Dynamic Asset Allocation*

Claus Munk†

This version: April 4, 2003

*Informal lecture notes originally prepared for Ph.D. course on Continuous-time Finance, Danish Doctoral Research Programme in Finance, Spring 2001. I appreciate comments and corrections from Heine Jepsen and Nicolai Nielsen.

†Department of Accounting and Finance, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark. Phone: 6550 3257. Fax: 6593 0726. E-mail: cmu@sam.sdu.dk. Internet homepage: http://www.sam.sdu.dk/~cmu

Contents

Chapter 1

Introduction

Financial markets offer opportunities to move money between different points in time and different states of the world. Investors must decide how much to invest in the financial markets and how to allocate that amount between the many, many available financial securities. Investors can change their investments as time passes and they will typically want to do so for example when they obtain new information about the prospective returns on the financial securities. Hence, they must figure out how to manage their portfolio over time. In other words, they must determine a dynamic investment or asset allocation strategy. The term asset allocation is sometimes used for the allocation of investments to major asset classes, e.g. stocks, bonds, and cash. In later chapters we will often focus on this decision, but we will use the term asset allocation interchangeably with the terms optimal investment or portfolio management.

It is intuitively clear that in order to determine the optimal investment strategy for an investor, we must make some assumptions about the objectives of the investor and about the possible returns on the financial markets. Different investors will have different motives for investments and hence different objectives. In Section 1.1 we will discuss the motives and objectives of different types of investors. We will focus on the asset allocation decisions of individual investors or households. Individuals invest in the financial markets to finance future consumption of which they obtain some felicity or utility. We discuss how to model the preferences of individuals in Chapter 2.

1.1 Investor classes and motives for investments

We can split the investors into individual investors (households; sometimes called retail investors) and institutional investors (includes both financial intermediaries – such as pension funds, insurance companies, mutual funds, and commercial banks – and manufacturing companies producing goods or services). Different investors have different objectives. Manufacturing companies probably invest mostly in short-term bonds and deposits in order to their liquidity needs and avoid the deadweight costs of raising small amounts of capital very frequently. They will rarely set up long-term strategies for investments in the financial markets and their financial investments is a very small part of the total investments.

Individuals can use their money either for consumption or savings. Here we use the term savings synonymously with financial investments so that it includes both deposits in banks and investments in stocks, bonds, and possibly other securities. Traditionally most individuals have saved in form of bank deposits and maybe government bonds, but in recent years there has been an increasing interest of individuals for investing in the stock market. Individuals typically save when they are young by consuming less than the labor income they earn, primarily in order to accumulate wealth they can use for consumption when they retire. Other motives for saving is to be able to finance large future expenditures (e.g. purchase of real estate, support of children during their education, expensive celebrations or vacations) or simply to build up a buffer for "hard times" due to unemployment, disability, etc. The objective of an individual investor is to maximize the utility of consumption throughout the life-time of the investor. We will discuss utility functions in Chapter 2.

A large part of the savings of individuals are indirect through pension funds and mutual funds. These funds are the major investors in today's markets. Some of these funds are non-profit funds that are owned by the investors in the fund. The objective of such funds should represent the objectives of the fund investors.

Let us look at pension funds. One could imagine a pension fund that determines the optimal portfolio of each of the fund investors and aggregates over all investors to find the portfolio of the fund. Each fund investor is then allocated the returns on his optimal portfolio, probably net of some servicing fee. The purpose of the fund is then simply to save transaction costs. A practical implementation of this is to let each investor allocate his funds among some pre-selected portfolios, for example a portfolio mimicking the overall stock market index, various portfolios of stock in different industries, one or more portfolios of government bonds (e.g. one in short-term and one in long-term bonds), portfolios of corporate bonds and mortgage-backed bonds, portfolios of foreign stocks and bonds, and maybe also portfolios of derivative securities and even non-financial portfolios of metals and real estate. Some pension funds operate in this way and there seems to be a tendency for more and more pension funds to allow investor discretion with regards to the way the deposits are invested.

However, in many pension funds some hired fund managers decide on the investment strategy. Often all the deposits of different fund members are pooled together and then invested according to a portfolio chosen by the fund managers (probably following some general guidelines set up by the board of the fund). Once in a while the rate of return of the portfolio is determined and the deposit of each investor is increased according to this rate of return less some servicing fee. In many cases the returns on the portfolio of the fund are distributed to the fund members using more complicated schemes. Rate of return guarantees, bonus accounts,.... The salary of the manager of a fund is often linked to the return on the portfolio he chooses and some benchmark portfolio(s). A rational manager will choose a portfolio that maximizes his utility and that portfolio choice may be far from the optimal portfolio of the fund members....

Mutual funds...

This lecture note will focus on the decision problem of an individual investor and aims to analyze and answer the following questions:

- What are the utility maximizing dynamic consumption and investment strategies of an individual?
- What is the relation between optimal consumption and optimal investment?
- How are financial investments optimally allocated to different asset classes, e.g. stocks and bonds?

- How are financial investments optimally allocated to single securities within each asset class?
- How does the optimal consumption and investment strategies depend on, e.g., risk aversion, time horizon, initial wealth, income, and asset price dynamics?
- Are the recommendations of investment advisors consistent with the theory of optimal investments?

1.2 Typical investment advice

1.3 An overview of the theory of optimal investments

1.4 The future of investment management and services

See Bodie (2003), Merton (2003)

1.5 Outline of the rest

1.6 Notation

Since we are going to deal simultaneously with many financial assets, it will often be mathematically convenient to use vectors and matrices. All vectors are considered column vectors. The superscript $^\top$ on a vector or a matrix indicates that the vector or matrix is transposed. We will use the notation $\mathbf{1}$ for a vector where all elements are equal to 1 – the dimension of the vector will be clear from the context. We will use the notation $\mathbf{e}_i$ for a vector $(0, \dots, 0, 1, 0, \dots, 0)^\top$ where the 1 is entry number i. Note that for two vectors $x = (x_1, \dots, x_d)^\top$ and $y = (y_1, \dots, y_d)^\top$ we have $x^\top y = \sum_{i=1}^d x_i y_i$. Also, $x^\top \mathbf{1} = \sum_{i=1}^d x_i$ and $\mathbf{e}_i^\top x = x_i$.

If $x = (x_1, \dots, x_n)$ and f is a real-valued function of x, then the (first-order) derivative of f with respect to x is the vector

$$f'(x) = \left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right)^\top .$$

This is also called the gradient of f. The second-order derivative of f is the $n \times n$ Hessian matrix

$$f''(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix} .$$

If x and a are n-dimensional vectors, then

$$\frac{\partial}{\partial x} \left(a^\top x \right) = a.$$

If x is an n-dimensional vector and A is a symmetric [i.e. $A = A^\top$] $n \times n$ matrix, then

$$\frac{\partial}{\partial x} \left(x^\top A x \right) = 2Ax.$$

Chapter 2

Preferences

There is a large literature on how to model the preferences of individuals for uncertain outcomes. The literature dates back at least to Daniel Bernoulli in 1738 (see English translation in Bernoulli (1954)), but was put on a firm formal setting by von Neumann and Morgenstern (1944). Some recent textbook presentations are given by Huang and Litzenberger (1988, Ch. 1) and Kreps (1990, Ch. 3). The short presentation below mainly follows that of Huang and Litzenberger.

2.1 Expected utility representation of preferences

2.1.1 Basic representation of preferences

- Assume a single consumption good and a one-period economy with uncertainty about the state $\omega \in \Omega$ at time 1 (end-of-period). The probabilities of the states of nature are given [objective].
- A **consumption plan** x is a specification of the number of units consumed in each state, $x = (x_\omega | \omega \in \Omega)$. The set of possible consumption plans is denoted X.
- Assume there are finitely many possible consumption levels, i.e. a finite set $Z \subseteq \mathbb{R}$ exist such that $x_\omega \in Z$ for all $\omega \in \Omega$ and all $x \in X$.
- A **preference relation** $\succeq$ on $\mathbf{P}$ is a binary relation satisfying
  (i) $p \succeq q$ and $q \succeq r \Rightarrow p \succeq r$ [transitivity]
  (ii) $\forall p, q \in \mathbf{P} : p \succeq q$ or $q \succeq p$ [completeness]
- Derived relations $\sim$, $\not\succeq$, and $\succ$ ...
- A probability distribution on Z is a function $p : Z \to [0,1]$ such that $p(z) \geq 0, \forall z \in Z$ and $\sum_{z \in Z} p(z) = 1$.
- Assume that the preferences of an individual can represented by a preference relation $\succeq$ on the set $\mathbf{P} \equiv \mathbf{P}(Z)$ of probability distributions (aka. lotteries) defined on Z. [...state-independence]
- A **utility function** is a function $\mathcal{U} : \mathbf{P} \to \mathbb{R}$ such that

$$p \succeq q \Leftrightarrow \mathcal{U}(p) \geq \mathcal{U}(q).$$

state ω	ω_1	ω_2	ω_3
state prob. π_ω	0.2	0.3	0.5
cons. plan 1, x_1	3	2	4
cons. plan 2, x_2	3	1	5
cons. plan 3, x_3	4	4	1
cons. plan 4, x_4	1	1	4

Table 2.1: The possible state-contingent consumption plans in the example.

cons. level z	1	2	3	4	5
cons. plan 1, p_1	0	0.3	0.2	0.5	0
cons. plan 2, p_2	0.3	0	0.2	0	0.5
cons. plan 3, p_3	0.5	0	0	0.5	0
cons. plan 4, p_4	0.5	0	0	0.5	0

Table 2.2: The probability distributions corresponding to the state-contingent consumption plans shown in Table 2.1. .

A preference relation can always be represented by a utility function (also if **P** is infinite).

A utility function is unique up to a strictly positive transformation.

- A **von Neumann-Morgenstern utility function** is a function $u : Z \to \mathbb{R}$ such that

$$p \succeq q \Leftrightarrow \underbrace{\sum_{z \in Z} p(z)u(z)}_{\mathrm{E}[u(\tilde{z}_p)]} \geq \underbrace{\sum_{z \in Z} q(z)u(z)}_{\mathrm{E}[u(\tilde{z}_q)]} .$$

Given a von Neumann-Morgenstern utility function u, a utility function $\mathcal{U}$ is defined by $\mathcal{U}(p) = \mathrm{E}[u(\tilde{z}_p)]$.

Example 2.1 Consider an economy with three possible states and four possible state-contingent consumption plans as illustrated in Table 2.1. Note that $Z = \{1, 2, 3, 4, 5\}$. The probability distributions corresponding to these consumption plans are then as shown in Table 2.2. Note that p_3 and p_4 are indistinguishable in this representation. □

2.1.2 Expected utility of consumption plans

Expected utility of consumption plan x corresponding to a probability distribution p:

$$\begin{aligned}\mathrm{E}[u(x)] &= \sum_{\omega\in\Omega} \pi_\omega u(x_\omega) \\ &= \sum_{z\in Z}\sum_{\omega:x_\omega=z} \pi_\omega u(x_\omega) \\ &= \sum_{z\in Z} u(z) \sum_{\omega\,:\,x_\omega=z} \pi_\omega \\ &= \sum_{z\in Z} u(z)p(z).\end{aligned}$$

2.1.3 Behavioral axioms and derived properties

Axiom 2.1 $\succeq$ *is a preference relation.*

Axiom 2.2 (Substitution/independence) *For all $p, q, r \in \mathbf{P}$ and all $a \in (0, 1]$:*

$$p \succ q \Rightarrow ap + (1-a)r \succ aq + (1-a)r.$$

Axiom 2.3 (Archimedean) *For all $p, q, r \in \mathbf{P}$ with $p \succ q \succ r$ there exist constants $a, b \in (0, 1)$ such that*

$$ap + (1-a)r \succ q \succ bp + (1-b)r.$$

From these basic axioms the following properties can be derived:

Theorem 2.1 *(1) $p \succ q$ and $0 \le a < b \le 1 \Rightarrow bp + (1-b)q \succ ap + (1-a)q$.*

(2) $p \succeq q \succeq r$ and $p \succ r \Rightarrow \exists! a^* \in [0, 1] : q \sim a^*p + (1-a^*)r$.

(3) $p \succ q$, $r \succ s$, and $a \in [0, 1] \Rightarrow ap + (1-a)r \succ aq + (1-a)s$.

(4) $p \sim q$ and $a \in [0, 1] \Rightarrow p \sim ap + (1-a)q$.

(5) $p \sim q$ and $a \in [0, 1] \Rightarrow ap + (1-a)r \sim aq + (1-a)r$ for all $r \in \mathbf{P}$.

(6) $\exists z^0, z_0 \in Z$ such that $P_{z^0} \succeq p \succeq P_{z_0}$ for all $p \in \mathbf{P}$.

2.1.4 Expected utility representation of preferences

Theorem 2.2 *Assume Z finite. A binary relation $\succeq$ has an expected utility representation if and only if $\succeq$ satisfies Axioms 1-3.*

Proof: First we prove the implication '$\Leftarrow$': Let z^0, z_0 be as in Property 6. If $P_{z^0} \sim P_{z_0}$ then $p \sim q$ for all $p, q \in \mathbf{P}$ and consequently any $u(z) = k, \forall z \in Z$, will do.

Assume now $P_{z^0} \succ P_{z_0}$. By Property 2, there exists for any p a unique $a_p \in [0, 1]$ such that $a_p P_{z^0} + (1-a_p)P_{z_0} \sim p$. Define the function $U : X \to \mathbb{R}$ by $U(p) = a_p$. Then

$$U(p) \ge U(q) \Leftrightarrow U(p)P_{z^0} + (1-U(p))P_{z_0} \succeq U(q)P_{z^0} + (1-U(q))P_{z_0} \Leftrightarrow p \succeq q.$$

U is linear:

$$\begin{aligned} ap + (1-a)q &\sim a\left[U(p)P_{z^0} + (1-U(p))P_{z_0}\right] + (1-a)\left[U(q)P_{z^0} + (1-U(q))P_{z_0}\right] \\ &\sim \left[aU(p) + (1-a)U(q)\right]P_{z^0} + \left[a(1-U(p)) + (1-a)(1-U(q))\right]P_{z_0}, \end{aligned}$$

hence $U(ap + (1-a)q) = aU(p) + (1-a)U(q)$. Define $u : Z \to \mathbb{R}$ by $u(z) = U(P_z)$. Then

$$U(p) = U\left(\sum_{z \in Z} p(z)P_z\right) = \sum_{z \in Z} p(z)U(P_z) = \sum_{z \in Z} p(z)u(z).$$

Next we prove the implication '⇒': Axiom 1 is simple. Axiom 2: if $\sum_{z \in Z} p(z)u(z) > \sum_{z \in Z} q(z)u(z)$ then

$$\sum_{z \in Z} \left(ap(z) + (1-a)r(z)\right)u(z) > \sum_{z \in Z} \left(aq(z) + (1-a)r(z)\right)u(z).$$

Axiom 3: Suppose $\sum_z p(z)u(z) > \sum_z q(z)u(z) > \sum_z r(z)u(z)$. Define

$$a = 1 - \frac{1}{2}\frac{\sum_z p(z)u(z) - \sum_z q(z)u(z)}{\sum_z p(z)u(z) - \sum_z r(z)u(z)} \in (0,1).$$

Then

$$\begin{aligned} \sum_z \left(ap(z) + (1-a)r(z)\right)u(z) &= \sum_z p(z)u(z) + (1-a)\left(\sum_z r(z)u(z) - \sum_z p(z)u(z)\right) \\ &= \sum_z p(z)u(z) - \frac{1}{2}\left(\sum_z p(z)u(z) - \sum_z q(z)u(z)\right) \\ &= \frac{1}{2}\left(\sum_z p(z)u(z) + \sum_z q(z)u(z)\right) \\ &> \sum_z q(z)u(z). \end{aligned}$$

Similarly for b. □

Theorem 2.3 *A von Neumann-Morgenstern utility function for a given preference relation is only determined up to a strictly positive affine transformation, i.e. if u is von Neumann-Morgenstern utility function for $\succeq$, then v will be so if and only if there exist constants $a > 0$ and b such that $v(z) = au(z) + b$ for all $z \in Z$.*

Suppose $\mathcal{U}$ is a utility function with an associated von Neumann-Morgenstern utility function u. If f is any strictly increasing transformation, then $V = f \circ \mathcal{U}$ is also a utility function for the same preferences, but $f \circ u$ is only the von Neumann-Morgenstern utility function for V if f is affine.

Infinite Z. What if Z is infinite, e.g. $Z = \mathbb{R}_+ \equiv [0, \infty)$? It can be shown that...
A preference relation $\succeq$ satisfies Axioms 1-3 + an Axiom 4 + "some technical conditions" ⇔ $\succeq$ has an expected utility representation.

Expected utility in this case: $\mathrm{E}[u(x)] = \int_Z u(z)p(z)\,dz$, where p is a probability density function derived from the consumption plan x.

Boundedness of expected utility. Suppose u is unbounded from above and $\mathbb{R}_+ \subseteq Z$. Then there exists $(z_n)_{n=1}^{\infty} \subseteq Z$ with $z_n \to \infty$ and $u(z_n) \geq 2^n$. Expected utility of consumption plan p with $p(z_n) = 1/2^n$:

$$\sum_{n=1}^{\infty} u(z_n)p(z_n) \geq \sum_{n=1}^{\infty} 2^n \frac{1}{2^n} = \infty.$$

If q, r are such that $p \succ q \succ r$, then the expected utility of q and r must be finite. But for no $b \in (0,1)$ do we have

$$q \succ bp + (1-b)r \qquad \text{[expected utility} = \infty\text{]}.$$

- no problem if Z is finite
- no problem if $\mathbb{R}_+ \subseteq Z$, u is concave, and consumption plans have finite expectations:

  u concave $\Rightarrow$ u is differentiable in some point b and

  $$u(z) \leq u(b) + u'(b)(z-b), \quad \forall z \in Z.$$

  If $x \in X$ has finite expectations, then

  $$\mathrm{E}[u(x)] \leq \mathrm{E}[u(b) + u'(b)(x-b)] = u(b) + u'(b)\,(\mathrm{E}[x] - b) < \infty.$$

2.1.5 Are the axioms reasonable?

Let us consider an example illustrating the so-called Allais Paradox. Suppose $Z = \{0, 1, 5\}$. Consider consumption plans

$$p_1 = (0,1,0) \qquad p_2 = (0.01, 0.89, 0.1) \qquad p_3 = (0.9, 0, 0.1) \qquad p_4 = (0.89, 0.11, 0).$$

Theory says: $p_1 \succ p_2 \Rightarrow p_4 \succ p_3$. Proof:

$$0.11(\$1) + 0.89\boxed{(\$\ 1)} \sim p_1 \succ p_2 \sim 0.11\left(\frac{1}{11}(\$0) + \frac{10}{11}(\$5)\right) + 0.89\boxed{(\$1)} \quad \Rightarrow$$

$$\underbrace{0.11(\$1) + 0.89\boxed{(\$\ 0)}}_{p_4 \sim} \succ 0.11\left(\frac{1}{11}(\$0) + \frac{10}{11}(\$5)\right) + 0.89\boxed{(\$0)} \sim \underbrace{0.9(\$0) + 0.1(\$5)}_{p_3 \sim}$$

But: people preferring p_1 to p_2 often choose p_3 over p_4. People tend to over-weight small probability events, e.g. ($0) in p_2.

Other "problems":

- the "framing" of possible choices seem to affect decisions
- models assume individuals have unlimited rationality

But... Still useful! No clear alternatives!

2.2 Risk aversion

2.2.1 Definitions

An individual is said to be **risk averse** if he prefers the expected value of a lottery to the lottery, i.e. for any consumption plan x with associated probability function p, $P_{\mathrm{E}[x]} \succeq p$. A risk averse individual will decline any mean-zero gamble.

Theorem 2.4 *An individual with a von Neumann-Morgenstern utility function u is risk averse $\Leftrightarrow$ u is concave.*

We will focus on greedy and risk averse investors so that the von Neumann-Morgenstern utility functions u we shall apply have $u' > 0$ and $u'' < 0$.

Consider a gamble (h_1, h_2) with probabilities $(\bar{p}, 1-\bar{p})$ such that $\bar{p}h_1 + (1-\bar{p})h_2 = 0$. For a risk averse individual we have

$$u\left(\bar{p}(z+h_1) + (1-\bar{p})(z+h_2)\right) = u(z) \geq \bar{p}u(z+h_1) + (1-\bar{p})u(z+h_2),$$

i.e. u is concave. And conversely.

Similarly: Risk lover, risk neutral...

The **certainty equivalent** of a consumption plan p is defined as the $z^* \in Z$ such that

$$u(z^*) = \sum_{z \in Z} u(z)p(z).$$

For $Z \subseteq \mathbb{R}$, z^* uniquely exists if u is continuous and strictly increasing.

For a risk averse individual we have $z^* < \sum_{z \in Z} zp(z)$. The risk compensation that a risk averse individual demands in order to participate in the lottery z is equal to $\sum_{z \in Z} zp(z) - z^*$.

The degree of risk aversion is associated with u'', but good measure should be invariant to strictly positive, affine transformations. This is satisfied by the Arrow-Pratt measures of risk aversion defined as follows. The **Absolute Risk Aversion** is given by $\mathrm{ARA}(z) = -\frac{u''(z)}{u'(z)}$. The **Relative Risk Aversion** is given by $\mathrm{RRA}(z) = -\frac{zu''(z)}{u'(z)} = z\,\mathrm{ARA}(z)$.

Consider a mean-zero gamble h around z. Then

$$E[u(z+h)] = u(z^*) = u\left(z - \pi(z,h)\right).$$

We can approximate the left-hand side by

$$E[u(z+h)] \approx \mathrm{E}\left[u(z) + hu'(z) + \frac{1}{2}h^2u''(z)\right] = u(z) + \frac{1}{2}\,\mathrm{Var}[h]u''(z)$$

and the right-hand side by

$$u\left(z - \pi(z,h)\right) \approx u(z) - \pi(z,h)u'(z).$$

Hence we can write the risk compensation as

$$\pi(z,h) \approx -\frac{1}{2}\,\mathrm{Var}[h]\frac{u''(z)}{u'(z)} = \frac{1}{2}\,\mathrm{Var}[h]\,\mathrm{ARA}(z).$$

Of course, the approximation is more accurate for "small" gambles.

We see that the absolute risk aversion $\mathrm{ARA}(z)$ is constant if and only if $\pi(z,h)$ is independent of z.

Probably $\pi(z,h)$ is decreasing in z, which implies decreasing absolute risk aversion (DARA utility). Note that

$$\mathrm{ARA}'(z) = -\frac{u'''(z)u'(z) - u''(z)^2}{u'(z)^2} = \left(\frac{u''(z)}{u'(z)}\right)^2 - \frac{u'''(z)}{u'(z)} < 0 \quad \Rightarrow \quad u'''(z) > 0,$$

that is, a positive third-order derivative of u is necessary for the utility u to exhibit decreasing absolute risk aversion.

The Arrow-Pratt risk aversion measures are not changed by increasing, affine transformations of U.

Loosely speaking, the absolute risk aversion $\mathrm{ARA}_U(W)$ measures the aversion to a fair gamble of a given dollar amount, such as a gamble where there is an equal probability of winning or loosing 1000 dollars. Since we expect that a wealthy investor will be less averse to that gamble than a poor investor, the absolute risk aversion is expected to be a decreasing function of wealth. The relative risk aversion $\mathrm{RRA}_U(W)$ measures the aversion to a fair gamble of a given percentage of wealth, such as a gamble where there is an equal probability of winning or loosing $0.05W$. Note that utility functions with constant or decreasing (or even modestly increasing) relative risk aversion will display decreasing absolute risk aversion.

2.2.2 Comparison of risk aversion between individuals

An individual with von Neumann-Morgenstern utility function u is said to be more risk averse than an individual with von Neumann-Morgenstern utility function v, if for any lottery $p \in \mathbf{P}$ and $\bar{z} \in Z$ with $\sum_{z \in Z} p(z)u(z) \geq u(\bar{z})$, we have $\sum_{z \in Z} p(z)v(z) \geq v(\bar{z})$.

Theorem 2.5 *u is more risk averse than v $\Leftrightarrow$ $\mathrm{ARA}^u(z) \geq \mathrm{ARA}^v(z)$, $\forall z \in Z$ $\Leftrightarrow$ there exists f strictly increasing and concave such that $u = f \circ v$.*

2.3 Frequently applied utility functions

Let us look at some concrete von Neumann-Morgenstern utility functions that are frequently applied:

CRRA utility. (Also known as power utility, isoelastic utility.) Utility functions $u(W)$ in this class are defined for $W \geq 0$:

$$u(W) = \frac{W^{1-\gamma}}{1-\gamma}, \tag{2.1}$$

where $\gamma > 0$. Note that

$$\mathrm{ARA}_u(W) = \frac{\gamma}{W}, \qquad \mathrm{RRA}_u(W) = \gamma.$$

The relative risk aversion is constant across wealth levels, hence the name CRRA (Constant Relative Risk Aversion) utility. Furthermore, $u'(0+) \equiv \lim_{W\to 0} u'(W) = \infty$ and $u'(\infty) \equiv \lim_{W\to\infty} u'(W) = 0$. Some authors assume a utility function of the form $u(W) = W^{1-\gamma}$, which only makes sense for $\gamma \in (0,1)$. However, empirical studies indicate that most investors have a relative risk aversion above 1.

Except for a constant, the utility function

$$u(W) = \frac{W^{1-\gamma} - 1}{1-\gamma}$$

is equal to the utility function specified in (2.1). The two utility functions are therefore equivalent in the sense that they generate the same optimal choices. The advantage in using the latter definition is that this function has a well-defined limit as $\gamma \to 1$. From l'Hôspital's rule we have that

$$\lim_{\gamma\to 1} \frac{W^{1-\gamma} - 1}{1-\gamma} = \lim_{\gamma\to 1} \frac{-W^{1-\gamma}\ln W}{-1} = \ln W,$$

which is the important special case of **logarithmic utility**. When we consider CRRA utility, we will assume the simpler version (2.1), but we will use the fact that we can obtain the optimal strategies of a log-utility investor as the limit of the optimal strategies of the general CRRA investor as $\gamma \to 1$.

HARA utility. (Also known as extended power utility.) Utility functions in this class are of the form

$$u(W) = \frac{\gamma}{1-\gamma}\left(\frac{aW}{\gamma} + \eta\right)^{1-\gamma} \tag{2.2}$$

with

$$\gamma \neq 0, \quad a > 0, \quad \frac{aW}{\gamma} + \eta > 0, \text{ and } \eta = 1 \text{ if } \gamma = \infty.$$

In this case

$$\mathrm{ARA}_u(W) = \frac{1}{\frac{W}{\gamma} + \frac{\eta}{a}}, \quad \mathrm{RRA}_u(W) = \frac{W}{\frac{W}{\gamma} + \frac{\eta}{a}}.$$

The absolute risk aversion is a hyperbolic function of W, hence the name HARA (Hyperbolic Absolute Risk Aversion) utility. Clearly, when $\eta = 0$ and $\gamma > 0$ we are back to CRRA utility. Therefore, CRRA utility functions belong to the HARA utility function class.

Applying the fact that increasing affine transformations do not change decisions, HARA utility functions can be divided into three different subclasses:

- $u(W) = \frac{(W-\bar{W})^{1-\gamma}}{1-\gamma}$ with $\gamma > 0$. The limit as $\gamma \to 1$ of the equivalent utility function $\frac{(W-\bar{W})^{1-\gamma}-1}{1-\gamma}$ is equal to the extended log utility function $u(W) = \ln(W - \bar{W})$. Utility is defined for $W \geq \bar{W}$ and $u'(\bar{W}) = \infty$, hence $\bar{W}$ is often referred to as a subsistence level of wealth/consumption. This makes sense only if $\bar{W} \geq 0$. We will refer to this subclass as **subsistence HARA utility** functions. For $\bar{W} = 0$ we recover the CRRA utility.

- $u(W) = -(\bar{W} - W)^{1-\gamma}$ with $\gamma < 0$. Utility is defined for $W \leq \bar{W}$, so that we can think of $\bar{W}$ as a satiation level. We could call this subclass **satiation HARA utility** functions.

- $u(W) = -e^{-aW}$ corresponding to the limit of (2.2) as $\gamma \to \infty$ and $\eta = 1$. This is the **(negative) exponential utility** which displays constant absolute risk aversion (CARA) – not very reasonable!

Since it is hard to imagine negative consumption, most attention has been given to the CRRA utility functions and (to a smaller extent) the non-CRRA subsistence HARA utility functions. There is also a technical advantage to subsistence HARA and to CRRA functions: Since $u'(\bar{W}+) = \infty$, an optimal solution will have the property that consumption/wealth W will be strictly above $\bar{W}$ with probability one. For example, with CRRA utility, we can ignore a non-negativity constraint on consumption since the constraint will never be binding. For computational purposes the negative exponential utility function is often used in connection with normally distributed returns, e.g. in one-period models as discussed below.

It is an empirical fact that even though consumption and wealth have increased tremendously over the years, the magnitude of real rates of return seems not to have changed significantly. This is consistent with agents having ("on average") close to CRRA utility, but not consistent with neither significantly increasing nor decreasing RRA utility. In addition, CRRA utility turns out to be very tractable. Consequently, it constitutes the most studied class of utility functions.

2.4 Preferences in multi-period settings

Above we implicitly considered preferences for consumption at one given future point in time. However, we can generalize the ideas and results to multi-period settings. Consider first a discrete time set $\{0, 1, 2, \ldots, T\}$. Then the appropriate consumption space is $Z = \{(z_0, z_1, z_2, \ldots, z_T)\}$ and consumption plans are represented by a probability p on Z.

Again, for finite Z, we have that for a preference relation $\succeq$ satisfying the Axioms 2.1–2.3 there exists a function $U : Z \to \mathbb{R}$ such that

$$p \succeq q \Leftrightarrow \sum_{z \in Z} U(z_0, z_1, \ldots, z_T) p\left(z = (z_0, z_1, \ldots, z_T)\right) \geq \sum_{z \in Z} U(z_0, z_1, \ldots, z_T) q\left(z = (z_0, z_1, \ldots, z_T)\right),$$

i.e. the consumption plans are ordered by expected utility. We can call U a multi-period von Neumann-Morgenstern utility function. Note that it depends on consumption at all dates. Again this result can be extended to the case of an infinite Z, e.g. $Z = \mathbb{R}_+^{T+1}$, but also to continuous-time settings where U will then be a function of the entire consumption process $c = (c_t)_{t \in [0,T]}$.

In multi-period settings it is important to know to which degree the investor is willing to shift consumption from one point in time to another. A measure of this is given by the **intertemporal elasticity of substitution**. The intertemporal elasticity of substitution between consumption at time t and time s (with $s > t$) is defined as

$$\mathrm{IES}_U(c_t, c_s) = -\frac{\frac{\partial U}{\partial c_s} \big/ \frac{\partial U}{\partial c_t}}{c_s / c_t} \frac{d\left(c_s / c_t\right)}{d\left(\frac{\partial U}{\partial c_s} \big/ \frac{\partial U}{\partial c_t}\right)}. \tag{2.3}$$

Often **time-additivity** is assumed so that the utility the agent gets from consumption in one period does not directly depend on what she consumed in earlier periods or what she plan to consume in later periods. For the discrete-time case, this means that

$$U(c_0, c_1, \ldots, c_T) = \sum_{t=0}^{T} u_t(c_t)$$

where each u_t qualifies as a von Neumann-Morgenstern utility function in a one-period setting. Still, when the agent has to choose her current consumption rate, she will take her prospects for future consumption into account. As we shall see, she will in fact try to smooth her consumption rates across time. We will sometimes allow for a utility from leaving wealth for bequest. Letting W_T denote the wealth level after consumption at time T and $\bar{u}$ the bequest utility function, the life-time utility is then given by

$$U(c_0, c_1, \ldots, c_T, W_T) = \sum_{t=0}^{T} u_t(c_t) + \bar{u}(W_T)$$

The continuous time analogue is

$$U(c, W_T) = \int_0^T u_t(c_t)\, dt + \bar{u}(W_T).$$

In addition we will typically assume that $u_t(c_t) = e^{-\delta t} u(c_t)$ for all t. This is to say that the direct utility the agent gets from a given consumption level is basically the same for all dates, but the agent prefers to consume any given number of goods sooner than later. This is modeled by the subjective time preference rate δ, which we assume to be constant over time and independent of the consumption level. For a unified notation we replace $\bar{u}(W_T)$ by $e^{-\delta T}\bar{u}(W_T)$. In sum, the life-time utility is typically assumed to be given by

$$U(c_0, c_1, \ldots, c_T, W_T) = \sum_{t=0}^{T} e^{-\delta t} u(c_t) + e^{-\delta T}\bar{u}(W_T)$$

in discrete-time models and

$$U(c, W_T) = \int_0^T e^{-\delta t} u(c_t)\, dt + e^{-\delta T}\bar{u}(W_T)$$

in continuous-time models.

These assumptions on preferences will simplify many computations and facilitate analytical solutions to optimal investment and consumption problems. However, it is important to realize that the time-additive specification does not follow from the basic axioms of choice under uncertainty, but is in fact a strong assumption, which most economists agree is not very realistic. For a time-additive utility specification the intertemporal elasticity of substitution becomes

$$\mathrm{IES}_u(c_t, c_s) = -\frac{u'(c_s)/u'(c_t}{c_s/c_t} \frac{d\,(c_s/c_t)}{d\,(u'(c_s)/u'(c_t))}. \tag{2.4}$$

When we let $s \to t$, we obtain the instantaneous elasticity of substitution at time t:

$$\mathrm{IES}_u(c_t) = -\frac{u'(c_t)}{u''(c_t)c_t}. \tag{2.5}$$

Note that this is equal to the inverse of the relative risk aversion, i.e.

$$\mathrm{IES}_u(c_t) = \frac{1}{\mathrm{RRA}_u(c_t)}. \tag{2.6}$$

This result is due to the assumption of time-additive utility. For the special case of power utility, we have $\mathrm{RRA}_u(c) = \gamma$ and hence $\mathrm{IES}_u(c) = 1/\gamma$.

The close link between IES and RRA is restrictive. IES and RRA measure two different aspects of preferences. The IES measures the willingness of the individual to substitute consumption over time, whereas the RRA measures the reluctance to substitute consumption across different states of the economy. There is nothing in the theory of choice under uncertainty that links these two concepts together. It is an unfortunate consequence of the assumption of time-additive utility.

According to Browning (1991), non-additive preferences were already discussed in the 1890 book "Principles of Economics" by Alfred Marshall. See Browning's paper for further references to the critique on intertemporally separable preferences. A relatively simple and therefore also quite tractable example of non-additive preferences is obtained by letting the utility associated with the choice of consumption at a given date may depend on past choices of consumption. This is modeled by replacing $u(c_t)$ by $u(c_t, h_t)$, where u is decreasing in h_t, which is a measure of the standard of living or the habit level of consumption, e.g. a weighted average of past consumption rates:

$$h_t = h_0 e^{-\beta t} + \alpha \int_0^t e^{-\beta(t-s)} c_s\, ds,$$

where h_0, α, and β are non-negative constants. High past consumption generates a desire for high current consumption, so that preferences display intertemporal complementarity. This is referred to as preferences with **habit formation**. In particular, models where $u(c_t, h_t)$ is assumed to be of the power-linear form,

$$u(c, h) = \frac{1}{1-\gamma}(c-h)^{1-\gamma}, \quad \gamma > 0, c \geq h,$$

turn out to be computationally tractable. See Section 9.1 on portfolio and consumption choice for investors with power-linear habit formation preferences.

Another non-additive model of preference is given by the so-called recursive preferences, suggested and discussed by, e.g., Kreps and Porteus (1978), Epstein and Zin (1989, 1991), and Weil (1989). The original motivation of this representation of preferences is that it allows individuals to have preferences for the timing of resolution of uncertainty, which is not the case with standard multi-period von Neumann-Morgenstern preferences. With recursive preferences the felicity U_t at some point in time t depends both on consumption and that date and expectations of next period's felicity. The most tractable, non-trivial specification is

$$U_t = \left[(1-\delta) c_t^{(1-\gamma)/\theta} + \delta \left(\mathrm{E}_t \left[U_{t+1}^{1-\gamma} \right] \right)^{1/\theta} \right]^{\theta/(1-\gamma)}, \quad \theta \equiv \frac{1-\gamma}{1-\frac{1}{\psi}}. \tag{2.7}$$

Here γ has the interpretation of the relative risk aversion and ψ the interpretation of the intertemporal elasticity of substitution. When $\gamma = 1/\psi$, we have $\theta = 1$, and it can be shown that the recursive equation above is satisfied by the standard time-additive power utility. The continuous-time equivalent of recursive utility is called stochastic differential utility and studied by, e.g., Duffie and Epstein (1992).

Despite the added realism of the non-standard preferences most researchers do assume time-additive preferences mainly due to the computational advantages they offer. We shall also do that in most of this lecture note. Some portfolio and consumption choice problems with non-standard preferences are discussed in Chapter 9.

Chapter 3

One-period models

3.1 The general one-period model

Given d risky assets with (stochastic) rates of return $R = (R_1, \ldots, R_d)^\top$ and a riskfree asset with a (certain) rate of return r over the period of interest. An investor with initial wealth W_0 who invests amounts $\theta = (\theta_1, \ldots, \theta_d)^\top$ in the risky assets and the remainder $\theta_0 = W_0 - \theta^\top \mathbf{1}$ in the riskfree asset will end up with wealth

$$W = W_0 + \theta^\top R + \theta_0 r = (1+r)w_0 + \theta^\top (R - r\mathbf{1})$$

at the end of the period. Letting $\pi_i = \theta_i / W_0$ denote the fraction of wealth invested in the i'th asset, we can rewrite the terminal wealth as

$$W = W_0 \left[1 + r + \pi^\top (R - r\mathbf{1})\right],$$

where $\pi = (\pi_1, \ldots, \pi_d)^\top$. The one-period utility-maximization problem is to choose π to maximize $\mathrm{E}[u(W)]$. Note that we ignore any consumption decision at the beginning of the planning period, i.e. we assume that the consumption decision has already been taken independently of the investment decision.

Without further assumptions one can show a number of interesting results on the optimal portfolio choice. We will state only a few and refer to Merton (1992, Ch. 2) for further properties of the general solution to this utility maximization problem.

Theorem 3.1 *An individual with strictly increasing and concave u will avoid any positive risky investment only if $\mathrm{E}[R_j] \le r$ for all j.*

Theorem 3.2 *Assume a single risky asset. The optimal risky investment $\theta = \theta(W_0)$ has the following properties:*

(i) if $\mathrm{ARA}(\cdot)$ is uniformly decreasing/increasing/constant, then θ is increasing/decreasing/constant in W_0

(ii) if $\mathrm{RRA}(\cdot)$ is uniformly decreasing/increasing/constant, then θ/W_0 is increasing/decreasing/constant in W_0.

3.2 Mean-variance analysis

Mean-variance analysis was introduced by Markowitz (1952, 1959). Our presentation is inspired by Huang and Litzenberger (1988, Ch. 3). Mean-variance analysis assumes that the portfolio choice of investors will depend only on the mean and variance of their end-of-period wealth and hence on the mean and variances of the portfolios investors can form. Before we go into the derivations of optimal portfolios, let us discuss the theoretical foundation of mean-variance analysis.

3.2.1 Theoretical foundation

In general an individual's utility of wealth will depend on all moments of wealth. This can be seen by the Taylor expansion of $u(W)$ around the expected wealth, $\mathrm{E}[W]$:

$$u(W) = u(\mathrm{E}[W]) + u'(\mathrm{E}[W])(W - \mathrm{E}[W]) + \frac{1}{2}u''(\mathrm{E}[W])(W - \mathrm{E})^2 + \sum_{n=3}^{\infty} \frac{1}{n!} u^{(n)}(\mathrm{E}[W])(W - \mathrm{E}[W])^n,$$

where $u^{(n)}$ is the n'th derivative of u. Taking expectations, we get

$$\mathrm{E}[u(W)] = u(\mathrm{E}[W]) + \frac{1}{2}u''(\mathrm{E}[W])\,\mathrm{Var}(W) + \sum_{n=3}^{\infty} \frac{1}{n!} u^{(n)}(\mathrm{E}[W])\,\mathrm{E}\left[(W - \mathrm{E}[W])^n\right].$$

A greedy and risk averse investor clearly prefers higher expected wealth and lower variance of wealth, other things equal, but the expected utility is also influenced by higher order moments. Of course, with quadratic utility, the derivatives of u of order 3 and higher are zero, so the higher order moments of wealth are irrelevant. However, quadratic utility is a very unrealistic model of investor preferences.

Mean-variance analysis is valid if the returns on the risky assets are multivariate normally distributed, $R \sim N(\mu, \Sigma)$, then the well-known mean-variance analysis applies. Here, μ is a vector of the expected rates of return on the risky assets, and Σ is the variance-covariance matrix of these rates of return, so that Σ_{ij} denotes the covariance between the returns on asset i and asset j. Given that the returns on all individual assets are normally distributed, the return on any portfolio – being a weighted average of the returns on the assets in the portfolio – will also be normally distributed. A portfolio characterized by the portfolio weights π has a return of $R^\pi \equiv \pi^\top R = \sum_{i=1}^{d} \pi_i R_i$, which is normally distributed with mean and variance given by

$$\mu(\pi) \equiv \mathrm{E}[R^\pi] = \pi^\top \mu = \sum_{i=1}^{d} \pi_i \mu_i, \tag{3.1}$$

$$\sigma^2(\pi) \equiv \mathrm{Var}[R^\pi] = \pi^\top \Sigma \pi = \sum_{i=1}^{d} \sum_{j=1}^{d} \pi_i \pi_j \Sigma_{ij}. \tag{3.2}$$

Consequently, the end-of-period wealth of each investor will also be normally distributed for any portfolio choice. All higher-order moments of wealth can be written in terms of mean and variance so that expected utility depends only on expected wealth and the variance of wealth.

