat{W}_t^\pi] / \Delta. \quad (4.46)$$

where $\kappa_r, \bar{r}, \kappa_\pi$ take values of the estimation results. These two estimates for $\rho_{r\pi}$, have to be consistent with each other. However, this is not usually the case. To plug the gap in this inconsistency of the estimation of $\rho_{r\pi}$, we implement an iterative estimation scheme: in the first step we fix $\rho_{r\pi}$ to be a value $\rho_{r\pi}^{(1)}$, say, 0, and estimate all other parameters by the maximum likelihood method and then calculate the estimated sample correlation $\hat{\rho}_{r\pi}^{(1)}$ as given in (4.46). Next, we compare $\rho_{r\pi}^{(1)}$ and $\hat{\rho}_{r\pi}^{(1)}$, if they are close to each other, we stop the iteration scheme, otherwise we set the initial value $\rho_{r\pi}^{(2)} = \hat{\rho}_{r\pi}^{(1)}$ for the second step and repeat entirely the above process. Under the assumption that the model begin estimated is the true model and the maximum likelihood estimator is consistent, this iterative scheme provides

a consistent estimator.

We implement the above iterative scheme with the initial correlation coefficient $\rho_{r\pi}^{(1)} = 0$. The sample correlation coefficient for the first iteration step is calculated as $\hat{\rho}_{r\pi}^{(1)} = -0.5476$. Taking this value as the correlation coefficient for the second step, the sample correlation coefficient is then calculated as $\hat{\rho}_{r\pi}^{(2)} = -0.5250$. We judge that these two values are close enough and stopped the iterative scheme at the second step.

The results of the parameter estimation are summarized in Table 4.6. The mean-reversion parameter $\kappa_\pi = 0.4016$ implies the estimated π_t with the dynamics (4.17) is a stationary process. The estimate corresponds to half-life around one and three quarter years (1.73 years). The π_t -sensitivity based on the estimated value is listed with different time to maturity in the lower panel in Table 4.6. It decreases with the time to maturity. The development of the nominal term structure, which is characterized by the decreasing term premia (the yield spread), can be explained mathematically by the increasing level of $A_n(\tau)/\tau$ and the decreasing sensitivity of $B_{n\pi}(\tau)/\tau$ in the yield formula 4.44. The corresponding values for $A_n(\tau)/\tau$ and $B_{n\pi}(\tau)/\tau$ can be obtained in Table 4.6.

In the lower panel in Table 4.6 we list the estimate for the scale of the measurement error $\hat{\sigma}_\varepsilon$ for each bond and its relative fitting error $\hat{\sigma}_\varepsilon/\text{SD}$. Figures 4.6 and 4.7 plot the estimated and market nominal yields for one year and ten years maturity respectively. Although the short-term yields are fitted satisfactorily, there is still room for improvement in the fitting of the long-term yields.

Log-Likelihood = 27479.20

	Estimates		t-stat.		
κ_π	0.4016		34.44		
g_π	0.0067		65.00		
λ_π	-1.5680		-22.36		
ξ_0	-0.0012		-4.53		
σ_{ϵ^r}	0.0025		58.87		
$\rho_{r\pi}$	-0.5476				

τ	1M	3M	6M	1Y	2Y
Mean	1.37%	1.47%	1.63%	1.84%	2.26%
SD	0.59%	0.66%	0.75%	0.75%	0.75%
$\frac{A(\tau)}{\tau}$	-0.05%	0.08%	0.27%	0.62%	1.23%
$\frac{B_{n\pi}(\tau)}{\tau}$ (Sensitivity)	98.34%	95.14%	90.60%	82.36%	68.74%
$\hat{\sigma}_{\epsilon^r}$	0.31%	0.21%	0.13%	0.13%	0.24%
$\hat{\sigma}_{\epsilon^r}/\text{SD}$	51.79%	31.45%	17.94%	17.08%	31.73%

τ	3Y	5Y	7Y	10Y	20Y
Mean	2.65%	3.31%	3.38%	4.17%	4.95%
SD	0.65%	0.47%	0.37%	0.32%	0.29%
$\frac{A(\tau)}{\tau}$	1.73%	2.52%	3.12%	3.77%	4.95%
$\frac{B_{n\pi}(\tau)}{\tau}$ (Sensitivity)	58.12%	43.11%	33.43%	24.45%	12.44%
$\hat{\sigma}_{\epsilon^r}$	0.26%	0.27%	0.26%	0.25%	0.34%
$\hat{\sigma}_{\epsilon^r}/\text{SD}$	39.80%	57.98%	71.86%	78.07%	117.34%

Table 4.6: Upper Panel: estimated parameters for nominal term structure; Lower Panel: statistics, fitting errors, and yield sensitivity

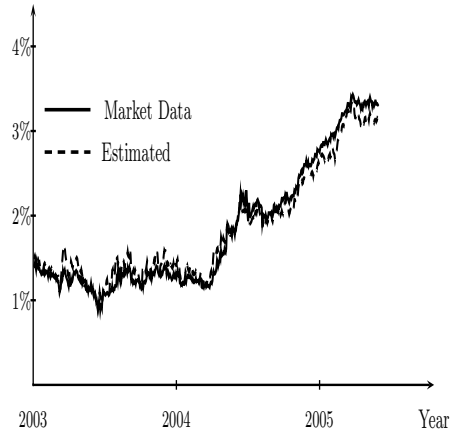


Figure 4.6: Market and Estimated 1Y bond yield



Figure 4.7: Market and Estimated 10Y bond yield

The estimated factors are plotted together with the nominal yields in Fig. 4.8 where the blue dashed curve is the estimated π_t and the black solid curve is the estimated r_t . During the filtering period the factor r_t fluctuates around its mean while the factor π_t displays its increasing trend. The increasing trajectory of the estimated factor π_t suggests that π_t is not a stationary process. We examine the mean-reversion of the π_t by implementing the OLS in a regression

$$\pi_{t+\Delta} - \pi_t = c_{\bar{\pi}} - \kappa_{\pi}\pi_t\Delta + g_{\pi}\Delta W_t$$

The OLS results are given in Table 4.7. It shows that the mean-reversion parameter κ_{π} takes a negative value and is not significantly different from zero. This result indicates that the filtered factor π_t should be treated as either a random walk or a nonstationary process. This result is considered with what is seen in Table 4.8 but is still different from the estimate $\kappa_{\pi} = 0.4016$ given in Table 4.6. Similarly to the parameter κ_r , the parameter κ_{π} is also related to two features, namely, the sensitivity of bond yields to the factors and the mean-reversion of the factor process. The inconsistent results for κ_{π} in Tables 4.6 and 4.7 indicate the necessity to modify the model as was suggested previously by the inconsistent results for κ_r .

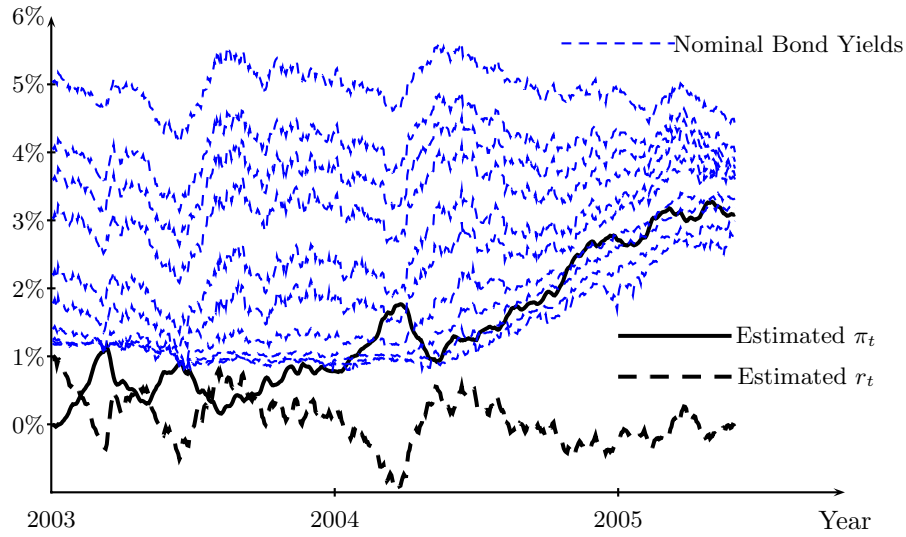


Figure 4.8: Nominal Yields and Estimated Factors

	Estimates	t-stat
κ_{π}	-0.0880	-0.40
c_{π}	4.6×10^{-5}	2.61

Table 4.7: OLS results for the filtered π_t

To give the estimation results shown in Figure 4.8 an economic interpretation, we recall that the Federal Reserve Bank kept on increasing the target Federal Fund Rate during this time and the market reacted to this policy of tightening with increasing short-term yields but stationary long-term yields. One possible interpretation is that the increase of the short-term yields, although they are driven by the Fed's open market operations, reflected mainly a disclosure of the potential for high inflation but not for an increasing long-run real interest rate. The results of the estimation here tell the same story where the increase of the short-term yields is mainly contributed by the increase of the expected inflation rate but not by the trend of the real interest rate.

As a validation check for the model estimation, we compare the instantaneous nominal interest rate given by the formula (4.40) based on the results

of estimation, and the corresponding market interest rates, the effective Federal Funds rate, which is an overnight interbank rate and is not included in the model estimation. The two rates are compared in Fig. 4.9 where we found the fit on the whole is quite reasonable and more satisfactory after the fourth Quarter 2003.

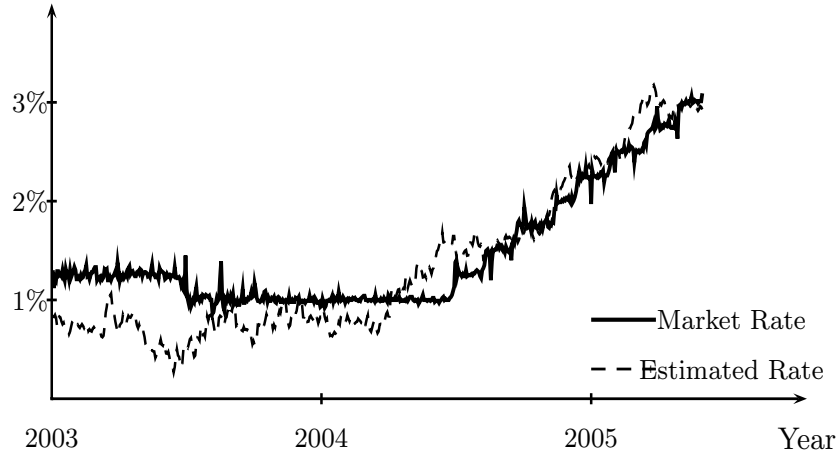


Figure 4.9: Federal Fund Rate and the Estimated Rate

4.4.3 A Modification of the Model

We pointed out in Section 4.4.2 that the term structure model does not have sufficient flexibility to allow the parameters κ_r and κ_π to fulfill their dual roles of mean-reversion and yield sensitivity parameters. The results of the estimation have indeed confirmed this inflexibility. In order to overcome this difficulty we propose a modification to the model.

The idea of the modification is to use two parameters to represent the two features respectively. We retain κ_r for the mean-reversion parameter and use a separate parameter for the sensitivity factor that we denote by κ_{rr} .

The impact of the parameter separation on the term structure model will be discussed in two parts in this subsection. In the first part we provide a modification of the model according to the parameter separation. In the second part we argue that the no-arbitrage principle is satisfied in a broader sense.

Due to the parameter separation, the coefficient $B_{rr}(\tau)$ in the formula (4.38) becomes

$$B_{rr}(\tau) = (1 - e^{\kappa_{rr}\tau})/\kappa_{rr} . \quad (4.47)$$

We apply this parameter separation also to the coefficients $B_{nr}(\tau)$ and $B_{n\pi}(\tau)$ then these two coefficients become

$$B_{nr}(\tau) = (1 - e^{\kappa_{nr}\tau})/\kappa_{nr} \quad (4.48)$$

$$B_{n\pi}(\tau) = (1 - e^{\kappa_{n\pi}\tau})/\kappa_{n\pi} , \quad (4.49)$$

where κ_{nr} and $\kappa_{n\pi}$ are the sensitivity parameters separated from the original parameters κ_π and κ_r respectively, which are retained for the mean-reversion parameters.

We adjust the coefficients $A_r(\tau)$ and $A_n(\tau)$ into

$$\begin{aligned} \frac{A_r(\tau)}{\tau} &= \left(1 - \frac{1}{\tau\kappa_{rr}} + \frac{e^{-\tau\kappa_{rr}}}{\tau\kappa_{rr}}\right) \left(\frac{\bar{r}\kappa_r - g_r(\lambda_r - \sigma_I\rho_{Ir})}{\kappa_{rr}}\right) \\ &\quad - \frac{g_r^2}{2\kappa_{rr}^2} \left(1 - 2\frac{1 - e^{-\kappa_{rr}\tau}}{\kappa_{rr}\tau} + \frac{1 - e^{-2\kappa_{rr}\tau}}{2\kappa_{rr}\tau}\right) \\ \frac{A_n(\tau)}{\tau} &= \left(1 - \frac{1}{\tau\kappa_{nr}} + \frac{e^{-\tau\kappa_{nr}}}{\tau\kappa_{nr}}\right) \frac{\bar{r}\kappa_r - g_r\lambda_r}{\kappa_{nr}} + \left(1 - \frac{1}{\tau\kappa_{n\pi}} + \frac{e^{-\tau\kappa_{n\pi}}}{\tau\kappa_{n\pi}}\right) \frac{\bar{\pi}\kappa_\pi - g_\pi\lambda_\pi}{\kappa_{n\pi}} \\ &\quad - \frac{g_r^2}{2\kappa_{nr}^2} \left(1 - 2\frac{1 - e^{-\kappa_{nr}\tau}}{\kappa_{nr}\tau} + \frac{1 - e^{-2\kappa_{nr}\tau}}{2\kappa_{nr}\tau}\right) - \frac{g_\pi^2}{2\kappa_{n\pi}^2} \left(1 - 2\frac{1 - e^{-\kappa_{n\pi}\tau}}{\kappa_{n\pi}\tau} + \frac{1 - e^{-2\kappa_{n\pi}\tau}}{2\kappa_{n\pi}\tau}\right) \\ &\quad - \frac{\sigma_{r\pi}}{\kappa_{nr}\kappa_{n\pi}} \left(1 - \frac{1 - e^{-\kappa_{nr}\tau}}{\kappa_{nr}\tau} - \frac{1 - e^{-\kappa_{n\pi}\tau}}{\kappa_{n\pi}\tau} + \frac{1 - e^{-(\kappa_{nr} + \kappa_{n\pi})\tau}}{(\kappa_{nr} + \kappa_{n\pi})\tau}\right) + \xi_0 . \end{aligned} \quad (4.51)$$

The adjusted expression (4.50) is obtained by replacing κ_r by κ_{rr} everywhere except in the term multiplying with \bar{r} in the second row of (4.50). The adjusted expression (4.51) is obtained in a similar way where κ_π and κ_r are replaced by κ_{nr} and $\kappa_{n\pi}$ respectively everywhere except in the term multiplying the \bar{r} and $\bar{\pi}$.

We discuss now how our central assumption, the no-arbitrage condition, is affected by these adjustments.

Property 9 *As a result of the adjustments (4.47), (4.48), (4.49), (4.50),*

and (4.51), the conditions (4.31) and (4.32) change into

$$\begin{aligned} \mu_n(t, \tau) - R_t &= -B_{nr}(\tau)g_r\lambda_r - B_{n\pi}(\tau)g_\pi\lambda_\pi \\ &\quad + B_{nr}(\tau)(\kappa_r - \kappa_{nr})r_t + B_{n\pi}(\tau)(\kappa_\pi - \kappa_{n\pi})\pi_t, \end{aligned} \quad (4.52)$$

$$\begin{aligned} \mu_I(t, \tau) - R_t &= -B_{rr}(\tau)g_r\lambda_r + \lambda_I\sigma_I \\ &\quad + B_{rr}(\tau)(\kappa_r - \kappa_{rr})r_t, \quad \forall \tau > 0. \end{aligned} \quad (4.53)$$

The third no-arbitrage condition (4.33) still remains satisfied under the parameter separation.

As shown in equation (4.52), the parameter separation adds two additional terms $(\kappa_r - \kappa_{nr})B_{nr}(\tau)r_t$ and $(\kappa_\pi - \kappa_{n\pi})B_{n\pi}(\tau)\pi_t$ to the original no-arbitrage condition (4.31). Similarly, there is one more additional term $B_{rr}(\tau)(\kappa_r - \kappa_{rr})r_t$ in equation (4.53).

We would claim that the new equations (4.52) and (4.53) with the additional terms still obey the no-arbitrage principle in a *broader* sense. Take for example a portfolio consisting of nominal bonds. Let $(\bar{\alpha}_1, \bar{\alpha}_2, \bar{\alpha}_3)$ be the weights of a replicated riskless portfolio, that is $\sum_{i=1}^3 \bar{\alpha}_i dP_{it}/P_{it}$ is free from uncertainty. The no-arbitrage condition (4.31) requires that the excess return of the replicated riskless portfolio over the riskless return is exactly equal to zero, that is,

$$\left(\sum_{i=1}^3 \bar{\alpha}_i \frac{dP_{it}}{P_{it}} \right) - R_t = 0.$$

This is a very strict version of the no-arbitrage principle, a less strict version would require only that the positive excess return

$$\left(\sum_{i=1}^3 \bar{\alpha}_i \frac{dP_{it}}{P_{it}} \right) - R_t > 0,$$

cannot be obtained *with probability one*.

Applying the condition (4.52), the excess return of the riskless portfolio is now given by

$$\left(\sum_{i=1}^3 \bar{\alpha}_i \frac{dP_{it}}{P_{it}} \right) - R_t = \sum_{i=1}^3 \bar{\alpha}_i \left(B_{nr}(\tau_i)(\kappa_r - \kappa_{nr})r_t + B_{n\pi}(\tau_i)(\kappa_\pi - \kappa_{n\pi})\pi_t \right).$$

Although one does not have the excess return of the riskless portfolio strictly equal to zero, one does not have a certain positive excess return either since

the factors r_t and π_t are not observable. Even though we can estimate the unobservable factors using some filtering method, we cannot fix the value with certainty. Moreover, according to the estimation results for the factors, the estimated r_t fluctuates around zero so the sign of the factors cannot be fixed. For these reasons, we claim that the no-arbitrage principle is still satisfied with the parameter separations in a broader sense.

4.4.4 Estimation of the Modified Model

This section is devoted to the estimation of the modified model. All empirical implementations are parallel to the estimation process in Section 4.4.2 but incorporate the parameter separations discussed in Section 4.4.3.

The Modified Term Structure of Real Yields

Applying the Kalman filter method to filter the real interest rate from the market real yields in the modified model, the observation equation now becomes

$$Y_r(t, t + \tau, r_t) = \frac{A_r(\tau)}{\tau} + \frac{1 - e^{\kappa_{rr}\tau}}{\kappa_{rr}\tau} r_t + \epsilon_t^\tau,$$

where $A_r(\tau)$ is given in (4.50) with the new independent sensitivity parameter κ_{rr} . The state equation remains the equation (4.16).

The result of the parameter estimation is given in Table 4.8. The estimated mean-reversion parameter $\kappa_r = 3.2467$, which is close to the OLS estimation result given in Table 4.5, while the sensitivity parameter $\kappa_{rr} = 0.1241$ in Table 4.8 has a very similar value with the estimate $\kappa_r = 0.1248$ in the original model given in Table 4.4. This result indicates that with the parameter separation that both the mean reversion and bond yield sensitivity can be more properly captured. Comparing the estimation results of the original model in Table 4.4 and the modified model in Table 4.8, apart from the mean-reversion parameters κ_r , the other parameters have similar values in both models. Furthermore, the two estimations have very similar fitting errors and very similar likelihood values.

Figure 4.10 compares the estimated factor r_t in the original and modified models. Although the mean-reversion parameter κ_r changes greatly in the modified model compared with that in the original model, the estimated factors have very similar trajectories. To determine whether it is statistically significant to consider the parameter separation, we employ the three

Log Likelihood = 10059.04

	Estimate	t-stat.	Est.(prev.)
κ_r (mean-rev.)	3.2467	1.58	(0.1248)
κ_{rr} (sensit.)	0.1241	12.85	0.1248
\bar{r}	0.0014	0.34	0.0040
g_r	0.0102	36.96	0.0101
λ_r^*	-0.1282	-0.67	-0.5161
σ_{ϵ^r}	0.0008	49.77	0.0008

$\lambda_r^* = \lambda_r - \sigma_I \rho_r$

τ	5Y	7Y	10Y
Mean	1.16%	1.56%	1.90%
SD	2.5e-3	2.6e-3	2.3e-3
$A_r(\tau)/\tau$	1.14%	1.48%	1.89%
$B(\tau)/\tau$	74%	67%	57%
$\hat{\sigma}_{\epsilon^r}$	8.55e-4	5.87e-4	7.80e-4
$\hat{\sigma}_{\epsilon^r}/\text{SD}$	35.01%	22.69%	33.37%

Table 4.8: Upper Panel: estimated parameters for the real yield formula/
Lower Panel: real yield statistics and price Sensitivities

information criteria: AIC, BIC, and HQIC¹¹ and give their values in Table 4.9. The strict BIC rejects the parameter separation while the AIC and the HQIC support the statistical improvement achieved by the parameter separation. In a rough sense, the statistical performance of the two models is equally good.

	original	modified
AIC	-11.112717	-11.114469
BIC	-11.097513	-11.096225
HQIC	-11.106883	-11.107468

Table 4.9: Comparing information criteria for the real yield models

We note that based on real yield data we can only determine the adjusted

¹¹For the definitions of the information criteria see the footnote for Table 4.3.



Figure 4.10: Comparing Estimated Real Rates

market price of risk $\lambda_r^* = \lambda_r - \sigma_I \rho_{Ir}$. The market price of real rate risk λ_r will be determined later in Section 4.4.5 after having obtained estimates σ_I and ρ_{Ir} .

The Modified Term Structure of Nominal Yields

In the modified nominal yield model, the observation equation now becomes

$$Y_n(t, t + \tau, r_t, \pi_t) = \frac{A_n(\tau)}{\tau} + \frac{1 - e^{-\kappa_{nr}\tau}}{\kappa_{nr}\tau} r_t + \frac{1 - e^{-\kappa_{n\pi}\tau}}{\kappa_{n\pi}\tau} \pi_t + \epsilon_t^r, \quad (4.54)$$

where $A_n(\tau)$ is given by (4.51). Recall that the agents input the information concerning r_t from the real bond market and need to filter out the expected inflation rate π_t . The state equation is still equation (4.17) for π_t .

As same as in Section 4.4.2, we adopt the iterative estimation process to make the correlation coefficient $\rho_{r\pi}$ coincide with the estimated correlation from the residuals $\hat{\rho}_{r\pi} = \mathbf{E}[\Delta \hat{W}_t^r \Delta \hat{W}_t^\pi] / \Delta$. We also start with zero initial correlation $\rho_{r\pi}^{(1)} = 0$ and end with the estimated correlation $\hat{\rho}_{r\pi}^{(1)} = -0.4990$ at the first step. For the second step of the estimation, we set $\rho_{r\pi}^{(2)} = \hat{\rho}_{r\pi}^{(1)} = -0.4990$ and obtain the estimated correlation $\hat{\rho}_{r\pi}^{(2)} = -0.4951$. We stop this iterative process at the second step, as same as the estimation for the original model since we judge the two correlation coefficients $\rho_{r\pi}^{(2)} = -0.4990$ and $\hat{\rho}_{r\pi}^{(2)}$ are close.

Table 4.10 gives the estimated results for the modified model.

Log-likelihood = 27789.48

	Estimates	t-stat.	Est.(prev.)
κ_π (mean-rev.)	-0.6060	-7.01	(0.4016)
$\kappa_{n\pi}$ (sensit.)	0.3495	34.64	0.4016
g_π	0.0064	56.84	0.0067
κ_{nr} (sensit.)	0.1770	65.33	(*) 0.1248
λ_π	-1.6864	-21.65	-1.5680
ξ_0	-0.0011	-5.60	-0.0012
σ_{ϵ^τ}	0.0024	51.11	0.0025
$\rho_{r\pi}$	-0.4960		-0.5476

(*) from Table 3.5

τ	1M	3M	6M	1Y	2Y
Mean	1.37%	1.47%	1.63%	1.84%	2.26%
SD	0.59%	0.66%	0.75%	0.75%	0.75%
$\frac{A(\tau)}{\tau}$	-0.04%	0.09%	0.28%	0.63%	1.25%
$\frac{B_{n\pi}(\tau)}{\tau}$ (Sensitivity)	98.55%	95.76%	91.75%	84.39%	71.95%
σ_{ϵ^τ}	0.30%	0.21%	0.14%	0.13%	0.22%
$\sigma_{\epsilon^\tau}/SD$	50.77%	31.70%	19.17%	16.71%	29.13%

τ	3Y	5Y	7Y	10Y	20Y
Mean	2.65%	3.31%	3.38%	4.17%	4.95%
SD	0.65%	0.47%	0.37%	0.32%	0.29%
$\frac{A(\tau)}{\tau}$	1.76%	2.56%	3.14%	3.75%	4.74%
$\frac{B_{n\pi}(\tau)}{\tau}$ (Sensitivity)	61.95%	47.26%	37.34%	27.74%	14.29%
σ_{ϵ^τ}	0.24%	0.24%	0.26%	0.26%	0.30%
$\sigma_{\epsilon^\tau}/SD$	36.17%	51.70%	68.55%	83.04%	103.33%

Table 4.10: Upper Panel: estimated parameters for the modified nominal term structure; Lower Panel: fitting errors and yield sensitivity

The sensitivity parameter $\kappa_{n\pi} = 0.3495$ has similar value to that of $\kappa_{n\pi} = 0.4016$ in the original model while the mean-reversion parameter has the estimated value $\kappa_\pi = -0.6060$. Comparing this estimate with the OLS estimation result κ_π given in Table 4.7, although they are not quite coincident, both indicate that the filtered factor π_t is not a stationary process. So, with the parameter separation, the model can reflect the two features,

mean-reversion and bond-yield sensitivity, more properly.

To determine the effect of the parameter separation on the estimates, we compare the estimation results of Table 4.10 and the previous results given in Table 4.6, which are also quoted in Table 4.10. In this table we can see that apart from the mean-reversion parameter κ_r , all other estimates have similar values. Because of the similarity of the estimates, the coefficients $A_n(\tau), B_{n\pi}(\tau)$ are very similar to the those based on the original model.

Fig 4.11 compares the filtered expected inflation rates π_t of the original and modified model, and we can see they are very close to each other.

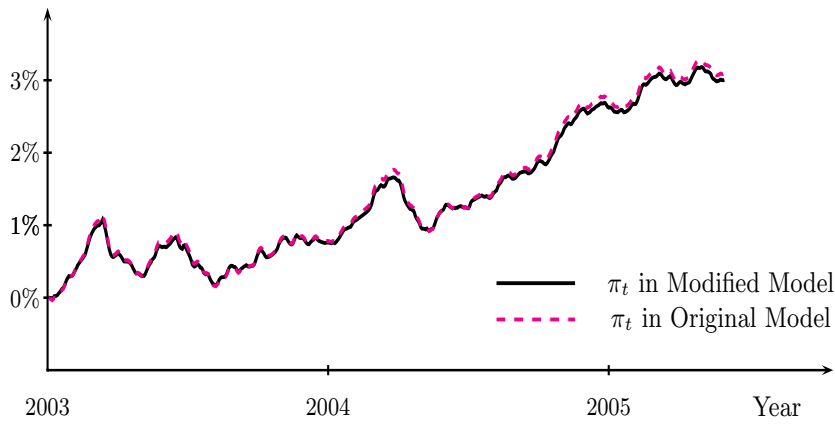


Figure 4.11: Estimated π_t in both models

The fitting errors given in the lower panel given in Table 4.10, roughly speaking, are slightly reduced by the separation¹². This reduction of the fitting errors leads to an increase of the likelihood value. Taking the information criteria¹³ to judge the statistical necessity for the parameter separations, Table 4.11 shows that all three information criteria support the statistical necessity of considering the parameter separation.

¹²Because of the precision limit, in order to compare fitting errors of Tables 4.10 and 4.6, it is better to compare the relative fitting errors σ_ϵ/SD in the last row of the lower panels.

¹³See the footnote by Table 4.9.

	original	modified
AIC	-30.358428	-30.709212
BIC	-30.321366	-30.661031
HIC	-30.345082	-30.691863

Table 4.11: Comparing Information Criteria

Some Concluding Remarks

We have seen much statistical evidence that suggests the use of the parameter separations for κ_r and κ_π , both features with respect to sensitivity and mean reversion can be represented much better and the three information criteria give higher values.

However, we still remark on a point that should be taken of for future research. In the modified model, the estimate of the mean-reversion parameter κ_r in Table 4.8 reflects the mean-reversion strength of r_t much more properly than the estimate in Table 4.4, but the statistical support (the reduction of the fitting error and the increase of the information criteria) seems not significant enough. This might be explained by the following fact that the t -statistic for the separated mean-reversion parameter κ_r in Table 4.8 is 1.58, which is not significant at the 5% level.

