Tilburg University

Essays on optimal hedging and investment strategies and on derivative pricing

van den Goorbergh, R.W.J.

Publication date:
2004

Essays on Optimal Hedging and Investment Strategies, and on Derivative Pricing

Essays on Optimal Hedging and Investment Strategies, and on Derivative Pricing

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit van Tilburg, op gezag van de rector magnificus, prof. dr. F.A. van der Duyn Schouten, in het openbaar te verdedigen ten overstaan van een door het college voor promoties aangewezen commissie in de aula van de Universiteit op woensdag 19 mei 2004 om 16.15 uur

door

Rob Willem Jean van den Goorbergh

geboren op 4 augustus 1976 te Goirle.

PROMOTOREN: prof. dr. F.A. de Roon
prof. dr. B.J.M. Werker

Voor mijn ouders

Dankwoord (Acknowledgments)

Dit proefschrift is het resultaat van het promotieonderzoek dat ik in september 1999 aan de Katholieke Universiteit Brabant ben begonnen en ruim vier jaar later aan de Universiteit van Tilburg heb voltooid. Ik ben dank verschuldigd aan tal van personen die de uitvoering van dit onderzoek mogelijk hebben gemaakt.

Allereerst gaat mijn dank uit naar mijn promotoren, Frans de Roon en Bas Werker, zonder wier voortdurende betrokkenheid en motivatie deze dissertatie geen kans van slagen had. Ik heb veel geleerd van onze discussies, zelfs al waren die, zeker wanneer we gedrieën van gedachten wisselden, soms wat overweldigend. Een eigenzinniger stel geesten had ik mij niet kunnen wensen.

Ik wil ook Peter Kort en Kuno Huisman bedanken, die met veel enthousiasme mijn eerste voorzichtige schreden als onderzoeker hebben begeleid. Verder dank ik Christian Genest, met wie Bas en ik in januari 2003 onder barre Quebecse weersomstandigheden de kiemen hebben gezaaid voor ons copulaproject. Theo Nijman wil ik bedanken voor zijn richtinggevende ideeën en nuttig commentaar.

Ik betuig mijn erkentelijkheid aan Bruno Gérard, Frank de Jong en Marno Verbeek voor hun bereidheid zitting te nemen in de promotiecommissie.

Ten slotte wil ik graag mijn familie, vrienden en collega's bedanken voor hun steun, belangstelling en gezelschap.

Ferronica, terima kasih.

Ik draag het proefschrift op aan mijn ouders. Immers, wat zullen we nog de dingen aanvangen die tot proefschriften leiden, als we het niet voor onze ouders doen?

Amsterdam, maart 2004

Contents

Chapter 1

Introduction

The work presented in this dissertation encompasses a wide range of topics within the general field of finance. It is a collection of studies on investment decisions and asset pricing issues, each motivated by its own specific considerations. The diversity of topics covered in the thesis is illustrated by the variety of financial assets and investment opportunities analyzed, which range from exchange-traded instruments such as stocks, bonds, and futures contracts, to over-the-counter derivatives and non-traded assets such as real options. Such diversity does not justify a unifying introductory chapter which would only repeat standard textbook material. The remainder of this introduction confines itself to presenting an overview of the contributions of each chapter.

Chapter 2, titled *Risk Aversion, Price Uncertainty, and Irreversible Investments*, provides a generalization of the theory of irreversible investment under uncertainty, or *real options theory*, by allowing for risk averse investors in the absence of complete markets. Until now this theory has only been developed in the cases of risk neutrality, or risk aversion in combination with complete markets; see the seminal work by McDonald and Siegel (1985, 1986) and Dixit and Pindyck (1994) for an overview. Within a general setting, we prove the existence of a unique critical output price that distinguishes price regions in which it is optimal for a risk averse investor to invest and price regions in which one should refrain from investing. We use a class of utility functions that exhibit non-increasing absolute risk aversion to examine the effects of risk aversion, price uncertainty, and other parameters on the optimal investment decision.

We find that, as one may expect, risk aversion reduces investment. Contrary to the risk neutral model, however, our results show that under risk aversion the investment threshold increases more than linearly with the investment outlay. Moreover, we show that a rise in price uncertainty increases the value of deferring irreversible investments. This effect is stronger for high levels of risk aversion. In addition, we provide, for the first time, closed-form comparative statics formulas for the risk neutral investor.

Chapter 3, titled *Economic Hedging Portfolios*, studies portfolios that investors hold to hedge economic risks. Using a model of state-dependent utility, we show that agents' economic hedging portfolios can be obtained by an intuitively appealing, risk aversion-weighted approximate replication of the economic risk variables using the investment opportunity set. This approach extends the usual unweighted hedging scheme obtained in the traditional mean-variance framework analyzed in, e.g., Mayers (1972) and Anderson and Danthine (1980, 1981).

Using an investment opportunity set of stock and bond portfolios, we show that agents across a broad range of levels of risk aversion are willing to pay significant compensations for hedges against inflation risk, real interest-rate risk, and dividend-yield risk. Furthermore, our results show that all economic risk variables we consider require significant hedging positions in one or more securities. Moreover, we analyze investors' speculative positions and find that hedges against economic risks may potentially explain the anomalies that have been found in stock markets as well as the term and default premiums in bond markets; see Fama and French (1992, 1993, 1995).

In Chapter 4, titled *Multivariate Option Pricing Using Dynamic Copula Models*, we examine the price behavior of multivariate options in the presence of association between the underlying assets. Multivariate options are derivatives written on two or more underlying assets, usually taking the form of calls (or puts) that give the right to buy (or sell) the best or worst performer of the underlyings. We model the association between the underlyings using parametric families of copulas which offer various alternatives to the commonly assumed normal dependence structure.

Contrary to earlier works on multivariate option pricing, the dependence structure is not treated as fixed, but as possibly varying over time. Incorporating the notion of "correlation breakdowns" (see, e.g., Boyer, Gibson and Loretan (1999) and Patton (2002a, 2002b)), the dependence between the underlyings is assumed to vary over time as a function of the volatilities of the assets. These dynamic copula models are applied to better-of-two-markets and worse-of-two-markets options on the S&P 500 and Nasdaq indexes. Results show that option prices implied by dynamic copula models differ substantially from prices implied by models that fix the dependence between the underlyings, particularly in times of high volatilities. Furthermore, the normal copula produces option prices that differ significantly from non-normal copula prices, irrespective of initial volatility levels. Within the class of non-normal copula families considered, option prices are robust with respect to the copula choice.

Chapter 5, titled *An Anatomy of Futures Returns: Risk Premiums and Trading Strategies*, analyzes trading strategies which capture the various risk premiums that have been distinguished in futures markets and documented by, e.g., Fama (1984), Fama and French (1987), Bessembinder (1992), Bessembinder and Chan (1992), Carter, Rausser and Schmitz (1983), and DeRoon, Nijman and Veld (1998, 2000). On the basis of a simple decomposition of futures returns, we show that the return on a short-term futures contract measures the spot-futures premium, while spreading strategies that go long in long-term contracts and short in short-term contracts isolate the term premiums. Using a broad cross-section of U.S. commodity and financial futures markets and a wide range of delivery horizons, we examine the components of futures risk premiums empirically by means of "passive" trading strategies which fix positions over time, and "active" trading strategies along the lines of Jegadeesh and Titman (1993) and Fama and French (1992, 1995), which allow for dynamic trading and are designed to exploit the predictable variation in futures returns.

We find that passive, short-term strategies do not yield abnormal returns, in contrast to passive spreading strategies, implying the presence of non-zero term premiums. Furthermore, we find that the term structure of futures yields has strong explanatory power for both spot and term premiums, which can be exploited using active trading strategies that go long in low-yield markets and short in high-yield markets. The profitability of these yield-based trading strategies is not due to systematic risk. However, we show that transaction costs may eliminate these gains, in particular for the spreading strategies which capture short-term premiums.

Furthermore, we find that spreading returns are predictable by net hedge demand, which we show can be also exploited by active trading, but only if transaction costs are relatively low. Finally, we document significant momentum in futures markets. However, we find no evidence that momentum strategies outperform benchmark portfolios.

A last precursory note concerns the intellectual property of the work presented in this dissertation. The chapters to follow are based on co-authored papers. This also explains the use of the first person plural throughout the dissertation.[1] Chapter 2 is based on a paper with Peter Kort and Kuno Huisman. Chapter 3 is based on joint work with Frans de Roon and Bas Werker. Chapter 4 originated from joint work with Christian Genest and Bas Werker. Finally, Chapter 5 is based on work with Frans de Roon and Theo Nijman.

[1] An exception is Chapter 4, which, due to the strong feelings of one of its conceivers, avoids first person writing altogether.

Chapter 2

Risk Aversion, Price Uncertainty, and Irreversible Investments

2.1 Introduction

How should investors decide whether and when to invest in uncertain, irreversible projects in the case of incomplete markets? And what is the effect of risk aversion on investment behavior? This chapter addresses these questions in the context of the real options theory developed by McDonald and Siegel (1985, 1986). They show that the conventional net present value rule to decide whether or not to invest in some uncertain project ignores the option value of postponing the investment.

Dixit and Pindyck (1994) give a textbook treatment of this new investment theory. They describe two closely related but essentially different mathematical tools to model investment decisions: dynamic programming and contingent claims analysis. The latter endogenously determines an investor's discount rate as an implication of the overall capital market equilibrium. Both risk neutrality and risk aversion can be dealt with within the contingent claims approach, but the approach requires the existence of a sufficiently rich set of markets of risky assets so that a dynamic portfolio of traded assets exactly replicates the payoff of the investment that is to be valued. This assumption of complete markets is in reality quite strong, especially for investments in non-traded assets such as investments in marketing or advertising, or the development of new products (see, e.g., Magill and Quinzii (1995)). Dynamic programming, however, makes no such demand; if risk cannot be traded in markets, the investor's objective function can simply reflect the decision maker's valuation of risk. Until now, dynamic programming has only been applied to the problem of irreversibility under the assumption of risk neutrality.

In this chapter we consider the economically relevant problem faced by risk averse investors who contemplate an irreversible investment in an asset whose payoff cannot be replicated by a dynamic portfolio of traded securities. Hence, in this (realistic) situation of incomplete markets, we are not able to use contingent claims analysis as a tool to solve the investment problem. Instead, we apply dynamic programming to an objective function that reflects risk aversion.

The purpose of this study is to generalize the approach of McDonald and Siegel (1986) and Dixit and Pindyck (1994) by allowing for risk aversion in an environment of incomplete markets. Our aim is to find out how the optimal investment decision is affected by risk aversion, investment size, price uncertainty, and other parameters.

Our main results are the following. First of all we prove that, within a general setting, a unique critical price level exists for which the risk averse investor is indifferent between investing and not investing. Second, we introduce a class of utility functions with the desirable property of non-increasing absolute risk aversion to examine the comparative statics of this critical price level with respect to risk aversion, investment size, price uncertainty, and other parameters. We find that risk aversion reduces investment, particularly if the investment size is large. Moreover, we find that a rise in uncertainty increases the value of deferring irreversible investments. This effect is stronger for high levels of risk aversion.

The remainder of the chapter is organized as follows. Section 2.2 formulates the investment problem. Section 2.3 describes the general solution of the investment problem. In Section 2.4 we introduce a class of utility functions which exhibit the desirable property of non-increasing absolute risk aversion. This class of utility functions allows us to numerically examine the comparative statics of the critical price level under risk aversion in Section 2.5. In addition, we provide analytical comparative statics formulas for the risk neutral investor. Section 5.6 concludes.

2.2 The investment problem

We use a set-up along the lines of Dixit and Pindyck (1994, pp. 185–186). Consider an infinitely-lived investor contemplating an irreversible, discrete investment opportunity with sunk cost $I > 0$. For simplicity we assume that once the investment is made, it produces one unit of output flow into the indefinite future with no variable costs of production. The output price P_t is assumed to follow a geometric Brownian motion,

$$dP_t = \alpha P_t dt + \sigma P_t dz_t, \tag{2.1}$$

where $\sigma > 0$ and z_t is a standard Wiener process. Let $P_0 = P \geq 0$ denote the current output price. The required amount of money I is borrowed at an instantaneous riskless rate of interest $r > 0$ which we assume to be constant and larger than α. Thus, if the investor decides to invest at time $t = 0$, then the instantaneous net cash flow accruing from the project at any time $t \geq 0$ is

$$ncf_t \equiv P_t - rI.$$

Note that since $P \geq 0$, the range of possible values for ncf is $[-rI, \infty)$.

We assume that the investor's preferences are intertemporally additive, and that they can be represented by an increasing, twice differentiable von Neumann-Morgenstern utility function $u(\cdot)$ which is defined over the instantaneous net cash flows and independent of time, $u : [-rI, \infty) \to \mathbb{R}$. Furthermore, we assume that utility flows are discounted at the riskless rate of return r. We shall consider both situations in which u reflects risk neutrality and situations in which u reflects risk aversion.

Our goal is to determine whether and when the investor should invest in the project. In making this investment decision it is important to not only take into account the expected utility of the net cash flows produced by the project, but also the real option value embedded in its irreversible nature. Once the investment has been made, it cannot be undone should prospects change for the worse. By deferring the investment, however, the investor can await new information that affects the desirability of the expenditure.

2.3 Valuing the investment opportunity

If the investor decides to invest at $t = 0$, the expected utility of the net cash flows produced by the project is given by

$$V(P) = E\int_{t=0}^{\infty} e^{-rt}u(ncf_t)dt.$$

As indicated by the notation, V depends on the current output price P of the project. According to the classical net present value (npv) rule, the investor would have to invest at $t = 0$ if P were such that $V(P)$ is positive, and refrain from investing otherwise. However, this approach disregards the option value of postponing the irreversible investment at time $t = 0$. Let $C(P)$ denote this option value. It is determined by the following Bellman equation:

$$C(P) = u(0)dt + e^{-rdt}E\left\{C(P + dP_t)\right\}, \tag{2.2}$$

that is, the option value of deferring the investment is equal to the sum of the utility of waiting during a time interval $[0, dt]$ in which no cash flow occurs, and the discounted expected future utility of waiting.

Without loss of generality we assume that $u(0) = 0$, thereby in effect associating net cash inflows with positive utility levels, and net cash outflows with negative utility levels. Using this convention, we apply Itô's Lemma to rewrite the right-hand side of (2.2) as[1]

$$C(P) + \left[\frac{1}{2}\sigma^2 P^2 C''(P) + \alpha P C'(P) - rC(P)\right] dt + o(dt).$$

Substitution of this expression into (2.2), dividing by dt, and letting dt approach zero yields a second-order differential equation which is solved by $C(P) = A_1P^{\beta_1} + A_2P^{\beta_2}$, where A_1 and A_2 are integration constants, and $\beta_1 > 1$ and $\beta_2 < 0$ are the roots of the quadratic equation $\frac{1}{2}\sigma^2\beta(\beta-1)+\alpha\beta-r = 0$. Clearly, the option to postpone the investment is worthless if the current output price is zero, i.e., $C(0) = 0$. Therefore A_2 must be zero, and hence,

$$C(P) = A_1P^{\beta_1}. \tag{2.3}$$

Note that $C(P)$ is increasing and convex in P.

[1] A quantity is said to be $o(dt)$ if $o(dt)/dt \to 0$ as $dt \downarrow 0$.

We can now characterize the optimal investment decision. The investor should undertake the investment if the expected utility of the cash flows accruing from the project exceeds the value of delaying it; otherwise, he should postpone the investment. Let P^* be the output price for which the investor is indifferent between investment and delay. Then

$$V(P^*) = C(P^*). \tag{2.4a}$$

Eq. (2.4a) is referred to as the value-matching condition. Furthermore, V and C should meet tangentially at P^*, that is,

$$V'(P^*) = C'(P^*), \tag{2.4b}$$

where V' and C' denote the partial derivatives of V and C with respect to P, respectively. Eq. (2.4b) is called the high-order contact or smooth-pasting condition. See Dixit and Pindyck (1994, pp. 130–132) for a discussion on smooth pasting.

Concerning existence and uniqueness of P^*, we were able to prove the following proposition.

Proposition 1 *Consider an investor who is either risk neutral or risk averse within the model outlined above. Then it holds that:*

1. *If there exists an output price P^* satisfying (2.4a) and (2.4b), it is unique.*

2. *Existence of P^* is guaranteed if the utility function is unbounded.*

Proof 1 *See Appendix A.*

Proposition 1 states that the existence of a critical output price implies its uniqueness. Hence, if there exists a critical output price, the optimal investment decision is tantamount to a simple investment rule: invest if $P > P^*$ and wait if $P < P^*$. If no critical output price exists, it is optimal never to invest, however high the current price level.

Figure 2.1 illustrates the investment decision graphically. It depicts V and C as functions of P. The critical output price is located at the point where V and C intersect. The functions are also tangent at this point. If the current output price is below this threshold, the investor defers the investment, and its value is equal to

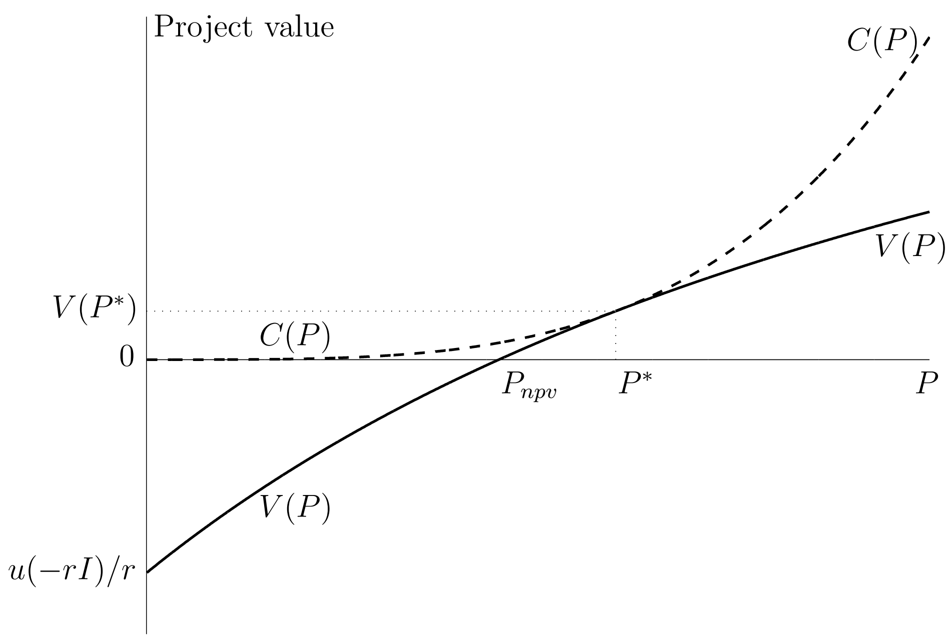

Figure 2.1: Graphical illustration of the optimal investment decision. The solid graph depicts V as a function of P. The dashed curve is C as a function of P. The critical output price P^* is located at the point where V and C are tangent and intersect. The *npv* critical price is located at the point where V intersects the P-axis.

the option value. If the current output price exceeds the threshold, the investment will be made, and its value is equal to the expected utility of its net cash flows.

Note that the *npv* critical output price (P_{npv}) is located at the point where the expected utility of the net cash flows produced by the project is equal to zero, i.e., $V(P_{npv}) = 0$. In fact, this is the relevant threshold if the investment project were reversible or when the investment decision is a now-or-never option. Clearly, the *npv* threshold is always smaller than the critical output price under irreversibility.

2.4 An example

In order to analyze the effects of changes in investors' attitudes toward risk on the optimal investment decision, we introduce the following utility function:

$$u(x) = (s - \eta)x + \eta(1 - e^{-x}), \tag{2.5}$$

where $s > 0$ and $\eta \in [0, s]$. This utility function is constructed as a linear combination of a risk neutral utility function and a constant absolute risk aversion (CARA) utility function with unit Arrow-Pratt measure (see, e.g., Mas-Colell, Whinston and Green (1995)). It is increasing and concave (for $\eta \neq 0$), and it meets the imposed normalization $u(0) = 0$. Moreover, it has the attractive feature that it incorporates risk neutrality as a special case (for $\eta = 0$). Hence, it allows us to compare the case of risk neutrality to the case of risk aversion.

Another important property of the utility function considered is that it exhibits non-increasing absolute risk aversion. The hypothesis of non-increasing absolute risk aversion was already propounded by Arrow (1970). It is supported by the empirical observation that the willingness to take small bets increases as individuals get wealthier. For $\eta \neq 0$, the Arrow-Pratt measure of absolute risk aversion is given by

$$R_A(x) = -\frac{u''(x)}{u'(x)} = \frac{1}{1 + \frac{s-\eta}{\eta}e^x}, \tag{2.6}$$

which is indeed decreasing in x. Another consequence of Eq. (2.6) is that the parameter η may be interpreted as a measure of the degree of risk aversion of the investor, since $R_A(x)$ is increasing in η for all x.