An obvious short-coming of the assumption of normally distributed returns is the possibility of rates of returns smaller than -100%, which is inconsistent with limited liability of securities. It also allows for negative end-of-period wealth and hence negative consumption with positive probability, which is clearly unreasonable. A promising alternative is to assume that the end-of-period prices of individual assets are lognormally distributed, ruling out negative prices and rates of return below 100%. The lognormal distribution is also fully described by its first two moments. Unfortunately, such an assumption is not tractable in a one-period setting since neither the value nor the return on a portfolio will then not be lognormally distributed (the lognormal distribution is not stable under addition).

3.2.2 Mean-variance analysis with only risky assets

Assume that the variance-covariance matrix Σ is non-singular, which is the case if none of the assets are redundant, i.e. no asset has a return which is linear combination of the returns of other assets. The inverse of Σ is denoted by Σ^{-1}. A portfolio is said to be **mean-variance efficient** if it has the minimum return variance among all the portfolios with the same mean return. Given the normality assumption on returns, greedy and risk averse investors will only choose among the mean-variance portfolios. Assuming that there are no portfolio constraints, we can find a mean-variance portfolio with expected return $\bar{\mu}$ by solving the quadratic minimization problem

$$\begin{aligned} &\min_{\pi} \frac{1}{2} \pi^\top \Sigma \pi \\ &\text{s.t. } \pi^\top \mu = \bar{\mu}, \\ &\qquad \pi^\top \mathbf{1} = 1. \end{aligned} \tag{3.3}$$

The '$\frac{1}{2}$' in the objective will be notationally convenient when we solve the problem. Clearly, the portfolio that minimizes half the variance will also minimize the variance.

We solve the problem by the Lagrange technique. Letting α and β denote the Lagrange multipliers of the two constraints, the Lagrangian is

$$\mathcal{L} = \frac{1}{2} \pi^\top \Sigma \pi + \alpha \left(\bar{\mu} - \pi^\top \mu \right) + \beta \left(1 - \pi^\top \mathbf{1} \right).$$

The first-order condition with respect to π is

$$\frac{\partial \mathcal{L}}{\partial \pi} = \Sigma \pi - \alpha \mu - \beta \mathbf{1} = 0,$$

which implies that

$$\pi = \alpha \Sigma^{-1} \mu + \beta \Sigma^{-1} \mathbf{1}. \tag{3.4}$$

The first-order conditions with respect to the multipliers simply give the two constraints to the minimization problem. Substituting the expression (3.4) for π into the two constraints, we obtain the equations

$$\begin{aligned} \alpha \mu^\top \Sigma^{-1} \mu + \beta \mathbf{1}^\top \Sigma^{-1} \mu &= \bar{\mu}, \\ \alpha \mu^\top \Sigma^{-1} \mathbf{1} + \beta \mathbf{1}^\top \Sigma^{-1} \mathbf{1} &= 1. \end{aligned}$$

Defining

$$A = \mu^\top \Sigma^{-1} \mu, \qquad B = \mu^\top \Sigma^{-1} \mathbf{1} = \mathbf{1}^\top \Sigma^{-1} \mu, \qquad C = \mathbf{1}^\top \Sigma^{-1} \mathbf{1},$$

we can write the solution to these two equations in α and β as

$$\alpha = \frac{C \bar{\mu} - B}{AC - B^2}, \qquad \beta = \frac{A - B \bar{\mu}}{AC - B^2}.$$

Substituting this into (3.4) we obtain

$$\pi = \pi(\bar{\mu}) \equiv \frac{C\bar{\mu} - B}{AC - B^2}\Sigma^{-1}\mu + \frac{A - B\bar{\mu}}{AC - B^2}\Sigma^{-1}\mathbf{1}. \tag{3.5}$$

Some tedious calculations show that the variance of the return on this portfolio is equal to

$$\sigma^2(\bar{\mu}) \equiv \pi(\bar{\mu})^\top \Sigma \pi(\bar{\mu}) = \frac{C\bar{\mu}^2 - 2B\bar{\mu} + A}{AC - B^2}.$$

We see that the combinations of variance and mean form a parabola in a (mean, variance)-diagram.

Traditionally the portfolios are depicted in a (standard deviation, mean)-diagram. The above relation can also be written as

$$\frac{\sigma^2(\bar{\mu})}{1/C} - \frac{(\bar{\mu} - B/C)^2}{D/C^2} = 1,$$

from which it follows that the optimal combinations of standard deviation and mean form a hyperbola in the (standard deviation, mean)-diagram. This hyperbola is called the **mean-variance frontier** of risky assets. The mean-variance efficient portfolios are sometimes called frontier portfolios.

Before we proceed let us clarify a point in the derivation above. We have assumed that $AC - B^2$ is non-zero. In fact, $AC - B^2 > 0$. To see this, first recall the following definition. A symmetric $d \times d$ matrix Σ is said to be *positive definite* if $\pi^\top \Sigma \pi > 0$ for any non-zero d-vector π. Since in our case $\pi^\top \Sigma \pi$ equals the variance of the portfolio π and all portfolios of risky assets will have a return with positive variance, the variance-covariance matrix Σ is indeed a positive definite matrix. A result in linear algebra says that the inverse Σ^{-1} is then also positive definite, i.e. $x^\top \Sigma^{-1} x > 0$ for any non-zero d-vector x. In particular we have $A > 0$ and $C > 0$. Also

$$A(AC - B^2) = (B\mu - A\mathbf{1})^\top \Sigma^{-1}(B\mu - A\mathbf{1}) > 0$$

and since $A > 0$ we must have $AC - B^2 > 0$.

The minimum-variance portfolio is the portfolio that has the minimum variance among all portfolios. We can find this directly by solving the constrained minimization problem

$$\begin{aligned} &\min_{\pi} \frac{1}{2}\pi^\top \Sigma \pi \\ &\text{s.t. } \pi^\top \mathbf{1} = 1 \end{aligned} \tag{3.6}$$

where there is no constraint on the mean portfolio return. Alternatively, we can minimize the variance $\sigma^2(\bar{\mu})$ over all $\bar{\mu}$. Taking the latter route, we find that the minimum variance is obtained when the mean return is $\bar{\mu}_{\min} = B/C$ and the minimum variance is given by $\sigma^2_{\min} = 1/C$. From (3.5) we get that the minimum-variance portfolio is

$$\pi_{\min} = \frac{1}{C}\Sigma^{-1}\mathbf{1} = \frac{1}{\mathbf{1}^\top \Sigma^{-1}\mathbf{1}}\Sigma^{-1}\mathbf{1}. \tag{3.7}$$

It can be shown that the portfolio

$$\pi_{\text{slope}} = \frac{1}{B}\Sigma^{-1}\mu = \frac{1}{\mathbf{1}^\top \Sigma^{-1}\mu}\Sigma^{-1}\mu \tag{3.8}$$

is the portfolio that maximizes the slope of a straight line between the origin and a point on the mean-variance frontier. Let us call it the maximum slope portfolio. This portfolio has mean A/B

and variance A/B^2. From (3.5) we see that any mean-variance optimal portfolio can be written as a linear combination of the maximum slope portfolio and the minimum-variance portfolio:

$$\pi(\bar{\mu}) = \frac{(C\bar{\mu} - B)B}{AC - B^2}\pi_{\text{slope}} + \frac{(A - B\bar{\mu})C}{AC - B^2}\pi_{\min}. \tag{3.9}$$

Note that the two multipliers of the portfolios sum to one. This is a **two-fund separation** result. If the investors can only form portfolios of the d risky assets with normally distributed returns, any greedy and risk-averse investor will choose a combination of two special portfolios or funds, namely the maximum slope portfolio and the minimum-variance portfolio. These two portfolios are said to generate the mean-variance frontier of risky assets. In fact, it can be shown that any other two frontier portfolios generate the entire frontier.

Exactly which combination of the two generating portfolios that a particular investor prefers is in general difficult to determine. For the unrealistic case of negative exponential utility (CARA) the optimal combination can be determined in closed form. For other, more reasonable, utility functions numerical optimization is necessary. Intuitively, the weight of the minimum-variance portfolio is increasing in the risk aversion of the investor.

Additional properties; alternative characterizations... Huang and Litzenberger (1988), Hansen and Richard (1987).

3.2.3 Mean-variance analysis with both risky assets and a riskless asset

A riskless asset corresponds to a point $(0, r)$ in the (standard deviation, mean)-diagram. The investors can combine any portfolio of risky assets with an investment in the riskless asset. The (standard deviation, mean)-pairs that can be obtained by such a combination forms a straight line between the point $(0, r)$ and the point corresponding to the portfolio of risky asset. Other things equal, greedy and risk-averse investors want high expected return and low standard deviation so they will move as far to the "north-east" as possible in the diagram. Consequently they will pick a point somewhere on the upward-sloping line that is tangent to the mean-variance frontier of risky assets and goes through the point $(0, r)$. The point where this line is tangent to the frontier of risky assets corresponds to a portfolio which we refer to as the **tangency portfolio**. This is a portfolio of risky assets only. It is the portfolio that maximizes the Sharpe ratio over all risky portfolios. The Sharpe ratio of a portfolio is the ratio $(\mu(\pi) - r)/\sigma(\pi)$ between the excess expected return of a portfolio and the standard deviation of the return. The tangency portfolio is given by the portfolio weights

$$\pi_{\tan} = \frac{1}{\mathbf{1}^{\top}\Sigma^{-1}(\mu - r\mathbf{1})}\Sigma^{-1}(\mu - r\mathbf{1}). \tag{3.10}$$

This upward-sloping straight line constitutes the mean-variance frontier of all assets. Again it is quite cumbersome to compute exactly which of these mean-variance efficient portfolios that a given investor prefers, except for the case of negative exponential utility. Again we have two-fund separation since all investors will combine just two funds, where one fund is simply the riskless asset and the other is the tangency portfolio.

Note that all investors will hold different risky assets in the same proportion to each other, i.e. for any $i, j \in \{1, \dots, d\}$ the ratio π_i/π_j is the same for all investors.

3.3 Critique of the one-period framework

- Investors typically get utility from consumption at different points in time and not simply the wealth level at one particular date.

- Even in the case where the investor only obtains utility from wealth at one date, she has the opportunity to change her portfolio over time, which she would normally do as new information arises (e.g. when stock prices and interest rates change) or simply because time passes. Investors live in a dynamic model and will take decisions dynamically. Of course, the existence of transaction costs is a reason for not changing the portfolio too frequently, but if we are really worried about transaction costs we should explicitly model that imperfection – the analysis of such models is quite difficult, however.

- Consumption and investment decisions are generally not to be separated from each other. Investments are meant to generate future consumption!

- The normality (or similar sufficient distributional) assumption employed in the mean-variance analysis and the CAPM is not reasonable, neither from a theoretical nor an empirical point of view. For example, the normal distribution allocates a strictly positive probability to a return below -100%, which cannot happen for investments in securities with limited liability.

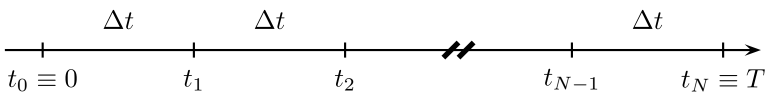
Chapter 4

Introduction to multi-period models

To study dynamic consumption and investment decisions, several papers have looked at multi-period, discrete-time models where the investor has the opportunity to consume and rebalance her portfolio at a number of fixed dates. Certainly this is a valuable extension of the single-period setting, but it is still a limitation that the investor can only change her decisions at pre-specified points in time and not react to new information arriving between these points in time. A continuous-time model seems more reasonable. Furthermore, the results on optimal consumption and investment strategies are typically clearer in continuous-time models than in discrete-time models, and the necessary mathematical computations are much more elegant in a continuous-time framework. Therefore, we will not give much attention to multi-period, discrete-time models. However, some aspects of the set-up of continuous-time models may be easier to understand if we start by looking at a discrete-time model and then take the limit as the period length goes to zero. The basic references for the discrete-time models are Samuelson (1969), Hakansson (1970), Fama (1970, 1976), and Ingersoll (1987, Ch. 11).

4.1 A multi-period, discrete-time framework

Consider the time line below:



At time $t_n = n\Delta t$ for $n = 0, 1, \dots, N-1$ the investor chooses a portfolio θ_{t_n} which is held unchanged until time t_{n+1} and a consumption rate c_{t_n} such that the total consumption in the interval $[t_n, t_{n+1})$ is $c_{t_n} \cdot \Delta t$. (We assume that there is a single consumption good so that c_{t_n} is one-dimensional.) This is subtracted from her wealth at time t_n. Of course, θ_{t_n} and c_{t_n} can only be based on the information known at time t_n, i.e. in mathematical terms they must be $\mathcal{F}_{t_n}$-measurable. We assume that there is no consumption or investment beyond time T, which we can think of as the time of death (assumed to be known in advance!). At time 0 the investor must choose the entire consumption rate process $c_0, c_{t_1}, \dots, c_{t_{N-1}}$ and the entire portfolio process $\theta_0, \theta_{t_1}, \dots, \theta_{t_{N-1}}$. In other words, she must choose the current values c_0 and θ_0 and for each future date t_n (with $n = 1, \dots, N-1$) she must choose a consumption rate $c_{t_n}(\omega)$ and $\theta_{t_n}(\omega)$ for each possible state of the world ω at day t_n.

We assume that the life-time utility of consumption and terminal wealth is given by

$$U(c_0, c_1, \dots, c_T, W_T) = \sum_{t=0}^{T} e^{-\delta t} u(c_t) + e^{-\delta T} \bar{u}(W_T)$$

as discussed in Section 2.4. The maximal obtainable expected life-time utility seen from time 0 is therefore

$$J_0 = \sup_{(c_{t_n}, \theta_{t_n})_{n=0}^{N-1}} \mathrm{E}\left[\sum_{n=0}^{N-1} e^{-\delta t_n} u(c_{t_n})\Delta t + e^{-\delta T} \bar{u}(W_T)\right],$$

where the supremum is taken over all budget-feasible consumption and investment strategies. Similarly, we define

$$J_{t_i} = \sup_{(c_{t_n}, \theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_i}\left[\sum_{n=i}^{N-1} e^{-\delta(t_n - t_i)} u(c_{t_n})\Delta t + e^{-\delta(T-t_i)} \bar{u}(W_T)\right], \quad i = 0, \dots, N-1, \tag{4.1}$$

where the subscript on the expectations operator denotes that the expectation is taken conditional on the information known to the agent at time t_i. J is often called the indirect or derived utility of wealth process or function, since it measures the highest attainable expected life-time utility the investor can derive from her current wealth in the current state of the world.

4.2 Wealth dynamics

Denote by $P_{t_n} = (P^1_{t_n}, \dots, P^d_{t_n})^\top$ the vector of prices of the d risky assets at time t_n. The nominal riskfree rate of return per year in the period $[t_n, t_{n+1})$ is denoted by r_{t_n}. This means that the riskfree rate of return over the period is $r_{t_n}\Delta t$. The value at time t_n of a dollar invested at time 0 and subsequently rolled over at the riskfree rate is denoted by $P^0_{t_n}$. We will refer to such a unit bank account as asset 0. For the purposes of deriving the budget constraint we will represent the portfolio by the number of units of each asset held. We let $N^i_{t_n}$ denote the number of units of asset $i = 0, 1, \dots, d$ held in the period $[t_n, t_{n+1})$. We will allow for the case where the agent earns income from other sources than his financial investments. We let y_{t_n} be the rate of income earned in the period $[t_n, t_{n+1})$ such that the entire income in this period is $y_{t_n} \cdot \Delta t$. We assume that the agent receives this amount at time t_n. Note that we do not model the labor supply decision resulting in this income, but take y_{t_n} as exogenously given.

The agent enters date t_n with a wealth of

$$W_{t_n} = \sum_{i=0}^{d} N^i_{t_{n-1}} P^i_{t_n}.$$

This is the value of her portfolio chosen in the previous period. She then receives income $y_{t_n} \cdot \Delta t$ and simultaneously has to choose the consumption rate c_{t_n} and the new portfolio $(N^0_{t_n}, N^1_{t_n}, \dots, N^d_{t_n})$.

The budget restriction on these choices is that

$$(y_{t_n} - c_{t_n})\,\Delta t = \sum_{i=0}^{d} \left[N^i_{t_n} - N^i_{t_{n-1}}\right] P^i_{t_n},$$

i.e. that income net of consumption equals the extra amount invested in the financial market. We then get that

$$\begin{aligned} W_{t_{n+1}} - W_{t_n} &= \sum_{i=0}^{d} N^i_{t_n} P^i_{t_{n+1}} - \sum_{i=0}^{d} N^i_{t_{n-1}} P^i_{t_n} \\ &= \sum_{i=0}^{d} N^i_{t_n} \left(P^i_{t_{n+1}} - P^i_{t_n} \right) + \sum_{i=0}^{d} \left[N^i_{t_n} - N^i_{t_{n-1}} \right] P^i_{t_n} \\ &= \sum_{i=0}^{d} N^i_{t_n} \left(P^i_{t_{n+1}} - P^i_{t_n} \right) + \left(y_{t_n} - c_{t_n} \right) \Delta t. \end{aligned}$$

Let $\theta^i_{t_n} = N^i_{t_n} P^i_{t_n}$ denote the amount invested in asset i and let $R^i_{t_n} = \left(P^i_{t_{n+1}} - P^i_{t_n} \right) / P^i_{t_n}$ denote the rate of return on asset i. Then the change in wealth can be rewritten as

$$W_{t_{n+1}} - W_{t_n} = \sum_{i=0}^{d} \theta^i_{t_n} R^i_{t_n} + \left(y_{t_n} - c_{t_n} \right) \Delta t.$$

With the vector notation $\theta_{t_n} = (\theta^1_{t_n}, \dots, \theta^d_{t_n})^\top$ and $R_{t_n} = (R^1_{t_n}, \dots, R^d_{t_n})^\top$, we get

$$W_{t_{n+1}} - W_{t_n} = \theta^0_{t_n} r_{t_n} \Delta t + \theta^\top_{t_n} R_{t_n} + \left(y_{t_n} - c_{t_n} \right) \Delta t.$$

Note that the only stochastic variable (seen from time t_n) on the right-hand side is the return vector R_{t_n}. Let us decompose the return into an expected and an unexpected part,

$$R_{t_n} = \mu_{t_n} \Delta t + \sigma_{t_n} \varepsilon_{t_n} \sqrt{\Delta t}. \tag{4.2}$$

Here μ_{t_n} is the vector of expected rates of return per year, ε_{t_n} is a vector of independent stochastic shocks all with mean zero and variance one, and σ_{t_n} is a matrix determining how the returns are affected by these shocks. The values of μ_{t_n} and σ_{t_n} are known at time t_n. The realization of the shock vector ε_{t_n} will be known at time $(n+1)\Delta t$, just before the consumption and portfolio decisions at that date are taken. It follows that, seen at time t_n, the variance-covariance matrix of R_{t_n} is given by $\sigma_{t_n} \sigma^\top_{t_n} \Delta t$. The elements in $\Sigma_{t_n} \equiv \sigma_{t_n} \sigma^\top_{t_n}$ are hence variances and covariances per year. The change in wealth can now be rewritten (yet another time) as

$$W_{t_{n+1}} - W_{t_n} = \left[\theta^0_{t_n} r_{t_n} + \theta^\top_{t_n} \mu_{t_n} + y_{t_n} - c_{t_n} \right] \Delta t + \theta^\top_{t_n} \sigma_{t_n} \varepsilon_{t_n} \sqrt{\Delta t}. \tag{4.3}$$

4.3 Dynamic programming in discrete-time models

In the definition of indirect utility in (4.1) the maximization is over both the current and all future consumption rates and portfolios. This is clearly a quite complicated maximization problem. We will now show that we can alternatively perform a sequence of simpler maximization problems.

This result is based on the following manipulations:

$$\begin{aligned}
J_{t_i} &= \sup_{(c_{t_n},\theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_i}\left[\sum_{n=i}^{N-1} e^{-\delta(t_n-t_i)}u(c_{t_n})\Delta t + e^{-\delta(T-t_i)}\bar{u}(W_T)\right] \\
&= \sup_{(c_{t_n},\theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_i}\left[u(c_{t_i})\Delta t + \sum_{n=i+1}^{N-1} e^{-\delta(t_n-t_i)}u(c_{t_n})\Delta t + e^{-\delta(T-t_i)}\bar{u}(W_T)\right] \\
&= \sup_{(c_{t_n},\theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_i}\left[u(c_{t_i})\Delta t + \mathrm{E}_{t_{i+1}}\left[\sum_{n=i+1}^{N-1} e^{-\delta(t_n-t_i)}u(c_{t_n})\Delta t + e^{-\delta(T-t_i)}\bar{u}(W_T)\right]\right] \\
&= \sup_{(c_{t_n},\theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_i}\left[u(c_{t_i})\Delta t + e^{-\delta\Delta t}\,\mathrm{E}_{t_{i+1}}\left[\sum_{n=i+1}^{N-1} e^{-\delta(t_n-t_{i+1})}u(c_{t_n})\Delta t + e^{-\delta(T-t_{i+1})}\bar{u}(W_T)\right]\right] \\
&= \sup_{c_{t_i},\theta_{t_i}} \mathrm{E}_{t_i}\left[u(c_{t_i})\Delta t + e^{-\delta\Delta t} \sup_{(c_{t_n},\theta_{t_n})_{n=i}^{N-1}} \mathrm{E}_{t_{i+1}}\left[\sum_{n=i+1}^{N-1} e^{-\delta(t_n-t_{i+1})}u(c_{t_n})\Delta t + e^{-\delta(T-t_{i+1})}\bar{u}(W_T)\right]\right]
\end{aligned}$$

Here, the first equality is simply due to the definition of indirect utility, the second equality comes from separating out the first term of the sum, the third equality is valid according to the law of iterated expectations, the fourth equality comes from separating out the discount term $e^{-\delta\Delta t}$, and the final equality is due to the fact the only the inner expectation depends on future consumption rates and portfolios. Noting that the inner supremum is by definition the indirect utility at time t_{i+1}, we arrive at

$$J_{t_i} = \sup_{c_{t_i},\theta_{t_i}} \mathrm{E}_{t_i}\left[u(c_{t_i})\Delta t + e^{-\delta\Delta t}J_{t_{i+1}}\right]. \tag{4.4}$$

This equation is called the **Bellman equation**, and the property is called the **dynamic programming property**. The decision to be taken at time t_n is split up in two: (1) the consumption and portfolio decision for the current period and (2) the consumption and portfolio decisions for all future periods. We take the decision for the current period assuming that we will make optimal decisions in all future periods. Note that this does not imply that the decision for the current period is taken independently from future decisions. We take into account the effect that our current decision has on the maximum expected utility we can get from all future periods. The expectation $\mathrm{E}_{t_i}\left[J_{t_{i+1}}\right]$ will depend on our choice of c_{t_i} and θ_{t_i}.

The dynamic programming property is the basis for a backward iterative solution procedure. First, we choose $c_{t_{N-1}}$ and $\theta_{t_{N-1}}$ to maximize

$$u(c_{t_{N-1}})\Delta t + e^{-\delta\Delta t}\,\mathrm{E}_{t_{N-1}}\left[\bar{u}(W_T)\right],$$

where

$$W_T = W_{t_{N-1}} + \left[\theta^0_{t_{N-1}} r_{t_{N-1}} + \theta^\top_{t_{N-1}}\mu_{t_{N-1}} + y_{t_{N-1}} - c_{t_{N-1}}\right]\Delta t + \theta^\top_{t_{N-1}}\sigma_{t_{N-1}}\varepsilon_{t_{N-1}}\sqrt{\Delta t}.$$

This is done for each possible state at time t_{N-1} and gives us $J_{t_{N-1}}$. Then we choose $c_{t_{N-2}}$ and $\theta_{t_{N-2}}$ to maximize

$$u(c_{t_{N-2}})\Delta t + e^{-\delta\Delta t}\,\mathrm{E}_{t_{N-2}}\left[J_{t_{N-1}}\right],$$

and so on until we reach time zero. Since we have to perform a maximization for each state of the world at every point in time, we have to make assumptions on the possible states at each point in time before we can implement the recursive procedure. The optimal decisions at any time are expected to depend on the wealth level of the agent at that date, but also on the value of other time-varying state variables that affect future returns on investment (e.g. the interest rate level) and future income levels. To be practically implementable only a few state variables can be incorporated. Also, these state variables must follow Markov processes so only the current values of the variables are relevant for the maximization at a given point in time. Note the similarity to the problem of determining the optimal exercise strategy of a Bermudan/American option. However, for that problem the decision to be taken is much simpler (exercise or not) than for the consumption/portfolio problem.

Under some simplifying assumptions on the precise form of the utility functions u and $\bar{u}$ and on the dynamics of asset returns and income, the backward iterative procedure yields an explicit solution to the maximization problem in the form of the optimal (possibly state- and time-dependent) consumption rate and portfolio process (and also the indirect utility of wealth J_t). Since we can obtain similar (and often clearer) results under similar assumptions in the more elegant and realistic continuous-time setting, we will not go into these discrete-time examples.

4.4 The basic continuous-time setting

The basic elements of mainstream continuous-time models can be seen as the limit of the multi-period discrete-time model elements. The basis is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with an associated filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ which is the formal model of the evolution of the relevant uncertainty for the investor.

The agent now has to choose a continuous-time process of consumption rates $c = (c_t)_{t\in[0,T]}$ and a continuous-time portfolio process $\theta = (\theta_t)_{t\in[0,T]}$. As before, θ_t is the d-dimensional vector of amounts invested at time t in the d risky assets. The remaining financial wealth $\theta^0_t = W_t - \theta_t^\top \mathbf{1} = W_t - \sum_{i=1}^d \theta_{it}$ is invested in the locally riskfree asset. A single consumption good is assumed and this good is used as a numeraire so that all prices are measured in units of this consumption good, i.e. in real terms. We will always require that $c_t \geq 0$ with probability one. We focus on unconstrained investors so that there are no constraints on the values θ_t may have, i.e. θ_t can have any value in $\mathbb{R}^d$; see references in Section 10.2 to problems with constraints on the portfolios, e.g. short-selling constraints or portfolio mix constraints. The stochastic variables c_t and θ_t must be $\mathcal{F}_t$-measurable, i.e. they can only depend on information available at time t. In other words, the processes c and θ are adapted. Other technical requirements should be added.[1] A consumption and investment strategy must also satisfy that the wealth process induced by the strategy always stays above a lower bound, say $-K$, where $K > 0$. This rules out doubling strategies, cf. the discussion in Duffie (2001, Ch. 6). In fact, we will typically require that wealth stays non-negative at all times. This is a natural requirement, at least for the case where the investor does not receive a minimum income from non-financial sources (labor). The set of all consumption and investment strategies that satisfy all these requirements on the interval $[t, T]$ is denoted by $\mathcal{A}_t$.

[1] The consumption process c must be an $\mathcal{L}^1$-process, i.e. $\int_0^T \|c_t\|\, dt < \infty$ with probability one. The portfolio strategy θ must satisfy that $\theta^\top \mu$ is an $\mathcal{L}^1$-process and that $\theta^\top \sigma$ is an $\mathcal{L}^2$-process, i.e. that $\int_0^T \|\theta_t^\top \sigma_t\|^2\, dt < \infty$ with probability one. Finally, θ must be a progressively measurable process which involves a bit more than just being adapted.

Preferences: The objective is to maximize the expected life-time utility which is assumed to be on the additively time-separable form

$$\mathrm{E}\left[\int_0^T e^{-\delta t} u(c_t)\, dt + e^{-\delta T} \bar{u}(W_T)\right], \tag{4.5}$$

where u and $\bar{u}$ are increasing and concave von Neumann-Morgenstern utility functions. We will assume that u and $\bar{u}$ are twice continuously differentiable on their domain. We will define the indirect utility process $J = (J_t)$ as

$$J_t = \sup_{(c,\theta)\in\mathcal{A}_t} \mathrm{E}_t\left[\int_t^T e^{-\delta(s-t)} u(c_s)\, ds + e^{-\delta(T-t)} \bar{u}(W_T)\right]. \tag{4.6}$$

An optimal consumption and investment strategy (c^*, θ^*) has the property that it provides at least as high an expected life-time utility as any other feasible strategy. In particular,

$$J_0 = \mathrm{E}\left[\int_0^T e^{-\delta t} u(c_t^*)\, dt + e^{-\delta T} \bar{u}(W_T^*)\right],$$

where W_T^* is the terminal wealth level that follows from the strategy (c^*, θ^*). In other words, when an optimal strategy exists the supremum in the definition of J is attained. Of course, J_0 will depend on the initial wealth W_0 of the investor. We shall assume that $J_0 < \infty$ for all $W_0 < \infty$. It can be shown that J_0 is an increasing and concave function of initial wealth W_0. See Exercise 4.1 at the end of the chapter.

Dynamics of prices and wealth: When the investor is about to choose consumption and investment strategies she has to deal with a number of variables that can evolve stochastically over time such as:

- the (locally) riskfree rate r_t (i.e. the short-term interest rate),
- the prices, the expected rates of returns, the variance-covariance matrix of rates of return on the risky assets,
- the expected rate of change and variation in her income rate,
- covariances or correlations between all these variables.

Of course, in a fuller model we should also include uncertainty e.g. about the time of death of the investor, relative prices of different consumption goods, etc., but we ignore these issues at this point.

We shall assume that all exogenous shocks to these variables can be represented by standard Brownian motions. A direct consequence is that we do not allow for any jumps in prices, except for points in time where the asset provides its owner with a lump-sum payment, e.g. a dividend payment of a stock or a coupon payment of a bond.[2] For simplicity, we assume that the assets provide no payments in the life of the investor and that the vector of risky asset prices P_t follows a stochastic process of the form

$$dP_t = \mathrm{diag}(P_t)\left[\mu_t\, dt + \sigma_t\, dz_t\right], \tag{4.7}$$

[2]See, e.g., Bardhan and Chao (1995), Wu (2000), and Jeanblanc-Picqué and Pontier (1990) for utility maximization problems involving jump processes.

where $z = (z_1, \dots, z_d)^\top$ is a d-dimensional standard Brownian motion, i.e. a vector of d independent one-dimensional standard Brownian motions. The term $\operatorname{diag}(P_t)$ denotes the $(d \times d)$-matrix with the vector P_t along the main diagonal and zeros off the diagonal. We can write this componentwise as

$$dP_{it} = P_{it} \left[\mu_{it}\, dt + \sum_{j=1}^{d} \sigma_{ijt}\, dz_{jt} \right], \quad i = 1, \dots, d.$$

The instantaneous rate of return on asset i is given by dP_{it}/P_{it}. The d-vector μ_t contains the expected rates of return and the $(d \times d)$-matrix σ_t measures the sensitivities of the risky asset prices with respect to exogenous shocks so that the $(d \times d)$-matrix $\sigma_t \sigma_t^\top$ contains the variance and covariance rates of instantaneous rates of return. We assume that σ_t is non-singular. Of course, μ and σ must be adapted to the information filtration $\mathbb{F} = (\mathcal{F}_t)$.[3] This way of modelling price dynamics in continuous-time can be seen as the limit of (4.2) when ε_{t_n} in that expression is assumed to be multivariate standard normally distributed.

Taking the limit of the wealth dynamics in (4.3) we get

$$dW_t = \left[\theta_t^0 r_t + \theta_t^\top \mu_t + y_t - c_t \right] dt + \theta_t^\top \sigma_t\, dz_t.$$

The amount invested in the (locally) riskfree asset can be expressed as total wealth minus the amounts invested in the risky assets,

$$\theta_t^0 = W_t - \theta_t^\top \mathbf{1}.$$

Substituting this into the wealth dynamics above, we obtain

$$dW_t = \left[r_t W_t + \theta_t^\top (\mu_t - r_t \mathbf{1}) + y_t - c_t \right] dt + \theta_t^\top \sigma_t\, dz_t.$$

We will frequently represent the investment strategy in terms of the fractions of wealth (the portfolio weight) invested in each asset. The portfolio weight vector is given by $\pi_t = \theta_t / W_t$ so that the weight of the riskfree asset is $1 - \pi_t^\top \mathbf{1} = 1 - \sum_{i=1}^{d} \pi_{it}$. This representation makes sense when $W_t > 0$, which also is a natural restriction at least if the investor has no non-financial income. In these terms the wealth dynamics can be rewritten as

$$dW_t = W_t \left[r_t + \pi_t^\top (\mu_t - r_t \mathbf{1}) \right] dt + \left[y_t - c_t \right] dt + W_t \pi_t^\top \sigma_t\, dz_t.$$

The investor can partially control the evolution of her wealth by her choice of (c, π), but is still subject to exogenous shocks (unless $\pi = 0$ and r_t and y_t are certain).

Since σ_t is assumed to be a non-singular $(d \times d)$-matrix, we can define the d-dimensional process $\lambda = (\lambda_t)$ by

$$\lambda_t = \sigma_t^{-1} (\mu_t - r_t \mathbf{1}),$$

so that

$$\mu_t = r_t \mathbf{1} + \sigma_t \lambda_t,$$

i.e. $\mu_{it} = r_t + \sum_{j=1}^{d} \sigma_{ijt} \lambda_{jt}$. λ has the interpretation of a vector of market prices of risk (corresponding to the shock process z) since it measures the excess rate of return relative to the standard

[3] Further technical requirements should be imposed, e.g. that the processes r, μ, and σ are progressively measurable, that $\operatorname{diag}(P_t)\mu_t$ is an $\mathcal{L}^1$-process, and that $\operatorname{diag}(P_t)\sigma_t$ is an $\mathcal{L}^2$-process; cf. footnote 1.

deviation. For example, if asset i is only sensitive to the first component of the exogenous shock z_t, it will have $\sigma_{i2t} = \cdots = \sigma_{idt} = 0$ and hence an expected rate of return of $\mu_{it} = r_t + \sigma_{i1t}\lambda_{1t}$ so that $\lambda_{1t} = (\mu_{it} - r_t)/\sigma_{i1t}$, where σ_{i1t} is identical to the volatility of the asset. We can now rewrite the price dynamics as

$$dP_t = \operatorname{diag}(P_t)\left[(r_t\mathbf{1} + \sigma_t\lambda_t)\, dt + \sigma_t\, dz_t\right].$$

The dynamics of wealth induced by a consumption and investment strategy (c, π) can be restated as

$$dW_t = W_t\left[r_t + \pi_t^\top \sigma_t \lambda_t\right] dt + \left[y_t - c_t\right] dt + W_t \pi_t^\top \sigma_t\, dz_t. \tag{4.8}$$

Solution techniques: There are two major questions to be answered: (i) Under which assumptions do optimal strategies exist, and (ii) How can optimal strategies (and the indirect utility function) be computed. In these notes we will focus on the second question. There are two major approaches for solving this type of optimization problems: the dynamic programming approach (also known as the stochastic control approach) and the martingale approach.

4.5 Dynamic programming in continuous-time models

In a previous section we introduced the dynamic programming approach in a discrete-time multi-period setting. Apparently, Merton (1969, 1971) was the first to apply the dynamic programming approach to a continuous-time optimal consumption/investment problem. The dynamic programming approach requires that a (possibly multi-dimensional) state variable exists so that this variable follows a Markov process and all relevant objects can be written as functions of this state variable and time. The theory of dynamic programming contains some results on the existence of optimal strategies, but they often require that all admissible strategies take values in a compact set, an assumption which is certainly unsuitable for most portfolio problems. Therefore, verification theorems are typically applied. This involves solving the Hamilton-Jacobi-Bellman (HJB) equation associated with the control problem. Under some technical conditions the solution to the HJB equation will give us both the optimal strategies and the indirect utility function. The HJB equation is a fully non-linear second-order partial differential equation. Despite the complexity of the equation explicit solutions have been found in many interesting settings, as we shall see in the following chapters.

In order to apply the dynamic programming approach in continuous-time there must exist a process $x = (x_t)$, possibly multi-dimensional, such that the pair (W_t, x_t) captures all relevant information for the agent's decision at time t. Basically, the pair of stochastic processes (W, x) must constitute a Markov system. If both r, λ, σ, λ, and y are constant (or at least deterministic functions of time), then the wealth process is by itself a Markov process and we need not add some x. We will refer to this situation as the case of constant investment opportunities. We study portfolio and consumption choice under that assumption in detail in Chapter 5. However, we do know that for example the short-term interest rate varies stochastically over time. If $r = (r_t)$ is in itself a Markov process, we should include r as a state variable, i.e. one of the elements of x should be r. Maybe a multiple state variables are needed to capture the interest rate dynamics. Then these variables should be included in x. We will study examples of such so-called stochastic investment opportunities in Chapters 6–8.

For simplicity we assume in the following that the agent receives no labor income, i.e. $y_t \equiv 0$. We assume further that there is a stochastically evolving state variable $x = (x_t)$ that captures the variations in r, μ, and σ over time, i.e.

$$r_t = r(x_t), \quad \mu_t = \mu(x_t, t), \quad \sigma_t = \sigma(x_t, t),$$

where r, μ, and σ now (also) denote sufficiently well-behaved functions. The variations in the state variable x determine the future expected returns and covariance structure in the financial market. The market price of risk is also given by the state variable:

$$\lambda(x_t) = \sigma(x_t, t)^{-1} \left(\mu(x_t, t) - r(x_t)\mathbf{1}\right).$$

Note that we have assumed that the short-term interest rate r_t and the market price of risk vector λ_t do not depend on calendar time directly. The fluctuations in r_t and λ_t over time are presumably not due to the mere passage of time, but rather due to variations in some more fundamental economic variables. In contrast, the expected rates of returns and the price sensitivities of some assets will depend directly on time, e.g. the volatility and the expected rate of return on a bond will depend on the time-to-maturity of the bond and therefore on calendar time. The income rate may also depend on time through life-cycle variations in labor income.

For simplicity we will first assume that x is one-dimensional in the following analysis and afterwards turn to the multi-dimensional state variables. The wealth process for a given portfolio and consumption strategy now evolves as

$$dW_t = W_t \left[r(x_t) + \pi_t^\top \sigma(x_t, t)\lambda(x_t)\right] dt - c_t\, dt + W_t \pi_t^\top \sigma(x_t, t)\, dz_t.$$

The state variable x is assumed to follow a one-dimensional diffusion process

$$dx_t = m(x_t)\, dt + v(x_t)^\top\, dz_t + \hat{v}(x_t)\, d\hat{z}_t, \tag{4.9}$$

where $\hat{z} = (\hat{z}_t)$ is a one-dimensional standard Brownian motion independent of $z = (z_t)$. Hence, if $\hat{v}(x_t) \neq 0$, there is an exogenous shock to the state variable that cannot be hedged by investments in the financial market. In other words, the financial market is incomplete. Conversely, if $\hat{v}(x_t)$ is identically equal to zero, the financial market is complete. We shall consider examples of both cases later. The d-vector $v(x_t)$ represents the sensitivity of the state variable with respect to the exogenous shocks to market prices. Note that the d-vector $\sigma(x, t)v(x)$ is the vector of instantaneous covariance rates between the returns on the risky assets and the state variable.

The pair (W_t, x_t) forms a two-dimensional Markov diffusion process that contains all the information the investor needs for making her consumption/investment decision. The indirect utility at time t is therefore $J_t = J(W_t, x_t, t)$, where the function J is given by

$$J(W, x, t) = \sup_{(c_s, \pi_s)_{s\in[t,T]}} \mathrm{E}_{W,x,t} \left[\int_t^T e^{-\delta(s-t)} u(c_s)\, ds + e^{-\delta(T-t)} \bar{u}(W_T)\right].$$

In a discrete-time approximation of this setting, it follows from (4.4) that

$$J(W, x, t) = \sup_{c_t \geq 0, \pi_t \in \mathbb{R}^d} \left\{u(c_t)\Delta t + e^{-\delta \Delta t}\, \mathrm{E}_{W,x,t}\left[J(W_{t+\Delta t}, x_{t+\Delta t}, t + \Delta t)\right]\right\},$$

where c_t and π_t is held fixed over the interval $[t, t+\Delta t)$. If we multiply by $e^{\delta \Delta t}$, subtract $J(W, x, t)$, and then divide by Δt, we get

$$\frac{e^{\delta \Delta t} - 1}{\Delta t} J(W, x, t) = \sup_{c_t \geq 0, \pi_t \in \mathbb{R}^d} \left\{ e^{\delta \Delta t} u(c_t) + \frac{1}{\Delta t} \mathrm{E}_{W,x,t} \left[J(W_{t+\Delta t}, x_{t+\Delta t}, t + \Delta t) - J(W, x, t) \right] \right\}.$$

When we let $\Delta t \to 0$, we have that

$$\frac{e^{\delta \Delta t} - 1}{\Delta t} \to \delta,$$

and that

$$\frac{1}{\Delta t} \mathrm{E}_{W,x,t} \left[J(W_{t+\Delta t} x_{t+\Delta t}, t + \Delta t) - J(W, x, t) \right] \tag{4.10}$$

will approach the drift of J at time t, which according to Itô's Lemma is given by

$$\begin{aligned} &\frac{\partial J}{\partial t}(W, x, t) + J_W(W, x, t) \left(W \left[r(x) + \pi_t^\top \sigma(x, t) \lambda(x) \right] - c_t \right) \\ &+ \frac{1}{2} J_{WW}(W, x, t) W^2 \pi_t^\top \sigma(x, t) \sigma(x, t)^\top \pi_t + J_x(W, x, t) m(x) \\ &+ \frac{1}{2} J_{xx}(W, x, t) (v(x)^\top v(x) + \hat{v}(x)^2) + J_{Wx}(W, x, t) W \pi_t^\top \sigma(x, t) v(x). \end{aligned}$$

The limit of (4.10) is therefore

$$\begin{aligned} \delta J(W, x, t) = \sup_{c \geq 0, \pi \in \mathbb{R}^d} \Big\{ u(c) &+ \frac{\partial J}{\partial t}(W, x, t) + J_W(W, x, t) \left(W \left[r(x) + \pi^\top \sigma(x, t) \lambda(x) \right] - c_t \right) \\ &+ \frac{1}{2} J_{WW}(W, x, t) W^2 \pi^\top \sigma(x, t) \sigma(x, t)^\top \pi + J_x(W, x, t) m(x) \\ &+ \frac{1}{2} J_{xx}(W, x, t) (v(x)^\top v(x) + \hat{v}(x)^2) \\ &+ J_{Wx}(W, x, t) W \pi^\top \sigma(x, t) v(x) \Big\}. \end{aligned} \tag{4.11}$$

This is called the **Hamilton-Jacobi-Bellman (HJB) equation** corresponding to the dynamic optimization problem. Subscripts on J denote partial derivatives, however we will write the partial derivative with respect to time as $\partial J / \partial t$ to distinguish it from the value J_t of the indirect utility process. The HJB equation involves the supremum over the feasible time t consumption rates and portfolios and is therefore a highly non-linear second-order partial differential equation.