In the following two subsections we will estimate the parameters for the realized price index and the stock return, which is also required to calculate the optimal intertemporal strategies.

4.4.5 Estimation of Realized Inflation Dynamics

We estimate the price index dynamics (3.1) based on market data. We employ the Consumer Price Index for all urban consumers (CPI-U) provided by the U.S. Department of Labor¹⁴, which are used to adjust the US TIPS.

Using Itô's Lemma, we transform the dynamics (3.1) into

$$d \ln I_t = \left(\pi_t - \frac{\sigma_I^2}{2} \right) dt + \sigma_I dW_t^I.$$

¹⁴<http://www.bls.gov/cpi/home.htm>

Discretising it using the Euler-Maruyama scheme, we obtain

$$\ln I_{t+\Delta} - \ln I_t = \left(\pi_t - \frac{\sigma_I^2}{2}\right)\Delta + \sigma_I(W_{t+\Delta}^I - W_t^I), \quad (4.55)$$

where we assume π_t follows the dynamics (4.17).

The annualized realized inflation $(\ln I_{t+\Delta} - \ln I_t)/\Delta$ is plotted in Fig. 4.12.

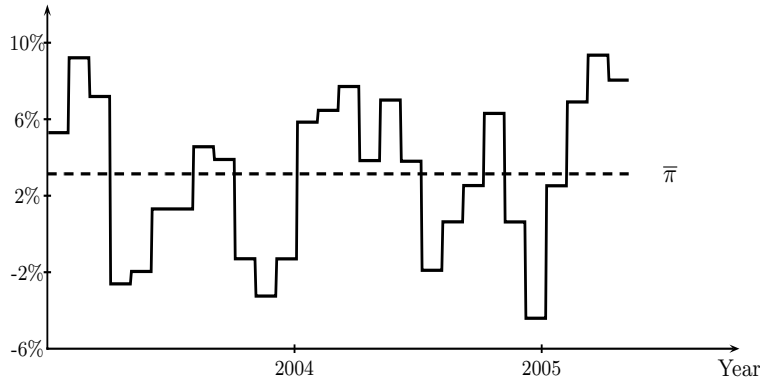


Figure 4.12: Realized and Filtered Annualized Inflation

To estimate the process π_t which is not unobservable and will be estimated through the time-discrete observation of the price index I_t , we face a filtering problem as encountered in the previous subsections. We still employ the Kalman filter method. In this case, the observation equation is given by the dynamics (4.55) and the state equation is the dynamics (4.17) of π_t .

The parameter estimation results are given in Table 4.12.

	Estimate	t-stat.
κ_π	0.4163	5.38
g_π	0.0000	0.00
$\bar{\pi}$	0.0315	4.18
σ_I	0.0115	11.47

Table 4.12: Estimation Results for the CPIU

The estimate result $g_\pi = 0.0$ suggests that the underlying factor π_t should remain constant at the level $\bar{\pi} = 3.149\%$. We show the expected $\pi_t = \bar{\pi}$,

for all t in Figure 4.12.

It is worth remarking that the estimation result for the expected inflation rate π_t here is different from that given in Figure 4.11 previously based on the nominal term structure model. The variable π_t in the both models incorporates the (instantaneous) inflation expectations. However, given the different context of both models, those estimations for π_t are based on different data set: the estimation here is based on the current realized price index, while the previous estimations in the nominal bond yield formula (4.21) in Sections 4.4.2 and 4.4.4 are based on the nominal and real bond yields with the time maturity stretching from one month until 20 years. Therefore, the variable π_t might have different interpretations. The result given in Figure 4.12 (constant $\bar{\pi}$) reflects the market expectations of the current price level while the result shown in Figure 4.11 reflect the market expectations of a long-term development of the inflation. We decide to keep both interpretations for π_t within the appropriate both context.

Following the result (4.41), the market price of the price index risk λ_I is given by

$$\lambda_I = -\frac{\xi_0}{\sigma_I} = \frac{0.0012}{0.0115} = 0.1043. \quad (4.56)$$

Next we calculate the correlation between W_t^I , W_t^r , and W_t^π . For W_t^r and W_t^π , we adopt the estimation results of the original model in Subsection 4.4.2 which are obtained on a daily basis. Since the estimated shock W_t^I is on a monthly basis, we accumulate W_t^r and W_t^π to monthly shocks by summing them up.

The sample correlations of the monthly shocks are calculated as $\rho_{Ir} = 0.0609$ and $\rho_{I\pi} = -0.0688$. Both correlations are quite low.

Having estimated the correlation ρ_{Ir} and using the result for λ_r^* in Table 4.8, we can calculate the market price of real interest rate risk by $\lambda_r = \lambda_r^* - \sigma_I \rho_{Ir} = -0.5168$.

4.4.6 Estimation of Stock Return Dynamics

For our intertemporal asset allocation problem, in addition to the bond assets modelled above, we also include one stock asset in the investment opportunity set. Assume that the dynamics of the stock price follow the

stochastic process

$$\frac{dP_S(t)}{P_S(t)} = (R_t + \lambda_S \sigma_S)dt + \sigma_S dW_t^S, \quad (4.57)$$

where σ_S, λ_S are positive constants. Applying the Itô formula to equation (4.57), we obtain the equivalent representation

$$d \ln P_S(t) = \left(R_t + \lambda_S \sigma_S - \frac{\sigma_S^2}{2} \right) dt + \sigma_S dW_t^S. \quad (4.58)$$

The model to be estimated is obtained by applying the Euler-Maruyama approximation method to the continuous-time dynamics (4.58) where the discretization interval is $\Delta t = 1/250$ for these daily data. The estimation of the parameters in the dynamics (4.58) is based on data of the daily S&P500 index from Jan. 02 2003 - May 31 2005 including 603 observations, which are plotted in Figure 4.13. The data can be found in “Finance Yahoo”. For the riskless rate R_t we adopt the Federal Funds rate whose values can be found in Section 4.2. Figure 4.14 shows the time series of the daily excess stock returns and Figure 4.15 shows their distribution. The parameters in (4.58) are estimated as $\sigma_S = 0.1391$ and $\lambda_S = 0.8669$.

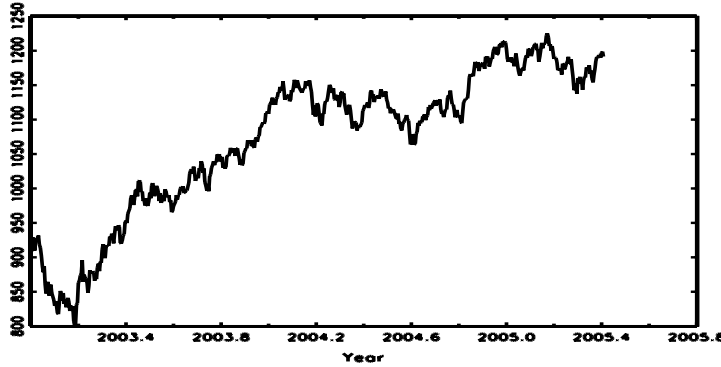


Figure 4.13: SP500 Index

For the asset allocation problem we will need to know the correlations between the shocks W_t^S and W_t^r , W_t^π and W_t^I . The estimated innovations W_t^r and W_t^π are adopted from the results of the original model given in Section 4.4.2. Based on the estimation results, the sample correlations are given by

$$\rho_{Sr} = 0.1744 \quad \rho_{S\pi} = -0.0221 \quad \rho_{SI} = -0.0587 \quad .$$

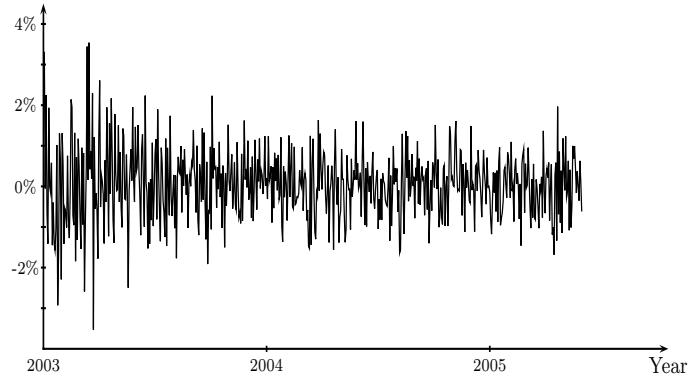


Figure 4.14: Daily Excess Returns(S&P500)

The correlation between the shocks W_t^S and W_t^I is calculated in a monthly basis.

4.5 Summary and Comparison of Models I and II

This chapter models the financial market in order to obtain the necessary information required for the intertemporal asset allocation strategies. Two models are used to model the bond market. The first model follows the data-oriented approach where the underlying factors are unspecified and have to be filtered out of the market nominal bond yields. The second model is based on a theoretically economical consideration where the instantaneous real interest rate and instantaneous expected inflation rate are specified *a priori* as the two factors. Summarizing the discussion of this chapter we compare the estimated factors of these two different kinds of the model in Figure 4.16, where the estimated real interest rate and the expected inflation in Model II are together plotted with the estimated factors in Model I labeled by "Factor1" and "Factor2". Surprisingly, we find that the trajectory of the estimated real interest rate in Model II shares some similarity with that of the estimated Factor 1 in Model I. While comparing the other remaining factors, we also find similarity between the trajectories of the estimated expected inflation rate in Model II and the estimated Factor II in Model I. This similarity in the estimation results is remarkable given the fact that one model (Model I) follows a data-oriented approach while the other (Model II) is based on a theoretical set-up.

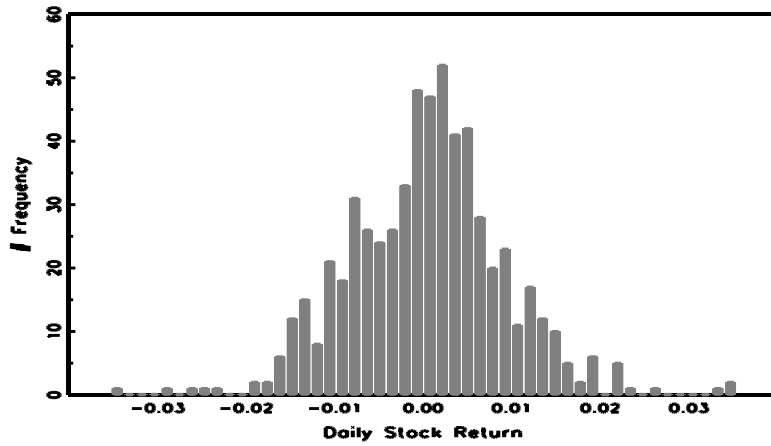


Figure 4.15: S&P500 Excess Daily Returns Distribution

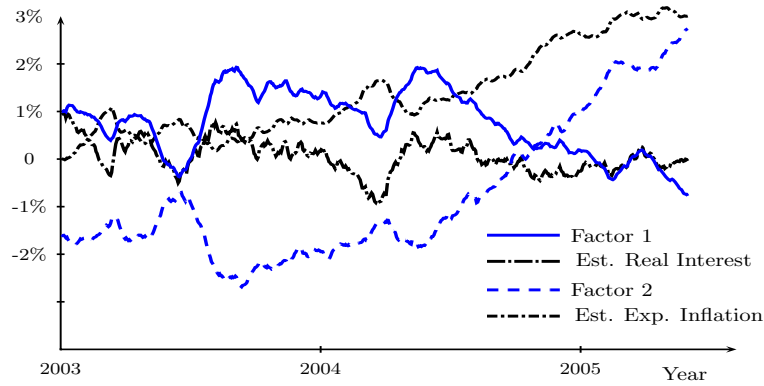


Figure 4.16: Comparison of the Estimated Factors in the Model I and Model II

Thus this chapter finishes its task of acquiring the requisite information for the intertemporal asset allocation problem. In the next Chapter we will give concrete asset allocation recommendations based on this information.

Chapter 5

Optimal Investment Strategies

In this chapter we give concrete intertemporal investment recommendations for the two models estimated in Chapter 4. The explicit formulas for the optimal investment strategies are derived from the general solution developed in Chapter 3 and the concrete recommendations of investment amounts are calculated by adopting the estimation results obtained in Chapter 4. We provide a simulation study to illustrate the advantage of considering intertemporal features while taking investment decisions. Risk aversion effects and horizon effects on the optimal investment strategies are studied both with and without inflation risk. In an investment environment exposed to inflation risk, inflation-indexed bonds turn out to be an efficient hedging asset for long-term investment plans.

5.1 How large should the Relative Risk Aversion be?

The relative risk aversion (RRA) parameter of a utility function $U(c)$ is defined as

$$\text{RRA} := -\frac{U''(c)c}{U'(c)}.$$

For the power utility function in our case as defined in (3.7) the RRA parameter is exactly the parameter γ . For an intertemporal portfolio decision, the RRA parameter γ determines the weights on the myopic portfolio, the intertemporal hedging term, and the inflation risk hedging term as given in

the formula (3.18). Here we provide a way to determine the RRA parameter based on market data.

Property 10 *Assume there are two assets: one risky and one risk-less assets available for investment. Let the risky asset have μ as the expected return and σ^2 as the variance. The risk-less asset has constant return R . If the RRA of an agent is equal to*

$$RRA \sim \frac{2(\mu - R)}{\sigma^2 + \mu^2}, \quad (5.1)$$

then the agent is indifferent between investing in the risky or the risk-less asset.

Using an equilibrium argument, a risky asset with the average return μ and the volatility σ and the riskless asset with the riskless return R can only coexist when the investors have the CRRA utility with the risk aversion parameter given in equation (5.1). Otherwise, the investors would only demand the one asset with higher utility and the other would vanish from the market. We still need to remark that this argument is based on a partial equilibrium perspective where the investment decisions are taken only between one risky asset and one risk-less asset.

In the following we provide the value of the market RRA parameter from the partial equilibrium point of view above based on the real market situations which have been obtained in Chapter 4. Table 5.1 gives the values of the market RRA as quoted in (5.1) based on the estimation results for Model I given in Section 4.3.3. We employ the results for the two-factor dynamics given in Table 4.2. We use $P_n(\tau)$ to denote the price of the nominal bond with time-to-maturity τ year. The excess return for the bond $P_n(\tau)$ is given by $-B(\tau)^\top \Gamma \lambda$ based on equation (4.9) and the variance is given by $B(\tau)^\top \Gamma \Gamma^\top B(\tau)$ which can be derived from the bond return formula (4.7).

	$P_n(1)$	$P_n(3)$	$P_n(10)$	$P_n(20)$
$\mu - R$	1.32%	3.19%	5.98%	7.30%
σ	0.0063	0.0155	0.0337	0.0520
σ^2	3.98×10^{-5}	0.0002	0.0011	0.0027
$\frac{\mu - R}{\sigma^2}$	330.90	132.64	52.81	26.96
$\frac{\mu - R}{\sigma}$ (Sharpe Ratio)	2.09	2.06	1.78	1.40
Market RRA	123.37	50.71	25.39	18.16

Table 5.1: Asset Returns and Risk Compensation for Model I

5.1. HOW LARGE SHOULD THE RELATIVE RISK AVERSION BE?93

This market RRA parameter can be analogously calculated given Model II and we give the values in Table 5.2. For Model II we have two additional kinds of asset: the inflation-indexed bond (IIB) with time-to-maturity τ years, denoted by $P_I(\tau)$, and the stock, denoted by P_S . The excess return $\mu - R$ for the nominal bond is calculated based on the no-arbitrage condition (4.31) and the volatility σ is given by

$$g_r^2 B_{nr}^2(\tau) + 2g_r g_\pi B_{nr}(\tau) B_{n\pi}(\tau) \rho_{r\pi} + g_\pi^2 B_{n\pi}^2(\tau),$$

which is the volatility of the return shock in the nominal bond return formula (4.19). The parameter values are taken from the estimation results given in Table 4.6 in Section 4.4.2. For the inflation-indexed bond, the excess return is based on the no-arbitrage condition (4.32) and the volatility is obtained by calculating the volatility of the return shock in the IIB return formula (4.28). The values of the stock can be obtained easily from the estimation results given in Section 4.4.6.

	$P_n(1)$	$P_n(3)$	$P_n(10)$	$P_n(20)$	$P_I(10)$	P_S
$\mu - R$	1.36%	3.14%	5.55%	6.45%	3.10%	12.06%
σ	0.0082	0.0218	0.0514	0.0674	0.0582	0.1391
σ^2	6.74×10^{-5}	0.0005	0.0026	0.0045	0.0034	0.0193
$\frac{\mu - R}{\sigma^2}$	201.26	66.03	21.48	14.23	9.15	6.23
$\frac{\mu - R}{\sigma}$ (Sharpe Ratio)	1.65	1.44	1.08	0.95	0.53	0.87
Market RRA	107.86	42.96	19.41	14.83	14.26	7.12

Table 5.2: Asset Returns and Risk Compensation for Model II with the parameter separations

Comparing the two Tables 5.1 and 5.2, we see that they give quite similar results for excess returns and Table 5.2 has higher estimated volatility for the four nominal bond assets. Therefore, Table 5.2 has lower return-risk measures for both cases $\frac{\mu - R}{\sigma^2}$ and $\frac{\mu - R}{\sigma}$ (the Sharpe Ratio). The market RRA parameters for the stock in Table 5.2 have typical sizes as usually quoted in the literature, for example, they are between 3 and 8 in the calibration results in Munk et al (2004), while they are much higher for the bond assets considered in the thesis. A high market RRA parameter indicates a higher profitability in relation to the risk measured by the variance σ^2 .

The summarized values for the market RRA give us same suggestions for

choosing a range for the theoretical RRA parameter γ in agents' utility function (3.7) for the concrete investment recommendations later.

Here we comment on the market RRA parameter, which is closely related to the risk compensation measure $\frac{\mu-R}{\sigma^2}$, by comparing it with the Sharpe ratio (defined as $\frac{\mu-R}{\sigma}$). Based on an elementary mathematical argument, we know that for a smaller volatility, say, $K^{-1}\sigma$, the Sharpe ratio increases K -fold while while the other risk compensation measure RRA parameter increases K^2 -fold. Because the RRA parameter relates closely to the latter risk measure, the RRA parameter exaggerates risk compensation more than the Sharpe ratio for an asset with smaller volatility σ . We do not go any further into discussing the question as to which risk measures are more proper since this issue is beyond the scope of this thesis.

5.2 Optimal Portfolio Allocation for Model I

In this section we will give an explicit formula for the optimal portfolio based on Model I in Section 4.3 by using the expectation operator representation (3.25) in Section 3.3. Here, we do not consider intermediate consumption, so $\epsilon_1 = 0$. Also, the price index I_t is set to be equal to 1 and the inflation risk is not considered in Model I.

Recall that the expectation operator representation (3.25) did not consider the measurement errors as did the empirical model (4.15) in Section 4.3 . In the first example in this section we provide a simulation study to examine the performance of the recommended investment strategies with measurement errors. In the second example, risk aversion effects and horizon effects on the investment strategies are studied.

5.2.1 Explicit Forms for Investment Strategies

Assume that the investment opportunity set includes n nominal bonds with distinct maturities dates $\bar{T}_1, \dots, \bar{T}_n$ whose yields follow the dynamics (4.3) where the factor X_t is an n -dimensional stochastic process satisfying the stochastic differential equation (4.6). Let $P_{it} := P(t, \bar{T}_i, X_t)$ denote the price of the i -th bond maturing at time \bar{T}_i . From (4.7), the bond return

dynamics can be represented in the vector form

$$\begin{pmatrix} \frac{dP_{1t}}{dP_{1t}} \\ \vdots \\ \frac{dP_{nt}}{P_{nt}} \end{pmatrix} = \mu_t dt + \Sigma_t dW_t^X,$$

where

$$\mu_t := \begin{pmatrix} \mu_P(\bar{T}_1 - t, X_t) \\ \vdots \\ \mu_P(\bar{T}_n - t, X_t) \end{pmatrix} \quad (5.2)$$

$$\Sigma_t := -\mathcal{B}_t \Gamma, \text{ with } \mathcal{B}_t := \begin{pmatrix} B_1(\bar{T}_1 - t) & \cdots & B_n(\bar{T}_1 - t) \\ \vdots & \ddots & \vdots \\ B_1(\bar{T}_n - t) & \cdots & B_n(\bar{T}_n - t) \end{pmatrix}. \quad (5.3)$$

The bond returns are assumed to satisfy the no-arbitrage condition (4.9), so the coefficients $B(\tau)$ and $A(\tau)$ in the formula (4.3) are given by the formulas (4.13) and (4.14).

In order to construct the optimal portfolio α_t^* , we use the formula (3.17) where we have first to obtain the value function $\Phi(t, T, X_t)$.

Property 11 *Using the expectation operator representation (3.25), the value function $\Phi(t, T, X_t)$ for the investment opportunity described above is given by*

$$\Phi(t, T, X_t) = \tilde{f}(t, T) e^{\frac{1-\gamma}{\gamma} B(T-t) X_t}, \quad (5.4)$$

where

$$\begin{aligned} \ln \tilde{f}(t, T) &= -\frac{\delta}{\gamma}(T-t) + \frac{1-\gamma}{2\gamma^2} \lambda^\top \lambda (T-t) + \frac{1-\gamma}{\gamma} \xi_0 (T-t) \\ &\quad + \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^T B(T-s)^\top \Gamma \lambda ds \\ &\quad + \frac{1}{2} \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^T B(T-s)^\top \Gamma \Gamma^\top B(T-s) ds. \end{aligned}$$

Recall the notation: λ is an $n \times 1$ vector in (4.9) representing the market price of risk, ξ_0 is the constant in the coefficient $A(\tau)$ given by (4.14), $B(\tau)^\top = \left(\frac{1}{\kappa_1}(1 - e^{-\kappa_1 \tau}), \dots, \frac{1}{\kappa_n}(1 - e^{-\kappa_n \tau})\right)$ as given in (4.13), and Γ is the volatility of the factor as given in (4.6).

Using the result of Property 11, the elasticity term of the factor in the optimal portfolio α_t^* given by (3.17) can be easily calculated and is given by

$$\frac{\Phi_X(t, T, X_t)}{\Phi(t, T, X_t)} = \frac{1 - \gamma}{\gamma} B(T - t), \quad (5.5)$$

so the vector of the optimal intertemporal portfolio weights α_t^* is now given explicitly by

$$\alpha_t^* = \frac{1}{\gamma} \underbrace{(\Sigma_t^\top)^{-1} \lambda}_{\text{I. } \alpha_t^{(M)}} + (1 - \frac{1}{\gamma}) \underbrace{(-\Sigma_t^\top)^{-1} \Gamma^\top B(T - t)}_{\text{II. } \alpha_t^{(I)}}. \quad (5.6)$$

Recall that in the general framework, the solution of the optimal portfolio weights can be decomposed into three parts: I. $\alpha^{(M)}$ the myopic portfolio (or the risk-return trade-off portfolio), II. $\alpha^{(I)}$ the intertemporal hedging term, and III. $\alpha^{(P)}$ the inflation hedging term as already demonstrated in equation (3.18). Comparing solution (5.6) with that in the general model, the third term of inflation hedging due to the stochastic price level does not appear here because Model I has a constant price level.

We define the *conservative* portfolio here as the intertemporal hedging portfolio,

$$\text{Conservative Portfolio} := \alpha_t^{(I)},$$

then the optimal portfolio α_t has the decomposition

$$\alpha_t = \frac{1}{\gamma} \text{Myopic Portfolio} + (1 - \frac{1}{\gamma}) \text{Conservative Portfolio}. \quad (5.7)$$

The name "conservative portfolio" comes from the fact that it is the portfolio held by the most conservative agents (considering the limit $\gamma = \infty$). A decomposition such as (5.7) is one example of the *general separation theorem* of Merton (1990)(p.490) that the optimal portfolio can be represented as a linear combination of three mutual funds: cash, the myopic and the conservative portfolio. The weights on each fund are determined by agents' risk aversion.

It is quite surprising to observe that the optimal portfolio weights (5.6) do not depend on the factor level X_t directly. This is due to the log-linear dependence of the value function (5.4) on the factor X_t . The mathematical reason for this is that the factor follows a mean-reverting Gaussian process so that it depends linearly on its past¹. Although the solution of the

¹For the precise formula see equation (8.72) in the proof of Property 11.

optimal portfolio weights α_t^* does not depend on the current factor level X_t , it appears only in an intertemporal framework and makes the optimal intertemporal portfolio different from the optimal static portfolio. The intertemporal hedging portfolio is affected by the mean reverting parameters κ_i .

From the expression (5.6), we know the required information for the construction of the optimal portfolio includes the mean-reversion parameter κ_i , the volatility of the underlying factors Γ , and the market price of factor risk λ . The intertemporal hedging effect is more significant,

- (i) when the investors are more risk-averse (large γ),
- (ii) when the investment horizon is long (large $T - t$),
- (iii) when the factor is more like a random walk process (small mean reversion speed κ_i), and
- (iv) when the mean-variance portfolio is not too dominant compared to the intertemporal hedging term (mathematically, we need to compare the scale of the market price of risk $|\lambda|$ with the volatility of the long term bond $B(T - t)^\top \Gamma$).

Furthermore, the optimal wealth based on the optimal portfolio evolves according to

$$\begin{aligned}
 \frac{dV_t^*}{V_t^*} &= R_t dt + \alpha_t^{*\top} \left((\mu - R_t \mathbf{1}) dt + \Sigma_t dW_t^X \right) & (5.8) \\
 &= R_t dt + \left(\left(\frac{1}{\gamma} \lambda^\top + \frac{1-\gamma}{\gamma} B(T-t)^\top \Gamma \right) \Sigma_t^{-1} \right) \left(\Sigma_t (\lambda dt + dW_t^X) \right) \\
 &= R_t dt + \frac{1}{\gamma} (\lambda^\top \lambda dt + \lambda^\top dW_t^X) + \frac{1-\gamma}{\gamma} B(T-t)^\top \Gamma (\lambda dt + dW_t^X).
 \end{aligned}$$

An important implication of the formula (5.8) for the wealth evolution is that the optimal wealth evolution is independent of the choice of bond assets, which means that it is independent of the time to maturities of the bonds in which the agents invest. A different choice of bond assets will give rise to a different volatility matrix Σ_t (recall the definition of Σ_t in (5.3)). We can see in the optimal wealth development (5.8) that the volatility matrix

Σ_t no longer appears. Only the market price of risk λ , the risk attitude of the agents γ , and the mean reversion parameters κ_i (appearing in $B_i(\tau)$) determine the optimal wealth process.

5.2.2 Examples: Intertemporal and Information Effects

The analytical solution for the optimal portfolios given above is based on the exact affine term structure (4.15) without measurement error $\epsilon^\tau_t = 0$. When applying the theoretical optimal strategies to the real world we need to take account of the measurement errors.

We develop an investment scenario and use simulation to determine the performance of the theoretical optimal strategies in the model with measurement errors. In the simulation example, we employ the two-factor model to simulate the bond price $P(t, \bar{T}_i, X_t)$ according to

$$P(t, \bar{T}_i, X_t) = e^{-A(\bar{T}-t) - \sum_{i=1,2} B_i(\bar{T}_i-t)^\top X_{it} - (\bar{T}_i-t)\epsilon_{it}}, \quad (5.9)$$

where all parameters take values from the estimation results of the two-factor model given in Table 4.2. For the investment opportunity set, we include two bonds with different time to maturity. At the initial time, the short-term matures in 3 years and the long-term bond in 10 years, so $T_1 = 3$ and $T_2 = 10$. As time evolves, the time to maturity $T_i - t$ decreases. Once the short-term bond matures, a new 3-year bond will be introduced into the investment set immediately. So, the maturities have the time schedule shown in Table 5.3.

	$0 \leq t < 3$	$3 \leq t < 6$	$6 \leq t < 9$	$9 \leq t \leq 10$
$\bar{T}_1 =$	3	6	9	12
$\bar{T}_2 =$	10	10	10	10

Table 5.3: Time schedule of the bond maturities

In our simulation study we consider four different investment strategies:

- **(S1)** The first strategy is the full-information two-factor intertemporal investment strategy. It is the best theoretical investment strategy. The agents adopting this strategy possess the full information of the model's price dynamics, which includes the number of the factors and the parameter values. The strategy is constructed by adopting the

formula (5.6) based on the two-factor model. After elementary operations, the strategy S1 at the time t , denoted by $\alpha_{S1}^*(t)$, is given by the formula

$$\begin{aligned} \alpha_{S1}^*(t) &= \frac{1}{\gamma} \begin{pmatrix} B_1(\bar{T}_1 - t) & B_1(\bar{T}_2 - t) \\ B_2(\bar{T}_1 - t) & B_2(\bar{T}_2 - t) \end{pmatrix}^{-1} \begin{pmatrix} \Gamma_{11} & \Gamma_{21} \\ 0 & \Gamma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \\ &\quad + \frac{1-\gamma}{\gamma} \begin{pmatrix} B_1(\bar{T}_1 - t) & B_1(\bar{T}_2 - t) \\ B_2(\bar{T}_1 - t) & B_2(\bar{T}_2 - t) \end{pmatrix}^{-1} \begin{pmatrix} B_1(10-t) \\ B_2(10-t) \end{pmatrix}, \end{aligned} \quad (5.10)$$

recall that $B_i(\tau) = (1 - e^{-\kappa_i \tau})/\kappa_i$, Γ_{ij} is the corresponding item in Γ , and \bar{T}_i for different t are given in Table 5.3. The agents know that the all parameter values $\mathcal{K}, \Gamma, \lambda$ and ξ_0 are given by the results of the two-factor model in Table 4.2.