Under this specification, the expected utility of net cash flows resulting from investing is given by

$$V(P) = (s - \eta)\left[\frac{P}{r-\alpha} - I\right] + \eta\left[\frac{1}{r} - e^{rI}G(P)\right],$$

where

$$G(P) \equiv E\int_{t=0}^{\infty} e^{-rt}e^{-P_t}dt.$$

An explicit expression for $G(P)$ can be obtained by writing down the dynamic programming-like recursion expression (cf. Dixit and Pindyck (1994, pp. 315–316)):

$$\begin{aligned} G(P) &= e^{-P}dt + e^{-rdt}E\{G(P + dP_t)\} \\ &= G(P) + \left[\frac{1}{2}\sigma^2P^2G''(P) + \alpha PG'(P) - rG(P) + e^{-P}\right]dt + o(dt), \end{aligned}$$

which implies a second-order differential equation whose solution reads

$$G(P) = \frac{1}{r} \times \frac{P^{\beta_1}\Psi_1(P) + P^{\beta_2}\Psi_2(P)}{1/\beta_1 - 1/\beta_2},$$

where

$$\begin{aligned} \Psi_1(P) &\equiv \int_{\nu=P}^{D_1} \nu^{-\beta_1-1}e^{-\nu}d\nu \\ \Psi_2(P) &\equiv \int_{\nu=D_2}^{P} \nu^{-\beta_2-1}e^{-\nu}d\nu, \end{aligned}$$

with integration constants $D_1 \geq 0$ and $D_2 \geq 0$. In Appendix B it is shown that $D_1 = \infty$ and $D_2 = 0$.

While the utility function considered allows for an explicit expression for $V(P)$, the corresponding critical output price P^* cannot be solved for analytically. However, it can easily be computed numerically by means of traditional search algorithms given the parameters of the model. Note in particular that if $\eta = 0$, that is if the investor is risk neutral, then P^* is equal to the investment threshold discussed in Dixit and Pindyck (1994, p. 186):

$$P^* = \frac{\beta_1}{\beta_1 - 1}(r - \alpha)I. \tag{2.7}$$

2.5 Comparative statics

In this section we examine the influence of the parameters of the model on the investment decision. First we derive the comparative statics for the risk neutral case. Subsequently we analyze the comparative statics for the utility function introduced in Section 2.4.

2.5.1 Risk neutrality

We start by examining the effect of a change in the investment cost I on the critical output price under risk neutrality. Recall that the critical output price is given by (2.7) in the risk neutral case. A first, trivial observation is that P^* is proportionally increasing with the investment cost I. Next, consider the other parameters of the model: α, r, and σ^2. The partial derivative of P^* with respect to $x \in \{\alpha, r, \sigma^2\}$ is equal to

$$\frac{\partial P*}{\partial x} = \frac{I}{\beta_1 - 1}\left[\beta_1 \frac{\partial(r-\alpha)}{\partial x} - \frac{r-\alpha}{\beta_1 - 1}\frac{\partial \beta_1}{\partial x}\right].$$

Using $\frac{1}{2}\sigma^2\beta_i(\beta_i - 1) = r - \alpha\beta_i$ for $i = 1, 2$, we find

$$\frac{\partial \beta_1}{\partial x} = \begin{cases} -\frac{\beta_1}{\frac{1}{2}\sigma^2(\beta_1-\beta_2)} & \text{for } x = \alpha \\ \frac{1}{\frac{1}{2}\sigma^2(\beta_1-\beta_2)} & \text{for } x = r \\ -\frac{r-\alpha\beta_1}{\frac{1}{2}\sigma^4(\beta_1-\beta_2)} & \text{for } x = \sigma^2, \end{cases}$$

and, hence, the comparative statics in the risk neutral case are given by

$$\begin{aligned} \frac{\partial P^*}{\partial \alpha} &= -\frac{\beta_1}{\beta_1 - \beta_2} I \\ \frac{\partial P^*}{\partial r} &= \frac{1 + \beta_1 - \beta_2}{\beta_1 - \beta_2} I \\ \frac{\partial P^*}{\partial \sigma^2} &= \frac{1}{2}\beta_1 \frac{1 - \beta_2}{\beta_1 - \beta_2} I. \end{aligned}$$

To the best of our knowledge, this is the first time these comparative statics results have been analytically derived; Dixit and Pindyck only compute them numerically for certain sets of parameter values.

In particular, we find the following bounds on the partial derivatives:

$$\begin{array}{ccccccc} & & -I & < & \frac{\partial P^*}{\partial \alpha} & < & 0 \\ 0 & < & I & < & \frac{\partial P^*}{\partial r} & < & \frac{1+\beta_1}{\beta_1} I \\ 0 & < & \frac{1}{2} I & < & \frac{\partial P^*}{\partial \sigma^2}. & & \end{array}$$

Thus, in the risk neutral model, an increase in the drift term α always reduces the critical output price. That is, the utility of postponing the investment always decreases because its growth rate is higher. In contrast, an increase in the interest or discount rate raises the critical output price. Apparently, the discouraging effects of a rise in interest payments and a reduction in present value dominate the accelerating effect of higher impatience on investment. Moreover, the effect of a change in the interest rate on the critical output price is always greater than on the *npv* critical price. Furthermore, an increase in the volatility also raises the investment threshold: uncertainty adds to the value of waiting.

2.5.2 Risk aversion

We now analyze the comparative statics under risk aversion using the utility function defined in Section 2.4. In order to assess the impact of a change of a parameter x on the threshold price, it is useful to define $\varphi(P) \equiv \beta_1 V(P) - PV'(P)$. From (2.3), (2.4a), and (2.4b) we have $\varphi(P^*) = 0$. Total differentiation of $\varphi(P^*) = 0$ gives

$$\frac{\partial \varphi}{\partial P} \frac{\partial P^*}{\partial x} + \frac{\partial \varphi}{\partial x} = 0,$$

where all partial derivatives are evaluated at P^*. This implies that the influence of a change in parameter x on P^* is measured by

$$\frac{\partial P^*}{\partial x} = -\frac{1}{\varphi'(P^*)} \left. \frac{\partial \varphi(P)}{\partial x} \right|_{P=P^*}.$$

As pointed out in Appendix A, the function φ is strictly increasing on $[0, \infty)$, so $\varphi'(P^*) > 0$. Hence, the sign of the partial derivative of P^* with respect to x is opposite to the sign of the partial derivative of φ with respect to x evaluated at P^*.

As in the risk neutral case, we start by analyzing the effect of a change in the investment cost on the threshold price. Monotonicity and concavity of u are sufficient

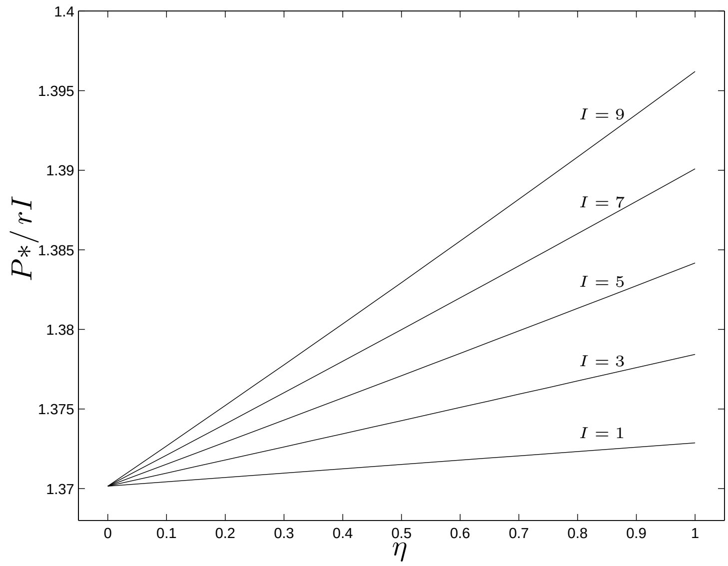

Figure 2.2: P^*/rI as a function of η for $I \in \{1, 3, 5, 7, 9\}$, $\alpha = 0$, $r = 0.05$, $\sigma = 0.1$, and $s = 1$.

to show that—not surprisingly—an increase in I raises the threshold price, while a decrease in I reduces it:

$$\frac{\partial \varphi(P)}{\partial I} = -rE \int_{t=0}^{\infty} e^{-rt} [\beta_1 u'(P_t - rI) - P_t u''(P_t - rI)] dt < 0,$$

and, hence, $\partial P^* / \partial I > 0$.

Such general statements are not possible with respect to the other parameters in the model, not even in the case of the utility function in Section 2.4. Therefore, we conduct a number of numerical analyses to find out the influence of these parameters on the optimal investment decision using this utility function. In particular, we are interested in the influence of a change in risk aversion.

Unless mentioned otherwise, we set $\alpha = 0$, $r = 0.05$, $\sigma = 0.1$, and $s = 1$. Figure 2.2 shows the threshold price to interest payment ratio P^*/rI as a function of the risk aversion parameter η for different values of the investment cost. For $\eta = 0$ the risk neutral case applies. In this case the threshold price is proportional to the investment cost. This implies that, for a given interest rate, the fraction P^*/rI is constant for different levels of I, which explains that in Figure 2.2 all curves coincide at $\eta = 0$. Figure 2.2 further shows that, for given I, P^* increases with η. This means

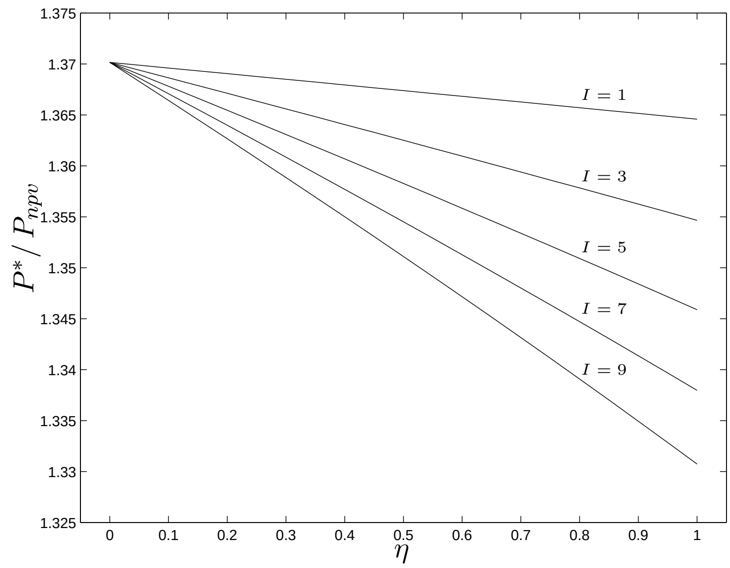

Figure 2.3: P^*/P_{npv} as a function of η for $I \in \{1, 3, 5, 7, 9\}$, $\alpha = 0$, $r = 0.05$, $\sigma = 0.1$, and $s = 1$.

that, the more risk averse the investor is, the higher must be the output price for investment to be optimal. We conclude that a risk averse investor is more cautious to invest. Moreover, this effect is reinforced by the size of the investment. This can be explained by the fact that concavity of the utility function implies that as the investment cost goes up, the disutility of a large negative cash flow becomes more and more important. Consequently, the larger the investment outlay, the more the investor needs to be compensated for by a higher critical output price relative to interest payments.

Figure 2.3 demonstrates that the wedge between P^* and the *npv* critical output price decreases with η. This means that the difference between the optimal investment decision and the decision based on the *npv* criterion shrinks the more risk averse the investor is. Hence, the error made by applying the *npv* rule, or, equivalently, the importance of irreversibility, becomes smaller under risk aversion. Again, the larger the investment cost, the stronger this effect becomes. The reason is that the large disutility of large investments plays a major role in case of a concave utility function, and this dominant factor affects P^* and P_{npv}.

Figure 2.4 shows P^* as a function of r. From this figure it can be concluded

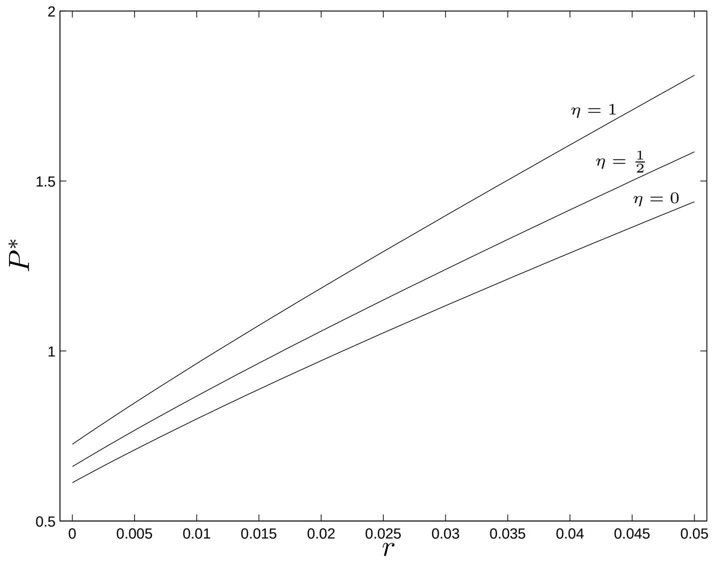

Figure 2.4: P^* as a function of r for $\eta \in \{0, \frac{1}{2}, 1\}$, $\alpha = 0$, $\sigma = 0.1$, $s = 1$, and $I = 1$.

that—as in the risk neutral case—the critical output price increases with r implying that it is less attractive to invest if the discount rate is larger. Another thing that emerges from Figure 2.4 is that risk aversion reinforces the influence of r on P^*. The reason is that, similarly to the dependence of the investor's utility on I, the disutility of large net cash outflows becomes more and more important for higher values of η.

Finally, we examine the effect of a change in price uncertainty on the investment decision. Figure 2.5 plots $P^*/\,rI$ as a function of σ for different levels of risk aversion. Clearly, an increase in the volatility of the output price causes the threshold price to grow. After all, the more uncertain the future revenues are, the more it pays to wait for more information concerning the development of output prices. Figure 2.5 shows that this effect intensifies under risk aversion. Interestingly, the effect becomes huge for high levels of risk aversion. Figure 2.6 demonstrates that the wedge between P^* and the *npv* critical output price widens with an increase in σ. Hence, the error made by deciding according to the *npv* rule exacerbates as price uncertainty rises. The figure shows that this effect can be quite substantial, but, as we already concluded from Figure 2.3, the effect is weaker when the level of risk aversion is higher.

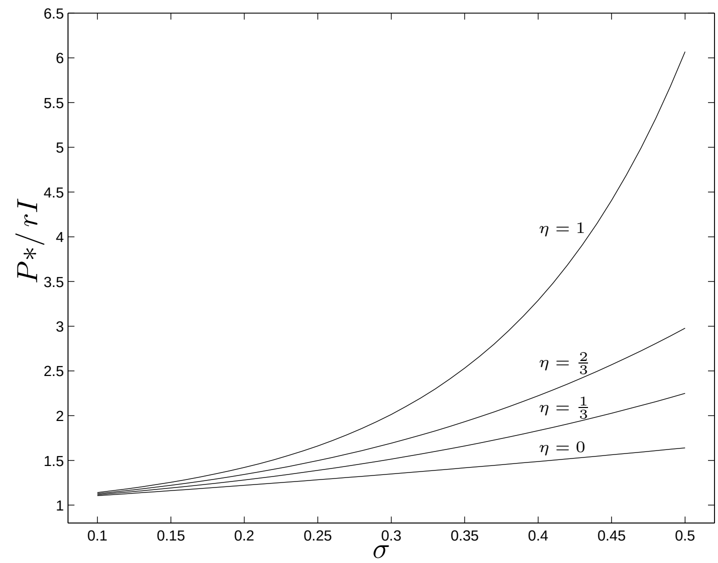

Figure 2.5: P^*/rI as a function of σ for $\eta \in \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, $\alpha = 0$, $r = 0.05$, $s = 1$, and $I = 1$.

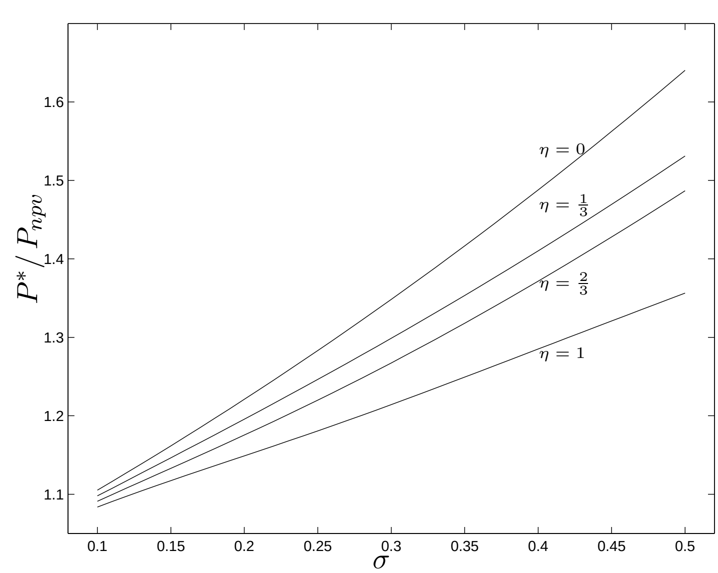

Figure 2.6: P^*/P_{npv} as a function of σ for $\eta \in \{0, \frac{1}{3}, \frac{2}{3}, 1\}$, $\alpha = 0$, $r = 0.05$, $s = 1$, and $I = 1$.

2.6 Conclusion

In this study we generalize the theory of irreversible investment under uncertainty by allowing for risk averse investors in a situation of incomplete markets. The model we use is similar to that of Dixit and Pindyck (1994, pp. 185–186), the only difference being that in their set-up the investment expenditure is immediately incurred, whereas in our model there is a flow of interest payments over the lifespan of the project. It is this adaptation that allows us to extend their model beyond risk neutrality using utility functions.

We have introduced a class of utility functions with non-increasing absolute risk aversion to examine the effects of risk aversion, price uncertainty, and other parameters on the optimal investment decision. We find that risk aversion reduces investment, particularly if the investment size is large. Moreover, we find that a rise in uncertainty increases the value of deferring irreversible investments, especially for high levels of risk aversion. Furthermore, we find that applying the net present value rule leads to better (although not optimal) decisions when the level of risk aversion is high. In addition, we provide closed-form comparative statics formulas for the risk neutral investor.

Finally, departing from the realistic situation of risk averse firms operating in an incomplete market setting, we list some ideas for further research. One of our main results is that risk aversion reduces the gap between the optimal decision and investing according to the net present value rule. Traditional real options theory shows that the gap is there because the option to wait for more information is valuable. Apparently, this option value is of less importance under risk aversion. It would be interesting to find out whether the gap shrinks even more when the behavior of competitors is taken into account, so that the incentive to preempt rivals will also play a role.

A second topic relates to investments in R&D. Since R&D investments create options (e.g., to produce cheaper or to commercialize patents), and option values increase with uncertainty, it is known from the real options literature (e.g., Dixit and Pindyck (1995)) that the incentive to invest in R&D goes up with uncertainty. It would be interesting to see to what extent this result still holds within our framework of risk aversion combined with incomplete markets.

A Proof of Proposition 1

In this appendix we show that there is a single, strictly positive critical output price (if it exists at all), whether the investor is risk neutral or risk averse. Risk aversion corresponds to a concave utility function; see, e.g., Mas-Colell et al. (1995, p. 187). This is equivalent to $u'' < 0$ since u is twice differentiable. Risk neutrality corresponds to a linear utility function. In that case, $u'' = 0$. In either case, we have $u' > 0$ and $u'' \leq 0$.

Assume there is a $P^* \in [0, \infty)$ such that (2.4a) and (2.4b) hold. Define $\varphi : [0, \infty) \to \mathbb{R}$, $\varphi(P) \equiv \beta_1 V(P) - PV'(P)$. Then, by construction, $\varphi(P^*) = 0$. Note that φ is strictly increasing on $[0, \infty)$:

$$\begin{aligned} \varphi'(P) &= (\beta_1 - 1)V'(P) - PV''(P) \\ &= E \int_{t=0}^{\infty} e^{-rt} [(\beta_1 - 1)u'(ncf_t) - P_t u''(ncf_t)] \frac{P_t}{P} dt \end{aligned}$$

is positive since $\beta_1 > 1$, $u' > 0$, and $u'' \leq 0$. Note furthermore, that $\varphi(0) = \beta_1 u(-rI)/r < \beta_1 u(0)/r = 0$ since $u' > 0$. Then, by continuity of the function φ, $P^* > 0$ and unique.

A sufficient condition for existence of P^* is unboundedness of the utility function. To see this, note that

$$\frac{\varphi(P)}{P} = E \int_{t=0}^{\infty} e^{-rt} \left[\beta_1 \frac{u(ncf_t)}{P_t} - u'(ncf_t) \right] \frac{P_t}{P} dt.$$

As $P \to \infty$, this ratio converges to $\frac{\beta_1 - 1}{r - \alpha} \lim_{x \to \infty} u'(x)$ which is positive if u is unbounded from above. Consequently, $\lim_{P \to \infty} \varphi(P) > 0$, which, together with $\varphi(0) < 0$, ensures there exists a $P^* \in (0, \infty)$ such that $\varphi(P^*) = 0$.