From the analysis above we will expect that the indirect utility function $J(W, x, t)$ solves the HJB equation for all possible values of W and x and all $t \in [0, T)$ and that it satisfies the terminal condition

$$J(W, x, T) = \bar{u}(W)$$

for all W and x. In the mathematical literature on stochastic control problems like the one we are looking at, there are a few results on when a solution to the HJB equation exists. However, these results are only valid under restrictive conditions, e.g. that the controls (c and π in our case) can only take values in a compact set. This is generally not true for the consumption/investment problems. We are mostly interested in finding a solution. Here, we can apply a verification result. Let us formulate the result for the problem with a one-dimensional state variable:

Theorem 4.1 *Assume $V(W,x,t)$ solves the HJB equation* (5.3) *and satisfies some technical conditions. Let $C(W,t)$ and $\Pi(W,t)$ be given by*

$$\begin{aligned}(C(W,t),\Pi(W,t)) = \underset{c\geq 0,\pi\in\mathbb{R}^d}{\arg\max}\Big\{ & u(c) + \frac{\partial V}{\partial t}(W,x,t) + V_W(W,x,t)\left(W\left[r(x)+\pi^\top\sigma(x,t)\lambda(x)\right]-c\right) \\ & + \frac{1}{2}V_{WW}(W,x,t)W^2\pi^\top\sigma(x,t)\sigma(x,t)^\top\pi + V_x(W,x,t)m(x) \\ & + \frac{1}{2}V_{xx}(W,x,t)(v(x)^\top v(x)+\hat{v}(x)^2) \\ & + V_{Wx}(W,x,t)W\pi^\top\sigma(x,t)v(x)\Big\}.\end{aligned} \tag{4.12}$$

If the strategies

$$c_t^* = C(W_t^*,x_t,t), \qquad \pi_t^* = \Pi(W_t^*,x_t,t),$$

where (W_t^*) is the wealth process that (c^*,π^*) induces, are feasible (i.e. $(c,\pi)\in\mathcal{A}_0$), then they are optimal, and V equals the indirect utility function, i.e.

$$J(W,x,t) = V(W,x,t) = \mathrm{E}_{W,x,t}\left[\int_t^T e^{-\delta(s-t)}u(c_s^*)\,ds + \bar{u}(W_T^*)\right].$$

For a formal proof and a precise statement of the technical conditions, see e.g. Theorem 11.2.2 in Øksendal (1998) or Theorem III.8.1 in Fleming and Soner (1993).

Suppose now that the state variable x is k-dimensional and follows the diffusion process

$$dx_t = m(x_t)\,dt + v(x_t)^\top\,dz_t + \hat{v}(x_t)\,d\hat{z}_t, \tag{4.13}$$

where m now is a k-vector valued function, v is a $(d\times k)$-matrix valued function[4], $\hat{v}$ is a $(k\times k)$-matrix valued function, and $\hat{z}$ is a k-dimensional standard Brownian motion independent of z. The basic derivation is the same as with a one-dimensional state variable, but the drift of J now becomes more complicated and so does the HJB equation:

$$\begin{aligned}\delta J(W,x,t) = \sup_{c\geq 0,\pi\in\mathbb{R}^d}\Big\{ & u(c) + \frac{\partial J}{\partial t}(W,x,t) + J_W(W,x,t)\left(W\left[r(x)+\pi^\top\sigma(x,t)\lambda(x)\right]-c\right) \\ & + \frac{1}{2}J_{WW}(W,x,t)W^2\pi^\top\sigma(x,t)\sigma(x,t)^\top\pi + J_x(W,x,t)^\top m(x) \\ & + \frac{1}{2}\operatorname{tr}\left(J_{xx}(W,x,t)[v(x)^\top v(x)+\hat{v}(x)\hat{v}(x)^\top]\right) \\ & + W\pi^\top\sigma(x,t)v(x)J_{Wx}(W,x,t)\Big\}.\end{aligned} \tag{4.14}$$

Now, J_{Wx} is a k-vector and J_{xx} a $(k\times k)$-matrix. The notation $\operatorname{tr}(A)$ stands for the trace of the square matrix A, which is defined as the sum of the diagonal elements, $\operatorname{tr}(A)=\sum_i A_{ii}$.

In the special case of constant investment opportunities, the indirect utility is given by $J_t = J(W_t,t)$ and the corresponding HJB equation is simply

$$\begin{aligned}\delta J(W,t) = \sup_{c\geq 0,\pi\in\mathbb{R}^d}\Big\{ & u(c) + \frac{\partial J}{\partial t}(W,t) + J_W(W,t)\left(W\left[r+\pi^\top\sigma\lambda\right]-c\right) \\ & + \frac{1}{2}J_{WW}(W,t)W^2\pi^\top\sigma\sigma^\top\pi\Big\}.\end{aligned} \tag{4.15}$$

[4] In this multi-dimensional setting it would be natural to write the dz_t-term in the state dynamics on the form $v(x_t)\,dz_t$, but this would conflict with our notation in the one-dimensional case, where we have used the term $v(x_t)^\top\,dz_t$.

4.6 The martingale approach to consumption-portfolio problems

The dynamic programming approach requires the existence of a finite-dimensional Markov process $x = (x_t)$ such that the indirect utility function of the investor can be written as $J_t = J(W_t, x_t, t)$. The dynamic programming approach does not allow many conclusions on problems where the PDE cannot be solved explicitly. For example, it is hard to tell whether an optimal strategy actually exists. In contrast, the martingale approach does not require additional assumptions on the stochastic processes that the investor cannot control beyond those outlined in Section 4.4. In particular, we do not have to assume that the interest rates, price variances etc. are fully described by a finite-dimensional Markov process.

We go back to the general model for risky asset prices stated in (4.7). We consider a complete market so that the variations in the riskfree rate of return r_t, expected rates of return μ_t, and variances and covariances defined by σ_t between rates of return are caused by the same d-dimensional standard Brownian motion z that affects the risky asset prices. Therefore, the market price of risk vector λ_t defined by

$$\lambda_t = \sigma_t^{-1} (\mu_t - r_t \mathbf{1})$$

summarizes the risk-return tradeoff of all risks. In a complete market there is a unique state price density process (a.k.a. the pricing kernel) $\zeta = (\zeta_t)$ given by

$$\zeta_t = \exp\left\{ -\int_0^t r_s\, ds - \int_0^t \lambda_s^\top\, dz_s - \frac{1}{2}\int_0^t \lambda_s^\top \lambda_s\, ds \right\}, \tag{4.16}$$

Note that the state-price density evolves as

$$d\zeta_t = -\zeta_t \left[r_t\, dt + \lambda_t\, dz_t \right]. \tag{4.17}$$

We also have a unique equivalent martingale measure (also known as the risk-neutral probability measure) $\mathbb{Q}$ defined by the Radon-Nikodym derivative $d\mathbb{Q}/d\mathbb{P} = \exp\{\int_0^T r_s\, ds\}\zeta_T$. We assume that λ is an $\mathcal{L}^2[0,T]$ process. The time zero price of a stochastic payoff X_T at some point T is given by

$$\mathrm{E}^{\mathbb{Q}}\left[e^{-\int_0^T r_s\, ds} X_T \right] = \mathrm{E}\left[\zeta_T X_T \right].$$

Similarly, the time t price is

$$\mathrm{E}_t^{\mathbb{Q}}\left[e^{-\int_t^T r_s\, ds} X_T \right] = \mathrm{E}_t\left[\frac{\zeta_T}{\zeta_t} X_T \right].$$

For simplicity we assume that the investor receives no income from non-financial sources. Then a natural constraint on the investor's choice of consumption and portfolio strategy (c, π) at time 0 is that

$$\mathrm{E}\left[\int_0^T \zeta_t c_t\, dt + \zeta_T W_T \right] \le W_0,$$

where W_T is the terminal wealth induced by (c, π) and W_0 is the initial wealth of the investor. This simply says that the time zero "price" of the strategy cannot exceed initial wealth available. This is shown rigorously in the following theorem. But first we recall from (4.8) that wealth evolves as

$$dW_t = W_t \left[r_t + \pi_t^\top \sigma_t \lambda_t \right] dt - c_t\, dt + W_t \pi_t^\top \sigma_t\, dz_t.$$

From this, (4.17), and Itô's Lemma we get that

$$d\left(\zeta_t W_t\right) = -\zeta_t c_t\, dt + \zeta_t W_t \left(\pi_t^\top \sigma_t - \lambda_t^\top \right) dz_t, \tag{4.18}$$

or equivalently

$$\zeta_t W_t + \int_0^t \zeta_s c_s \, ds = W_0 + \int_0^t \zeta_s W_s \left(\pi_s^\top \sigma_s - \lambda_s^\top\right) dz_s. \tag{4.19}$$

Theorem 4.2 *If (c, π) is a feasible strategy, then*

$$\mathrm{E}\left[\int_0^T \zeta_t c_t \, dt + \zeta_T W_T\right] \leq W_0,$$

where W_T is the terminal wealth induced by (c, π).

Proof: Define the stopping times $(\tau_n)_{n\in\mathbb{N}}$ by

$$\tau_n = T \wedge \inf\left\{t \in [0,T] \,\middle|\, \int_0^t \|\zeta_s W_s \left[\pi_s^\top \sigma_s - \lambda_s\right]\|^2 \, ds \geq n\right\}.$$

Then the stochastic integral on the right-hand side of (4.19) is a martingale on $[0, \tau_n]$. Taking expectations in (4.19) leaves us with

$$\mathrm{E}\left[\zeta_{\tau_n} W_{\tau_n}\right] + \mathrm{E}\left[\int_0^{\tau_n} \zeta_t c_t \, dt\right] = W_0.$$

Letting $n \uparrow \infty$, we have $\tau_n \uparrow T$, and it can be shown by use of Lebesgue's monotone convergence theorem that

$$\mathrm{E}\left[\int_0^{\tau_n} \zeta_t c_t \, dt\right] \to \mathrm{E}\left[\int_0^T \zeta_t c_t \, dt\right].$$

Furthermore, Fatou's lemma can be applied to show that

$$\liminf_{n\to\infty} \mathrm{E}\left[\zeta_{\tau_n} W_{\tau_n}\right] \geq \mathrm{E}\left[\zeta_T W_T\right].$$

The claim now follows. □

The idea of the martingale approach is to focus on the *static* optimization problem

$$\sup_{(c,W)} \mathrm{E}\left[\int_0^T e^{-\delta t} u(c_t) \, dt + e^{-\delta T} \bar{u}(W)\right], \tag{4.20}$$

$$\text{s.t. } \mathrm{E}\left[\int_0^T \zeta_t c_t \, dt + \zeta_T W\right] \leq W_0$$

rather than the original *dynamic* problem

$$\sup_{(c,\pi)} \mathrm{E}\left[\int_0^T e^{-\delta t} u(c_t) \, dt + e^{-\delta T} \bar{u}(W_T)\right], \tag{4.21}$$

$$\text{s.t. } dW_t = W_t \left[r_t + \pi_t^\top \sigma_t \lambda_t\right] dt - c_t \, dt + W_t \pi_t^\top \sigma_t \, dz_t. \tag{4.22}$$

In the static problem the agent chooses the terminal wealth directly, whereas in the dynamic problem the terminal wealth follows from the portfolio strategy (and the consumption strategy). For the terminal wealth variable W, the agent is allowed to choose among the non-negative, integrable and $\mathcal{F}_T$-measurable random variables. This approach was suggested by Karatzas, Lehoczky, and Shreve (1987) and Cox and Huang (1989, 1991). Some preliminary aspects were addressed by Pliska (1986).

The Lagrangian for the constrained optimization problem (4.20) is given by

$$\begin{aligned}\mathcal{L} &= \mathrm{E}\left[\int_0^T e^{-\delta t} u(c_t)\,dt + e^{-\delta T}\bar{u}(W)\right] + \psi\left(W_0 - \mathrm{E}\left[\int_0^T \zeta_t c_t\,dt + \zeta_T W\right]\right)\\ &= \psi W_0 + \mathrm{E}\left[\int_0^T \left(e^{-\delta t}u(c_t) - \psi\zeta_t c_t\right)\,dt + \left(e^{-\delta T}\bar{u}(W) - \psi\zeta_T W\right)\right],\end{aligned}$$

where ψ is a Lagrange multiplier. We maximize the last expectation by maximizing $\left(e^{-\delta T}\bar{u}(W) - \psi\zeta_T W\right)$ with respect to W for each possible value of ζ_T and maximizing $\left(e^{-\delta t}u(c_t) - \psi\zeta_t c_t\right)$ with respect to c_t for each t and each possible value of ζ_t. This results in the first-order conditions

$$e^{-\delta t}u'(c_t) = \psi\zeta_t, \qquad e^{-\delta T}\bar{u}'(W) = \psi\zeta_T,$$

where ψ is then chosen such that the inequality constraint holds as an equality. Let $I_u(\cdot)$ denote the inverse of the marginal utility function $u'(\cdot)$ and $I_{\bar{u}}(\cdot)$ the inverse of $\bar{u}'(\cdot)$. Define

$$\mathcal{H}(\psi) = \mathrm{E}\left[\int_0^T \zeta_t I_u(\psi e^{\delta t}\zeta_t)\,dt + \zeta_T I_{\bar{u}}(\psi e^{\delta T}\zeta_T)\right]. \tag{4.23}$$

Since marginal utility is decreasing, this is also the case for the inverse of marginal utility and hence also for the function $\mathcal{H}$. We will assume that $\mathcal{H}(\psi)$ is finite for all $\psi > 0$. Under this assumption, $\mathcal{H}$ has an inverse denoted by $\mathcal{Y}$. The next theorem says that the optimal policy in the static problem is feasible and optimal in the dynamic problem.

Theorem 4.3 *The optimal consumption rate is given by*

$$c_t^* = I_u\left(\mathcal{Y}(W_0)e^{\delta t}\zeta_t\right).$$

Under the optimal portfolio strategy the terminal wealth level is

$$W^* = I_{\bar{u}}\left(\mathcal{Y}(W_0)e^{\delta T}\zeta_T\right).$$

The wealth process under the optimal policy is given by

$$W_t^* = \frac{1}{\zeta_t}\,\mathrm{E}_t\left[\int_t^T \zeta_s c_s^*\,ds + \zeta_T W^*\right]. \tag{4.24}$$

Proof: First note that, by concavity of $u(\cdot)$ and $\bar{u}(\cdot)$,

$$\begin{aligned} u(I_u(z)) - u(c) &\geq z\left(I_u(z) - c\right), \quad \forall c, z > 0,\\ \bar{u}(I_{\bar{u}}(z)) - \bar{u}(W) &\geq z\left(I_{\bar{u}}(z) - W\right), \quad \forall W, z > 0.\end{aligned}$$

Hence, for any feasible strategy (c, π) with associated terminal wealth W, we have that

$$\begin{aligned}&\mathrm{E}\left[\int_0^T e^{-\delta t}\left(u(c_t^*) - u(c_t)\right)\,dt + e^{-\delta T}\left(\bar{u}(W^*) - \bar{u}(W)\right)\right]\\ &\qquad \geq \mathrm{E}\left[\int_0^T \mathcal{Y}(W_0)\zeta_t\left(c_t^* - c_t\right)\,dt + \mathcal{Y}(W_0)\zeta_T\left(W^* - W\right)\right]\\ &\qquad \geq 0,\end{aligned}$$

where the last inequality follows from the fact that, by Theorem 4.2,

$$\mathrm{E}\left[\int_0^T \zeta_t c_t\, dt + \zeta_T W\right] \leq W_0,$$

and, per construction,

$$\mathrm{E}\left[\int_0^T \zeta_t c_t^*\, dt + \zeta_T W^*\right] = W_0.$$

Thus, if there is a portfolio strategy π^* such that (c^*, π^*) is feasible and gives a terminal wealth of W^*, then the strategy (c^*, π^*) will be optimal. Define the process W^* by (4.24). Obviously,

$$\zeta_t W_t^* + \int_0^t \zeta_s c_s^*\, ds = \mathrm{E}_t\left[\int_0^T \zeta_s c_s^*\, ds + \zeta_T W_T^*\right]$$

defines a martingale, so by the martingale representation theorem, an adapted $\mathcal{L}^2[0,T]$ process η exists such that

$$\zeta_t W_t^* + \int_0^t \zeta_s c_s^* = W_0 + \int_0^t \eta_s^\top\, dz_s. \tag{4.25}$$

Define a portfolio process π by

$$\pi_t = (\sigma_t^\top)^{-1}\left(\frac{\eta_t}{W_t^* \zeta_t} + \lambda_t\right)$$

(with the remaining wealth $W_t^*(1 - \pi_t^\top \mathbf{1})$ invested in the bank account). A comparison of (4.25) and (4.19) shows that this strategy together with the consumption strategy c^* indeed yield a terminal wealth of W^* and the process (W_t^*) is the wealth process corresponding to this strategy. □

Note that the indirect utility at time 0 as a function of initial wealth W_0 is

$$\begin{aligned} J(W_0) &= \mathrm{E}\left[\int_0^T e^{-\delta t} u(c_s^*)\, ds + e^{-\delta T}\bar{u}(W^*)\right] \\ &= \mathrm{E}\left[\int_0^T e^{-\delta t} u\left(I_u(\mathcal{Y}(x) e^{\delta t}\zeta_t)\right)\, dt + e^{-\delta T}\bar{u}\left(I_{\bar{u}}(\mathcal{Y}(x) e^{\delta T}\zeta_T)\right)\right]. \end{aligned}$$

We shall demonstrate how to apply the martingale approach in Section 6.7. The martingale approach is in many aspects more elegant and it is better suited for answering the existence question under general conditions. However, the existence of an optimal portfolio strategy is based on the martingale representation theorem, which in itself does not give an explicit representation of the optimal portfolio, nor a way to compute it. In some settings the martingale approach can give an abstract characterization of both the optimal consumption and portfolio strategy even for non-Markov dynamics, but in order to obtain explicit expressions for the optimal strategies the setting is typically specialized to a Markov setting. So far, there are only a few examples of explicit solutions computed with the martingale approach where the solution could not have been found by an application of the dynamic programming approach. (See Munk and Sørensen (2003) for one example.) However, in some of the relatively simple problems, such as the complete markets case studied by Cox and Huang (1989), it can be shown that the optimal portfolio policies can be found by solving a partial differential equation (PDE), which has a simpler structure than the HJB equation.

Above we have assumed complete financial markets. If the variations in r_t, μ_t, and σ_t are influenced by other exogenous shocks than prices, then we face an incomplete market. There will be no unique market price of risk on the extra shock components and hence there will be a multitude of state-price densities ζ^ν consistent with no-arbitrage. In this case the appropriate static budget constraint is that

$$\mathrm{E}\left[\int_0^T \zeta_t^\nu c_t \, dt + \zeta_T^\nu W_T\right] \leq W_0,$$

for all possible state-price densities ζ^ν. Of course, this complicates the analysis considerably. The interested reader is referred to Karatzas, Lehoczky, Shreve, and Xu (1991), Cvitanić and Karatzas (1992), He and Pearson (1991), and Cuoco (1997).

4.7 Exercises

EXERCISE 4.1 Show that the indirect utility, J_t, defined in (4.6) is an increasing and concave function of wealth, W_t. *Hint:* To show concavity, let (c_1, θ_1) be the optimal strategy with initial wealth W_1 and let (c_2, θ_2) be the optimal strategy with initial wealth W_2. Here, c_i is the consumption rate and θ_i the vector of dollar amounts invested in the risky assets. The corresponding terminal wealth levels are denoted W_{1T} and W_{2T}, respectively. For any $\alpha \in (0,1)$, you should first show that the strategy $(\alpha c_1 + (1-\alpha)c_2, \alpha\theta_1 + (1-\alpha)\theta_2)$ is a feasible strategy with initial wealth $\alpha W_{1t} + (1-\alpha)W_{2t}$ that results in the terminal wealth $\alpha W_{1T} + (1-\alpha)W_{2T}$. Then apply that u and $\bar{u}$ are assumed concave.

Chapter 5

Asset allocation with constant investment opportunities

In this chapter we will consider the relatively simple case in which the short-term interest rate r, the expected rates of return μ, and the volatility matrix σ of the risky assets are all assumed to be constant through time. The market price of risk vector λ is therefore also a constant. We shall also assume that the investor has no other income, i.e. $y = 0$. This is the problem originally considered by Merton (1969). A direct consequence of these additional assumptions is that the risky asset price processes in (4.7) become geometric Brownian motions so that future risky asset prices are lognormally distributed, as is well-known from the Black-Scholes model for stock option pricing; see, e.g., Hull (2003). In this case the wealth dynamics for a given consumption strategy c and a given portfolio weight process π is

$$dW_t = W_t \left[r + \pi_t^\top \sigma \lambda \right] dt - c_t \, dt + W_t \pi_t^\top \sigma \, dz_t, \tag{5.1}$$

and the indirect utility function (sometimes called the value function) is a function of only current wealth and time

$$J(W,t) = \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathrm{E}_{W,t} \left[\int_t^T e^{-\delta(s-t)} u(c_s) \, ds + e^{-\delta(T-t)} \bar{u}(W_T) \right], \tag{5.2}$$

where $\mathrm{E}_{W,t}$ denotes the expectations operator given $W_t = W$ (and given the chosen consumption and investment strategies). We will apply the dynamic programming approach and try to solve the HJB equation associated with the utility maximization problem. From (4.15), we have that the HJB equation is given by

$$\begin{aligned} \delta J(W,t) = \sup_{c\geq 0, \pi\in\mathbb{R}^d} \Big\{ & u(c) + \frac{\partial J}{\partial t}(W,t) + J_W(W,t) \left(W \left[r + \pi^\top \sigma \lambda \right] - c \right) \\ & + \frac{1}{2} J_{WW}(W,t) W^2 \pi^\top \sigma \sigma^\top \pi \Big\}. \end{aligned} \tag{5.3}$$

We will first consider general utility functions and later specialize to CRRA utility for which explicit solutions can be obtained.

5.1 General utility function

Now let us try to solve our consumption and investment problem by an application of the verification theorem, Theorem 4.1, i.e. by solving the HJB equation (5.3). Maximizing the right-hand side of the HJB equation with respect to c gives the first-order condition

$$u'(c) = J_W(W,t), \tag{5.4}$$

where we have used the fact that the non-negativity constraint on consumption will not be binding under the assumption that marginal utility is infinite for zero consumption (or even at a positive subsistence level of consumption). This optimality condition is called the *envelope condition* and says that the marginal utility from currently consuming one unit more in optimum must equal the marginal utility from investing that unit optimally. This is an intuitive optimality condition for intertemporal choice. If we by I denotes the inverse of marginal utility $u'(c)$, we can write our candidate for the optimal consumption strategy as

$$c_t^* = C(W_t^*, t),$$

where

$$C(W,t) = I(J_W(W,t)). \tag{5.5}$$

The (unconstrained) maximization with respect to π gives the first-order condition

$$J_W(W,t)W\sigma\lambda + J_{WW}(W,t)W^2\sigma\sigma^\top\pi = 0.$$

Isolating π we get

$$\pi = -\frac{J_W(W,t)}{WJ_{WW}(W,t)}(\sigma^\top)^{-1}\lambda,$$

so that our candidate for the optimal investment strategy can be written as

$$\pi_t^* = \Pi(W_t^*, t),$$

where

$$\Pi(W,t) = -\frac{J_W(W,t)}{WJ_{WW}(W,t)}(\sigma^\top)^{-1}\lambda = -\frac{J_W(W,t)}{WJ_{WW}(W,t)}(\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1}). \tag{5.6}$$

Note that the fraction $-J_W(W,t)/[WJ_{WW}(W,t)]$ is the relative risk tolerance (i.e. the inverse of the relative risk aversion) of the indirect utility function. The optimal risky investments is therefore the relative risk tolerance of the investor times a vector that is the same for all investors (assuming they have the same perceptions about σ, μ, and r), namely the inverse of the variance-covariance matrix multiplied by the vector of excess expected rates of return. The second-order conditions for a maximum is satisfied since J is concave in W and u is concave in c.

Substituting the candidate optimal values of c and π back into the HJB equation and gathering terms, we get the second order PDE

$$\begin{aligned}\delta J(W,t) = u\big(I(J_W(W,t))\big) - J_W(W,t)I(J_W(W,t)) + \frac{\partial J}{\partial t}(W,t)\\ + rWJ_W(W,t) - \frac{1}{2}\lambda^\top\lambda\frac{J_W(W,t)^2}{J_{WW}(W,t)}.\end{aligned} \tag{5.7}$$

If this PDE has a solution $J(W,t)$ such that the strategy defined by (5.5) and (5.6) is feasible (satisfies the technical conditions), then we know from the verification theorem that this strategy is indeed the optimal consumption and investment strategy and the function $J(W,t)$ is indeed the indirect utility function. We shall sometimes consider problems with no utility from intermediate consumption, i.e. $u \equiv 0$. In that case, it is of course optimal not to consume, and it is relatively easy to see that the first two terms of the right-hand side of (5.7) will vanish, i.e. the equation simplifies to

$$\delta J(W,t) = \frac{\partial J}{\partial t}(W,t) + rWJ_W(W,t) - \frac{1}{2}\lambda^\top \lambda \frac{J_W(W,t)^2}{J_{WW}(W,t)}. \tag{5.8}$$

In the following sections we shall obtain simple, closed-form solutions for problems with CRRA and logarithmic utility. In Exercise 5.3 at the end of the chapter we will consider the problem with a subsistence HARA utility function, where a simple solution also can be obtained. Semi-explicit solutions for other utility functions have been given by Karatzas, Lehoczky, Sethi, and Shreve (1986). Merton (1971, Sec. 6) claimed to have found a solution for the general class of HARA functions (2.2) but as noted by Sethi and Taksar (1988), this solution does not satisfy the non-negativity constraints on wealth and consumption.

Without further computations we can already note an important result: With constant r, μ, and σ, **two-fund separation** obtains in the continuous-time setting. This is obvious from the optimal investment strategy in (5.6).

Theorem 5.1 (Two-fund separation) *In a financial market with constant r, μ, and σ, the optimal investment strategy of any unconstrained investor with time-separable utility of the form* (4.5) *and no non-financial income is a combination of the riskfree asset and a single portfolio of risky assets given by the weights*

$$\pi^{tan} = \frac{1}{\mathbf{1}^\top (\sigma^\top)^{-1}\lambda}(\sigma^\top)^{-1}\lambda. \tag{5.9}$$

The investor will invest the fraction $-\frac{J_W(W,t)}{WJ_{WW}(W,t)}\mathbf{1}^\top\left(\sigma^\top\right)^{-1}\lambda$ of her wealth in the risky fund and the remaining wealth in the riskfree asset.

The portfolio $\pi^{\tan}$ is almost indistinguishable from the tangency portfolio (3.10) of the one-period mean-variance analysis, but in the continuous-time case the relevant expected rates of return and variances and covariances are measured over the next infinitesimal period of time. With this little modification of the interpretation we can again look at the investment problem graphically in a (standard deviation,mean)-diagram as we are used to from the static one-period setting. Also note that the necessary assumption of lognormal prices is much more realistic than the normality assumption in the one-period model. Analogous to the one-period setting, the two-fund separation result above is the basis for a capital market equilibrium result, which in the continuous-time case is referred to as the Intertemporal Capital Asset Pricing Model (ICAPM) or the Continuous-time CAPM; see, e.g., Merton (1973b), Merton (1992, Ch. 15), Duffie (2001), and Cochrane (2001) for more on equilibrium asset pricing.

5.2 CRRA utility function

We will now focus on the case where the utility function exhibits constant relative risk aversion. We are interesting in three types of problems:

(1) utility from consumption only,

(2) utility from terminal wealth only,

(3) utility both from consumption and terminal wealth.

We can solve all three problems simultaneously by introducing two indicator parameters ε_1 and ε_2 that are either 1 or 0. We let

$$u(c) = \varepsilon_1 \frac{c^{1-\gamma}}{1-\gamma}, \quad \bar{u}(W) = \varepsilon_2 \frac{W^{1-\gamma}}{1-\gamma}.$$

The three situations above corresponds to (1) $\varepsilon_1 = 1, \varepsilon_2 = 0$, (2) $\varepsilon_1 = 0, \varepsilon_2 = 1$, and (3) $\varepsilon_1 = \varepsilon_2 = 1$. The indirect utility function is

$$J(W,t) = \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathrm{E}_{W,t}\left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \frac{c_s^{1-\gamma}}{1-\gamma}\, ds + \varepsilon_2 e^{-\delta(T-t)} \frac{W_T^{1-\gamma}}{1-\gamma}\right].$$

For $\varepsilon_1 = 1$, the marginal utility for consumption is $u'(c) = c^{-\gamma}$ with inverse $I(a) = a^{-1/\gamma}$. Consequently, we have that

$$u(I(a)) = \frac{I(a)^{1-\gamma}}{1-\gamma} = \frac{a^{1-1/\gamma}}{1-\gamma}$$

and

$$u(I(a)) - aI(a) = \frac{a^{1-1/\gamma}}{1-\gamma} - a^{1-1/\gamma} = \frac{\gamma}{1-\gamma} a^{1-1/\gamma}.$$

The first two terms on the right-hand side of Eq. 5.7 are equal to $\frac{\gamma}{1-\gamma} J_W^{1-1/\gamma}$. Without consumption these terms are not present. We can capture both cases by writing these two terms as $\varepsilon_1 \frac{\gamma}{1-\gamma} J_W^{1-1/\gamma}$. Therefore, the HJB equation with or without intermediate consumption implies that

$$\delta J(W,t) = \varepsilon_1 \frac{\gamma}{1-\gamma} J_W(W,t)^{1-\frac{1}{\gamma}} + \frac{\partial J}{\partial t}(W,t) + rW J_W(W,t) - \frac{1}{2}\lambda^\top \lambda \frac{J_W(W,t)^2}{J_{WW}(w,t)}. \tag{5.10}$$

The terminal condition is that $J(W,T) = \varepsilon_2 W^{1-\gamma}/(1-\gamma)$.

Due to the linearity of the wealth dynamics in (5.1) it seems reasonable to conjecture that if the strategy (c^*, π^*) is optimal with time t wealth W and the corresponding wealth process W^*, then the strategy (kc^*, π^*) will be optimal with time t wealth kW and the corresponding wealth process kW^*. If this is true, then

$$\begin{aligned} J(kW,t) &= \mathrm{E}_t\left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \frac{(kc_s^*)^{1-\gamma}}{1-\gamma}\, ds + \varepsilon_2 e^{-\delta(T-t)} \frac{(kW_T^*)^{1-\gamma}}{1-\gamma}\right] \\ &= k^{1-\gamma}\, \mathrm{E}_t\left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \frac{(c_s^*)^{1-\gamma}}{1-\gamma}\, ds + \varepsilon_2 e^{-\delta(T-t)} \frac{(W_T^*)^{1-\gamma}}{1-\gamma}\right] \\ &= k^{1-\gamma} J(W,t), \end{aligned}$$

i.e. the indirect utility function $J(W,t)$ is homogeneous of degree $1-\gamma$ in the wealth W. Inserting $k = 1/W$ and rearranging, we get

$$J(W,t) = \frac{g(t)^\gamma W^{1-\gamma}}{1-\gamma},$$

where $g(t)^\gamma = (1-\gamma)J(1,t)$. From the terminal condition $J(W,T) = \varepsilon_2 W^{1-\gamma}/(1-\gamma)$, we have that $g(T)^\gamma = \varepsilon_2$, hence $g(T) = \varepsilon_2^{1/\gamma} = \varepsilon_2$ for ε_2 equal to zero or one.

The relevant derivatives of our guess $J(W,t)$ are

$$J_W(W,t) = g(t)^\gamma W^{-\gamma}, \quad J_{WW}(W,t) = -\gamma g(t)^\gamma W^{-\gamma-1},$$
$$\frac{\partial J}{\partial t}(W,t) = \frac{\gamma}{1-\gamma} g(t)^{\gamma-1} g'(t) W^{1-\gamma}.$$

Substituting into (5.10) and gathering terms, we get

$$\left\{\left(\frac{\delta}{1-\gamma}-r-\frac{1}{2\gamma}\lambda^\top\lambda\right)g(t)-\frac{\varepsilon_1\gamma}{1-\gamma}-\frac{\gamma}{1-\gamma}g'(t)\right\}g(t)^{\gamma-1}W^{1-\gamma}=0.$$

Since this equation should hold for all W and all $t\in[0,T)$, the term in the brackets must be equal to zero for all t, i.e. the function g must satisfy the ordinary differential equation

$$g'(t)=\frac{1}{\gamma}\left(\delta-r(1-\gamma)-\frac{1-\gamma}{2\gamma}\lambda^\top\lambda\right)g(t)-\varepsilon_1$$

with the terminal condition $g(T)=\varepsilon_2$. It can be checked that the solution is given by

$$g(t)=\frac{1}{A}\left(\varepsilon_1+[\varepsilon_2A-\varepsilon_1]\,e^{-A(T-t)}\right),$$

where A is the constant

$$\begin{aligned}A&=\frac{\delta-r(1-\gamma)}{\gamma}-\frac{1}{2}\frac{1-\gamma}{\gamma^2}\lambda^\top\lambda\\&=\frac{\delta-r(1-\gamma)}{\gamma}-\frac{1}{2}\frac{1-\gamma}{\gamma^2}(\mu-r\mathbf{1})^\top(\sigma\sigma^\top)^{-1}(\mu-r\mathbf{1})\end{aligned}\tag{5.11}$$

We summarize the solution in the following theorem:

Theorem 5.2 *Assume that the constant A defined in* (5.11) *is positive. For the CRRA utility maximization problem in a market with constant r, μ, and σ, we then have that the indirect utility function is given by*

$$J(W,t)=\frac{g(t)^\gamma W^{1-\gamma}}{1-\gamma}\tag{5.12}$$

with

$$g(t)=\frac{1}{A}\left(\varepsilon_1+\left[\varepsilon_2^{1/\gamma}A-\varepsilon_1\right]e^{-A(T-t)}\right).\tag{5.13}$$

The optimal investment strategy is given by

$$\Pi(W,t)=\frac{1}{\gamma}(\sigma^\top)^{-1}\lambda=\frac{1}{\gamma}(\sigma\sigma^\top)^{-1}(\mu-r\mathbf{1}).\tag{5.14}$$

If the agent has utility from intermediate consumption ($\varepsilon_1=1$), her optimal consumption rate is

$$C(W,t)=\frac{1}{g(t)}W=A\left(1+[\varepsilon_2A-1]\,e^{-A(T-t)}\right)^{-1}W.\tag{5.15}$$

A similar result was first demonstrated by Merton (1969). The condition $A>0$ ensures that consumption is positive for any horizon $T>0$. The condition is equivalent to

$$\delta>(1-\gamma)\left(r+\frac{1}{2\gamma}(\mu-r\mathbf{1})^\top(\sigma\sigma^\top)^{-1}(\mu-r\mathbf{1})\right).\tag{5.16}$$

In any reasonable model δ and r will be positive. The above condition will then clearly hold for $\gamma>1$, but for a given time preference rate δ, it may fail for γ very close to zero, i.e. for nearly risk-neutral investors.

The optimal consumption strategy is to consume a time-varying fraction of wealth. It is easy to show that when $\varepsilon_2=1$, the consumption/wealth ratio approaches one as $t\to T$, whereas $c/W\to\infty$ for $t\to T$ when $\varepsilon_2=0$.

The optimal investment strategy consists of keeping the fraction of wealth invested in each asset constant over time. Note that this requires continuous rebalancing of the portfolio since the prices of individual assets vary all the time. Consider an asset which enters the optimal portfolio with a positive weight. If the price of this asset increases more than the other assets in the portfolio, the fraction of wealth made up by that asset will increase. Hence, the investor should reduce the number of units of that particular asset. So the optimal investment strategy is a "sell winners, buy losers" strategy. The higher the risk aversion coefficient γ, the lower the investment in the risky assets and the higher the investment in the riskfree asset. The investment strategy is independent of the horizon of the investor. Under the given assumptions, all CRRA investors will hold different risky assets in the same proportion to each other, i.e. for any $i, j \in \{1, \ldots, d\}$ the ratio π_i/π_j is the same for all investors. This is exactly as in the traditional one-period mean-variance analysis. The optimal strategy is to be further analyzed in an exercise at the end of these notes. Inserting the optimal strategy into the general expression for the dynamics of wealth, we find that

$$dW_t^* = W_t^* \left[\left(r + \frac{1}{\gamma} \lambda^\top \lambda - \varepsilon_1 g(t)^{-1} \right) dt + \frac{1}{\gamma} \lambda^\top \, dz_t \right]. \tag{5.17}$$

Therefore, optimal wealth evolves as a geometric Brownian motion (although with a time-dependent drift). Future values of wealth are lognormally distributed. In particular, wealth stays positive.

For the case where the agent only gets utility from terminal wealth ($\varepsilon_1 = 0, \varepsilon_2 = 1$), the function g reduces to $g(t) = e^{-A(T-t)}$ so that the indirect utility function can be written as

$$J(W, t) = \frac{1}{1-\gamma} e^{-A(T-t)} W^{1-\gamma}.$$

The optimal investment strategy is unaltered. Exactly the same portfolio should be hold whether or not the agent has utility from intermediate consumption. With constant investment opportunities and time-additive CRRA utility there is no clear link between investment and consumption. Of course, wealth will evolve differently over time if the agent withdraws money for consumption. Consequently, *ceteris paribus*, the value of the portfolio and the number of units held of the different assets will be different (smaller) with utility from intermediate consumption.

5.3 Logarithmic utility

The solution for the case of logarithmic utility is obtained by a similar procedure. This is the subject of Exercise 5.2 at the end of the chapter. The indirect utility function is here defined as

$$J(W, t) = \sup_{(c_s, \pi_s)_{s \in [t,T]}} \mathrm{E}_{W,t} \left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \ln c_s \, ds + \varepsilon_2 e^{-\delta(T-t)} \ln W_T \right].$$

The result is:

Theorem 5.3 *For the logarithmic utility maximization problem in a market with constant r, μ, and σ, we have that the indirect utility function is given by*

$$J(W, t) = g(t) \ln W + h(t), \tag{5.18}$$

with

$$g(t) = \frac{\varepsilon_1}{\delta} \left(1 - e^{-\delta(T-t)} \right) + \varepsilon_2 e^{-\delta(T-t)} \tag{5.19}$$

and $h(t)$ is the solution to the ordinary differential equation

$$h'(t) = \delta h(t) + \varepsilon_1 \ln g(t) - \left(r + \frac{1}{2}\lambda^\top \lambda \right) g(t) + \varepsilon_1, \quad h(T) = 0. \tag{5.20}$$

The optimal investment strategy is given by

$$\Pi(W,t) = (\sigma^\top)^{-1}\lambda = (\sigma\sigma^\top)^{-1}(\mu - r\mathbf{1}), \tag{5.21}$$

and if the agent has utility from intermediate consumption ($\varepsilon_1 = 1$) the optimal consumption strategy is

$$C(W,t) = g(t)^{-1}W = \delta \left(1 + [\varepsilon_2\delta - 1]\, e^{-\delta(T-t)} \right)^{-1} W. \tag{5.22}$$

Since $h(t)$ does not affect the optimal strategy we do not care about its precise form. Note that if we take the limit of $g(t)$ defined in Eq. (5.13) as $\gamma \to 1$, we get the expression given in Eq. 5.19. Also note that the optimal strategy for the logarithmic utility case can be obtained by taking limits of the optimal strategy for the CRRA case as $\gamma \to 1$.

5.4 Discussion of the optimal investment strategy

Many empirical studies have documented that in the past century long-term stock investments have in most cases outperformed (i.e. have given a higher return than) a long-term bond investment. Over short investment horizons, the dominance of stock investments is less clear. Referring to these empirical facts, many investment consultants recommend that long-term investors should place a large part of their wealth in stocks and then gradually shift from stocks to bonds as they get older and their investment horizon shrinks. This recommendation conflicts with the optimal portfolio strategy we have derived above. According to our analysis, the optimal portfolio weights of CRRA investors are independent of the investment horizon. Is this because our model of the financial asset prices is inconsistent with the empirical facts mentioned before? The answer is no. To see this let us consider the simplest case with a single stock (representing the stock index) with price dynamics

$$dP_t = P_t \left[\mu\, dt + \sigma\, dz_t \right],$$

where μ and σ as well as the interest rate r are constants. This implies that the probability that a stock investment outperforms a riskless investment over a period of T years is equal to

$$\text{Prob}\left(\frac{P_T}{P_0} > e^{rT} \right) = N\left(\frac{(\mu - r - \sigma^2/2)\sqrt{T}}{\sigma} \right),$$

where $N(\cdot)$ is the cumulative distribution function for a standard normally distributed random variable.