- **(S2)** The second strategy is the full-information *mean-variance efficient* (MVE) investment strategy. Agents adopting this strategy also have the full information of the price dynamics as those adopting best theoretical investment strategy S1, but they follow the mean-variance efficient (MVE) strategy. The strategy is constructed by using the MVE portfolio, which is the first term in the formula (5.6) based on the two-factor model. So, this strategy can be represented by

$$\alpha_{S2}^*(t) = \frac{1}{\gamma} \begin{pmatrix} B_1(\bar{T}_1 - t) & B_1(\bar{T}_2 - t) \\ B_2(\bar{T}_1 - t) & B_2(\bar{T}_2 - t) \end{pmatrix}^{-1} \begin{pmatrix} \Gamma_{11} & \Gamma_{21} \\ 0 & \Gamma_{22} \end{pmatrix}^{-1} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \quad (5.11)$$

where the agents also know the parameter values are given in Table 4.2. Recall that the strategy S2 is the best strategy when the investment environment is static. It is also in line with the conventional consideration of portfolio decisions based on the trade-off between return and risk.

- **(S3)** The third strategy is a *partial information* MVE strategy. The agents adopting this strategy have no information about the bond price dynamics. They adopt the same two-factor MVE strategy as the agents adopting S2, but they have to use the original formula given in (3.17), namely,

$$\alpha_{S3}^*(t) = \frac{1}{\gamma} (\Sigma_t \Sigma_t^\top)^{-1} (\mu_t - R_t \mathbf{1}) \quad (5.12)$$

since they do not have information about the bond price dynamics. Their strategy to find proxies for Σ_t and μ_t is to use sample statistics of the bond daily returns. We set the learning period as one year. So, at time t the agents collect the last 50 weekly bond returns for

the two bonds maturing at \bar{T}_1 and \bar{T}_2 over the last year $[t, t - 1]$ and subsequently calculate the sample mean and sample covariance of these daily bond returns. The sample mean minus the average riskless returns of R_t is the proxy for $(\mu_t - R_t \mathbf{1})\Delta$ and the sample covariance matrix is the proxy for $\Sigma_t \Sigma_t^\top \Delta$.

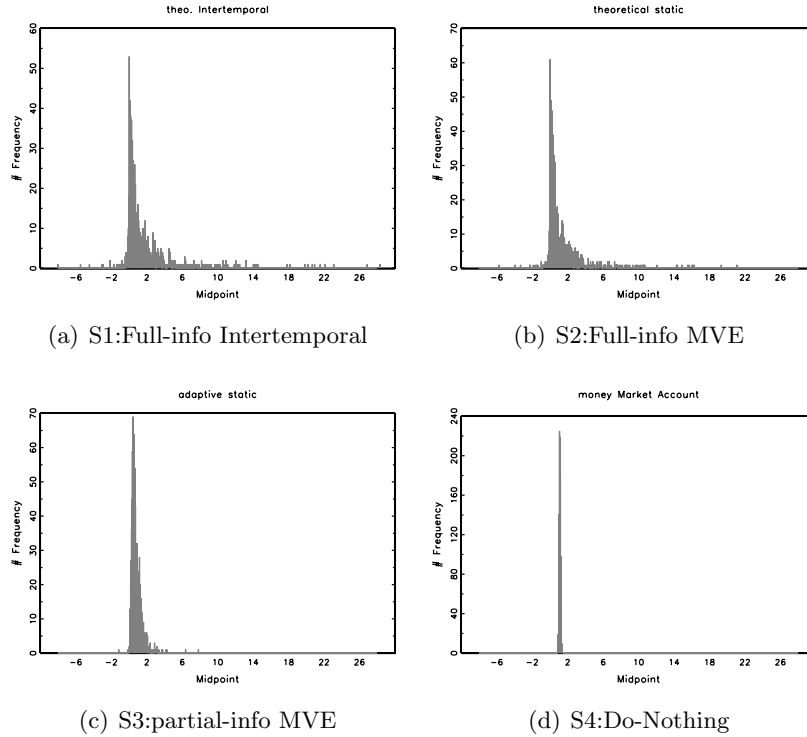
- **(S4)** With the fourth strategy the agents keep all their wealth as money and earn the riskless instantaneous interest rate. This is just the value of the money market account and serves as a reference value.

The risk aversion parameter γ is taken as 7, 15 and 30. The time step for our simulation is $1/50$, which corresponds to a week. At the beginning of the investment horizon, the agents are endowed with one unit of wealth. For each simulation example we repeat the investment scenario 1,000 times. The criteria used to evaluate performance of the strategies are the average and variation of the distribution of the final wealth and the expected final utility. Figures 5.1 – 5.3 provide the final wealth distribution of 1000 simulations and Table 5.4 summarize the statistics.

Strategy	S1	S2	S3	S4
aver. Wealth				
$\gamma = 7$	1.75	1.34	0.79	1.11
$\gamma = 15$	7.35	5.45	0.98	1.11
$\gamma = 30$	4.55	3.32	1.04	1.11
st. deviation of wealth				
$\gamma = 7$	3.24	2.39	0.61	0.08
$\gamma = 15$	3.60	2.44	0.30	0.09
$\gamma = 30$	1.07	0.63	0.13	0.09
aver. Utility				
$\gamma = 7$	-2.07×10^{24}	-9.73×10^{16}	-2.39×10^5	-0.09
$\gamma = 15$	-2.54×10^{-6}	-2.81×10^{-5}	-45.25	-0.03
$\gamma = 30$	-6.78×10^{-14}	-1.47×10^{-11}	-7.04	-0.02

Table 5.4: Example, Statistics of Wealth Distributions, Model I

Strategies S1 and S2 with the full information about the intertemporal feature for the bond price dynamics perform significantly better than the partial-information MVE strategies S3 and the passive strategy S4 for all

Figure 5.1: Example: Final Wealth Distribution, $\gamma = 7$, Model I

three level of γ regarding the wealth distribution. We can see the informational advantage especially clearly when comparing the full-information MVE strategy S2 and the partial-information MVE strategy S3. Both of them are based on the same formula (5.12) but the proxies for the volatility matrix Σ_t and μ_t are different. The investors implementing Strategy S2 have the information about the intertemporal feature of the asset returns so that they can use the more correct formula (5.11). Table 5.4 shows the possession of this information can lead to high profit. Especially, the mean and the volatility of the bond returns vary with the time to maturity. Therefore, the sample mean $\hat{\mu}$ and the sample volatility $\hat{\Sigma}$ of the bond return used for Strategy S3 cannot represent the model μ_t and Σ_t properly. This leads to a quite bad performance of the partial-information strategy S3. Even the passive Strategy S4 can under this situation outperform S3.

In the simulations negative results for final wealth occur, so we adjust the

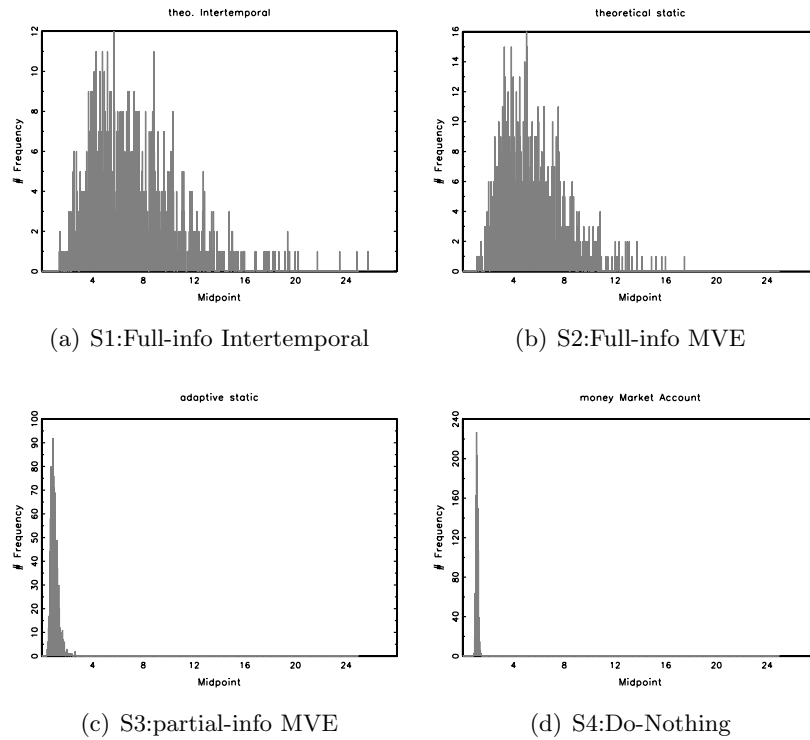


Figure 5.2: Example: Final Wealth Distribution, $\gamma = 15$, Model I

utility function². The average values utility following Strategies S1 and S2 for $\gamma = 7$ are very low because many samples of the realized final wealth take negative value. For $\gamma = 7$, it is very advantageous not to follow any investment strategies because Strategy S4 outperforms the other smart strategies when comparing the average utility of the final wealth. This is because the final wealth of the other strategies take negative values in some samples due to the existence of measurement errors.

Comparing further the two full-information strategies S1 and S2, in Figures 5.1 – 5.3, we can see that the intertemporal S1 is superior to the static S2, in Table 5.4 S1 gives a higher average utility and higher average wealth than

²The adjustment we use is given by

$$\begin{aligned} \tilde{U}(w) &= U(w), & \text{for } w \geq \bar{b} \\ &= U(\bar{b}) + U'(\bar{b})(w - \bar{b}), & \text{for } w < \bar{b}. \end{aligned}$$

Here we take $\bar{b} = 1e^{-6}$.

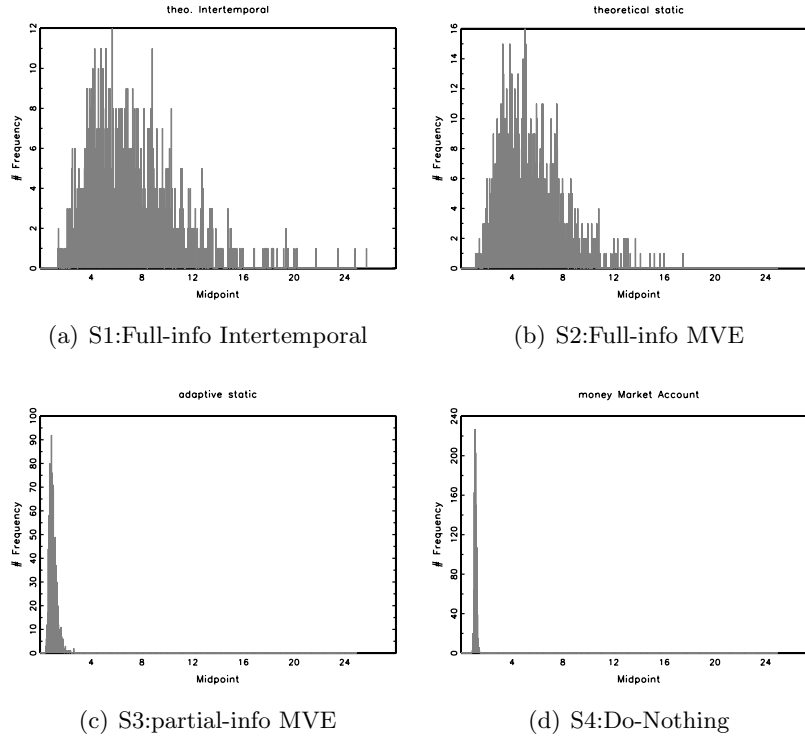


Figure 5.3: Example: Final Wealth Distribution, $\gamma = 30$, Model I

the S3 for all cases of γ .

We found it is interesting that the performance of Strategy S3 become better: the average level of wealth increases while the standard deviation decreases. For the two full-information strategies S1 and S2, the dependence of the average level and the standard deviation of the final wealth on the risk aversion is quite non-linear: for S1 and S2 the highest level are achieved for $\gamma = 15$ where the wealth distributions also have highest standard deviation. Further, comparing the final wealth distributions for $\gamma = 15$ and $\gamma = 30$, we can observe that for Strategies S1 and S2, the more conservative agents $\gamma = 30$ obtain an investment result with a smaller spread (a lower variance) but also a lower return on average than the more aggressive agents. The reason again could be due to the measurement errors.

In this simulation study, We found that the information about the factors plays an important role in obtaining good portfolio performance. Agents

who have information on the dynamics of the underlying factors have a clear advantage over those who do not have this information. This advantage is especially significant when considering an asset with time varying return dynamics where the sample mean and volatility cannot represent the model mean and volatility properly. The bad performance of the partial-information strategy S3 supports this argument. Comparing the two full-information strategies, as expected, the intertemporal strategies are superior to the myopic strategies.

5.2.3 Examples: Risk Aversion and Horizon Effects

Using the explicit formula (5.6), we can give concrete investment recommendations. The recommendations are based on real market situation that is represented by the estimation results given in Table 4.2 in Section 4.3. We employ the results with the two-factor dynamics and therefore the investment opportunity set consists of two nominal bonds: we choose one three-year bond and one ten-year bonds. The investment horizon is set to be 10 years.

The first example studies the effect of the risk aversion parameter (the RRA parameter) on the optimal portfolio weights. In the third column in Table 5.5 we give the optimal portfolio strategies by choosing a typical RRA $\gamma = 4$. Recall the strategies are given in proportion to wealth, so the number 49.59 reads that the optimal portfolio decision should hold 49.59 times of the whole wealth.

	I. $\alpha^{(M)}$	II. $\alpha^{(I)}/\text{Conserv.}$	$\alpha^*(\gamma = 4)$
NB3Y	198.36	0	49.59
NB10Y	-32.44	1	-7.36
Money	-165.91	0	-41.23

Table 5.5: Optimal Weights against Risk Aversion

We observe extraordinary amounts of buying of the short-term bond and short selling of the long term bond for the myopic portfolio $\alpha^{(M)}$ as found in the first column in Table 5.5. It can be understood partly by looking at market features shown in Table 5.1, where we can read that the risk compensation measure $\frac{\mu-R}{\sigma^2}$ of the three-year bond ($= 132.64$) is much higher than that of the ten-year bond ($= 52.81$). This indicates that the short-term bond has higher profitability. However, in the real world, such extreme “optimal”

portfolio strategies are not useful. We need to consider some real market to reduce those extreme strategies. We will provide a study by considering the short-sale constraints in Chapter 5.

Figure 5.4 plots the theoretical optimal portfolio weights in each asset against the risk aversion parameter from 20 to 500.

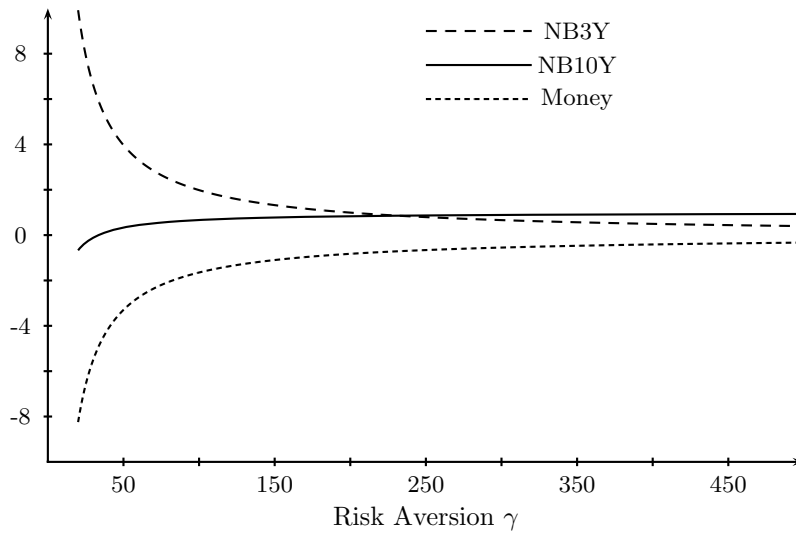


Figure 5.4: Optimal Portfolio Weights for Model I, against γ

The intertemporal hedging portfolio given in the second column in Table 5.5 suggests investing all wealth in the ten-year bond, which matures at the end of the investment horizon. As demonstrated in the portfolio decomposition (5.7), the most conservative agent with $\gamma = \infty$ holds the intertemporal hedging portfolio. The underlying intuition is that only this strategy can guarantee a fixed payout that the most conservative agents will require. This result can be easily verified by expanding the intertemporal hedging term $\alpha^{(I)}$ in equation (5.6). In general, the most conservative agents will invest all their wealth on the long-term bond maturing at the end of the investment horizon. Wachter (2003) has derived the same result by use of the static variational method. From Table 5.5 we can tell that the holders of the myopic and conservative portfolios prefer different assets. The holders of the myopic portfolio prefers to short the long-term bond and money in order to buy the more profitable short-term bond while the holders of the conservative portfolio only buys the long-term bond for its certain payout.

So, we can observe a reversal of the holding positions of the bonds in Figure 5.4 when the value of the risk aversion parameter becomes large enough.

The second example studies the horizon effect on the investment decision. We fix the risk aversion parameter $\gamma = 70$ and vary the investment horizon. Agents invest in a three-year bond and a long-term bond which will mature at the end of the investment horizon. Figure 5.5 plots the optimal investment weights against the investment horizon from 4 to 50 years.

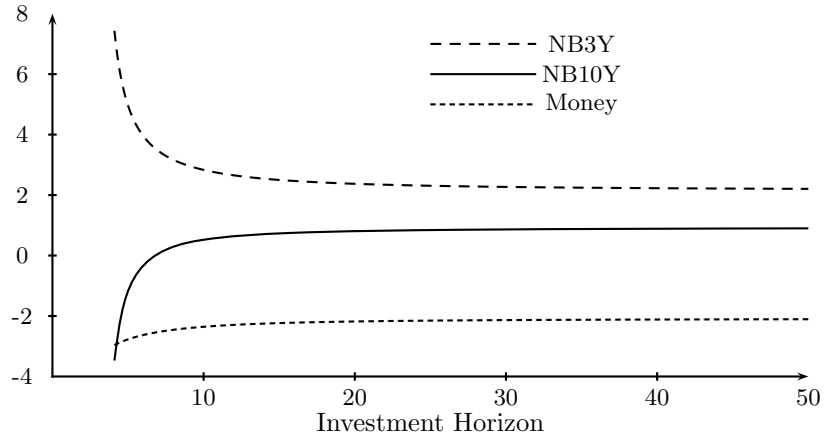


Figure 5.5: Optimal Investment Weights for Model I, Horizon Effect

As the length of the horizon increases, the holding of the long-term bond increases. Since the risk aversion γ is not infinity, agents still demand the short-term bond because of its profitability, but the holding decreases with the length of horizon. As the investment horizon increases further, the optimal investment weights will converge to the values given in Table 5.6.

	NB3Y	NB10Y	Money
limit weights	2.17	0.92	-2.09

Table 5.6: Limits of the Investment Weights when $\tau \rightarrow \infty$

5.3 Optimal Portfolio Allocation for Model II

In this section we will give the explicit formula for the intertemporal investment strategies based on Model II given in Section 4.4.

5.3.1 Analytical Solution

We include four risky assets in the investment opportunity set, namely, two nominal bonds with different maturity dates T_1, T_2 , one inflation-indexed bond (IIB) maturing at T_3 , and one stock. Recall that the returns of the nominal bonds follow the dynamics given in (4.19), where the expressions of the coefficients $A_n(\tau)$, $B_{nr}(\tau)$, and $B_{n\pi}(\tau)$ are given in equations (4.51), (4.48), and (4.49) respectively. The return of the IIB has the dynamics given in (4.28), where the coefficients $A_r(\tau)$, $B_{rr}(\tau)$ are given in (4.50) and (4.47). The stock price is assumed to follow the dynamics (4.58). Following the classic no-arbitrage argument, for example, see Chiarella (2004), these three bond assets form a complete market, which means, that the nominal bonds and IIBs with other maturities can be replicated by the given three bonds.

The returns of the assets in the investment opportunity set are summarized by the stochastic differential equation system

$$\begin{pmatrix} dP_n(t, T_1)/P_n(t, T_1) \\ dP_n(t, T_2)/P_n(t, T_2) \\ dP_I(t, T_3)/P_I(t, T_3) \\ dP_S(t)/P_S(t) \end{pmatrix} = \mu_t dt + \Sigma_t dW_t \quad (5.13)$$

where

$$\begin{aligned} dW_t &:= (dW_t^r, dW_t^\pi, dW_t^I, dW_t^S)^\top, \text{ and} \\ \Sigma_t &:= \begin{pmatrix} -B_{nr}(T_1 - t)g_r & -B_{n\pi}(T_1 - t)g_\pi & 0 & 0 \\ -B_{nr}(T_2 - t)g_r & -B_{n\pi}(T_2 - t)g_\pi & 0 & 0 \\ -B_{rr}(T - t)g_r & 0 & \sigma_I & 0 \\ 0 & 0 & 0 & \sigma_S \end{pmatrix}. \end{aligned} \quad (5.14)$$

According to the no-arbitrage conditions (4.31), (4.32), and the stock dynamics (4.57), the expected instantaneous returns in the vector process μ_t in equation (5.13) are given by

$$\mu_t = R_t \mathbf{1} + \Sigma_t \lambda \quad (5.15)$$

where R_t is given in equation (4.40), $\mathbf{1} := (1, 1, 1, 1)^\top$ and $\lambda := (\lambda_r, \lambda_\pi, \lambda_I, \lambda_S)^\top$.

We remark that we employ the original no-arbitrage conditions (4.31), (4.32) because they support a closed form solution for the optimal intertemporal portfolio.

Here we give the explicit expression for the optimal intertemporal portfolio for the foregoing investment scenario according to the result given in (3.18).

Property 12 *Using the expectation operator representation (3.25), the value function $\Phi(t, T, X_t)$ for the investment opportunity described above is given by*

$$\Phi(t, T, r_t, \pi_t) = e^{\frac{1-\gamma}{\gamma} B_r(T-t)r_t} \Psi(t, T), \quad (5.16)$$

where

$$\begin{aligned} & \Psi(t, T) \\ = & \exp \left(j(T-t) + \frac{1-\gamma}{\gamma} (T-t - B_r(T-t)) (\bar{r} + \hat{z}_1 \frac{g_r}{\kappa_r}) \right. \\ & \left. + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} \right)^2 \left(\frac{g_r}{\kappa_r} \right)^2 \left(T-t - 2B_r(T-t) + \frac{1 - e^{-2\kappa_r(T-t)}}{2\kappa_r} \right) \right), \end{aligned}$$

where

$$j = -\frac{\delta}{\gamma} + \frac{1-\gamma}{2\gamma^2} \lambda^\top \mathcal{R}_{AA}^{-1} \lambda + \frac{(1-\gamma)\sigma_I^2}{2\gamma^2} - \frac{1-\gamma}{\gamma^2} \lambda_I \sigma_I \quad (5.17)$$

$$z = \frac{1-\gamma}{\gamma} \begin{pmatrix} \lambda_r - \sigma_I \rho_{Ir} \\ \lambda_\pi - \sigma_I \rho_{I\pi} \end{pmatrix}, \quad (5.18)$$

and

$$B_r(T-t) = \frac{1 - e^{\kappa_r(T-t)}}{\kappa_r}. \quad (5.19)$$

The notation \hat{z}_1 is the first element in \hat{z} where

$$\hat{z} := \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} := \mathcal{C}^{-1} z$$

with \mathcal{C} the lower-triangular Cholesky decomposition of \mathcal{R}_{XX} ($\mathcal{C}\mathcal{C}^\top = \mathcal{R}_{XX}$). For the investment environment described above, $W_t^X = (W_t^r, W_t^\pi)^\top$, so

$$\mathcal{R}_{XX} = \begin{pmatrix} 1 & \rho_{r\pi} \\ \rho_{r\pi} & 1 \end{pmatrix}.$$

After having obtained the value function Φ , we still need to obtain the factor elasticity Φ_X/Φ .

Property 13 *The factor elasticities are given by*

$$\begin{pmatrix} \frac{\Phi_r}{\Phi} \\ \frac{\Phi_\pi}{\Phi} \\ \frac{\Phi}{\Phi} \end{pmatrix} = \begin{pmatrix} \frac{1-\gamma}{\gamma} \frac{B_r(T-t)}{T-t} \\ 0 \end{pmatrix}. \quad (5.20)$$

Property 13 is proved simply by differentiating $\Phi(t, T, r_t, \pi_t)$ given in (5.16).

The parameter κ_r here is the mean-reverting parameter for the real interest rate r_t . It is worth noticing that the value function $\Phi(t, T, r_t, \pi_t)$ does not depend on the level of the expected inflation rate π_t . To understand this result we can consider the following argument. The value function symbolizes the best achievement of the long-term expected utility of the *real* consumption. Recalling the real wealth evolution (3.9) and rewriting it using the no-arbitrage condition (5.15), we obtain the evolution dynamics

$$\begin{aligned} \frac{dv_t}{v_t} &= (R_t - \pi_t + \sigma_I^2)dt \\ &+ (\Sigma_t^\top \alpha_t)^\top (\lambda - \sigma_I \mathcal{R}_{AI})dt + (\Sigma_t^\top \alpha_t)^\top dW_t - \sigma_I dW_t^I, \end{aligned} \quad (5.21)$$

where the factors involve the real wealth evolution in the term $R_t - \pi_t$. Considering further the riskless rate formula (4.40), this term is replaced by

$$R_t - \pi_t = r_t + \xi_0,$$

so that the real wealth evolution now is characterized only by one of the factors, r_t . A more detailed and technical explanation can be found in the proof of Property 13 in the Appendix 8.3.

Using the result of Property 13, we obtain the optimal strategies of the intertemporal investment plan.

Property 14 *The optimal investment proportions are given by*

$$\alpha_t := \begin{pmatrix} \alpha_{1t} \\ \alpha_{2t} \\ \alpha_{3t} \\ \alpha_{4t} \end{pmatrix} = \frac{1}{\gamma} \underbrace{(\Sigma_t^\top)^{-1} \mathcal{R}_{AA}^{-1} \lambda}_{I. \alpha_t^{(M)}} + \underbrace{\left(1 - \frac{1}{\gamma}\right) (\Sigma_t^\top)^{-1} \begin{pmatrix} -g_r \frac{B_r(T-t)}{T-t} \\ 0 \\ 0 \\ 0 \end{pmatrix}}_{II. \alpha_t^{(I)}} + \underbrace{\left(1 - \frac{1}{\gamma}\right) (\Sigma_t^\top)^{-1} \begin{pmatrix} 0 \\ 0 \\ \sigma_I \\ 0 \end{pmatrix}}_{III. \alpha_t^{(P)}}, \quad (5.22)$$

where $B_r(T-t)$ is as same as (5.19).

We remark that the order of the investment proportions $(\alpha_{1t}, \alpha_{2t}, \alpha_{3t}, \alpha_{4t})^\top$ is identical with the order in the equation system (5.13) so that

α_{1t} represents the investment proportion in the nominal bond maturing at T_1 ,
 α_{2t} represents the investment proportion in the nominal bond maturing at T_2 ,
 α_{3t} represents the investment proportion in the IIB maturing at T_3 , and
 α_{4t} represents the investment the stock

respectively.

The three-part structure of the optimal intertemporal portfolio $(\alpha_t^{(M)}, \alpha_t^{(I)}, \alpha_t^{(P)})$ under inflation risk in Property 14 has already been mentioned in the formula (3.17) in the general framework.

We lay out in more detail the intertemporal hedging term and the inflation hedging term in the following property

Property 15 *The intertemporal and inflation hedging portfolios are given by*

$$\alpha_t^{(I)} = \begin{pmatrix} \mathcal{D}^{-1} B_{n\pi}(\tau_2) B_r(\tau) \\ -\mathcal{D}^{-1} B_{n\pi}(\tau_1) B_r(\tau) \\ 0 \\ 0 \end{pmatrix}, \quad \alpha_t^{(P)} = \begin{pmatrix} -\mathcal{D}^{-1} B_{n\pi}(\tau_2) B_{rr}(\tau_3) \\ \mathcal{D}^{-1} B_{n\pi}(\tau_1) B_{rr}(\tau_3) \\ 1 \\ 0 \end{pmatrix}, \quad (5.23)$$

where $\tau = T - t$, $\tau_i = T_i - t$ for $i = 1, 2, 3$ and

$$\mathcal{D} := \det \begin{pmatrix} B_{nr}(\tau_1) & B_{nr}(\tau_2) \\ B_{n\pi}(\tau_1) & B_{n\pi}(\tau_2) \end{pmatrix}. \quad (5.24)$$

According the result of Property 15, because all coefficients $B_{**}(\tau_i)$ are positive within Model II, the sign of the hedging positions in the intertemporal hedging portfolio $\alpha^{(I)}$ and the inflation hedging portfolio $\alpha^{(P)}$ depend on the sign of the determinant \mathcal{D} . We can characterize the conditions for the sign of the determinant \mathcal{D} in Property 16

Property 16 *For $\tau_1 < \tau_2$, we have*

$$\mathcal{D} \begin{matrix} > \\ < \end{matrix} 0 \iff \begin{matrix} \kappa_r > \kappa_\pi \\ \kappa_r < \kappa_\pi \end{matrix}.$$

Similarly to in Model I we define the *conservative* portfolio as the sum of the intertemporal hedging and price hedging terms,

$$\text{Conservative Portfolio} := \alpha_t^{(I)} + \alpha_t^{(P)}, \quad (5.25)$$

so we obtain also the decomposition for the optimal portfolio α_t :

$$\alpha_t = \frac{1}{\gamma} \text{Myopic Portfolio} + \left(1 - \frac{1}{\gamma}\right) \text{Conservative Portfolio} . \quad (5.26)$$

According to Property 15, the conservative portfolio is given by

Property 17 *The conservative portfolio investing in the assets given in the stochastic differential equation system (5.13) is given by*

$$\text{Conservative Portfolio} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} ,$$

which is obtained simply by adding the two hedging portfolios up.