B Determination of D_1 and D_2

To determine the integration constant $D_1 \geq 0$, first note from the definition of G that $\lim_{P \to \infty} G(P) = 0$. This implies

$$\lim_{P \to \infty} \left[P^{\beta_1} \Psi_1(P) + P^{\beta_2} \Psi_2(P) \right] = 0. \tag{2.8}$$

Notice that $\lim_{P\to\infty} P^{\beta_2} = 0$ since $\beta_2 < 0$, and $\lim_{P\to\infty} \Psi_2(P) = \int_{\nu=D_2}^{\infty} \nu^{-\beta_2-1}e^{-\nu}d\nu \leq \int_{\nu=0}^{\infty} \nu^{-\beta_2-1}e^{-\nu}d\nu = \Gamma(-\beta_2)$ which is finite because $\beta_2 < 0$.[2] As a consequence, $\lim_{P\to\infty} P^{\beta_2}\Psi_2(P) = 0$. Therefore, in view of (2.8), it should hold that

$$\lim_{P\to\infty} P^{\beta_1}\Psi_1(P) = 0. \tag{2.9}$$

Suppose that $D_1 < \infty$. Then $\lim_{P\to\infty} \Psi_1(P) = -\int_{\nu=D_1}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu < 0$. Also, $\lim_{P\to\infty} P^{\beta_1} = \infty$ since $\beta_1 > 0$. Hence, $\lim_{P\to\infty} P^{\beta_1}\Psi_1(P) = -\infty$, which contradicts (2.9). Therefore, a necessary condition for (2.9) to hold is that $D_1 = \infty$. To see that this condition is also sufficient, consider

$$\lim_{P\to\infty} P^{\beta_1}\Psi_1(P) = \lim_{P\to\infty} \frac{\int_{\nu=P}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu}{P^{-\beta_1}}.$$

We can apply l'Hôpital's rule to this limit, for $\lim_{P\to\infty} \int_{\nu=P}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu = 0$ and $\lim_{P\to\infty} P^{-\beta_1} = 0$:

$$\begin{aligned}\lim_{P\to\infty} P^{\beta_1}\Psi_1(P) &= \lim_{P\to\infty} \frac{-P^{-\beta_1-1}e^{-P}}{-\beta_1 P^{-\beta_1-1}}\\ &= 0,\end{aligned}$$

so that, indeed, $D_1 = \infty$ is sufficient for (2.9), and thus (2.8) holds.

As for the other integration constant $D_2 \geq 0$, a similar reasoning holds. First, observe from the definition of G that $\lim_{P\downarrow 0} G(P) = \frac{1}{r}$. This implies

$$\lim_{P\downarrow 0} \left[P^{\beta_1}\Psi_1(P) + P^{\beta_2}\Psi_2(P)\right] = \frac{1}{\beta_1} - \frac{1}{\beta_2}. \tag{2.10}$$

Now that we know $D_1 = \infty$, consider

$$\lim_{P\downarrow 0} P^{\beta_1}\Psi_1(P) = \lim_{P\downarrow 0} \frac{\int_{\nu=P}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu}{P^{-\beta_1}},$$

to which we can apply l'Hôpital's rule, because of the fact that $\lim_{P\downarrow 0} P^{-\beta_1}$ and $\lim_{P\downarrow 0} \int_{\nu=P}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu = \int_{\nu=0}^{\infty} \nu^{-\beta_1-1}e^{-\nu}d\nu = \Gamma(-\beta_1)$ are both equal to ∞ since β_1 is positive:

$$\begin{aligned}\lim_{P\downarrow 0} P^{\beta_1}\Psi_1(P) &= \lim_{P\downarrow 0} \frac{-P^{-\beta_1-1}e^{-P}}{-\beta_1 P^{-\beta_1-1}}\\ &= \frac{1}{\beta_1}.\end{aligned}$$

[2] $\Gamma(\cdot)$ denotes the Euler gamma function, $\Gamma(a) \equiv \int_{\nu=0}^{\infty} \nu^{a-1}e^{-\nu}d\nu$, $a \in \mathbb{R}$.

Therefore, in view of (2.10), it should hold that

$$\lim_{P\downarrow 0} P^{\beta_2}\Psi_2(P) = -\frac{1}{\beta_2}. \tag{2.11}$$

Suppose $D_2 > 0$. Then $\lim_{P\downarrow 0}\Psi_2(P) = -\int_{\nu=0}^{D_2} \nu^{-\beta_2-1}e^{-\nu}d\nu < 0$. In addition, $\lim_{P\downarrow 0} P^{\beta_2} = \infty$. Hence, $\lim_{P\downarrow 0} P^{\beta_2}\Psi_2(P) = -\infty$, which contradicts (2.11). Therefore, a necessary condition for (2.11) to hold is that $D_2 = 0$. To see that this is also sufficient, consider

$$\lim_{P\downarrow 0} P^{\beta_2}\Psi_2(P) = \lim_{P\downarrow 0} \frac{\int_{\nu=0}^{P} \nu^{-\beta_2-1}e^{-\nu}d\nu}{P^{-\beta_2}}.$$

Again, l'Hôpital's rule can be applied, because $\lim_{P\downarrow 0}\int_{\nu=0}^{P} \nu^{-\beta_2-1}e^{-\nu}d\nu = 0$ and $\lim_{P\downarrow 0} P^{-\beta_2} = 0$:

$$\begin{aligned} \lim_{P\downarrow 0} P^{\beta_2}\Psi_2(P) &= \lim_{P\downarrow 0} \frac{P^{-\beta_2-1}e^{-P}}{-\beta_2 P^{-\beta_2-1}} \\ &= -\frac{1}{\beta_2}, \end{aligned}$$

so that, indeed, $D_2 = 0$ is sufficient for (2.11), and thus (2.10) holds.

Chapter 3

Economic Hedging Portfolios

3.1 Introduction

The purpose of the study presented in this chapter is to estimate and interpret the composition of hedging portfolios that investors hold on account of various economic risks. Furthermore, the study estimates and tests the significance of the hedging costs associated with these economic hedging portfolios.

We use a model of state-dependent preferences to show that economic hedging portfolios can be obtained as combinations of traded assets which mimic as far as possible the economic risk variables to which investors are exposed. The weights in these mimicking portfolios turn out to be a function of the level of risk aversion of investors. The weighting scheme implies that the composition of economic hedging portfolios is investor-specific, as is the associated premium investors pay—at least, if the risk variables under consideration cannot be perfectly replicated. This will, of course, typically be the case, as we generally observe an incomplete securities market, which makes it impossible to hedge all sources of risk perfectly.

Portfolios and premiums associated with economic risks have been studied by several authors in various contexts. For example, Breeden, Gibbons and Litzenberger (1989) test the consumption-based CAPM using a portfolio that has maximum correlation with consumption growth. Vassalou (2003) constructs a mimicking portfolio to proxy news related to future GDP growth to explain the cross-section of equity returns. Balduzzi and Kallal (1997) tighten the variance bounds of Hansen and Jagannathan (1991) using hedging portfolios for various economic risk variables. And Balduzzi and Robotti (2001) use the minimum-variance kernel of Hansen and Jagannathan to estimate economic risk premiums.

In all of these papers, the mimicking portfolios are constructed by means of an ordinary least squares projection of the risk variables on a set of security returns. As a consequence, portfolio weights and hedging costs are identical for all agents in these studies. In this study, however, hedging is achieved by a *weighted* least squares projection of the risk variables on the security returns, in which the weights depend on investors' appetite for risk, making the composition of hedging portfolios and the implied cost of hedging individual specific.

We derive these risk aversion-weighted hedging portfolios from a model of state-dependent preferences, in which economic risk variables enter the investor's utility function in addition to the return on financial wealth. In this framework, we define an investor's economic hedging portfolio as the difference between the expected utility maximizing investment portfolio and a portfolio constructed on the basis of the return on financial wealth only, i.e., in the absence of economic risk exposures. Using a linear approximation of the investor's first order optimality conditions, we show that the resulting hedging portfolio weights are in fact approximately equal to the regression coefficients in a weighted least squares regression of the economic risk variable on the available asset returns, in which the weights are proportional to the second derivative of the utility function.[1] The implied hedging cost is then the compensation investors are willing to pay for investing in a hedged position instead of a zero-exposure portfolio, in terms of expected return forgone.

Our approach is related to the literature on nonmarketable risks. Nonmarketable risks arise from positions in nontraded claims such as human capital (Mayers 1972) and commodities (Stoll 1979). As is well-known from mean-variance investment analysis with nonmarketable risks, an investor's optimal portfolio holdings can be split up into speculative demand (i.e., the standard Markowitz portfolio choice) and hedging demand due to the nonmarketable risks to which the investor is exposed. This hedging demand is an ordinary least squares projection of the nontraded risk

[1]Similar ideas have been applied by DeRoon, Nijman and Werker (2003) in the context of currency hedging for international stock portfolios.

onto the traded security returns. In fact, a more general utility framework would produce a non-orthogonal projection similar to the one in our state-dependent utility approach.

In the empirical analysis, we focus on economic risk variables that have been found to command risk premiums in empirical studies of multi-beta and of multi-factor models. We consider the inflation rate, real interest rate, term spread, default spread, dividend yield, and consumption growth. Similar variables have been used by, for instance, Chen, Roll and Ross (1986), Burmeister and McElroy (1988), Ferson and Harvey (1991), Campbell (1996), and Balduzzi and Kallal (1997). The possibilities for hedging these economic risks will, of course, depend on which traded assets are included in the analysis. We focus on a number of equity and bond factors of which it is well-known that they induce significant risk premiums. We include the Fama-French-Carhart factors in our set of securities to represent the stock market, and we use a two-factor model to represent the bond market. Using these priced risk factors, of several of which it is as yet not clear how they are related to economic fundamentals, allows us to explore the possibility that they are induced by an underlying hedging demand for economic risks.

We find that several stock-market and bond-market portfolios provide hedges for economic risks for a wide range of levels of risk aversion. In particular, inflation risk and real interest-rate risk can be partially hedged using corporate bonds; the term factor provides a good hedge for term-structure risk; and default risk can be partially hedged using bond portfolios. Bond portfolios, in combination with the equity market and momentum portfolios, also provide a good hedge against dividend-yield risk. Finally, the size portfolio appears to be useful for hedging consumption-growth risk.

Not all hedging instruments are equally useful in every situation, however. For some levels of risk aversion, portfolio adjustments are required in a particular security, while for other levels, no such adjustments are needed. For instance, a relatively risk tolerant investor may hedge against inflation risk by taking a short position in the momentum portfolio, while this is not true for relatively risk averse agents. Hence, introducing a risk aversion-dependent weighting scheme in the hedging problem can indeed lead to different hedging instruments being important for different types of investors, which is in contrast with what the more restrictive mean-variance analysis predicts.

Furthermore, we find that both inflation risk, real interest risk, and dividend-yield risk imply statistically and economically significant hedging costs, while there is no evidence of a compensation for hedging default risk, consumption-growth risk, or term-structure risk.

Finally, using a decomposition of investment portfolios into speculative and hedging demand, we find that deviations from two-fund separation, i.e. investments in only the risk-free asset and the market portfolio, can be attributed to hedges against economic risks. Our results show that the size factor can be attributed to hedges against consumption-growth risk; that the term factor in bond markets is related to hedges against real interest-rate, term-structure, default, and dividend-yield risk; and that the default factor in bond markets is related to hedges against default, dividend-yield, and consumption-growth risk. However, a complete explanation of anomalies remains elusive, as we find that part of investors' demand for assets is due to speculative motives.

The structure of the remainder of the chapter is as follows. Section 3.2 describes the model and its implications for investors' hedging demand due to economic risks as well as the associated risk premiums. In Section 5.3 we discuss the data on securities returns and economic risk variables which are used in Section 4.3 to estimate and test the significance of risk premiums and hedging portfolios associated with economic risks. Furthermore, we investigate whether hedging motives can explain the premiums on the Fama-French portfolios. Section 5.6 concludes.

3.2 Hedging economic risks

Assume that K risky securities are traded, and a risk-free one. Let R_t denote the K-vector of gross returns on the risky securities from date $t-1$ to date t, and let $R_{f,t-1}$ be the gross risk-free rate of return from date $t-1$ to date t. Under the law of one price, there exist stochastic discount factors or pricing kernels M_t that satisfy

$$E_{t-1}\left[M_t R_t\right] = \iota_K \tag{3.1}$$

and

$$E_{t-1}\left[M_t\right] = \frac{1}{R_{f,t-1}}, \tag{3.2}$$

with ι_K being a K-vector of ones, and where E_{t-1} denotes the conditional expectation given all information up to date $t-1$; see, e.g., Cochrane (2001). If, furthermore, there are no arbitrage opportunities, then there is at least one such pricing kernel which is strictly positive almost surely.

It is well-known that stochastic discount factors or pricing kernels can be thought of as investors' marginal utility. Consider a risk-averse investor who maximizes the expected utility of the gross return on his wealth, $R_{W,t}$, by choosing his investments in the $K+1$ available securities according to

$$\begin{array}{c} \max_w E_{t-1}\left[u(R_{W,t})\right] \\ \text{s.t. } R_{W,t} = R_{f,t-1} + w^\top R_t^e, \end{array} \tag{3.3}$$

where $R_t^e \equiv R_t - R_{f,t-1}\iota_K$ is the K-vector of excess returns on the risky securities. Note that the w's need not sum to one. The first order conditions of problem (3.3) imply that a valid stochastic discount factor is

$$\frac{u'(R_{f,t-1} + w_0^\top R_t^e)}{R_{f,t-1}E_{t-1}[u'(R_{f,t-1} + w_0^\top R_t^e)]}, \tag{3.4}$$

with marginal utility being evaluated at the optimal portfolio choice w_0. Note that positive marginal utility implies the absence of arbitrage opportunities.

We extend this simple portfolio problem by allowing for state-dependent utility, in which sources of risk other than the uncertain security returns may affect the investor's utility. Typically, these sources of risk are investor-specific. In principle, they can be anything from human capital and illiquid equity to health risk and the weather. In this study, however, we focus on a set important (macro)economic risk variables such as inflation, the term spread, and consumption growth, following, for example, Chen et al. (1986), Ferson and Harvey (1991), Campbell (1996), and Balduzzi and Kallal (1997).

To be more precise, let y_t be an economic risk variable, and write an investor's state-dependent utility as $\mathcal{U}(R_{W,t}; y_t)$. Hence, the investor's utility does not only depend on the return on his invested wealth, but also on the realization of the economic risk variable. We assume that the risk variable enters the individual's utility function linearly:

$$\mathcal{U}(R_{W,t}; y_t) = u(R_{W,t} - qy_t), \tag{3.5}$$

where q is a parameter reflecting the extent to which the investor cares about the economic risk under consideration. Relation (3.5) can also be interpreted as a linearization of $\mathcal{U}(R_{W,t}; y_t)$, with $q = -\mathcal{U}_y(R_{W,t}; y_t)/\mathcal{U}_R(R_{W,t}; y_t)$.

To motivate this specification, consider the rate of inflation of the investor's consumption as an economic risk variable, and assume, for now, that q equals unity. Then the argument of the utility function can be interpreted as the individual's real return on wealth (taking R_W to be the nominal return on wealth). Depending on the investor's inclination to look at real returns rather than nominal returns, parameter q may assume other values. In particular, $q = 0$ may be interpreted as the investor being prone to money illusion.

More generally, any economic risk may affect the individual's utility of wealth. For instance, an interest rate shock can have an effect on the investor's utility of wealth, perhaps through his positions in non-tradable assets, such as a mortgage. Similarly, default risk can affect utility as bankruptcy jeopardizes one's labor income. Furthermore, a change in dividend yield may cause one's investment opportunity set to shift (dividend-yield risk), as well as an unanticipated fall in consumption growth (business cycle risk).

We will refer to q as the individual's *exposure* to the economic risk, by analogy with the literature on non-marketed securities mentioned in the introduction. Note that in case of zero exposure, the utility function reduces to the one considered in (3.3). In case of non-zero exposure, however, the economic risk will affect the investor's portfolio choice, and, hence, give rise to hedging demand.

The portfolio choice problem now becomes:

$$\begin{array}{c} \max_w E_{t-1}\left[u(R_{W,t} - qy_t)\right] \\ \text{s.t. } R_{W,t} = R_{f,t-1} + w^\top R_t^e, \end{array} \tag{3.6}$$

and the corresponding first order conditions read

$$E_{t-1}[u'(R_{f,t-1} + w_1^\top R_t^e - qy_t)R_t^e] = 0_K, \tag{3.7}$$

where w_1 denotes the vector of optimal portfolio weights. We take a first order Taylor series approximation around the optimal portfolio in case of no exposure, i.e., around $w_1 = w_0$ and $q = 0$, to obtain

$$\begin{aligned} 0_K &= E_{t-1}[u'(R_{f,t-1} + w_1^\top R_t^e - qy_t)R_t^e] \\ &\approx E_{t-1}[u'(R_{f,t-1} + w_0^\top R_t^e)R_t^e] \\ &\quad + E_{t-1}[u''(R_{f,t-1} + w_0^\top R_t^e)R_t^e((w_1 - w_0)^\top R_t^e - qy_t)] \\ &= 0_K - E_{t-1}[R_t^e \Omega_t R_t^{e\top}](w_1 - w_0) + E_{t-1}[R_t^e \Omega_t y_t]q, \end{aligned} \quad (3.8)$$

where $\Omega_t \equiv -u''(R_{f,t-1} + w_0^\top R_t^e) > 0$, and the last equality follows from the first order conditions of problem (3.3). Hence, the difference in optimal portfolio weights per unit of exposure is

$$\frac{w_1 - w_0}{q} \approx E_{t-1}[R_t^e \Omega_t R_t^{e\top}]^{-1} E_{t-1}[R_t^e \Omega_t y_t]. \quad (3.9)$$

Formula (3.9) tells us how an individual's investment portfolio should be reallocated on account of his exposure to economic risk; some assets will require additional investment, while others will require less. Hence, this portfolio of incremental (dis)investments constitutes the investor's hedging demand associated with the economic risk variable under consideration. Accordingly, we refer to (3.9) as an investor's *economic hedging portfolio.*[2]

To further elaborate on this hedging interpretation, note that the expression on the right-hand side of equation (3.9) is equal to the vector of regression coefficients in a weighted least squares regression of the economic risk variable on the excess returns R_t^e, in which the weight is given by the negative of the second derivative of the utility function evaluated at the zero-exposure optimum:

$$y_t = \delta^\top R_t^e + \varepsilon_t, \quad (3.10)$$

where $E_{t-1}[R_t^e \Omega_t \varepsilon_t] = 0_K$ and $\delta = (w_1 - w_0)/q$. This regression is, in effect, an approximate replication of the economic risk variable using the set of traded securities; the investor hedges his exposure to the economic risk by taking an offsetting

[2]Note that this economic hedging portfolio does not have the interpretation of a "pure" hedge as in Anderson and Danthine (1981), in the sense that it minimizes the variance of the return on wealth. In our more general expected utility framework we cannot speak of such a pure hedge, as other moments of the distribution matter as well.

position in a portfolio that mimics the economic risk variable best. By investing in this economic hedging portfolio, the investor essentially minimizes the weighted expected squared hedging error ε_t:

$$\min_{\delta} E_{t-1}[\Omega_t \varepsilon_t^2]. \tag{3.11}$$

Weighting by the concavity of the utility function implies that for utility functions with an upward sloping second derivative (like, for instance, power utility), large negative returns on wealth get a large weight, while large positive returns get a small weight. This makes sense intuitively, since risk-averse investors will want their hedge against economic risk to work best when wealth is low, whereas the quality of the hedge is less important to the investor when wealth is high.

It is well-known that in a traditional mean-variance framework, hedging demand is independent of the level of risk aversion. Hence, for mean-variance investors, the weight is constant, and the hedging problem reduces to an ordinary least squares projection. Ergo, in this special case, heterogeneity of risk preferences is not an issue. Many theoretical papers, including Mayers (1972), Stoll (1979), Anderson and Danthine (1980), Anderson and Danthine (1981), and Hirshleifer (1989), effectively adopt this restrictive assumption. Moreover, Balduzzi and Kallal (1997) and Balduzzi and Robotti (2001) also make use of unweighted hedging.[3] However, weighted hedging is important for non-mean variance utility, as our results show.

Given the above analysis, it is natural to define the *implied hedging cost* associated with the economic risk variable as the expected excess return on the corresponding economic hedging portfolio:

$$\lambda_{t-1} \equiv \delta^\top E_{t-1}[R_t^e]. \tag{3.12}$$

The implied hedging cost is the expected return an investor with preferences described by u is willing to give up to hold a position that is hedged against economic risk. Equivalently, it is the required compensation for an investor providing the hedge in terms of additional expected return.

[3]Anderson and Danthine (1981) do mention the possibility of a general expected utility formulation, but they do not explore the issue further. Neither do they examine the empirical implications of weighted hedging.

Balduzzi and Kallal (1997) and Balduzzi and Robotti (2001) refer to the implied hedging cost as an economic risk premium. The term risk premium, however, suggests the existence of an equilibrium price of economic risk that is the same for all agents. Clearly, the implied hedging cost does not have an equilibrium interpretation, since the underlying economic risk is typically not traded. Rather, the implied hedging cost is a compensation for economic risk that is required by an individual investor. For this reason we avoid the use of the term risk premium.[4]

In Section 4.3 we examine the implied hedging costs associated with several economic risk variables using investments in both stocks and bonds. Furthermore, we analyze the composition of the underlying hedging portfolios.

3.3 Description of the data

This section describes the data used in the empirical analysis. The data is at a monthly frequency, and the period considered is August 1960 through December 2001, giving a total of 497 months.

3.3.1 Securities returns

The set of traded securities we consider includes the three factor portfolios of Fama and French (1992)—market, size, and book-to-market value—as well as the momentum portfolio of Carhart (1997). These factors have been found to explain the premiums on stocks. Furthermore, following Fama and French (1993), we include two bond-market factors: a term factor (the difference between a long-term government bond return and the one-month T-bill rate) and a default factor (the difference between the return on a portfolio of long-term corporate bonds and a long-term government bond return). The one-month T-bill rate is used as a proxy for the risk-free

[4] Balduzzi and Kallal (1997) and Balduzzi and Robotti (2001) do recognize that the implied hedging cost depends on (marginal) utility and, hence, the selected pricing kernel. In fact, Balduzzi and Kallal (1997) analyze the *bounds* on economic risk premiums, for given levels of the pricing-kernel variance. Moreover, they compare these bounds to the kernel of a representative consumer with power utility. Balduzzi and Robotti (2001) use a very specific kernel (the minimum-variance kernel of Hansen and Jagannathan), which leads to premiums that are equal for all agents.

rate.