Figure 5.1 illustrates the relation between the outperformance probability and the investment horizon. The curves differ with respect to the presumed expected rate of return on the stock, i.e. μ, whereas the interest rate is 4% and the volatility of the stock is 20% for all curves. Empirical studies indicate that U.S. stocks over a 100-year period have had an average excess rate of return of 8-9% per year. A μ-value of 15% corresponds to an expected excess rate of return of 9% per year since $0.15 - 0.04 - (0.20)^2/2 = 0.09$. However, it should be emphasized that historical estimates of expected rates of return, volatilities, and correlations are not necessarily good predictors of the

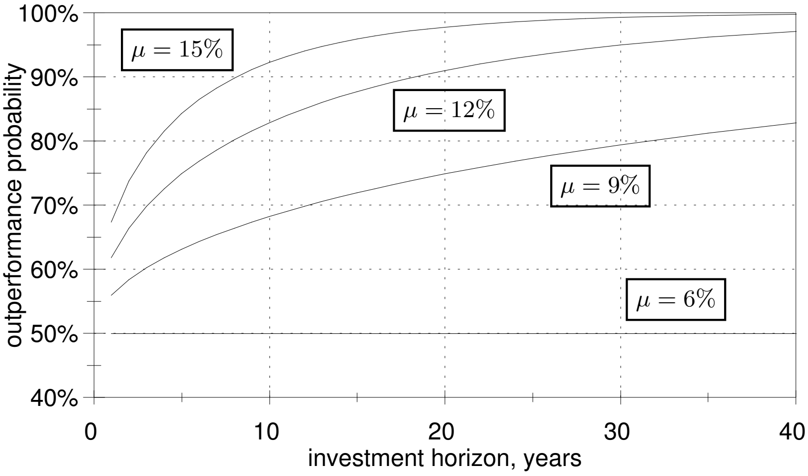


Figure 5.1: Outperformance probabilities. The figure shows the probability that a stock investment outperforms a riskless investment over different investment horizons. For all curves the riskless interest rate is 4%, and the volatility of the stock is 20%. Each of the curves correspond to the value of the parameter μ which is shown besides the curve.

future values of these quantities. In particular, the value of the excess expected rate of return on the stock market is frequently discussed both among practitioners and academics. There are several reasons to believe that the average return on the US stock market over the past century is higher than what the stock market is currently offering in terms of expected returns. This discussion is also closely linked to the so-called equity premium puzzle. See, e.g., Mehra and Prescott (1985), Weil (1989), Welch (2000), and Mehra (2003); Shiller (2000), Campbell and Shiller (2001), Ibbotson and Chen (2003). Probably the curves labeled $\mu = 9\%$ and $\mu = 12\%$ are more representative of the current investment opportunities. In any case, it is tempting to conclude from the graph that long-term investors should invest more in stocks than short-term investors. Why does the optimal portfolio derived previously not reflect this property?

It is important to realize that the optimal decision cannot be based just on the probabilities of gains and losses. After all most individuals will reject a gamble with a 99% probability of winning 1 dollar and a 1% probability of losing a million dollars. The magnitudes of gains and losses are also important for the optimal investment decision. Table 5.1 shows the probability that a stock investment will provide a return which is 0, 25, 50, 75, and 100 percentage points lower than the riskless return over the same period. (The numbers in the row labeled 0% are equal to 100% minus the outperformance probabilities shown in Figure 5.1.) Over a 10-year period the return on a riskless investment at a rate of 4% per year is

$$\left(e^{0.04 \cdot 10} - 1\right) \cdot 100\% \approx 49.1\%.$$

The table shows that with a 22.2% probability a stock investment over a 10-year period will give

Excess return on bond	1 year	10 years	40 years
0%	44.0%	31.8%	17.1%
25%	6.4%	22.2%	16.1%
50%	0.0%	13.1%	15.1%
75%	0.0%	5.7%	14.0%
100%	0.0%	1.3%	13.0%

Table 5.1: Underperformance probabilities. The table shows the probability that a stock investment over a period of 1, 10, and 40 years provides a percentage return which is at least 0, 25, 50, 75, or 100 percentage points lower than the riskless return. The numbers are computed using the parameter values $\mu = 9\%$, $r = 4\%$, and $\sigma = 20\%$.

a return which is lower than $49.1\% - 25\% = 24.1\%$, and there is a 5.7% probability that the stock return will be lower than $49.1\% - 75\% = -25.9\%$. Over a 40-year period the riskless return is 395%. There is a 13% probability that a stock investment will give a return which is at least 100 percentage points lower, i.e. lower than 295%. Over longer periods the probability that stocks underperform bonds is lower, but the probability of extremely bad stock returns is larger than over short periods. The expected excess return on the stock increases with the length of the investment horizon, but so does the variance of the return. Any risk-averse investor has to consider this trade-off. For a CRRA investor in our simple financial model, the two effects offset each other exactly so that the optimal portfolio is independent of the investment horizon.

5.5 Exercises

EXERCISE 5.1 Consider the optimal consumption and investment strategy for a CRRA investor (with no labor income) in a market with constant r, μ, and σ, cf. Theorem 5.2. How does the optimal strategy depend on time and the parameters of the model? (You may assume that only one risky asset is traded.)

EXERCISE 5.2 Give a proof of Theorem 5.3.

EXERCISE 5.3 Assume a financial market with a constant riskfree rate r and risky assets with constant μ and σ. Consider an investor with no income from non-financial sources and an indirect utility function

$$J(W,t) = \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathrm{E}_{W,t}\left[\int_t^T e^{-\delta(s-t)} u(c_s)\, ds\right],$$

where u now is a subsistence HARA function,

$$u(c) = \frac{(c-\bar{c})^{1-\gamma}}{1-\gamma}$$

with $\bar{c}$ being the subsistence level of consumption. What is the optimal consumption and investment strategy for this investor? Compare with the standard CRRA solution.

EXERCISE 5.4 Consider a market with a single consumption good and d risky assets with price dynamics

$$dP_t = \mathrm{diag}(P_t)\,[\mu\, dt + \sigma\, dz_t]\,,$$

where μ is a constant d-vector and σ is a constant $(d\times d)$-matrix. In contrast with the standard model, we will consider the case where no riskfree asset is traded.

We look at an agent with horizon T and a life-time expected utility function

$$\mathrm{E}\left[\int_0^T e^{-\delta s} u(c_s)\, ds + e^{-\delta T} \bar{u}(W_T)\right].$$

The agent must choose a consumption strategy $c = (c_t)_{t \geq 0}$, where we as usual require that $c_t \geq 0$, and an investment strategy $\pi = (\pi_t)_{t \geq 0}$. Here π_t is the d-vector $\pi_t = (\pi_t^1, \dots, \pi_t^d)$ where π_t^i denotes the fraction of wealth that the agent has invested in asset i at time t. Due to the fact that there is no riskfree asset, a feasible strategy must satisfy $\pi_t^\top \mathbf{1} = 1$, i.e. $\sum_{i=1}^d \pi_t^i = 1$.

(a) Write down the dynamics of the agent's wealth for a given consumption and investment strategy (c, π).

Define the indirect utility function

$$J(W, t) = \sup_{(c,\pi)} \mathrm{E}_{W,t}\left[\int_t^T e^{-\delta(s-t)} u(c_s)\, ds + e^{-\delta(T-t)} \bar{u}(W_T)\right],$$

where the supremum is taken over all feasible strategies.

(b) Write the HJB equation corresponding to the problem.

(c) Use the first-order conditions from the HJB equation to express the optimal consumption and investment choice in terms of the unknown indirect utility function J. *Hint: If ν denotes the Lagrange multiplier for the constraint $\pi^\top \mathbf{1} = 1$, you should get that*

$$\pi^* = -\frac{J_W(W,t)}{W J_{WW}(W,t)} (\sigma\sigma^\top)^{-1}\mu - \frac{\nu}{W^2 J_{WW}(W,t)} (\sigma\sigma^\top)^{-1}\mathbf{1}. \qquad (**)$$

(d) Show that there is two-fund separation in this setting with the two mutual funds given by the portfolios

$$\pi^{\tan} = \frac{1}{\mathbf{1}^\top (\sigma\sigma^\top)^{-1}\mu} (\sigma\sigma^\top)^{-1}\mu$$

and

$$\pi^{\mathrm{mv}} = \frac{1}{\mathbf{1}^\top (\sigma\sigma^\top)^{-1}\mathbf{1}} (\sigma\sigma^\top)^{-1}\mathbf{1}.$$

(e) Show that π^{mv} is the minimum-variance portfolio, i.e. the solution to the problem $\min_\pi \pi^\top \sigma\sigma^\top \pi$ s.t. $\pi^\top \mathbf{1} = 1$.

(f) Use (**) and the condition $\mathbf{1}^\top \pi^* = 1$ to determine ν.

In the remainder of this exercise we assume that the agent has CRRA utility, i.e.

$$u(c) = \frac{1}{1-\gamma} c^{1-\gamma}, \qquad \bar{u}(W) = \frac{1}{1-\gamma} W^{1-\gamma},$$

where $\gamma > 0$ and $\gamma \neq 1$.

(g) Show that the indirect utility function is given by $J(W,t) = \frac{1}{1-\gamma} h(t) W^{1-\gamma}$, where

$$h(t) = A^{-\gamma} \left(1 + (A-1)e^{-A[T-t]}\right)^\gamma,$$

and

$$A = \frac{\delta}{\gamma} - \frac{1-\gamma}{2\gamma}\left(\mu^\top (\sigma\sigma^\top)^{-1}\mu - \gamma k^2 \mathbf{1}^\top (\sigma\sigma^\top)^{-1}\mathbf{1}\right),$$

$$k = \frac{1}{\mathbf{1}^\top (\sigma\sigma^\top)^{-1}\mathbf{1}} \left(1 - \frac{1}{\gamma}\mathbf{1}^\top (\sigma\sigma^\top)^{-1}\mu\right).$$

Provide explicit expressions for the optimal consumption and investment in terms of wealth and time. Show how the wealth of the investor will evolve if she follows the optimal consumption and investment strategy.

(h) Compare the optimal consumption and investment strategy found above with the optimal strategy for the case where a riskfree asset is traded.

Chapter 6

Asset allocation with stochastic investment opportunities: the general case

In the previous chapter we analyzed the optimal investment/consumption decision under the assumption of constant investment opportunities, i.e. constant interest rates, expected rates of return, volatilities, and correlations. However, it is well-documented that some, if not all, of these quantities vary over time in a stochastic manner. This situation is referred to as a **stochastic investment opportunity set**. In this chapter we will study the dynamic investment/consumption choice in a general financial market with stochastic investment opportunities. In the next chapter we will then focus on concrete models in which, for example, interest rates or expected excess stock returns follow some specific dynamics.

The main effect of allowing investment opportunities to vary over time is easy to explain. Risk-averse investors with time-additive utility are reluctant to substitute consumption over time, as discussed in Section 2.4. To keep consumption stable across states and time, a (sufficiently) risk-averse investor will therefore choose a portfolio with high positive returns in states with relatively bad future investment opportunities (or bad future labor income) and conversely. This is what is known as **intertemporal hedging**. The optimal investment strategy will thus be different than in the case with constant investment opportunities. From this argument, we also see that there will be a close link between the optimal consumption strategy and the intertemporal hedging part of the optimal investment strategy.

In the rest of this chapter we will formalize these issues in a general modeling framework. We will continue to assume that the investor receives no non-financial income, i.e. no labor income, and refer to Chapter 8 for the extension to the case with labor income. In Section 6.7 we will attack the problem with the martingale approach, but in the other sections we shall apply the dynamic programming approach, i.e. we focus on solving the Hamilton-Jacobi-Bellman equation associated with the utility maximization problem.

6.1 One-dimensional state variable

As in Section 4.5 we assume that there is a stochastically evolving state variable $x = (x_t)$ that captures the variations in r, μ, and σ over time. The variations in the state variable x determine the future expected returns and covariance structure in the financial market. For simplicity we will in this section consider the case where x is one-dimensional and then turn to the multi-dimensional case in the following section.

6.1.1 General utility functions

The dynamics of the d risky asset prices is in this setting given by

$$\begin{aligned} dP_t &= \operatorname{diag}(P_t)\left[\mu(x_t,t)\,dt + \sigma(x_t,t)\,dz_t\right] \\ &= \operatorname{diag}(P_t)\left[\left(r(x_t)\mathbf{1} + \sigma(x_t,t)\lambda(x_t)\right)\,dt + \sigma(x_t,t)\,dz_t\right]. \end{aligned} \tag{6.1}$$

We assume that x follows a one-dimensional diffusion process

$$dx_t = m(x_t)\,dt + v(x_t)^\top\,dz_t + \hat{v}(x_t)\,d\hat{z}_t, \tag{6.2}$$

where $\hat{z}$ is a one-dimensional standard Brownian motion independent of z. If $\hat{v}(x_t) \neq 0$, the market is incomplete; otherwise, it is complete. The wealth evolves as

$$dW_t = W_t\left[r(x_t) + \pi_t^\top \sigma(x_t,t)\lambda(x_t)\right]\,dt - c_t\,dt + W_t\pi_t^\top\sigma(x_t,t)\,dz_t,$$

and the indirect utility function is defined by

$$J(W,x,t) = \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathrm{E}_{W,x,t}\left[\int_t^T e^{-\delta(s-t)}u(c_s)\,ds + e^{-\delta(T-t)}\bar{u}(W_T)\right].$$

The HJB equation associated with this problem is

$$\begin{aligned} \delta J(W,x,t) = \sup_{c\geq 0,\pi\in\mathbb{R}^d} \Big\{ & u(c) + \frac{\partial J}{\partial t}(W,x,t) + J_W(W,x,t)\left(W\left[r(x) + \pi^\top\sigma(x,t)\lambda(x)\right] - c_t\right) \\ & + \frac{1}{2}J_{WW}(W,x,t)W^2\pi^\top\sigma(x,t)\sigma(x,t)^\top\pi + J_x(W,x,t)m(x) \\ & + \frac{1}{2}J_{xx}(W,x,t)(v(x)^\top v(x) + \hat{v}(x)^2) \\ & + J_{Wx}(W,x,t)W\pi^\top\sigma(x,t)v(x)\Big\} \end{aligned} \tag{6.3}$$

with the terminal condition $J(W,x,T) = \bar{u}(W)$.

The first order condition with respect to c is

$$u'(c) = J_W(W,x,t)$$

so that the (candidate) optimal consumption strategy is

$$c_t^* = C(W_t^*, x_t, t),$$

where

$$C(W,x,t) = I(J_W(W,x,t)) \tag{6.4}$$

and, as before, $I(\cdot)$ is the inverse of $u'(\cdot)$.

The first order condition with respect to π is different than with constant investment opportunities:

$$J_W(W,x,t)W\sigma(x,t)\lambda(x) + J_{WW}(W,x,t)W^2\sigma(x,t)\sigma(x,t)^\top\pi + J_{Wx}(W,x,t)W\sigma(x,t)v(x) = 0$$

so that the candidate optimal portfolio is

$$\pi_t^* = \Pi(W_t^*, x_t, t),$$

where

$$\Pi(W,x,t) = -\frac{J_W(W,x,t)}{WJ_{WW}(W,x,t)}\left(\sigma(x,t)^\top\right)^{-1}\lambda(x) - \frac{J_{Wx}(W,x,t)}{WJ_{WW}(W,x,t)}\left(\sigma(x,t)^\top\right)^{-1}v(x). \tag{6.5}$$

As the horizon shrinks, the indirect utility function $J(W,x,t)$ approaches the terminal utility function $\bar{u}(W)$ which is independent of the state x. Consequently, the derivative $J_{Wx}(W,x,t)$ and hence the last term of the portfolio will approach zero as $t \to T$. In other words, very short-term investors do not hedge. The last term will also disappear for "non-instantaneous" investors in two special cases:

(1) $J_{Wx}(W,x,t) \equiv 0$: The state variable does not affect the marginal utility of the investor. As we shall see below this is always true for investors with logarithmic utility. Such an investor is *not interested in hedging* changes in the state variable.

(2) $v(x) \equiv 0$: The state variable is uncorrelated with instantaneous returns on the traded assets. In this case the investor is *not able to hedge* changes in the state variable.

In all other cases the state variable induces an additional term to the optimal portfolio relative to the case of constant investment opportunities. We now have the following important result:

Theorem 6.1 (Three-fund separation) *All investors will combine (1) the locally riskfree asset ("the bank account"), (2) the tangency portfolio given by the weights*

$$\pi_t^{tan} = \frac{1}{\mathbf{1}^\top\left(\sigma(x,t)^\top\right)^{-1}\lambda(x)}\left(\sigma(x,t)^\top\right)^{-1}\lambda(x),$$

and (3) the hedge portfolio given by the weights

$$\pi_t^{hdg} = \frac{1}{\mathbf{1}^\top\left(\sigma(x,t)^\top\right)^{-1}v(x)}\left(\sigma(x,t)^\top\right)^{-1}v(x).$$

Note that the composition of the two risky funds varies over time due to fluctuations in the state variable. It is no longer true that all investors will hold different risky assets in the same proportion, i.e. the fractions π_i/π_j will be investor-specific since different investors may put different weights on the two portfolios of risky assets. The tangency portfolio has the same interpretation as previously. The position in the portfolio π^{hdg} is the change in the optimal investment strategy due to the stochastic variations in the investment opportunity set, hence the name "hedge portfolio". It can be shown that the portfolio π^{hdg} is the portfolio with the maximal absolute correlation with the state variable. In that sense it is the portfolio that is best at hedging changes in the state variable. In a complete market the maximal correlation is one and the hedge portfolio basically replicated the dynamics of the state variable.

Let us focus for a moment on the case with a single risky asset so that both $\sigma(x,t)$ and $v(x)$ are scalars. The hedge term in π_t^* can then be written as $-J_{Wx}/[W J_{WW}] \cdot v/\sigma$. Note that $J_{WW} < 0$ by concavity. If v and σ have the same sign, then the return of the risky asset will be positively correlated with changes in the state variable. In this case we see that the hedge demand on the asset is positive if marginal utility J_W is increasing in x so that $J_{Wx} > 0$. This makes good sense: relative to the situation with a constant investment opportunity set, the agent will devote a larger fraction of wealth to a risky asset that has a high return in states of the world where marginal utility is high. Conversely, if v and σ have opposite signs so that they are negatively correlated.

Here is another interpretation of the portfolio strategy (following Ingersoll (1987, p. 282)):

Theorem 6.2 *The optimal portfolio strategy π^* is the one that minimizes fluctuations in consumption over time among all portfolio strategies with the same expected rate of return as π^*.*

Proof: The expected rate of return on the optimal portfolio in (6.5) is

$$\mu^*(x,t) = r(x) + (\pi_t^*)^\top (\mu(x,t) - r(x)\mathbf{1}).$$

The consumption rate is given by

$$c_t^* = C(W_t, x_t, t).$$

An application of Itô's Lemma yield

$$\begin{aligned} dc_t^* = \ldots\, dt + \big(& C_W(W_t, x_t, t) W_t \pi_t^\top \sigma(x_t, t) \\ & + C_x(W_t, x_t, t) v(x_t)^\top \big)\, dz_t + C_x(W_t, x_t, t) \hat{v}(x_t)\, d\hat{z}_t, \end{aligned}$$

where we leave the drift term unspecified and the subscripts on C denote partial derivatives. It follows that the instantaneous variance rate of consumption is equal to

$$\begin{aligned} \sigma_c^2 \equiv\; & C_W(W,x,t)^2 W^2 \pi^\top \sigma(x,t)\sigma(x,t)^\top \pi + C_x(W,x,t)^2 \left(v(x)^\top v(x) + \hat{v}(x)^2 \right) \\ & + 2 C_W(W,x,t) C_x(W,x,t) W \pi^\top \sigma(x,t) v(x). \end{aligned}$$

Now consider the problem of minimizing σ_c^2 over all portfolios π that have an expected rate of return equal to $\mu^*(x,t)$, i.e. portfolios π with $r(x) + \pi^\top \sigma(x,t)\lambda(x) = \mu^*(x,t)$. Forming the Lagrangian

$$\mathcal{L} = \sigma_c^2 + \psi \left[\mu^*(x,t) - r(x) - \pi^\top \sigma(x,t)\lambda(x) \right]$$

we find the optimality condition

$$\pi^{**} = \frac{\psi}{2 C_W(W,x,t)^2 W^2} \left(\sigma(x,t)^\top \right)^{-1} \lambda(x) - \frac{C_x(W,x,t)}{W C_W(W,x,t)} \left(\sigma(x,t)^\top \right)^{-1} v(x).$$

Differentiating the envelope condition $u'(C(W,x,t)) = J_W(W,x,t)$ along the optimal consumption path with respect to W we get

$$u''(C(W,x,t)) C_W(W,x,t) = J_{WW}(W,x,t)$$

and by differentiating with respect to x we get

$$u''(C(W,x,t)) C_x(W,x,t) = J_{Wx}(W,x,t).$$

Hence,

$$\frac{C_x(W,x,t)}{WC_W(W,x,t)} = \frac{J_{Wx}(W,x,t)}{WJ_{WW}(W,x,t)}$$

so that the second terms in π^* and π^{**} are identical. The first term in π^{**} is proportional to the first term in π^* and since π^{**} is chosen such that it has the same expected rate of return as π^*, the first terms must also coincide. In total, $\pi^{**} = \pi^*$, which was to be shown. □

On the other hand, if we minimize the instantaneous variance of wealth, i.e. $\sigma_W^2 = \pi^\top \sigma(x,t)\sigma(x,t)^\top \pi$, over all portfolios π having the same expected rate of return as π^*, we get

$$\pi^{**} = \psi \left(\sigma(x,t)^\top\right)^{-1} \lambda(x).$$

This only involves the tangency portfolio. We can conclude that the investor is concerned about fluctuations over time in consumption, not wealth.

Above, we discussed the general expressions for the optimal consumption and investment strategy in the presence of a state variable. But these were expressed in terms of the unknown indirect utility function. How do we proceed to find concrete solutions?

Substituting the candidate optimal values of c and π back into the HJB equation and gathering terms, we get the second order PDE

$$\begin{aligned}\delta J(W,x,t) &= u\left(I(J_W(W,x,t))\right) - J_W(W,x,t)I(J_W(W,x,t)) + \frac{\partial J}{\partial t}(W,x,t) \\ &\quad + r(x)WJ_W(W,x,t) - \frac{1}{2}\frac{J_W(W,x,t)^2}{J_{WW}(W,x,t)}\lambda(x)^\top\lambda(x) \\ &\quad + J_x(W,x,t)m(x) + \frac{1}{2}J_{xx}(W,x,t)\left(v(x)^\top v(x) + \hat{v}(x)^2\right) \\ &\quad - \frac{1}{2}\frac{J_{Wx}(W,x,t)^2}{J_{WW}(W,x,t)}v(x)^\top v(x) - \frac{J_W(W,x,t)J_{Wx}(W,x,t)}{J_{WW}(W,x,t)}\lambda(x)^\top v(x).\end{aligned} \tag{6.6}$$

If this PDE has a solution $J(W,x,t)$ such that the strategy defined by (6.4) and (6.5) is feasible (satisfies the technical conditions), then we know from the verification theorem that this strategy is indeed the optimal consumption and investment strategy and the function $J(W,x,t)$ is indeed the indirect utility function. With no utility from intermediate consumption, i.e. $u \equiv 0$, the first two terms of the right-hand side of (6.6) vanish.

6.1.2 Specific utility functions

Although the PDE (6.6) looks very complicated, closed-form solutions can be found for a number of interesting model specifications. For CRRA utility and logarithmic utility we can derive the general form of the indirect utility function by an argument used earlier. Kogan and Uppal (2000) apply a so-called perturbation analysis to study the differences between the optimal strategies for a non-log CRRA investor ($\gamma \neq 1$) and a log investor ($\gamma = 1$).

CRRA utility

Consider the indirect utility function with CRRA utility:

$$J(W,x,t) = \sup_{(c_s,\pi_s)_{s\in[t,T]}} \mathrm{E}_{W,x,t}\left[\varepsilon_1 \int_t^T e^{-\delta(s-t)}\frac{c_s^{1-\gamma}}{1-\gamma}\,ds + \varepsilon_2 e^{-\delta(T-t)}\frac{W_T^{1-\gamma}}{1-\gamma}\right],$$

where ε_1 and ε_2 are either zero or one. We set up a conjecture for the form of J using the same arguments as we did in the case of constant investment opportunities. Due to the linearity of the wealth dynamics it seems reasonable to guess that if the strategy (c^*, π^*) is optimal with time t wealth W and state x and the corresponding wealth process W^*, then the strategy (kc^*, π^*) will be optimal with time t wealth kW and state x and the corresponding wealth process kW^*. If this is true, then

$$\begin{aligned} J(kW, x, t) &= \mathrm{E}_t \left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \frac{(kc_s^*)^{1-\gamma}}{1-\gamma}\, ds + \varepsilon_2 e^{-\delta(T-t)} \frac{(kW_T^*)^{1-\gamma}}{1-\gamma} \right] \\ &= k^{1-\gamma}\, \mathrm{E}_t \left[\varepsilon_1 \int_t^T e^{-\delta(s-t)} \frac{(c_s^*)^{1-\gamma}}{1-\gamma}\, ds + \varepsilon_2 e^{-\delta(T-t)} \frac{(W_T^*)^{1-\gamma}}{1-\gamma} \right] \\ &= k^{1-\gamma} J(W, x, t), \end{aligned}$$

i.e. the indirect utility function is homogeneous of degree $1-\gamma$ in the wealth level. Inserting $k = 1/W$ and rearranging, we get

$$J(W, x, t) = \frac{g(x,t)^\gamma W^{1-\gamma}}{1-\gamma}, \tag{6.7}$$

where $g(x,t)^\gamma = (1-\gamma)J(1,x,t)$. From the terminal condition $J(W,x,T) = \varepsilon_2 W^{1-\gamma}/(1-\gamma)$, we have that $g(x,T)^\gamma = \varepsilon_2$, which is equivalent to $g(x,T) = \varepsilon_2$ when ε_2 is either zero or one.

The relevant derivatives of J are

$$\begin{aligned} J_W(W,x,t) &= g(x,t)^\gamma W^{-\gamma}, \\ J_{WW}(W,x,t) &= -\gamma g(x,t)^\gamma W^{-\gamma-1}, \\ J_x(W,x,t) &= \frac{\gamma}{1-\gamma} g(x,t)^{\gamma-1} g_x(x,t) W^{1-\gamma}, \\ J_{xx}(W,x,t) &= -\gamma g(x,t)^{\gamma-2} g_x(x,t)^2 W^{1-\gamma} + \frac{\gamma}{1-\gamma} g(x,t)^{\gamma-1} g_{xx}(x,t) W^{1-\gamma}, \\ J_{Wx}(W,x,t) &= \gamma g(x,t)^{\gamma-1} g_x(x,t) W^{-\gamma}, \\ \frac{\partial J}{\partial t}(W,x,t) &= \frac{\gamma}{1-\gamma} g(x,t)^{\gamma-1} \frac{\partial g}{\partial t}(x,t) W^{1-\gamma}. \end{aligned}$$

The optimal investment strategy becomes

$$\Pi(W,x,t) = \frac{1}{\gamma} \left(\sigma(x,t)^\top\right)^{-1} \lambda(x) + \frac{g_x(x,t)}{g(x,t)} \left(\sigma(x,t)^\top\right)^{-1} v(x), \tag{6.8}$$

and the optimal consumption strategy is

$$C(W,x,t) = \varepsilon_1 \frac{W}{g(x,t)}, \tag{6.9}$$

which, of course, is zero if the investor obtains no utility from intermediate consumption. It is optimal to consume a time- and state-dependent fraction of wealth. The optimal fractions of wealth allocated to the various risky assets are independent of the level of wealth, but depend on the state and time.

Inserting the derivatives above into (6.6) and simplifying, we get that $g(x,t)$ must solve the

PDE

$$
\begin{aligned}
0 = \varepsilon_1 &- \left(\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma} r(x) - \frac{1-\gamma}{2\gamma^2} \lambda(x)^\top \lambda(x) \right) g(x,t) \\
&+ \frac{\partial g}{\partial t}(x,t) + \left(m(x) + \frac{1-\gamma}{\gamma} \lambda(x)^\top v(x) \right) g_x(x,t) \\
&+ \frac{1}{2} g_{xx}(x,t) \left(v(x)^\top v(x) + \hat{v}(x)^2 \right) - \frac{1}{2}(1-\gamma)\hat{v}(x)^2 \frac{g_x(x,t)^2}{g(x,t)}
\end{aligned}
\tag{6.10}
$$

with the terminal condition $g(x,T) = \varepsilon_2$. This PDE has a nice solution in a large class of interesting models – see Sections 6.4 and 6.5 below.

Logarithmic utility

Applying the same procedure on the problem with log utility, we get

$$J(W,x,t) = f(t) \ln W + h(x,t)$$

where $h(x,t)$ must satisfy a certain PDE. Since the cross derivative $J_{Wx}(W,x,t) = 0$, the optimal risky portfolio reduces to

$$\Pi(W,x,t) = \left(\sigma(x,t)^\top\right)^{-1} \lambda(x). \tag{6.11}$$

We can conclude that **a logarithmic investor does not hedge stochastic variations in the investment opportunity set**. She behaves myopically, i.e. as in a static one-period framework. Optimal consumption is again given by

$$C(W,x,t) = \frac{W}{f(t)}. \tag{6.12}$$

Letting $\Pi_0(W,x,t)$ denote the fraction of wealth optimally invested in the instantaneously riskless asset, we can summarize the entire investment strategy as

$$\begin{pmatrix} \Pi_0(W,x,t) \\ \Pi(w,x,t) \end{pmatrix} = \begin{pmatrix} 1 - \mathbf{1}^\top \left(\sigma(x,t)^\top\right)^{-1} \lambda(x) \\ \left(\sigma(x,t)^\top\right)^{-1} \lambda(x) \end{pmatrix} \tag{6.13}$$

This portfolio is sometimes referred to as the **log portfolio** or the **growth-optimal portfolio**, since it is also the portfolio with the highest expected average compound growth rate of portfolio value which is defined as $\frac{1}{T-t} \ln\left(W_T/W_t\right)$.

6.2 Multi-dimensional state variable

Suppose now that the state variable x is k-dimensional and follows the diffusion process

$$dx_t = m(x_t)\,dt + v(x_t)^\top\,dz_t + \hat{v}(x_t)\,d\hat{z}_t, \tag{6.14}$$

where m now is a k-vector valued function, v is a $(d \times k)$-matrix valued function, $\hat{v}$ is a $(k \times k)$-matrix valued function, and $\hat{z}$ is a k-dimensional standard Brownian motion independent of z. The HJB equation becomes

$$
\begin{aligned}
\delta J(W,x,t) = \sup_{c \geq 0, \pi \in \mathbb{R}^d} \Big\{ & u(c) + \frac{\partial J}{\partial t}(W,x,t) + J_W(W,x,t)\left(W\left[r(x) + \pi^\top \sigma(x,t)\lambda(x)\right] - c\right) \\
&+ \frac{1}{2} J_{WW}(W,x,t) W^2 \pi^\top \sigma(x,t)\sigma(x,t)^\top \pi + J_x(W,x,t)^\top m(x) \\
&+ \frac{1}{2} \operatorname{tr}\left(J_{xx}(W,x,t)[v(x)^\top v(x) + \hat{v}(x)\hat{v}(x)^\top]\right) \\
&+ W\pi^\top \sigma(x,t) v(x) J_{Wx}(W,x,t) \Big\}.
\end{aligned}
\tag{6.15}
$$

This multi-dimensional setting was briefly introduced in Section 4.5.

Again, the (candidate) optimal consumption strategy is

$$c_t^* = C(W_t^*, x_t, t),$$

where

$$C(W, x, t) = I(J_W(W, x, t)).$$

The first order condition with respect to π implies that the (candidate) optimal portfolio is

$$\pi^* = \Pi(W_t^*, x_t, t),$$

where

$$\Pi(W, x, t) = -\frac{J_W(W, x, t)}{W J_{WW}(W, x, t)} \left(\sigma(x, t)^\top\right)^{-1} \lambda(x) - \left(\sigma(x, t)^\top\right)^{-1} v(x) \frac{J_{Wx}(W, x, t)}{W J_{WW}(W, x, t)}.$$

We can split up the last term into k terms, one for each element of the state variable:

$$\begin{aligned}\Pi(W, x, t) = &-\frac{J_W(W, x, t)}{W J_{WW}(W, x, t)} \left(\sigma(x, t)^\top\right)^{-1} \lambda(x) \\ &- \sum_{j=1}^{k} \left(\sigma(x, t)^\top\right)^{-1} \begin{pmatrix} v_{1j}(x) \\ v_{2j}(x) \\ \vdots \\ v_{dj}(x) \end{pmatrix} \frac{J_{Wx_j}(W, x, t)}{W J_{WW}(W, x, t)}.\end{aligned} \tag{6.16}$$

Each of the term in the sum has the interpretation as a fund hedging changes in one element of the state variable. Therefore, we have $(k+2)$**-fund separation**: all investors are satisfied with access to trade in the riskfree asset, the tangency portfolio, and k hedge funds.

Substituting the candidate optimal values of c and π back into the HJB equation and gathering terms, we get the second-order PDE

$$\begin{aligned}\delta J(W, x, t) = &\, u\left(I(J_W(W, x, t))\right) - J_W(W, x, t) I(J_W(W, x, t)) + \frac{\partial J}{\partial t}(W, x, t) \\ &+ r(x) W J_W(W, x, t) - \frac{1}{2} \frac{J_W(W, x, t)^2}{J_{WW}(W, x, t)} \lambda(x)^\top \lambda(x) \\ &+ J_x(W, x, t)^\top m(x) + \frac{1}{2} \operatorname{tr}\left(J_{xx}(W, x, t)[v(x)^\top v(x) + \hat{v}(x)\hat{v}(x)^\top]\right) \\ &- \frac{1}{2 J_{WW}(W, x, t)} J_{Wx}(W, x, t)^\top v(x) v(x)^\top J_{Wx}(W, x, t) \\ &- \lambda(x)^\top v(x) \frac{J_W(W, x, t) J_{Wx}(W, x, t)}{J_{WW}(W, x, t)}.\end{aligned} \tag{6.17}$$

Again the first two terms on the right-hand side are not present when the agent has no utility from intermediate consumption.

With CRRA utility, a qualified guess on the solution is

$$J(W, x, t) = \frac{g(x, t)^\gamma W^{1-\gamma}}{1-\gamma}, \tag{6.18}$$

which indeed is a solution to the HJB equation if the function $g(x, t)$ solves the PDE

$$\begin{aligned}0 = &\, \varepsilon_1 - \left(\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma} r(x) - \frac{1-\gamma}{2\gamma^2} \lambda(x)^\top \lambda(x)\right) g(x, t) + \frac{\partial g}{\partial t}(x, t) \\ &+ \left(m(x) + \frac{1-\gamma}{\gamma} v(x)^\top \lambda(x)\right)^\top g_x(x, t) + \frac{1}{2} \operatorname{tr}\left(g_{xx}(x, t)\left[v(x)^\top v(x) + \hat{v}(x)\hat{v}(x)^\top\right]\right) \\ &- \frac{1}{2}(1-\gamma) g(x, t)^{-1} g_x(x, t)^\top \hat{v}(x) \hat{v}(x)^\top g_x(x, t)\end{aligned} \tag{6.19}$$

with the terminal condition $g(x, T) = \varepsilon_2$. The optimal investment strategy is

$$\Pi(W, x, t) = \frac{1}{\gamma} \left(\sigma(x, t)^\top\right)^{-1} \lambda(x) + \frac{1}{g(x, t)} \left(\sigma(x, t)^\top\right)^{-1} v(x) g_x(x, t), \tag{6.20}$$

and with intermediate consumption the optimal consumption rate is given by

$$C(W, x, t) = \frac{W}{g(x, t)}. \tag{6.21}$$

6.3 What risks are to be hedged?

It may appear from the analysis above that investors would want to hedge all variables affecting r_t, μ_t, and σ_t, but this is actually not so. We will show that the only risks the agent will want to hedge are those affecting r_t and λ_t.

We will think of the investor choosing the "volatility vector of wealth" $\varphi_t = \sigma_t^\top \pi_t$ directly rather than π_t. In these terms wealth evolves as

$$dW_t = W_t \left[r_t + \varphi_t^\top \lambda_t\right] dt - c_t \, dt + W_t \varphi_t^\top \, dz_t.$$

The indirect utility function is

$$J_t = \sup_{(c,\varphi)} \mathrm{E}_t \left[\int_t^T e^{-\delta(s-t)} u(c_s) \, ds + e^{-\delta(T-t)} \bar{u}(W_T) \right].$$

Note that this optimization problem does not involve μ_t or σ_t. Assuming now that there is a variable x_t so that

$$r_t = r(x_t), \quad \lambda_t = \lambda(x_t),$$

then $J_t = J(W_t, x_t, t)$ and we can use the dynamic programming approach.

For the multidimensional x we will get the optimal wealth volatility vector

$$\varphi_t = -\frac{J_W(W, x, t)}{W J_{WW}(W, x, t)} \lambda(x_t) - v(x) \frac{J_{Wx}(W, x, t)}{W J_{WW}(W, x, t)}. \tag{6.22}$$

Hence, the optimal portfolio strategy is

$$\pi_t = -\frac{J_W(W, x, t)}{W J_{WW}(W, x, t)} \left(\sigma_t^\top\right)^{-1} \lambda(x_t) - \left(\sigma_t^\top\right)^{-1} v(x) \frac{J_{Wx}(W, x, t)}{W J_{WW}(W, x, t)}. \tag{6.23}$$

We can conclude from this analysis that the investor will only hedge the variables that affect the short-term interest rate and the market prices of risk (this is of course only true within the present framework; e.g. an investor with stochastic income will also want to hedge the income risk). Stochastic variations in μ_t and σ_t are only interesting to the extent that they cause stochastic variations in the market price of risk! One could imagine a market where volatilities vary stochastically but expected rates of return follow the variations in volatilities so that the market price of risk is constant over time. In such a market no agent would hedge the variations in volatilities and expected rates of return. Similar observations were made by Detemple, Garcia, and Rindisbacher (2003) and Munk and Sørensen (2003). The volatility matrix σ_t of the risky assets becomes relevant when the agent wants to find a portfolio π_t that will generate the desired wealth volatility vector φ_t.

In fact, the statement above can be strengthened slightly. Look at the PDE (6.19). Suppose that both r and $\lambda^\top \lambda$ and, hence, the entire coefficient of $g(x,t)$ are independent of x. Then the function $g(t)$ that satisfies the ordinary differential equation

$$0 = \varepsilon_1 - \left(\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma} r(t) - \frac{1-\gamma}{2\gamma^2} \lambda(t)^\top \lambda(t) \right) g(t) + g'(t)$$

and $g(T) = \varepsilon_2^{1/\gamma}$ will also satisfy the PDE (6.19). So in this case the solution $g(x,t)$ to (6.19) is independent of x. Consequently, the hedge term in (6.20) disappears. In other words, the investor will only hedge stochastic variations that affect the short-term interest rate r_t and the squared market prices of risk[1]

$$\lambda_t^\top \lambda_t = (\mu_t - r_t \mathbf{1})^\top \left(\sigma_t \sigma_t^\top\right)^{-1} (\mu_t - r_t \mathbf{1}).$$

Nielsen and Vassalou (2000) show that this result is also true for non-Markov dynamics of prices and non-CRRA preferences of terminal wealth (and, consequently, it also holds for time-additive non-CRRA utility of consumption). We summarize this in the following theorem:

Theorem 6.3 *Investors with time-additive utility functions will only hedge stochastic variations in the short-term interest rate r_t and in the squared market prices of risk $\lambda_t^\top \lambda_t$.*

There is a very intuitive interpretation of this result, which we can see after a few computations: The tangency portfolio is in general given by [see (5.9)]

$$\pi_t^{\tan} = \frac{1}{\mathbf{1}^\top \left(\sigma_t^\top\right)^{-1} \lambda_t} \left(\sigma_t^\top\right)^{-1} \lambda_t. \tag{6.24}$$

The expected excess rate of return on the tangency portfolio is

$$\left(\pi_t^{\tan}\right)^\top (\mu_t - r_t \mathbf{1}) = \frac{1}{\mathbf{1}^\top \left(\sigma_t^\top\right)^{-1} \lambda_t} \lambda_t^\top \lambda_t.$$

The volatility (instantaneous standard deviation) of the tangency portfolio is

$$\sqrt{\left(\pi_t^{\tan}\right)^\top \sigma_t \sigma_t^\top \pi_t^{\tan}} = \frac{1}{\mathbf{1}^\top \left(\sigma_t^\top\right)^{-1} \lambda_t} \sqrt{\lambda_t^\top \lambda_t}.$$

The slope of the instantaneous capital market line is therefore equal to $\sqrt{\lambda_t^\top \lambda_t}$. (In a setting with a single risky asset, $\lambda_t = (\mu_t - r_t)/\sigma_t$ and $\sqrt{\lambda_t^\top \lambda_t} = \sqrt{\lambda_t^2} = \lambda_t$.) In a static framework the optimal portfolio is determined by the position of the capital market line, i.e. (1) the intercept which is equal to the riskfree rate of return and (2) the slope which is the Sharpe ratio of the tangency portfolio. It is therefore natural that investors in a dynamic framework only are concerned about the variations in these two variables.

6.4 Closed-form solution for CRRA utility: affine models with one state variable

In this and the following subsection we will look at models in which the optimal portfolio and consumption strategies of a CRRA investor can be derived in closed-form. In some of these cases

[1]Examples where $\lambda^\top \lambda$ is constant, but λ itself is not, can be given [see Nielsen and Vassalou (2000)], but seem rather contrived.

we can obtain explicit solutions, in other cases the solution involves time-dependent functions that can be found by numerically solving ordinary differential equations. Many of our concrete examples in the following chapters are special cases of these models. In this section we will discuss so-called affine models, while the next section focuses on the so-called quadratic models. The results presented are similar to those obtained by Liu (1999). For notational simplicity we shall assume that the state variable is one-dimensional with dynamics given by (6.2). We will briefly discuss solutions to problems with a multi-dimensional state variable in Section 6.6.

6.4.1 Utility from terminal wealth only

Let us first consider the case with utility from terminal wealth only, i.e. $\varepsilon_1 = 0$ and $\varepsilon_2 = 1$. In that case, the PDE (6.10) that we have to solve reduces slightly to

$$\begin{aligned} 0 = &-\left(\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma} r(x) - \frac{1-\gamma}{2\gamma^2}\lambda(x)^\top\lambda(x)\right) g(x,t) \\ &+ \frac{\partial g}{\partial t}(x,t) + \left(m(x) + \frac{1-\gamma}{\gamma} v(x)^\top \lambda(x)\right) g_x(x,t) \\ &+ \frac{1}{2} g_{xx}(x,t)\left(v(x)^\top v(x) + \hat{v}(x)^2\right) - \frac{1}{2}(1-\gamma)\hat{v}(x)^2 \frac{g_x(x,t)^2}{g(x,t)} \end{aligned} \tag{6.25}$$

with terminal condition $g(x,T) = 1$.