This result means that, in an investment environment with inflation risk, the most risk averse investors put all the wealth in the IIB which matures at the end of the investment horizon. This is the extension of the case illustrated in Table 5.5 where the most conservative investors only buy the nominal bond maturing at the end of the horizon when the investment environment is free of inflation risk. Those two results are based on the same intuition that the most conservative investors require a certain payout at the end of the investment. It is clear that the IIB, instead of the nominal bond, guarantees a certain payout when the investment is exposed to inflation risk.

As a comparison we also provide the optimal intertemporal portfolio without an investment opportunity in the IIBs.

Property 18 *The factor elasticities for the intertemporal investment decision without an investment opportunity in IIBs are identical to those given in (5.20) with an investment opportunity in IIBs.*

Property 18 claims that the formulas for the factor elasticity for the value function are the same regardless of the inclusion of the IIBs in the investment opportunity set. We might understand this result using the same intuition for Property 13. We provide more detailed and technical details in the proof of this Property in the Appendix 8.3 where we can see that although the optimal decisions of the investment problem with and without the IIBs will achieve different expected utility, that means, their value functions will take different values, how the value functions depend on the real interest rate r_t are the same therefore they have the same outcome of the factor

elasticity. Having obtained the formula of the factor elasticity, the solution of the optimal investment weights is then simply applied.

Property 19 *The optimal portfolio weights in the case without the investment opportunity in the IIBs are given by*

$$\alpha_t^* = \underbrace{\frac{1}{\gamma} (\Sigma_t^\top)^{-1} \mathcal{R}_{AA}^{-1} \begin{pmatrix} \lambda_r \\ \lambda_\pi \\ \lambda_S \end{pmatrix}}_{I. \alpha_t^{(M)}} + (1 - \frac{1}{\gamma}) (\Sigma_t^\top)^{-1} \underbrace{\begin{pmatrix} -g_r \frac{B_r(T-t)}{T-t} \\ 0 \\ 0 \end{pmatrix}}_{II. \alpha_t^{(I)}} + (1 - \frac{1}{\gamma}) (\Sigma_t^\top)^{-1} \sigma_I \mathcal{R}_{AA}^{-1} \underbrace{\begin{pmatrix} \rho_{rI} \\ \rho_{\pi I} \\ \rho_{SI} \end{pmatrix}}_{III. \alpha_t^{(P)}}, \quad (5.27)$$

where $B_r(T-t)$ as same as (5.19).

Without the investment opportunity in the IIBs, the risk of the stochastic price index W_t^I can only be hedged by its correlations with the other risky assets, as shown in the third term III. $\alpha^{(P)}$ in the formula (5.27). Without the IIBs, the financial market exposed to the inflation risk is incomplete, no assets can give a certain payout. Therefore, there is no longer certain strategy for the most risk averse agents and they can only partially hedge the systematic risk by utilization of correlations of asset returns.

Since the factor elasticity without IIB as given in Property 18 is the same as that with IIB, and since the intertemporal hedging term II. $\alpha^{(I)}$ in the optimal portfolio (5.27) is closely related to the factor elasticity, we can expect that the intertemporal hedging term in the case without IIB is very similar to that with IIB.

Property 20 *The intertemporal hedging portfolio in the case without IIB is given by*

$$\alpha_t^{(I)} = \begin{pmatrix} \mathcal{D}^{-1} B_{n\pi}(\tau_2) B_r(\tau) \\ -\mathcal{D}^{-1} B_{n\pi}(\tau_1) B_r(\tau) \\ 0 \end{pmatrix}, \quad (5.28)$$

where $\tau = T - t$, $\tau_i = T_i - t$ for $i = 1, 2$ and \mathcal{D} is defined as (5.24).

5.3.2 Example: The Hedging Effect of Inflation-Index Bonds

This section provides concrete investment recommendations for the strategies including investing IIBs. We are interested in studying hedging effect of the IIBs.

We consider four risky assets in the investment opportunity set: a three-year nominal bond (NB3Y), a 10-year nominal bond (NB10Y), a 10-year

IIB and a stock whose return dynamics are summarized in (5.13) in Section 5.3.1. The parameter values for this example are adopted from the estimation results obtained in Sections 4.4.2, 4.4.5, and 4.4.6. We summarize the relevant parameter values for the optimal investment strategies in Table 5.7. Figure 5.6 plots the optimal portfolio weights against the risk aversion

$$\begin{aligned}
 \kappa_r &= 0.1241, & \bar{r} &= 0.0040, & g_r &= 0.0101 \\
 \kappa_\pi &= 0.4016, & \xi_0 &= -0.0012, & g_\pi &= 0.0067 \\
 \\
 \sigma_S &= 0.1391, & \sigma_I &= 0.0115, \\
 \\
 \lambda_r &= -0.5168, & \lambda_\pi &= -1.5681, \\
 \lambda_I &= 0.1014, & \lambda_S &= 0.8669, \\
 \\
 \rho_{\pi r} &= -0.5082, \\
 \rho_{I r} &= 0.0609, & \rho_{I \pi} &= -0.0688, \\
 \rho_{S r} &= 0.1744, & \rho_{S \pi} &= -0.0221, & \rho_{S I} &= -0.0587
 \end{aligned}$$

Table 5.7: Parameter summary for Model II

parameter $\gamma \in [4, 1000]$. The investment horizon is ten years. In Fig. 5.6 all positions decrease in absolute value when the agents' risk aversion becomes larger with the only one exception of the IIB. To understand this result we recall the portfolio decomposition (5.22) and present the weights of each portfolio in Table 5.8. As the risk aversion γ increases, the optimal portfolio converges to the conservative portfolio as shown in (5.26). According to Property 17, the conservative portfolio invest all the wealth in the IIB. Further, we look at the intertemporal and inflation hedging portfolios in the conservative portfolio. The sign of the intertemporal hedging position is explained by Properties 15 and 16. In our case we have $\kappa_r < \kappa_\pi$ from the estimation result, so the holders of the intertemporal hedging portfolio prefer a long position in the long-term bond and a short position in the short-term bond. The exact amounts are given in Table 5.8. Table 5.8 shows that the myopic portfolio I. $\alpha^{(M)}$ has very extreme positions for the two nominal bonds. This might be explained by the two following facts. First, the risk compensation ratios $(\mu - R)/\sigma^2$ in Table 5.2 of these two nominal bonds ($= 66.03$ and $= 21.48$) are much higher than those of the IIB ($= 9.15$) and the stock ($= 6.23$). Relating this risk compensation ratio to the myopic

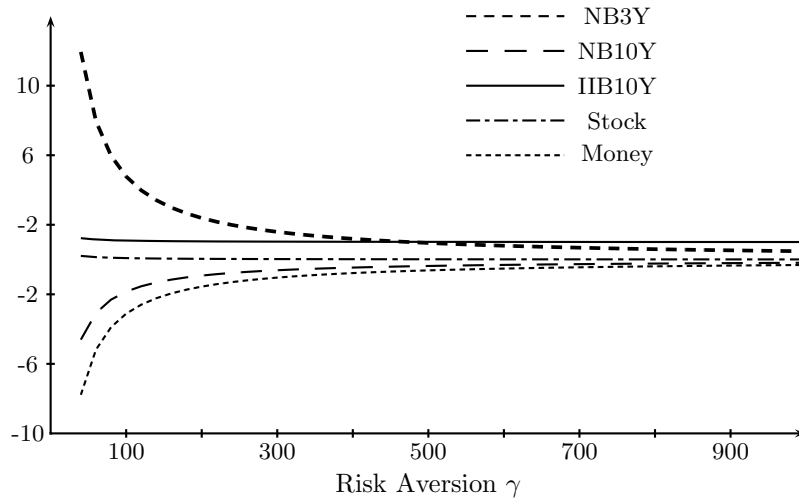


Figure 5.6: Optimal Portfolio Weights for Model II with IIB

	I. $\alpha^{(M)}$	II. $\alpha^{(I)}$	III. $\alpha^{(P)}$	Conserv
NB3Y	477.72	-3.64	3.64	0.00
NB10Y	-184.27	2.59	-2.59	0.00
IIB10Y	10.20	0.00	1.00	1.00
Stock	8.42	0.00	0.00	0.00
Money	-311.08	2.04	-1.04	0.00

Table 5.8: Decomposition of Portfolio for Model II with IIB

portfolio

$$\text{Myopic Portfolio} = (\Sigma_t \mathcal{R}_{AA} \Sigma_t^\top)^{-1} (\mu_t - R_t \mathbf{1})$$

in the portfolio formula in (3.17), the risk compensation represent the myopic portfolio when the market consists of only that one asset. (Recall the partial equilibrium argument in Section 5.1.) So, from the view of the risk compensation, agents hold large amounts of nominal bonds. Second, the correlations between the bonds are quite high as given in

$$\text{Cor}(\text{NB3}, \text{NB10}) = 0.92 \quad \text{Cor}(\text{NB3}, \text{IIB10}) = 0.81 \quad \text{Cor}(\text{NB10}, \text{IIB10}) = 0.97 .$$

The high correlation between the two nominal bond provides an excellent opportunity to get rid the return risk by a "long one and short the other" strategy. Although the IIB is also highly correlated with the long-term

nominal bond, it has a more moderate position as given in Table 5.8 because the IIB is not only considered for hedging the return risk but also for hedging (realized) inflation risk.

The optimal portfolio strategies without the opportunity to invest in IIBs are shown in Fig. 5.7. The message from the figure is clear: without the investment opportunity in IIBs, more risk averse agents revert to demanding the long-term bond.

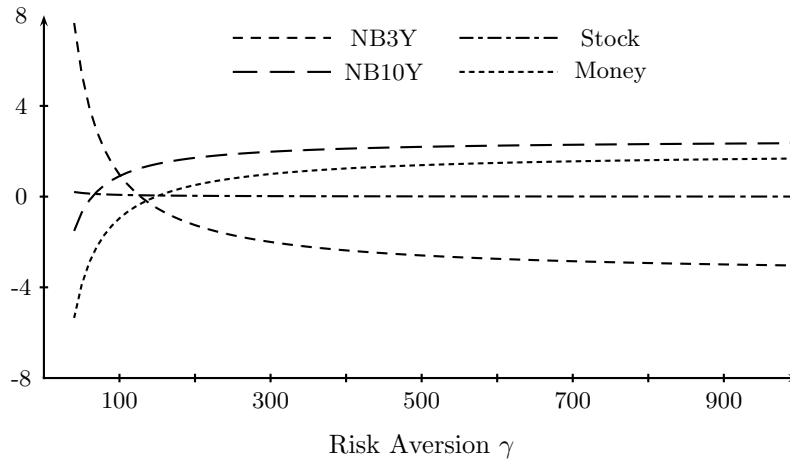


Figure 5.7: Optimal Portfolio Weights for Model II without IIB

We give exact values of each portfolios in Table 5.9.

	I/Myopic	II	III	Conserv
NB1Y	442.17	-3.64	0.151	-3.49
NB20Y	-115.60	2.59	-0.076	2.51
Stock	8.36	0.00	-0.006	-0.006
Cash	-290.94	2.04	0.931	1.97

Table 5.9: Portfolio Decomposition without IIB

Comparing between the intertemporal and the inflation hedging portfolios, the first one dominates in the conservative portfolio. The intertemporal hedging portfolio has a long position in the long-term bond and a short-position in short-term bond because $\kappa_r < \kappa_p i$ according to Property 19 and

Property 16. Recall Property 20 that the holding amounts the two nominal bonds in the intertemporal hedging portfolio are just the same as those in the case with IIB given in Property 15. The inflation hedging portfolio is relatively weak where without IIBs agents can only hedge the (realized) inflation risk through the correlation between asset returns and the price index change.

Both examples in our intertemporal framework, with and without IIBs, can explain the investment puzzle raised by Canner, Mankiw and Weil (1997) where the bond-to-stock ratio increases with risk aversion. In our examples, the stock has no hedging function at all in the case with IIBs and a very weak hedging function in the case without IIBs. Therefore the investment portion in stock decreases with increasing risk aversion and the bond-to-stock ratio goes up.

We also wish to examine the investment horizon effects. The risk aversion is fixed at $\gamma = 70$ and the investment horizon goes from 4 to 30 years. We let the IIB and the long-term nominal bond mature when the investment ends. Figure 5.8 shows that in the case with IIB, positions in absolute value in the both nominal bonds decrease when the investment horizon increases, while those in the IIB and stock remain constant. This can be directly explained by Property 14. We can also obtain the limit positions $\bar{\alpha}_i$ where $\tau_2 = \infty, \tau_3 = \infty$ and they are given by

$$\bar{\alpha}_1 = 5.16 \quad \bar{\alpha}_2 = -1.42 \quad \bar{\alpha}_3 = 1.13 \quad \bar{\alpha}_4 = 0.12 \quad \bar{\alpha}_5 = -3.99 .$$

The horizon effect for the case without IIB is shown in Figure 5.9. The amount of short-term bond demanded decreases when the horizon increases. The stock demand is still kept as constant while the position of the long-term bond change sign when the horizon becomes longer. We also provide the limit positions

$$\bar{\alpha}_1 = 2.30 \quad \bar{\alpha}_2 = 0.62 \quad \bar{\alpha}_3 = 0.11(\text{stock}) \quad \bar{\alpha}_4 = -2.03(\text{money}).$$

Our result is different to that of Brennan and Xia (2002) because they fixed the bond maturity while varying the horizon length.

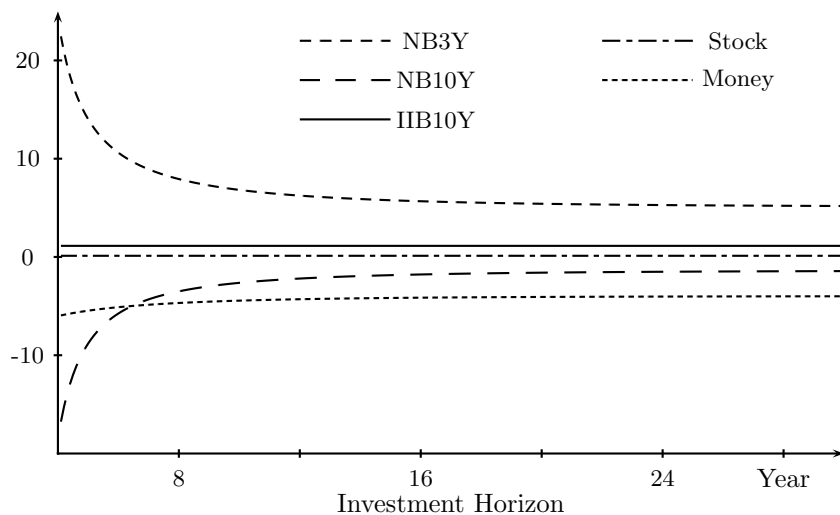


Figure 5.8: Optimal Weights for Model II, Horizon Effect, with IIB

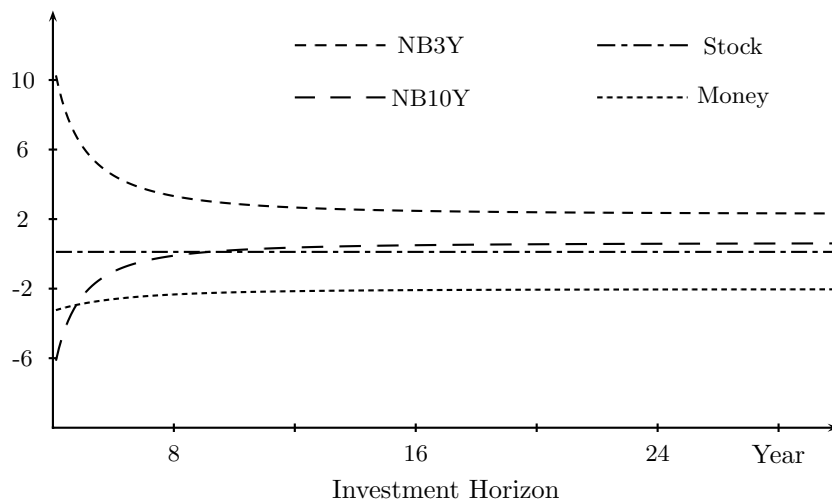


Figure 5.9: Optimal Weights for Model II, Horizon Effect, without IIB

5.4 Summary

Our simulation results show that taking account of the intertemporal features of the investment strategy does really bring about extra profit when compared to the strategies that ignore it. General speaking, the risk aversion parameter turns out to be an important characteristic of intertemporal portfolios. The less risk averse agents are more concerned with the risk-return trade-off, while the more risk averse agents give priority to certainty of the payout for the cases both with and without inflation risk. When considering hedging strategies, the presence of inflation risk does matter. In a world without inflation risk, the nominal bond maturing at the final day is an ideal hedging asset because it can provide a certain payout when the investment ends, as mentioned in Wachter (2003). However, when the investment is exposed to inflation risk, the role of this long-term nominal bond will be taken over by the IIB maturing at the final day based on the same reasoning. Furthermore, when the IIBs are not available for hedging inflation risk, agents will go back to demand the long-term bonds in our case.

Similar to Campbell and Viceira (2001), and Brennan and Xia (2002), the positions of the bond holding or the short positions are large, especially in the myopic portfolios. Such recommendations would not be practical because such an extreme investment strategy as 100 times the whole wealth, could not be accepted in real world situations. This leads us to incorporate into our modelling framework real market features, such as short-sale constraints, in order to reduce the investment recommendations to within a reasonable range. We will consider several short-sale constraints and study their impact in the next chapter by means of computational methods.

Chapter 6

Portfolio Strategies under Short-Sale Constraints

In this chapter we shall allow for short-sale constraints in the asset allocation problem, in order to account for the real-world trading environment. The intertemporal control problem can readily be handled when short-sale constraints are considered. However it turns out that to solve the intertemporal control problem, in this case, one has to resort to computational methods. Based on the Markov chain approximation (MCA) method proposed by Kushner (1977,1999), we develop a backward iteration scheme to evaluate the value function for finite time horizon.

Short-sale constrains will change wealth dynamics along the investment path. We wish to consider three different short-sale constraints (SSC).

SSC-1. imposes an additional payment of a short-sale commission η for each unit of short position

SSC-2. excludes short positions in risky assets, and

SSC-3. excludes all short positions and also the possibility of borrowing money.

For the first short-sale constraint SSC-1, the real wealth evolution has been given in equation (3.10) in Section 3.1, which is

$$\begin{aligned} \frac{dv_t}{v_t} = & \left[(R_t - \psi_t - \pi_t + \sigma_I^2) + \sum_{i=0}^m \min(0, \alpha_{it}) \eta \right. \\ & \left. + \alpha_t^\top (\mu_t - R_t \mathbf{1} - \sigma_I \Sigma_t \mathcal{R}_I) \right] dt + \alpha^\top \Sigma_t dW_t - \sigma_I dW_t^I. \end{aligned}$$

For the SSC-2 constraint the investors are not allowed to short any risky assets therefore all holding positions $\{\alpha_{1t}, \dots, \alpha_{nt}\}$ are restricted to be positive. For the SSC-3 constraint not only the holding positions of the risky assets but also the money holding α_{0t} must take positive values.

6.1 The Backward Markov Chain Approximation Method

The Markov chain approximation (MCA) method solves the continuous-time stochastic control problem by approximating the original controlled process by a finite-state controlled process. A finite-state controlled process is obtained by discretising the time space, by approximating the Wiener process by symmetric random walks and by using state space grids.

At first we discretise the time space. For a finite-state process, actions take place only at discrete time points $\{k\Delta\}_{k=0,1,\dots,N_\Delta}$, where N_Δ is the greatest natural number less than T/Δ . The corresponding discrete-time model can be found in the Appendix.

To apply the MCA method, first we approximate the continuous-time value function (3.6) by the the discrete-time value function

$$J(0, T, v_0, X_0) := \max_{\substack{c_{k\Delta}, \alpha_{k\Delta}, \\ k = 1, \dots, TN_\Delta - 1}} \mathbf{E}_0 \left[\epsilon_1 \sum_{k=0}^{TN_\Delta - 1} e^{-\delta k\Delta} U(c_{k\Delta}) + e^{-\delta T} U(v_T) \right], \quad (6.1)$$

given the initial states (v_0, X_0) . The aim in applying the MCA method is to evaluate numerically the discrete-time value function (6.1).

For our backward iteration scheme we define the value function on the sub-period $[k\Delta, T]$ given the initial states $(X_{k\Delta}, v_{k\Delta})$ by

$$J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta}) := \max_{\substack{c_{k'\Delta}, \alpha_{k'\Delta}, \\ k' = k, \dots, TN_\Delta - 1}} \mathbf{E}_{k\Delta} \left[\epsilon_1 \sum_{k'=k}^{TN_\Delta - 1} e^{-\delta k'\Delta} U(c_{k'\Delta}) + e^{-\delta T} U(v_T) \right], \quad (6.2)$$

where k can be any number from $\{0, 1, \dots, N_\Delta\}$.

Due to the natural time structure, the later asset allocation decisions do not affect the earlier dynamics. So, for each subperiod optimization prob-

lem, we have the following iterative formula between the time points $k\Delta$ and $(k+1)\Delta$:

$$\begin{aligned} & J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta}) \\ &= \max_{c_{k\Delta}, \alpha_{k\Delta}} \left\{ \epsilon_1 e^{-\delta(k\Delta)} U(c_{k\Delta})\Delta + \mathbf{E}_{k\Delta} [J^\Delta((k+1)\Delta, T, v_{(k+1)\Delta}, X_{(k+1)\Delta})] \right\}. \end{aligned} \quad (6.3)$$

The first iteration begins with the initial value function

$$J^\Delta(T, T, v_T, X_T) \equiv e^{-\delta T} U(v_T)$$

inserted in the RHS of the iteration formula. The value function (6.1) evaluated as with k stepping back to 0. Precise details of the iteration scheme are given Appendix 8.1.2.

Besides the time discretisation, we approximate the Wiener processes with a binominal tree. The increment $W_{(k+1)\Delta}^X - W_{k\Delta}^X$ is approximated by a n -dimensional symmetric random walk¹ u_k^X with the same covariance \mathcal{R}_{XX} .

To obtain a finite-state controlled process we still need to discretize the state space. Also, we employ the truncation technique of Camilli and Falcone (1995), which truncates the control problem on a compact state space which is "large enough". Then, we take cuboidal grids on the chosen compact set and use the multilinear interpolation for the value function as described in Gruene (2001).

Conditions for applying the MCA method are discussed in Appendix 8.1.4. We also remark that the backward iteration scheme (6.3) is for the finite time horizon problem. For an infinite time horizon problem the Jacobi iteration scheme is adopted. The details are provided in Appendix 8.1.3.

6.2 Optimal Portfolio Application with Short-Sale Constraints

The backward iteration scheme for the MCA method is applied to study the impact of the short-sale constraints on the optimal intertemporal portfolios. Our computational scheme consists of two parts. In the first part,

¹Each component of u_k^X has the probability distribution $\mathcal{P}(u_{ki}^X = \pm\sqrt{\Delta}) = \frac{1}{2}$.

we investigate the performance of the backward MCA method with different time steps and grid sizes for the investment problem without short-sale constraints. In this case the analytical solutions are available so that we can determine the discretization errors for the MCA method. In the second part, we apply the MCA method to find optimal strategies under the short-sale constraints SSC-1, SSC-2, and SSC-3 described in the beginning of this Chapter. The software we use to implement the numerical algorithm is GAUSS with the two application packages OP_3.1_DOS for optimization and CO_2.0_DOS for constrained optimization.

The parameters used for the numerical study are given in Table 6.1, and are suggested by the estimation results in the subsections 4.4.2 and 4.4.4 in Chapter 4.

$$\begin{aligned} \kappa_r &= 3.00, & \bar{r} &= 0.0014, & g_r &= 0.01 \\ \kappa_\pi &= 0.50, & \bar{\pi} &= 0.035, & g_\pi &= 0.0064 \\ \\ \sigma_S &= 0.14, & \sigma_I &= 0.01, \\ \\ \lambda_r &= -0.13, & \lambda_\pi &= -0.57, \\ \lambda_I &= 0.64, & \lambda_S &= 0.87, \\ \\ \rho_{\pi r} &= -0.02, \\ \rho_{I r} &= 0.07, & \rho_{I \pi} &= -0.02, \\ \rho_{S r} &= 0.17, & \rho_{S \pi} &= 0.10, & \rho_{S I} &= -0.06 \end{aligned}$$

Table 6.1: Parameters Used for Simulation Examples with Short-Sale Constraints

We truncate our control problem on the compact set $-2\% \leq r_t \leq 2\%$. The invariant distribution of the process r_t defined in (4.16) has a standard deviation of 0.004082 ($= g_r^2/2\kappa_r$) and according to this distribution the probability of being out of this compact set is extremely low ($9.6 \cdot 10^{-7}$). The investment horizon is set to be 5 years. The two nominal bonds for investment have 2 and 5 years maturity. When the first bond matures at the end of the second and the fourth years, another 2-year bond will be introduced immediately. The relative risk aversion parameter is set at $\gamma = 4.0$. The subjective discount factor δ is chosen as $\delta = 0.02$. We consider the case without intermediate consumption $\epsilon_1 = 0$.

6.2. OPTIMAL PORTFOLIO APPLICATION WITH SHORT-SALE CONSTRAINTS 123

We solve this stochastic control problem with different time steps $\Delta = 0.5, 0.05, 0.005$ and different grid sizes $\Delta r = 0.002, 0.0002$. Table 6.2 gives the average errors of the value function of the numerical solutions compared to the corresponding theoretical solution given in Property 12. Average errors of the optimal portfolio decisions with respect to the theoretical solution given in Property 14. The convergence criterion for the gradients in the numerical optimization was set at 10^{-8} .

	$\Delta_{state} = 0.002$	$\Delta_{state} = 0.0002$
$\Delta_{time} = 0.5$	$-1102 \cdot 10^{-5}$	$-1012 \cdot 10^{-5}$
$\Delta_{time} = 0.05$	$-104 \cdot 10^{-5}$	$-104 \cdot 10^{-5}$
$\Delta_{time} = 0.005$	$-4 \cdot 10^{-5}$	$-17 \cdot 10^{-5}$

Table 6.2: Average Errors of the Value Function with Parameters given in Table 6.1

		$\Delta_{state} = 0.002$	$\Delta_{state} = 0.0002$
$\Delta_{time} = 0.5$	α_1	48%	49%
	α_2	11%	12%
	α_3	13%	13%
	α_4	47%	47%
$\Delta_{time} = 0.05$	α_1	1.72%	1.80%
	α_2	0.18%	0.22%
	α_3	0.87%	0.86%
	α_4	3.74%	2.80%
$\Delta_{time} = 0.005$	α_1	4.70%	1.24%
	α_2	1.95%	0.56%
	α_3	0.10%	0.10%
	α_4	0.26%	0.26%

Table 6.3: Average Relative Errors of the Portfolio Decisions with Parameters given in Table 6.1

We can see in Tables 6.2 and 6.3 that the performance of the numerical algorithm improves when the time step reduces while it does not improve much when the grid size decreases. When comparing between different time

steps, by the refinement of the time step from 0.05 to 0.005, Table 6.2 shows the fitting of the value function becomes better while the fitting of the optimal portfolios in Table 6.3 does not improve much. Considering the trade-off between numerical precision and calculation cost we have decided to choose the time step $\Delta_{time} = 0.05$ and the grid size $\Delta_{state} = 0.002$ and implement our numerical process to solve for investment strategies with short-sale constraints.

Recall the three short-sale constraints:

- SSC-1.** Short-sale commissions: investors have to pay an additional η units of commission for each unit short position. In our example η is equal to 0.0002.
- SSC-2.** Short-sale exclusion for all risky assets
- SSC-3.** Short-sale exclusion for all assets including the money market account.

Table 6.4 gives the average investment proportions under the short-sale constraints. Recall from Property 14 that the theoretical values of α are independent of the state variable r_t .

	α_1	α_2	α_3	α_4
Theoretical values	-13.86	23.74	18.63	1.86
Numerical values	-13.62	23.70	18.79	1.93
SSC-1 commissions	0.00	13.93	18.29	1.92
SSC-2 exclusion, risky	6.32	9.53	18.29	1.93
SSC-3 exclusion, all	0.00	0.00	0.00	1.00

Table 6.4: Effect of Short-Sale Constraints on Portfolio Allocations

All short-sale constraints change significantly the decisions concerning the optimal portfolio positions. When the short-sale commission is introduced as in the example, we can see in the line “SSC-1 commissions” the agents do not purchase the short-term bond, which the agents would sell short if there were no commissions. If the short-sale possibility is excluded for the risky assets as shown in “SSC-2”, the agents reduce their holding in the nominal bonds while keeping their positions in the inflation-indexed bond and the stock. If now the short-sale possibility is excluded for all assets, our agents only wish to hold the stock.