The market (*RM–RF*), size (*SMB*), book-to-market value (*HML*), and momentum (*UMD*) portfolios are from Kenneth French' data library.[5] The bond factors (*TERM* and *DEF*) are constructed using long-term government and corporate bond series from Ibbotson and Associates, and the risk-free rate (*RF*) which is also from French.

Table 3.1 reports summary statistics for the securities data.[6] These are very much in line with the results reported by other authors. The assets considered cover a fairly wide range of average returns. The market risk premium was about 49 basis points per month on average in our sample period, which corresponds to about 6 percent annually. Only the risk-free rate exhibits strong positive autocorrelation; the risky returns are typically not very autocorrelated. The size portfolio and the term factor are positively correlated with the market portfolio, while book-to-market value has a sizeable negative correlation with the market. The bond-market factors, *DEF* and *TERM*, are strongly negatively correlated, which is due to the fact that they are constructed using the same long-term government bond.

3.3.2 Economic risk variables

We consider six (macro)economic risk variables that have also been used in previous studies. See, for example, Chen et al. (1986), Burmeister and McElroy (1988), Ferson and Harvey (1991), Campbell (1996), Balduzzi and Kallal (1997), and Balduzzi and Robotti (2001). They are:

1. *Inflation* (*INF*): The monthly net rate of inflation.

2. *Real interest* (*RI*): The monthly real net return on a one-month T-bill.

[5] These are acronyms for "small minus big" (*SMB*), "high minus low" (*HML*), and "up minus down" (*UMD*).

[6] The risky securities we consider are all zero-cost portfolios, but some of them are financed at the risk-free rate, while others are financed using other short positions. Nevertheless, we can take R^e to be equal to the selected vector of excess returns, and the analysis of Section 3.2 continues to apply. The only difference is in the interpretation of the portfolio weights. In particular, the fraction of wealth invested in the risk-free rate is one minus the fractions invested in *RM–RF* and *TERM*, and the fraction of wealth invested in the long-term government bond is equal to the difference between the portfolio weights invested in *TERM* and *DEF*.

3. *Term spread* (*TS*): The yield spread between a long- and a short-term government bond.

4. *Default spread* (*DS*): The difference in yields between corporate bonds rated Baa by Moody's Investor Service and Aaa corporate bonds.

5. *Dividend yield* (*DIV*): The monthly dividend yield on the S&P 500 composite.

6. *Consumption growth* (*CG*): Monthly real per-capita consumption growth of du-rables, nondurables, and services.

The monthly inflation rate is provided by Ibbotson and Associates and is computed as the relative change of the consumer price index for all urban consumers. The monthly real interest rate is the CRSP one-month T-bill rate deflated by *INF*, the inflation rate. The default spread and the term spread are constructed using government bond-yield series (10-year and 1-year) and corporate bond-yield series (Baa and Aaa) obtained from the Federal Reserve Statistical Release. The dividend yield and consumption growth series are obtained from Datastream.

Summary statistics for the six economic risk variables are provided in Table 3.2. Note that the risk variables are much less volatile than the security returns, and that they are typically highly autocorrelated. Only consumption growth is negatively autocorrelated at the first lag, which is consistent with previous research (e.g., Balduzzi and Kallal (1997)). A clear pattern emerges from the correlation matrix of the risk variables. In particular, note the high negative correlation between the inflation rate and the real risk-free rate, which is not surprising given that the real risk-free rate is equal to the nominal risk-free rate less the rate of inflation, and the nominal risk-free rate is relatively constant over our sample period. Also note the strong positive correlations between the dividend yield on the one hand, and the default spread and the inflation rate on the other, as well as the negative correlation between the term spread and the inflation rate.

3.4 Hedging portfolios and implied hedging costs

In this section, we compute and interpret the hedging portfolios and implied hedging costs associated with the six economic risk variables under scrutiny using the available set of traded assets. As a first step in the analysis, we estimate a vector autoregressive (VAR) model for the "raw" economic risk variables, and use the residuals as our actual economic risk variables, as in Campbell (1996). The reason for this is that we are only interested in hedging the unanticipated components of economic risks (shocks); any anticipated part can be hedged trivially using the risk-free asset.

Table 3.3 reports the coefficients in a first-order VAR, as well as the standard deviations and the correlations of innovations to the system. Many variables enter significantly with either positive or negative signs in the forecasting equations. In particular, the regression coefficients on the dependent variables' own lags are all highly significant due to the substantial autocorrelation in the economic variables. The autocorrelation is most pronounced in the term spread, the default spread, and dividend yield, explaining the high R^2 in those regressions. The innovations in inflation and the real interest rate are highly negatively correlated, while the correlations of the other innovations are on average less than 10 percent.

The economic hedging portfolios and their corresponding hedging costs are estimated in two steps. In the first step, we use the generalized method of moments (GMM) of Hansen (1982) to estimate the optimal zero-exposure portfolio weights for investors in a standard constant relative risk aversion or power utility framework. The power utility function is given by $u(x) = x^{1-\gamma}/(1-\gamma)$, where $\gamma > 0$ is the parameter of risk aversion which we allow to vary. These zero-exposure portfolio weights are subsequently used in the second step in the weighted least squares regression to obtain estimates of the hedging portfolios and the implied hedging costs. This procedure involves an errors-in-variables problem and requires an adjustment of the standard errors. The econometric details are given in the appendix.

3.4.1 Implied hedging costs

The hedging costs associated with the economic risk variables are in Panel A of Table 3.4. These hedging costs are measured in units of risk, or Sharpe ratios, as in Balduzzi and Kallal (1997); that is, the vector-autoregressive residuals are scaled by their standard deviations, so that they can be compared to each other. To get an idea of the order of magnitude of these implied hedging costs, or their economic significance, consider that the monthly market Sharpe ratio is about 0.11 for the period under scrutiny.

Our results show that investors are willing to pay for inflation shocks and innovations in dividend yield only. This holds for all types of agents, with levels of risk aversion ranging from $\gamma = 1$ to $\gamma = 20$. The estimated hedging costs for inflation and dividend yield are both statistically and economically significant. Both are negative, indicating that investors must forgo expected return if they want to hedge long exposures to these economic risks. Note that the implied hedging cost for inflation decreases as we consider more risk averse investors. This result may at first sight seem counterintuitive, however the magnitude of the hedging cost is in fact not determined by the level of risk aversion directly, but rather by its effect on the weight, Ω, in the hedging problem. Neither size nor sign of this effect can be predicted without examination of the data. An increase in the level of risk aversion, which makes the investor put more weight on low returns than on high returns, apparently decreases the slope in the hedging regression (in absolute value) and thus reduces the associated hedging cost for the case of inflation risk.[7] Contrary to inflation, the hedging cost associated with dividend yield seems to be independent of people's attitudes toward risk. We find no evidence for significant risk compensations for other economic risk variables.

Apart from a single risk exposure, agents may very well be exposed to several economic risks simultaneously. This implies that hedging portfolios are constructed to hedge for multiple risks. The resulting hedging costs are linear combinations of the hedging costs in Panel A of Table 3.4. These then constitute the price for simultaneously hedging for several economic risks.

An alternative way of analyzing risk premiums is to look at an innovation in isolation, disregarding innovations in other risk variables. To achieve this, we follow Campbell (1996) and Balduzzi and Kallal (1997) by orthogonalizing the VAR-residuals using a Cholesky decomposition of their variance-covariance matrix. The first innovation, the one in the rate of inflation, is unaffected by this procedure; the other innovations are. The orthogonalized innovation in the real interest rate

[7] Note that there is also no reason why the implied hedging cost should increase with risk aversion. In fact, in a mean-variance framework, the hedging cost is independent of the investor's risk aversion.

is equal to the part of the original real interest rate innovation orthogonal to the innovation in inflation; the orthogonalized innovation in term-structure risk is equal to the part of the original innovation in term-structure risk orthogonal to the innovations in inflation and the real interest rate; et cetera. The variables are ordered in such a way that the orthogonalized innovations are easily interpretable. For instance, the orthogonalized innovation in the real interest rate is a change in the real interest rate that is not caused by a change in the inflation rate. Hence, it amounts to a shock in the nominal rate.

The hedging costs related to the orthogonalized economic innovations are in Panel B of Table 3.4. We find that both inflation risk and dividend yield still require economically and statistically significant hedging costs for a broad range of levels of risk aversion. In addition, shocks to the real interest rate that are unrelated to inflation surprises also require a negative and significant hedging cost for all types of investors considered. This cost is quite sizeable for relatively risk tolerant agents, but gets smaller for higher levels of risk aversion.

3.4.2 Economic hedging portfolios

Table 3.5 reports the hedging portfolios underlying the hedging costs associated with each of the economic risk variables. Several securities provide hedges for economic risks. For instance, a significant long position in the default portfolio is required to hedge against inflation risk. That is, investors prone to inflation risk should reduce their investment in government bonds and buy corporate bonds. This is because when inflation is higher than anticipated, the return on the default factor is high. This result holds for a broad range of levels of risk aversion.

Observe, however, that a hedge against inflation requires other portfolio adjustments as well. For instance, the momentum portfolio appears to be a useful hedging instrument for relatively risk tolerant agents, while the market portfolio provides inflation protection for relatively risk averse investors. This shows that differences in risk aversion can have such an effect on the weighting in the hedging problem, that some securities turn out to be good hedges for certain types of agents, while others do well for other types of agents.

Note furthermore, that the total (dis)investment in the hedging portfolios depends on the investor's exposure to inflation risk. Naturally, if the exposure is zero, no hedging is needed, while in case of a non-zero exposure, some portfolio adjustments are required. Table 3.5 shows the hedging portfolios for an investor with unit exposure to innovations in the economic risk variables. Hence, an investor with relative risk aversion of one who cares about real returns instead of nominal returns, i.e., an investor with $\gamma = 1$ and $q = 1$, should increase his investment in the default portfolio by almost 5 percentage points. For investors with $\gamma = 1$ and $q = .5$, the adjustment is half this amount.

For investors who face real interest-rate risk, a short position in the default portfolio is required. Hence, an exposure to the real interest rate can be offset by a disinvestment in corporate bonds and an investment in government bonds, as the return on the default portfolio is low. Moreover, risk averse investors ($\gamma = 1$) should increase their position in the momentum factor.

Furthermore, the term portfolio provides a good hedge for term-structure risk across all levels of risk aversion. That is, when there is a shock in the interest rate differential, investors with an exposure to term-structure risk ought to use the term factor as a hedging instrument. For instance, investors whose portfolio is adversely affected by a high long-term interest rate and a low short-term interest rate, perhaps due to a mortgage loan and a savings account, can hedge the risk of a high interest rate differential by increasing their investment in long-term bonds and decreasing their investment in T-bills, because the excess return on long-term bonds is expected to be higher at such times.

As for hedges against default risk, most investors (if they face an exposure to this economic risk) seem to be best off taking long positions in the term portfolio and the default portfolio. Hence, corporate bonds appear to perform best when the risk of default is high, that is, if the yield spread rises.

For all levels of risk aversion considered, the hedging portfolio associated with dividend yield requires significant short positions in the market and momentum portfolios as well as long positions in the term and default portfolios. Note that an unanticipated increase in the dividend yield usually coincides with an unexpected drop in stock prices. Hence, investors can offset such a price drop by selling the market and momentum portfolios and buying relatively cheap bonds.

Finally, hedging consumption-growth risk calls for a disinvestment in the size portfolio. Hence, if an investor fears an unexpected drop in consumption growth, he had better avoid small company stocks, and increase his investment in large company stocks.

Note that while it is true that the magnitude and significance of hedging positions typically dies down with increasing risk aversion, many individual security positions remain economically and statistically significant even for high levels of risk aversion. It is also interesting to note that most hedging portfolio weights have the same sign for different risk-aversion levels. Only rarely do the weights change sign, and when they do, statistical significance disappears.

3.4.3 Speculative versus hedging demand

We can extend our model by not only considering investors' hedging demand, but also their speculative demand for assets, analogously to the mean-variance case, as studied in, e.g., Anderson and Danthine (1980, 1981). In case of mean-variance investors, one may break down investors' total demand for assets into a pure speculation component, which is equivalent to the position of an investor with no exposure to exogenous risk, and a pure hedge component, which is equal to the position of an infinitely risk-averse investor. In our more general setting, we cannot make this distinction, as investors' hedges against economic risk will in general depend on the concavity of their utility functions, and hence not be "pure" in the sense of Anderson and Danthine (1981). Nevertheless, we can separate investors' demand for assets due to speculative motives, and their demand for assets due to hedging, with both components being risk-aversion dependent.

For reasons that will become clear shortly, we define speculative demand as the set of (dis)investments in the available assets relative to a position in just the risk-free asset and the market portfolio. Hence, we look at the portfolio choice problem from the point of view of an agent who invests according to the premise of two-fund separation which follows from the capital asset pricing model. Moreover, we consider the possibility that this investor has an exposure to one or more economic risks. The economic hedging demand that is induced by this exposure may potentially shed light on the Fama-French anomalies, to the extent that the Fama-French premiums may in fact be explained by hedges against economic risk.

As suggested before, let us consider an agent who invests only in the risk-free asset and the market portfolio. Write the vector of available risky excess returns, R^e, as

$$R^e = \begin{bmatrix} R^e_m \\ R^e_F \end{bmatrix},$$

where R^e_m denotes the excess return on the market portfolio and R^e_F are the returns on the (other) Fama-French portfolios. Consider the optimal portfolio choice in case of an exposure q to economic risk with unrestricted investment opportunities (i.e., the investor may choose all available traded assets), and the optimal portfolio choice in case of no exposure with the restricted investment opportunity set (the CAPM investor):

$$\begin{aligned} w_1 &= \arg\max_{w} E[u(R_f + w^\top R^e - qy)] \\ w_{m,0} &= \arg\max_{w_m} E[u(R_f + w_m R^e_m)]. \end{aligned}$$

The corresponding first-order conditions are

$$\begin{aligned} 0_K &= E[u'(R_f + w_1^\top R^e - qy)R^e] \\ 0 &= E[u'(R_f + w_0^\top R^e)R^e_m], \end{aligned}$$

where $w_0 \equiv (w_{m,0}, 0^\top_{K-1})^\top$. A first-order Taylor expansion around $w_1 = w_0$ and $q = 0$ implies:

$$\begin{aligned} w_1 - w_0 \approx\ & E[R^e\Omega_0 R^{e\top}]^{-1} E[u'(R_f + w_0^\top R^e)R^e] \\ & + E[R^e\Omega_0 R^{e\top}]^{-1} E[R^e\Omega_0 y]q, \end{aligned} \tag{3.13}$$

where $\Omega_0 = -u''(R_f + w_0^\top R^e)$. Note that the weight Ω_0 will be different from the weight obtained in Section 3.2, since the restricted model will imply a different optimal zero-exposure portfolio choice than the unrestricted model. The first term on the right-hand side of (3.13) can now be interpreted as speculative demand, and the second term can be interpreted as an economic hedging component. Note that speculative demand depends on the term $E[u'(R_f + w_0^\top R^e)R^e]$, which is proportional to the generalized Jensen measure $E[M_0 R^e]$, where $M_0 \equiv u'(R_f + w_0^\top R^e)/R_f E[u'(R_f + w_0^\top R^e)]$ is the stochastic discount factor of an investor restricted to the market portfolio and the risk-free asset. It measures the attractiveness of new investments relative to a set of reference assets (in this case the market portfolio and the risk-free asset); a positive value indicates that the investor can improve his expected utility by going long in the new investment, whereas a negative value implies a short position. See, for instance, Glosten and Jagannathan (1994) and Chen and Knez (1996). Since the market portfolio is our point of reference, the first element of this vector is equal to zero.

We find significant speculative demand for the various assets relative to the CAPM portfolio. The market portfolio, the book-to-market value portfolio, and the momentum portfolio all require significant additional investments on account of speculative motives. As expected, the size of this speculative demand decreases with risk aversion, but the effect remains statistically significant. Hence, we conclude that the CAPM does not hold. We do not find significant speculative demand for the size portfolio and the two bond portfolios.

Table 3.6 reveals that speculative motives are not the only reason for people to diverge from the CAPM. There is significant hedging demand for almost all assets by various agents. For instance, investors with a unit exposure to inflation risk require a short position of about three quarters of a percent in the market portfolio. This result is quite robust across different levels of risk aversion. In fact, most hedging positions are, contrary to the hedging portfolios in Table 3.5, independent of risk aversion.

Note that only the market portfolio provides a good hedge in case of an inflation-risk exposure, whereas in the unrestricted case, the momentum and default portfolios do, too. Apparently, the weighting scheme implied by the CAPM makes these assets less useful as hedging instruments than they are for the unrestricted investor. Different weighting schemes are also the reason for the differences we find in the hedging portfolios associated with real interest-rate risk. For the CAPM case, a long position in the term portfolio is required; in the unrestricted case, a hedge is obtained by shorting the default portfolio (and for relatively risk tolerant investors, going long in the momentum portfolio). For the other risks, we find that by and large the same assets turn up as useful hedging instruments as in the unrestricted case. Hence, restricting the investment opportunity set does not have an important effect on the hedging portfolios in those cases. An exception is consumption-growth risk, in which case, in addition to a long position in the size portfolio, a long position in the default portfolio is called for. The only asset which does not show up as a useful hedging instrument in Table 3.6 is the momentum portfolio.

The results in Table 3.6 have interesting implications for the *raison d'être* of the Fama-French risk premiums. While we find that some of investors' demand for assets is due to speculative motives, part of the reason why they deviate from the CAPM is attributable to economic risks. In particular, we find no significant speculative demand for the size portfolio, the term portfolio, and the default portfolio. Therefore, investments in these portfolios can be entirely explained by hedging. Our results suggest that the size premium is related to hedges against consumption-growth risk; that the term premium in bond markets is caused by hedging against real interest-rate, term-structure, default, and dividend-yield risk; and that the default premium in bond markets is related to hedging against default, dividend-yield, and consumption-growth risk. The anomalies in the investments for which we do find significant speculative demand can still be partly explained by economic risks. Only the momentum portfolio seems to be unrelated to any of the economic risks included in the analysis.

3.5 Conclusion

In this chapter we estimate and interpret the composition of portfolios that investors hold to hedge various economic risks. We also consider the implied hedging costs associated with these economic hedging portfolios for various types of agents. We wish to stress that these hedging portfolios are individual specific. Using a model of state-dependent utility, we show that agents' economic hedging portfolios can be obtained by a risk aversion-weighted least squares regression of the economic risk variables onto the available risky security returns, as opposed to the unweighted hedging demand one obtains in the traditional mean-variance framework.

We find that agents across a broad range of levels of risk aversion are willing to pay (or demand) significant compensations for hedges against three sources of economic risk: inflation risk, real interest-rate risk, and dividend-yield risk. Furthermore, our results show that all economic risk variables we consider require a significant hedging adjustment with respect to one or more traded securities. Some of these securities prove to be useful hedging instruments across different types of investors, whereas others only serve as hedges for particular levels of risk aversion, which demonstrates the empirical relevance of risk aversion-weighted hedging.

Furthermore, we contribute to the discussion on asset pricing anomalies by examining whether the Fama-French premiums can be attributed to economic hedging motives. While we cannot conclude that book-to-market and momentum anomalies are (solely) due to reasons of economic hedging, we do find that the size effect found in stock markets as well as the term and default premiums found in bond markets, may potentially be explained by hedges against economic risk, most notably by hedges against real interest-rate risk, default risk, term-structure risk, and consumption-growth risk.

A Econometric issues

This appendix discusses the estimators of several key parameters in this chapter and their limiting distributions. They are w, the zero-exposure portfolio weights; δ, the hedging portfolio weights; λ, the implied hedging cost; and α, the speculative demand for risky assets.

Let R_f denote the risk-free rate, and let R^e be the K-vector of excess returns. k of the K risky assets are basis assets, $k \leq K$. Let R_b^e denote the excess returns on these basis assets. In Section 3.4.3 we take the excess return on the market portfolio as the basis asset.

1. Zero-exposure portfolio weights

Using standard GMM notation, the pricing errors are

$$e(\theta) = cu'\left(R_f + w^\top R_b^e\right) \begin{bmatrix} R_f \\ R_b^e \end{bmatrix} - \begin{bmatrix} 1 \\ 0_k \end{bmatrix},$$

where $\theta = (c, w^\top)^\top$. The moment conditions read $0_{k+1} = E[e(\theta)] \equiv g(\theta)$. The GMM-estimator is then given by $0_{k+1} = E_T[e(\hat{\theta})] \equiv g_T(\hat{\theta})$, where E_T denotes the sample average. Note that

$$\sqrt{T}(\hat{\theta} - \theta) \simeq -\left(\frac{\partial g_T(\theta)}{\partial \theta^\top}\right)^{-1} \sqrt{T} E_T[e(\theta)].$$

Hence, under standard regularity conditions, the limiting distribution of $\hat{\theta}$ is given by

$$\sqrt{T}(\hat{\theta} - \theta) \longrightarrow N\left(0, \left(\frac{\partial g(\theta)}{\partial \theta^\top}\right)^{-1} A \left(\frac{\partial g^\top(\theta)}{\partial \theta}\right)^{-1}\right),$$

where $A = var[e(\theta)]$.