Let us consider when a solution of the form

$$g(x,t) = e^{-\frac{\delta}{\gamma}(T-t) + A_1(T-t) + A_2(T-t)x}$$

will work, where A_1 and A_2 are real-valued deterministic functions of time. From the terminal condition we must have that $A_1(0) = A_2(0) = 0$. The relevant derivatives of g can be written as

$$\begin{aligned} g_x(x,t) &= A_2(T-t) g(x,t), \quad g_{xx}(x,t) = A_2(T-t)^2 g(x,t), \\ \frac{\partial g}{\partial t}(x,t) &= \left(\frac{\delta}{\gamma} - A_1'(T-t) - A_2'(T-t)x\right) g(x,t). \end{aligned}$$

Inserting these into (6.25) and dividing by $g(x,t)$, we get

$$\begin{aligned} 0 = &\frac{1-\gamma}{\gamma} r(x) + \frac{1-\gamma}{2\gamma^2}\lambda(x)^\top\lambda(x) - A_1'(T-t) - A_2'(T-t)x \\ &+ \left(m(x) + \frac{1-\gamma}{\gamma} v(x)^\top\lambda(x)\right) A_2(T-t) + \frac{1}{2}\left(v(x)^\top v(x) + \gamma\hat{v}(x)^2\right) A_2(T-t)^2. \end{aligned} \tag{6.26}$$

If $r(x)$, $\lambda(x)^\top\lambda(x)$, $m(x)$, $v(x)^\top\lambda(x)$, $v(x)^\top v(x)$, and $\hat{v}(x)^2$ are all affine[2] functions of x, then we can find two ordinary differential equations for A_1 and A_2. In order to see this, suppose that

$$r(x) = r_0 + r_1 x, \tag{6.27}$$

$$m(x) = m_0 + m_1 x, \tag{6.28}$$

$$\hat{v}(x) = \sqrt{\hat{v}_0 + \hat{v}_1 x} \tag{6.29}$$

[2] A real-valued function is said to be an affine function of the k-vector x, if it can be written as $a_1 + a_2^\top x$, where a_1 is a constant scalar and a_2 is a constant k-vector (possibly zero so that a constant is also included in the set of affine functions). A vector- or matrix-valued function is said to be affine if all its elements are affine.

for some constants $r_0, r_1, m_0, m_1, \hat{v}_0$, and $\hat{v}_1$. Of course, we should have that $\hat{v}_0+\hat{v}_1 x \geq 0$ for all possible values of x, which is easily satisfied if either $\hat{v}_0$ or $\hat{v}_1$ are zero and the other parameter is positive. The term $\lambda(x)^\top \lambda(x)$ will be affine in x if each element of the vector $\lambda(x) = (\lambda_1(x), \dots, \lambda_d(x))^\top$ is of the form $\lambda_i(x) = \sqrt{\lambda_{i0} + \lambda_{i1} x}$ since then

$$\lambda(x)^\top \lambda(x) = \sum_{i=1}^d \lambda_i(x)^2 = \sum_{i=1}^d (\lambda_{i0} + \lambda_{i1} x) = \left(\sum_{i=1}^d \lambda_{i0} \right) + \left(\sum_{i=1}^d \lambda_{i1} \right) x \equiv \Lambda_0 + \Lambda_1 x. \tag{6.30}$$

Similarly, the term $v(x)^\top v(x)$ will be affine in x if each element of the vector $v(x) = (v_1(x), \dots, v_d(x))^\top$ is of the form $v_i(x) = \sqrt{v_{i0} + v_{i1} x}$. Then we have

$$v(x)^\top v(x) = \sum_{i=1}^d v_i(x)^2 = \sum_{i=1}^d (v_{i0} + v_{i1} x) = \left(\sum_{i=1}^d v_{i0} \right) + \left(\sum_{i=1}^d v_{i1} \right) x \equiv V_0 + V_1 x. \tag{6.31}$$

In addition, we must have that $v(x)^\top \lambda(x)$ is affine in x. With the specifications of $\lambda(x)$ and $v(x)$ just given, we have

$$v(x)^\top \lambda(x) = \sum_{i=1}^d v_i(x) \lambda_i(x) = \sum_{i=1}^d \sqrt{(v_{i0} + v_{i1} x)(\lambda_{i0} + \lambda_{i1} x)}.$$

This will only be affine in x if for each i we have either $v_{i0} = \lambda_{i0} = 0$ or $v_{i1} = \lambda_{i1} = 0$. To encompass all possible situations let us write

$$v(x)^\top \lambda(x) = K_0 + K_1 x, \tag{6.32}$$

where K_0 and K_1 are real-valued parameters. If we substitute (6.27)–(6.32) into (6.26) and use the fact that (6.26) must hold for all values of x and all t, we obtain a system of two ordinary differential equations for A_1 and A_2:

$$0 = \frac{1-\gamma}{\gamma} r_0 + \frac{1-\gamma}{2\gamma^2} \Lambda_0 - A_1'(\tau) + \left(m_0 + \frac{1-\gamma}{\gamma} K_0 \right) A_2(\tau) + \frac{1}{2} \left(V_0 + \gamma \hat{v}_0 \right) A_2(\tau)^2, \tag{6.33}$$

$$0 = \frac{1-\gamma}{\gamma} r_1 + \frac{1-\gamma}{2\gamma^2} \Lambda_1 - A_2'(\tau) + \left(m_1 + \frac{1-\gamma}{\gamma} K_1 \right) A_2(\tau) + \frac{1}{2} \left(V_1 + \gamma \hat{v}_1 \right) A_2(\tau)^2. \tag{6.34}$$

These equations are to be solved with the initial conditions $A_1(0) = A_2(0) = 0$. First (6.34) is solved for $A_2(\tau)$. Since $A_1(\tau) = A_1(\tau) - A_1(0) = \int_0^\tau A_1'(s)\, ds$, we can afterwards compute $A_1(\tau)$ as

$$\begin{aligned} A_1(\tau) = &\left(\frac{1-\gamma}{\gamma} r_0 + \frac{1-\gamma}{2\gamma^2} \Lambda_0 \right) \tau + \left(m_0 + \frac{1-\gamma}{\gamma} K_0 \right) \int_0^\tau A_2(s)\, ds \\ &+ \frac{1}{2} \left(V_0 + \gamma \hat{v}_0 \right) \int_0^\tau A_2(s)^2\, ds. \end{aligned} \tag{6.35}$$

We summarize these findings in the following theorem.

Theorem 6.4 *Assume that $r(x)$, $\lambda(x)^\top \lambda(x)$, $m(x)$, $v(x)^\top \lambda(x)$, $v(x)^\top v(x)$, and $\hat{v}(x)^2$ are all affine functions of x and given by (6.27)–(6.32). For an investor with CRRA utility from terminal wealth only, the indirect utility function is then given by*

$$J(W, x, t) = e^{-\delta(T-t)} \frac{\left(e^{A_1(T-t) + A_2(T-t)x} \right)^\gamma W^{1-\gamma}}{1-\gamma}, \tag{6.36}$$

where A_2 is the solution to the ordinary differential equation (6.34) *with the initial condition $A_2(0) = 0$ and A_1 is given by* (6.35)*. The optimal investment strategy is given by*

$$\Pi(W, x, t) = \frac{1}{\gamma} \left(\sigma(x,t)^\top\right)^{-1} \lambda(x) + \left(\sigma(x,t)^\top\right)^{-1} v(x) A_2(T-t). \tag{6.37}$$

In some cases, A_1 and A_2 can be computed explicitly, see for example Section 7.1. In other cases, the ordinary differential equations for A_1 and A_2 can be solved quickly and accurately by numerical methods. Note the close connection between the analysis above and the analysis for so-called affine models of the term structure of interest rates, see e.g. Duffie and Kan (1996) and Dai and Singleton (2000).

6.4.2 Utility from consumption (and possibly terminal wealth)

Above we found closed-form solutions for the indirect utility function and the optimal investment strategy for a CRRA investor with utility from terminal wealth. These solutions apply both to the complete market case (where $\hat{v}(x) \equiv 0$) and the incomplete market case. With utility from intermediate consumption, the PDE (6.19) we have to solve for the function $g(x,t)$ is slightly more complicated due to the presence of the constant $\varepsilon_1 = 1$. It turns out that in order to find a solution of the same form as for utility of terminal wealth only, we must restrict ourselves to the complete market case, where $\hat{v}(x) \equiv 0$ so that the last term in (6.19) vanishes. The PDE reduces to

$$\begin{aligned} 0 = 1 &- \left(\frac{\delta}{\gamma} - \frac{1-\gamma}{\gamma} r(x) - \frac{1-\gamma}{2\gamma^2} \lambda(x)^\top \lambda(x)\right) g(x,t) + \frac{\partial g}{\partial t}(x,t) \\ &+ \left(m(x) + \frac{1-\gamma}{\gamma} v(x)^\top \lambda(x)\right) g_x(x,t) + \frac{1}{2} v(x)^\top v(x) g_{xx}(x,t) \end{aligned} \tag{6.38}$$

with terminal condition $g(x,T) = \varepsilon_2$, where ε_2 is either zero or one.

Considering the solution to the case with terminal wealth only, we try a solution of the form

$$\begin{aligned} g(x,t) = &\int_t^T \exp\left\{-\frac{\delta}{\gamma}(s-t) + A_1(s-t) + A_2(s-t)x\right\} ds \\ &+ \varepsilon_2 \exp\left\{-\frac{\delta}{\gamma}(T-t) + A_1(T-t) + A_2(T-t)x\right\}, \end{aligned}$$

where we must have $A_1(0) = A_2(0) = 0$ to satisfy the terminal condition. Let us for notational simplicity write

$$h(x,\tau) = -\frac{\delta}{\gamma}\tau + A_1(\tau) + A_2(\tau)x.$$

The relevant derivatives of $g(x,t)$ are now

$$g_x(x,t) = \int_t^T A_2(s-t) e^{h(x,s-t)}\, ds + \varepsilon_2 A_2(T-t) e^{h(x,T-t)}, \tag{6.39}$$

$$g_{xx}(x,t) = \int_t^T A_2(s-t)^2 e^{h(x,s-t)}\, ds + \varepsilon_2 A_2(T-t)^2 e^{h(x,T-t)}, \tag{6.40}$$

$$\begin{aligned} \frac{\partial g}{\partial t}(x,t) = &\int_t^T \left(\frac{\delta}{\gamma} - A_1'(s-t) - A_2'(s-t)x\right) e^{h(x,s-t)}\, ds - 1 \\ &+ \varepsilon_2 \left(\frac{\delta}{\gamma} - A_1'(T-t) - A_2'(T-t)x\right) e^{h(x,T-t)}, \end{aligned} \tag{6.41}$$

where we have used Leibnitz' rule for computing the derivative with respect to time.[3] Substituting these derivatives into (6.38), we get that if the functions A_1 and A_2 are so that

$$\begin{aligned} 0 = \frac{1-\gamma}{\gamma} r(x) + \frac{1-\gamma}{2\gamma^2} \lambda(x)^\top \lambda(x) - A_1'(\tau) - A_2'(\tau) x \\ + \left(m(x) + \frac{1-\gamma}{\gamma} v(x)^\top \lambda(x) \right)^\top A_2(\tau) + \frac{1}{2} v(x)^\top v(x) A_2(\tau)^2 \end{aligned} \tag{6.42}$$

for all $\tau > 0$, we indeed have a solution. This is exactly as (6.26) except that the $\hat{v}$ terms are now not present. We can therefore make a similar conclusion. In particular, if r, m, v, and λ satisfy (6.27), (6.28), (6.30), (6.31), and (6.32), we have that A_2 solves the ordinary differential equation

$$0 = \frac{1-\gamma}{\gamma} r_1 + \frac{1-\gamma}{2\gamma^2} \Lambda_1 - A_2'(\tau) + \left(m_1 + \frac{1-\gamma}{\gamma} K_1 \right) A_2(\tau) + \frac{1}{2} V_1 A_2(\tau)^2 \tag{6.43}$$

with the initial condition $A_2(0) = 0$, and A_1 can be computed as

$$A_1(\tau) = \left(\frac{1-\gamma}{\gamma} r_0 + \frac{1-\gamma}{2\gamma^2} \Lambda_0 \right) \tau + \left(m_0 + \frac{1-\gamma}{\gamma} K_0 \right) \int_0^\tau A_2(s)\, ds + \frac{1}{2} V_0 \int_0^\tau A_2(s)^2\, ds. \tag{6.44}$$

Concerning the optimal consumption strategy, recall the general result in (6.9).

Theorem 6.5 *Assume a complete financial market ($\hat{v}(x) \equiv 0$), where $r(x)$, $\lambda(x)^\top \lambda(x)$, $m(x)$, $v(x)\lambda(x)$, and $v(x)v(x)^\top$ are all affine functions of x and given by* (6.27), (6.28), (6.30), (6.31), *and* (6.32). *For an investor with CRRA utility from intermediate consumption and possibly terminal wealth, the indirect utility function is given by*

$$J(W, x, t) = \frac{1}{1-\gamma} \left(\int_t^T e^{h(x, s-t)}\, ds + \varepsilon_2 e^{h(x, T-t)} \right)^\gamma W^{1-\gamma}, \tag{6.45}$$

where

$$h(x, \tau) = -\frac{\delta}{\gamma} \tau + A_1(\tau) + A_2(\tau) x,$$

A_2 solves the ordinary differential equation (6.43) *with $A_2(0) = 0$, and A_1 is given by* (6.44). *The optimal investment strategy is given by*

$$\begin{aligned} \Pi(W, x, t) = \frac{1}{\gamma} \left(\sigma(x,t)^\top \right)^{-1} \lambda(x) \\ + \left(\sigma(x,t)^\top \right)^{-1} v(x) \frac{\int_t^T e^{h(x, s-t)} A_2(s-t)\, ds + \varepsilon_2 e^{h(x, T-t)} A_2(T-t)}{\int_t^T e^{h(x, s-t)}\, ds + \varepsilon_2 e^{h(x, T-t)}}. \end{aligned} \tag{6.46}$$

The optimal consumption strategy is given by

$$C(W, x, t) = \left(\int_t^T e^{h(x, s-t)}\, ds + \varepsilon_2 e^{h(x, T-t)} \right)^{-1} W. \tag{6.47}$$

[3] According to Leibnitz' rule, we have

$$\frac{\partial}{\partial t} \left(\int_{a(t)}^{b(t)} f(t, s)\, ds \right) = \int_{a(t)}^{b(t)} \frac{\partial f}{\partial t}(t, s)\, ds + b'(t) f(t, b(t)) - a'(t) f(t, a(t)),$$

assuming a and b are differentiable functions.

Let us look at the hedge term of the optimal investment strategy. Assuming for simplicity that $\varepsilon_2 = 0$, we can rewrite the hedge term as

$$\left(\sigma(x,t)^\top\right)^{-1} v(x) \frac{\int_t^T e^{h(x,s-t)} A_2(s-t)\,ds}{\int_t^T e^{h(x,s-t)}\,ds} = \left(\sigma(x,t)^\top\right)^{-1} v(x) \int_t^T w(x,s-t) A_2(s-t)\,ds,$$

where we have defined $w(x,s-t) = e^{h(x,s-t)} / \int_t^T e^{h(x,s-t)}\,ds$. Since $w(x,s-t) > 0$ and $\int_t^T w(x,s-t)\,ds = 1$, we may interpret the hedging demand of an investor with utility of consumption and a time horizon of T as a weighted average of the hedging demands of investors with time horizons of $s \in [t,T]$ and utility of terminal wealth only. If A_2 is either monotonically increasing or decreasing (as will be the case in many concrete settings), there will exist a $T^* \in [t,T]$ such that

$$\int_t^T w(x,s-t) A_2(s-t)\,ds = A_2(T^* - t),$$

in which case we can represent the hedging demand as $\left(\sigma(x,t)^\top\right)^{-1} v(x) A_2(T^* - t)$. Since this is exactly the hedging demand of an investor with time horizon T^* and utility of terminal wealth only, we may interpret T^* as the **effective time horizon** of the investor with time horizon T and utility of consumption. Note the similarity to the concept of duration for fixed-income securities, cf. Munk (2003a).

6.5 Closed-form solution for CRRA utility: quadratic models with one state variable

The assumptions of the affine models cover some interesting settings, but not all. In this section we shall see that under another set of assumptions on the market parameter functions r, m, v, λ, and $\hat{v}$, we obtain an exponential-quadratic expression for the function $g(x,t)$. In Section 7.2, we will study an important example which is covered by these assumptions. As for the affine models, we distinguish between the case with utility from terminal wealth only and the case with utility from intermediate consumption.

6.5.1 Utility from terminal wealth only

Let us consider when the PDE (6.25) has a solution of the form

$$g(x,t) = e^{-\frac{\delta}{\gamma}(T-t) + A_1(T-t) + A_2(T-t)x + \frac{1}{2}A_3(T-t)x^2},$$

where A_1, A_2, and A_3 are real-valued deterministic functions of time. From the terminal condition we must have that $A_1(0) = A_2(0) = A_3(0) = 0$. The relevant derivatives of g are

$$\begin{aligned}
g_x(x,t) &= \left(A_2(T-t) + A_3(T-t)x\right) g(x,t),\\
g_{xx}(x,t) &= \left(A_3(T-t) + A_2(T-t)^2 + 2A_2(T-t)A_3(T-t)x + A_3(T-t)^2x^2\right) g(x,t),\\
\frac{\partial g}{\partial t}(x,t) &= \left(\frac{\delta}{\gamma} - A_1'(T-t) - A_2'(T-t)x - \frac{1}{2}A_3'(T-t)x^2\right) g(x,t).
\end{aligned}$$

Inserting these into (6.25), dividing by $g(x,t)$, and replacing $T-t$ by τ, we get

$$\begin{aligned}0 &= \frac{1-\gamma}{\gamma}r(x) + \frac{1-\gamma}{2\gamma^2}\lambda(x)^\top\lambda(x) - A_1'(\tau) - A_2'(\tau)x - \frac{1}{2}A_3'(\tau)x^2 \\ &+ \left(m(x) + \frac{1-\gamma}{\gamma}v(x)^\top\lambda(x)\right)(A_2(\tau) + A_3(\tau)x) + \frac{1}{2}(1-\gamma)\hat{v}(x)^2A_3(\tau) \\ &+ \frac{1}{2}\left(v(x)^\top v(x) + \gamma\hat{v}(x)^2\right)\left(A_2(\tau)^2 + A_3(\tau) + 2A_2(\tau)A_3(\tau)x + A_3(\tau)^2x^2\right).\end{aligned} \tag{6.48}$$

To ensure that we only have powers of x of order zero, one, and two, we can allow (i) $r(x)$ and $\lambda(x)^\top\lambda(x)$ to be quadratic[4] in x, (ii) $m(x)$ and $v(x)^\top\lambda(x)$ can be affine in x, while (iii) $v(x)^\top v(x)$ and $\hat{v}(x)^2$ have to be constant. Therefore, write $v(x) = v = (v_1,\dots,v_d)^\top$, $\hat{v}(x) = \hat{v}$, and

$$r(x) = r_0 + r_1x + r_2x^2, \tag{6.49}$$

$$m(x) = m_0 + m_1x, \tag{6.50}$$

$$\lambda_i(x) = \lambda_{i0} + \lambda_{i1}x \tag{6.51}$$

for some constants $r_0, r_1, r_2, m_0, m_1, m_2, \lambda_{i0}, \lambda_{i1}, \lambda_{i2}$. Consequently,

$$\begin{aligned}\lambda(x)^\top\lambda(x) &= \sum_{i=1}^d \lambda_i(x)^2 = \left(\sum_{i=1}^d \lambda_{i0}^2\right) + 2\left(\sum_{i=1}^d \lambda_{i0}\lambda_{i1}\right)x + \left(\sum_{i=1}^d \lambda_{i1}^2\right)x^2 \\ &\equiv \Lambda_0 + \Lambda_1x + \Lambda_2x^2,\end{aligned} \tag{6.52}$$

$$v(x)^\top\lambda(x) = \sum_{i=1}^d v_i(x)\lambda_i(x) = \left(\sum_{i=1}^d v_i\lambda_{i0}\right) + \left(\sum_{i=1}^d v_i\lambda_{i1}\right)x \equiv K_0 + K_1x. \tag{6.53}$$

If we substitute (6.49)–(6.53) into (6.48) and use the fact that (6.48) must hold for all values of x and all t, we obtain a system of three ordinary differential equations for A_1, A_2, and A_3:

$$\begin{aligned}0 &= \frac{1-\gamma}{\gamma}r_0 + \frac{1-\gamma}{2\gamma^2}\Lambda_0 - A_1'(\tau) + \left(m_0 + \frac{1-\gamma}{\gamma}K_0\right)A_2(\tau) \\ &+ \frac{1}{2}\left(v^\top v + \gamma\hat{v}^2\right)A_2(\tau)^2 + \frac{1}{2}\left(v^\top v + \hat{v}^2\right)A_3(\tau),\end{aligned} \tag{6.54}$$

$$\begin{aligned}0 &= \frac{1-\gamma}{\gamma}r_1 + \frac{1-\gamma}{2\gamma^2}\Lambda_1 - A_2'(\tau) + \left(m_0 + \frac{1-\gamma}{\gamma}K_0\right)A_3(\tau) \\ &+ \left(m_1 + \frac{1-\gamma}{\gamma}K_1\right)A_2(\tau) + \left(v^\top v + \gamma\hat{v}^2\right)A_2(\tau)A_3(\tau),\end{aligned} \tag{6.55}$$

$$0 = \frac{1-\gamma}{\gamma}r_2 + \frac{1-\gamma}{2\gamma^2}\Lambda_2 - \frac{1}{2}A_3'(\tau) + \left(m_1 + \frac{1-\gamma}{\gamma}K_1\right)A_3(\tau) + \frac{1}{2}\left(v^\top v + \gamma\hat{v}^2\right)A_3(\tau)^2, \tag{6.56}$$

These equations are to be solved with the initial conditions $A_1(0) = A_2(0) = A_3(0) = 0$. First (6.56) can be solved for A_3, then (6.55) can be solved for A_2. Finally, we can compute $A_1(\tau)$ as

$$\begin{aligned}A_1(\tau) &= \left(\frac{1-\gamma}{\gamma}r_0 + \frac{1-\gamma}{2\gamma^2}\Lambda_0\right)\tau + \left(m_0 + \frac{1-\gamma}{\gamma}K_0\right)\int_0^\tau A_2(s)\,ds \\ &+ \frac{1}{2}\left(v^\top v + \gamma\hat{v}^2\right)\int_0^\tau A_2(s)^2\,ds + \frac{1}{2}\left(v^\top v + \hat{v}^2\right)\int_0^\tau A_3(s)\,ds.\end{aligned} \tag{6.57}$$

We summarize these findings in the following theorem.

[4] A real-valued function is said to be a quadratic function of the k-vector x, if it can be written as $a_1 + a_2^\top x + x^\top a_3 x$, where a_1 is a constant scalar, a_2 is a constant k-vector, and a_3 is a constant $(k\times k)$-matrix (either a_2 or a_3 or both can be zero so that a constant and an affine function are also considered quadratic. A vector- or matrix-valued function is said to be quadratic if all its elements are quadratic.

Theorem 6.6 *Assume that $v(x) = v$, $\hat{v}(x) = \hat{v}$, and that $r(x)$, $m(x)$, and $\lambda(x)$ are given as in* (6.49)–(6.51)*. For an investor with CRRA utility from terminal wealth only, the indirect utility function is then given by*

$$J(W,x,t) = e^{-\delta(T-t)} \frac{\left(e^{A_1(T-t)+A_2(T-t)x+\frac{1}{2}A_3(T-t)x^2}\right)^\gamma W^{1-\gamma}}{1-\gamma}, \tag{6.58}$$

where A_2 and A_3 solve the ordinary differential equations (6.55) *and* (6.56) *with $A_2(0) = A_3(0) = 0$, and A_1 is given by* (6.57)*. The optimal investment strategy is given by*

$$\Pi(W,x,t) = \frac{1}{\gamma}\left(\sigma(x,t)^\top\right)^{-1}\lambda(x) + \left(\sigma(x,t)^\top\right)^{-1} v(x)\left(A_2(T-t) + A_3(T-t)x\right). \tag{6.59}$$

Note the close connection to the so-called quadratic models of the term structure of interest rates, see e.g. Leippold and Wu (2000) and Ahn, Dittmar, and Gallant (2002).

6.5.2 Utility from consumption (and possibly terminal wealth)

For a complete market we can generalize the above results to encompass investors with utility from intermediate consumption. In this case $\hat{v} \equiv 0$ so the relevant equations for A_1, A_2, and A_3 reduce to

$$\begin{aligned} 0 &= \frac{1-\gamma}{\gamma}r_1 + \frac{1-\gamma}{2\gamma^2}\Lambda_1 - A_2'(\tau) + \left(m_0 + \frac{1-\gamma}{\gamma}K_0\right)A_3(\tau) \\ &\quad + \left(m_1 + \frac{1-\gamma}{\gamma}K_1\right)A_2(\tau) + v^\top v A_2(\tau)A_3(\tau), \end{aligned} \tag{6.60}$$

$$0 = \frac{1-\gamma}{\gamma}r_2 + \frac{1-\gamma}{2\gamma^2}\Lambda_2 - \frac{1}{2}A_3'(\tau) + \left(m_1 + \frac{1-\gamma}{\gamma}K_1\right)A_3(\tau) + \frac{1}{2}v^\top v A_3(\tau)^2, \tag{6.61}$$

$$\begin{aligned} A_1(\tau) &= \left(\frac{1-\gamma}{\gamma}r_0 + \frac{1-\gamma}{2\gamma^2}\Lambda_0\right)\tau + \left(m_0 + \frac{1-\gamma}{\gamma}K_0\right)\int_0^\tau A_2(s)\,ds \\ &\quad + \frac{1}{2}v^\top v\int_0^\tau A_2(s)^2\,ds. \end{aligned} \tag{6.62}$$

Analogously to the analysis of affine models, we can draw the following conclusion:

Theorem 6.7 *Assume that the market is complete ($\hat{v}(x) \equiv 0$), that $v(x) = v$, and that $r(x)$, $m(x)$, and $\lambda(x)$ are given as in* (6.49)–(6.51)*. For an investor with CRRA utility from intermediate consumption and possibly terminal wealth, the indirect utility function is then given by*

$$J(W,x,t) = \frac{1}{1-\gamma}\left(\int_t^T e^{h(x,s-t)}\,ds + \varepsilon_2 e^{h(x,T-t)}\right)^\gamma W^{1-\gamma}, \tag{6.63}$$

where

$$h(x,\tau) = -\frac{\delta}{\gamma}\tau + A_1(\tau) + A_2(\tau)x + \frac{1}{2}A_3(\tau)x^2,$$

A_2 and A_3 solve the ordinary differential equations (6.55) *and* (6.56) *with $A_2(0) = A_3(0) = 0$, and A_1 is given by* (6.57)*. The optimal investment strategy is given by*

$$\begin{aligned} \Pi(W,x,t) &= \frac{1}{\gamma}\left(\sigma(x,t)^\top\right)^{-1}\lambda(x) \\ &\quad + \left(\sigma(x,t)^\top\right)^{-1} v(x)\frac{\int_t^T e^{h(x,s-t)}\left(A_2(s-t) + A_3(s-t)x\right)ds}{\int_t^T e^{h(x,s-t)}\,ds + \varepsilon_2 e^{h(x,T-t)}} \\ &\quad + \left(\sigma(x,t)^\top\right)^{-1} v(x)\varepsilon_2\frac{e^{h(x,T-t)}\left(A_2(T-t) + A_3(T-t)x\right)}{\int_t^T e^{h(x,s-t)}\,ds + \varepsilon_2 e^{h(x,T-t)}}. \end{aligned} \tag{6.64}$$

The optimal consumption strategy is given by

$$C(W,x,t) = \left(\int_t^T e^{h(x,s-t)}\, ds + \varepsilon_2 e^{h(x,T-t)} \right)^{-1} W. \tag{6.65}$$

6.6 Closed-form solution for CRRA utility: multiple state variables

In the preceding two subsections, we have stated conditions under which we can find closed-form solutions for the optimal strategies of CRRA investors, when the dynamics of the investment opportunity set can be represented by a one-dimensional state variable. It will come as no surprise for the experienced reader that these results generalize to settings with a multi-dimensional state variable. We get exactly the same results as in Theorems 6.4–6.7 except that the A_2-function is now vector-valued and the A_3-function is matrix-valued. We get a larger system of differential equations to solve. For example, in the affine case with a k-dimensional state variable we have to simultaneously solve k ordinary differential equations for the k components of A_2.

Finally, there are cases in which the solution function $g(x,t)$ is the exponential of the sum of a function which is affine in some of the individual state variables and quadratic in the others. For example, with a two-dimensional state variable $x = (x_1, x_2)^\top$, we will under some conditions get a solution of the form

$$g(x_1,x_2,t) = \exp\left\{ A_1(T-t) + A_2(T-t)x_1 + A_3(T-t)x_2 + \frac{1}{2}A_4(T-t)x_2^2 \right\},$$

and, consequently, the investment strategy

$$\begin{aligned}\Pi(W,x,t) &= \frac{1}{\gamma}\left(\sigma(x,t)^\top\right)^{-1}\lambda(x) \\ &\quad + \left(\sigma(x,t)^\top\right)^{-1}\left[v_1(x)A_2(T-t) + v_2(x)\left(A_3(T-t) + A_4(T-t)x_2\right)\right],\end{aligned}$$

where v_i is the d-vector of sensitivities of x_i with respect to the "traded" risks dz_t.

6.7 Applying the martingale approach

As discussed in Section 4.6 portfolio/consumption problems can also be analyzed using the so-called martingale approach instead of the dynamic programming approach used above. Recall that the martingale approach does not require a Markov system of state variables, but can, in principle, be used for any price dynamics. The application of the martingale approach is considerably more complex for incomplete markets, so we assume a complete market setting. According to Theorem 4.3 the optimal consumption rate is given by

$$c_t^* = I_u\left(\mathcal{Y}(W_0)e^{\delta t}\zeta_t\right)$$

and the optimal level of terminal wealth level is

$$W^* = I_{\bar{u}}\left(\mathcal{Y}(W_0)e^{\delta T}\zeta_T\right).$$

The wealth process under the optimal policy is given by

$$W_t^* = \frac{1}{\zeta_t}\,\mathrm{E}_t\left[\int_t^T \zeta_s c_s^*\, ds + \zeta_T W^*\right]. \tag{6.66}$$

Let us again consider the case of CRRA utility

$$u(c) = \frac{c^{1-\gamma}}{1-\gamma}, \qquad \bar{u}(W) = \frac{W^{1-\gamma}}{1-\gamma}.$$

Then

$$u'(c) = c^{-\gamma}, \qquad \bar{u}'(W) = W^{-\gamma}$$

with inverse functions

$$I_u(z) = z^{-\frac{1}{\gamma}}, \qquad I_{\bar{u}}(z) = z^{-\frac{1}{\gamma}}.$$

Consequently, the function $\mathcal{H}$ defined in (4.23) can be computed as

$$\begin{aligned}\mathcal{H}(\psi) &= \mathrm{E}\left[\int_0^T \zeta_t e^{-\frac{\delta}{\gamma}t}\psi^{-\frac{1}{\gamma}}\zeta_t^{-\frac{1}{\gamma}}\,dt + \zeta_T e^{-\frac{\delta}{\gamma}T}\psi^{-\frac{1}{\gamma}}\zeta_T^{-\frac{1}{\gamma}}\right] \\ &= \psi^{-\frac{1}{\gamma}}\left(\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right]\right)\end{aligned}$$

with inverse function

$$\mathcal{Y}(W_0) = W_0^{-\gamma}\left(\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right]\right)^{\gamma}.$$

Therefore, the optimal consumption policy is

$$\begin{aligned}c_t^* &= e^{-\frac{\delta}{\gamma}t}\mathcal{Y}(W_0)^{-\frac{1}{\gamma}}\zeta_t^{-\frac{1}{\gamma}} \\ &= e^{-\frac{\delta}{\gamma}t}W_0\zeta_t^{-\frac{1}{\gamma}}\left(\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right]\right)^{-1},\end{aligned}$$

and the optimal terminal wealth level is

$$\begin{aligned}W^* &= e^{-\frac{\delta}{\gamma}T}\mathcal{Y}(W_0)^{-\frac{1}{\gamma}}\zeta_T^{-\frac{1}{\gamma}} \\ &= e^{-\frac{\delta}{\gamma}T}W_0\zeta_T^{-\frac{1}{\gamma}}\left(\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right]\right)^{-1}.\end{aligned}$$

The indirect utility function becomes

$$\begin{aligned}J(W_0) &= \mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\frac{1}{1-\gamma}\mathcal{Y}(W_0)^{1-\frac{1}{\gamma}}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\frac{1}{1-\gamma}\mathcal{Y}(W_0)^{1-\frac{1}{\gamma}}\zeta_T^{1-\frac{1}{\gamma}}\right] \\ &= \frac{\mathcal{Y}(W_0)^{1-\frac{1}{\gamma}}}{1-\gamma}\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right] \\ &= \frac{W_0^{1-\gamma}}{1-\gamma}\left(\mathrm{E}\left[\int_0^T e^{-\frac{\delta}{\gamma}t}\zeta_t^{1-\frac{1}{\gamma}}\,dt + e^{-\frac{\delta}{\gamma}T}\zeta_T^{1-\frac{1}{\gamma}}\right]\right)^{\gamma}.\end{aligned}$$

Note that we found a solution of the same form with the dynamic programming approach, cf. (6.7).

Munk and Sørensen (2003) provide a characterization of the optimal portfolio in this general setting. Define the process $g = (g_t)$ by

$$g_t = \mathrm{E}_t\left[\int_t^T e^{-\frac{\delta}{\gamma}(s-t)}\left(\frac{\zeta_s}{\zeta_t}\right)^{1-1/\gamma}ds + e^{-\frac{\delta}{\gamma}(T-t)}\left(\frac{\zeta_T}{\zeta_t}\right)^{1-1/\gamma}\right].$$

g is positive and adapted, and if we write the dynamics of g as

$$dg_t = g_t\left[\mu_t^g\,dt + (\sigma_t^g)^\top\,dz_t\right],$$

we have the following result due to Munk and Sørensen:

Theorem 6.8 *In the setting above, the indirect utility at time t can be expressed as*

$$J_t = \frac{g_t^{\gamma} W_t^{1-\gamma}}{1-\gamma},$$

the optimal consumption strategy is

$$c_t = \frac{W_t}{g_t},$$

and the optimal portfolio strategy is

$$\pi_t = \frac{1}{\gamma} \left(\sigma_t^{\top}\right)^{-1} \lambda_t + \left(\sigma_t^{\top}\right)^{-1} \sigma_t^{g}.$$

This result can be verified by computing the dynamics of wealth when the strategy (c, π) described in the theorem is followed. Since this wealth dynamics will match the optimal wealth process in (6.66), the strategy is optimal.

The above theorem is a natural generalization of the results obtained in Markov settings using the dynamic programming approach. The hedge term of the portfolio is matching the volatility of the process g which is important for consumption. Looking at the definition of g, we can see that only variations in the state-price density, i.e. in interest rates and market prices of risk, will be hedged. This is also in line with findings in Markov set-ups. Of course, σ^g has to be identified in order for this result to be of practical relevance. This is possible in many concrete cases, primarily cases with Markov dynamics where the dynamic programming approach also applies, but Munk and Sørensen consider a relevant and non-trivial example with non-Markov dynamics. For investors with logarithmic utility ($\gamma = 1$), we see that the process (g_t) is deterministic so that the volatility σ^g is zero. The optimal portfolio of a log investor is therefore $\pi_t = \frac{1}{\gamma} \left(\sigma_t^{\top}\right)^{-1} \lambda_t$ as has already been shown for Markov settings.

Chapter 7

Asset allocation with stochastic investment opportunities: concrete cases

In the previous chapter we discussed optimal portfolio and consumption strategies for CRRA investors in settings with a general state variable x. In this section we will look at some concrete cases of practical interest. First, we consider models where interest rates are stochastic and market prices of risk are at most dependent on interest rates. We focus on one-factor models of interest rate dynamics, where the short-term interest rates is the state variable. Next, we consider a model where interest rates and asset price volatilities are constant, but the market price of stock market risk varies stochastically over time. In particular, the excess rate of return on the stock market is assumed to follow a mean-reverting process around some long-term average.

7.1 Stochastic interest rates

It is an empirical fact that both nominal and real interest rates (at least of finite maturity) vary stochastically over time. It is therefore natural to include the short-term interest rate r_t as a state variable. This was first done in a portfolio-choice context by Merton (1973c) who considered a general one-factor dynamics for r_t, but he was not able to go beyond the general characterization of the investment strategy in (6.5). From our general analysis in the previous chapter we can see that we will be able to get explicit solutions for affine and quadratic short-rate models of the term structure of interest rates including the well-known models of Vasicek (1977) and Cox, Ingersoll, and Ross (1985); see for example Munk (2003a) for a comprehensive analysis of dynamic models of the term structure of interest rates. We can also apply the results of the previous chapter on cases where the dynamics of the term structure of interest rate is given by a multi-factor affine or quadratic model. Due to the diversity and accessibility of bond markets it is natural to consider a complete market setting where interest rate risks are fully hedgeable. Exercise 7.1 at the end of the chapter discusses the optimal investment problem with stochastic interest rates when no bonds are traded.

We will focus on determining the optimal bond/stock mix so we assume that only a single stock is traded. We interpret this stock as the entire stock market index. The results can be generalized to the case with multiple stocks. Investors with non-log utility will hedge variations in interest rates. Bonds carry a build-in hedge against interest rate risk since bond prices are inversely related to interest rates. Over a period where interest rates have fallen, indicating that future investment opportunities are worsened, bond prices have risen and generated a positive return. The converse is also true.

7.1.1 One-factor Vasicek interest rate dynamics

Following Vasicek (1977), assume that r_t follows the Ornstein-Uhlenbeck process

$$dr_t = \kappa[\bar{r} - r_t]\, dt - \sigma_r\, dz_{1t}, \tag{7.1}$$

with an associated constant market price of risk λ_1. From the Vasicek model we know that the price of a zero-coupon bond with maturity $\bar{T}$ is given by

$$B_t^{\bar{T}} = e^{-a(\bar{T}-t)-b(\bar{T}-t)r_t}, \tag{7.2}$$

where

$$\begin{aligned} b(\tau) &= \frac{1}{\kappa}\left(1 - e^{-\kappa\tau}\right), \\ a(\tau) &= \left(\bar{r} + \frac{\lambda_1\sigma_r}{\kappa} - \frac{\sigma_r^2}{2\kappa^2}\right)(\tau - b(\tau)) + \frac{\sigma_r^2}{4\kappa} b(\tau)^2. \end{aligned}$$

From Itô's Lemma it follows that the dynamics of the zero-coupon bond price is

$$dB_t^{\bar{T}} = B_t^{\bar{T}}\left[\left(r_t + \lambda_1\sigma_r b(\bar{T}-t)\right)\, dt + \sigma_r b(\bar{T}-t)\, dz_{1t}\right],$$

and similarly the dynamics of any bond (and other fixed-income securities) is of the form

$$dB_t = B_t\left[(r_t + \lambda_1\sigma_B(r_t,t))\, dt + \sigma_B(r_t,t)\, dz_{1t}\right]. \tag{7.3}$$

It is well-known that any bond (or other fixed-income security) can be generated from an appropriate dynamic investment strategy in the bank account and in just one (arbitrary) bond (or other long-lived term structure derivative). Let us for the present take an arbitrary bond with price B_t and dynamics given by (7.3).

The price of the single stock (representing the stock market index) is assumed to follow the process

$$dS_t = S_t\left[(r_t + \psi\sigma_S)\, dt + \rho\sigma_S\, dz_{1t} + \sqrt{1-\rho^2}\sigma_S\, dz_{2t}\right].$$

The parameter ρ is the correlation between bond market returns and stock market returns, σ_S is the volatility of the stock, and ψ is the Sharpe ratio of the stock which we assume constant.