To analyze further the effect of the commission fees on the portfolio decisions, we provide simulation results with different fees in Table 6.5. As expected, the higher is the commission fee, the smaller is the short position α_1 (of the short-run bond). We observe also that large positions in the long-run bond α_2 decrease with the commission fees. Whereas the positions of the inflation-indexed bonds and the stock are not changed by the introduction of the commission fees.

comm. fees	α_1	α_2	α_3	α_4
$\eta = 0$	-13.62	23.70	18.79	1.93
$\eta = 0.0001$	-5.65	18.00	18.52	1.92
$\eta = 0.0002$	0.00	13.93	18.29	1.92

Table 6.5: Effect of Commission Fees on Portfolio Decisions

We analyze the SSC-3 portfolio decision further. In our numerical example the stock has a slightly higher market price of risk $\lambda_S = 0.87$ and a significantly higher volatility $\sigma_S = 0.14$ in comparison with the other volatilities g_r, g_π in Table 6.1. Therefore, the stock has a relatively higher excess return $\lambda_S \sigma_S$. Hence, a risk-friendly investor would wish to invest in the stock. Given this consideration we would expect the stock holding of a risk averse agent to be smaller. We increase the agents' risk aversion from $\gamma = 4$ to $\gamma = 15$ and $\gamma = 45$. The optimal strategies, both with and without the short-sale exclusion, are given in Table 6.6. Under short-sale exclusion for all assets, the agents increase their holding of the inflation-indexed bonds, due to the fact that they are considered as a hedging asset, and so the agents correspondingly decrease their stock holding.

Risk Aversion	SSC-3	α_1	α_2	α_3	α_4
$\gamma = 4$	without	-13.86	23.74	18.63	1.86
$\gamma = 15$	without	-3.69	6.33	5.70	0.50
$\gamma = 45$	without	-1.23	2.11	2.57	0.17
$\gamma = 4$	with SSE	0.00	0.00	0.00	1.00
$\gamma = 15$	with SSE	0.00	0.09	0.52	0.39
$\gamma = 45$	with SSE	0.00	0.05	0.82	0.13

Table 6.6: Risk Aversion and Short-Sale Exclusion (SSE).

6.3 Summary

This chapter solves the decision problem of intertemporal portfolios under inflation risk. Several short-sale constraints are considered and the optimal intertemporal portfolios under short-sale constraints are solved by means of a numerical method – the backward MCA (Markov chain approximation) method. In the case without short-sale constraints we have an analytical solution. Using this knowledge we can choose an “optimal” discretization of time and state spaces, with a view to both precision and numerical cost for implementing the numerical method. We find that all three short-sale constraints have a significant impact on the intertemporal portfolio decisions.

Chapter 7

Conclusions

The objective of this dissertation has been to give optimal intertemporal investment strategy recommendations when the investment environment is exposed to inflation risk. We have extended Merton's continuous-time model of the intertemporal asset allocation problem to accommodate a stochastic price index. One of important features of this extension is that the stochastic price index gives rise to both real terms and nominal terms. So, in the extended model the agents maximize their expected life-time utility of consumption in real terms while agents' investment activities in the financial market and the evolution of agents' wealth are evaluated in nominal terms.

We have extended Merton's solution method using dynamic programming to solve the intertemporal asset allocation problem in the extended model with the stochastic price index. By use of the Feynman-Kac formula to solve the HJB equation that arises in the dynamic programming approach, we have developed an expectation operator representation for the value function. Also, we have provided another way than that of Cox and Huang (1989) to link the dynamic programming and the static variational method of Cox and Huang (1985) through the expectation operator representation under some assumptions.

Bonds with different time to maturities, a stock and inflation-indexed bonds have been considered in the investment opportunity set of the agents. The risk of interest rate variation and inflation risk are systematic risk sources in the investment environment. Two different kinds of term structure models of interest rates have been provided within the Gaussian Duffie-Kan model framework. The first is a data-oriented model where the underlying factors

affecting bond pricing are not specified *a priori* but are filtered from market bond yields in the empirical study. The second model is a theory-oriented model where the two underlying factors are specified as the instantaneous real interest rate and the instantaneous expected inflation rate. The two models have been estimated in Chapter 4. Using two different models to analyze the same bond yields, we have found that the *econometric* factors in the first model share some similarity of the economic factors, which are the real interest rate and expected inflation rate, in the second model. Another finding is that the market bond yield data suggest adding more flexibility in the second model so that the model can better capture two important features: the mean reversion of the factors and the bond yield sensitivity with respect to the factors.

Based on the result of the expectation operator representation in Chapter 3 and the estimation results of the market data in Chapter 4, we were able to give explicit forms for the optimal intertemporal investment strategies in Chapter 5. Based on the first model, we have provided a simulation example to illustrate the advantage of considering the intertemporal feature in the investment environment over the conventional risk-return trade-off strategy for constructing bond portfolios. Moreover we found, by conducting bond asset management, that the market sample mean and variance might not provide the required information for constructing the optimal investment strategies, because the dynamics of bond returns vary with their time to maturity.

We have also examined the risk aversion effect and horizon effect on optimal intertemporal portfolios. Agents with different degrees of risk aversion in the model have different preferences for the assets. The less risk averse agents determine their investment strategies by relying more on the risk-return trade-off while the more risk averse agents are more concerned about certainty of the final payment at the end of the investment horizon. In the case without inflation risk, the most risk averse agents only invest in the long-term bond maturing at the final date, which guarantees a certain payout at the end of the investment. When the investment environment is exposed to inflation risk, we find that the most risk averse agents now only hold the inflation-indexed bond maturing also at the final date, but not the long-term nominal bond since the final payment is affected by the uncertain price index development. In, this case, when the market does not provide the opportunity to invest in inflation-indexed bond, the most risk averse agents prefer to hold long-term bonds. With regard to the horizon effect,

when the horizon increases, the demand of the "conservative" assets, which are preferred by the more risk averse agents, increases.

The extreme scales of the optimal investment strategies suggest strongly the need to include real market frictions into the modelling framework in order to give more reasonable investment recommendations. Chapter 6 has studied the impact of short-sale constraints on the intertemporal investment problem. Since closed form solutions are no longer available, we developed a computational algorithm based on the Markov Chain Approximation method. By introduction of the short-sale commissions, the short positions are reduced. We find that the short-sale exclusion has a significant effect on the portfolio decisions. In the current market situation, aggressive agents invest more in stocks and conservative agents invest more in inflation-indexed bonds. The short-sale exclusion reduces the demand for nominal bonds in the presence of inflation risk.

Many points raised in the dissertation can be studied further in future research. The general Duffie-Kan model including the squared root volatility in the term structure model, more complicated and realistic stock asset models with stochastic volatility, time-varying risk premia and jumps, can be incorporated into the intertemporal asset allocation problem. The optimal strategy for the model with the parameter separation, either in closed form, or by use of computational algorithm needs to be developed further. Also, we wish to consider further more realistic market frictions such as transactions costs in future research.

Chapter 8

Appendix

8.1 The Discrete-Time Counterpart

Here we provide the discrete-time counterpart for the continuous-time model introduced in Chapter 3 in this dissertation. We take the view that the correspondence between the continuous-time and discrete-time frameworks can help in understanding of the intertemporal optimization in the continuous-time framework via the method of dynamic programming. Also, since the computational algorithm is based on discrete-time dynamics, this section serves as a background for developing the computational algorithm.

For a finite-state process, actions take place only at discrete time points $\{k\Delta\}_{k=0,1,\dots,N_\Delta}$. We choose Δ so that $N_\Delta := 1/\Delta$ is a natural number. The transition of the factor X_t in (3.2) is approximated by the Euler-Maruyama scheme and is denoted by

$$X_{(k+1)\Delta} = X_{k\Delta} + F(X_{k\Delta})\Delta + G(X_{k\Delta})(W_{(k+1)\Delta}^X - W_{k\Delta}^X). \quad (8.1)$$

The existence and the stationarity of the factor process, corresponding to Assumption 1 and Assumption 2 in Section 3.1, are also assumed here.

We let $\hat{X}(X_t)$ denote the discrete-time factor evolution, so that

$$\hat{X}(X_t) := X_{t+\Delta}. \quad (8.2)$$

8.1.1 The Discrete-Time Model

We are interested in the real wealth dynamics. The discrete-time real wealth dynamics are obtained by applying the Euler-Maruyama approximation in

(3.9),

$$\begin{aligned}
\frac{v_{t+\Delta}}{v_t} &= 1 + (R(X_t) - \psi_t - \pi(X_t) + \sigma_I^2)\Delta - \sigma_I \Delta W_{t+\Delta}^I & (8.3) \\
&\quad + \alpha_t^\top (\mu(X_t, t) - R(X_t)\mathbf{1} - \sigma_I \Sigma(X_t, t)\mathcal{R}_{AI})\Delta + \alpha_t^\top \Sigma(X_t, t)\Delta W_{t+\Delta} \\
&=: \hat{v}_v(X_t, \psi_t, \alpha_t, t).
\end{aligned}$$

for $t = k\Delta$, where $\Delta W_{t+\Delta} = W_{t+\Delta} - W_t$ and $\Delta W_{t+\Delta}^I = W_{t+\Delta}^I - W_t^I$. We let $\hat{v}_v(X_t, \psi_t, \alpha_t, t)$ to denote the real wealth development in proportion to its current level. By observing equation (8.3) we know that the development $\frac{v_{t+\Delta}}{v_t}$ depends on the factor level X_t , the asset allocation decisions ψ_t and α_t , and the current time t but not on its own level v_t and the price index I_t .

Based on (8.3), the (absolute) wealth development, denoted by \hat{v} , can be obtained easily

$$\hat{v}(v_t, X_t, \psi_t, \alpha_t, t) := v_{t+\Delta} = v_t \cdot \hat{v}_v(X_t, \psi_t, \alpha_t, t). \quad (8.4)$$

About no-arbitrage principle in the discrete-time model, we employ exactly Assumption 3 and Assumption 4 for its continuous-time counterpart in Section 3.1. So, the no-arbitrage condition for the discrete-time model is given by

$$\lambda_t := \lambda(X_t) = \Sigma_t^{-1}(\mu_t - R_t\mathbf{1}) \quad (8.5)$$

and it is exact the same as that given in (3.5). The same argument for this principle for the continuous-time model, see Chiarella (2004) can be also employed here.

Inserting the no-arbitrage condition (3.5) into the wealth dynamics (8.3), we obtain

$$\begin{aligned}
\frac{v_{t+\Delta}}{v_t} &= 1 + (R(X_t) - \psi_t - \pi(X_t) + \sigma_I^2)\Delta - \sigma_I \Delta W_{t+\Delta}^I & (8.6) \\
&\quad + \tilde{\alpha}_t^\top \left((\lambda(X_t) - \sigma_I \mathcal{R}_{AI})\Delta + \Delta W_{t+\Delta} \right),
\end{aligned}$$

where $\tilde{\alpha}_t$ is a transformed portfolio defined by

$$\tilde{\alpha}_t := \Sigma^\top(X_t, t)\alpha_t. \quad (8.7)$$

It turns out that the RHS of the equation (8.6) does not depend on t . The reason of the independence of the wealth dynamics on t is the no-arbitrage condition (8.5). Investigating the wealth dynamics (8.6) more carefully, we see the time dependence of the wealth dynamics arises the time-varying drift coefficient $\mu(X_t, t)$ and diffusion coefficient $\Sigma(X_t, t)$ in the asset return process vanishes due to the no-arbitrage conditions (8.5), which replaces the relating terms of the drift and diffusions coefficients by a term consisting of the market price of risk $\lambda(X_t)$ which depends only on X_t by assumption but not on t . The portfolio α_t is transformed into $\tilde{\alpha}_t$ as given in equation (8.7). If there are no transaction constraints, the portfolio α_t (investment proportions in the *risky* assets) is not subject to any restrictions. Then, wealth development reached by any given portfolio decision α_t can be also reached by $\tilde{\alpha}_t$ with the transformation (8.7), so this transformation does not affect the optimization result. As a consequence, the optimized wealth development does not depend on t . In other words, the optimized wealth development does not depend on the bond maturity dates.

Using this to rewrite the discrete-time wealth evolution (8.4), we obtain

$$v_{t+\Delta} =: \hat{v}(v_t, X_t, \psi_t, \tilde{\alpha}_t), \quad \text{with } t = k\Delta. \quad (8.8)$$

The consumption and portfolio decisions are only revised at the discrete-time points and remain constant over $[k\Delta, (k+1)\Delta)$, therefore the objective function of the given finite-state process can be written as

$$\max_{\substack{c_{k\Delta}, \alpha_{k\Delta}, \\ k = 1, \dots, TN_\Delta - 1}} \mathbf{E}_0 \left[\epsilon_1 \sum_{k=0}^{TN_\Delta - 1} e^{-\delta k\Delta} U(c_{k\Delta}) + e^{-\delta T} U(v_T) \right], \quad (8.9)$$

with $c_{k\Delta} = \psi_{k\Delta} v_{k\Delta}$.

8.1.2 The Backward Iteration Formula

We employ the notations in Chapter 6.1. The function $J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta})$ as given in (6.2) is defined as the value function for a partial optimization problem over the subperiod $[k\Delta, T]$ given the initial states $(X_{k\Delta}, v_{k\Delta})$ and is expressed as

$$J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta}) := \max_{\substack{c_{k'\Delta}, \alpha_{k'\Delta}, \\ k' = k, \dots, TN_\Delta - 1}} \mathbf{E}_{k\Delta} \left[\epsilon_1 \sum_{k'=k}^{TN_\Delta - 1} e^{-\delta k'\Delta} U(c_{k'\Delta}) + e^{-\delta T} U(v_T) \right].$$

The backward iteration solution is implemented by solving the sub-period optimization problem sequentially. Due to the natural time structure, we can solve the sub-period optimization problem backwards, which starts with $t = T - \Delta$, follows with $t = T - 2\Delta$, then proceeding backwards until $t = 0$. This optimization process can be represented by the following iterative formula

$$\begin{aligned}
& J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta}) \\
&= \max_{c_{k\Delta}, \alpha_{k\Delta}} \left\{ \epsilon_1 e^{-\delta(k\Delta)} U(c_{k\Delta}) \Delta + \mathbf{E}_{k\Delta} [J^\Delta((k+1)\Delta, T, v_{(k+1)\Delta}, X_{(k+1)\Delta})] \right\} \\
&= \max_{c_{k\Delta}, \alpha_{k\Delta}} \left\{ \epsilon_1 e^{-\delta(k\Delta)} U(c_{k\Delta}) \Delta \right. \\
&\quad \left. + \mathbf{E}_{k\Delta} [J^\Delta((k+1)\Delta, T, \hat{v}(v_{k\Delta}, X_{k\Delta}, \psi_{k\Delta}, \alpha_{k\Delta}, k\Delta), \hat{X}(X_{k\Delta}))] \right\},
\end{aligned} \tag{8.10}$$

where \hat{v} and \hat{X} represent the real wealth evolution and the factor evolution defined in (8.4) and (8.2).

The first iteration is indexed by $k = TN_\Delta - 1$ and the initial value function

$$J^\Delta(T, T, v_T, X_T) \equiv e^{-\delta T} U(v_T) \tag{8.11}$$

is inserted on the RHS of the iteration formula.

We can decompose the value function J^Δ further into a multiplicative form through a simple algebraic operation.

Property 21 *If (i) the utility function is of the CRRA class, and (ii) the growth rate of real wealth is independent of the real wealth level, then the discrete-time value function defined in (6.2) has the multiplicative form*

$$J^\Delta(k\Delta, T, v_{k\Delta}, X_{k\Delta}) = e^{-\delta k\Delta} U(v_{k\Delta}) \Phi^\Delta(k\Delta, T, X_{k\Delta}), \tag{8.12}$$

for all $k = 0, 1, \dots, TN_\Delta$, with

$$\begin{aligned}
& \Phi^\Delta(k\Delta, T, X_{k\Delta}) \\
&:= (1 - \gamma) \max_{\psi_{k\Delta}, \alpha_{k\Delta}} \left\{ \epsilon_1 U(\psi_{k\Delta}) \Delta \right. \\
&\quad \left. + e^{-\delta\Delta} \mathbf{E}_{k\Delta} [U(\hat{v}_v(X_{k\Delta}, \psi_{k\Delta}, \alpha_{k\Delta}, k\Delta)) \Phi^\Delta((k+1)\Delta, T, \hat{X}(X_{k\Delta}))] \right\},
\end{aligned} \tag{8.13}$$

where \hat{X} and \hat{v}_v represent the real wealth evolution relative to its current level and the factor evolution defined in (8.2) and (8.3). The iteration is

defined through a backward scheme from $t = (k + 1)\Delta$ to $t = k\Delta$. For the first iteration with $k = TN_\Delta - 1$ we have

$$\Phi^\Delta(T, T, X_T) \equiv 1 \quad (8.14)$$

on the RHS of the iteration scheme (8.13).

Proof

We prove this property by working backwards.

The initial function for the iteration formula (8.11) satisfies the multiplicative form (8.12).

Now we assume the multiplicative form is satisfied for $t = (k + 1)\Delta$ and insert this form into the iteration formula (8.10). Because the utility function is of the CRRA class and the real wealth change rate is independent of its level, we can rewrite the iteration formula as

$$\begin{aligned} & J^\Delta(t, T, v_t, X_t) \\ = & \max_{\psi_t, \alpha_t} \left\{ \epsilon_1 e^{-\delta t} U(c_t) \Delta \right. \\ & \left. + \mathbf{E}_t [e^{-(t+\Delta)} U(v_t \hat{v}_v(X_t, \psi_t, \alpha_t, t)) \Phi^\Delta(t + \Delta, T, \hat{X}(X_t))] \right\} \\ = & e^{-\delta t} U(v_t) (1 - \gamma) \max_{\psi_t, \alpha_t} \left\{ \epsilon_1 U(\psi_t) \Delta \right. \\ & \left. + e^{-\Delta} \mathbf{E}_t [U(\hat{v}_v(X_t, \psi_t, \alpha_t, t)) \Phi^\Delta(t + \Delta, T, \hat{X}(X_t))] \right\}, \end{aligned}$$

where $t = k\Delta$.

□

The multiplicative form (8.12) has the same form as (3.13) in the continuous-time model. By comparing them, we can readily see that they are based on the same reasoning.

The advantage of the multiplicative form is that we need only to iterate Φ^Δ without considering v_t . This greatly reduces the computational burden.

8.1.3 The Jacobi Iteration for the Infinite Time Horizon Problem

In order to solve a control problem with an infinite-time horizon, we need a different solution concept to the backward recursive method applied to the finite horizon problems because we do not have a finite final time with which

we can start the backward iteration.

The value function for the infinite-horizon asset allocation problem is denoted as

$$J_{\infty}^{\Delta}(v_t, X_t) = \max_{c_{t+k\Delta}, \alpha_{t+k\Delta}, k=0,1,\dots} \mathbf{E}_t \left[\sum_{k=0}^{\infty} e^{-\delta\Delta k} U(c_{t+k\Delta}) \right], \quad (8.15)$$

for given initial real wealth v_t , initial state X_t at time t . The discrete-time dynamics are denoted by $v_{t+\Delta} = \hat{v}(v_t, X_t, \psi_t, \tilde{\alpha}_t)$ and $X_{t+\Delta} = \hat{X}(X_t)$, where $\tilde{\alpha}_t$ is a transformed portfolio described in the transformed wealth dynamics (8.7) in the Appendix. We note that the dynamics are autonomous therefore the value function J_{∞}^{Δ} in the definition (8.15) is independent of the starting time t . Also, this discrete-time value function has a discount schedule different to that of the continuous-time value function (3.11) where the utility in the discrete-time model is discounted to the time point t while in the continuous-time model the utility is discounted to the time point 0.

Under some reasonable assumptions¹, the value function satisfies the *Bellman optimality condition*

$$J_{\infty}^{\Delta}(v, X) = \max_{\psi, \tilde{\alpha}} \left\{ U(\psi v) \Delta + e^{-\delta\Delta} \mathbf{E}_t [J_{\infty}^{\Delta}(\hat{v}(v, X, \psi, \tilde{\alpha}), \hat{X}(X))] \right\}. \quad (8.16)$$

for a given starting wealth v and an initial state X at time point t .

According the Bellman optimality condition (8.16) we can solve the value function using the *Jacobi iteration*. Let \mathcal{T} denote the operator on the functional space $\mathbb{R}_+ \times \mathbb{R}^n \rightarrow \mathbb{R}$ giving by the mapping

$$\mathcal{T}(I)(v, X) := \max_{\psi, \tilde{\alpha}} \left\{ U(\psi v) \Delta + e^{-\delta\Delta} \mathbf{E}_t [I(\hat{v}(v, X, \psi, \tilde{\alpha}), \hat{X}(X))] \right\}.$$

Then, the value function is the fixed point of the operator \mathcal{T}

$$\mathcal{T} J_{\infty}^{\Delta}(v, X) = J_{\infty}^{\Delta}(v, X).$$

The Jacobi iteration is defined as

$$I^{(k+1)}(v, X) := \mathcal{T}(I^{(k)})(v, X). \quad (8.17)$$

Under some assumptions² the iterations converge to the solution

$$J_{\infty}^{\Delta}(v, X) = \lim_{k \rightarrow \infty} I^{(k)}(v, X),$$

and the limit is independent of the choice of the initial function $I^{(0)}$.

¹For details see Stocky and Lucas (1989)

²See Camilli and Falcone (1995)

8.1.4 Remarks on the Convergence of the Algorithms

Kushner and Dupuis (2000) give conditions under which the MCA method converges to its continuous-time solution as the time step and the grid size converge to zero, see pp. 70-71, p.276. Essentially these conditions are

- C1.** The approximating finite-state processes are “locally consistent”.
- C2.** The optimal control policy has a “relaxed control representation”.
- C3.** The drift and diffusion coefficients of the state variables are bounded and continuous.
- C4.** The space of control variables is compact and the utility function U is continuous and bounded

The local consistency C1 defined in p.71 Kushner and Dupuis (2001) requires that the approximating finite-state processes are close to the original process. It is automatically satisfied if we consider the Euler-Maruyama scheme. Condition C2 requires the optimal control policy to have some “nice” property such that it can be approximated by a piecewise constant and finite-valued control policy with an arbitrarily small penalty on the value function, see pp.86-87 Kushner and Dupuis (2001). We will see this convergence later in the numerical examples. With regard to Condition C3 recall that the state variables are (X_t, v_t) . The variable t satisfies C3 directly. For X_t it is also satisfied using the truncated problem of Camilli and Falcone (1995) where X_t is confined to a compact set. It is difficult to require C3 for the last state variable v_t if we do not put any constraint on the portfolio decision α_t . Then the agents are allowed to hold extreme positions which may cause extreme wealth movements. However, a rational agent will not take extreme positions but optimize her/his asset allocation according to the utility function. For our case where α_t can be solved analytically we know how to choose a compact set for α_t which includes all the maxima. Considering the stochastic control problem on this compact set the conditions C3 and C4 can then be satisfied. In the case of short-sale constraints, Conditions C3 and C4 are satisfied easily when we exclude short-sale possibilities.

8.1.5 Self-Financing and Nominal Wealth Dynamics

In this subsection we derive the discrete-time *nominal* wealth development (2.5) in the introductory example based on the self-financing constraint.

This derivation can go further to obtain the nominal wealth development (3.8) in the continuous-time framework, where we can see the evolution of the financial events more clearly. In the discrete-time model the wealth development must satisfy a *self-financing budget constraint*. The continuous-time dynamics (3.8) are obtained when the discretization step Δ goes to zero.

For notation convenience we express here the time index in brackets instead of as a subindex. Let $P_i(t)$ denote the i -th asset price realization at time t and $N_i(t)$ be the number of shares at time t . Events are assumed to happen in the following sequence. When entering the period $[t, t + \Delta]$, the agents still hold the shares from the last period $[t - \Delta, t]$. Then, the prices of risky asset return $P_i(t)$ are realized and thereafter the agents decide their consumption $C(t)$ and the holding of the shares for this period $[t, t + \Delta]$.

Their decisions have to satisfy the *self-financing budget constraint*

$$\sum_{i=0}^n N_i(t)P_i(t) + C(t)\Delta = \sum_{i=0}^n N_i(t - \Delta)P_i(t) := V(t) \quad (8.18)$$

which expresses the fact that the investment and the consumption decisions can be financed by their wealth $V(t)$, which has the value of their asset holdings with the realized prices. The consumption $C(t)$ represents the consumption amount for one unit time, so the consumption for the period $[t, t + \Delta]$ is equal to $C(t)\Delta$.

Under the self-financing constraint (8.18), we derive the wealth dynamics according to the agents' decisions according to

$$\begin{aligned} V(t + \Delta) - V(t) &= \sum_{i=0}^n N_i(t)P_i(t + \Delta) - \sum_{i=0}^n N_i(t - \Delta)P_i(t) \\ &= \sum_{i=0}^n N_i(t)(P_i(t + \Delta) - P_i(t)) + \sum_{i=0}^n (N_i(t) - N_i(t - \Delta))P_i(t) \\ &= \left(\sum_{j=0}^n N_j(t)P_j(t) \right) \sum_{i=0}^n \alpha_i(t) \frac{P_i(t + \Delta) - P_i(t)}{P_i(t)} - C(t)\Delta, \end{aligned}$$

where

$$\alpha_i(t) = \frac{N_i(t)P_i(t)}{\sum_{j=0}^n N_j(t)P_j(t)} = \frac{N_i(t)P_i(t)}{V(t) - C(t)\Delta} \quad (8.19)$$

represents the investment proportion with respect to the wealth after consumption $V(t) - C(t)\Delta$. The second equality of the equation (8.19) is due

to the self-financing constraint (8.18).

Rearranging we obtain

$$V(t + \Delta) = (V(t) - C(t)\Delta) \left(1 + \sum_{i=0}^m \alpha_i(t) \frac{P_i(t + \Delta) - P_i(t)}{P_i(t)} \right) \quad (8.20)$$

$$\Rightarrow \frac{V(t + \Delta) - V_t}{V(t) - C(t)\Delta} = -\psi(t)\Delta + \sum_{i=0}^m \alpha_i(t) \frac{P_i(t + \Delta) - P_i(t)}{P_i(t)}, \quad (8.21)$$

where $\psi(t) := \frac{C(t)}{V(t) - C(t)\Delta}$ is the consumption proportion with respect to the wealth after consumption.

For the transition from the discrete-time model to its continuous-time counterpart as Δ goes to zero, we have $V(t) - C(t)\Delta \sim V(t)$ and so obtain (3.8).

8.1.6 Sequential Optimization

We saw already in Section 8.1.2 that the life-time optimization problem (6.1) can be solved backwards step for step. Here we will show that asset allocation decisions for every step can be taken sequentially further: first the agents decide their investment plan, then their consumption plan. For the investment decisions, the agents maximize their *intertemporal expected utility* of the portfolio return.

We consider only the case without inflation risk. That means, the price index is constant equal to one and the wealth level and the consumption are all considered in nominal terms. We follow the wealth dynamics (8.20), which leads to a marginally difference to the discrete-time dynamics in Section 8.1.2 as mentioned above.

Analogous to (6.2), the value functions on the subperiod $[k\Delta, T]$ now depending on nominal consumption are defined by

$$J^\Delta(k\Delta, T, V_{k\Delta}, X_{k\Delta}) := \max_{\substack{C_{k'\Delta}, \alpha_{k'\Delta}, \\ k' = k, \dots, TN_\Delta - 1}} \mathbf{E}_{k\Delta} \left[\epsilon_1 \sum_{k'=k}^{TN_\Delta - 1} e^{-\delta k'\Delta} U(C_{k'\Delta}) + e^{-\delta T} U(V_T) \right], \quad (8.22)$$

where k is any number from $\{0, 1, \dots, TN_\Delta - 1\}$.

Applying the Euler-Maruyama scheme on the price dynamics (3.3), we obtain

$$\frac{P_i(t + \Delta) - P_i(t)}{P_i(t)} = \mu_i(X_t, t)\Delta + \Sigma_i(X_t, t)(W_{t+\Delta} - W_t).$$

This means that the risky asset returns $\frac{P_i(t+\Delta) - P_i(t)}{P_i(t)}$ depend on the state X_t , the time point t and the uncertainty $\Delta W_{t+\Delta} := W_{t+\Delta} - W_t$.

Following this, the portfolio gross return is a function of those determinants of the asset returns $(X_t, t, \Delta W_{t+\Delta})$ and the portfolio decision α_t in addition. Let Π denote the mapping of the determinants to the portfolio gross return, which we mean mathematically

$$\Pi(t + \Delta, X_t, \Delta W_{t+\Delta}, \alpha_t) = 1 + \sum_{i=0}^m \alpha_i(t) \frac{P_i(t + \Delta) - P_i(t)}{P_i(t)}.$$

Using this to rewrite the expression (8.20) for the nominal wealth development, we obtain

$$V(t + \Delta) = (V(t) - C(t)\Delta)\Pi(t + \Delta, X_t, \Delta W_{t+\Delta}, \alpha_t). \quad (8.23)$$

Using the similar proof idea as Property 21, we can have a multiplicative form also for the nominal value function.