2. Hedging portfolio weights

Let $\Omega(\theta) = -u''(R_f + w^\top R_b^e)$, and $\varepsilon(\theta) = y - \delta(\theta)^\top R^e$, where

$$\delta(\theta) = E\left[R^e \Omega(\theta) R^{e\top}\right]^{-1} E\left[R^e \Omega(\theta) y\right].$$

Estimate $\Omega\left(\theta\right)$ by the plug-in estimator $\Omega(\hat{\theta}) \equiv \hat{\Omega}$, and $\delta = \delta(\theta)$ by

$$\hat{\delta} = \hat{\delta}(\hat{\theta}) = E_T[R^e \hat{\Omega} R^{e\top}]^{-1} E_T[R^e \hat{\Omega} y].$$

Note that

$$\begin{aligned}\hat{\Omega} &\simeq \Omega\left(\theta\right) + \frac{\partial \Omega\left(\theta\right)}{\partial \theta^\top}(\hat{\theta} - \theta) \\ &\simeq \Omega\left(\theta\right) - \frac{\partial \Omega\left(\theta\right)}{\partial \theta^\top}\left(\frac{\partial g_T\left(\theta\right)}{\partial \theta^\top}\right)^{-1} E_T\left[e\left(\theta\right)\right].\end{aligned}$$

Hence,

$$\sqrt{T}(\hat{\delta} - \delta) \simeq E_T[R^e \Omega(\theta) R^{e\top}]^{-1} \sqrt{T} E_T[\zeta(\theta)],$$

where

$$\zeta(\theta) = R^e \Omega(\theta) \varepsilon(\theta) - E\left[R^e \frac{\partial \Omega(\theta)}{\partial \theta^\top} \varepsilon(\theta)\right]^{-1} \left(\frac{\partial g\left(\theta\right)}{\partial \theta^\top}\right)^{-1} e(\theta),$$

and so the limiting distribution of $\hat{\delta}$ is

$$\sqrt{T}(\hat{\delta} - \delta) \longrightarrow N\left(0, E[R^e \Omega(\theta) R^{e\top}]^{-1} B E[R^e \Omega(\theta) R^{e\top}]^{-1}\right),$$

where $B = var[\zeta(\theta)]$.

3. Implied hedging cost

The implied hedging cost, λ, is defined as

$$\lambda = \lambda\left(\theta\right) = E[R^{e\top} \delta\left(\theta\right)] = E[R^e]^\top \delta.$$

It can be estimated by

$$\hat{\lambda} = \hat{\lambda}(\hat{\theta}) = E_T[R^e]^\top \hat{\delta}.$$

Notice

$$\begin{aligned}
\sqrt{T}(\hat{\lambda}-\lambda) &= \sqrt{T}\left(E_T[R^e]^\top\hat{\delta} - E[R^e]^\top\delta\right)\\
&= \sqrt{T}\left(\hat{\delta}^\top(E_T[R^e]-E[R^e]) + E[R^e]^\top(\hat{\delta}-\delta)\right)\\
&= \begin{bmatrix} \hat{\delta} \\ E\left[R^e\right] \end{bmatrix}^\top \sqrt{T}\begin{bmatrix} E_T[R^e]-E[R^e] \\ \hat{\delta}-\delta \end{bmatrix}\\
&\simeq \begin{bmatrix} \hat{\delta} \\ E\left[R^e\right] \end{bmatrix}^\top \sqrt{T}\begin{bmatrix} E_T[R^e]-E[R^e] \\ E_T[R^e\Omega\left(\theta\right)R^{e\top}]^{-1}E_T\left[\zeta\left(\theta\right)\right] \end{bmatrix}\\
&= \begin{bmatrix} \hat{\delta} \\ E\left[R^e\right] \end{bmatrix}^\top \begin{bmatrix} I_K & 0 \\ 0 & E_T[R^e\Omega\left(\theta\right)R^{e\top}]^{-1} \end{bmatrix}\times\\
&\quad \sqrt{T}\begin{bmatrix} E_T[R^e-E[R^e]] \\ E_T[\zeta\left(\theta\right)] \end{bmatrix}.
\end{aligned}$$

Hence, the limiting distribution of $\hat{\lambda}$ is given by

$$\sqrt{T}(\hat{\lambda}-\lambda) \longrightarrow N(0, a^\top HCHa),$$

where

$$\begin{aligned}
a &= \begin{bmatrix} \delta \\ E\left[R^e\right] \end{bmatrix}\\
H &= \begin{bmatrix} I_K & 0 \\ 0 & E[R^e\Omega\left(\theta\right)R^{e\top}]^{-1} \end{bmatrix}\\
C &= var\begin{bmatrix} R^e \\ \zeta\left(\theta\right) \end{bmatrix}.
\end{aligned}$$

4. Speculative demand

Speculative demand is defined as

$$\alpha = \alpha\left(\theta\right) = E[R^e\Omega\left(\theta\right)R^{e\top}]^{-1}E[u'(R_f + w^\top R_b^e)R^e],$$

and it is estimated by

$$\hat{\alpha} = \hat{\alpha}(\hat{\theta}) = E_T[R^e\hat{\Omega}R^{e\top}]^{-1}E_T[u'(R_f + \hat{w}^\top R_b^e)R^e].$$

Using a first-order Taylor approximation, we obtain

$$u'(R_f + \hat{w}^\top R_b^e) \simeq u'(R_f + w^\top R_b^e) + u''(R_f + w^\top R_b^e) R_b^{e\top}(\hat{w} - w).$$

Let $\phi = \phi(\theta) = E[u'(R_f + w^\top R_b^e)R^e]$, and $\hat{\phi} = \hat{\phi}(\hat{\theta}) = E_T[u'(R_f + \hat{w}^\top R_b^e)R^e]$. Then

$$\begin{aligned}
\sqrt{T}(\hat{\phi} - \phi) &= \sqrt{T}\left(E_T[u'(R_f + \hat{w}^\top R_b^e)R^e] - E[u'(R_f + w^\top R_b^e)R^e]\right) \\
&\simeq \sqrt{T}\left(E_T[u'(R_f + w^\top R_b^e)R^e] - E[u'(R_f + w^\top R_b^e)R^e]\right) \\
&\quad + E_T[R^e u''(R_f + w^\top R_b^e)R_b^{e\top}]\sqrt{T}(\hat{w} - w) \\
&= \begin{bmatrix} I_K & 0 & -E_T[R^e \Omega R_b^{e\top}] \end{bmatrix} \sqrt{T} \begin{bmatrix} E_T[\eta(\theta)] \\ \hat{\theta} - \theta \end{bmatrix},
\end{aligned}$$

where $\eta = \eta(\theta) = u'(R_f + w^\top R_b^e)R^e - E[u'(R_f + w^\top R_b^e)R^e]$. Hence,

$$\begin{aligned}
\sqrt{T}(\hat{\alpha} - \alpha) &\simeq E[R^e \Omega(\theta) R^{e\top}]^{-1} \begin{bmatrix} I_K & 0 & -E[R^e \Omega(\theta) R_b^{e\top}] \end{bmatrix} \times \\
&\quad \begin{bmatrix} I_K & 0 \\ 0 & -\left(\frac{\partial g_T(\theta)}{\partial \theta^\top}\right)^{-1} \end{bmatrix} \sqrt{T} \begin{bmatrix} E_T[\eta(\theta)] \\ E_T[e(\theta)] \end{bmatrix} \\
&\longrightarrow N\left(0, PQRDR^\top Q^\top P\right)
\end{aligned}$$

where

$$\begin{aligned}
P &= E[R^e \Omega(\theta) R^{e\top}]^{-1} \\
Q &= \begin{bmatrix} I_K & 0 & -E[R^e \Omega(\theta) R_b^{e\top}] \end{bmatrix} \\
R &= \begin{bmatrix} I_K & 0 \\ 0 & -\left(\frac{\partial g(\theta)}{\partial \theta^\top}\right)^{-1} \end{bmatrix} \\
D &= var \begin{bmatrix} \eta(\theta) \\ e(\theta) \end{bmatrix}.
\end{aligned}$$

B Tables

Table 3.1: **Summary Statistics for Security Returns**

The sample period is August 1960 through December 2001. Mean returns and standard deviations are in percentage points per month. *RM–RF* is the return on the market portfolio in excess of the risk-free rate, *SMB* is the return on the size portfolio, *HML* is the return on the book-to-market value portfolio, *UMD* is return on the momentum portfolio, *TERM* is the return on a long-term government bond in excess of the risk-free rate, *DEF* is the return on a long-term corporate bond less the return on a long-term government bond, and *RF* is the risk-free rate. Corr_t is the autocorrelation at lag t.

Panel A: Means, Standard Deviations, and Autocorrelations

Variable	Mean	Std. Dev.	Corr_1	Corr_2	Corr_3	Corr_6	Corr_{12}
RM–RF	0.491	4.453	0.061	-0.054	-0.003	-0.032	0.010
SMB	0.160	3.217	0.075	0.033	-0.105	0.080	0.136
HML	0.446	2.934	0.125	0.067	0.043	0.063	0.028
UMD	0.879	3.874	-0.025	-0.056	-0.030	0.088	0.114
TERM	0.125	2.765	0.060	0.001	-0.107	0.041	-0.017
DEF	0.014	1.192	-0.170	-0.078	-0.020	-0.034	-0.024
RF	0.485	0.216	0.945	0.909	0.885	0.825	0.714

Panel B: Correlations Matrix

Variable	*RM–RF*	*SMB*	*HML*	*UMD*	*TERM*	*DEF*	*RF*
RM–RF	1	0.301	-0.422	-0.026	0.278	0.082	-0.106
SMB		1	-0.298	0.005	-0.091	0.151	-0.046
HML			1	-0.161	0.002	0.024	0.043
UMD				1	0.041	-0.191	-0.012
TERM					1	-0.473	0.022
DEF						1	-0.059
RF							1

Table 3.2: **Summary Statistics for Economic Risk Variables**

The sample period is August 1960 through December 2001. Means and standard deviations are in percentage points per month. *INF* is the monthly net rate of inflation, *RI* is the monthly real net risk-free rate, *TS* is the yield spread between long- and short-term government bonds, *DS* is the yield spread between Baa and Aaa corporate bonds, *DIV* is dividend yield on the S&P 500 composite, and *CG* is monthly real per-capita consumption growth of durables, nondurables, and services. Corr_t is the autocorrelation at lag t.

Panel A: Means, Standard Deviations, and Autocorrelations

Variable	Mean	Std. Dev.	Corr_1	Corr_2	Corr_3	Corr_6	Corr_{12}
INF	0.361	0.318	0.642	0.548	0.510	0.469	0.521
RI	0.124	0.274	0.481	0.362	0.351	0.310	0.436
TS	0.062	0.086	0.959	0.894	0.836	0.704	0.501
DS	0.082	0.037	0.971	0.932	0.903	0.826	0.679
DIV	0.288	0.091	0.997	0.994	0.990	0.968	0.905
CG	0.207	0.571	-0.204	0.012	0.019	0.051	-0.005

Panel B: Correlations Matrix

Variable	*INF*	*RI*	*TS*	*DS*	*DIV*	*CG*
INF	1	-0.743	-0.422	0.234	0.414	-0.225
RI		1	0.118	0.206	-0.014	0.157
TS			1	0.100	-0.091	0.107
DS				1	0.660	-0.034
DIV					1	-0.030
CG						1

Table 3.3: **First-Order VAR of the Economic Risk Variables**

Regression coefficients of a first-order vector autoregression of the economic risk variables. The sample period is August 1960 through December 2001. Standard errors are in parentheses. Acronyms are defined in Table 3.2. Standard deviations of the innovations are in percentage points per month.

Panel A: Regression Coefficients

Depend.	Regressors						
Variable	*INF*	*RI*	*TS*	*DS*	*DIV*	*CG*	R^2
INF	0.723	0.253	-0.521	-0.707	0.542	0.049	0.777
	(0.086)	(0.086)	(0.165)	(0.465)	(0.165)	(0.019)	
RI	0.146	0.588	0.323	0.839	-0.411	-0.047	0.390
	(0.089)	(0.089)	(0.170)	(0.480)	(0.170)	(0.019)	
TS	0.018	0.024	0.973	0.118	-0.058	-0.001	0.950
	(0.009)	(0.009)	(0.017)	(0.048)	(0.017)	(0.002)	
DS	0.011	0.011	-0.003	0.927	0.002	-0.002	0.991
	(0.003)	(0.003)	(0.006)	(0.017)	(0.006)	(0.001)	
DIV	0.004	0.002	-0.003	-0.020	1.000	-0.001	1.000
	(0.002)	(0.002)	(0.005)	(0.013)	(0.005)	(0.001)	
CG	-0.675	-0.446	0.207	0.515	0.608	-0.255	0.198
	(0.207)	(0.205)	(0.394)	(1.114)	(0.394)	(0.044)	

Panel B: Standard Deviations and Correlation Matrix of VAR Innovations

Variable	Std. Dev.	*INF*	*RI*	*TS*	*DS*	*DIV*	*CG*
INF	0.228	1	-0.957	-0.056	-0.077	0.124	-0.157
RI	0.235		1	-0.058	0.081	-0.126	0.153
TS	0.024			1	0.182	0.063	-0.063
DS	0.008				1	0.174	-0.038
DIV	0.007					1	-0.023
CG	0.545						1

Table 3.4: **Implied Hedging Costs**

Implied hedging costs associated with innovations in risk variables, measured in units of risk (i.e. standard deviation). The sample period is August 1960 through December 2001. Standard errors are in parentheses. Acronyms are defined in Tables 3.1 and 3.2.

Panel A: First-Order VAR Innovations

Risk aver.	*INF*	*RI*	*TS*	*DS*	*DIV*	*CG*
$\gamma = 1$	-0.0669**	0.0321	0.0856	0.0037	-0.0356**	0.0409
	(0.0286)	(0.0261)	(0.0789)	(0.0343)	(0.0141)	(0.0260)
$\gamma = 2$	-0.0587**	0.0286	0.0622	0.0083	-0.0354***	0.0354
	(0.0251)	(0.0235)	(0.0566)	(0.0266)	(0.0134)	(0.0225)
$\gamma = 5$	-0.0487**	0.0227	0.0507	0.0075	-0.0348***	0.0285
	(0.0243)	(0.0233)	(0.0484)	(0.0255)	(0.0135)	(0.0222)
$\gamma = 20$	-0.0425*	0.0193	0.0413	0.0072	-0.0341**	0.0237
	(0.0246)	(0.0235)	(0.0427)	(0.0242)	(0.0137)	(0.0223)

Panel B: Orthogonalized First-Order VAR Innovations

Risk aver.	*INF*	*RI*	*TS*	*DS*	*DIV*	*CG*
$\gamma = 1$	-0.0669**	-0.1104**	0.0426	-0.0075	-0.0323*	0.0351
	(0.0286)	(0.0461)	(0.0694)	(0.0468)	(0.0171)	(0.0272)
$\gamma = 2$	-0.0587**	-0.0956***	0.0239	0.0014	-0.0330**	0.0297
	(0.0251)	(0.0347)	(0.0504)	(0.0342)	(0.0153)	(0.0233)
$\gamma = 5$	-0.0487**	-0.0828***	0.0174	0.0023	-0.0331**	0.0236
	(0.0243)	(0.0299)	(0.0433)	(0.0312)	(0.0151)	(0.0224)
$\gamma = 20$	-0.0425*	-0.0738***	0.0114	0.0036	-0.0328**	0.0192
	(0.0246)	(0.0267)	(0.0387)	(0.0287)	(0.0152)	(0.0221)

***/**/* indicates significance at the 1/5/10% level.

Table 3.5: **Economic Hedging Portfolios**

Hedging portfolio weights (in percentage points) due to a unit exposure to innovations in economic risk variables. The sample period is August 1960 through December 2001. Standard errors are in parentheses. Acronyms are defined in Tables 3.1 and 3.2.

Risk aversion	Security	*INF*		*RI*		*TS*		*DS*		*DIV*		*CG*	
$\gamma = 1$	*RM–RF*	-0.56	(0.42)	-0.13	(0.40)	0.16	(0.12)	-0.01	(0.02)	-0.03***	(0.01)	1.64*	(0.86)
	SMB	-0.18	(0.45)	-0.11	(0.45)	0.24*	(0.14)	-0.04	(0.03)	0.00	(0.01)	4.52***	(1.35)
	HML	-0.94	(0.66)	0.35	(0.63)	0.27*	(0.15)	-0.03	(0.03)	-0.03**	(0.01)	2.39	(1.70)
	UMD	-0.94***	(0.30)	0.77**	(0.35)	-0.10*	(0.06)	0.02*	(0.01)	0.00	(0.00)	-0.25	(0.48)
	TERM	-0.22	(0.82)	0.37	(0.91)	0.42***	(0.15)	0.05	(0.03)	0.04**	(0.02)	-0.78	(1.32)
	DEF	4.76***	(1.61)	-3.61*	(1.97)	0.07	(0.47)	0.17**	(0.08)	0.04*	(0.02)	-4.14	(3.22)
$\gamma = 2$	*RM–RF*	-0.72**	(0.35)	0.16	(0.33)	0.10	(0.08)	0.00	(0.02)	-0.03***	(0.01)	1.30*	(0.74)
	SMB	-0.19	(0.45)	-0.03	(0.42)	0.14*	(0.09)	-0.03*	(0.02)	0.00	(0.01)	4.21***	(1.11)
	HML	-0.90	(0.57)	0.46	(0.54)	0.17*	(0.10)	-0.01	(0.02)	-0.03**	(0.01)	2.03	(1.34)
	UMD	-0.65**	(0.31)	0.42	(0.35)	-0.05	(0.04)	0.01	(0.01)	0.00	(0.00)	-0.16	(0.42)
	TERM	-0.24	(0.59)	0.45	(0.64)	0.30***	(0.09)	0.08***	(0.03)	0.03**	(0.01)	-1.05	(1.17)
	DEF	3.52***	(1.16)	-2.95**	(1.37)	0.27	(0.26)	0.14***	(0.05)	0.05**	(0.02)	-0.49	(2.53)
$\gamma = 5$	*RM–RF*	-0.77**	(0.34)	0.28	(0.34)	0.08	(0.06)	0.00	(0.01)	-0.02***	(0.01)	1.12	(0.74)
	SMB	-0.22	(0.43)	0.04	(0.41)	0.10	(0.07)	-0.02*	(0.01)	0.00	(0.01)	3.91***	(1.02)
	HML	-0.81	(0.54)	0.45	(0.53)	0.13	(0.08)	-0.01	(0.02)	-0.03**	(0.01)	1.81	(1.25)
	UMD	-0.38	(0.36)	0.18	(0.39)	-0.03	(0.05)	0.00	(0.01)	0.00	(0.01)	-0.33	(0.52)
	TERM	-0.28	(0.56)	0.50	(0.58)	0.25***	(0.07)	0.09***	(0.03)	0.03**	(0.01)	-1.16	(1.15)
	DEF	2.77***	(1.05)	-2.41**	(1.20)	0.30	(0.21)	0.14***	(0.05)	0.06**	(0.02)	1.14	(2.47)
$\gamma = 20$	*RM–RF*	-0.77**	(0.34)	0.32	(0.34)	0.07	(0.06)	0.00	(0.01)	-0.02***	(0.01)	0.98	(0.74)
	SMB	-0.22	(0.43)	0.06	(0.41)	0.08	(0.06)	-0.02	(0.01)	0.00	(0.01)	3.75***	(0.98)
	HML	-0.73	(0.53)	0.41	(0.52)	0.11	(0.07)	0.00	(0.02)	-0.03**	(0.01)	1.62	(1.22)
	UMD	-0.25	(0.39)	0.07	(0.41)	-0.03	(0.05)	0.00	(0.01)	0.00	(0.01)	-0.44	(0.57)
	TERM	-0.30	(0.55)	0.51	(0.57)	0.23***	(0.07)	0.09***	(0.03)	0.03**	(0.01)	-1.10	(1.15)
	DEF	2.40**	(1.05)	-2.13*	(1.17)	0.31	(0.19)	0.13***	(0.05)	0.06**	(0.02)	1.90	(2.47)

***/**/* indicates significance at the 1/5/10% level.

Table 3.6: **CAPM Economic Hedging Portfolios**

CAPM hedging portfolio weights (in percentage points) due to a unit exposure to innovations in economic risk variables. The sample period is August 1960 through December 2001. Standard errors are in parentheses. Acronyms are defined in Tables 3.1 and 3.2.