To get this into the notation applied so far, we rewrite the price dynamics as

$$\begin{aligned} \begin{pmatrix} dB_t \\ dS_t \end{pmatrix} = \begin{pmatrix} B_t & 0 \\ 0 & S_t \end{pmatrix} \Bigg[\left(r_t\mathbf{1} + \begin{pmatrix} \sigma_B(r_t,t) & 0 \\ \rho\sigma_S & \sqrt{1-\rho^2}\sigma_S \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \right) dt \\ + \begin{pmatrix} \sigma_B(r_t,t) & 0 \\ \rho\sigma_S & \sqrt{1-\rho^2}\sigma_S \end{pmatrix} \begin{pmatrix} dz_{1t} \\ dz_{2t} \end{pmatrix} \Bigg], \end{aligned} \tag{7.4}$$

where

$$\lambda_2 = (\psi - \rho\lambda_1)/\sqrt{1-\rho^2}. \tag{7.5}$$

We are therefore in a complete market model with a single state variable ($x = r$). We can rewrite the dynamics of r as

$$dr_t = \kappa[\bar{r} - r_t]\, dt + \left(-\sigma_r, 0\right) dz_t,$$

where $z = (z_1, z_2)^\top$. In this model the state variable has an affine drift and a constant volatility, and the market price of risk vector $\lambda = (\lambda_1, \lambda_2)^\top$ is also constant. Hence, Theorem 6.4 applies with CRRA utility from terminal wealth only and Theorem 6.5 applies with CRRA utility from intermediate consumption and possibly terminal wealth. In the notation used there, we have

$$\begin{aligned} r_0 &= 0, & r_1 &= 1, \\ m_0 &= \kappa\bar{r}, & m_1 &= -\kappa, \\ \Lambda_0 &= \lambda_1^2 + \lambda_2^2, & \Lambda_1 &= 0, \\ \hat{v}_0 &= 0, & \hat{v}_1 &= 0, \\ V_0 &= \sigma_r^2, & V_1 &= 0, \\ K_0 &= -\sigma_r\lambda_1, & K_1 &= 0. \end{aligned}$$

The ordinary differential equation (6.34) therefore reduces to

$$A_2'(\tau) = \frac{1-\gamma}{\gamma} - \kappa A_2(\tau),$$

which with the initial condition $A_2(0) = 0$ has the unique solution

$$A_2(\tau) = \frac{1-\gamma}{\kappa\gamma}\left(1 - e^{-\kappa\tau}\right) = \frac{1-\gamma}{\gamma} b(\tau). \tag{7.6}$$

Next, A_1 follows from (6.35):

$$\begin{aligned} A_1(\tau) &= \frac{1-\gamma}{2\gamma^2}\left(\lambda_1^2 + \lambda_2^2\right)\tau + \left(\kappa\bar{r} - \frac{1-\gamma}{\gamma}\sigma_r\lambda_1\right)\int_0^\tau A_2(s)\, ds + \frac{1}{2}\sigma_r^2 \int_0^\tau A_2(s)^2\, ds \\ &= \frac{1-\gamma}{2\gamma^2}\left(\lambda_1^2 + \lambda_2^2\right)\tau + \frac{1-\gamma}{\gamma}\left(\kappa\bar{r} - \frac{1-\gamma}{\gamma}\sigma_r\lambda_1\right)\int_0^\tau b(s)\, ds + \frac{(1-\gamma)^2}{2\gamma^2}\sigma_r^2\int_0^\tau b(s)^2\, ds \\ &= \frac{1-\gamma}{\gamma}\left(\frac{1-\gamma}{2\gamma}\frac{\sigma_r^2}{2\kappa^2} + \bar{r} - \frac{1-\gamma}{\gamma}\frac{\sigma_r\lambda_1}{\kappa}\right)(\tau - b(\tau)) - \frac{(1-\gamma)^2}{4\kappa\gamma^2}\sigma_r^2 b(\tau)^2 \\ &\quad + \frac{1-\gamma}{2\gamma^2}\left(\lambda_1^2 + \lambda_2^2\right)\tau, \end{aligned} \tag{7.7}$$

where we have used that

$$\int_0^\tau b(s)\, ds = \frac{1}{\kappa}\left(\tau - b(\tau)\right), \qquad \int_0^\tau b(s)^2\, ds = \frac{1}{\kappa^2}\left(\tau - b(\tau)\right) - \frac{1}{2\kappa}b(\tau)^2.$$

For the case with utility from terminal wealth only we have from Theorem 6.4 that the optimal investment strategy is

$$\Pi(W, r, t) \equiv \begin{pmatrix} \Pi_B(W, r, t) \\ \Pi_S(W, r, t) \end{pmatrix} = \frac{1}{\gamma}\left(\sigma(t)^\top\right)^{-1}\lambda + \left(1 - \frac{1}{\gamma}\right)\left(\sigma(t)^\top\right)^{-1}\begin{pmatrix} \sigma_r \\ 0 \end{pmatrix} b(T - t). \tag{7.8}$$

We can see that the hedge portfolio only involves the bond, not the stock, which should not come as a surprise since bonds seem more appropriate for hedging interest rate risk than stocks. The higher the risk aversion γ, the lower the investment in the tangency portfolio and the higher the investment in the hedge bond. The inverse of the transposed volatility matrix is

$$\begin{pmatrix} \sigma_B(r, t) & \rho\sigma_S \\ 0 & \sqrt{1-\rho^2}\sigma_S \end{pmatrix}^{-1} = \frac{1}{\sqrt{1-\rho^2}\sigma_B(r, t)\sigma_S}\begin{pmatrix} \sqrt{1-\rho^2}\sigma_S & -\rho\sigma_S \\ 0 & \sigma_B(r, t) \end{pmatrix}$$

so that we can write out the fraction of wealth invested in the stock and the bond as

$$\Pi_S(W,r,t) = \frac{\lambda_2}{\gamma\sigma_S\sqrt{1-\rho^2}}, \tag{7.9}$$

$$\Pi_B(W,r,t) = \frac{1}{\gamma\sigma_B(r,t)}\left(\lambda_1 - \frac{\rho}{\sqrt{1-\rho^2}}\lambda_2\right) + \left(1-\frac{1}{\gamma}\right)\frac{\sigma_r b(T-t)}{\sigma_B(r,t)}. \tag{7.10}$$

If the bond in the portfolio is the zero-coupon bond maturing at the end of the investment horizon of the investor, i.e. at time T, then $\sigma_B(r,t) = \sigma_r b(T-t)$, and we see that the hedge term simply consists of a fraction $1-1/\gamma$ in the zero-coupon bond. This is a natural choice of hedge instrument since it is exactly the truly riskfree asset for an investor only interested in time T wealth. The log utility investor ($\gamma = 1$) does not hedge. The hedge position of a less risk averse investor ($\gamma < 1$) is negative, while a more risk averse investor ($\gamma > 1$) takes a long position in the bond in order to hedge interest rate risk. An infinitely risk averse investor ($\gamma \to \infty$) will invest her entire wealth in the zero-coupon bond maturing at T.

If we continue to use the zero-coupon bond maturing at T as the bond instrument, we see from (7.8) that we can write the risky part of the optimal investment strategy as

$$\Pi(W,r,t) \equiv \begin{pmatrix}\Pi_B(W,r,t)\\ \Pi_S(W,r,t)\end{pmatrix} = \frac{1}{\gamma}\left(\sigma(t)^\top\right)^{-1}\lambda + \left(1-\frac{1}{\gamma}\right)\begin{pmatrix}1\\0\end{pmatrix}.$$

Consequently, the fraction of wealth invested in the bank account (i.e. the *locally* riskfree asset) is

$$\begin{aligned}\Pi_0(W,r,t) &= 1 - \Pi_B(W,r,t) - \Pi_S(W,r,t)\\ &= 1 - \frac{1}{\gamma}\mathbf{1}^\top\left(\sigma(t)^\top\right)^{-1}\lambda - \left(1-\frac{1}{\gamma}\right)\\ &= \frac{1}{\gamma}\left(1 - \mathbf{1}^\top\left(\sigma(t)^\top\right)^{-1}\lambda\right).\end{aligned}$$

Note that the term in the parenthesis is exactly what a log investor would hold in the bank account. The entire investment strategy can be written as

$$\begin{pmatrix}\Pi_0\\ \Pi_B\\ \Pi_S\end{pmatrix} = \frac{1}{\gamma}\begin{pmatrix}\Pi_0^{\log}\\ \Pi_B^{\log}\\ \Pi_S^{\log}\end{pmatrix} + \left(1-\frac{1}{\gamma}\right)\begin{pmatrix}0\\1\\0\end{pmatrix}. \tag{7.11}$$

The strategy is hence a simple combination of the log investor's portfolio and the zero-coupon bond maturing at the investment horizon of the investor. Note that as the risk aversion γ increases, the position in stocks will decrease, while the position in bonds will increase. Hence, the bond/stock ratio increases with risk aversion consistent with popular advice. However, the allocation to stock is still independent of the investment horizon which conflicts with traditional advice that the stock weight should increase with the investment horizon.

With utility from intermediate consumption only, it follows from Theorem 6.5 that the hedge term of the optimal bond investment strategy is

$$\left(1-\frac{1}{\gamma}\right)\frac{\sigma_r}{\sigma_B(r_t,t)}\frac{\int_t^T e^{-\frac{\delta}{\gamma}(s-t)+A_1(s-t)+\frac{1-\gamma}{\gamma}b(s-t)r}b(s-t)\,ds}{\int_t^T e^{-\frac{\delta}{\gamma}(s-t)+A_1(s-t)+\frac{1-\gamma}{\gamma}b(s-t)r}\,ds}, \tag{7.12}$$

where $\sigma_B(r_t,t)$ again represents the volatility of the bond chosen for implementing the strategy. It can be shown that the time t volatility of a coupon bond paying a continuous coupon at a deterministic rate $K(s)$ up to time T is given by

$$\sigma_B(r,t) = \frac{\int_t^T K(s) B_t^s \sigma_r b(s-t)\, ds}{\int_t^T K(s) B_t^s\, ds}.$$

Hence, we can interpret the time t interest rate hedge as the fraction $1 - 1/\gamma$ of wealth invested in a bond with continuous coupon

$$K(s) = e^{a(s-t)+A_1(s-t)-\frac{\delta}{\gamma}(s-t)+\frac{1}{\gamma}b(s-t)r}.$$

Munk and Sørensen (2003) show that this coupon is closely related to the expected consumption rate at time s. For an investor with utility from consumption over the entire period $[t, T]$, the zero-coupon bond maturing at T is no longer the truly riskfree asset. Since the investor is interested in payments at all dates in $[t, T]$, he hedges interest rate risk by investing in a combination of all zero-coupon bonds maturing in this interval, i.e. in some sort of coupon bond.

This problem was studied by Sørensen (1999) and Bajeux-Besnainou, Jordan, and Portait (2000) for utility of terminal wealth only. Korn and Kraft (2001) discuss some technical issues related to the application of the verification theorem to this problem.

7.1.2 One-factor CIR dynamics

Consider the same set-up as above except that the short-term interest rate now is assumed to follow the square-root process

$$dr_t = \kappa[\bar{r} - r_t]\, dt - \sigma_r \sqrt{r_t}\, dz_{1t} \tag{7.13}$$

with the market price of risk on term structure derivatives given by $\lambda_1(r,t) = \lambda_1 \sqrt{r_t}/\sigma_r$ as suggested by Cox, Ingersoll, and Ross (1985). Again, zero-coupon bond prices are of the form

$$B_t^{\bar{T}} = e^{-a(\bar{T}-t)-b(\bar{T}-t)r_t},$$

but a and b are now given by

$$\begin{aligned} b(\tau) &= \frac{2(e^{\nu\tau}-1)}{(\nu+\hat{\kappa})(e^{\nu\tau}-1)+2\nu}, \\ a(\tau) &= -\frac{2\kappa\bar{r}}{\sigma_r^2}\left(\frac{1}{2}(\hat{\kappa}+\nu)\tau + \ln\frac{2\nu}{(\nu+\hat{\kappa})(e^{\nu\tau}-1)+2\nu}\right), \end{aligned}$$

where $\hat{\kappa} = \kappa - \lambda_1$ and $\nu = \sqrt{\hat{\kappa}^2 + 2\sigma_r^2}$. The price evolves as

$$dB_t^{\bar{T}} = B_t^{\bar{T}}\left[\left(r_t + b(\bar{T}-t)\lambda_1 r_t\right)\, dt + b(\bar{T}-t)\sigma_r\sqrt{r_t}\, dz_{1t}\right]. \tag{7.14}$$

We assume that the investor can also trade in a single stock with price S_t evolving as

$$dS_t = S_t\left[(r_t + \psi(r_t)\sigma_S)\, dt + \rho\sigma_S\, dz_{1t} + \sqrt{1-\rho^2}\sigma_S\, dz_{2t}\right]. \tag{7.15}$$

Here σ_S is a positive constant, and z_2 is a one-dimensional standard Brownian motion independent of z_1 so that the constant ρ is the instantaneous correlation between stock returns and bond returns. We assume that the market price of risk associated with z_2 is a constant λ_2 so that

$$\psi(r) = \rho\frac{\lambda_1}{\sigma_r}\sqrt{r} + \sqrt{1-\rho^2}\lambda_2. \tag{7.16}$$

Again we have an affine, complete market model of the type studied in Section 6.4. In this case we have

$$\begin{aligned} r_0 &= 0, & r_1 &= 1, \\ m_0 &= \kappa\bar{r}, & m_1 &= -\kappa, \\ \Lambda_0 &= \lambda_2^2, & \Lambda_1 &= \frac{\lambda_1^2}{\sigma_r^2}, \\ \hat{v}_0 &= 0, & \hat{v}_1 &= 0, \\ V_0 &= 0, & V_1 &= \sigma_r^2, \\ K_0 &= 0, & K_1 &= -\lambda_1. \end{aligned}$$

The ordinary differential equation (6.34) for A_2 becomes

$$0 = -\left(1 - \frac{1}{\gamma}\right)\left(1 + \frac{\lambda_1^2}{2\gamma\sigma_r^2}\right) - A_2'(\tau) - \left(\kappa - \left(1 - \frac{1}{\gamma}\right)\lambda_1\right)A_2(\tau) + \frac{1}{2}\sigma_r^2 A_2(\tau)^2,$$

which with the initial condition $A_2(0) = 0$ has the unique solution

$$A_2(\tau) = -\left(1 - \frac{1}{\gamma}\right)A_2^*(\tau), \tag{7.17}$$

where

$$A_2^*(\tau) = \frac{2\left(1 + \frac{\lambda_1^2}{2\gamma\sigma_r^2}\right)\left(e^{\bar{\nu}\tau} - 1\right)}{\left(\bar{\nu} + \bar{\kappa}\right)\left(e^{\bar{\nu}\tau} - 1\right) + 2\bar{\nu}}, \tag{7.18}$$

and we have introduced the additional auxiliary parameters

$$\begin{aligned} \bar{\kappa} &= \kappa - \left(1 - \frac{1}{\gamma}\right)\lambda_1, \\ \bar{\nu} &= \sqrt{\bar{\kappa}^2 + 2\sigma_r^2\left(1 - \frac{1}{\gamma}\right)\left(1 + \frac{\lambda_1^2}{2\gamma\sigma_r^2}\right)}. \end{aligned}$$

A_1 can then be computed from (6.35):

$$\begin{aligned} A_1(\tau) &= \frac{1-\gamma}{2\gamma^2}\lambda_2^2\tau + \kappa\bar{r}\int_0^\tau A_2(s)\,ds \\ &= -\frac{1}{2\gamma}\left(1 - \frac{1}{\gamma}\right)\lambda_2^2\tau + \frac{2\kappa\bar{r}}{\sigma_r^2}\left(\frac{1}{2}\left(\bar{\nu} + \bar{\kappa}\right)\tau + \ln\frac{2\bar{\nu}}{\left(\bar{\nu} + \bar{\kappa}\right)\left(e^{\bar{\nu}\tau} - 1\right) + 2\bar{\nu}}\right), \end{aligned} \tag{7.19}$$

It follows that the optimal investment strategy for an investor with CRRA utility from terminal wealth only is

$$\Pi_B(W, r, t) = \frac{1}{\gamma\sigma_B(r,t)}\left(\frac{\lambda_1}{\sigma_r}\sqrt{r} - \frac{\rho\lambda_2}{\sqrt{1-\rho^2}}\right) + \left(1 - \frac{1}{\gamma}\right)\frac{\sigma_r\sqrt{r}}{\sigma_B(r,t)}A_2^*(T-t), \tag{7.20}$$

$$\Pi_S(W, r, t) = \frac{\lambda_2}{\gamma\sigma_S\sqrt{1-\rho^2}}. \tag{7.21}$$

If the bond instrument used is the zero-coupon bond maturing at the end of the investor's horizon, we have $\sigma_B(r,t) = \sigma_r\sqrt{r}b(T-t)$, and the hedge component will simplify to $(1-1/\gamma)A_2^*(t-t)/b(T-t)$. As opposed to the Vasicek case we do not have $A_2(T-t) = \frac{1-\gamma}{\gamma}b(T-t)$, i.e. $A_2^*(T-t)$ is generally different from $b(T-t)$. This implies that the optimal hedge consists of investing the time-varying fraction $(1-1/\gamma)A_2^*(T-t)/b(T-t)$ in the zero-coupon bond maturing at the end of the investor's horizon. A similar result was obtained by Grasselli (2000) using the martingale approach (discussed in a section below) for the case of utility from terminal wealth only.

For an investor with CRRA utility of intermediate consumption only, the fraction of wealth optimally invested in the stock is the same as above, while the fraction of wealth optimally invested in the bond instrument changes to

$$\Pi_B(W,r,t) = \frac{1}{\gamma\sigma_B(r,t)}\left(\frac{\lambda_1}{\sigma_r}\sqrt{r} - \frac{\rho\lambda_2}{\sqrt{1-\rho^2}}\right) + \left(1-\frac{1}{\gamma}\right)\frac{\sigma_r\sqrt{r}}{\sigma_B(r,t)}\frac{\int_t^T A_2^*(s-t)e^{-\frac{\delta}{\gamma}(s-t)+A_1(s-t)-(1-\frac{1}{\gamma})A_2^*(s-t)r}\,ds}{\int_t^T e^{-\frac{\delta}{\gamma}(s-t)+A_1(s-t)-(1-\frac{1}{\gamma})A_2^*(s-t)r}\,ds}. \tag{7.22}$$

7.1.3 A numerical example

We will take historical estimates of mean returns, standard deviations, and correlations as representative of future investment opportunities. These estimates are taken from Dimson, Marsh, and Staunton (2002). All returns are measured per year. The historical average real return on the U.S. stock market is $\mu_S = 8.7\%$ with a standard deviation of $\sigma_S = 20.2\%$, while the average real return on bonds is $\mu_B = 2.1\%$ with a standard deviation of $\sigma_B = 10.0\%$. The average real U.S. short-term interest rate is $\bar{r} = 1.0\%$. The correlation between stock returns and bond returns is $\rho = 0.2$. Different bonds will have different average returns and different standard deviation of the return. Similarly, the correlation between the return on a bond and the return on the stock market index may not be identical for all bonds. It is not clear exactly what bond or bond index, the above estimates are based on, but we will throughout the assignment assume that the estimates for μ_B and σ_B apply to a 10-year zero-coupon bond.

The volatility matrix of the bond and the stock is

$$\sigma = \begin{pmatrix} 0.1 & 0 \\ 0.0404 & 0.1979 \end{pmatrix}.$$

The (average) Sharpe ratio of the bond is $\lambda_1 = (2.1 - 1.0)/10.0 = 0.11$ and the (average) Sharpe ratio of the stock market is $\psi = (8.7 - 1.0)/20.2 \approx 0.3812$. Using (7.5) this corresponds to a market price of risk of $\lambda_2 \approx 0.3666$ on the exogenous shock that only affects the stock market. The variance-covariance matrix of returns in $\Sigma = \sigma\sigma^\top$. According to (3.10), the tangency portfolio of the bond and the stock is given by

$$\pi^{\tan} = \begin{pmatrix} \pi_B^{\tan} \\ \pi_S^{\tan} \end{pmatrix} = \begin{pmatrix} 0.1596 \\ 0.8404 \end{pmatrix},$$

so that the bond/stock ratio is approximately 0.19. Recall that this will be true for all time-additive agents that believe investment opportunities are constant over time. The tangency portfolio has a mean return of 7.65% and a standard deviation of 17.37%.

CRRA investors ignoring the fluctuations of interest rates will choose a portfolio of risky assets given by $\pi = \frac{1}{\gamma}[\mathbf{1}^\top(\sigma^\top)^{-1}\lambda]\pi^{\tan}$, where γ is the relative risk aversion of the agent. The portfolio is independent of the investment horizon. In Table 7.1 we show the portfolio allocation for various γ-values. The numbers in the column "tangency" denotes the fraction of wealth invested in the tangency portfolio. This investment is divided into the bond and the stock in the following two

γ	tangency	bond	stock	cash	exp. return	volatility
0.5	4.4079	0.7034	3.7045	-3.4079	0.3030	0.7655
1	2.2039	0.3517	1.8522	-1.2039	0.1565	0.3827
2	1.1020	0.1758	0.9261	-0.1020	0.0832	0.1914
2.2039	1.0000	0.1596	0.8404	0.0000	0.0765	0.1737
3	0.7346	0.1172	0.6174	0.2654	0.0588	0.1276
4	0.5510	0.0879	0.4631	0.4490	0.0466	0.0957
5	0.4408	0.0703	0.3704	0.5592	0.0393	0.0765
6	0.3673	0.0586	0.3087	0.6327	0.0344	0.0638
8	0.2755	0.0440	0.2315	0.7245	0.0283	0.0478
10	0.2204	0.0352	0.1852	0.7796	0.0246	0.0383
20	0.1102	0.0176	0.0926	0.8898	0.0173	0.0191
50	0.0441	0.0070	0.0370	0.9559	0.0129	0.0077
200	0.0110	0.0018	0.0093	0.9890	0.0107	0.0019

Table 7.1: Portfolio weights for CRRA investors ignoring interest rate fluctuations.

columns. The cash position is determined residually so that weights sum to one. The last two columns show the instantaneous expected rate of return and volatility of the portfolio. In Figure 7.1 the curved line shows the mean-variance efficient portfolios of risky assets, i.e. the combinations of expected returns and volatility that can be obtained by combining the bond and the stock. The straight line corresponds to the optimal portfolios for investors assuming constant investment opportunities with an interest rate equal to the long-term average.

Now let us look at investors that realize that interest rates vary over time and consequently alter their investment strategy (except for log-utility investors). First, we assume that the real short-term interest rate r_t follows the one-factor Vasicek model so that the analysis and results of Section 7.1.1 applies. The long-term average interest rate is $\bar{r} = 1.0\%$ and we take a short-rate volatility of $\sigma_r = 5\%$, which is also consistent with the U.S. historical estimate. We use the same values of the market prices of risk as above. We set the value of the mean reversion rate $\kappa = 0.4965$ so that the volatility of a 10-year zero-coupon bond according to the model is equal to the historical estimate of 10.0%. The current short rate is assumed to equal the long-term level, $r_t = \bar{r}$.

Let us first consider investors with utility of terminal wealth only. Their optimal portfolios are given by (7.9) and (7.10). Table 7.2 shows the optimal portfolios for CRRA investors with different combinations of risk aversion and investment horizon. The numbers under the column heading 'hedge' are $(1 - 1/\gamma)\sigma_r b(T)/b(10)$, which is the hedge demand for the 10-year zero-coupon bond which the investors are allowed to trade in. While the weight on the tangency portfolio and thus the stock is independent on the investment horizon, this is not true for the weight on the hedge portfolio and hence not true for the total weight on the bond and on cash. The ratio of the bond weight to the stock weight is shown in the column 'bond/stock'. The bond-stock ratio increases considerably with the risk aversion and, for investors with $\gamma > 1$, with the investment horizon. The investor with a horizon of T will want to hedge interest rate risk by investing in the T-period zero-coupon bond. That bond is replicated by a portfolio of $b(T)/b(10)$ units of the 10-year zero-coupon bond and a cash position. Since b is increasing in T, the hedge demand for

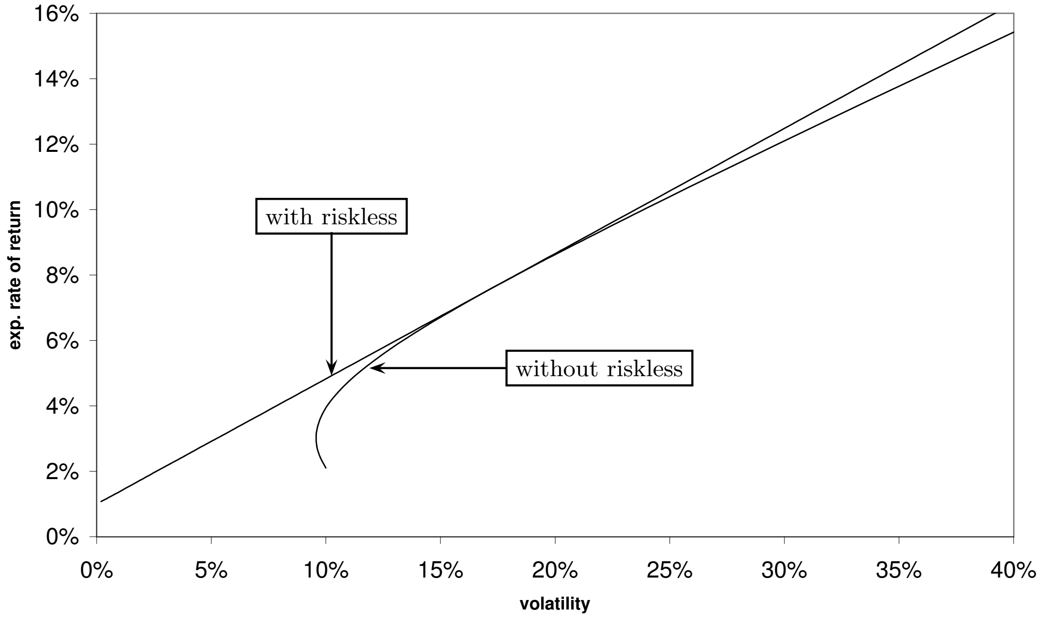


Figure 7.1: The mean-variance frontiers with and without the riskless asset.

the 10-year bond increases with the horizon T. It is important to emphasize that the portfolio weights on the bond and thus the bond/stock ratio will depend on the maturity (and payment schedule) of the bond, the investor is trading in. In particular, a recommendation of a particular bond weight or bond/stock ratio should always be accompanied by a specification of what bond the recommendation applies to.

Next, we consider investors with utility from terminal consumption and no utility from terminal wealth. In this case the hedge term in the bond weight (7.10) is replaced by (7.12). Now the hedge demand depends on the current interest rate level, which we assume is equal to the long-term average of 1%. Table 7.3 shows the optimal portfolios for investors with a 1-year and a 30-year horizon. We see the same overall picture as for investors with utility from terminal wealth only, but for a given investment horizon the hedge demand for bond and hence the bond/stock ratio are smaller with utility of consumption since the optimal bond for hedging has a smaller duration then the investment horizon.

Let us now compare the current mean/variance tradeoff chosen by different investors. As discussed above, CRRA investors that either have a zero (or very, very short) investment horizon or do not take interest rate risk into account will pick a portfolio that corresponds to a point on the straight line in Figure 7.2. This is the instantaneous mean-variance efficient frontier. Similarly, each of the other curves corresponds to the combinations chosen by CRRA investors with a given non-zero horizon who take interest rate risk into account. Since these curves lie to the right of the instantaneous mean-variance frontier, all these investors could obtain a higher instantaneous expected rate of return for the same volatility by choosing a different portfolio. But the long-term investors are willing to sacrifice some expected return in the short-term in order to hedge changes in interest rates and place themselves in a better position if interest rates should decline.

horizon	γ	tangency	hedge	bond	stock	$\frac{\text{bond}}{\text{stock}}$	cash	exp. return	volatility
$T = 1$	0.5	4.4079	-0.3941	0.3093	3.7045	0.08	-3.0138	0.2986	0.7551
	1	2.2039	0.0000	0.3517	1.8522	0.19	-1.2039	0.1565	0.3827
	2	1.1020	0.1970	0.3729	0.9261	0.40	-0.2990	0.0854	0.1979
	5	0.4408	0.3153	0.3856	0.3704	1.04	0.2439	0.0428	0.0908
	10	0.2204	0.3547	0.3899	0.1852	2.11	0.4249	0.0286	0.0592
	20	0.1102	0.3744	0.3920	0.0926	4.23	0.5154	0.0214	0.0467
$T = 5$	0.5	4.4079	-0.9229	-0.2195	3.7045	-0.06	-2.4850	0.2928	0.7442
	1	2.2039	0.0000	0.3517	1.8522	0.19	-1.2039	0.1565	0.3827
	2	1.1020	0.4615	0.6373	0.9261	0.69	-0.5634	0.0883	0.2094
	5	0.4408	0.7383	0.8087	0.3704	2.18	-0.1791	0.0474	0.1207
	10	0.2204	0.8306	0.8658	0.1852	4.67	-0.0510	0.0338	0.1010
	20	0.1102	0.8768	0.8943	0.0926	9.66	0.0130	0.0270	0.0950
$T = 10$	0.5	4.4079	-1.0000	-0.2966	3.7045	-0.08	-2.4079	0.2920	0.7429
	1	2.2039	0.0000	0.3517	1.8522	0.19	-1.2039	0.1565	0.3827
	2	1.1020	0.5000	0.6758	0.9261	0.73	-0.6020	0.0887	0.2112
	5	0.4408	0.8000	0.8703	0.3704	2.35	-0.2408	0.0481	0.1256
	10	0.2204	0.9000	0.9352	0.1852	5.05	-0.1204	0.0345	0.1074
	20	0.1102	0.9500	0.9676	0.0926	10.45	-0.0602	0.0278	0.1022
$T = 30$	0.5	4.4079	-1.0070	-0.3036	3.7045	-0.08	-2.4009	0.2919	0.7428
	1	2.2039	0.0000	0.3517	1.8522	0.19	-1.2039	0.1565	0.3827
	2	1.1020	0.5035	0.6794	0.9261	0.73	-0.6055	0.0888	0.2114
	5	0.4408	0.8056	0.8760	0.3704	2.37	-0.2464	0.0482	0.1261
	10	0.2204	0.9063	0.9415	0.1852	5.08	-0.1267	0.0346	0.1080
	20	0.1102	0.9567	0.9743	0.0926	10.52	0.0669	0.0278	0.1028

Table 7.2: Portfolio weights for CRRA investors who assume Vasicek interest rate dynamics and have utility from terminal wealth only.

Table 7.4 shows the optimal portfolios for investors with a constant relative risk aversion equal to 2, but with different investment horizons. Here we can clearly see the effect of the investment horizon on the optimal bond holdings and the bond/stock ratio. Relative to the extreme short-term investor, long-term investors have the same stock weight but shifts wealth from cash to bonds. If we look at the instantaneous risk/return trade-off, the longer-term investors choose more risky portfolios, i.e. they take on more short-term risk. But the main point is that long-term investors do not choose their portfolio according to the short-term risk/return trade-off.

Next, we want to investigate how sensitive the asset allocation choice is to the assumed interest rate model. We do that by computing the optimal portfolios when interest rates follow the CIR model (7.13). We want to make a reasonably fair comparison between the two models. For that purpose we choose $\sigma_r = 0.5$ in the CIR model so that the average short rate volatility is $\sigma_r \sqrt{\bar{r}} = 0.05$ as in the Vasicek model. We set $\bar{\lambda}_1 = 0.55$ and $\lambda_2 = 0.3666$ so that the model is consistent with the estimated mean stock and bond returns when $r = \bar{r}$ is used to compute the Sharpe ratios of

horizon	γ	tangency	hedge	bond	stock	$\frac{\text{bond}}{\text{stock}}$	cash	exp. return	volatility
$T = 1$	0.5	4.4079	-0.2253	0.4781	3.7045	0.1290	-3.1826	0.3005	0.7593
	1	2.2039	0.0000	0.3517	1.8522	0.1899	-1.2039	0.1565	0.3827
	2	1.1020	0.1114	0.2872	0.9261	0.3101	-0.2134	0.0845	0.1949
	5	0.4408	0.1787	0.2490	0.3704	0.6722	0.3805	0.0413	0.0835
	10	0.2204	0.2013	0.2365	0.1852	1.2766	0.5783	0.0269	0.0481
	20	0.1102	0.2126	0.2302	0.0926	2.4856	0.6772	0.0197	0.0324
$T = 30$	0.5	4.4079	-0.9624	-0.2590	3.7045	-0.0699	-2.4455	0.2924	0.7436
	1	2.2039	0.0000	0.3517	1.8522	0.1899	-1.2039	0.1565	0.3827
	2	1.1020	0.4425	0.6184	0.9261	0.6677	-0.5445	0.0881	0.2084
	5	0.4408	0.7241	0.7944	0.3704	2.1445	-0.1649	0.0473	0.1195
	10	0.2204	0.8228	0.8579	0.1852	4.6319	-0.0432	0.0337	0.1002
	20	0.1102	0.8731	0.8907	0.0926	9.6176	0.0167	0.0269	0.0946

Table 7.3: Portfolio weights for CRRA investors who assume Vasicek interest rate dynamics and have utility from terminal wealth only.

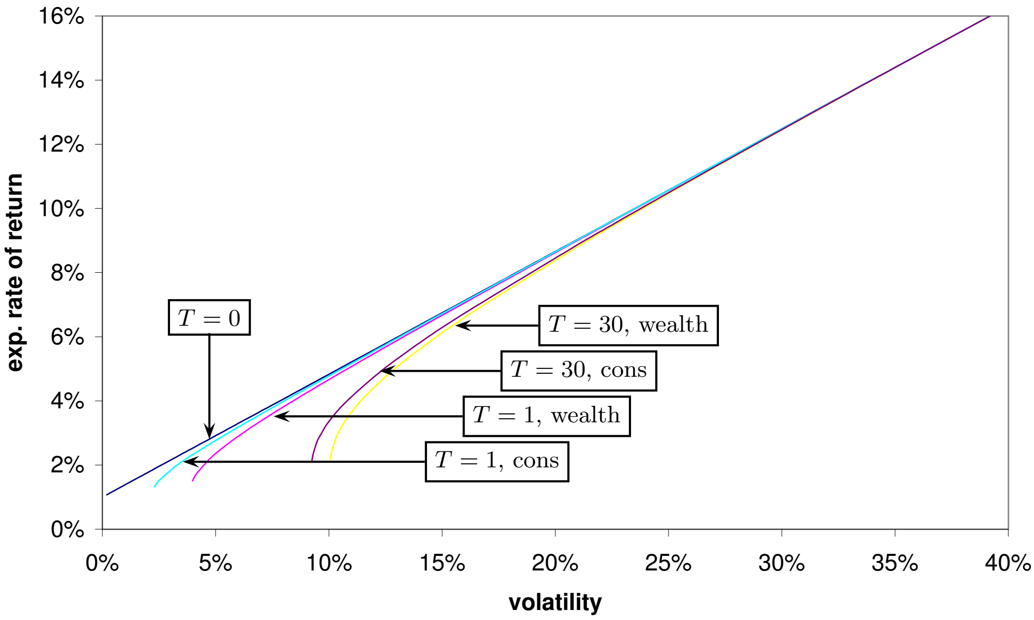


Figure 7.2: The optimal combinations of current expected rate of return and volatility for CRRA investors who assumes that interest rates follow the Vasicek model.

horizon	tangency	hedge	bond	stock	$\frac{\text{bond}}{\text{stock}}$	cash	exp. return	volatility
$T = 0$	1.1020	0	0.1758	0.9261	0.19	-0.1020	0.0832	0.1914
$T = 1$, wealth	1.1020	0.1970	0.3729	0.9261	0.40	-0.2990	0.0854	0.1979
$T = 5$, wealth	1.1020	0.4615	0.6373	0.9261	0.69	-0.5634	0.0883	0.2094
$T = 10$, wealth	1.1020	0.5000	0.6758	0.9261	0.73	-0.6020	0.0887	0.2112
$T = 30$, wealth	1.1020	0.5035	0.6794	0.9261	0.73	-0.6055	0.0888	0.2114
$T = 1$, cons.	1.1020	0.1114	0.2872	0.9261	0.3101	-0.2134	0.0845	0.1949
$T = 30$, cons.	1.1020	0.4425	0.6184	0.9261	0.6677	-0.5445	0.0881	0.2084

Table 7.4: Portfolio weights for investors with a constant relative risk aversion of $\gamma = 2$.

				Vasicek model			CIR model		
horizon	γ	tangency	stock	hedge	bond	cash	hedge	bond	cash
$T = 1$	0.5	4.4079	3.7045	-0.3941	0.3093	-3.0138	-0.6374	0.0660	-2.7705
	1	2.2039	1.8522	0.0000	0.3517	-1.2039	0	0.3517	-1.2039
	2	1.1020	0.9261	0.1970	0.3729	-0.2990	0.2482	0.4241	-0.3502
	5	0.4408	0.3704	0.3153	0.3856	0.2439	0.3653	0.4357	0.1939
	10	0.2204	0.1852	0.3547	0.3899	0.4249	0.3979	0.4331	0.3817
	20	0.1102	0.0926	0.3744	0.3920	0.5154	0.4129	0.4305	0.4769
$T = 30$	0.5	4.4079	3.7045	-1.0070	-0.3036	-2.4009	-1.0066	-0.3033	-2.4012
	1	2.2039	1.8522	0.0000	0.3517	-1.2039	0	0.3517	-1.2039
	2	1.1020	0.9261	0.5035	0.6794	-0.6055	0.5012	0.6771	-0.6032
	5	0.4408	0.3704	0.8056	0.8760	-0.2464	0.8012	0.8715	-0.2420
	10	0.2204	0.1852	0.9063	0.9415	-0.1267	0.9010	0.9362	-0.1214
	20	0.1102	0.0926	0.9567	0.9743	0.0669	0.9509	0.9685	-0.0611

Table 7.5: Portfolio weights with Vasicek or CIR dyanmics for CRRA investors with utility from terminal wealth only.

the bond market ($\lambda_1(r)$) and the stock market ($\psi(r)$ in (7.16). The mean reversion rate is set at $\kappa = 0.7994$ so that the volatility of a 10-year zero-coupon bond according to the model is equal to the historical estimate of 10.0%. The optimal portfolio in the CIR setting depends on the current interest rate level. In the computations we put this equal to the long-term average of 1%.

In Table 7.5 we list the optimal portfolios for investors with CRRA utility of terminal wealth both for the Vasicek and the CIR setting. We consider an investor with a 1-year horizon and an investor with a 30-year horizon. The stock weight is identical in the two models. The hedge demand for bonds and hence the total bond demand (and the cash position) do depend on the interest rate model, but the differences are relatively small. Although we have tried to keep the two interest rate models comparable, the yield curves of the two models are far from identical. The long-term yield is 1.04% in the Vasicek model and 1.60% in the CIR model. With a current short rate of 1%, the Vasicek yield curve is humped, whereas the CIR yield curve is uniformly increasing. Basically, it seems to be hard to align the two models for near-zero interest rates as in this example. The differences in portfolio weights should be evaluated with this in mind.

7.1.4 Other studies with stochastic interest rates

Brennan and Xia (2000) study a two-factor Vasicek interest rate model with utility from terminal wealth only.

Brennan, Schwartz, and Lagnado (1997) apply the two-factor Brennan-Schwartz interest rate dynamics in a model that also has stochastic dividends on stocks. They study the effect the length of the investment horizon has for an investor with utility from terminal wealth only. Due to the complexity of their model they must resort to numerical solution techniques.

Munk and Sørensen (2003) study the asset allocation problem when the term structure of interest rates evolve according to models in the Heath-Jarrow-Morton (HJM) class. As shown by Heath, Jarrow, and Morton (1992), any dynamic interest rate model is fully specified by the current term structure and the forward rate volatilities. Therefore the HJM modeling framework is natural when comparing the separate effects of the current term structure and the dynamics of the term structure on the optimal interest rate hedging strategy. Term structure models in the HJM class are not necessarily Markovian, but the class includes the well-known Markovian models such as the Vasicek model. To cover the non-Markovian models the authors apply the martingale approach to solve the utility maximization problem instead of the dynamic programming approach. Within the HJM framework one may fix the current yield curve and vary its future dynamics to gauge the effect of the interest rate dynamics. As in all term structure models one can fix the dynamics and vary the initial yield curve (for absolute pricing models, such as the Vasicek and CIR models, not all initial yield curves are possible). The paper compares the optimal portfolio and consumption strategies for a standard one-factor Vasicek and a three-factor model where the term structure can exhibit three kinds of changes: A parallel shift, a slope change, and a curvature change. The authors find that the form of the initial term structure is of crucial importance for the certainty equivalents of future consumption and, hence, important for the relevant interest rate hedge, while the specific dynamics of the term structure is of minor importance. Of course, further studies of this kind is needed to find out whether this conclusion is generally valid.

Further references: Campbell and Viceira (2001), Wachter (1999)

7.2 Stochastic excess returns

Several empirical studies provide evidence of mean reversion in stock returns so that expected stock returns are high after a period of low realized returns and *vice versa*. Some recent papers have studied the implications for portfolio decisions of this deviation from the traditional setting with constant investment opportunities. Both Kim and Omberg (1996) and Wachter (2002b) obtain closed-form expressions for the optimal investment strategy in a set-up with a constant riskfree interest rate r and a single risky asset (representing the stock market) with price P_t evolving as

$$dP_t = P_t \left[(r + \sigma\lambda_t) \; dt + \sigma \, dz_t \right], \tag{7.23}$$

where the volatility σ is assumed to be a positive constant, but the market price of risk λ_t follows a mean-reverting process. Note that in this setting the market price of risk is identical to the Sharpe ratio of the stock. Kim and Omberg (1996) consider an investor with a CRRA utility of terminal wealth only, which allows them to let λ_t have an undiversifiable risk component. On the other hand, Wachter (2002b) considers a time-separable CRRA utility function of consumption, so to obtain explicit solutions she assumes that the market price of risk is perfectly (negatively) correlated with the price level. Wachter argues that the assumption of a correlation of -1 is empirically not unreasonable. To allow for non-perfect correlation we write the dynamics of λ as

$$d\lambda_t = \kappa\left[\bar{\lambda} - \lambda_t\right] dt + \rho\sigma_\lambda\, dz_t + \sqrt{1-\rho^2}\sigma_\lambda\, d\hat{z}_t. \tag{7.24}$$

All constants are assumed positive, except the correlation parameter ρ. The market price of risk is assumed to follow an Ornstein-Uhlenbeck process with long-term average $\bar{\lambda}$, mean reversion speed κ, and volatility σ_λ. It can be shown that the future stock price P_T with the above specification is given by

$$\begin{aligned} P_T = P_t \exp\Bigg\{ &\left(r - \frac{\sigma^2}{2} + \sigma\bar{\lambda}\right)(T-t) + \sigma b(T-t)\left(\lambda_t - \bar{\lambda}\right) \\ &+ \sigma \int_t^T \left(1 + \rho\sigma_\lambda b(T-s)\right) dz_s + \sigma\sigma_\lambda\sqrt{1-\rho^2}\int_t^T b(T-s)\, d\hat{z}_s \Bigg\}, \end{aligned} \tag{7.25}$$

where $b(\tau) = (1 - e^{-\kappa\tau})/\kappa$. Consequently, P_T is lognormally distributed (given P_t):

$$\begin{aligned} \ln\frac{P_T}{P_t} \sim N\Bigg(&\left(r - \frac{\sigma^2}{2} + \sigma\bar{\lambda}\right)(T-t) + \sigma b(T-t)\left(\lambda_t - \bar{\lambda}\right), \\ &\left(1 + \frac{2\rho\sigma_\lambda}{\kappa} + \frac{\sigma_\lambda^2}{\kappa^2}\right)(T-t) - \left(\frac{2\rho\sigma_\lambda}{\kappa} + \sigma_\lambda^2\right) b(T-t) - \frac{\sigma_\lambda^2}{2\kappa} b(T-t)^2 \Bigg). \end{aligned} \tag{7.26}$$

This is a model where the market price of risk is the only state variable, i.e. $x = \lambda$. Since λ is an affine function of itself, we see from our general analysis in Sections 6.4–6.5 that we need an exponential-quadratic expression for the g-function.