Property 22 *If (i) the utility function is of the CRRA class and (ii) the nominal wealth dynamics is given by the expression (8.23), then the nominal value function defined in (8.22) has the following multiplicative form*

$$J^\Delta(k\Delta, T, V_{k\Delta}, X_{k\Delta}) = e^{-\delta k\Delta} U(V_{k\Delta}) \Theta^\Delta(k\Delta, T, X_{k\Delta}), \quad (8.24)$$

for all $k = 0, 1, \dots, TN_\Delta$, where

$$\begin{aligned} & \Theta^\Delta(t, T, X_t) \\ & := (1 - \gamma) \max_{\psi_t, \alpha_t} \left\{ \epsilon_1 U(\psi_t)\Delta + e^{-\delta\Delta} U(1 - \psi_t \Delta) (1 - \gamma) H(t + \Delta, T, \alpha_t, X_t) \right\}, \end{aligned} \quad (8.25)$$

for $t = k\Delta$. The function $H(t, T, \alpha_t, X_t)$ is defined by

$$\begin{aligned} & H(t, T, \alpha_t, X_t) \\ & := \mathbf{E}_t \left[\Pi(t + \Delta, X_t, \Delta W_{t+\Delta}, \alpha_t)^{1-\gamma} \Theta^\Delta(t + \Delta, T, \hat{X}(X_t)) \right] \\ & = (1 - \gamma) \mathbf{E}_t \left[U \left(\Pi(t + \Delta, X_t, \Delta W_{t+\Delta}, \alpha_t) \right) \Theta^\Delta(t + \Delta, T, \hat{X}(X_t)) \right] \end{aligned} \quad (8.26)$$

and $\psi_t := \frac{C_t}{V_t}$ ³. The iteration is defined through a backward scheme from $t = (k+1)\Delta$ to $t = k\Delta$. For the first iteration with $k = TN_\Delta - 1$ we have

$$\Theta^\Delta(T, T, X_T) \equiv 1$$

on the RHS of the iteration scheme (8.25).

Proof Use the similar idea of the proof of Property 21.

□

The function $H(t, T, X_t, \alpha_t)$ defined in the equation (8.26) can be interpreted as the *intertemporal expected utility* of the portfolio return.

For the case considering intermediate consumption $\epsilon_1 = 1$, we show that the decision problem (8.25) where ψ_t, α_t are determined, can be solved sequentially: first choosing the investment plan, then the consumption plan.

Property 23 *Under the same assumptions as in Property 22, the optimization in (8.25) can be taken sequentially*

$$\begin{aligned} & \max_{\psi_t, \alpha_t} \left\{ U(\psi_t)\Delta + e^{-\delta\Delta}U(1 - \psi_t\Delta)(1 - \gamma)H(t, T, X_t, \alpha_t) \right\} \quad (8.27) \\ & = \max_{\psi_t} \left\{ U(\psi_t)\Delta + e^{-\delta\Delta}U(1 - \psi_t\Delta)(1 - \gamma) \max_{\alpha_t} H(t, T, X_t, \alpha_t) \right\}. \end{aligned}$$

Proof

In the optimization problem on the LHS of the equation (8.27) the portfolio decision α_t appears only in the term $H(t, T, X_t, \alpha_t)$. Therefore α_t must be an extreme solution of $H(t, T, X_t, \alpha_t)$. The question now that which extreme solution of α_t – the maximum or the minimum solution – maximizes the expected utility of the portfolio return on the LHS of equation (8.27). We show in the following that α_t is the maximum solution.

First, we let

$$\tilde{H} = (1 - \gamma)H(t, T, X_t, \alpha_t) \quad (8.28)$$

for some given X_t, α_t and solve the consumption decision in terms of \tilde{H} . The consumption decision ψ_t^* is the maximum solution of the following expression

$$\psi_t^* = \arg \max \left\{ U(\psi_t)\Delta + e^{-\delta\Delta}U(1 - \psi_t\Delta)\tilde{H} \right\}.$$

The maximum solution ψ_t^* has to satisfy the FOC

$$(\psi_t^*)^{-\gamma}\Delta - e^{-\delta\Delta}(1 - \psi_t^*\Delta)^{-\gamma}\Delta H = 0 \quad (8.29)$$

³This consumption ratio is different than that in the wealth dynamics (8.20). This difference vanishes when $\Delta \rightarrow 0$.

and has a negative second order derivative. The second order derivative is equal to

$$\begin{aligned} & -\gamma \left((\psi_t^*)^{-\gamma-1} \Delta + e^{-\delta\Delta} (1 - \psi_t^* \Delta)^{-\gamma-1} H \Delta^2 \right) \\ = & -\gamma (\psi_t^*)^{-\gamma-1} \frac{\Delta}{1 - \psi_t^* \Delta}, \end{aligned}$$

which is negative since γ and ψ_t^* are all positive.

Using the FOC (8.29), we solve the optimal consumption ψ_t^* in terms of \tilde{H} as given in the following

$$\psi_t^*(\tilde{H}) = \frac{(e^{-\delta\Delta} \tilde{H})^{-1/\gamma}}{1 + \Delta (e^{-\delta\Delta} \tilde{H})^{-1/\gamma}}. \quad (8.30)$$

It is important to note that $\psi_t^*(\tilde{H})$ is a decreasing function in \tilde{H} .

With the solution (8.30) we can express the “optimized” objective function on the LHS of the equation (8.27) in terms of \tilde{H}

$$\begin{aligned} & U(\psi_t^*) \Delta + e^{-\delta\Delta} U(1 - \psi_t^* \Delta) \tilde{H} \quad (8.31) \\ = & \frac{\psi_t^*(\tilde{H})^{1-\gamma}}{1-\gamma} \Delta + e^{-\delta\Delta} \frac{(1 - \psi_t^*(\tilde{H}) \Delta)^{1-\gamma}}{1-\gamma} \tilde{H} \\ = & \frac{1}{1-\gamma} \left(\psi_t^*(\tilde{H})^{-\gamma} \psi_t^*(\tilde{H}) \Delta + \psi_t^*(\tilde{H})^{-\gamma} (1 - \psi_t^*(\tilde{H}) \Delta) \right) = \frac{\psi_t^*(\tilde{H})^{-\gamma}}{1-\gamma}. \end{aligned}$$

The first equality in the equation (8.31) is obtained by using the FOC (8.29).

Returning the expression of \tilde{H} using (8.28), the result of the reformulation of the objective function (8.31) stipulates that the portfolio decision α_t should maximize the following expression

$$\frac{\psi_t^* \left((1-\gamma) H(t, T, X_t, \alpha_t) \right)^{-\gamma}}{1-\gamma}.$$

We consider the following two cases:

- For the case $0 < \gamma < 1$, we have

$$\begin{aligned} & \alpha_t \max \frac{\psi_t^* \left((1 - \gamma) H(t, T, X_t, \alpha_t) \right)^{-\gamma}}{1 - \gamma} \\ \Leftrightarrow & \alpha_t \max \psi_t^* \left((1 - \gamma) H(t, T, X_t, \alpha_t) \right)^{-\gamma} \\ \Leftrightarrow & \alpha_t \min \psi_t^* \left(H(t, T, X_t, \alpha_t) \right) \\ \Leftrightarrow & \alpha_t \max H(t, T, X_t, \alpha_t) . \end{aligned}$$

- For the case $\gamma > 1$, we have

$$\begin{aligned} & \alpha_t \max \frac{(1 - \gamma) \psi_t^* \left(H(t, T, X_t, \alpha_t) \right)^{-\gamma}}{1 - \gamma} \\ \Leftrightarrow & \alpha_t \min \psi_t^* \left((1 - \gamma) H(t, T, X_t, \alpha_t) \right)^{-\gamma} \\ \Leftrightarrow & \alpha_t \max \psi_t^* \left((1 - \gamma) H(t, T, X_t, \alpha_t) \right) \\ \Leftrightarrow & \alpha_t \min (1 - \gamma) H(t, T, X_t, \alpha_t) \\ \Leftrightarrow & \alpha_t \max H(t, T, X_t, \alpha_t) . \end{aligned}$$

For the both cases, the investment decision α_t should maximize the intertemporal expected utility $H(t, T, X_t, \alpha_t)$ in order to achieve the maximization task (8.27).

□

8.2 Some Basic Results

Theorem 2 (Feynman-Kac Formula) *Let X_t be the solution of the stochastic differential equation (SDE)*

$$dX_t = \hat{F}_t dt + \hat{G}_t d\hat{W}_t \quad (8.32)$$

the infinitesimal generator of which is given by

$$\hat{\mathcal{D}}_t = \hat{F}_t^\top \frac{\partial}{\partial x} + \frac{1}{2} \sum_{i,j=1}^n \hat{G}_{it} \hat{G}_{jt}^\top \frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} .$$

Let $h : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$, $g : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$, and $\Psi : \mathbb{R}^d \times \mathbb{R}_+ \rightarrow \mathbb{R}$. If $\Psi(x, t)$ satisfies the PDE

$$\frac{\partial}{\partial t} \Psi(x, t) + \hat{\mathcal{D}}_t \Psi(x, t) + h(x, t) \Psi(x, t) + l(x, t) = 0, \quad (8.33)$$

subject to the boundary condition

$$\Psi(x, T) = \omega(x), \quad (8.34)$$

then

$$\Psi(x, t) = \hat{\mathbf{E}}_{t,x} \left[\omega(X_T) e^{\int_t^T h(X_s, s) ds} + \int_t^T l(X_s, s) e^{\int_t^s h(X_u, u) du} ds \right], \quad (8.35)$$

where $\hat{\mathbf{E}}_{t,x}$ is the expectation operator with respect to the stochastic process $X_s, s \geq t$ satisfying the SDE (8.32) with initial position $X_t = x$.

For the proof for Feymann-Kac formula, see, for example, Korn and Korn(2001).
□

In the following, a standard multidimensional Wiener process means the Wiener process with mutual independent components.

Theorem 3 (Girsanov's Transformation with standard Wiener processes)

Let W_t be an $n \times 1$ -dimensional standard \mathcal{P} -Wiener process. Let $a(s) = (a_1(s), \dots, a_n(s))^\top$ and $a_i(s) : \mathbb{R}_+ \rightarrow \mathbb{R}$ for $i = 1, \dots, n$ with

$$\mathbf{E}[\exp(\int_0^T a(s)^\top a(s) ds)] < \infty.$$

Let $\tilde{\mathcal{P}}$ be a new measure defined by

$$\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} = \exp\left(\int_0^T a(s)^\top dW_s - \frac{1}{2} \int_0^T a(s)^\top a(s) ds\right).$$

Then, a new process defined by

$$\tilde{W}_s := W_s - \int_0^s a(u) du \quad (8.36)$$

is an $n \times 1$ standard $\tilde{\mathcal{P}}$ -Wiener process.

Proof See Theorem 5.1 on p191 in Karatzas and Shreve (1991).

For our purpose we need the Girsanov's Transformation with correlated Wiener processes.

Theorem 4 (Girsanov's Transformation with correlated Wiener processes)

Let W_t be an $n \times 1$ -dimensional \mathcal{P} -Wiener process with the correlation matrix $\mathcal{R}dt = dW_t dW_t^\top$. Let $a(s) = (a_1(s), \dots, a_n(s))^\top$ and $a_i(s) : \mathbb{R}_+ \rightarrow \mathbb{R}$ for $i = 1, \dots, n$ with

$$\mathbf{E}[\exp(\int_0^T a(s)^\top a(s) ds)] < \infty. \quad (8.37)$$

Let $\tilde{\mathcal{P}}$ be a new measure defined by

$$\frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} = \exp\left(\int_0^T a(s)^\top \mathcal{R}^{-1} dW_s - \frac{1}{2} \int_0^T a(s)^\top \mathcal{R}^{-1} a(s) ds\right). \quad (8.38)$$

Then, the new process \tilde{W}_s defined by

$$\tilde{W}_s := W_s - \int_0^s a(u) du \quad (8.39)$$

is a $n \times 1$ $\tilde{\mathcal{P}}$ -Wiener process.

Proof

For the positive definite correlation matrix \mathcal{R} , there exists a linear transformation \mathcal{T} such that $\mathcal{T}\mathcal{T}^\top = \mathcal{R}$.

Let $\hat{W}_s := \mathcal{T}^{-1}W_s$. Then \hat{W}_s is a standard n -dimensional \mathcal{P} -Wiener process.

Rewriting the Radon-Nikodym derivative (8.38) into

$$\begin{aligned} \frac{d\tilde{\mathcal{P}}}{d\mathcal{P}} &= \exp\left(\int_0^T a(s)^\top \mathcal{R}^{-1} dW_s - \frac{1}{2} \int_0^T a(s)^\top \mathcal{R}^{-1} a(s) ds\right) \\ &= \exp\left(\int_0^T (\mathcal{T}^{-1}a(s))^\top d\hat{W}_s - \frac{1}{2} \int_0^T (\mathcal{T}^{-1}a(s))^\top \mathcal{T}^{-1}a(s) ds\right) \end{aligned}$$

and applying the Girsanov's Theorem 3, we can have that

$$\mathcal{T}^{-1}W_s - \int_0^s \mathcal{T}^{-1}a(u) du = \mathcal{T}^{-1}(W_s - \int_0^s a(u) du) =: \mathcal{T}^{-1}\tilde{W}_s$$

is an n -dimensional standard $\tilde{\mathcal{P}}$ Wiener process. Thus, the new process \tilde{W} defined in (8.39) is an $\tilde{\mathcal{P}}$ -Wiener process and with the same correlation

coefficient \mathcal{R} .

□

Theorem 5 (Novikov Condition) *Let*

$$\Lambda_t := \exp \left(\int_0^t a(s)^\top \mathcal{R}^{-1} dW_s - \frac{1}{2} \int_0^t a(s)^\top \mathcal{R}^{-1} a(s) ds \right). \quad (8.40)$$

If the Novikov condition (8.37) is satisfied, then $(\Lambda_t)_{t \in [0, T]}$ is a \mathcal{P} -martingale.

Proof See p.198 Karatzas and Shreve (1991).

8.3 Proofs

Proof of Property 1

Comparing the HJB equation (3.23) and the partial differential equation used by the Feynmann-Kac formula (8.33), we need to find a new measure $\tilde{\mathcal{P}}^X$ under that the process X_t satisfies the stochastic differential equation

$$dX_t = (F_t + G_t z_t) dt + G_t d\tilde{W}_t^X, \quad (8.41)$$

with \tilde{W}_t^X is an n -dimensional $\tilde{\mathcal{P}}^X$ Wiener process.

To this end we construct the new measure by the Radon-Nikodym derivative

$$\frac{d\tilde{\mathcal{P}}^X}{d\mathcal{P}} = \exp \left(\int_0^T z_s^\top \mathcal{R}_{XX}^{-1} dW_s^X - \frac{1}{2} \int_0^T z_s^\top \mathcal{R}_{XX}^{-1} z_s ds \right). \quad (8.42)$$

Then, using the result of Theorem 4, the process \tilde{W}_s^X defined by

$$\tilde{W}_s^X := W_s^X - \int_0^s z_u du$$

is an n -dimensional $\tilde{\mathcal{P}}^X$ -Wiener process. In other words, the original Wiener process W_s^X is a Wiener process with drift z_s

$$dW_s^X = d\tilde{W}_s^X + z_s ds \quad (8.43)$$

under the new measure $\tilde{\mathcal{P}}^X$. It turns out that the factor X_s satisfying the original SDE (3.2) now satisfies the SDE (8.41) because

$$dX_s = F_s ds + G_s dW_s^X = F_s ds + G_s (d\tilde{W}_s^X + z_s ds) = \text{RHS of (8.41)}.$$

Now we can apply the Feynman-Kac Formula in Theorem 2 under the new measure $\tilde{\mathcal{P}}^X$. Based on result of Theorem 2, the equation $\Phi(t, T, x)$ satisfying the SDE (3.23) and the boundary condition (3.15) is solved by the expectation expression

$$\Phi(t, T, x) = \tilde{\mathbf{E}}_{t,x} [e^{\int_t^T h_s ds} + \epsilon_1 \int_t^T e^{\int_t^s h_u du} ds],$$

where $\tilde{\mathbf{E}}_{t,x}$ is the expectation operator with respect the process X_s satisfying the SDE (8.41) with the initial value $X_t = x$. The two expectation operators are related by the equation

$$\tilde{\mathbf{E}}_{t,x}[w] = \mathbf{E}_{t,x}[w \frac{d\tilde{\mathcal{P}}}{d\mathcal{P}}],$$

see p.191 Karatzas and Shreve (1991), where w is any \mathcal{F}_T -measurable function.

Using our previous notation given in (3.26), the Radon-Nikodym derivative $d\tilde{\mathcal{P}}/d\mathcal{P}$ can be denoted by Λ_T . Thus, the solution $\Phi(t, T, x)$ can be rewritten as

$$\Phi(t, T, x) = \mathbf{E}_{t,x} [e^{\int_t^T h_s ds} \Lambda_T + \epsilon_1 (\int_t^T e^{\int_t^s h_u du} ds) \Lambda_T], \quad (8.44)$$

from which we obtain the first term on the RHS of equation (3.25).

The second term on the RHS of equation (8.44) can be rewritten further as

$$\begin{aligned} & \mathbf{E}_{t,x} \left[\left(\int_t^T e^{\int_t^s h_u du} ds \right) \Lambda_T \right] = \mathbf{E}_{t,x} \left[\int_t^T (e^{\int_t^s h_u du} \Lambda_T) ds \right] \\ &= \int_t^T \mathbf{E}_{t,x} [e^{\int_t^s h_u du} \Lambda_T] ds \\ &= \int_t^T \mathbf{E}_{t,x} [e^{\int_t^s h_u du} \mathbf{E}[\Lambda_T | \mathcal{F}_s]] ds = \int_t^T \mathbf{E}_{t,x} [e^{\int_t^s h_u du} \Lambda_s] ds \\ &= \mathbf{E}_{t,x} \left[\int_t^T e^{\int_t^s h_u du} \Lambda_s ds \right], \end{aligned}$$

where the first equality is because Λ_T can be considered as constant along the integral with respect to ds . The second equality is based on the usual Fubini Theorem to interchange the two integral operators $\mathbf{E}_{t,x}$ and \int_t^T , see,

for example p.53 in Klebaner (2005). The third equality is due to the iteration law of the conditional expectation, see for example Chung (1974). The fourth equality is because the process $(\Lambda_s)_{s \in [0, T]}$ is a martingale according to Theorem 5 when the Novikov condition (3.24) is satisfied. The last equality is an application of the usual Fubini Theorem again. The the second term on the RHS of equation (8.44) is obtained.

□

Proof of Property 2

Comparing the optimal strategy obtained by the Feynman-Kac formula (3.25)

$$\Phi(t, T, x) = \mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T e^{\int_t^s h_u du} \Lambda_s ds + e^{\int_t^T h_s ds} \Lambda_T \right]$$

and that solved by the martingale method (3.32)

$$\Phi(t, T, X_t) = \mathbf{E}_{t,x} \left[\epsilon_1 \int_t^T e^{-\frac{\delta(s-t)}{\gamma}} H(t, s)^{1-\frac{1}{\gamma}} ds + e^{-\frac{\delta(T-t)}{\gamma}} H(t, T)^{1-\frac{1}{\gamma}} \right],$$

we observe that the two solutions have a similar structure and we can prove the equivalence of these two solutions by proving

$$\mathbf{E}_{t,X_t} \left[e^{\int_t^s h_u du} \Lambda_s \right] = \mathbf{E}_{t,X_t} \left[e^{-\frac{\delta(s-t)}{\gamma}} H(t, s)^{1-\frac{1}{\gamma}} \right], \quad (8.45)$$

for $s \in [t, T]$.

It is convenient to prove this property by considering standard (orthogonal) Wiener processes. Let \mathcal{C} be the Cholesky decomposition of the covariance matrix \mathcal{R}_{AA} , that is, \mathcal{C} is a lower triangular matrix with $\mathcal{C}\mathcal{C}^\top = \mathcal{R}_{AA}$. So,

$$\hat{W}_u := \mathcal{C}^{-1} W_u$$

is an m -dimensional standard Wiener process because $\mathcal{C}^{-1}\mathcal{R}_{AA}(\mathcal{C}^{-1})^\top$ is a unit matrix.

Under the assumption A5 in Section 3.1 the risk sources of the factor are included in the set of the risk sources of the asset returns, that is, $W_u^X \subseteq W_u$. Recall the notations used in Section 3.1 where the m sources of the factor uncertainty W_u^X are the first m components of the asset return uncertainty so that $W_u = (W_{1u}^X, \dots, W_{nu}^X, W_{(n+1)u}, \dots, W_{mu})^\top$. We also let $W_u^O = (W_{(n+1)u}, \dots, W_{mu})^\top$.

Let \mathcal{C}^{XA} consist of the first n rows in the matrix \mathcal{C}^{-1} and \mathcal{C}^{OA} consist of the rest $m - n$ rows. We let also $\hat{W}_u^X := \mathcal{C}^{XA}W_u$ and $\hat{W}_u^O := \mathcal{C}^{OA}W_u$. Expressing it in the vector form, we have

$$\mathcal{C}^{-1}W_u = \begin{pmatrix} \mathcal{C}^{XA} \\ \mathcal{C}^{OA} \end{pmatrix} W_u = \begin{pmatrix} \hat{W}_u^X \\ \hat{W}_u^O \end{pmatrix} =: \hat{W}_u.$$

We note that $\hat{W}_u^X \perp \hat{W}_u^O$ under the construction.

In the similar way we transform the uncertainty source of the factor. Let \mathcal{C}_X be the Cholesky decomposition of the covariance matrix \mathcal{R}_{XX} , where \mathcal{C}_X is a lower triangular matrix with $\mathcal{C}_X(\mathcal{C}_X)^\top = \mathcal{R}_{XX}$.

We note that

$$(\mathcal{C}_X)^{-1}W_u^X = \mathcal{C}^{XA}W_u = \hat{W}_u^X$$

because \mathcal{C} and \mathcal{C}_X are both the lower-triangular Cholesky decompositions⁴.

For the market price of risk, correspondingly, we let $\hat{\lambda}_u$

$$\hat{\lambda}_u := \mathcal{C}^{-1}\lambda_u,$$

which represents the corresponding market price of risk for the standard Wiener process \hat{W}_u . We can also decompose $\hat{\lambda}_u$ as $\hat{\lambda}_u = (\hat{\lambda}_u^X, \hat{\lambda}_u^O)^\top$, where $\hat{\lambda}_u^X$ is an $n \times 1$ process representing the market price for the uncertainty source \hat{W}_u^X and $\hat{\lambda}_u^O$ is an $(m - n) \times 1$ process representing the market price for the uncertainty source \hat{W}_u^O .

Now we are ready to prove the equality (8.45). We begin with the LHS. For a constant price index we have the terms $\pi(X_t)$ and σ_I equal zero, so the process h_u given by (3.22) becomes

$$h_u = -\frac{\delta}{\gamma} + \frac{1 - \gamma}{\gamma}R_u + \frac{1 - \gamma}{2\gamma^2}\lambda_u^\top \mathcal{R}_{AA}^{-1}\lambda_u$$

⁴Intuitively, the lower-triangular orthogonal transformation is proceeded as follows: the first component of \hat{W}_u is the same as the first component of W_u , the second component of \hat{W}_u is the part of the second component of W_u which is independent to the first component W_u , the third component of the \hat{W}_u is the part of the third component of W_u which is independent to the first and second component of W_u , and so on. Since W_u^X are the first m components of W_u , the lower-triangular orthogonal decomposition of W_u^X must be the as same as the lower-triangular orthogonal decomposition of the first m components of W_u .

where $R_u := R(X_u)$ and $\lambda_u := \lambda(X_u)$. With the same reason, the Radon-Nikodym derivative Λ_s given by equation (3.26) becomes

$$\begin{aligned} \Lambda_s &= \exp\left(\frac{1-\gamma}{\gamma} \int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} \mathcal{R}_{XX}^{-1} dW_u^X\right) \\ &\quad \times \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} \mathcal{R}_{XX}^{-1} \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \lambda_u du\right). \end{aligned}$$

Inserting the two equations above into the LHS of equation (8.45) and using the notations based on the orthogonal Wiener processes given above, we have

$$\begin{aligned} &\text{LHS of the equality (8.45)} \\ &= \mathbf{E}_{t,x} \left[\exp\left(\int_t^s \left(-\frac{\delta}{\gamma} + \frac{1-\gamma}{\gamma} R_u + \frac{1-\gamma}{2\gamma^2} \lambda_u^\top \mathcal{R}_{AA}^{-1} \lambda_u\right) du\right) \right. \\ &\quad \exp\left(\frac{1-\gamma}{\gamma} \int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} \mathcal{R}_{XX}^{-1} dW_u^X\right) \\ &\quad \left. \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^s \lambda_u^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} \mathcal{R}_{XX}^{-1} \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \lambda_u du\right) \right] \\ &= \mathbf{E}_{t,x} \left[\exp\left(-\frac{\delta}{\gamma}(s-t) + \frac{1-\gamma}{\gamma} \int_t^s R_u du + \frac{1-\gamma}{2\gamma^2} \int_t^s \hat{\lambda}_u^\top \hat{\lambda}_u du\right) \right. \\ &\quad \exp\left(\frac{1-\gamma}{\gamma} \int_t^s \hat{\lambda}_u^\top \begin{pmatrix} \mathcal{I}_X \\ 0 \end{pmatrix} d\hat{W}_u^X\right) \\ &\quad \left. \exp\left(-\frac{1}{2} \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^s \hat{\lambda}_u^\top \begin{pmatrix} \mathcal{I}_X \\ 0 \end{pmatrix} \begin{pmatrix} \mathcal{I}_X & 0 \end{pmatrix} \hat{\lambda}_u du\right) \right]. \end{aligned} \quad (8.46)$$

The second equality is based on the orthogonal transformation defined above and also the identity

$$\mathcal{C}^{-1} \mathcal{R}_{AX} (\mathcal{C}_X^{-1})^\top = \mathbf{Cov}\left((\mathcal{C}^{-1} W_1)(\mathcal{C}_X^{-1} W_1^X)^\top\right) = \mathbf{Cov}(\hat{W}_1(\hat{W}_1^X)^\top) = \begin{pmatrix} \mathcal{I}_X \\ 0 \end{pmatrix},$$

where \mathcal{I}_X is the n -dimensional unit matrix and the notation 0 above denotes $(m-n) \times n$ -zero matrix.

Using elementary matrix operations, we obtain the decomposition of the market price of risk

$$\hat{\lambda}_u^\top \hat{\lambda}_u = (\hat{\lambda}_u^X)^\top \hat{\lambda}_u^X + (\hat{\lambda}_u^O)^\top \hat{\lambda}_u^O$$

and the stochastic integrals

$$\int_t^s \hat{\lambda}_u^\top d\hat{W}_u = \int_t^s (\hat{\lambda}_u^X)^\top d\hat{W}_u^X + \int_t^s (\hat{\lambda}_u^O)^\top d\hat{W}_u^O.$$

Using these equations to rewrite the expression (8.46), we obtain

$$\begin{aligned}
& \text{LHS the equality (8.45)} \\
= & \mathbf{E}_{t, X_t} \left[\exp \left(-\frac{\delta}{\gamma}(s-t) + \frac{1-\gamma}{\gamma} \int_t^s R_u du + \frac{1-\gamma}{2\gamma} \int_t^s \hat{\lambda}_u^\top \hat{\lambda}_u du \right) \right. \\
& \exp \left(\frac{1-\gamma}{\gamma} \int_t^s (\hat{\lambda}_u)^\top d\hat{W}_u - \frac{1-\gamma}{\gamma} \int_t^s (\hat{\lambda}^O)_u^\top d\hat{W}_u^O \right) \\
& \left. \exp \left(\frac{1}{2} \left(\frac{1-\gamma}{\gamma} \right)^2 \int_t^s (\hat{\lambda}^O)_u^\top \hat{\lambda}_u^O du \right) \right] \quad (8.47) \\
= & \mathbf{E}_{t, X_t} \left[e^{-\frac{\delta}{\gamma}(s-t)} H(t, s)^{1-\frac{1}{\gamma}} \right] \\
& \mathbf{E}_{t, X_t} \left[\exp \left(-\frac{1-\gamma}{\gamma} \int_t^s (\hat{\lambda}^O)_u^\top d\hat{W}_u^O + \frac{1}{2} \left(\frac{1-\gamma}{\gamma} \right)^2 \int_t^s (\hat{\lambda}^O)_u^\top \hat{\lambda}_u^O du \right) \right] \\
= & \left(\text{RHS of the equality (8.45)} \right) \mathbf{E}_{t, X_t} \left[\exp \left(\left(\frac{1-\gamma}{\gamma} \right)^2 \int_t^s (\hat{\lambda}^O)_u^\top \hat{\lambda}_u^O du \right) \right].
\end{aligned}$$

The second equality in equation (8.47) is because \hat{W}_t^O is orthogonal to \hat{W}_t^X .