Risk aversion	Security	*INF*		*RI*		*TS*		*DS*		*DIV*		*CG*	
$\gamma = 1$	*RM–RF*	-0.73**	(0.30)	0.36	(0.31)	0.05	(0.04)	0.00	(0.01)	-0.03***	(0.01)	0.96	(0.65)
	SMB	-0.11	(0.37)	-0.11	(0.36)	0.05	(0.04)	-0.01	(0.01)	0.00	(0.01)	2.86***	(0.82)
	HML	-0.54	(0.40)	0.23	(0.40)	0.09*	(0.05)	-0.01	(0.02)	-0.02**	(0.01)	1.46*	(0.86)
	UMD	0.38	(0.32)	-0.48	(0.33)	-0.04	(0.04)	-0.01	(0.01)	0.00	(0.01)	-0.92	(0.63)
	TERM	-0.69	(0.49)	1.02*	(0.53)	0.17**	(0.07)	0.10***	(0.03)	0.03*	(0.02)	-1.00	(0.97)
	DEF	0.86	(1.04)	-0.55	(1.12)	0.20	(0.15)	0.15***	(0.05)	0.07**	(0.03)	5.67**	(2.55)
$\gamma = 2$	*RM–RF*	-0.75**	(0.30)	0.40	(0.31)	0.05	(0.04)	-0.01	(0.01)	-0.02***	(0.01)	0.90	(0.68)
	SMB	-0.12	(0.38)	-0.09	(0.36)	0.05	(0.04)	-0.01	(0.01)	0.00	(0.01)	2.83***	(0.80)
	HML	-0.52	(0.40)	0.23	(0.41)	0.09*	(0.05)	-0.01	(0.02)	-0.02**	(0.01)	1.35	(0.86)
	UMD	0.35	(0.32)	-0.41	(0.33)	-0.04	(0.04)	-0.01	(0.01)	0.00	(0.01)	-0.90	(0.62)
	TERM	-0.64	(0.49)	0.93*	(0.53)	0.17**	(0.07)	0.11***	(0.03)	0.03*	(0.02)	-1.02	(1.01)
	DEF	0.87	(1.05)	-0.61	(1.13)	0.22	(0.15)	0.16***	(0.05)	0.06**	(0.03)	5.63**	(2.47)
$\gamma = 5$	*RM–RF*	-0.74**	(0.30)	0.41	(0.31)	0.05	(0.04)	-0.01	(0.01)	-0.02**	(0.01)	0.85	(0.69)
	SMB	-0.12	(0.38)	-0.08	(0.37)	0.05	(0.04)	-0.01	(0.01)	0.00	(0.01)	2.82***	(0.80)
	HML	-0.49	(0.40)	0.22	(0.41)	0.09*	(0.05)	-0.01	(0.02)	-0.02**	(0.01)	1.29	(0.86)
	UMD	0.33	(0.32)	-0.38	(0.33)	-0.04	(0.04)	-0.01	(0.01)	0.00	(0.01)	-0.89	(0.62)
	TERM	-0.62	(0.49)	0.89*	(0.53)	0.16**	(0.07)	0.11***	(0.03)	0.02	(0.02)	-1.02	(1.02)
	DEF	0.87	(1.06)	-0.64	(1.14)	0.23	(0.14)	0.16***	(0.04)	0.06**	(0.03)	5.56**	(2.43)
$\gamma = 20$	*RM–RF*	-0.74**	(0.30)	0.42	(0.31)	0.05	(0.04)	-0.01	(0.01)	-0.02**	(0.01)	0.80	(0.70)
	SMB	-0.11	(0.38)	-0.09	(0.37)	0.04	(0.04)	-0.01	(0.01)	0.00	(0.01)	2.81***	(0.80)
	HML	-0.47	(0.40)	0.21	(0.41)	0.09*	(0.04)	-0.01	(0.02)	-0.02**	(0.01)	1.20	(0.86)
	UMD	0.32	(0.32)	-0.37	(0.33)	-0.05	(0.03)	-0.01	(0.01)	0.00	(0.01)	-0.91	(0.62)
	TERM	-0.60	(0.49)	0.86	(0.52)	0.16**	(0.07)	0.11***	(0.03)	0.02	(0.01)	-0.94	(1.03)
	DEF	0.84	(1.05)	-0.64	(1.13)	0.24*	(0.14)	0.15***	(0.04)	0.06**	(0.03)	5.56**	(2.42)

***/**/* indicates significance at the 1/5/10% level.

Chapter 4

Multivariate Option Pricing Using Dynamic Copula Models

4.1 Introduction

In today's economy, multivariate (or rainbow) options are viewed as excellent tools for hedging the risk of multiple assets. These options, which are written on two or more underlying securities or indexes, usually take the form of calls (or puts) that give the right to buy (or sell) the best or worst performer of a number of underlying assets. Other examples include forward contracts whose payoff is equal to that of the best or worst performer of its underlyings, and spread options on the difference between the prices of two assets.

One of the key determinants in the valuation of multivariate options is the dependence between the underlying assets. Consider for instance a bivariate call-on-max option, namely a contract that gives the holder the right to purchase the more valuable of two underlying assets for a pre-specified strike price. Intuitively, the value of such an option should be smaller if the underlyings tend to move together than when they move in opposite directions. More generally, the dependence between the underlyings could change over time. Accounting for time variation in the dependence structure between assets should prove helpful in providing a more realistic valuation of multivariate options.

Over the years, various generalizations of the Black–Scholes (1973) Brownian motion framework have been used to model multivariate option prices. Examples include Margrabe (1978), Stulz (1982), Johnson (1987), Reiner (1992), and Shimko (1994). In these papers, the dependence between assets is modelled by their correlation. However, unless asset returns are well represented by a multivariate normal distribution, correlation is often an unsatisfactory measure of dependence; see, for instance, Embrechts, McNeil and Straumann (2002). Furthermore, it is a stylized fact of financial markets that correlations observed under ordinary market conditions differ substantially from correlations observed in hectic periods. In particular, asset prices have a greater tendency to move together in bad states of the economy than in quiet periods; see, for instance, Boyer et al. (1999) and Patton (2002a, 2002b) and references therein. These "correlation breakdowns," associated with economic downturns, suggest a dynamic model of the dependence structure of asset returns.

In this chapter, the relation between multivariate option prices and the dependence structure of the underlying financial assets is modelled dynamically through copulas. A copula is a multivariate distribution function each of whose marginals is uniform on the unit interval. It has been known since the work of Sklar (1959) that any multivariate continuous distribution function can be uniquely factored into its marginals and a copula. Thus while correlation measures dependence through a single number, the dependence between multiple assets is fully captured by the copula. From a practical point of view, the advantage of the copula-based approach to modelling is that appropriate marginal distributions for the components of a multivariate system can be selected by any desired method, and then linked through a copula or family of copulas suitably chosen to represent the dependence prevailing between the components.

The use of copulas to price multivariate options is not new. For example, in Rosenberg (1999), univariate options data are used to estimate marginal risk-neutral densities, which are linked with a Plackett copula to obtain a bivariate risk-neutral density from which bivariate claims are valuated. This semiparametric procedure uses a particular identifying assumption on the risk-neutral correlation to fix the copula parameter. Cherubini and Luciano (2002) extend Rosenberg's work by considering other families of copulas. In Rosenberg (2003), a risk-neutral bivariate distribution is estimated from nonparametric estimates of the marginal distributions and a nonparametric estimate of the copula.

An innovating feature of the present study, however, is that, contrary to earlier works on multivariate option pricing, the dependence structure of the underlying assets is not treated as fixed, but rather as possibly varying over time. Taking into account this time variation is important because it may influence option prices. This chapter proposes a model for the time variation of the dependence structure, in which a parametric copula is specified whose dependence parameter is allowed to change with the volatilities of the underlying assets. A distinct advantage of the parametric approach is that while the model may be misspecified, the robustness of the conclusions can easily be verified by repeating the analysis for as many different copula families as desired.

A similar dynamic-copula approach has already been used in the foreign exchange market literature by Patton (2002a), who found time variation to be significant in a copula model for asymmetric dependence between two exchange rates where the dependence parameter followed a ARMA-type process. While Patton's goal was to study the effect of asymmetric dependence on portfolio returns, the objective of the present work is very different. The main focus here is on the effect of time variation in the underlying dependence structure on the price of multivariate options.

In the empirical study presented herein, multivariate options on two important American equity index returns are considered: the S&P 500 and the Nasdaq. An analysis of the results suggests that allowing for time variation in the dependence structure of the underlyings produces substantially different option prices than under constant dependence, particularly in times of increased volatility. Moreover, option prices implied by a normal dynamic dependence structure differ significantly from option prices implied by non-normal dynamic dependence structures. These findings suggest that unless the dependence between the S&P 500 and Nasdaq stock indexes is well described by a normal copula, alternative copula families should be considered. Option prices turned out to be robust among the alternative—i.e., non-normal—copula models considered in this study.

The remainder of this chapter is organized as follows. Section 5.2 describes the payoff structure of better-of-two-markets and worse-of-two-markets claims, and explains in detail the proposed dynamic-dependence option valuation scheme. The empirical results are presented in Section 4.3, and conclusions are given in Section 4.4.

4.2 Option pricing with time-varying dependence

Multivariate options come in a wide variety of payoff schemes. The most commonly traded options of this kind are basket options on a portfolio of assets, such as index options. Other examples include spread options, some of which are traded on commodity exchanges (see, for example, Rosenberg (1998)), or dual-strike and multivariate-digital options.

This study concentrates on European-type options on the best (worst) performer of several assets, sometimes referred to as outperformance (underperformance) options. As these are typically traded over the counter, price data are not available. Therefore, valuation models cannot be tested empirically. However, a robustness study comparing models with different assumptions remains feasible, and this is the objective pursued herein. While the study described in the sequel is restricted to options on better- and worse-of-two-markets claims, the technique is sufficiently general to analyze the aforementioned alternative multivariate options as well, and may thus be of wider interest.

One can distinguish four types of better-of-two-markets or worse-of-two-markets claims: call options on the better performer, put options on the worse performer, call options on the worse performer, and put options on the better performer. These may be referred to as call-on-max, put-on-min, call-on-min, and put-on-max options, respectively. Their payoffs at maturity are:

$$\begin{aligned}
\text{call on max} &: \quad \max\{\max(R_1, R_2) - E, 0\}, \\
\text{put on min} &: \quad \max\{E - \min(R_1, R_2), 0\}, \\
\text{call on min} &: \quad \max\{\min(R_1, R_2) - E, 0\}, \\
\text{put on max} &: \quad \max\{E - \max(R_1, R_2), 0\},
\end{aligned}$$

where R_i is the return at maturity on index $i \in \{1, 2\}$, and E denotes the exercise price of the option.

The proposed scheme for valuating these options is as follows. First, each of the two objective marginal distributions of the daily index returns underlying a type of claim is modelled, and their risk-neutral counterparts are derived. Next, a parametric family of copulas is chosen to fix the joint risk-neutral distribution of the index returns. The fair value of the option is then determined by taking the discounted expected value of the option's payoff under the risk-neutral distribution.

The specification chosen for the objective marginal distributions is from Duan (1995). It is general enough to capture volatility clustering, a stylized fact of equity returns for which there is overwhelming empirical evidence at the daily frequency, while still providing a relatively easy transformation to risk-neutral distributions. Each of the objective marginal distributions of the index returns is modelled by a GARCH(1,1) process. It is repeated here for the sake of completeness; see Bollerslev (1986). For $i \in \{1, 2\}$,

$$\begin{aligned} R_{i,t+1} &= \mu_i + \eta_{i,t+1}\,, \\ \eta_{i,t+1}|\mathcal{I}_t &\sim \mathcal{N}(0, h_{i,t})\,, \\ h_{i,t+1} &= \omega_i + \beta_i h_{i,t} + \alpha_i \eta_{i,t+1}^2, \end{aligned}$$

where $\omega_i > 0$, $\beta_i > 0$, and $\alpha_i > 0$. Here, $\mathcal{I}_t$ denotes the information available at time t. However, it must be stressed that, in the light of Sklar's theorem, in principle *any* choice for the marginal distributions is consistent with the copula approach. The vast collection of alternatives that have been used by other authors to model univariate index return distributions includes (variants of) continuous-time geometric Brownian motion of Black and Scholes (1973), and the discrete-time binomial model of Cox, Ross and Rubinstein (1979). Again, the GARCH specification that is employed here is appealing as it allows for an easy change of measure in addition to being able to capture volatility clustering. In particular, Duan (1995) shows that, under certain conditions, the change of measure comes down to a change in the drift.

An alternative, nonparametric approach is to use univariate option price data to obtain arbitrage-free estimates of the marginal risk-neutral densities, as in Ait-Sahalia and Lo (1998). This route is taken by Rosenberg (2003). Clearly, an advantage of this approach is that it does not impose restrictions on the asset return processes or on the functional form of the risk-neutral densities. However, this flexibility comes at the cost of imprecise estimates, especially if the distributions are time-varying.

The second step in the proposed valuation scheme is to fix the joint risk-neutral distribution of the index returns by choosing a copula. A set of well-known one-parameter copula families is considered for this purpose. They are the Frank, Gumbel–Hougaard, Plackett, Galambos, and normal families. Their cumulative distribution functions are given in Appendix A. For all of these copulas, there is a one-to-one relation between the dependence parameter—denoted θ—and Kendall's nonparametric measure of association. For any copula C_θ, Kendall's tau is related to θ in the following way:

$$\tau(\theta) = 4\mathrm{E}C_\theta(U, V) - 1, \tag{4.1}$$

where (U, V) is distributed as C_θ, and E denotes the expectation operator with respect to U and V. Appendix B displays closed-form formulas for the population value of Kendall's tau for some of the copula models under consideration.

This relation suggests a natural way to estimate the copula. An estimate of θ is readily obtained by computing the sample version of tau on a (sub)sample of paired index-return observations, inverting Relation (4.1), and plugging in the sample tau.[1] This method-of-moment type procedure yields a rank-based estimate of the association parameter which is consistent, under the assumption that the selected family of copulas describes accurately the dependence structure of the equity indexes. Other methods could be used without fundamentally altering this approach, e.g., inversion of Spearman's rho, or the maximum pseudo-likelihood method.

The proposed technique assumes that the objective and risk-neutral copulas are identical. Rosenberg (2003) makes this assumption as well. If multivariate option price data were available, this assumption could be tested or the appropriate risk-neutral copula could be estimated. Only data on prices of multivariate claims would reveal information about the risk-neutral dependence structure. Information about the risk-neutral dependence structure can never be extracted from univariate op-

[1]The sample version of Kendall's tau is defined as follows. Let $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ be a random sample of n observations from a vector (X, Y) of continuous random variables. Two distinct pairs (X_i, Y_i) and (X_j, Y_j) are said to be concordant if $(X_i - X_j)(Y_i - Y_j) > 0$, and discordant if $(X_i - X_j)(Y_i - Y_j) < 0$. Kendall's tau for the sample is then defined as $t = (c - d)/(c + d)$, where c denotes the number of concordant pairs, and d is the number of discordant pairs.

tion prices—which *are* available—as these only bear relevance to the risk-neutral marginal processes, and not to the joint risk-neutral process. Identification of the multivariate density requires knowledge of both the marginal densities and the dependence function that links them together.

Time variation in the copula is modelled by allowing the parameter of dependence parameter to evolve through time according to a particular equation. The forcing variables in this equation are the conditional volatilities of the underlying assets. These are also the forcing variables that are typically chosen to model time-varying correlations; see, e.g., the BEKK model introduced by Engle and Kroner (1995). Additional motivation is provided by the evidence on correlation breakdowns, which suggests that financial markets exhibit high dependence in periods of high volatility. Patton (2002a) proposes an ARMA-type process linking the dependence parameter to absolute differences in return innovations, which is another way to capture the same idea.

To be more specific, let τ_t be Kendall's measure of association at time t, and let $h_{i,t}$ be the objective conditional variance estimate at time t of underlying index return $i \in \{1, 2\}$ implied by Duan's GARCH option pricing model. It is assumed that

$$\tau_t = \gamma(h_{1,t}, h_{2,t}) \tag{4.2}$$

for some function $\gamma(\cdot, \cdot)$ to be specified later. This conditional measure of association governs the degree of dependence for the risk-neutral copula under consideration.

The proposed valuation scheme is implemented using Monte Carlo simulations. Pairs of random variates are drawn from the copula implied by the estimated conditional risk-neutral measure of association, which are then transformed to return innovations using Duan's GARCH model. Subsequently, the payoffs implied by these innovations are averaged and discounted at the risk-free rate. The result then constitutes the fair value of the option. Algorithms for random variable generation from the non-normal copulas are given in Genest and MacKay (1986), Genest (1987), Ghoudi, Khoudraji and Rivest (1998), and Nelsen (1999). For the normal copula, a straightforward Cholesky decomposition may be used.

4.3 Pricing options on two equity indexes

The dynamic-dependence valuation scheme outlined in Section 5.2 is applied to better-of-two-markets and worse-of-two-markets options on the S&P 500 and Nasdaq indexes. A sample consisting of pairs of daily returns on the S&P 500 and Nasdaq from January 1, 1993 to August 30, 2002 was obtained from Datastream. The sample size is $T = 2422$. The maximum likelihood estimates of the GARCH parameters for the marginal index return processes may be found in Table 4.1. The values for α and β nearly add up to one. These estimates are in line with previously reported values.

Figure 4.1 depicts the time series of the estimated standardized GARCH innovations $(\eta_{1,t+1}/\sqrt{h_{1,t}}, \eta_{2,t+1}/\sqrt{h_{2,t}})$ for the last 250 trading days in the sample. (For clarity, the picture is restricted to a subsample; other episodes show a similar pattern.) Note that outliers typically occur simultaneously and in the same direction. This positive dependence between the two series is even more apparent from Figure 4.2, which displays the support set of the empirical copula of the standardized return innovations. This scatter plot consists of the observed pairs of ranks (divided by $T + 1$) for the estimated standardized GARCH innovations of the two markets. Under regularity conditions, the empirical copula function converges to the true (here, objective) copula function; see Van der Vaart and Wellner (1996). Notice the pronounced positive dependence, particularly in the tails. The sample version of Kendall's tau for the entire sample amounts to 0.60, confirming positive dependence. Figure 4.3 gives an impression of how this dependence measure of the standardized return innovations evolves over time. It shows rolling-window estimates of Kendall's tau using window sizes of two months, i.e., Kendall's tau at day t is computed using the 20 trading days prior to day t, day t itself, and the 20 trading days after day t. While the estimates show considerable variation, a slightly upward trend over the sample period is discernable.

The time variation in the copula is governed by Equation (4.2). It models the dependence measure as a function of the conditional volatilities of the index returns. The following specification of this function is proposed:

$$\gamma(h_1, h_2) = \gamma_0 + \gamma_1 \log \max(h_1, h_2). \tag{4.3}$$

To motivate this specification, recall that the evidence on correlation breakdowns suggests that increased dependence occurs in hectic periods. Hence, theory predicts a positive value of γ_1. The maximum operator reflects that hectic periods in either market may cause dependence to go up. Since volatility in both markets is highly dependent, the actual specification is likely not to affect the results in the present section too much. The parameters γ_0 and γ_1 were estimated by regressing the rolling-window estimates of Kendall's tau on the estimated log maximum conditional volatility. This is illustrated in Figure 4.4. The slope coefficient, γ_1, was estimated at 0.063; positive, as expected. The estimated dependence measure implied by these parameter estimates,

$$\gamma(h_{1,t}, h_{2,t}) = \gamma_0 + \gamma_1 \log \max(h_{1,t}, h_{2,t}),$$

was then used to fix the conditional risk-neutral copula at time t. Return innovations were sampled from this conditional copula to compute the price of the option. In total, the Monte Carlo study was based on $100,000$ replications, leading to simulation errors in the order of magnitude of 1 basis point for one-month maturity claims.

Clearly, the option price depends on the initial levels of volatility of the underlyings. Prices for three levels of initial volatility were computed: low, medium, and high volatility, where medium volatility is defined as the estimated unconditional variance $\omega/(1-\beta-\alpha)$, and low and high volatility are one-fourth of and four times this amount, respectively. Furthermore, different maturities were considered, ranging from one day to one month (i.e., 20 trading days). The strike price was set at levels between .98 and 1.02. Finally, the risk-free rate was assumed to be 4 percent per annum.

The results show that allowing for time varying dependence leads to different option prices than under static dependence, in particular in times of high volatility. This is illustrated in Figure 4.5 which displays, for various copula parametrizations, the price (measured in basis points) of a one-month put-on-max option as a function of the exercise price implied by dynamic dependence, and compares it to the option price under three levels of static dependence: low, medium, and high static dependence. The medium level of dependence is equal to the average measure of dependence found in the sample, 0.60; the low and high levels are 0.40 and 0.80, respectively. Note that a static model for the dependence structure, which uses the sample measure of dependence of 0.60, underestimates the option price generated by the dynamic model considerably for all copula parametrizations and over the entire range of strike prices considered. The difference is significant since the 95% confidence intervals of the price estimates do not overlap. In the interest of clarity, confidence intervals are not displayed here, but available from the authors upon request. Note that the prices implied by dynamic copulas are between the high and the medium static-dependence prices, suggesting that the dynamic model implies a dependence that is on average stronger than in the medium static-dependence case. Interestingly, price differences between the dynamic and static model vanish as initial volatilities are at a medium level; see Figure 4.6. The same holds for low initial volatilities (not shown), again, across a broad range of copula families and strike prices.

It is also interesting to compare option prices produced by different dynamic copula families. It turns out that prices implied by the normal copula deviate substantially from prices implied by the other copula families. Outside the normal class, the copula choice appears to be irrelevant. This suggests that unless the dependence between index returns can be described by a normal model, alternative specifications should be considered. These findings are illustrated in Figures 4.7 and 4.8 which depict dynamic-dependence one-month call-on-max and put-on-min option prices respectively, as a function of their strike under medium initial volatilities. The prices implied by the normal copula are significantly lower than the prices implied by the other copulas across the whole range of strike prices. The effect is there at other maturities as well. The difference between normal and non-normal prices is also found for high and low initial volatility levels. The differences are less significant for call-on-min and put-on-max options.

4.4 Conclusions

This chapter studies the relation between multivariate options prices and the dependence structure of the underlying assets. A copula-based model was proposed for the valuation of claims on multiple assets. A novel feature of the proposed model is that, contrary to earlier works on multivariate option pricing, the dependence structure is not taken as fixed, but rather as potentially varying with time. The time variation in the dependence structure was modelled using various parametric copulas by letting the copula parameter depend on the conditional volatilities of the underlyings.