Let us first consider the case with CRRA utility from terminal wealth only. In that case we can allow for any correlation $\rho \in [-1, 1]$. In the notation of Section 6.5 we have

$$\begin{aligned} r_0 &= r, & r_1 &= r_2 = 0, \\ m_0 &= \kappa\bar{\lambda}, & m_1 &= -\kappa, \\ \Lambda_0 &= \Lambda_1 = 0, & \Lambda_2 &= 1, \\ K_0 &= 0, & K_1 &= \rho\sigma_\lambda, \\ v^\top v &= \rho^2\sigma_\lambda^2, & \hat{v}^2 &= (1-\rho^2)\sigma_\lambda^2. \end{aligned}$$

Equation (6.56) yields the ordinary differential equation

$$A_3'(\tau) = \sigma_\lambda^2(\rho^2 + \gamma(1-\rho^2))A_3(\tau)^2 - 2\left(\kappa - \frac{1-\gamma}{\gamma}\rho\sigma_\lambda\right)A_3(\tau) + \frac{1-\gamma}{\gamma^2},$$

which we must solve with the initial condition $A_3(0) = 0$. Define the auxiliary parameters

$$\begin{aligned} \bar{\kappa} &= \kappa - \frac{1-\gamma}{\gamma}\rho\sigma_\lambda, \\ q &= \bar{\kappa}^2 - \sigma_\lambda^2\left(\rho^2 + \gamma(1-\rho^2)\right)\frac{1-\gamma}{\gamma^2}. \end{aligned}$$

Assuming that q is positive[1], the solution to the differential equation is

$$A_3(\tau) = \frac{1-\gamma}{\gamma^2}\frac{1 - e^{-2\sqrt{q}\tau}}{2\sqrt{q} - \left(\sqrt{q} - \bar{\kappa}\right)\left(1 - e^{-2\sqrt{q}\tau}\right)}. \tag{7.27}$$

[1]This condition will be satisfied except for "extreme" combinations of κ, σ_λ, ρ, and γ. A discussion of the solution if this condition is not satisfied can be found in Kim and Omberg (1996).

Equation (6.55) yields the equation

$$A_2'(\tau) = \kappa\bar{\lambda}A_3(\tau) - \bar{\kappa}A_2(\tau) + \sigma_\lambda^2 \left(\rho^2 + \gamma(1-\rho^2)\right) A_2(\tau)A_3(\tau),$$

which, with the initial condition $A_2(0) = 0$ and the expression for A_3 found above, has the unique solution

$$A_2(\tau) = \frac{1-\gamma}{\gamma^2} \frac{\kappa\bar{\lambda}\left(1 - e^{-\sqrt{q}\tau}\right)^2}{\sqrt{q}\left[2\sqrt{q} - \left(\sqrt{q} - \bar{\kappa}\right)\left(1 - e^{-2\sqrt{q}\tau}\right)\right]}. \tag{7.28}$$

Finally, A_1 can be determined from (6.57):

$$\begin{aligned}
A_1(\tau) &= \frac{1-\gamma}{\gamma} r\tau + \kappa\bar{\lambda}\int_0^\tau A_2(s)\,ds + \frac{1}{2}\sigma_\lambda^2\left(\rho^2 + \gamma(1-\rho^2)\right)\int_0^\tau A_2(s)^2\,ds \\
&\quad + \frac{1}{2}\sigma_\lambda^2 \int_0^\tau A_3(s)\,ds \\
&= \frac{1-\gamma}{\gamma} r\tau + \frac{1-\gamma}{2\gamma^2}\left(\frac{\kappa^2\bar{\lambda}^2}{q} + \frac{\sigma_\lambda^2}{\sqrt{q} + \bar{\kappa}}\right)\tau \\
&\quad + \frac{1-\gamma}{2\gamma^2}\frac{\kappa^2\bar{\lambda}^2}{q\sqrt{q}} \frac{\left(\sqrt{q} - 2\bar{\kappa}\right)e^{-2\sqrt{q}\tau} + 4\bar{\kappa}e^{-\sqrt{q}\tau} - 2\bar{\kappa} - \sqrt{q}}{2\sqrt{q} - \left(\sqrt{q} - \bar{\kappa}\right)\left(1 - e^{-2\sqrt{q}\tau}\right)} \\
&\quad - \frac{1}{2[\rho^2 + \gamma(1-\rho^2)]} \ln\left(\frac{2\sqrt{q} - \left(\sqrt{q} - \bar{\kappa}\right)\left(1 - e^{-2\sqrt{q}\tau}\right)}{2\sqrt{q}}\right),
\end{aligned} \tag{7.29}$$

where the last equality follows from long and tedious calculations.

From Theorem 6.6 we have that for an investor with CRRA utility of terminal wealth only, the fraction of wealth optimally invested in stock is

$$\Pi(W, \lambda, t) = \frac{1}{\gamma}\frac{\lambda}{\sigma} + \frac{\rho\sigma_\lambda}{\sigma}\left(A_2(T-t) + A_3(T-t)\lambda\right). \tag{7.30}$$

It can be shown that for $\gamma > 1$, $A_2(\tau)$ and $A_3(\tau)$ are negative and decreasing. If the current value of the market price of risk is positive and the correlation is negative (consistent with empirical observations), it follows that the hedge term of the optimal portfolio is positive and increasing with the horizon of the investor. An investor with a long horizon should therefore invest a larger fraction of wealth in stocks than an investor with the same risk aversion, but a shorter horizon. This is consistent with typical recommendations of investment advisors.

With utility from intermediate consumption and possibly terminal wealth, we must assume that either $\rho = 1$ or $\rho = -1$. We will stick to the latter, more realistic case. The restriction $\rho = -1$ affects all the functions A_1, A_2, and A_3 due to the presence of ρ in $\bar{\kappa}$ and q. For notational simplicity let us consider an investor with utility stemming only from intermediate consumption, i.e. $\varepsilon_2 = 0$. From Theorem 6.7, we get that the optimal investment strategy is

$$\Pi(W, \lambda, t) = \frac{1}{\gamma}\frac{\lambda}{\sigma} - \frac{\sigma_\lambda}{\sigma}\frac{\int_t^T e^{h(\lambda, s-t)}\left(A_2(s-t) + A_3(s-t)\lambda\right)\,ds}{\int_t^T e^{h(\lambda, s-t)}\,ds}, \tag{7.31}$$

where

$$h(\lambda, \tau) = -\frac{\delta}{\gamma}\tau + A_1(\tau) + A_2(\tau)\lambda + \frac{1}{2}A_3(\tau)\lambda^2,$$

and we must insert $\rho = -1$ in the expressions of the A_i's. Again it can be shown that, for $\gamma > 1$ and $\lambda > 0$, the hedging component is positive and increasing with the time horizon T. With intermediate consumption the horizon effect on the stock investment is dampened relative to the case with utility from terminal wealth only since the "effective" investment horizon is lower than T.

The optimal consumption rate is

$$C(W, \lambda, t) = \left(\int_t^T e^{h(\lambda, s-t)} \, ds \right)^{-1} W. \tag{7.32}$$

It can be shown that the consumption/wealth ratio is increasing in λ when $\lambda > 0$ and $\gamma > 1$. To see this note that the derivative of the wealth/consumption ratio with respect to λ is $\int_t^T e^{h(\lambda,s-t)} \left(A_2(s-t) + A_3(s-t)\lambda \right) ds$, which also enters the hedging demand. In fact, whenever the hedging demand is positive, the wealth/consumption ratio will be decreasing in λ, and the consumption/wealth ratio will therefore be increasing in λ. The intuition for this result is as follows: An increase in λ indicates better future investment opportunities. This gives an income effect that induces higher current consumption. On the other hand, investments are then more profitable – there is a substitution effect. With $\gamma > 1$, the income effect dominates. To keep consumption stable across states, the investor must choose a portfolio which gives positive returns in states with relatively bad future investment opportunities, i.e. low λ. Since with $\rho = -1$ stocks have high returns exactly when λ is low, the investor will hold more stocks relative to the case with constant investment opportunities.

Further references: Barberis (2000)

7.3 Stochastic volatility

As discussed in Section 6.3, stochastic volatility is only an issue to the extent that it affects the market prices of risk. It seems natural that expected excess rate of returns increase with volatility so that an assumption on a constant market price of risk is more realistic than an assumption on a constant excess rate of return as assumed by some authors.

Liu (1999) considers some examples involving a stock having a stochastic volatility and an expected excess rate of return which is not proportional to the level of the volatility, i.e. the market price of risk varies with volatility. His examples are within the framework that allows explicit solutions. A simple model with a single risky stock and a constant interest rate is the following:

$$\frac{dP_t}{P_t} = (r + \eta V_t) \, dt + \sqrt{V_t} \, dz_{1t}, \tag{7.33}$$

$$dV_t = \kappa \left(\bar{V} - V_t \right) dt + \rho \sigma_V \sqrt{V_t} \, dz_{1t} + \sqrt{1-\rho^2} \sigma_V \sqrt{V_t} \, dz_{2t}, \tag{7.34}$$

where V_t is the instantaneous variance rate of the stock, i.e. the square of the volatility. The market price of risk, i.e. the Sharpe ratio of the stock, is $\eta \sqrt{V_t}$. The market is incomplete since the volatility risk is not perfectly hedgeable. The reader can verify that this model fits into the affine framework so that Theorem 6.4 will give the optimal investment strategy for an investor with CRRA utility of terminal wealth only.

Chacko and Viceira (2002) consider a quite spurious model with stochastic volatility that does not fit into the cases where we have explicit solutions. They find explicit, *approximate* solutions.

See also Kraft (2003).

7.4 Exercises

EXERCISE 7.1 Consider a financial market where the only two assets traded are (1) a bank account with a rate of return of r_t and (2) a risky asset with price P_t following the geometric Brownian motion,

$$dP_t = P_t \left[\mu\, dt + \sigma\, dz_t\right].$$

The short-term interest rate is assumed to follow a Vasicek process:

$$dr_t = \kappa \left[\bar{r} - r_t\right] dt + \rho\sigma_r\, dz_t + \sqrt{1-\rho^2}\sigma_r\, d\hat{z}_t.$$

(a) Describe the model!

We look at an investor with CRRA utility of terminal wealth only,

$$J(W, r, t) = \sup_{\pi} \mathrm{E}_{W,r,t}\left[\frac{W_T^{1-\gamma}}{1-\gamma}\right],$$

where the process π denotes the fraction of wealth invested in the risky asset.

(b) State the HJB equation corresponding to this problem.

(c) Find the first-order condition for π.

(d) Show that the indirect utility function is of the form

$$J(W, r, t) = \frac{1}{1-\gamma}\left[\exp\left\{A_1(T-t) + A_2(T-t)r + \frac{1}{2}A_3(T-t)r^2\right\}\right]^{\gamma} W^{1-\gamma}.$$

What can you say about the functions A_i?

(e) Find the optimal portfolio strategy. Compare it with the solution for constant r.

EXERCISE 7.2 Consider an economy with a single agent. The agent owns a production plant that generates units of the consumption good of the economy. The agent can choose to withdraw consumption goods from the production or reinvest them in the production process. The productivity of her plant depends on a state variable Y_t that follows the process

$$dY_t = (b - \kappa Y_t)\, dt + k\sqrt{Y_t}\, dz_t, \quad Y_0 = y,$$

where b, κ and k are positive constants with $2b > k^2$. Let $c_t \geq 0$ denote the rate by which the agent withdraws consumption goods from the production plant and let X_t^c be the value of the plant at time t given the consumption process c. We assume that

$$dX_t^c = \left(X_t^c h Y_t - c_t\right) dt + X_t^c \epsilon\sqrt{Y_t}\, dz_t, \quad X_0^c = x,$$

where h and ϵ are positive constants with $h > \epsilon^2$. The agent has a log utility of consumption over her life-time T, so that the indirect utility function is

$$V(x, y, t) = \sup_{c} \mathrm{E}_{x,y,t}\left[\int_t^T e^{-\delta(s-t)} \ln c_s\, ds\right].$$

(a) State the HJB equation corresponding to the problem and find the first-order condition for the optimal consumption rate.

(b) Verify that the function

$$V(x, y, t) = A_1(t) \ln x + A_2(t) y + A_3(t)$$

satisfies the HJB equation and find ordinary differential equations that the functions A_1, A_2 and A_3 must solve. Show that $A_1(t) = \frac{1}{\delta}(1 - e^{-\delta(T-t)})$. Find an explicit expression for the optimal consumption rate, c_t^*.

(c) We know from the martingale approach that the state-price density ζ_t satisfies $\psi\zeta_t = u'(c_t^*, t)$, where ψ is a constant, and where $u(c,t) = e^{-\delta t}\ln c$ in our case. Use this and the expression for optimal consumption to show that

$$\zeta_t = \frac{1}{\psi} e^{-\delta t} \frac{A_1(t)}{X_t^*},$$

where X_t^* is the optimal value of the production plant, i.e. $X_t^* = X_t^{c^*}$. Apply Itô's Lemma in order to find the dynamics of ζ_t.

(d) We also know that

$$d\zeta_t = -\zeta_t \left[r_t \, dt + \lambda_t \, dz_t \right],$$

where r_t is the short-term interest rate. Conclude that $r_t = (h - \epsilon^2)Y_t$. Show that the dynamics of r_t is on the form

$$dr_t = \kappa[\bar{r} - r_t] \, dt + \sigma_r \sqrt{r_t} \, dz_t,$$

where κ, $\bar{r}$ and σ_r are positive constants. Appreciate this result!

Chapter 8

Non-financial risks

8.1 Labor income

In the general description of the continuous-time model in Section 4.4 we allowed for the case where the agent receives income from non-financial sources at a rate y_t. But in all the problems studied until now we have assumed $y \equiv 0$. We shall refer to income from non-financial sources as labor income although this may in general include gifts, welfare payments, etc. In this section we will study the influence of labor income on optimal portfolio and consumption choice. From (4.8) wealth evolves as

$$dW_t = W_t \left[r_t + \pi_t^\top \sigma_t \lambda_t \right] dt + \left[y_t - c_t \right] dt + W_t \pi_t^\top \sigma_t \, dz_t.$$

We take a Markovian framework so that we can apply the dynamic programming approach. We consider both the case where the labor income rate is exogenously given and the case where the labor supply decision of the agent is taken into account.

8.1.1 Exogenous labor income

Most studies of the effect of labor income on consumption and portfolio choice assume an exogenous process of the labor income rate such as

$$dy_t = y_t \left[\alpha(y_t, t) \, dt + \xi(y_t, t)^\top \, dz_t + \hat{\xi}(y_t, t) \, d\hat{z}_t \right].$$

If $\hat{\xi} \neq 0$, the income risk is not fully hedgeable in the financial market, which seems to be the realistic situation. However, this is a more difficult problem to analyze, so let us first look at the complete market case.

In the complete market case where $\hat{\xi} \equiv 0$, the income stream is fully hedgeable and can be valued as any financial asset. The time t value of the income stream $(y_s)_{s \in [t,T]}$ must be

$$\begin{aligned} H(x, y, t) &= \mathrm{E}^{\mathbb{Q}}_{x,y,t} \left[\int_t^T e^{-\int_t^s r(x_u) \, du} y_s \, ds \right] \\ &= \mathrm{E}_{x,y,t} \left[\int_t^T \exp \left\{ - \int_t^s r(x_u) \, du - \int_t^s \lambda(x_u)^\top \, dz_u - \frac{1}{2} \int_t^s \lambda(x_u)^\top \lambda(x_u) \, du \right\} y_s \, ds \right], \end{aligned}$$

where $\mathbb{Q}$ is the risk-neutral probability measure, and x is a state variable affecting the short-term interest rate r and the market price of risk vector λ. We refer to $H(x, y, t)$ as the **human wealth** of the agent at time t. In this situation we can think of the agent "selling" his future income at the financial market in the exchange of the payment $H(x, y, t)$ so that he has a total wealth of $W + H(x, y, t)$ to invest. He will invest in a financial portfolio such that the riskiness of his total position of financial investments and labor income is similar to the riskiness of his optimal financial portfolio in the absence of labor income.

For example, consider the classical setting with a constant interest rate r and a constant market price of risk λ. We have from Theorem 5.2 that without labor income it is optimal for a CRRA utility investor to invest the proportions

$$\Pi(W, t) = \frac{1}{\gamma} \left(\sigma^{\top}\right)^{-1} \lambda$$

in the risky assets. With the optimal investment strategy the wealth will evolve as

$$dW_t = \dots\, dt + W_t \frac{1}{\gamma} \lambda^{\top}\, dz_t,$$

cf. (5.17). An investor with a human wealth of

$$H(y, t) = \mathrm{E}_{y,t}\left[\int_t^T \exp\left\{-r(s-t) - \lambda^{\top}(z_s - z_t) - \frac{1}{2}\lambda^{\top}\lambda(s-t)\right\} y_s\, ds\right]$$

has a total wealth of $W_t + H(y, t)$, where W_t still denotes the financial wealth. Such an investor will seek to invest such that the dynamics of total wealth is

$$d\left(W_t + H(y_t, t)\right) = \dots\, dt + \left(W_t + H(y_t, t)\right) \frac{1}{\gamma} \lambda^{\top}\, dz_t.$$

By Itô's Lemma, the dynamics of human wealth is

$$dH(y_t, t) = \dots\, dt + H_y(y_t, t) y_t \xi(y_t, t)^{\top}\, dz_t.$$

So the dynamics of the optimally invested financial wealth must be given by

$$\begin{aligned} dW_t &= \dots\, dt + \left(W_t + H(y_t, t)\right) \frac{1}{\gamma} \lambda^{\top}\, dz_t - H_y(y_t, t) y_t \xi(y_t, t)^{\top}\, dz_t \\ &= \dots\, dt + \left[\left(W_t + H(y_t, t)\right) \frac{1}{\gamma} \lambda \; - H_y(y_t, t) y_t \xi(y_t, t)\right]^{\top} dz_t. \end{aligned}$$

This is the case for a portfolio π_t that satisfies

$$W_t \pi_t^{\top} \sigma_t = \left[\left(W_t + H(y_t, t)\right) \frac{1}{\gamma} \lambda - H_y(y_t, t) y_t \xi(y_t, t)\right]^{\top},$$

i.e. the optimal fractions of financial wealth invested in the risky financial assets are given by the vector $\pi_t = \Pi(W_t, y_t, t)$, where

$$\Pi(W, y, t) = \frac{1}{\gamma} \frac{W + H(y, t)}{W} \left(\sigma_t^{\top}\right)^{-1} \lambda - \frac{H_y(y, t) y_t}{W} \left(\sigma_t^{\top}\right)^{-1} \xi(y, t). \tag{8.1}$$

The indirect utility function of the investor with constant relative risk aversion γ is

$$J(W, y, t) = \frac{1}{1-\gamma} g(t)^{\gamma} \left(W + H(y, t)\right)^{1-\gamma}, \tag{8.2}$$

where $g(t)$ is a deterministic function of time. These findings can be verified by setting up the appropriate Hamilton-Jacobi-Bellman equation and substituting the above expression for $J(W, y, t)$.

In particular, if labor income is deterministic, ξ will be zero, and the optimal portfolio reduces to

$$\Pi(W, y, t) = \frac{1}{\gamma} \frac{W + H(y,t)}{W} \left(\sigma_t^{\top}\right)^{-1} \lambda. \tag{8.3}$$

Since human wealth is increasing in the horizon T and decreasing as time t goes, we see that, under these restrictive assumptions, it is optimal for younger investors to have a higher fraction of financial wealth invested in risky assets than more mature investors. This is consistent with popular investment advice – but not with the explanation that usually accompanies the advice!

Let us look at a small numerical example with a single risky financial asset representing the stock index.[1] Consider an investor with a financial wealth of 500,000 dollars and a risk aversion of $\gamma = 2$. Assume that the riskless interest rate is $r = 4\%$, the expected rate of return on stocks is $\mu = 10\%$, and the volatility of the stock is $\sigma = 20\%$. (The market price of risk is $\lambda = (\mu - r)/\sigma = 0.3$.) It is optimal for the investor to have 75% of his total wealth invested in stocks and 25% in the riskless asset (bonds). If he has no labor income, this is also the optimal allocation of his financial wealth.

Let us first assume that the investor has a labor income with a present value of 500,000 dollars and, hence, a total wealth of one million. It is then optimal to have a total position of 750,000 dollars in stocks and 250,000 dollars in the riskless asset. How the financial wealth is to be allocated depends on the riskiness of his labor income. In Table 8.1 we consider three cases:

(a) If the labor income is completely riskless, it is equivalent to a position of 0 dollars in stocks and 500,000 dollars in the riskless asset. To obtain the desired overall riskiness, he has to allocate his financial wealth of 500,000 by investing 750,000 dollars in stocks and -250,000 dollars in the riskless asset. This corresponds to a stock investment 150% of the financial wealth, financed in part by borrowing 50% of the financial wealth. The certain labor income corresponds to the returns of a riskless investment. Hence the financial wealth (and more) has to be invested in stocks to achieve the wanted balance between risky and riskless returns.

(b) If the labor income is quite risky and corresponds to an equal combination of stocks and bonds, the entire financial wealth (100%) is to be invested in stocks.

(c) If the labor income is extremely risky and corresponds to a 100% investment in stocks, the financial wealth is to be split equally between stocks and bonds.

Clearly, the optimal allocation of financial wealth is highly dependent on the risk profile of labor income.

Next, let us consider an investor with the same risk aversion, but a longer investment horizon and, consequently, a higher capitalized labor income, namely 1,500,000 dollars. Table 8.2 shows the allocation of the financial wealth that is needed to obtain the desired 75-25 split between risky and riskless returns. Comparing with Table 8.1 we see that the younger investor in Table 8.2 will have a significantly higher fraction of financial wealth invested in stocks than the older investor in Table 8.1, except for the case where the income is extremely uncertain. The optimal stock weight in the portfolio is clearly depending on the investment horizon.

According to empirical studies, the correlation between labor income and stock prices is very small for most individuals. In that case, labor income resembles a riskless investment more than a

[1] See also Jagannathan and Kocherlakota (1996) for a related example.

	Stock investment		Bond investment	
Riskless income	0	(0%)	500,000	(100%)
Financial inv.	750,000	(150%)	-250,000	(-50%)
Total position	750,000	(75%)	250,000	(25%)
Quite risky income	250,000	(50%)	250,000	(50%)
Financial inv.	500,000	(100%)	0	(0%)
Total position	750,000	(75%)	250,000	(25%)
Very risky income	500,000	(100%)	0	(0%)
Financial inv.	250,000	(50%)	250,000	(50%)
Total position	750,000	(75%)	250,000	(25%)

Table 8.1: Investments with a relatively short horizon. The table shows the optimal investment strategy for three types of labor income. The financial wealth is 500,000 and the capitalized labor income is 500,000 corresponding to a relatively short investment horizon.

	Stock investment		Bond investment	
Riskless income	0	(0%)	1,500,000	(100%)
Financial inv.	1,500,000	(300%)	-1,000,000	(-200%)
Total position	1,500,000	(75%)	500,000	(25%)
Quite risky income	750,000	(50%)	750,000	(50%)
Financial inv.	750,000	(150%)	-250,000	(-50%)
Total position	1,500,000	(75%)	500,000	(25%)
Very risky income	1,500,000	(100%)	0	(0%)
Financial inv.	0	(0%)	500,000	(100%)
Total position	1,500,000	(75%)	500,000	(25%)

Table 8.2: Investments with a relatively long horizon. The table shows the optimal investment strategy for three types of labor income. The financial wealth is 500,000 and the capitalized labor income is 1,500,000 corresponding to a relatively long investment horizon.

stock investment, and the fraction of financial wealth invested in stocks should increase with the length of the investment horizon. However, for some investors the labor income may be highly correlated with the stock market and in that case the weight of stocks in the financial portfolio should decrease with the length of the horizon.

Note that although the labor income of a given individual is not significantly correlated with the overall stock market, it may be correlated with a specific stock. One could imagine that the labor income of an employee of a corporation was positively correlated with the price of the company's stocks and maybe also with stock prices of other companies in the same industry. If this is true, the labor income will to some extent replace a financial investment in these stocks. Consequently, the individual should invest less of his financial wealth in these stocks. Following this line of thought, a pension fund with members in a given industry should perhaps *underinvest* in the stocks of the corporations in which the members work - simply to give the members a better diversified total position.

As seen in the example, the optimal strategy outlined above may involve extensive borrowing of young investors that anticipate high future income rates. In practice, investors cannot actually sell their future income stream as slavery is forbidden these days. Moreover, young investors will find it extremely difficult to borrow substantive amounts for risky stock investments putting up only anticipated future income as collateral. This can be explained by the moral hazard and adverse selection features of labor income. In reality the income rate is not exogenously given, but reflects the abilities and the effort of the investor.

Some models take these problems partially into account by still assuming an exogenous income process, but restrict the agent to consumption and investment strategies that have the property that financial wealth W_t always stays positive. The future income stream will then have a lower value than in the unrestricted, complete market case. See Duffie and Zariphopoulou (1993), Duffie, Fleming, Soner, and Zariphopoulou (1997), Koo (1998), and Munk (2000). For example, Duffie, Fleming, Soner, and Zariphopoulou (1997) and Munk (2000) study the case with a single risky asset with price process

$$dP_t = P_t \left[\mu \, dt + \sigma \, dz_t\right],$$

constant r, μ, and σ, and where the income rate follows the geometric Brownian motion

$$dy_t = y_t \left[\alpha \, dt + \rho \sigma_y \, dz_t + \sqrt{1-\rho^2}\sigma_y \, d\hat{z}_t\right].$$

Here ρ is the correlation between the asset price and the labor income. The agent must keep financial wealth positive, $W_t > 0$, so that she faces a liquidity constraint. Furthermore, she faces undiversifiable income risk. The numerical results of Munk (2000) show that the implicit value the agent associates with her income stream can be considerably less than without the liquidity constraint and the undiversifiable part of the income risk, especially if she has a high preference for current consumption and a low current financial wealth. The results indicate that the reduction in human wealth is mainly due to the liquidity constraint, while the undiversifiability is of minor importance.

8.1.2 Endogenous labor supply and income

Bodie, Merton, and Samuelson (1992) endogenize the labor supply decision of the agent. Let us look at a version of their model. Let ω_t denote the wage rate, which is assumed to follow the geometric Brownian motion

$$d\omega_t = \omega_t \left[m\, dt + v^\top\, dz_t \right]. \tag{8.4}$$

In particular, the wage rate is spanned by the financial securities traded. Let φ_t denote the fraction of time working so that the total labor income over the interval $[t, t+dt]$ is $\varphi_t \omega_t\, dt$. Assuming a constant interest rate and a constant market price of risk, the wealth of the investor will then follow

$$dW_t = \left(rW_t + W_t \pi_t^\top \sigma \lambda - c_t + \varphi_t \omega_t \right) dt + W_t \pi_t^\top \sigma_t\, dz_t.$$

Assume a Cobb-Douglas type utility of consumption and leisure,

$$u(c, \varphi) = \frac{1}{1-\gamma} \left[c^\theta (1-\varphi)^{1-\theta} \right]^{1-\gamma},$$

where θ is a constant between 0 and 1. The indirect utility function is now defined as

$$J(W, \omega, t) = \sup_{c,\pi,\varphi} \mathrm{E}_{W,\omega,t} \left[\int_t^T e^{-\delta(s-t)} \frac{1}{1-\gamma} \left[c_s^\theta (1-\varphi_s)^{1-\theta} \right]^{1-\gamma} ds \right].$$

Upon substitution into the HJB equation associated with this problem, it can be verified after long and tedious calculations that the indirect utility function is given in closed-form by

$$J(W, \omega, t) = \frac{1}{1-\gamma} \theta^{\theta(1-\gamma)} (1-\theta)^{(1-\theta)(1-\gamma)} G(t) \omega^{-(1-\theta)(1-\gamma)} \left(W + \omega F(t) \right)^{1-\gamma}, \tag{8.5}$$

where

$$\begin{aligned}
G(t) &= \frac{1}{k} \left(1 - e^{-k(T-t)} \right), \\
F(t) &= \frac{1}{r - m + v^\top \lambda} \left(1 - e^{-(r-m+v^\top\lambda)(T-t)} \right), \\
k &= \frac{\delta}{\gamma} - r\frac{1-\gamma}{\gamma} - \frac{1-\gamma}{2\gamma^2} \lambda^\top \lambda + \frac{1-\gamma}{\gamma}(1-\theta) \left[m + \frac{1-\gamma}{\gamma} v^\top \lambda - \frac{1}{2\gamma} \left(1 - \theta(1-\gamma) \right) v^\top v \right].
\end{aligned}$$

The optimal strategies are $c_t^* = C(W_t, \omega_t, t)$, $\varphi_t^* = \Phi(W_t, \omega_t, t)$, and $\pi_t^* = \Pi(W_t, \omega_t, t)$, where

$$C(W, \omega, t) = \frac{\theta}{G(t)} \left(W + \omega F(t) \right), \tag{8.6}$$

$$\Phi(W, \omega, t) = 1 - \frac{1-\theta}{G(t)} \frac{W + \omega F(t)}{\omega}, \tag{8.7}$$

$$\Pi(W, \omega, t) = \frac{1}{\gamma} \frac{W + \omega F(t)}{W} \left(\sigma^\top \right)^{-1} \lambda - \frac{(1-\theta)(1-\gamma)}{\gamma} \frac{W + \omega F(t)}{W} \left(\sigma^\top \right)^{-1} v - \frac{F(t)\omega}{W} \left(\sigma^\top \right)^{-1} v. \tag{8.8}$$

Here $\omega_t F(t)$ denotes the time t value of the maximum labor income that the agent can receive. To see this note that the future wage rate is

$$\omega_s = \omega_t \exp\left\{ \left(m - \frac{1}{2} v^\top v \right)(s-t) + v^\top (z_s - z_t) \right\}.$$

Working at a maximum rate, $\varphi_s \equiv 1$ for all $s \in [t, T]$, the time t value of future labor income is

$$\begin{aligned}
\mathrm{E}_{\omega,t}&\left[\int_t^T \exp\left\{-r(s-t) - \lambda[z_s - z_t] - \frac{1}{2}\lambda^2(s-t)\right\}\omega_s\, ds\right] \\
&= \omega_t \int_t^T \mathrm{E}_{\omega,t}\left[\exp\left\{\left(m - r - \frac{1}{2}\lambda^\top\lambda - \frac{1}{2}v^\top v\right)(s-t) + (v-\lambda)^\top (z_s - z_t)\right\}\right] ds \\
&= \omega_t \int_t^T e^{(m-r-v^\top\lambda)(s-t)}\, ds \\
&= \omega_t F(t).
\end{aligned}$$

We can think of the agent having a human wealth of $\omega F(t)$ and then buying leisure at the unit price ω_t. Note that as a consequence of the Cobb-Douglas type utility, the relation between optimal consumption and leisure, $c_t^*/(1-\varphi_t^*)$, is equal to $\omega_t\theta/(1-\theta)$, i.e. the relative price of the two "goods" multiplied by their relative importance in the utility function.

To study the effect of labor supply flexibility on optimal investments let us look at an agent who once and for all fixes a constant labor supply rate $\bar{\varphi}$. For a given supply $\bar{\varphi}$, the agent finds the optimal consumption and investment strategies by solving the optimization problem

$$\begin{aligned}
J(W,\omega,t;\bar{\varphi}) &= \sup_{c,\pi} \mathrm{E}_{W,\omega,t}\left[\int_t^T e^{-\delta(s-t)}\frac{1}{1-\gamma}\left[c_s^\theta(1-\bar{\varphi})^{1-\theta}\right]^{1-\gamma} ds\right] \\
&= \theta\,(1-\bar{\varphi})^{(1-\theta)(1-\gamma)} \sup_{c,\pi} \mathrm{E}_{W,\omega,t}\left[\int_t^T e^{-\delta(s-t)}\frac{1}{\theta(1-\gamma)} c_s^{\theta(1-\gamma)}\, ds\right].
\end{aligned}$$

The supremum in the last expression equals the indirect utility of an investor with a constant relative risk aversion of $1-\theta(1-\gamma)$ and an exogenously given labor income at the rate $y_t = \bar{\varphi}\omega_t$. Clearly the present value of future labor income will be $H(y_t,t) = \bar{\varphi}\omega_t F(t)$, where $F(t)$ is given above. In analogy with the analysis for an exogenous income rate in the previous subsection, we get

$$J(W,\omega,t;\bar{\varphi}) = \theta\,(1-\bar{\varphi})^{(1-\theta)(1-\gamma)} \frac{1}{1-\theta(1-\gamma)} g(t)^{\theta(1-\gamma)} \left(W + \bar{\varphi}\omega F(t)\right)^{1-\theta(1-\gamma)},$$

and the optimal portfolio for a given $\bar{\varphi}$ is given by

$$\Pi(W,\omega,t;\bar{\varphi}) = \frac{1}{1-\theta(1-\gamma)} \frac{W + \bar{\varphi}\omega F(t)}{W} \left(\sigma_t^\top\right)^{-1}\lambda - \frac{F(t)}{W}\left(\sigma_t^\top\right)^{-1} v.$$

The optimal value of $\bar{\varphi}$ is found by maximizing $J(W,\omega,0;\bar{\varphi})$.

For easy comparison let us assume a deterministic wage rate, $v \equiv 0$. Then the optimal portfolio of the agent with flexible labor supply is

$$\Pi(W,\omega,t) = \frac{1}{\gamma}\frac{W + \omega F(t)}{W}\left(\sigma^\top\right)^{-1}\lambda,$$

while the optimal portfolio of the agent with fixed labor supply at a rate $\bar{\varphi}$ is

$$\Pi(W,\omega,t;\bar{\varphi}) = \frac{1}{1-\theta(1-\gamma)} \frac{W + \bar{\varphi}\omega F(t)}{W}\left(\sigma^\top\right)^{-1}\lambda.$$

First note that the portfolio weights of the two agents have the same sign. There are two differences between these two expressions: the relevant risk aversion coefficient and the valuation of future income. With flexible supply the labor income enters as the maximum value of future wages, which can only be obtained by working all the time. On the other hand, the total risk aversion γ is applied for the flexible supplier instead of the consumption risk aversion $1-\theta(1-\gamma)$ applied for the fixed supplier. Let us consider assets with positive portfolio weights. If $\gamma < 1$, then $\gamma < 1-\theta(1-\gamma)$, and hence the flexible supplier will unambiguously invest more in the risky assets. If γ is sufficiently larger than 1, the relation between the portfolio weights is ambiguous and will depend on the exact parameter values, the remaining life-time, and the fixed labor supply rate. For moderately risk-averse investors at an early stage in their working life, the financial investments of the flexible labor supplier tend to be more risky than those of the fixed labor supplier. The intuition is that investors incurring losses on their financial investments may compensate by working harder and drive up labor income. Labor supply flexibility serves as a kind of insurance. Changes of labor supply have the largest effect on capitalized labor income for young investors. The flexibility of labor supply may therefore amplify the horizon effect of labor income on risky investments which is present already for an exogenously given labor income stream. With an uncertain wage rate spanned by the risky financial assets, this conclusion seems to hold as long as the wage rate is not "too risky", cf. the discussion in Bodie, Merton, and Samuelson (1992). Apparently, the effects of labor supply flexibility have not been studied in the more reasonable incomplete market setting, where the wage rate is not fully diversifiable.

8.1.3 Further references on labor income in portfolio and consumption choice

Cocco, Gomes, and Maenhout (1998), Svensson and Werner (1993), Kenc (1999), El Karoui and Jeanblanc-Picqué (1998), Constantinides, Donaldson, and Mehra (2002), Cocco (1999), Cocco (2000), Cuoco (1997), He and Pagès (1993), Koo (1995)

8.2 Inflation

In the models considered so far we have implicitly assumed either that the financial assets are real assets in the sense that they pay out in consumption units *or* that the price of the consumption good is constant over time. In this section we allow for stochastic consumer prices in a market where the traded bonds are nominal in the sense that they pay out in monetary units. It is sometimes claimed that stocks are appropriate for hedging inflation uncertainty so that the real returns on stocks are quite stable relative to the real returns on long-term nominal bonds. This could explain the popular advice that long-term investors should invest more in stocks than short-term investors.

If only nominal bonds are traded, the optimal investment strategy of an investor with utility of terminal wealth only is to combine the mean-variance portfolio and the portfolio that has the highest correlation with the return on an indexed bond with a maturity equal to the remaining horizon. The hedge portfolio generally involves both stocks and nominal bonds, the precise mix will be determined by the correlation structure. If inflation uncertainty is modest, nominal bonds are good substitutes for real bonds (true in the U.S. for the period 1983-2000; not true for 1950-1982) and nominal bonds will dominate the hedge portfolio. Estimates on U.S. data approx. 1950–2000 show that the stock index is slightly *positively* correlated with the real interest rate. Hence the stock will enter the hedge portfolio with a negative weight unlike the popular advice.

General aspects of the portfolio choice problem with uncertain inflation are discussed by Munk and Sørensen (2002). The effects of uncertain inflation on portfolio choice have been studied in concrete settings by e.g. Brennan and Xia (2002), Munk, Sørensen, and Vinther (2002), and Campbell and Viceira (2001). Both Brennan and Xia (2002) and Munk, Sørensen, and Vinther (2002) consider investors with CRRA utility of wealth at the end of a finite horizon, whereas Campbell and Viceira (2001) allow for intermediate consumption and a more general recursive utility specification in an infinite horizon setting. The infinite horizon assumption, however, makes it difficult to address effects due to investors having different investment horizons. In both Brennan and Xia (2002) and Campbell and Viceira (2001) the real interest rate is described by a one-factor Vasicek model and the expected inflation dynamics is given by an Ornstein-Uhlenbeck process. The term structure of nominal interest rates is therefore described by a two-factor model. Munk, Sørensen, and Vinther (2002) differ slightly by assuming a one-factor Vasicek model for the nominal interest rates, while the implied term structure of real interest rates is described by a two-factor model. In the model of Munk, Sørensen, and Vinther it is impossible to replicate a real bond by trading in any number of nominal bonds whereas this is possible in the other models. The main conclusions of Brennan and Xia (2002) and Munk, Sørensen, and Vinther (2002) are very close, however. For concreteness, let us follow the set-up of Munk, Sørensen, and Vinther.

We consider the investment problem of an investor who has CRRA utility of terminal (time T) real wealth only. As before γ represents the relative risk aversion of the agent. The investor can hold cash (i.e. a money market bank account), nominal bonds, and stocks. The nominal interest rate dynamics is described by an Ornstein-Uhlenbeck process,

$$dr_t = \kappa(\bar{r} - r_t)dt - \sigma_r dz_{1t}, \tag{8.9}$$

cf. Vasicek (1977) and Section 7.1. The dynamics of the price B_t of any bond (or other fixed-income securities) is of the form

$$dB_t = B_t \left[(r_t + \lambda_1 \sigma_B(r_t, t))\, dt + \sigma_B(r_t, t)\, dz_{1t} \right], \tag{8.10}$$

where λ_1 is the market price of risk induced by the exogenous shock process z_1. The stock index (with dividends reinvested) is assumed to evolve according to the stochastic differential equation

$$dS_t = S_t \left[(r_t + \psi \sigma_S)\, dt + \rho_{BS}\sigma_S\, dz_{1t} + \sqrt{1-\rho_{BS}^2}\sigma_S\, dz_{2t} \right].$$

The parameter ρ_{BS} is the correlation between bond market returns and stock market returns, σ_S is the volatility of the stock, and ψ is the Sharpe ratio of the stock which we assume constant. In total, the dynamics of nominal asset prices can be written as

$$\begin{pmatrix} dB_t \\ dS_t \end{pmatrix} = \begin{pmatrix} B_t & 0 \\ 0 & S_t \end{pmatrix} \left[\left(r_t \mathbf{1} + \begin{pmatrix} \sigma_B(r_t,t) & 0 \\ \rho_{BS}\sigma_S & \sqrt{1-\rho_{BS}^2}\sigma_S \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \right) dt \right. \\ \left. + \begin{pmatrix} \sigma_B(r_t,t) & 0 \\ \rho_{BS}\sigma_S & \sqrt{1-\rho_{BS}^2}\sigma_S \end{pmatrix} \begin{pmatrix} dz_{1t} \\ dz_{2t} \end{pmatrix} \right], \tag{8.11}$$

where $\lambda_2 = (\psi - \rho_{BS}\lambda_1)/\sqrt{1-\rho_{BS}^2}$. Letting $\pi = (\pi_B, \pi_S)^\top$ denote the fractions of nominal wealth invested in the bond and the stock, the nominal wealth W_t will evolve as

$$dW_t = W_t \left[(r_t + \pi_t^\top \sigma_t \lambda)\, dt + \pi_t^\top \sigma_t \begin{pmatrix} dz_{1t} \\ dz_{2t} \end{pmatrix} \right],$$

where

$$\sigma_t = \begin{pmatrix} \sigma_B(r_t,t) & 0 \\ \rho_{BS}\sigma_S & \sqrt{1-\rho_{BS}^2}\sigma_S \end{pmatrix}, \qquad \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix}.$$

The nominal price of the real consumption good in the economy at time t is denoted by I_t. The real price of any asset in the economy is thus determined by deflating by the price index I_t. For example, the real price of the stock is given by S_t/I_t. The dynamics of the nominal price of the consumption good is given by the following system of differential equations:

$$\frac{dI_t}{I_t} = i_t\,dt + \sigma_{I1}\,dz_{1t} + \sigma_{I2}\,dz_{2t} + \sigma_{I3}\,dz_{3t}, \tag{8.12}$$

and

$$di_t = \beta(\bar{i} - i_t)\,dt + \sigma_{i1}\,dz_{1t} + \sigma_{i2}\,dz_{2t} + \sigma_{i3}\,dz_{3t} + \sigma_{i4}\,dz_{4t}, \tag{8.13}$$

where i_t is the expected rate of inflation, $\bar{i}$ describes the long-run mean of the rate of inflation, β describes the degree of mean-reversion, and the volatility coefficients σ_{Ik} and σ_{ik} are all constant. Define $\sigma_I^2 = \sigma_{I1}^2 + \sigma_{I2}^2 + \sigma_{I3}^2$ and $\sigma_i^2 = \sigma_{i1}^2 + \sigma_{i2}^2 + \sigma_{i3}^2 + \sigma_{i4}^2$. The instantaneous variance rates of the price index and the expected inflation rate are then $\sigma_I^2 I_t^2$ and σ_i^2, respectively. Changes in the nominal price index and the inflation rate are correlated with the stock index return and interest rates. Let us denote the covariance rate between the return on the stock index and the price level by $\sigma_{SI} = \sigma_S[\rho_{BS}\sigma_{I1} + \sqrt{1-\rho_{BS}^2}\sigma_{I2}]$. Similarly, the covariance rate between the return on the stock index and the expected inflation rate is denoted by $\sigma_{Si} = \sigma_S[\rho_{BS}\sigma_{i1} + \sqrt{1-\rho_{BS}^2}\sigma_{i2}]$, the covariance rate between the return on the bond and the price level is $\sigma_{BI} = \sigma_B\sigma_{I1}$, the covariance rate between the return on the bond and the inflation rate is $\sigma_{Bi} = \sigma_B\sigma_{i1}$, and the covariance rate between the price level and the inflation rate is $\sigma_{Ii} = \sigma_{I1}\sigma i1 + \sigma_{I2}\sigma_{i2} + \sigma_{I3}\sigma_{i3}$.