If the risk sources of the asset returns $(W_s)_{s \in [0, T]}$ can be spanned by the risk sources of the factor $(W_s^X)_{s \in [0, T]}$, that is, $W_s \subseteq W_s^X$ then we have $W_t^X \equiv W_t$ because we assume $W_s^X \subseteq W_s$ due to the assumption A5 in Section 3.1. So, we do not have \hat{W}^O and equation (8.47) turns out to be the LHS of equation (8.45) is equal to the RHS of the equation (8.45).
□

We would like to remark that the two solution processes use Girsanov's transformation. However, the Feynman-Kac formula transforms measure in the space of the factor risks W_t^X while the martingale method transforms measure in the space of all asset uncertainty W_t . Therefore, the correlation of the other uncertainty W_t^O and W_t^X still affects the value function $\Phi(t, T, X_t)$.

Proof of Property 3

There are three parts of this proof. In the first part we characterize the invariant transformation of the parameters under an affine transformation of the factors. The characterization is summarized in a lemma. After the lemma we will show in the second part that the normalization conditions stated in the Property 3 do not restrict the general parameterization. In the third part we will show there is only one parameter representation cor-

responds to one data generation process.

An **invariant transformation** of the parameters is a transformation of the following parameters: the parameters $\mathcal{K}, \theta, \Gamma$ in (4.6) and those ξ_0, ξ_1 in (4.11) and (4.10) under the transformation of the factors X_t which supports the same data generating process $Y_n(t, t + \tau, X_t)$ described previously in the formula (4.12) with the transformed factors. In addition to this requirement, an invariant transformation of the parameters has to guarantee that the transformed parameters satisfy the no-arbitrage conditions (4.11) and (4.10). Here we consider a full-rank affine transformation

$$X_t^{\mathcal{L}} := \mathcal{L}X_t + \bar{\Theta}. \quad (8.48)$$

We need to remark that the characterization of the invariant transformation of the parameters are already stated in Dai and Singleton (2000). Here we provide a more detailed proof.

Lemma 23.1 (*Dai and Singleton (2000)*) *The invariant transformation of the parameters $(\mathcal{K}, \theta, \Gamma, \xi_0, \xi_1)$ with respect to the factor transformation $X_t^{\mathcal{L}} = \mathcal{L}X_t + \bar{\Theta}$ is*

$$(\mathcal{L}\mathcal{K}\mathcal{L}^{-1}, \mathcal{L}\theta + \bar{\Theta}, \mathcal{L}\Gamma, \xi_0 - \xi_1^{\top}\mathcal{L}^{-1}\bar{\Theta}, (\mathcal{L}^{\top})^{-1}\xi_1). \quad (8.49)$$

Proof

The first three invariant parameter transformation can be determined easily. We denote $\mathcal{K}^{\mathcal{L}}, \theta^{\mathcal{L}}, \Gamma^{\mathcal{L}}$ as new parameters for the new factor dynamics

$$dX_t^{\mathcal{L}} = \mathcal{K}^{\mathcal{L}}(\theta^{\mathcal{L}} - X_t^{\mathcal{L}})dt + \Gamma^{\mathcal{L}}dW_t^X. \quad (8.50)$$

Under the factor transformation (8.48), the new factor dynamics can be transformed into

$$\begin{aligned} dX_t^{\mathcal{L}} &= \mathcal{L}dX_t = \mathcal{L}\mathcal{K}(\theta - X_t)dt + \mathcal{L}\Gamma dW_t^X \\ &= \mathcal{L}\mathcal{K}(\theta - \mathcal{L}^{-1}(\mathcal{L}X_t + \bar{\Theta}) + \mathcal{L}^{-1}\bar{\Theta})dt + \mathcal{L}\Gamma dW_t^X \\ &= (\mathcal{L}\mathcal{K}\mathcal{L}^{-1})(\mathcal{L}\theta + \bar{\Theta} - X_t^{\mathcal{L}})dt + \mathcal{L}\Gamma dW_t^X. \end{aligned} \quad (8.51)$$

Identifying the two dynamics (8.50) and (8.51), we obtain

$$\begin{aligned} \mathcal{K}^{\mathcal{L}} &= \mathcal{L}\mathcal{K}\mathcal{L}^{-1}, \\ \theta^{\mathcal{L}} &= \mathcal{L}\theta + \bar{\Theta}, \\ \Gamma^{\mathcal{L}} &= \mathcal{L}\Gamma. \end{aligned} \quad (8.52)$$

Let $B^{\mathcal{L}}(\tau)$, $A^{\mathcal{L}}(\tau)$ be the new coefficients under the transformation $X_t^{\mathcal{L}}$. The invariant transformation requires, on the one hand, the model of the bond yields to remain invariant

$$Y(t, t + \tau, X_t) = \frac{A(\tau)}{\tau} + \frac{B(\tau)^\top}{\tau} X_t \equiv \frac{A^{\mathcal{L}}(\tau)}{\tau} + \frac{B^{\mathcal{L}}(\tau)^\top}{\tau} X_t^{\mathcal{L}}. \quad (8.53)$$

Replacing the new factor with its definition (8.48) in the equivalence (8.53), we obtain the following equalities for the new coefficients $B^{\mathcal{L}}(\tau)$ and $A^{\mathcal{L}}(\tau)$

$$B^{\mathcal{L}}(\tau)^\top = B(\tau)^\top \mathcal{L}^{-1} \quad (8.54)$$

$$A(\tau) = A^{\mathcal{L}}(\tau) + B(\tau)^\top \mathcal{L}^{-1} \bar{\Theta}. \quad (8.55)$$

On the other hand, the invariant transformation also requires the new coefficient $B^{\mathcal{L}}(\tau)$ to satisfy the no-arbitrage equation (4.10) introduced in Chapter 3 with the new parameters given in (8.52)

$$\frac{d}{d\tau} B^{\mathcal{L}}(\tau) = -(\mathcal{K}^{\mathcal{L}})^\top B^{\mathcal{L}}(\tau) + \xi_1^{\mathcal{L}} = -(\mathcal{L}^{-1})^\top \mathcal{K}^\top \mathcal{L}^\top B^{\mathcal{L}}(\tau) + \xi_1^{\mathcal{L}}.$$

Multiplying \mathcal{L}^\top on the both sides, we obtain

$$\frac{d}{d\tau} \left(\mathcal{L}^\top B^{\mathcal{L}}(\tau) \right) = -\mathcal{K}^\top \mathcal{L}^\top B^{\mathcal{L}}(\tau) + \mathcal{L}^\top \xi_1^{\mathcal{L}}.$$

This equation can be rewritten further to

$$\frac{d}{d\tau} B(\tau) = -\mathcal{K}^\top B(\tau) + \xi_1, \quad (8.56)$$

due to the fact $\mathcal{L}^\top B^{\mathcal{L}}(\tau) \equiv B(\tau)$ from the equality (8.54).

Identifying the new differential equation (8.56) with the original one (4.10), the new parameter $\xi_1^{\mathcal{L}}$ has to satisfy

$$\mathcal{L}^\top \xi_1^{\mathcal{L}} = \xi_1. \quad (8.57)$$

With the same reason, the coefficient $A^{\mathcal{L}}(\tau)$ has to satisfy the equation (4.11) with the new parameters (8.52) and the new coefficient $B^{\mathcal{L}}(\tau)$

$$\frac{d}{d\tau} A^{\mathcal{L}}(\tau) = (\mathcal{K}^{\mathcal{L}} \theta^{\mathcal{L}} - \Gamma^{\mathcal{L}} \lambda)^\top B^{\mathcal{L}}(\tau) - \frac{1}{2} \sum_{i,j=1}^n B_i^{\mathcal{L}}(\tau) B_j^{\mathcal{L}}(\tau) \Gamma_i^{\mathcal{L}} \Gamma_j^{\mathcal{L}} + \xi_0^{\mathcal{L}}. \quad (8.58)$$

We observe that

$$\sum_{i,j=1}^n B_i^{\mathcal{L}}(\tau) B_j^{\mathcal{L}}(\tau) \Gamma_i^{\mathcal{L}} \Gamma_j^{\mathcal{L}} = (B^{\mathcal{L}}(\tau)^\top \Gamma^{\mathcal{L}}) (B^{\mathcal{L}}(\tau)^\top \Gamma^{\mathcal{L}})^\top = (B(\tau)^\top \Gamma) (B(\tau)^\top \Gamma)^\top.$$

Using this fact and the expression of the new parameters (8.52), the differential equation (8.58) can be transformed further into

$$\begin{aligned} \frac{d}{d\tau} A^{\mathcal{L}}(\tau) &= (\mathcal{K}\theta + \mathcal{K}\mathcal{L}^{-1}\bar{\Theta} - \Gamma\lambda)^\top B(\tau) - \frac{1}{2} \sum_{i,j=1}^n B_i(\tau) B_j(\tau) \Gamma_i \Gamma_j + \xi_0^{\mathcal{L}} \\ &= \frac{d}{d\tau} A(\tau) + (\mathcal{K}\mathcal{L}^{-1}\bar{\Theta})^\top B(\tau) + \xi_0^{\mathcal{L}} - \xi_0. \end{aligned} \quad (8.59)$$

The second equality above is obtained by using the original no-arbitrage condition (4.11). We note that the risk price λ remains unchanged under the factor transformation because we keep the original factor uncertainty W_t^X . The risk price is the compensation for bearing the uncertainty W_t^X .

Differentiating both sides of (8.55) and then replacing $\frac{d}{d\tau} B(\tau)$ by the original no-arbitrage condition (4.10), we have

$$\begin{aligned} \frac{d}{d\tau} A(\tau) &= \frac{d}{d\tau} A^{\mathcal{L}}(\tau) + \frac{d}{d\tau} B(\tau)^\top \mathcal{L}^{-1} \bar{\Theta} \\ &= \frac{d}{d\tau} A^{\mathcal{L}}(\tau) + (-B(\tau)^\top \mathcal{K} + \xi_1^\top) \mathcal{L}^{-1} \bar{\Theta}. \end{aligned} \quad (8.60)$$

Identifying the two equations (8.59) and (8.60), the new parameter $\xi_0^{\mathcal{L}}$ has to satisfy

$$\xi_0^{\mathcal{L}} = \xi_0 - \xi_1^\top \mathcal{L}^{-1} \bar{\Theta}. \quad (8.61)$$

□

Proof of Property 3 (continued)

Now we will prove for any given admissible parameters $(\mathcal{K}, \theta, \Gamma, \xi_0, \xi_1)$, there exists exactly one parameter representation which satisfies the conditions (i) – (iv) in Property 3. Recall the admissible \mathcal{K} is positive definite with different (positive) eigenvalues.

First, we want to find \mathcal{L}^* and $\bar{\Theta}^*$ for the factor transformation $X^{\mathcal{L}^*} = \mathcal{L}^* X_t + \bar{\Theta}^*$ so that the transformed parameters using the corresponding invariant transformation (8.49) satisfy the conditions (i) – (iii). From an operational point of view, we need to rotate, rescale and shift the factors such that the corresponding transformed parameters can satisfy the conditions (i) – (iii).

Rotation:

Since \mathcal{K} is positive definite, there exists a unique full-rank transformation $\hat{\mathcal{L}}$ with $\hat{\mathcal{L}}^{-1} = \hat{\mathcal{L}}^\top$ such that $\hat{\mathcal{L}}\mathcal{K}\hat{\mathcal{L}}^{-1}$ is diagonal.

Rescaling:

Let d_1, \dots, d_n be constants defined by

$$\begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{pmatrix} := (\hat{\mathcal{L}}^\top)^{-1} \xi_1 .$$

Let \mathcal{D} be the diagonal matrix with the elements d_1, \dots, d_n on the diagonal. By the construction of \mathcal{D} , we can have

$$((\mathcal{D}\hat{\mathcal{L}})^\top)^{-1} \xi_1 = \mathcal{D}^{-1} (\hat{\mathcal{L}}^\top)^{-1} \xi_1 = \begin{pmatrix} d_1^{-1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n^{-1} \end{pmatrix} \begin{pmatrix} d_1 \\ \vdots \\ d_n \end{pmatrix} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} . \quad (8.62)$$

Define $\mathcal{L}^* := \mathcal{D}\hat{\mathcal{L}}$ and $\bar{\Theta}^* := -\mathcal{L}^*\theta$ (shifting). We show that the parameter under the invariant transformation based on $(\mathcal{L}^*, \bar{\Theta}^*)$ satisfy the conditions (i) – (iii).

Condition (i) is satisfied because $\mathcal{K}^{\mathcal{L}^*}$ is diagonal with the following reformulation

$$\mathcal{K}^{\mathcal{L}^*} = \mathcal{L}^* \mathcal{K} (\mathcal{L}^*)^{-1} = \mathcal{D} \hat{\mathcal{L}} \mathcal{K} \hat{\mathcal{L}}^\top \mathcal{D} .$$

Condition (ii) is satisfied because

$$\theta^{\mathcal{L}^*} = \mathcal{L}^* \theta + \bar{\Theta}^* = \mathcal{L}^* \theta - \mathcal{L}^* \theta = 0 .$$

Condition (iii) is satisfied because, by following the result of the transformation of ξ_1 given by the expression (8.57), the new ξ_1 satisfies

$$\xi_1^{\mathcal{L}^*} = (\mathcal{L}^*)^\top \xi_1 = ((\mathcal{D}\hat{\mathcal{L}})^\top)^{-1} \xi_1 = (1, 1, \dots, 1)^\top$$

due to the calculation (8.62).

For Condition (iv) we note that we still have flexibility to rotate the factor uncertainty W_t^X . The rotation is represented by $\mathcal{T}W_t^X$ where \mathcal{T} is an $n \times n$ -matrix with $\mathcal{T}^{-1} = \mathcal{T}^\top$. This rotation does not affect the factor dynamics when we change the diffusion coefficient correspondingly

$$\Gamma^{\mathcal{L}^*} W_t^X = (\Gamma^{\mathcal{L}^*} \mathcal{T}^{-1})(\mathcal{T}W_t^X) .$$

We can find \mathcal{T} such that $\Gamma^{\mathcal{L}^*} \mathcal{T}^{-1}$ is lower triangular. Note that by rotating W_t^X we need to rotate the market price of risk $\mathcal{T} \lambda_t$ correspondingly. Under this readjustment of the market price of risk, the no-arbitrage conditions (4.9) remains the same because of $\Gamma \mathcal{T}^{-1} \mathcal{T} \lambda_t = \Gamma \lambda_t$. Also the data generating process (4.12) has the same coefficients $A(\tau), B(\tau)$ under the rotation. The rotation affects only $A(\tau)$, see the equation (4.11) where we can have the invariant $\Gamma_i \mathcal{T}^\top \mathcal{T} \Gamma_j = \Gamma_i \Gamma_j$.

For the uniqueness of the parameter representation, we note that for any given admissible parameters $(\mathcal{K}, \theta, \Gamma, \xi_0, \xi_1)$ this transformation $(\mathcal{L}^*, \bar{\Theta}^*)$ is uniquely determined by the eigen decomposition with the transformation $\hat{\mathcal{L}}$. The eigen decomposition is unique up to permutations of the eigenvalues.
□

Proof of Property 4

Because \mathcal{K} is diagonal, we can solve every component of the coefficient $B(\tau)$ separately. Together with Condition (iii) in Property 3, the i -th component of $B(\tau)$ has to satisfy

$$\frac{d}{d\tau} B_i(\tau) = -\kappa_i B_i(\tau) + 1.$$

We can check that (4.13) is the solution.

The solution given in (4.14) can be checked easily as the solution of $A(\tau)$ satisfying (4.11) after $B(\tau)$ has been solved.

□

Proof of Property 5

Replacing μ_I by the expression (4.29), the no-arbitrage equality (4.32) can be rewritten as

$$\begin{aligned} \mu_I(t, \tau) - R_t & \tag{8.63} \\ = -\lambda_r B_{rr}(\tau) g_r + \lambda_I \sigma_I & = (\mu_r(t, \tau) + \pi_t - B_{rr}(\tau) g_r \sigma_I \rho_{I_r}) - R_t. \end{aligned}$$

Rearranging the second line above and using the third no-arbitrage equality (4.33), then we have the equation

$$\begin{aligned} \mu_r(t, \tau) - (R_t - \pi_t + \lambda_I \sigma_I) & \stackrel{(4.33)}{=} \mu_r(t, \tau) - r_t \\ = -B_{rr}(\tau) g_r (\lambda_r - \sigma_I \rho_{I_r}). & \tag{8.64} \end{aligned}$$

□

Proof of Property 6

First we solve the coefficients $B_{rr}(\tau)$ and $A_r(\tau)$ for the second part of the statement (ii).

Adopting the condition (8.64), then replacing μ_r in the equality (8.64) by use of the expression (4.27), and rearranging it, we obtain

$$0 = \left(\frac{d}{d\tau} B_{rr}(\tau) + B_{rr}(\tau) \kappa_r - 1 \right) r_t + \frac{d}{d\tau} A_r(\tau) - B_{rr}(\tau) (\kappa_r \bar{r} - \lambda_r g_r + \sigma_I \rho_{lr} g_r) + \frac{1}{2} g_r^2 B_{rr}(\tau)^2. \quad (8.65)$$

Since r_t can take any arbitrary value, using the method of collecting coefficients, the equation above can hold if and only if

$$\frac{d}{d\tau} B_{rr}(\tau) + B_{rr}(\tau) \kappa_r - 1 = 0, \quad (8.66)$$

$$\frac{d}{d\tau} A_r(\tau) - B_{rr}(\tau) (\kappa_r \bar{r} - \lambda_r g_r + \sigma_I \rho_{lr} g_r) + \frac{1}{2} g_r^2 B_{rr}(\tau)^2 = 0. \quad (8.67)$$

We can check easily that the expression (4.38) solves $B_{rr}(\tau)$ in equation (8.66) and the expression (4.39) solves $A_r(\tau)$ in equation (8.67). More details about the solution technique can be found, for example, in Chiarella (2004).

The first part the model is a multi-factor Gaussian model. The solution is similar to the second part. The solution process can be found, for example, in Brigo and Mercurio (2001).

□

Proof of Property 7

Property (i) holds because $A'_n(0) = \xi_0$ and $B'_{nr}(0) = B'_{n\pi}(0) = 1$. Property (ii) follows directly from (4.33).

□

Proof of Property 8

Since shifting the factor $\hat{\pi}_t$ by adding a constant $\bar{\pi}$ is one special case of the affine transformation in Lemma 23.1 with $\mathcal{L} = 1$ and $\bar{\Theta} = \bar{\pi}$, the result (4.45) is immediately obtained by using the invariant relation (8.61) in Lemma 23.1, where $\xi_1 = 1$ according to the model identification condition (iii) in Property 3. However, we like to give a more direct (but not rigorous) derivation.

Let $A_n(\tau)$, $B_{nr}(\tau)$, and $B_{n\pi}(\tau)$ be the coefficients in (4.37), (4.35) and (4.36) for setting $\bar{\pi} = 0$. and let $A_n^{\bar{\pi}}(\tau)$, $B_{nr}^{\bar{\pi}}(\tau)$, and $B_{n\pi}^{\bar{\pi}}(\tau)$ be those coefficients for taking $\bar{\pi}$ as an arbitrary constant. By using the expression (4.37), we can obtain the relation

$$\frac{A_n^{\bar{\pi}}(\tau)}{\tau} = \frac{A_n(\tau)}{\tau} + \left(1 - \frac{B_{n\pi}^{\bar{\pi}}(\tau)}{\tau}\right)\bar{\pi} + \xi_0^{\bar{\pi}} - \xi_0.$$

Similarly, by using (4.35) and (4.36), we can have $B_{nr}^{\bar{\pi}}(\tau) = B_{nr}(\tau)$, and $B_{n\pi}^{\bar{\pi}}(\tau) = B_{n\pi}(\tau)$.

Let $\hat{\pi}_t$ be the factor when its mean $\bar{\pi}$ is set to be zero. The shifted factor π_t is obtained by $\pi_t = \hat{\pi}_t + \bar{\pi}$. Using all mentioned transformations above, we can express the nominal bond yield based on the shifted factor π_t by

$$\begin{aligned} Y_n(t, T) &= \frac{A_n^{\bar{\pi}}(\tau)}{\tau} + \frac{B_{nr}^{\bar{\pi}}(\tau)}{\tau} r_t + \frac{B_{n\pi}^{\bar{\pi}}(\tau)}{\tau} \pi_t \\ &= \frac{A_n(\tau)}{\tau} - \frac{B_{n\pi}(\tau)}{\tau} \bar{\pi} + \frac{B_{nr}(\tau)}{\tau} r_t + \frac{B_{n\pi}(\tau)}{\tau} (\hat{\pi}_t + \bar{\pi}) + (\xi_0^{\bar{\pi}} - \xi_0 + \bar{\pi}). \end{aligned}$$

We require that the nominal yield formula based on the shifted is equivalent to the original formula given by (4.21)

$$Y_n(t, T) = \frac{A_n(\tau)}{\tau} + \frac{B_{nr}(\tau)}{\tau} r_t + \frac{B_{n\pi}(\tau)}{\tau} \hat{\pi}_t,$$

Then we must have the constraint between the constants

$$\xi_0^{\bar{\pi}} = \xi_0^0 - \bar{\pi}.$$

□

Proof of Property 9

We show how we obtain the adjustment (4.50) for the coefficient $A_r(\tau)$. The idea behind the adjustment (4.50) is to prevent arbitrage possibility in the model as discussed in Section 4.4.3.

From the proof of Property 6, the coefficients $A_r(\tau)$ and $B_{rr}(\tau)$ supporting the no-arbitrage equality (4.32) have to satisfy the equalities (8.66) and (8.67).

The coefficient $B_{rr}(\tau)$ with the parameter separation as given in equation (4.47) does not satisfy equation (8.66) but the equation

$$\frac{d}{d\tau} B_{rr}(\tau) + B_{rr}(\tau) \kappa_r - 1 = (\kappa_r - \kappa_{rr}) B_{rr}(\tau). \quad (8.68)$$

For the other equation (8.67), however, we can set the adjusted coefficient $A_r(\tau)$ as given in (4.50) so that the equation (8.67) still holds. It can be checked easily by integrating the both sides of the equality (8.67).

The adjustment of the coefficient $A_n(\tau)$ follows the similar way.

With the change (8.68), the equation (8.65) is changed into

$$\begin{aligned} & \left(\frac{d}{d\tau} B_{rr}(\tau) + B_{rr}(\tau) \kappa_r - 1 \right) r_t \\ & + \frac{d}{d\tau} A_r(\tau) - B_{rr}(\tau) (\kappa_r \bar{r} - \lambda_r g_r + \sigma_I \rho_{lr} g_r) + \frac{1}{2} g_r^2 B_{rr}(\tau)^2 \\ = & B_{rr}(\tau) (\kappa_r - \kappa_{rr}) r_t . \end{aligned}$$

Applying the formula (4.27) for $\mu_r(t, \tau)$, we can rewrite it further into

$$\mu_r(t, \tau) - r_t = -B_{rr}(\tau) g_r (\lambda_r - \sigma_I \rho_{lr}) + B_{rr}(\tau) (\kappa_r - \kappa_{rr}) r_t . \quad (8.69)$$

Using the transformation from equation (8.64) back to equation (8.63), we can obtain

$$\mu_I(t, \tau) - R_t = -\lambda_r B_{rr}(\tau) g_r + \lambda_I \sigma_I + B_{rr}(\tau) (\kappa_r - \kappa_{rr}) r_t .$$

□

Proof of Property 10

For the agent who is indifferent between investing in the risky asset and the risk-less asset, the utility of the risk-less return is equal to the expected utility of the risky return. Let V be the initial wealth and ε denote the uncertainty with $\mathbf{E}[\varepsilon] = 0$ and $\mathbf{Var}[\varepsilon] = \sigma^2$. The indifference is represented by

$$\begin{aligned} u(V(1+R)) &= \mathbf{E}[u(V(1+\mu+\varepsilon))] \\ \sim u(V) + u'(V)VR &\sim u(V) + u'(V)V\mathbf{E}[\mu+\varepsilon] + \frac{1}{2}u''(V)V^2\mathbf{E}[(\mu+\varepsilon)^2] \\ &= u(V) + u'(V)V\mu + \frac{1}{2}u''(V)V^2(\mu^2 + \sigma^2) . \end{aligned}$$

Some straightforward calculation and rearrangement yields

$$-u'(V)(\mu - R) \sim \frac{1}{2}u''(V)(\mu^2 + \sigma^2) .$$

The property is proved by applying the definition of relative risk aversion (RRA).

□

Proof of Property 11

This property will be proved based on the expectation operator representation (3.25) in Property 1.

In the case with a constant price level, we have $\sigma_I = 0$ and $\pi(X_t) \equiv 0$. For our investment opportunity set consisting only of the nominal bonds, the sources of the return uncertainty are expanded by the sources of the factor innovations W_t^X as shown in (4.7), therefore for the application of the formulas (3.21) for z_t and (3.22) for h_t , we have W_t is identical to W_t^X and $\mathcal{R}_{XX} = \mathcal{R}_{AX} = \mathcal{R}_{AA}$. Recall that the components of the n -dimensional factor innovations W_t^X are independently distributed with each other, therefore \mathcal{R}_{XX} is a unit matrix.

Based on the foregoing discussion, the auxiliary functions z_t given by (3.21) and h_t given by (3.22) are now have the expressions

$$\begin{aligned} z_t &= \frac{1-\gamma}{\gamma} \lambda, \\ h_t &= -\frac{\delta}{\gamma} + \frac{1-\gamma}{\gamma} (\xi_0 + \sum_{i=1}^n X_{it}) + \frac{1-\gamma}{2\gamma^2} \lambda^\top \lambda, \end{aligned}$$

we can obtain

$$\Phi(t, T, x) = \mathbf{E}_{t,x} [\exp(\Psi(t, T))], \quad (8.70)$$

where we use $\Psi_t(t, T)$ to denote

$$\begin{aligned} \Psi(t, T) &:= -\frac{\delta}{\gamma}(T-t) + \frac{1-\gamma}{2\gamma^2} \lambda^\top \lambda(T-t) + \frac{1-\gamma}{\gamma} \int_t^T R_s ds \\ &\quad + \frac{1-\gamma}{\gamma} \lambda^\top (W_T^X - W_t^X) - \frac{(1-\gamma)^2}{2\gamma^2} \lambda^\top \lambda(T-t) \\ &= -\frac{\delta}{\gamma}(T-t) + \frac{1-\gamma}{2\gamma} \lambda^\top \lambda(T-t) + \frac{1-\gamma}{\gamma} \xi_0(T-t) \\ &\quad + \frac{1-\gamma}{\gamma} \int_t^T \left(\sum_{i=1}^n X_{is} ds + \lambda^\top dW_t^X \right). \end{aligned} \quad (8.71)$$

The second equality in equation (8.71) is due to the formula of the model riskless rate R_s , as already given in (4.4)

$$R_s = \xi_0 + \xi_1^\top X_s$$

and the model identification condition (iii) in Property 3. In the subindex of \mathbf{E} t is the initial time and the vector $x = (x_1, \dots, x_n)^\top$ is the initial value of the stochastic process $X_t = x$.

Since the matrix \mathcal{K} is diagonal due to the identification restriction (i) in Property 3, the underlying process (4.6) can be expressed componentwise as

$$dX_{is} = \kappa_i(\theta_i - X_{is})ds + \Gamma_i dW_s^X.$$

The solution of the stochastic differential equation above is given by⁵

$$X_{is} = e^{-\kappa_i(s-t)}X_{it} + \int_t^s e^{-\kappa_i(s-u)}\Gamma_i dW_u^X. \quad (8.72)$$

Thus, the last term of the equation (8.71) becomes

$$\begin{aligned} \int_t^T X_{is}ds &= \int_t^T e^{-\kappa_i(s-t)}X_{it}ds + \int_t^T \int_t^s e^{-\kappa_i(s-u)}\Gamma_i dW_u^X ds \\ &= \frac{1}{\kappa_i}(1 - e^{-\kappa_i(T-t)})X_{it} + \int_t^T \int_u^T e^{-\kappa_i(s-u)}ds\Gamma_i dW_u^X \\ &= B_i(T-t)X_{it} + \int_t^T B_i(T-u)\Gamma_i dW_u^X. \end{aligned}$$

Using this result to rewrite equation (8.71), we obtain

$$\begin{aligned} \Psi(t, T) &= -\frac{\delta}{\gamma}(T-t) + \frac{1-\gamma}{2\gamma}\lambda^\top\lambda(T-t) + \frac{1-\gamma}{\gamma}\xi_0(T-t) \\ &\quad + \frac{1-\gamma}{\gamma}B(T-t)X_t + \frac{1-\gamma}{\gamma} \int_t^T (B(T-u)^\top\Sigma + \lambda^\top) dW_u^X. \end{aligned}$$

It follows that $\Psi(t, T)$ is normally distributed with the expectation

$$\mathbf{E}\Psi(t, T) = -\frac{\delta}{\gamma}(T-t) + \frac{1-\gamma}{2\gamma}\lambda^\top\lambda(T-t) + \frac{1-\gamma}{\gamma}\xi_0(T-t) + \frac{1-\gamma}{\gamma}B(T-t)X_t$$

and the variance

$$\mathbf{Var}\Psi(t, T) = \left(\frac{1-\gamma}{\gamma}\right)^2 \int_t^T (B(T-u)^\top\Sigma + \lambda^\top)(B(T-u)^\top\Sigma + \lambda^\top)^\top ds.$$

Using the well-known result concerning the expected value of the exponential of a normally distributed random variable, we obtain from (8.70) that

$$\Phi(t, T, x) = \mathbf{E}_{t,x}[e^{\Psi(t,T)}] = e^{\mathbf{E}\Psi(t,T) + \frac{1}{2}\mathbf{Var}\Psi(t,T)},$$

⁵See Kloeden and Platen (1992).

which is equivalent to the expression (5.4) in Property 11.