This dynamic copula model was applied to better- and worse-of-two-markets options on the S&P 500 and Nasdaq indexes for a variety of copula parametrizations. Option prices implied by the dynamic model turned out to differ substantially from prices implied by a model that fixes the dependence between the underlying indexes, especially in high-volatility market conditions. Hence, the application suggests that time variation in the dependence between the S&P 500 and the Nasdaq is important for the price of options on these indexes. A comparison of option prices computed from different copula families shows that the normal family produces prices that differ significantly from the ones implied by the non-normal alternatives. These findings suggests that if the dependence between the index returns is not well represented by a normal copula, alternative copulas need to be considered. The empirical relevance of such alternatives is apparent given the evidence of non-normality in financial markets.

A One-parameter copula families

The table below displays several one-parameter copula families.

Frank	$C_\theta(u,v) = \frac{1}{\theta}\log\left\{1 + \frac{(e^{\theta u}-1)(e^{\theta v}-1)}{e^{\theta}-1}\right\}$
Gumbel–Hougaard	$C_\theta(u,v) = \exp\left\{-\left(\lvert\log u\rvert^\theta + \lvert\log v\rvert^\theta\right)^{\frac{1}{\theta}}\right\}$
Plackett	$C_\theta(u,v) = \frac{1+(\theta-1)(u+v)-\sqrt{[1+(\theta-1)(u+v)]^2-4uv\theta(\theta-1)}}{2(\theta-1)}$
Galambos	$C_\theta(u,v) = uv\exp\left\{\left(\lvert\log u\rvert^\theta + \lvert\log v\rvert^\theta\right)^{\frac{1}{\theta}}\right\}$
Normal	$C_\theta(u,v) = N_\theta(\Phi^{-1}(u), \Phi^{-1}(v))$

Note: Φ is the standard (univariate) normal distribution function, and N_θ denotes the standard bivariate normal distribution function with correlation coefficient θ.

B Kendall's tau

The table below provides expressions—closed-form if available—of the relation between Kendall's tau and the parameter of the families considered in Appendix A.

Frank	$\tau(\theta) = 1 - 4\left\{D_1(-\theta) - 1\right\}/\theta$
Gumbel-Hougaard	$\tau(\theta) = 1 - 1/\theta$
Plackett	$\tau(\theta) = 4\int_0^1\int_0^1 C_\theta(u,v)dC_\theta(u,v) - 1$
Galambos	$\tau(\theta) = \frac{\theta+1}{\theta}\int_0^1\left(\frac{1}{t^{1/\theta}} + \frac{1}{(1-t)^{1/\theta}} - 1\right)^{-1}dt$
Normal	$\tau(\theta) = \frac{2}{\pi}\arcsin\theta$

Note: D_1 denote the first-order Debye function, $D_1(-\theta) = \frac{1}{\theta}\int_0^\theta \frac{t}{e^t-1}dt + \frac{\theta}{2}$.

C Tables and figures

Table 4.1: Maximum likelihood estimates of the GARCH parameters for the marginal index return processes. Figures in brackets are robust quasi-maximum likelihood standard errors.

Parameter	S&P 500		Nasdaq	
$\mu \times 10^2$	0.0674	(0.0168)	0.0812	(0.0246)
$\omega \times 10^5$	0.0680	(0.0398)	0.1895	(0.0987)
β	0.9258	(0.0220)	0.8906	(0.0309)
α	0.0680	(0.0198)	0.1015	(0.0288)

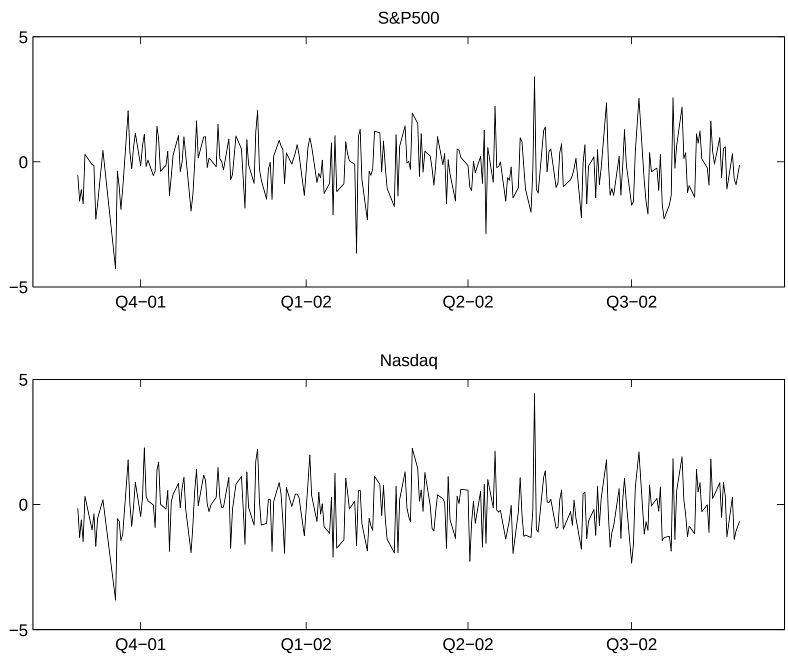

Figure 4.1: Daily standardized GARCH innovations for S&P 500 and Nasdaq for the last 250 trading days in the sample.

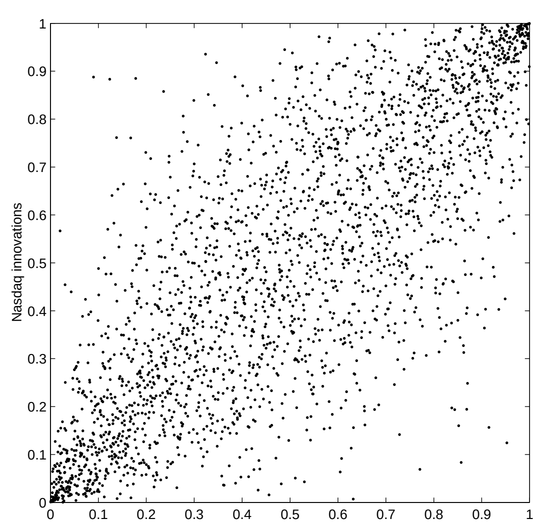

Figure 4.2: Support set of the empirical copula of the standardized GARCH innovations.

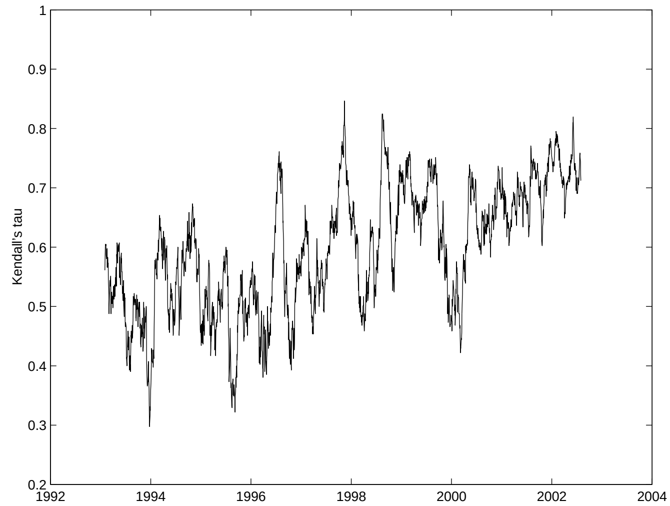

Figure 4.3: Rolling-window estimates of Kendall's tau for the standardized return innovations using a window size of 41 trading days.

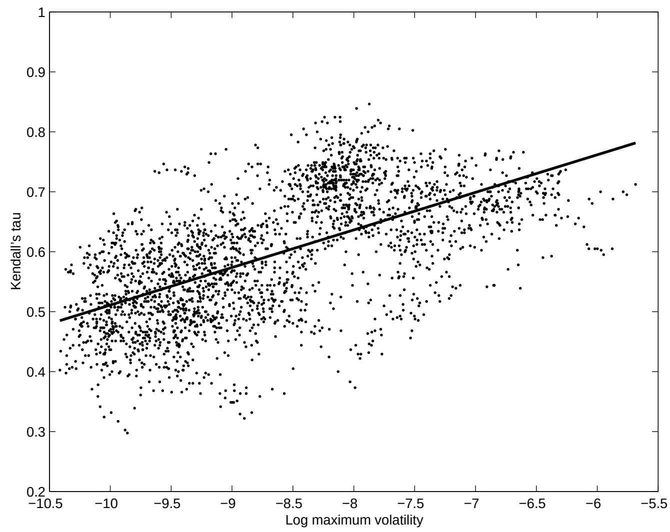

Figure 4.4: Regression of rolling-window estimates of Kendall's tau for the standardized return innovations on the logarithm of the maximum return volatility.

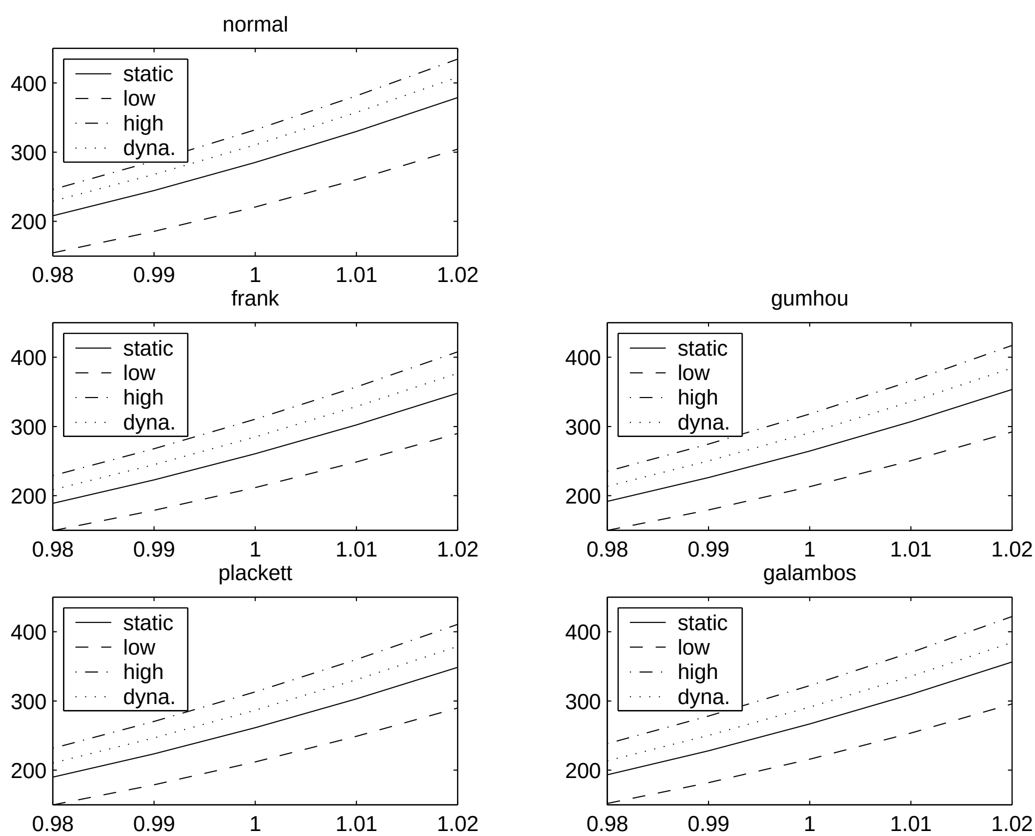

Figure 4.5: One-month maturity put-on-max prices as a function of the strike under high initial volatilities for dynamic dependence and for low, medium, and high static dependence for various copulas.

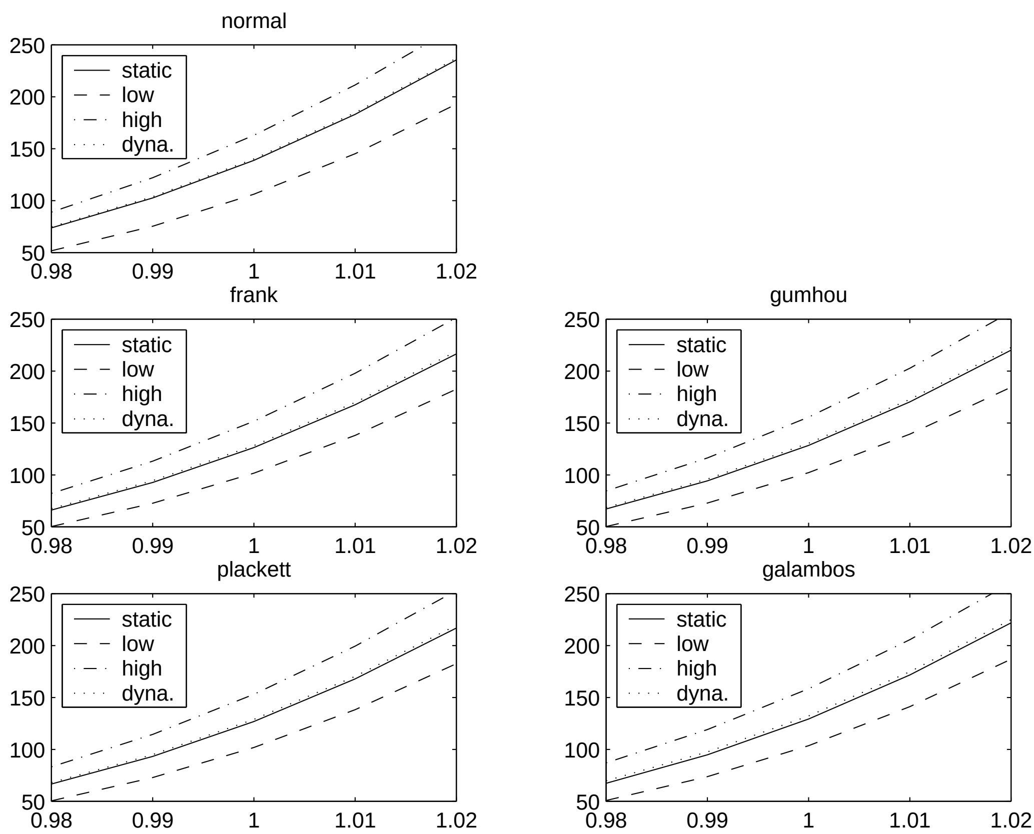

Figure 4.6: One-month maturity put-on-max prices as a function of the strike under medium initial volatilities for dynamic dependence and for low, medium, and high static dependence for various copulas.

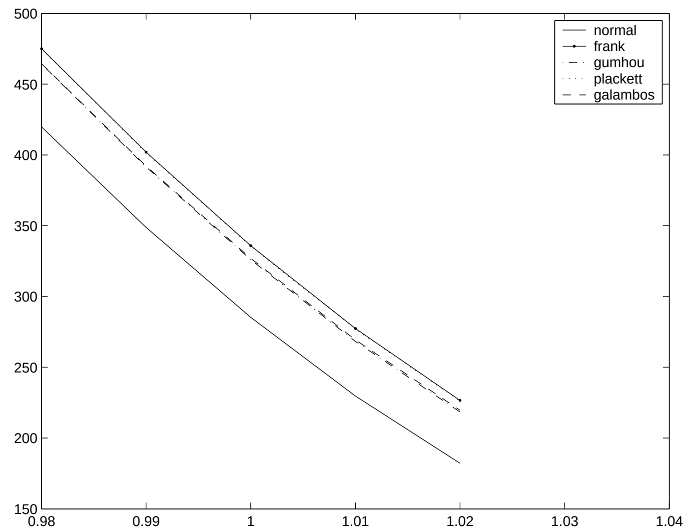

Figure 4.7: One-month call-on-max prices as a function of the strike under dynamic dependence and medium initial volatilities for various copula models.

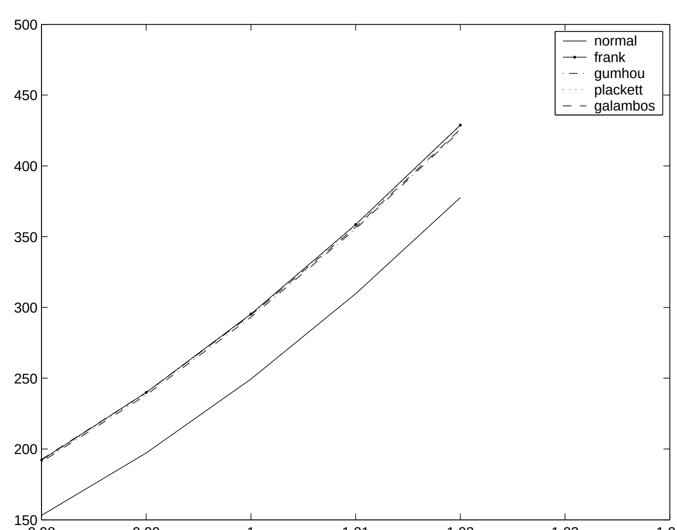

Figure 4.8: One-month put-on-min prices as a function of the strike under dynamic dependence and medium initial volatilities for various copula models.

Chapter 5

An Anatomy of Futures Returns: Risk Premiums and Trading Strategies

5.1 Introduction

Futures contracts are known to demand risk premiums in various ways. First, as the price of a futures contract will converge to the spot price of the underlying asset, we can expect that the risk factors that drive the underlying asset returns will also generate risk premiums in the corresponding futures returns. These spot-futures premiums have been analyzed for instance by Bessembinder (1992), who investigates whether futures markets and asset markets are integrated and finds that premiums for systematic risk factors in equity markets and 22 different futures markets are very similar. Although Dusak (1973) finds that for three different agricultural contracts the CAPM-beta is basically zero, Jagannathan (1985) finds that for the same three contracts the consumption-based CAPM does imply significant risk premiums and finds market prices of risk that coincide with those found in equity markets. Bessembinder and Chan (1992) report that instrumental variables known to possess forecast power in equity and bond markets also possess forecast power for prices in agricultural, metals, and currency futures markets. This evidence of predictability is consistent with the existence of time-varying risk premiums in futures markets.

Second, using a simple cost-of-carry relationship between the spot and the futures price, the term structure of futures prices depends on the term structure of the cost of carry, or yield.[1] Similarly to the term structure of interest rates, the term structure of yields can be expected to contain term premiums that show up in the expected futures returns. DeRoon, Nijman and Veld (1998) analyze the yields of five different futures contracts and show that they contain term premiums that lead to predictable variation in returns on spreading strategies (i.e., combined long and short positions in futures contracts on the same underlying asset but with different maturities.) Earlier research by Fama (1984) and Fama and French (1987) has shown that the level of the yield also contains information about the spot-futures premium. This implies that the yield is not only relevant because it gives rise to term premiums, but also because it is linked to the spot-futures premium. Furthermore, Bessembinder, Coughenour, Seguin and Smoller (1995) find a negative relation between futures yields and the spot price of the underlying asset, which is indicative of an anticipated mean reversion in asset prices.

Finally, without differentiating with respect to futures' maturities, there is an extensive literature that shows that the net hedge demand for futures contracts induces risk premiums in futures markets. This is known as the hedging pressure effect. Although the description of the hedging pressure effect dates back to Keynes (1930) and Hicks (1939), the empirical relevance of the effect has only been documented during the last two decades in Carter et al. (1983), Chang (1985), Bessembinder (1992), and DeRoon, Nijman and Veld (2000). These studies find that the net position of hedgers in futures indeed results in significant and time-varying risk premiums, an effect that is especially strong in commodity futures markets, and to a lesser extent in financial futures markets. Also, as DeRoon et al. (2000) show, there appear to be spillover effects of hedging pressure from one market to another caused by cross-hedging.

This paper analyzes trading strategies that intend to capture the various premiums in futures markets. We study the cross-section of futures returns over different markets and different delivery horizons, and link the differences in returns to the

[1] The yield of a futures contract is defined as the annualized percentage spread between the futures price and the spot price of the underlying asset.

various risk premiums that have been distinguished in these markets. The paper therefore provides a link between different parts of the futures literature, and it translates futures premiums into implementable trading strategies.

We start by analyzing the unconditional mean returns of futures contracts. This amounts to an analysis of static futures-only strategies. These strategies are of interest in themselves because they provide an understanding of passive strategies that may serve as a benchmark for hedge funds and commodity trade advisors (CTAs) that are active in those markets. Although such passive strategies serve a purpose of their own, it may very well be that the underlying factors are already captured by equity and bond markets. Therefore, we also analyze the performance of passive futures strategies relative to equity, bond, and currency benchmarks, similar to the ones used by Fung and Hsieh (1997, 2000) in analyzing the performance of hedge funds and CTAs. To the extent that futures returns and asset returns are generated by the same risk factors (as documented by Bessembinder (1992)), we may expect that there will be no outperformance of equity and bond portfolios by passive futures trading. Indeed, we find that, generally, unconditional mean returns are zero after correcting for these benchmarks. We do, however, find evidence of non-zero average returns for passive spreading strategies which go long in long-term contracts and short in the nearest-to-maturity contract.

We proceed by analyzing active trading strategies that exploit the predictable variation in futures returns. We study predictability from three sources: the term structure of futures yields, the hedging pressure effect, and past returns or momentum. The forecast power of yields, previously documented in Fama (1984) and Fama and French (1987), is re-examined using contracts that cover a wider range of futures contracts and of the term structure of futures prices than before. Next, we investigate whether the hedging pressure effect can explain the variation in spot and term premiums. Finally, we examine whether futures returns are forecastable by past returns. We find that futures yields across a wide range of maturities have substantial forecast power for both short and spreading returns. These returns are also predictable by past hedging pressure, while momentum is only present in spreading returns.