The real wealth of the investor at time t is $w_t = W_t/I_t$, which by Itô's Lemma has the dynamics

$$\begin{aligned} dw_t &= \frac{1}{I_t}\,dW_t - \frac{W_t}{I_t^2}\,dI_t - \frac{1}{I_t^2}(dW_t)(dI_t) + \frac{W_t}{I_t^3}(dI_t)^2 \\ &= w_t\left[\left(r_t - i_t + \sigma_I^2 + \pi_t^\top\sigma_t\lambda - \pi_t^\top\sigma_t\begin{pmatrix}\sigma_{I1}\\ \sigma_{I2}\end{pmatrix}\right)dt + \pi_t^\top\sigma_t\begin{pmatrix}dz_{1t}\\ dz_{2t}\end{pmatrix} - (\sigma_{I1},\sigma_{I2},\sigma_{I3})^\top\begin{pmatrix}dz_{1t}\\ dz_{2t}\\ dz_{3t}\end{pmatrix}\right]. \end{aligned}$$

The variables w, r, and i form a Markov system and provide sufficient information for the decisions of the investor. Hence, the indirect utility is given as a function $J(w,r,i,t)$. Defining

$$\mu_w = r + \pi^\top\sigma_t\lambda - i_t + \sigma_I^2 - \pi^\top\begin{pmatrix}\sigma_{SI}\\ \sigma_{BI}\end{pmatrix}, \qquad \sigma_w^2 = \pi^\top\sigma_t\sigma_t^\top\pi + \sigma_I^2 - 2\pi^\top\begin{pmatrix}\sigma_{SI}\\ \sigma_{BI}\end{pmatrix},$$

$$\sigma_{wr} = -\sigma_r\left(\pi_B\sigma_B - \sigma_{I1}\right), \qquad \sigma_{wi} = \pi^\top\begin{pmatrix}\sigma_{Si}\\ \sigma_{Bi}\end{pmatrix} - \sigma_{Ii},$$

we can write the Hamilton-Jacobi-Bellman equation associated with the problem of maximizing the expected utility as

$$\begin{aligned} \sup_{\pi=(\pi_B,\pi_S)\in\mathbb{R}^2} \Big\{ & \mu_w w J_w + \kappa(\bar{r} - r)J_r + \beta(\bar{i} - i)J_i + \frac{1}{2}\sigma_w^2 w^2 J_{ww} \\ & + \frac{1}{2}\sigma_r^2 J_{rr} + \frac{1}{2}\sigma_i^2 J_{ii} + \sigma_{wr} w J_{wr} + \sigma_{wi} w J_{wi} + \sigma_{ri} J_{ri} + \frac{\partial J}{\partial t} \Big\} = 0. \end{aligned} \tag{8.14}$$

The boundary condition is $J(w, r, i, T) = w^{1-\gamma}/(1-\gamma)$. The first order condition of the maximization problem in (8.14) provides the following characterization of the optimal risky asset proportions π:

$$\pi = \begin{pmatrix} \pi_B \\ \pi_S \end{pmatrix} = \frac{-J_w}{wJ_{ww}} \left(\sigma_t^\top\right)^{-1} \lambda + \frac{J_{wr}}{wJ_{ww}} \frac{\sigma_r}{\sigma_B} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left(1 + \frac{J_w}{wJ_{ww}}\right) \left(\sigma_t \sigma_t^\top\right)^{-1} \begin{pmatrix} \sigma_{BI} \\ \sigma_{SI} \end{pmatrix} - \frac{J_{wi}}{wJ_{ww}} \left(\sigma_t \sigma_t^\top\right)^{-1} \begin{pmatrix} \sigma_{Bi} \\ \sigma_{Si} \end{pmatrix}. \tag{8.15}$$

The expression in (8.15) provides a general characterization of the optimal portfolio weights in the specific market setting. The first term in (8.15) is the usual speculative portfolio which is optimal for an investor with log utility. The other three terms in (8.15) describe how the investor optimally hedges changes in the opportunity set. The second term describes the hedge against the nominal interest rate and was also found in Section 7.1. The last two terms in (8.15) describe how the investor hedges against short-run unexpected inflation and changes in future inflation rates, respectively.

With the assumed "affine" dynamics of r and i, it will come as no surprise that the indirect utility function of the CRRA investor is given by

$$J(w, r, i, t) = \frac{1}{1-\gamma} g(r, i, t)^\gamma w^{1-\gamma}, \tag{8.16}$$

where

$$\begin{aligned} g(r, i, t) &= e^{A_1(T-t) + A_2(T-t)r + A_3(T-t)i}, \\ A_2(\tau) &= \frac{1-\gamma}{\gamma} b(\tau) \equiv \frac{1-\gamma}{\gamma} \frac{1}{\kappa} \left(1 - e^{-\kappa\tau}\right), \\ A_3(\tau) &= -\frac{1-\gamma}{\gamma} A_3^*(\tau) \equiv -\frac{1-\gamma}{\gamma} \frac{1}{\beta} \left(1 - e^{-\beta\tau}\right), \end{aligned}$$

and A_1 can be found explicitly, but is not important for the optimal portfolio choice. By substitution of the relevant derivatives into (8.15), the vector of optimal risky asset allocations at time t is given by:

$$\pi = \frac{1}{\gamma} \left(\sigma_t^\top\right)^{-1} \lambda + \left(1 - \frac{1}{\gamma}\right) \frac{\sigma_r b(T-t)}{\sigma_B} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \left(1 - \frac{1}{\gamma}\right) \left(\sigma_t \sigma_t^\top\right)^{-1} \left[\begin{pmatrix} \sigma_{BI} \\ \sigma_{SI} \end{pmatrix} + \begin{pmatrix} \sigma_{Bi} \\ \sigma_{Si} \end{pmatrix} A_3^*(T-t) \right]. \tag{8.17}$$

The residual $1 - \pi^\top \mathbf{1} = 1 - \pi_S - \pi_B$ is invested in the bank account.

The optimal portfolio weights for CRRA investors are linear combinations of the speculative portfolio and the different hedge portfolios. In particular, for investors with the same investment horizon T the optimal portfolios are linear combinations of the speculative portfolio and a single hedge portfolio; the relative risk tolerance, $1/\gamma$, describes the weights on the two relevant portfolios.

As discussed above, the second term in (8.17) describes the hedge against changes in the nominal interest rate and consists entirely of a position in the bond. As noted in Section 7.1, the occurrence of this hedge term implies that the bond/stock ratio will increase with the risk aversion consistent with popular recommendations. On the other hand, the last hedge term in (8.17) describes the inflation hedge and involves the stock. This term is depending on the investment horizon through the positive and increasing function $A_3^*(T-t)$. In particular, the parameter β determines the difference on the stock allocations for myopic and long term investors with the same relative risk aversion. If β is small, changes in the expected inflation rate are relatively permanent, and horizon effects may be significant. However, whether this horizon effect implies more or fewer stocks for the long-term investor depends on the sign of the covariance σ_{Si} between stock returns and inflation, that is whether the stock serves as a relatively good substitute for the real bond that should ideally be used for hedging changes in real rates in a complete market setting. Moreover, while the last term in (8.17) can potentially explain the typically recommended horizon dependence for stocks, it may also change the ratio between bonds and stocks.

Munk, Sørensen, and Vinther calibrate the model using historical US data from the period 1951–2001. The estimation is based on maximum likelihood and an application of the Kalman filter. The point estimate of the covariance parameter σ_{Si} is slightly negative so that the optimal stock weight for $\gamma > 1$ is slightly *decreasing* with the investment horizon in contrast to popular investment advice. The stock index is, in fact, positively correlated with the real interest rate[2] and is therefore a bad substitute for the relevant real bond that should ideally be used as the instrument for hedging long term inflation risk and real interest rate risk. However, when the capital market parameters are allowed to vary within intervals of plus-minus two standard deviations on the estimates (which could reflect reasonable uncertainty on the parameter estimates), the theoretical asset allocation results can closely mimic popular asset allocation advice. In particular, the model can generate both a bond/stock ratio which is increasing in the risk aversion coefficient and a stock investment that increases with the length of the investment horizon. The recommendations are quantitatively very difficult to match, however.

8.3 Multiple and/or durable consumption goods

References: Several perishable: Breeden (1979)

With durable: Grossman and Laroque (1990), Hindy and Huang (1993), Detemple and Giannikos (1996), Cuoco and Liu (2000), Damgaard, Fuglsbjerg, and Munk (2002)

Housing: Brueckner (1997), Cocco (1999), Cocco (2000), Flavin and Yamashita (1998)

8.4 Uncertain time of death

References: See Richard (1975)

[2]Under the assumptions of the model, the real short-term interest rate is given by the nominal interest rate minus the expected inflation rate plus a constant.

Chapter 9

Non-standard assumptions on investors

9.1 Preferences with habit formation

It has long been recognized by economists that preferences may not be intertemporally separable. According to Browning (1991), this idea dates back to the 1890 book "Principles of Economics" by Alfred Marshall. See Browning's paper for further references to the critique on intertemporally separable preferences. In particular, the utility associated with the choice of consumption at a given date may depend on past choices of consumption. This is modeled by replacing $u(c_t, t)$ by $u(c_t, h_t, t)$, where u is decreasing in h_t, which is a measure of the standard of living or the habit level of consumption, e.g. a weighted average of past consumption rates:

$$h_t = h_0 e^{-\beta t} + \alpha \int_0^t e^{-\beta(t-s)} c_s \, ds,$$

where h_0, α, and β are non-negative constants. High past consumption generates a desire for high current consumption, so that preferences display intertemporal complementarity. As additional motivation for such preferences, note that several papers have documented the importance of allowing for habit formation in utilities when it comes to equilibrium asset pricing. Empirical facts that seem puzzling relative to models with a representative agent having time-separable utility can be resolved by introducing habit formation into the utility function. For example, Constantinides (1990) and Sundaresan (1989) demonstrate that models with habit formation can obtain a high equity premium with low risk aversion. Campbell and Cochrane (1999) and Wachter (2002a) construct representative agent models with habit formation that are consistent with observed variations in expected returns on stocks and bonds over time. Detemple and Zapatero (1991) also study asset pricing implications of habit formation preferences.[1]

Sundaresan (1989), Constantinides (1990), and Ingersoll (1992) all derive the optimal strategies for an investor with an infinite time horizon under the assumption of a constant investment opportunity set. In addition, Ingersoll (1992) considers a finite-horizon investor with log utility.

Detemple and Zapatero (1992) derive conditions under which optimal policies exist for an investor with habit persistence in preferences. They are able to characterize the optimal consumption strategy in a general setting, but, except for the case of deterministic investment opportunities,

[1] Both Campbell and Cochrane (1999) and Wachter (2002a) consider utility with *external* habit formation in the sense that the agent does not take into account the effect that the choice of current consumption has on future habit levels. In the other papers referred to, these effects are considered.

they state the optimal portfolio in terms of an unknown stochastic process that comes out of the martingale representation theorem. Detemple and Karatzas (2001) provide a similar analysis for a preference structure that also involves habit formation but is more general in several respects.

Schroder and Skiadas (2002) show that the general decision problem of an investor with habit persistence in preferences who can trade in a given financial market is equivalent to the decision problem of an investor who does not exhibit habit formation, and who can trade in a financial market with more complex dynamics of investment opportunities.

Munk (2002) gives a precise characterization of the optimal portfolio in a general complete market setting and derive explicit results in concrete settings with stochastic investment opportunities. The assumed objective is

$$J_t = \sup_{(c,\pi)\in\mathcal{A}(t)} \mathrm{E}_t \left[\int_t^T e^{-\delta(s-t)} u(c_s, h_s)\, ds \right], \tag{9.1}$$

where $\mathcal{A}(t)$ denotes the set of feasible consumption and portfolio strategies over the period $[t, T]$, and the "instantaneous" utility function $u(c, h)$ is assumed to be power-linear,

$$u(c,h) = \frac{1}{1-\gamma} (c-h)^{1-\gamma}, \tag{9.2}$$

where the constant $\gamma > 0$ is a risk aversion parameter. With this specification the consumption rate is required to exceed the habit level, so that the habit level plays the role of a minimum or subsistence consumption rate determined by past consumption rates. Let us briefly summarize the main findings of that paper without going into the modeling details:

Mean-reverting stock returns. Stock returns are assumed to be predictable in the sense that the market price of risk follows a mean-reverting process. Interest rates are assumed constant. Under the assumption of perfect negative correlation between the stock price and the market price of risk, Munk finds an explicit solution for the optimal strategies. This is a generalization of the results of Wachter (2002b), cf. Section 7.2, who assumes time-separable utility. The optimal fraction of wealth invested in stocks is the sum of a myopic demand and a (positive) hedge demand. Habit persistence has different effects on these two components, but in our numerical examples the differences are very small. It is argued that, contrary to the case of time-additive utility, the optimal fraction of wealth invested in stocks is not necessarily monotonically decreasing over the life of an investor with habit persistence in preferences for consumption. Finally, relative to the case of constant expected returns, mean reverting returns support a higher consumption rate, but in the numerical examples the increase is considerably smaller for investors with habit persistence than investors without.

Stochastic interest rates. The short-term interest rate is assumed to follow a square-root process as suggested by Cox, Ingersoll, and Ross (1985) with the market prices of risk being fully determined by the interest rate level. The assets available for investment are a stock (index), cash (i.e. the bank account), and a single bond (without loss of generality). While the optimal stock portfolio weight can be found in closed form, the optimal allocation to the bond and cash as well as the optimal consumption rate involve a time and interest rate dependent function which is the solution to a relatively simple partial differential equation (PDE). With time-additive preferences the PDE has an explicit solution, cf. Section 7.1, but with habit preferences the PDE must be solved numerically. The bond portfolio weight has all three components identified in the general model: a myopic term, a hedge term, and a term ensuring that the future consumption at least reaches the habit level. The stock portfolio weight, on the other hand, has only the myopic component. The numerical experiments shown in the paper verify that habit formation have very different effects on stock and bond investments and show that the effects on consumption are ambiguous.

Labor income. The agent is assumed to receive a continuous stream of labor income. The income stream has two effects. Firstly, the initial wealth is to be increased by the present value of the future income stream, which implies that a larger fraction of financial wealth is to be invested in the risky assets. Habit persistence in preferences dampens this effect. Secondly, a labor income stream is implicitly equivalent to a stream of returns on a financial portfolio, so the explicit investment strategy must be adjusted accordingly. This adjustment is independent of the preference parameters and, hence, unaffected by habit persistence. Except for extreme habit persistence and very low present value of income (relative to financial wealth), the effects of labor income seem to dominate the effects of habit persistence.

In sum, habit persistence dampens the speculative investments of investors due to the fact that some funds must be reserved for the purpose of ensuring that consumption in the future will meet the habit level. The hedge investments may be affected differently by habit persistence, but in the numerical examples given by Munk (2002) the differences are small. The main effect on the relative allocations to different assets stems from the fact that some assets (bonds and cash) are better investment objects than others (stocks) when it comes to ensuring that future consumption will not fall below the habit level.

Further references: Hindy, Huang, and Zhu (1997)

9.2 Recursive utility

Schroder and Skiadas (1999) and Campbell and Viceira (1999, 2001) study consumption and portfolio decisions with so-called recursive utility or stochastic differential utility...

9.3 Other objective functions

Portfolio choice problems of portfolio managers whose compensation depends on the performance of the portfolio chosen and a benchmark portfolio. The compensation may include option elements. See Carpenter (2000), Browne (1999).

9.4 Consumption and Portfolio Choice for Non-price takers

References: See Cuoco and Cvitanić (1998), Başak (1997)

9.5 Non-Utility Based Portfolio Choice

References: See Cover (1991), Jamshidian (1992)

9.6 Allowing for Bankruptcy

References: See Lehoczky, Sethi, and Shreve (1983), Sethi, Taksar, and Presman (1992), Presman and Sethi (1996)

Chapter 10

Trading and information imperfections

10.1 Model/parameter uncertainty or incomplete information

References: See Brennan (1998), Barberis (2000), Gennotte (1986), Karatzas and Xue (1991)

10.2 Trading constraints

References: See Bardhan (1994), Cuoco (1997), Cvitanić (1996), Cvitanić and Karatzas (1992), Fleming and Zariphopoulou (1991), Grossman and Vila (1991), He and Pearson (1991), Shirakawa (1994), Teplá (2000, 2001), Xu and Shreve (1992a), Xu and Shreve (1992b), Zariphopoulou (1992), Zariphopoulou (1994)

Value-at-risk constraints: Başak and Shapiro (2001), Cuoco, He, and Issaenko (2002), Cuoco and Liu (2002)

Drawdown constraints: Cvitanić and Karatzas (1995), Grossman and Zhou (1993),

10.3 Transaction Costs

References: See Magill and Constantinides (1976), Constantinides (1986), Davis and Norman (1990), Balduzzi and Lynch (1999), Cvitanić and Karatzas (1996), Shreve and Soner (1994), Duffie and Sun (1990), Taksar, Klass, and Assaf (1988)

Chapter 11

Computational Methods

References: See Detemple, Garcia, and Rindisbacher (2003), Cvitanić, Goukasian, and Zapatero (2000), Fitzpatrick and Fleming (1991), Munk (1997), Munk (2003b)

Bibliography

Ahn, D.-H., R. F. Dittmar, and A. R. Gallant (2002). Quadratic Term Structure Models: Theory and Evidence. *The Review of Financial Studies 15*(1), 243–288.

Bajeux-Besnainou, I., J. V. Jordan, and R. Portait (2000). An Asset Allocation Puzzle: Comment. Working paper.

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

Barberis, N. (2000). Investing for the Long Run when Returns are Predictable. *The Journal of Finance 55*, 225–264.

Bardhan, I. (1994). Consumption and Investment under Constraints. *Journal of Economic Dynamics and Control 18*, 909–929.

Bardhan, I. and X. Chao (1995). Martingale Analysis for Assets with Discontinuous Returns. *Mathematics of Operations Research 20*(1), 243–256.

Başak, S. (1997). Consumption Choice and Asset Pricing with a Non-Price-Taking Agent. *Economic Theory 10*, 437–462.

Başak, S. and A. Shapiro (2001). Value-at-Risk Based Risk Management: Optimal Policies and Asset Prices. *The Review of Financial Studies 14*, 371–405.

Bernoulli, D. (1954). Exposition of a New Theory on the Measurement of Risk. *Econometrica 22*(1), 23–36. Translation form the 1738 version.

Bodie, Z. (2003). Thoughts on the Future: Life-Cycle Investing in Theory and Practice. *Financial Analysts Journal 59*(1), 24–29.

Bodie, Z., R. C. Merton, and W. F. Samuelson (1992). Labor Supply Flexibility and Portfolio Choice in a Life Cycle Model. *Journal of Economic Dynamics and Control 16*, 427–449.

Breeden, D. T. (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics 7*, 265–296.

Brennan, M. J. (1998). The Role of Learning in Dynamic Portfolio Decisions. *European Finance Review 1*(3), 295–306.

Brennan, M. J., E. S. Schwartz, and R. Lagnado (1997). Strategic Asset Allocation. *Journal of Economic Dynamics and Control 21*(8–9), 1377–1403.

Brennan, M. J. and Y. Xia (2000). Stochastic Interest Rates and the Bond-Stock Mix. *European Finance Review 4*(2), 197–210.

Brennan, M. J. and Y. Xia (2002). Dynamic Asset Allocation under Inflation. *The Journal of Finance 57*(3), 1201–1238.

Browne, S. (1999). Beating a Moving Target: Optimal Portfolio Strategies for Outperforming a Stochastic Benchmark. *Finance and Stochastics 3*, 275–294.

Browning, M. (1991). A Simple Nonadditive Preference Structure for Models of Household Behavior over Time. *Journal of Political Economy 99*(3), 607–637.

Brueckner, J. K. (1997). Consumption and Investment Motives and the Portfolio Choices of Homeowners. *Journal of Real Estate Finance and Economics 15*(2), 159–180.

Campbell, J. Y. and J. H. Cochrane (1999). By Force of Habit: A Consumption-Based Explanation of Aggregate Stock Market Behavior. *Journal of Political Economy 107*, 205–251.

Campbell, J. Y. and R. J. Shiller (2001). Valuation Ratios and the Long-Run Stock Market Outlook: An Update. Forthcoming in Advances in Behavioral Finance, Volume II.

Campbell, J. Y. and L. M. Viceira (1999). Consumption and Portfolio Decisions when Expected Returns are Time Varying. *The Quarterly Journal of Economics 114*, 433–495.

Campbell, J. Y. and L. M. Viceira (2001). Who Should Buy Long-Term Bonds? *American Economic Review 91*(1), 99–127.

Carpenter, J. (2000). Does Option Compensation Increase Managerial Risk Appetite? *The Journal of Finance 55*(5), 2311–2331.

Chacko, G. and L. M. Viceira (2002, January). Dynamic Consumption and Portfolio Choice with Stochastic Volatility in Incomplete Markets. Working paper, Harvard University.

Cocco, J. F. (1999, January). Owner-Occupied Housing, Permanent Income, and Portfolio Choice. Working paper, Dept. of Economics, Harvard University.

Cocco, J. F. (2000, September). Hedging House Price Risk With Incomplete Markets. Working paper, London Business School.

Cocco, J. F., F. J. Gomes, and P. J. Maenhout (1998, February). Consumption and Portfolio Choice over the Life-Cycle. Working paper, Dept. of Economics, Harvard University.

Cochrane, J. H. (2001). *Asset Pricing*. Princeton University Press.

Constantinides, G. M. (1986). Capital Market Equilibrium with Transaction Costs. *Journal of Political Economy 94*(4), 842–862.

Constantinides, G. M. (1990). Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy 98*, 519–543.

Constantinides, G. M., J. B. Donaldson, and R. Mehra (2002). Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle. *Quarterly Journal of Economics 117*, 269–296.

Cover, T. (1991). Universal Portfolios. *Mathematical Finance 1*(1), 1–29.

Cox, J. C. and C.-f. Huang (1989). Optimal Consumption and Portfolio Policies when Asset Prices Follow a Diffusion Process. *Journal of Economic Theory 49*, 33–83.

Cox, J. C. and C.-f. Huang (1991). A Variational Problem Arising in Financial Economics. *Journal of Mathematical Economics 20*, 465–487.

Cox, J. C., J. E. Ingersoll, Jr., and S. A. Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica 53*(2), 385–407.

Cuoco, D. (1997). Optimal Consumption and Equilibrium Prices with Portfolio Constraints and Stochastic Income. *Journal of Economic Theory 71*(1), 33–73.

Cuoco, D. and J. Cvitanić (1998). Optimal Consumption Choices for a "Large" Investor. *Journal of Economic Dynamics and Control 22*, 401–436.

Cuoco, D., H. He, and S. Issaenko (2002, September). Optimal Dynamic Trading Strategies with Risk Limits. Working paper.

Cuoco, D. and H. Liu (2000). Optimal Consumption of a Divisible Durable Good. *Journal of Economic Dynamics and Control 24*(4), 561–613.

Cuoco, D. and H. Liu (2002, November). An Analysis of VaR-based Capital Requirements. Working paper.

Cvitanić, J. (1996). Optimal Trading under Constraints. Lecture notes, Department of Statistics, Columbia University.

Cvitanić, J., L. Goukasian, and F. Zapatero (2000, May). Monte Carlo Computation of Optimal Portfolios in Complete Markets. Working paper, University of Southern California.

Cvitanić, J. and I. Karatzas (1992). Convex Duality in Constrained Portfolio Optimization. *The Annals of Applied Probability 2*(4), 767–818.

Cvitanić, J. and I. Karatzas (1995). On Portfolio Optimization under "Drawdown" Constraints. In M. H. A. Davis, D. Duffie, W. H. Fleming, and S. E. Shreve (Eds.), *Mathematical Finance*, Volume 65 of *The IMA Volumes in Mathematics and Its Applications*, pp. 35–45. Springer-Verlag.

Cvitanić, J. and I. Karatzas (1996). Hedging and Portfolio Optimization under Transactions Costs: A Martingale Approach. *Mathematical Finance 6*, 133–165.

Dai, Q. and K. J. Singleton (2000). Specification Analysis of Affine Term Structure Models. *The Journal of Finance 55*(5), 1943–1978.

Damgaard, A., B. Fuglsbjerg, and C. Munk (2002). Optimal Consumption and Investment Strategies with a Perishable and an Indivisible Durable Consumption Good. Working paper, University of Southern Denmark. Forthcoming in *Journal of Economic Dynamics and Control*.

Davis, M. H. A. and A. R. Norman (1990). Portfolio Selection with Transaction Costs. *Mathematics of Operations Research 15*(4), 676–713.

Detemple, J., R. Garcia, and M. Rindisbacher (2003). A Monte-Carlo Method for Optimal Portfolios. *The Journal of Finance 58*(1), 401–446.

Detemple, J. and I. Karatzas (2001, July). Non-Addictive Habits: Optimal Consumption-Portfolio Policies. Working paper, Boston University and Columbia University.

Detemple, J. B. and C. I. Giannikos (1996). Asset and Commodity Prices with Multi-Attribute Durable Goods. *Journal of Economic Dynamics and Control 20*, 1451–1504.

Detemple, J. B. and F. Zapatero (1991). Asset Prices in An Exchange Economy with Habit Formation. *Econometrica 59*, 1633–1658.

Detemple, J. B. and F. Zapatero (1992). Optimal Consumption-Portfolio Policies with Habit Formation. *Mathematical Finance 2*(4), 251–274.

Dimson, E., P. Marsh, and M. Staunton (2002). *Triumph of the Optimists: 101 Years of Global Investment Returns*. Princeton University Press.

Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third ed.). Princeton University Press.

Duffie, D. and L. G. Epstein (1992). Stochastic Differential Utility. *Econometrica 60*(2), 353–394.

Duffie, D., W. Fleming, H. M. Soner, and T. Zariphopoulou (1997). Hedging in Incomplete Markets with HARA Utility. *Journal of Economic Dynamics and Control 21*(4–5), 753–782.

Duffie, D. and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance 6*(4), 379–406.

Duffie, D. and T.-s. Sun (1990). Transactions Costs and Portfolio Choice in a Discrete-Continuous-Time Setting. *Journal of Economic Dynamics and Control 14*, 35–51.

Duffie, D. and T. Zariphopoulou (1993). Optimal Investment with Undiversifiable Income Risk. *Mathematical Finance 3*(2), 135–148.

El Karoui, N. and M. Jeanblanc-Picqué (1998). Optimization of Consumption with Labor Income. *Finance and Stochastics 2*(4), 409–440.

Epstein, L. G. and S. E. Zin (1989). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. *Econometrica 57*(4), 937–969.

Epstein, L. G. and S. E. Zin (1991). Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis. *Journal of Political Economy 99*, 263–286.

Fama, E. F. (1970). Multiperiod Consumption-Investment Decisions. *American Economic Review 60*(1), 163–174. Correction: Fama (1976).

Fama, E. F. (1976). Multiperiod Consumption-Investment Decisions: A Correction. *American Economic Review 66*(4), 723–724.

Fitzpatrick, B. G. and W. H. Fleming (1991). Numerical Methods for an Optimal Investment-Consumption Model. *Mathematics of Operations Research 16*(4), 823–841.

Flavin, M. and T. Yamashita (1998, January). Owner-Occupied Housing and the Composition of the Household Portfolio over the Life Cycle. Working paper 6389, National Bureau of Economic Research.

Fleming, W. H. and H. M. Soner (1993). *Controlled Markov Processes and Viscosity Solutions*, Volume 25 of *Applications of Mathematics*. New York: Springer-Verlag.

Fleming, W. H. and T. Zariphopoulou (1991). An Optimal Investment/Consumption Model with Borrowing. *Mathematics of Operations Research 16*(4), 802–822.

Gennotte, G. (1986). Optimal Portfolio Choice under Incomplete Information. *The Journal of Finance 41*, 733–746.

Grasselli, M. (2000, April). HJB Equations with Stochastic Interest Rates and HARA Utility Functions. Working paper, CREST, Malakoff Cedex, France.

Grossman, S. J. and G. Laroque (1990). Asset Pricing and Optimal Portfolio Choice in the Presence of Illiquid Durable Consumption Goods. *Econometrica 58*(1), 25–51.

Grossman, S. J. and J.-L. Vila (1991). Optimal Dynamic Trading with Leverage Constraints. *Journal of Financial and Quantitative Analysis 27*(2), 151–168.

Grossman, S. J. and Z. Zhou (1993). Optimal Investment Strategies for Controlling Drawdowns. *Mathematical Finance 3*(3), 241–276.

Hakansson, N. H. (1970). Optimal Investment and Consumption Strategies Under Risk for a Class of Utility Functions. *Econometrica 38*(5), 587–607.

Hansen, L. P. and S. F. Richard (1987). The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models. *Econometrica 55*, 587–614.

He, H. and H. F. Pagès (1993). Labor Income, Borrowing Constraints, and Equilibrium Asset Prices. *Economic Theory 3*, 663–696.

He, H. and N. D. Pearson (1991). Consumption and Portfolio Policies with Incomplete Markets and Short-Sale Constraints: The Infinite Dimensional Case. *Journal of Economic Theory 54*, 259–304.

Hindy, A. and C.-f. Huang (1993). Optimal Consumption and Portfolio Rules with Durability and Local Substitution. *Econometrica 61*(1), 85–121.

Hindy, A., C.-f. Huang, and H. Zhu (1997). Optimal Consumption and Portfolio Rules With Durability and Habit Formation. *Journal of Economic Dynamics and Control 21*(2–3), 525–550.

Huang, C.-f. and R. H. Litzenberger (1988). *Foundations for Financial Economics*. Prentice-Hall.

Hull, J. C. (2003). *Options, Futures, and Other Derivatives* (Fifth ed.). Prentice-Hall, Inc.

Ibbotson, R. G. and P. Chen (2003). Long-Run Stock Returns. *Financial Analysts Journal 59*(1), 88–98.

Ingersoll, Jr., J. E. (1987). *Theory of Financial Decision Making*. Rowman & Littlefield.

Ingersoll, Jr., J. E. (1992). Optimal Consumption and Portfolio Rules with Intertemporally Dependent Utility of Consumption. *Journal of Economic Dynamics and Control 16*, 681–712.

Jagannathan, R. and N. R. Kocherlakota (1996). Why Should Older People Invest Less in Stocks Than Younger People? *Federal Reserve Bank of Minneapolis Quarterly Review 20*(3), 11–23.

Jamshidian, F. (1992). Asymptotically Optimal Portfolios. *Mathematical Finance 2*(2), 131–150.

Jeanblanc-Picqué, M. and M. Pontier (1990). Optimal Portfolio for a Small Investor in a Market Model with Discontinuous Prices. *Applied Mathematics and Optimization 22*, 287–310.

Karatzas, I., J. P. Lehoczky, S. P. Sethi, and S. E. Shreve (1986). Explicit Solution of a General Consumption/Investment Problem. *Mathematics of Operations Research 11*(2), 261–294.

Karatzas, I., J. P. Lehoczky, and S. E. Shreve (1987). Optimal Portfolio and Consumption Decisions for a "Small Investor" on a Finite Horizon. *SIAM Journal on Control and Optimization 25*(6), 1557–1586.

Karatzas, I., J. P. Lehoczky, S. E. Shreve, and G.-L. Xu (1991). Martingale and Duality Methods for Utility Maximization in an Incomplete Market. *SIAM Journal on Control and Optimization 29*(3), 702–730.

Karatzas, I. and X.-X. Xue (1991). A Note on Utility Maximization under Partial Observations. *Mathematical Finance 1*(2), 57–70.

Kenc, T. (1999, June). On the Determinants of Life-Cycle Risk-Taking. Working paper, University of Durham.

Kim, T. S. and E. Omberg (1996). Dynamic Nonmyopic Portfolio Behavior. *The Review of Financial Studies 9*(1), 141–161.

Kogan, L. and R. Uppal (2000, June). Risk Aversion and Optimal Portfolio Policies in Partial and General Equilibrium Economies. Technical report, The Wharton School and University of British Columbia.

Koo, H. K. (1995). Consumption and Portfolio Selection with Labor Income: Evaluation of Human Capital. Working paper, Olin School of Business, Washington University.

Koo, H. K. (1998). Consumption and Portfolio Selection with Labor Income: A Continuous Time Approach. *Mathematical Finance 8*(1), 49–65.

Korn, R. and H. Kraft (2001). A Stochastic Control Approach to Portfolio Problems with Stochastic Interest Rates. *SIAM Journal on Control and Optimization 40*(4), 1250–1269.

Kraft, H. (2003, February). Optimal Portfolios with Stochastic Volatility. Working paper, Frauenhofer ITWM, Kaiserslautern.

Kreps, D. M. (1990). *A Course in Microeconomic Theory*. Harvester Wheatsheaf.

Kreps, D. M. and E. Porteus (1978). Temporal Resolution of Uncertainty and Dynamic Choice Theory. *Econometrica 46*, 185–200.

Lehoczky, J. P., S. P. Sethi, and S. E. Shreve (1983). Optimal Consumption and Investment Policies Allowing Comsumption Constraints and Bankruptcy. *Mathematics of Operations Research 8*(4), 613–636.

Leippold, M. and L. Wu (2000, April). Quadratic Term Structure Models. Working paper, University of St. Gallen and Fordham University.

Liu, J. (1999, August). Portfolio Selection in Stochastic Environments. Working paper, Stanford University.

Magill, M. J. P. and G. M. Constantinides (1976). Portfolio Selection with Transactions Costs. *Journal of Economic Theory 13*, 245–263.

Markowitz, H. (1952). Portfolio Selection. *The Journal of Finance 7*, 77–91.

Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investment*. Wiley.

Mehra, R. (2003). The Equity Premium: Why Is It a Puzzle? *Financial Analysts Journal 59*(1), 54–69.

Mehra, R. and E. C. Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics 15*, 145–162.

Merton, R. C. (1969). Lifetime Portfolio Selection Under Uncertainty: The Continuous-Time Case. *Review of Economics and Statistics 51*, 247–257. Reprinted as Chapter 4 in Merton (1992).

Merton, R. C. (1971). Optimum Consumption and Portfolio Rules in a Continuous-Time Model. *Journal of Economic Theory 3*, 373–413. Erratum: Merton (1973a). Reprinted as Chapter 5 in Merton (1992).

Merton, R. C. (1973a). Erratum. *Journal of Economic Theory 6*, 213–214.

Merton, R. C. (1973b). An Intertemporal Capital Asset Pricing Model. *Econometrica 41*(5), 867–887.

Merton, R. C. (1973c). An Intertemporal Capital Asset Pricing Model. *Econometrica 41*(5), 867–887. Reprinted in an extended form as Chapter 15 in Merton (1992).

Merton, R. C. (1992). *Continuous-Time Finance*. Padstow, UK: Basil Blackwell Inc.

Merton, R. C. (2003). Thoughts on the Future: Theory and Practice in Investment Management. *Financial Analysts Journal 59*(1), 17–23.

Munk, C. (1997). Numerical Methods for Continuous-Time, Continuous-State Stochastic Control Problems. Publications from Department of Management 97/11, Odense University.

Munk, C. (2000). Optimal Consumption-Investment Policies with Undiversifiable Income Risk and Liquidity Constraints. *Journal of Economic Dynamics and Control 24*(9), 1315–1343.

Munk, C. (2002). Portfolio and Consumption Choice with Stochastic Investment Opportunities and Habit Formation in Preferences. Working paper, University of Southern Denmark.

Munk, C. (2003a). Fixed-Income Analysis: Securities, Pricing, and Risk Management. Lecture notes, University of Southern Denmark.

Munk, C. (2003b). Numerical Methods for Continuous-Time, Continuous-State Stochastic Control Problems: Experiences from Merton's Problem. *Applied Mathematics and Computation 136*(1), 47–77.

Munk, C. and C. Sørensen (2002). Optimal Real Consumption and Investment Strategies in Dynamic Stochastic Economies. Working paper, University of Southern Denmark and Copenhagen Business School.

Munk, C. and C. Sørensen (2003). Optimal Consumption and Investment Strategies with Stochastic Interest Rates. Working paper, University of Southern Denmark and Copenhagen Business School.

Munk, C., C. Sørensen, and T. N. Vinther (2002). Dynamic Asset Allocation Under Mean-Reverting Returns, Stochastic Interest Rates and Inflation Uncertainty. Working paper.

Nielsen, L. T. and M. Vassalou (2000, January). Portfolio Selection with Randomly Time-Varying Moments: The Role of the Instantaneous Capital Market Line. Working paper.

Øksendal, B. (1998). *Stochastic Differential Equations* (Fifth ed.). Springer-Verlag.

Pliska, S. R. (1986). A Stochastic Calculus Model of Continuous Trading: Optimal Portfolios. *Mathematics of Operations Research 11*(2), 371–382.

Presman, E. and S. Sethi (1996). Distribution of Bankruptcy Time in a Consumption/Investment Problem. *Journal of Economic Dynamics and Control 20*, 471–477.

Richard, S. F. (1975). Optimal Consumption, Portfolio and Life Insurance Rules for an Uncertain Lived Individual in a Continuous Time Model. *Journal of Financial Economics 2*, 187–203.

Samuelson, P. A. (1969). Lifetime Portfolio Selection by Dynamic Stochastic Programming. *Review of Economics and Statistics 51*, 239–246.

Schroder, M. and C. Skiadas (1999). Optimal Consumption and Portfolio Selection with Stochastic Differential Utility. *Journal of Economic Theory 89*, 68–126.

Schroder, M. and C. Skiadas (2002). An Isomorphism between Asset Pricing Models with and without Linear Habit Formation. *The Review of Financial Studies 15*(4), 1189–1221.

Sethi, S. P. and M. I. Taksar (1988). A Note on Merton's "Optimum Consumption and Portfolio Rules in a Continuous-Time Model". *Journal of Economic Theory 46*, 395–401.

Sethi, S. P., M. I. Taksar, and E. L. Presman (1992). Explicit Solution of a General Consumption/Portfolio Problem with Subsistence Consumption and Bankruptcy. *Journal of Economic Dynamics and Control 16*, 747–768.

Shiller, R. J. (2000). *Irrational Exuberance*. Princeton, New Jersey: Princeton University Press.

Shirakawa, H. (1994). Optimal Consumption and Portfolio Selection with Incomplete Markets and Upper and Lower Bound Constraints. *Mathematical Finance 4*(1), 1–24.

Shreve, S. E. and H. M. Soner (1994). Optimal Investment and Consumption with Transaction Costs. *The Annals of Applied Probability 4*(3), 609–692.

Sørensen, C. (1999). Dynamic Asset Allocation and Fixed Income Management. *Journal of Financial and Quantitative Analysis 34*, 513–531.

Sundaresan, S. M. (1989). Intertemporally Dependent Preferences and the Volatility of Consumption and Wealth. *The Review of Financial Studies 2*(1), 73–89.

Svensson, L. E. O. and I. M. Werner (1993). Nontraded Assets in Incomplete Markets. *European Economic Review 37*, 1149–1168.

Taksar, M., M. J. Klass, and D. Assaf (1988). A Diffusion Model for Optimal Portfolio Selection in the Presence of Brokerage Fees. *Mathematics of Operations Research 13*(2), 277–294.

Teplá, L. (2000). Optimal Portfolio Policies with Borrowing and Shortsale Constraints. *Journal of Economic Dynamics and Control 24*(11-12), 1623–1639.

Teplá, L. (2001). Optimal Investment with Minimum Performance Constraints. *Journal of Economic Dynamics and Control 25*(10), 1629–1645.

Vasicek, O. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics 5*, 177–188.

von Neumann, J. and O. Morgenstern (1944). *Theory of Games and Economic Behavior*. New Jersey: Princeton University Press.

Wachter, J. A. (1999, October). Risk Aversion and Allocation to Long-Term Bonds. Working paper, Harvard University.

Wachter, J. A. (2002a). Habit Formation and Returns on Bonds and Stocks. Working paper, Stern School of Business, NYU.

Wachter, J. A. (2002b). Portfolio and Consumption Decisions under Mean-Reverting Returns: An Exact Solution for Complete Markets. *Journal of Financial and Quantitative Analysis 37*(1), 63–91.

Weil, P. (1989). The Equity Premium Puzzle and the Risk-free Rate Puzzle. *Journal of Mathematical Economics 24*(3), 401–421.

Welch, I. (2000). Views of Financial Economists on the Equity Premium and Other Issues. *Journal of Business 73*(4), 501–537.

Wu, L. (2000, April). Jumps and Dynamic Asset Allocation. Working paper, Graduate School of Business, Fordham University, New York.

Xu, G.-L. and S. E. Shreve (1992a). A Duality Method for Optimal Consumption and Investment Under Short-Selling Prohibition. I. General Market Coefficients. *The Annals of Applied Probability 2*(1), 87–112.

Xu, G.-L. and S. E. Shreve (1992b). A Duality Method for Optimal Consumption and Investment Under Short-Selling Prohibition. II. Constant Market Coefficients. *The Annals of Applied Probability 2*(2), 314–328.

Zariphopoulou, T. (1992). Investment-Consumption Models with Transaction Fees and Markov-Chain Parameters. *SIAM Journal on Control and Optimization 30*(3), 613–636.

Zariphopoulou, T. (1994). Consumption-Investment Models with Constraints. *SIAM Journal on Control and Optimization 32*(1), 59–85.