□

Proof of Property 13

The key of the proof is to apply the expectation operator representation (3.25) to solve the value function $\Phi(t, T, r_t, \pi_t)$, where now the factor set consists of the instantaneous interest rate r_t and the expected instantaneous inflation rate π_t . According to the dynamics of the asset returns included in the investment opportunity set, the equations (3.22) and (3.21) are now given by

$$h_t = \frac{1-\gamma}{\gamma} r_t + j_t, \quad (8.73)$$

$$j_t = -\frac{\delta}{\gamma} + \frac{1-\gamma}{\gamma} (-\lambda_I \sigma_I + \sigma_I^2) + \frac{1-\gamma}{2\gamma^2} \lambda^\top \mathcal{R}_{AA}^{-1} \lambda + \frac{(1-\gamma)^3}{2\gamma^2} \sigma_I^2 \mathcal{R}_{IA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} - \frac{(1-\gamma)^2}{\gamma^2} \lambda^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I - \frac{1-\gamma}{2} \sigma_I^2, \quad (8.74)$$

$$z_t = \frac{1-\gamma}{\gamma} \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \lambda - \frac{(1-\gamma)^2}{\gamma} \mathcal{R}_{XA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I - (1-\gamma) \mathcal{R}_{XI} \sigma_I. \quad (8.75)$$

Comparing the the equation h_t above with (3.22), the difference is due to the no-arbitrage equality (4.40).

It is easy to observe that j_t (8.74) and z_t (8.75) are actually constants because of the constant market price of risk and constant correlation matrices. To stress this, we omit the subindex t .

An remarkable feature of the solution structure is that the second factor π_t does not appear in the equations (8.73) and (8.75) anymore due to the replacement based on the arbitrage equality (4.33). So we can expect that the value function $\Phi(t, T, r_t, \pi_t)$ will be independent of π_t .

We note in (8.74) that $\mathcal{R}_{IA} \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} = 1$ and $\lambda^\top \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} \sigma_I = \lambda_I \sigma_I$. This is because

$$\mathcal{R}_{AA}^{-1} \mathcal{R}_{AX} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathcal{R}_{AA}^{-1} \mathcal{R}_{AI} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}. \quad (8.76)$$

Recall the matrix \mathcal{R}_{AA} is the correlation matrix of uncertainty sources of the asset returns, which are $W_t^r, W_t^\pi, W_t^I, W_t^S$, and \mathcal{R}_{AX} is that of the asset returns and factors W_t^r, W_t^π , so \mathcal{R}_{AX} consists of the first two columns of \mathcal{R}_{AA} and \mathcal{R}_{AI} is exactly the third columns of \mathcal{R}_{AA} . That explains the equations

(8.76).

Using the matrix identities above to rewrite (8.74), we can obtain the result (5.17).

In the expression for z in (8.75) we have

$$\mathcal{R}_{XA}\mathcal{R}_{AA}^{-1}\lambda = \begin{pmatrix} \lambda_r \\ \lambda_\pi \end{pmatrix},$$

and

$$\mathcal{R}_{XA}\mathcal{R}_{AA}^{-1}\mathcal{R}_{AI} = \mathcal{R}_{XI} = \begin{pmatrix} \rho_r I \\ \rho_\pi I \end{pmatrix}.$$

Using these two equalities above we obtain (5.18).

Because z is constant, the Radon-Nikodym derivative (3.26) can be rewritten as

$$\mathbf{E}_t[\Lambda_T] = \exp\left(z^\top \mathcal{R}_{XX}^{-1}(W_T^X - W_t^X) - \frac{1}{2}z^\top \mathcal{R}_{XX}^{-1}z(T-t)\right). \quad (8.77)$$

Using the notation $\mathcal{C}\mathcal{C}^\top = \mathcal{R}_{XX}$ to rewrite (8.77) and letting

$$\hat{z} = \mathcal{C}^{-1}z = \begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix}, \quad \hat{W}_t^X = \mathcal{C}^{-1}W_t^X = \begin{pmatrix} \hat{W}_{1t}^X \\ \hat{W}_{2t}^X \end{pmatrix},$$

we have

$$\mathbf{E}_t[\Lambda_T] = \exp\left(\hat{z}^\top (\hat{W}_T^X - \hat{W}_t^X) - \frac{1}{2}\hat{z}^\top \hat{z}(T-t)\right).$$

Note that \hat{W}_t^X is an orthogonal Wiener process because $\mathbf{Var}[\hat{W}_1^X] = \mathcal{C}^{-1}\mathcal{R}_{XX}\mathcal{C}^{-1\top} = I_n$.

The solution for r_t is given by⁶

$$r_s = e^{-\kappa_r(s-t)}r_t + \bar{r}(1 - e^{-\kappa_r(s-t)}) + g_r \int_t^s e^{-\kappa_r(s-u)}dW_u^r.$$

Using this solution and Fubini's theorem, we calculate

$$\begin{aligned} \int_t^T r_s ds &= (r_t - \bar{r}) \int_t^T e^{-\kappa(s-t)} ds + \bar{r}(T-t) + g_r \int_t^T \int_u^T e^{-\kappa(s-u)} ds dW_u^r \\ &= B_r(t, T)r_t + \bar{r}(T-t - B_r(t, T)) + g_r \int_t^T B_r(u, T)dW_u^r, \end{aligned} \quad (8.78)$$

⁶See for example Kloeden and Platen (1992).

where

$$B_r(t, T) = \frac{1}{\kappa_r}(1 - e^{-\kappa_r(T-t)}) .$$

Summarizing all the above calculations we can rewrite $\Phi(t, T, r_t)$ as

$$\Phi(t, T, r_t) = \mathbf{E}_{t,x}[\exp \mathcal{Y}(t, T)] ,$$

where

$$\begin{aligned} & \mathcal{Y}(t, T) \\ := & \frac{1-\gamma}{\gamma} B_r(T-t)r_t + \frac{1-\gamma}{\gamma} \bar{r}(T-t - B_r(T-t)) + h(T-t) - \frac{1}{2} \hat{z}^\top \hat{z}(T-t) \\ & + \int_t^T \left(\frac{1-\gamma}{\gamma} g_r B_r(T-u) + \hat{z}_1 \right) d\hat{W}_{1u}^X + \hat{z}_2 (\hat{W}_{2T}^X - \hat{W}_{1t}^X) . \end{aligned} \quad (8.79)$$

Note that $\mathcal{Y}(t, T)$ is normally distributed with the mean and the variance given by

$$\begin{aligned} \mathbf{E}_{t,x}[\mathcal{Y}(t, T)] &= \frac{1-\gamma}{\gamma} B_r(T-t)r_t + \frac{1-\gamma}{\gamma} \bar{r}(T-t - B_r(T-t)) + h(T-t) \\ &\quad - \frac{1}{2} \hat{z}^\top \hat{z}(T-t) , \\ \mathbf{Var}_{t,x}[\mathcal{Y}(t, T)] &= \int_t^T \left(\frac{1-\gamma}{\gamma} g_r B_r(T-u) + \hat{z}_1 \right)^2 du + \hat{z}_2^2 (T-t) . \end{aligned}$$

Using the equality

$$\mathbf{E}_{t,x}[\exp(\mathcal{Y}(t, T))] = \exp\left(\mathbf{E}_{t,x}[\mathcal{Y}(t, T)] + \frac{1}{2} \mathbf{Var}_{t,x}[\mathcal{Y}(t, T)]\right) ,$$

we obtain the result (5.16).

□

Proof of Property 14

The first step is to insert the model specific parameters and constants into the optimal portfolio solution (3.18) in the general framework, where we use the substitutions (8.76) again. The rest of the proof is to apply the result (5.20) and then the result (5.22) can be obtained.

□

Proof of Property 15

This property can be easily proved by providing the inverse of the asset volatility matrix Σ_t^\top given in (5.14)

$$(\Sigma_t^\top)^{-1} = \begin{pmatrix} -\frac{B_{n\pi}(\tau_2)}{g_r \mathcal{D}} & \frac{B_{nr}(\tau_2)}{g_\pi \mathcal{D}} & -\frac{B_{rr}(\tau_3)B_{nr}(\tau_2)}{\sigma_I \mathcal{D}} & 0 \\ -\frac{B_{n\pi}(\tau_1)}{g_r \mathcal{D}} & \frac{B_{nr}(\tau_1)}{g_\pi \mathcal{D}} & -\frac{B_{rr}(\tau_3)B_{nr}(\tau_1)}{\sigma_I \mathcal{D}} & 0 \\ 0 & 0 & \frac{1}{\sigma_I} & 0 \\ 0 & 0 & 0 & \frac{1}{\sigma_S} \end{pmatrix}$$

where

$$\mathcal{D} := \det \begin{pmatrix} B_{nr}(\tau_1) & B_{nr}(\tau_2) \\ B_{n\pi}(\tau_1) & B_{n\pi}(\tau_2) \end{pmatrix}.$$

□

Proof of Property 18

The proof goes analogously to the proof of Property 13. The difference to the previous proof is that now different correlation matrices \mathcal{R}_{AA} , \mathcal{R}_{AI} , and \mathcal{R}_{AX} are inserted in the expressions (8.73), (8.74) and (8.75). The asset return innovations have now three sources W_t^r , W_t^π , and W_t^S . The innovation of the price index W_t^I does not appear in the set of asset return uncertainty due to the exclusion of the IIBs.

The substitution of the different correlation matrices leads a change of the constant j and z given in (8.74) and (8.75) but not change the basic form given in (8.73) in terms of the factor r_t . So, the value function in this case will share the same form given in (5.16) and therefore has the same expression of the factor elasticity (5.20).

□

Proof of Property 19

The result (5.27) is obtained simply by inserting the model specific constants into the general solution (3.18) and then applying the result of Property (18).

□

Proof of Property 20

This property can be easily proved by providing the inverse of the asset volatility matrix Σ_t^\top given in (5.14)

$$(\Sigma_t^\top)^{-1} = \begin{pmatrix} -\frac{B_{n\pi}(\tau_2)}{g_r \mathcal{D}} & \frac{B_{nr}(\tau_2)}{g_\pi \mathcal{D}} & 0 \\ -\frac{B_{n\pi}(\tau_1)}{g_r \mathcal{D}} & \frac{B_{nr}(\tau_1)}{g_\pi \mathcal{D}} & 0 \\ 0 & 0 & \frac{1}{\sigma_S} \end{pmatrix},$$

where \mathcal{D} is given in (5.24).

□

The Kalman Filter

We employ the maximum likelihood estimation based on the Kalman filter to estimate the real interest rate.

The Kalman filter is applied to a model of *state space expression*⁷ which consists of a *measurement equation*

$$y_t = Z_t X_t + d_t + \varepsilon_t , \quad (8.80)$$

and a *transition equation*

$$X_t = T_t X_{t-1} + c_t + R_t \eta_t . \quad (8.81)$$

The variable of interest y_t is observable and is explained by an observable component d_t and an *unobservable state variable* X_t which follows the dynamics (8.81). The Kalman filter is an algorithm to formulate the best linear projection of X_t on the observed variables y_t and d_t .

⁷See Harvey(1990) or and Hamilton(1994).

Bibliography

- [1] Akaike, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, AC 19:716–723, 1974.
- [2] Amin, K.I. and R. Jarrow. Pricing foreign currency options under stochastic interest rates. *Journal of International Money and Finance*, 10:310–329, 1991.
- [3] Babbs, S.H. and K.B. Nowman. Kalman filtering of generalized Vasicek term structure models. *Journal of Financial and Quantitative Analysis*, 34(1):115–130, March 1999.
- [4] Black, F. and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973.
- [5] Brace, A., D. Gatarek, and M. Musiela. The market model of interest rate dynamics. *Mathematical Finance*, 7(2):127–155, 1997.
- [6] Brennan, M.J. and E.S. Schwartz. A continuous time approach to the pricing of bonds. *Journal of Banking and Finance*, 3:133–155, 1979.
- [7] Brennan, M.J., E.S. Schwartz, and R. Lagnado. Strategic asset allocation. *Journal of Economic Dynamics and Control*, 21:1377–1403, 1997.
- [8] Brennan, M.J., A.W. Wang, and Y. Xia. Estimation and test of a simple model of intertemporal capital asset pricing. *The Journal of Finance*, 59(4):1743–75, August 2004.
- [9] Brennan, M.J. and Y. Xia. Dynamic asset allocation under inflation. *The Journal of Finance*, 57(3):1201–1238, 2002.
- [10] Brigo, D. and F. Mercurio. *Interest Rate Models, Theory and Practice*. Springer, 2001.

- [11] Camilli, F. and M. Falcone. An approximation scheme for the optimal control of diffusion processes. *Mathematical Modelling and Numerical Analysis*, 29(1):97–122, 1995.
- [12] A.W. Lo Campbell, J.W. and C. MacKinlay. *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- [13] Campbell, J.Y. , G. Chacko, J. Rodriguez, and L.M. Viceira. Strategic asset allocation in a continuous-time VAR model. *Journal of Economic Dynamics and Control*, 28:2195–2214, 2004.
- [14] Campbell, J.Y., Y.L. Chan, and L.M. Viceira. A multivariate model of strategic asset allocation. *Journal of Financial Economics*, 67:41–80, 2003.
- [15] Campbell, J.Y. and L.M. Viceira. Consumption and portfolio decisions when expected returns are time varying. *The Quarterly Journal of Economics*, 114:433–496, May 1999.
- [16] Campbell, J.Y. and L.M. Viceira. Who should buy long term bonds? *American Economic Review*, 9:99–127, 2001.
- [17] Canner, N, N.G. Mankiw, and D.N.Weil. An asset allocation puzzle. *The American Economic Review*, 87:181–191, 1997.
- [18] Chiarella, C. *An Introduction to Derivative Security Pricing, Lecture Notes, UTS*, 2004.
- [19] Chung, K.L. *A Course in Probability Theory*. Stanford University Press, 1974.
- [20] Cox, J.C. , J.E. Ingersoll, and S.A. Ross. An intertemporal general equilibrium model of asset prices. *Econometrica*, 53(2):363–384, March 1985.
- [21] Cox, J.C. , J.E. Ingersoll, and S.A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, March 1985.
- [22] Cox, J.C. and C-F. Huang. Optimal consumption and portfolio policies when asset prices follow a diffusion process. *Journal of Economic Theory*, 49:33–83, 1989.
- [23] Dai, Q. and K. Singleton. Specification analysis of affine term structure models. *Journal of Finance*, 55:1943–78, 2000.

- [24] de Jong, F. Time series and cross-section information in affine term structure models. *Journal of Business and Economic Statistics*, 18(3):300–314, 2000.
- [25] Duffie, D. and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6(4):397–406, 1996.
- [26] Epstein, L. and S. Zin. Substitution, risk aversion, and the temporal behavior of consumption and asset returns: An empirical investigation. *Journal of Political Economy*, 99(2):263–286, 1991.
- [27] Gruene, L. An adaptive grid scheme for the discrete Hamilton-Jacobi-Bellman equation. *Numerische Mathematik*, 75:319–337, 1997.
- [28] Gruene, L. *Lecture Note on Numerische Optimale Steuerung*. Frankfurt University, <http://www.math.uni-frankfurt.de/numerik/lehre/gruene/optctrl01/skript.pdf>, 2001.
- [29] Gruene, L. and W. Semmler. Solving asset pricing models with stochastic dynamic programming. Working Papers No.54, Center for Empirical Macroeconomics, University Bielefeld, Germany, 2003.
- [30] Hamilton, J. *Time Series Analysis*. Princeton, 1994.
- [31] Hannan, E.J. and B.G.Quinn. The determination of the order of an autoregression. *Journal of the Royal Statistical Society, B*, 41:190–195, 1979.
- [32] Harrison, J.M. and D.M. Kreps. Martingales and arbitrage in multi-period securities markets. *Journal of Economic Theory*, 20:381–408, 1979.
- [33] Harrison, J.M. and S. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and Their Applications*, 11:215–260, 1981.
- [34] Harvey, A.C. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge, 1990.
- [35] Heath, D, R. Jarrow, and A. Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60(1):77–105, 1992.
- [36] Hicks, J. R. *Value and Capital*. Oxford University Press, 1939.

- [37] Hull, J. and A. White. Numerical procedures for implementing term structure models II: Two-factor models. *Journal of Derivatives*, 2:37–48, 1994.
- [38] Jarrow, R. and Y. Yildirim. Pricing treasury inflation protected securities and related derivations using an HJM model. *Journal of Financial and Quantitative Analysis*, 38(2):337–358, 2003.
- [39] Jegadeesh, N. and G.G. Pennacchi. The behavior of interest rates implied by the term structure of eurodollar futures. *Journal of Money, Credit and Banking*, 28(3):426–446, 1996.
- [40] Kamien, M.I. and N.L. Schwartz. *Dynamic Optimization*. North-Holland, 1991.
- [41] Karatzas, I., J.P. Lechoczky, S.E. Shreve, and G. Xu. Martingale and duality methods for utility maximization in an incomplete market. *SIAM J. Control and Optimization*, 29(3):702–730, 1991.
- [42] Karatzas, I., J.P. Lehoczky, and S.E.Shreve. Optimal portfolio and consumption decisions for a small investor on a finite horizon. *SIAM Journal of Control and Optimization*, 25(6):1557–1586, 1987.
- [43] Karatzas, I. and S.E. Shreve. *Bronian Motion and Stochastic Calculus*. Springer, 1991.
- [44] Kim, T.S. and E. Omberg. Dynamic nonmyopic portfolio behavior. *The Review of Financial Studies*, 9(1):141–161, 1996.
- [45] Klebaner, F. *Introduction to Stochastic Calculus with Applications*. Imperial College Press, 2005.
- [46] Kloeden, P.E. and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1992.
- [47] Korn, R. and E. Korn. *Option Pricing and Portfolio Optimization*. American Mathematical Society, 2000.
- [48] H.J. Kushner. Consistency issues for numerical methods for variance control, with application to optimization in finance. *IEEE Trnsactions on Automatic Control*, 44(12):2283–2296, 1999.
- [49] Kushner, H.J. *Probability Methods for Approximations in Stochastic Control and for Elliptic Equations*. Academic Press, New York, 1977.

- [50] Kushner, H.J. and P. Dupuis. *Numerical Methods for Stochastic Control Problems in Continuous Time*. Springer, second edition, 2000.
- [51] Lintner, J. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47:13–37, 1965.
- [52] Liu, J. Portfolio selection in stochastic environments. *Review of Financial Studies*, forthcoming.
- [53] Liu, J. and J. Pan. Dynamic derivative strategies. *Journal of Financial Economics*, 69:401–430, 2003.
- [54] Longstaff, F.A. and E.S. Schwartz. Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance*, 47(4):1259–1281, 1992.
- [55] Markowitz, H.M. *Portfolio Selection: Efficient Diversification of Investments*. Wiley, New York, 1959.
- [56] Merton, R.C. Optimum consumption and portfolio rules in a continuous-time model. *Journal of Economic Theory*, 3:373–413, 1971.
- [57] Merton, R.C. An intertemporal capital asset pricing model. *Econometrica*, 41(5):867–887, 1973.
- [58] Merton, R.C. *Continuous Time Finance*. Blackwell, 1990.
- [59] Modigliani, F. and R. Sutch. Innovations in interest rate policy. *American Economic Review*, 56:178–197, 1966. Papers and Proceedings.
- [60] Munk, C. The Markov chain approximation approach for numerical solution of stochastic control problems: Experiences from Merton’s problems. *Applied Mathematics and Computation*, 136(1):47–77, 2003.
- [61] Munk, C., C. Sørensen, and T.N. Vinther. Dynamic asset allocation under mean-reverting returns, stochastic interest rates and inflation uncertainty: Are popular recommendations consistent with rational behavior? *International Review of Economics and Finance*, 13(141–166), 2004.
- [62] Oksendal, B. *Stochastic Differential Equations*. Springer, 6 edition, 2000.

- [63] Pliska, S.R. A stochastic calculus model of continuous trading: Optimal portfolios. *Mathematics of Operations Research*, 11(2):371–382, 1986.
- [64] Rebonato, R. *Interest-Rate Option Models*. Wiley, 1998.
- [65] Richard, S.F. An arbitrage model of the term structure of interest rates. *Journal of Financial Economics*, 6:33–57, 1978.
- [66] Samuelson, P.A. Lifetime portfolio selection by dynamic stochastic programming. *The Review of Economics and Statistics*, 51:239–246, 1969.
- [67] Santos, M.S. and J. Vigo-Aguiar. Analysis of a numerical dynamic programming algorithm applied to economic models. *Econometrica*, 66(2):409–426, 1998.
- [68] Schwartz, G. Estimating the dimension of a model. *Annals of Statistics*, 6:461–464, 1978.
- [69] Sharpe, W.F. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.
- [70] Stokey, N.L. and Jr. R.E. Lucas. *Recursive Methods in Economic Dynamics*. Harvard University Press, 1989.
- [71] Tapiero, C.S and A. Sulem. Computational aspects in applied stochastic control. *Computational Economics*, 7:109–146, 1994.
- [72] Tobin, J. Liquidity preference as behavior towards risk. *Review of Economic Studies*, 25:65–86, 1958.
- [73] Vasicek, O. An equilibrium characterization of the term structure. *Journal of financial economics*, 5:177–188, 1977.
- [74] Wachter, J.A. Portfolio and consumption decisions under mean-reverting returns: An exact solution for complete markets. *Journal of Financial and Quantitative Analysis*, 37(1):63–91, March 2002.
- [75] Wachter, J.A. Risk aversion and allocation to long-term bonds. *Journal of Economic Theory*, 112:325–333, 2003.

May 2007

Curriculum Vitae

Chih-Ying Hsiao

Center for Empirical Macroeconomics
Bielefeld University, Germany
email: chsiao@uni-bielefeld.de

Personal Information

Place of Birth: Taiwan
Citizenship: Taiwanese
Gender: Female

Education

2007 PhD in Economics, Bielefeld University, Germany.

PhD thesis: *Intertemporal Asset Allocation under Inflationary Risk* under supervision of Professor Willi Semmler and Professor Carl Chiarella.

2001 Diploma in Mathematics with degree “excellent”,
Bielefeld University, Germany.

Final Exams in Topics: Probability Theory, Mathematical Statistics and Stochastics, Numerical Methods and Econometrics. Diploma Thesis: *Study of Maximum Likelihood Estimation for Drift Parameters in Stochastic Differential Equations*.

1994 BSc in Mathematics, National Taiwan University, Taiwan

Professional Experience

2006 (Feb.–Dec.) *Research Associate* in School of Finance and Economics, University of Technology, Sydney, under the supervision of Professor Carl Chiarella working on the projects ” *The implementation of optimal investment decision strategies in finance*” and ” *The implementation of continuous time portfolio optimization strategies with a focus on interest rate term structure models and numerical aspects*” funded by the ARC Discovery Grant DP0452996.

2002 – 2005 *Lecturer* in the Center of Empirical Macroeconomics, Bielefeld University, Germany (currently on leave).

Courses taught include:

- Seminar series for “*Empirical research on financial markets and economic activities*” (with Professor W. Semmler) in SS02, WS02/03, SS04.
- Assistant teaching for Lecture ”Financial markets and economic activities” WS03, SS04.
- Workshop for “*Introduction to financial econometrics with GAUSS programming*” in WS02/03, SS04.
- Lecture “*Intertemporal optimization in economics and finance*”.

1998 – 2001 *Research Assistant* in Center of Empirical Macroeconomics, Bielefeld University, Germany.

Research subjects: Modelling interest rates, estimation for continuous-time and discrete-time models.

Refereed Publications

R1 Chiarella, C. and C.Y. Hsiao (2006), “*The impact of Short-Sale Constraints on Asset Allocation Strategies via the Backward Markov Chain Approximation Method*”, in *Computational Economics* Vol. 28, No. 2, September, 2006.

R2 Chiarella, C., C.Y. Hsiao and W. Semmler (2007), “*Intertemporal As-*

set Allocation when Underlying Factors are Unobservable” in *Computational Economics*, accepted 30. October, accessible online first under the link “<http://www.springerlink.com/content/3737x74812778565/>”.

- R3** Chiarella, C., C.Y. Hsiao and W. Semmler (2006), “*Intertemporal Asset Allocation under Inflation Risk*”, under review at *Computational Statistics and Data Analysis*.

Working Papers

- W1** Hsiao, C.Y. and W. Semmler (2002), “Modelling Short Term Interest Rates”, CEM Working Papers No.39
(<http://www.wiwi.uni-bielefeld.de/~cem/>)
- W2** Chiarella, C, Hsiao, C.Y. and W. Semmler (2004), “Strategic Asset Allocation with Arbitrage-Free Bond Market using Dynamic Programming”, *CEM Working Paper No. 76*
(<http://www.wiwi.uni-bielefeld.de/~cem/>)
- W3** Chen, P. and C.Y. Hsiao (2004), “Testing Weak Exogeneity in Cointegrated System”,
<http://ideas.repec.org/p/ecm/feam04/537.html>
- W4** Chen, P. and C.Y. Hsiao (2005), “Time series causal models”, *Bielefeld University*.
<http://www.wiwi.uni-bielefeld.de/%7Efrohn/Mitarbeiter/chenpu/chenpu/TScausal1.pdf>
- W5** Chen, P. and C.Y. Hsiao (2005), “Observational differentiability in simultaneous equations models”, *Bielefeld University*.
- W6** Chen, P., P. Flaschel and C.Y. Hsiao (2005), “Causal analysis of wage price spiral”, *Bielefeld University*.
<http://www.wiwi.uni-bielefeld.de/%7Efrohn/Mitarbeiter/chenpu/chenpu/TSWP.pdf>
- W7** Chen, P. and C.Y. Hsiao (2006), “What happens to Japan when China catches cold? A causal analysis of the Chinese growth and the Japanese growth”, revised and re-submitted with “*Japan and the World Economics*”.

W8 Hsiao, C.Y., and P. Chen (2006) "Nonarbitrage in the Chinese term structure" *Bielefeld University*.
<http://www.wiwi.uni-bielefeld.de/%7Efrohn/Mitarbeiter/chenpu/chenpu/TStex1.pdf>

Conference and Seminar Presentations

- Hsiao, C.Y. and W. Semmler, "Continuous-Time Maximum Likelihood Estimation", *Workshop on Financial Markets and the Macroeconomy, Bielefeld, Germany, December 1999*
- Hsiao, C.Y. and W. Semmler, "Maximum Likelihood Estimation of SDE Dynamics Based on Discrete Time Data – How well does the Euler Method Perform?" *Center for Nonlinear Dynamics in Economics and Finance Workshop, Amsterdam, Netherlands, January 2001*
- Hsiao, C.Y. and W. Semmler, "Estimating Diffusion Process Based on Discrete Time Observations", *9th Annual Symposium of the Society of Nonlinear Dynamics and Econometrics, Atlanta, March 2001*
- Chiarella, C., C.Y. Hsiao and W. Semmler, "Strategic Asset Allocation with an Arbitrage-Free Bond Market using Dynamic Programming", *10th International Conference on Computing in Economics and Finance, Amsterdam, July 2004*.
- Chiarella, C., C.Y. Hsiao and W. Semmler, "Strategic Asset Allocation under Inflation", *Quantitative Methods in Finance Conference 2004, Sydney, Dec. 2004*.
- Chiarella, C., C.Y. Hsiao and W. Semmler, "Intertemporal Asset Allocation when the Underlying Factors are Unobservable", *13th Annual Symposium of the Society of Nonlinear Dynamics and Econometrics, London, March 2005*.
Also in *Quantitative Economics Doctorate, Bielefeld, Germany, April 2005*,
and in *Seminar in Institute for Physics, University Oldenburg, Germany, Mai 2005*.
- Chiarella, C., C.Y. Hsiao and W. Semmler, "Intertemporal Asset Allocation with Inflation-Indexed Bonds", *11th International Conference on Computing in Economics and Finance, Washington D.C, June 2005*.

- Chiarella, C., C.Y. Hsiao and W. Semmler, “Intertemporal Investment Strategies under Inflation Risk”, *Seminar Series in School of Banking and Finance in UNSW, Sydney, Sept. 2006*.
- Chiarella, C., C.Y. Hsiao and W. Semmler, “Intertemporal Asset Allocation under Inflation Risk”, *Quantitative Methods in Finance Conference 2006, Sydney, Dec. 2006*.

Other Experiences

- Member in International Students Organization, Bielefeld University, 1999.
- Designer of one project in ”Orientation-Phase” for new coming students for the faculty of Economics and Business Administration, 2002.

Additional Information

- Computational Skills: GAUSS (proficient), Fortran (competent).
- Mathematical and Statistical Programming: Mathematica, R (both competent).
- Editing Software: Latex, Word (both proficient)
- General Software Skills: MS-Offices, Linux.

- Native Language: Chinese; fluent in German and English.