In order to exploit forecastability of futures returns, we use active trading strategies along the lines of Jegadeesh and Titman (1993) and Fama and French (1992, 1995). These trading strategies sort futures markets every period on a particular characteristic into groups, and then take long positions in one group and short positions in another. For instance, the information in the yields is used to construct a portfolio of long positions in a group of low-yield futures markets, and short positions in a group of high-yield futures markets. The returns on the nearest-to-maturity contracts in this periodically updated spreading portfolio exploit the spot-futures premium, while the returns on longer-maturity contracts also capture the term premium. Using information variables such as yields, hedging pressure, and past returns in futures markets is similar in nature to using information variables such as dividend yields or price-earnings ratios in equity markets. As with the passive strategies, we also analyze the performance of the active strategies relative to equity, bond, and currency benchmarks. Our results show that predictability in both spot and spreading returns can be exploited using yield-based trading strategies. Strategies based on past hedging pressure also outperform benchmark portfolios. Finally, in contrast with results in equity markets, momentum strategies do not appear to pay in futures markets. Our findings seem to hold up under a number of robustness tests.

The paper proceeds as follows. Section 5.2 shows a simple decomposition of futures returns that enables us to isolate the different elements of the expected futures returns. Moreover, it describes how we construct active trading strategies on the basis of predictable futures returns. Section 5.3 describes the data and provides empirical results for the passive futures strategies. Section 5.4 analyzes the active strategies based on futures yields, hedging pressure, and momentum. In Section 5.5 we present the conclusions and examine the robustness of the empirical results.

5.2 Methodology

5.2.1 A decomposition of futures returns

We start our analysis with a simple decomposition of futures returns that highlights the different premiums that are present in futures markets. Denoting by S_t the spot price of the underlying asset, and by $F_t^{(n)}$ the futures price for delivery at time $t+n$, we use the storage model or cost-of-carry relation, which dates back to Working

(1949) and Brennan (1958), to define the *yield* $y_t^{(n)}$:

$$F_t^{(n)} = S_t \exp\{y_t^{(n)} \times n\}. \tag{5.1}$$

Thus, $y_t^{(n)}$ is the per-period yield for maturity n, analogous to the n-period interest rate. It is also the slope of the term structure of (log) futures prices, as is readily seen by solving (5.1) for $y_t^{(n)}$. This yield consists of the n-period interest rate, and possibly other items such as dividend yields, foreign interest rates, storage costs, and convenience yields, depending on the nature of the underlying asset.

From the one-period expected log-spot return we define the spot risk premium $\pi_{s,t}$ as the expected spot return in excess of the one-period yield,

$$\begin{aligned} E_t\left[r_{s,t+1}\right] &= E_t\left[\ln(S_{t+1}) - \ln(S_t)\right] = E_t\left[s_{t+1} - s_t\right] \\ &= y_t^{(1)} + \pi_{s,t}, \end{aligned} \tag{5.2}$$

where we take expectations E_t conditional on the information available at time t and use lower cases to denote log prices. The spot premium $\pi_{s,t}$ can be interpreted as the expected return in excess of the short-term yield, similar to stock returns in excess of the short-term interest rate and dividend yield. It is easy to show that the spot premium is also the expected return of the short-term futures contract, $r_{f,t+1}^{(1)}$, i.e., the return on the futures contract that matures at time $t+1$. This follows from applying the cost-of-carry relation in (5.1) to such a contract and from the fact that the futures price converges to the spot price at the delivery date:

$$\begin{aligned} E_t[r_{f,t+1}^{(1)}] &= E_t[s_{t+1} - f_t^{(1)}] \\ &= E_t[s_{t+1} - s_t - y_t^{(1)}] = \pi_{s,t}. \end{aligned} \tag{5.3}$$

Next, we define a term premium $\pi_{y,t}^{(n)}$ similarly to DeRoon et al. (1998), as the (expected) deviation from the expectations hypothesis of the term structure of yields:

$$ny_t^{(n)} = y_t^{(1)} + (n-1)E_t[y_{t+1}^{(n-1)}] - \pi_{y,t}^{(n)}. \tag{5.4}$$

DeRoon et al. (1998) estimate the term premiums for five different futures contracts, using one-factor models for the yields similar to the Vasicek-model and the Cox-Ingersoll-Ross model for the term structure of interest rates. Without imposing any structure on the term structure of yields, it is important to note that the term premium $\pi_{y,t}^{(n)}$ also shows up in the expected return on a futures contract for delivery at time $t+n$. This follows from the log return on such a contract and applying the cost-of-carry relation again. Using the definitions of $\pi_{s,t}$ and $\pi_{y,t}^{(n)}$ in (5.2) and (5.4) it is easily seen that the expected one-period futures return for a contract that matures at time $t+n$ is:

$$\begin{aligned} E_t[r_{f,t+1}^{(n)}] &= E_t[f_{t+1}^{(n-1)} - f_t^{(n)}] \\ &= \pi_{s,t} + \pi_{y,t}^{(n)} \equiv \pi_{f,t}^{(n)}. \end{aligned} \tag{5.5}$$

Thus, the expected one-period return on an n-period futures contract consists of the futures premium $\pi_{f,t}^{(n)}$ only, which can be separated in a spot premium $\pi_{s,t}$ and a term premium $\pi_{y,t}^{(n)}$. Notice that it follows immediately from (5.3) that $\pi_{y,t}^{(1)} = 0$, i.e., the short term futures contract does not contain a term premium.

This decomposition of the futures premium into a spot premium and a term premium is a useful starting point for our analysis. From (5.3) we have that the spot premium can be identified with a long position in a short-term futures contract. Using spreading strategies it is also possible to isolate the term premium. Combining a long position in an n-period futures contract with a short position in an m-period futures contract on the same underlying asset, the expected return on this portfolio is

$$E_t[r_{f,t+1}^{(n)} - r_{f,t+1}^{(m)}] = \pi_{y,t}^{(n)} - \pi_{y,t}^{(m)}. \tag{5.6}$$

If $m = 1$, i.e., if we combine a long position in a long-term contract with a short position in the short-term contract, then the expected return on the spreading strategy is generated by one term premium $\pi_{y,t}^{(n)}$ only. Otherwise the expected return is a combination of two term premiums.

The decomposition in (5.5) is important, because the two risk premiums $\pi_{s,t}$ and $\pi_{y,t}^{(n)}$ are likely to compensate for different risk factors. For instance, in case of index futures, $\pi_{s,t}$ reflects equity market risk, whereas $\pi_{y,t}^{(n)}$ reflects interest rate risk. In case of oil futures the spot premium reflects the oil price risk, whereas the term premium mainly reflects the risk that is present in the convenience yield. Therefore, we will focus on short-term futures trading strategies and on spreading strategies in order to capture the expected returns generated by the different risk factors, i.e., to capture both the spot premiums and the term premiums.

5.2.2 Predictability and active trading strategies

We now show how predictable variation in futures returns can be used to construct simple, implementable active trading strategies. Suppose that the spot-futures premium $\pi_{s,t}$ in a particular market can be forecast by an instrument x_t, observable at time t, through the following simple linear relation:

$$\pi_{s,t} = \alpha + \beta x_t, \tag{5.7}$$

and suppose that β is positive. As mentioned before, the spot premium is the expected return on the short-term futures contract; see Equation (5.3). Thus, investors could take a long position in the short-term contract whenever the instrument has a high value, and a short position otherwise. Such an active trading strategy would yield a return that is on average higher than the return on a passive strategy which is long in the contract at any given time. If markets are efficient, this higher expected return compensates for additional risk involved in the active strategy.

Similar trading strategies can be constructed if term premiums can be explained by observable variables. In that case, we use the fact that the n-th term premium $\pi_{y,t}^{(n)}$ is the expected return on a spreading strategy which takes a long position in the n-period futures contract combined with a short position in the short-term contract; see Equation (5.6). Denoting again by x_t the forecast variable, and assuming a positive relation between the term premium and the forecast variable, a simple active trading strategy would be to take a long position in the long-term contract combined with a short position in the short-term contract whenever the instrument has a high value, and a short position in the long-term contract combined with a long position in the short-term contract otherwise.

In this paper we focus on predicability of returns using instruments which are observed in all futures markets, such as futures yields, hedgers' positions, and past returns. This allows us to construct trading strategies along the lines of Jegadeesh (1990), Lehmann (1990), and Jegadeesh and Titman (1993) which operate in multiple markets. These studies analyze the returns on momentum strategies in equity markets. Momentum strategies are spreading strategies which buy stocks that have performed well in the past and sell stocks that have performed poorly in the past. Similarly, trading strategies based on, for instance, futures yields can be formed by ranking futures markets on their yields at a given point in time, and taking positions in high-yield contracts (for instance, the one-third highest yield contracts) combined with offsetting positions in low-yield contracts (the one-third lowest yield contracts). At a later date, the futures portfolio is updated by sorting the markets again on the then prevailing yields, and adapting positions accordingly; and so on. Note that this type of strategy depends on the rank order statistics of the forecast variable (in this case the futures yields), and, hence, requires the use of market-specific instruments. Forecast variables which are not directly related to futures markets, such as the equity and bond market variables used by Bessembinder and Chan (1992), are not applicable here.

In Section 5.4 we analyze the time variation of both components of the futures premium using futures yields, hedging pressure, and past returns as explanatory variables, and we examine if any explanatory power found can be exploited using the simple type of trading strategy sketched above. First, however, we analyze the performance of passive futures strategies.

5.3 Data, descriptive statistics, and passive trade

We analyze a data set consisting of semi-monthly observations of 23 U.S. futures markets over the interval January 1986 to December 2000 obtained from the Futures Industry Institute (FII) Data Center. Using the classification of Duffie (1989), the data can be divided into 16 commodity futures contracts and seven financial futures contracts. The commodities include grains (wheat, corn, and oats), soybean complex (soybeans, soybean oil, and soybean meal), livestock (live cattle, feeder cattle, and live hogs), energy (crude oil and heating oil), metals (gold, silver, and platinum), and foodstuffs (coffee and sugar). The financial contracts include interest rates (Eurodollars), foreign currencies (Swiss francs, British pounds, Japanese yen, and Canadian dollars), and equity indices (S&P 500 and NYSE composite). These markets have relatively large trading volumes and provide a broad cross-section of futures markets. Details about the delivery months and the exchanges where these futures contracts are traded are in Table 5.1.

Following common practice in the literature (see, for example, Fama and French (1987), Bessembinder (1992), Bailey and Chan (1993), Bessembinder et al. (1995) and DeRoon et al. (2000)), we construct continuous series of futures returns by using rollover strategies. For the nearest-to-maturity series a position is taken in the nearest-to-maturity contract until the delivery month, at which time the position changes to the contract with the following delivery month, which is then the nearest-to-maturity contract. In this way we are able to derive return series for second nearby contracts, third nearby contracts, et cetera. Prices of futures observed in the delivery month are excluded from the analysis to avoid obligatory delivery of the physical asset. At least four different return series exist for each contract, up till 12 series for the oil contracts. Depending on the delivery dates during the year, the different series are for delivery one to three months apart. We obtain a maximum of 376 observations per series.

Since the delivery dates are more than two weeks apart for all contracts, and since for many futures the delivery dates are not evenly spread over the year, it is not possible to get the exact short futures returns on regular time intervals. Assuming that the term premium is relatively unimportant for the nearest-to-maturity contracts, we use the returns on those contracts as a proxy for the short futures returns, $s_{t+1} - f_t^{(1)}$. The first column of Table 5.2 gives the average returns of the nearest-to-maturity contracts for the different futures. These are estimates of the unconditional spot-futures premiums $E[\pi_{s,t}]$. Except for oats, which has an estimated premium of -15.5 percent on an annual basis, and the equity indices, which require compensations of 8.6 and 7.8 percent, the hypothesis that the mean short futures return is zero cannot be rejected for any of the futures markets at the 5 percent level, indicating that most of the markets considered do not demand significant spot premiums.[2] Similar evidence is found in, e.g., Bessembinder (1992), Bessembinder and Chan (1992), and DeRoon et al. (2000) who also study broad cross-sections of futures markets using various sample periods that only partially overlap with our sample period. However, as Bessembinder and Chan (1992) point

[2] All statistical tests were conducted using White's (1980) heteroskedasticity-consistent standard errors, unless stated differently.

Table 5.1: **Futures exchanges and Delivery Months**

Contract	Exchange	Delivery months
Commodities		
Grains		
Wheat	Chicago Board of Trade	3 5 7 9 12
Corn	Chicago Board of Trade	
Oats	Chicago Board of Trade	3 5 7 9 12[a]
Oil & Meal		
Soybeans	Chicago Board of Trade	1 3 5 7–9 11
Soybean oil	Chicago Board of Trade	1 3 5 7–10 12
Soybean meal	Chicago Board of Trade	1 3 5 7–10 12
Livestock		
Live cattle	Chicago Mercantile Exchange	2 4 6 8 10 12
Feeder cattle	Chicago Mercantile Exchange	1 3–5 8–11
Live (lean) hogs	Chicago Mercantile Exchange	2 4 6–8 10 12
Energy		
Crude oil	New York Mercantile Exchange	All
Heating oil	New York Mercantile Exchange	All
Metals		
Gold	Commodity Exchange, Inc.	2 4 6 8 10 12[b]
Silver	Commodity Exchange, Inc.	2 4 6 8 10 12[b c]
Platinum	New York Mercantile Exchange	1 4 7 10
Foodstuffs		
Coffee C	Coffee, Sugar & Cocoa Exchange	3 5 7 9 12
Sugar #11	Coffee, Sugar & Cocoa Exchange	1[d] 3 5 7 9[e] 10
Financials		
Interest Rates		
Eurodollars	International Monetary Market	3 6 9 12[f]
Foreign Currencies		
Swiss franc	International Monetary Market	3 6 9 12
Pound Sterling	Chicago Mercantile Exchange	3 6 9 12
Japanese yen	International Monetary Market	3 6 9 12
Canadian dollar	International Monetary Market	3 6 9 12
Indices		
S&P 500	International Monetary Market	3 6 9 12
NYSE Composite	New York Futures Exchange	3 6 9 12

[a] November 2000 and January 2001 contracts also traded; [b] All delivery months traded in 1995–2000; [c] Except November 1998; [d] January contracts traded until 1990; [e] September contracts traded until 1987; [f] All delivery months traded in November 1995–June 2001.

out, "while zero-mean returns are consistent with the absence of risk premia, they are also consistent with the existence of time-varying risk premia." Hence, the fact that returns are zero on average does not preclude non-zero conditional premiums.

The next columns of Table 5.2 show the average returns on passive spreading strategies which combine a long position in a longer-maturity contract with a short position in the nearest-to-maturity contract. Using (5.6) and assuming that the term premium on the short contract is approximately zero, the average returns on the spreading strategies give us estimates of the unconditional term premiums $E[\pi_{y,t}^{(n)}]$ for various maturities. Significant term premiums are found for many markets, in particular grains, soybean complex, heating oil, and Eurodollar futures. For many futures there is also a clear pattern in the average spreading returns, implying an average term structure of futures prices that is either upward or downward sloping. Except for the financial futures, the estimated term premiums often have the opposite sign of the corresponding spot premiums. As is clear from (5.5), an estimate of the total unconditional futures premium is obtained by adding the average short return to the average spreading return.

The standard deviations also show a clear structure over the different maturities, where the volatility of the spreading strategies is always increasing in the maturity of the contract. The volatility of the short-term futures contract is always higher than than the volatility of the spreading strategies for the same underlying asset, implying that spot price risk is larger than yield or basis risk. However, for many commodity markets the yield or basis risk is as high as the spot price risk of the index futures and even higher than the spot price risk of the Eurodollar futures as well as some currency futures.

Thus, Table 5.2 illustrates the relevance of both spot premiums and term premiums as components of the average returns on passive, futures-only strategies. We analyze the underlying factors that determine these premiums by examining the relative performance of these passive strategies with respect to several benchmarks. First, we test whether the returns can be explained by the Capital Asset Pricing Model. We consider as a benchmark the return on the MSCI U.S. equity index in excess of the risk-free rate as measured by the one-month Eurodollar deposit rate. The first column of Table 5.3 gives Jensen's unconditional measure of performance—

Jensen's alpha—for the nearest-to-maturity contracts. Apart from the index futures, the nearest-to-maturity alphas do not differ much from their unconditional means. Indeed, the corresponding CAPM-betas (not reported here) are close to zero. This is consistent with Dusak's (1973) finding that for wheat, corn, and soybean futures systematic risk is basically zero. As expected, the CAPM captures the factors underlying the spot premium in the index futures well, with betas close to 1.0 and alphas indistinguishable from zero. However, the significant spot premium in the short-term oats contract cannot be explained by domestic equity market risk.

The alphas of the spreading strategies are reported in the next columns of Table 5.3. The futures markets that showed significant term premiums also have non-zero alphas, which are similar to their unconditional means. Indeed, the CAPM-betas for the spreading strategies are basically zero, implying that the term premiums in futures markets cannot be accounted for by the market portfolio. Most futures show an upward or downward sloping term structure of Jensen's alphas.

As an alternative to the CAPM we consider a six-factor model which includes, apart from the excess returns on the MSCI U.S. equity index, five other benchmarks. They are: the excess returns on non-U.S. equities (from MSCI), U.S. and non-U.S. government bonds (from J.P. Morgan), emerging market stocks (from IFC), and the U.S. dollar (from the U.S. Federal Reserve). These benchmarks are similar to the ones used by Fung and Hsieh (1997, 2000) in analyzing the performance of hedge funds and CTAs. The remaining columns of Table 5.3 present the unconditional multi-factor alphas for the short futures returns and the returns on the spreading strategies. By and large the same pattern emerges; non-zero alphas are found in the same markets as before (grains, soybean oil and meal, heating oil, silver, and Eurodollars), and they are of the same sign and order of magnitude as in the CAPM case.

To sum up, Table 5.3 demonstrates that passive rollover trading strategies, which go long in the nearest-to-delivery futures contract, do not outperform or are not outperformed by the market portfolio, except in one or two cases. There is somewhat more evidence that passive, short-term trading produces abnormal returns relative to a set of equity, bond, and currency benchmarks. Passive spreading strategies, which capture the term structure of futures prices, do yield abnormal returns in a significant number of markets, both with respect to the market and multiple benchmarks.

5.4 Active trading strategies

We now turn to an analysis of the time variation in the spot and term premiums of futures returns. Our goal is to examine whether the predictable variation in either component can be exploited in simple, active trading strategies explained in Section 5.2.2. Three sources of predictability are considered: futures yields, hedging pressure, and past returns.

5.4.1 Yield-based strategies

Using (5.1), the yield on the m-th nearby futures contract is defined as the spread between the m-th nearby log futures price and the log spot price of the underlying asset, divided by the remaining time to maturity,

$$y_t^{(m)} = \frac{f_t^{(m)} - s_t}{T^{(m)} - t}, \tag{5.8}$$

where $T^{(m)}$ is the delivery date of the m-th nearby contract.[3] Since the moment of settlement within the delivery month is often at the option of one of the contract participants or not easily determined due to market-specific regulations, we cannot measure the time to maturity of the contract exactly. To solve this problem, we assume that contracts are settled at the 15-th of each delivery month. This assumption may potentially result in some measurement error, in particular for the nearest-to-delivery contracts, since the relative effect of errors will be largest on the shortest maturity, whereas it vanishes for longer-maturity contracts. It is important to note, however, that the results for the yield-based trading strategies are not likely to be affected by the exact measurement of the futures yields, since only the order statistics of the relative yields—not their nominal value—play a role in the trading strategies.

[3]Note the difference in notation with Section 5.2. The number in brackets now refers to the order of maturity, not the actual time to maturity of the contract.

Table 5.4 shows the average annualized yields of the first to the sixth nearby contract for every futures market along with the standard deviations. Upward as well as downward sloping term structures are common in futures markets, apparently independent of the classification given in Table 5.1. Yields tend to be larger in absolute value and more variable for agricultural futures (grains, soybean complex, livestock, and foodstuffs) than for energy and metal futures. Financial futures have even smaller yields and show the least variability. For most commodity futures, there is either an upward or a downward sloping term structure of yields, while index and currency futures show a constant term structure.

The theory of storage—which predicts that a futures' yield equals the interest rate plus the marginal storage cost, less the marginal convenience yield from holding the underlying asset—can help us interpret these figures. Convenience yields and storage costs are important for many commodities, and they are likely to be more important and variable for agricultural futures than for energy and metal futures; see, e.g., Bessembinder et al. (1995). For the currency and index futures, no storage cost or convenience yield is likely to be included in the yield to maturity. Theory predicts that the yield on currency futures is equal to the differential between domestic and foreign interest rates. For instance, for Japanese yen futures, the positive mean yield implies that U.S. interest rates were, on average, higher than Japanese interest rates by about 3.0 percent per year. The relatively constant term structure of yields observed for the currency futures implies that there have been, on average, little differences between interest rate differentials across different maturities. For index futures, the yield on the n-th nearby contract reflects the domestic interest rate of the same maturity. The flat yield term structure implies that the term structure of interest rates was relatively flat on average.

Documenting predictability from yields

Previous research has examined the forecast power of yields for futures returns. Fama (1984) shows that the current short-term futures-spot differential, or basis, i.e., the numerator in (5.8), has power to predict the future change in futures prices in a number of currency futures markets. Fama and French (1987) find that the short-term basis in agricultural and metal markets also contains information about the variation in futures premiums, both the spot-futures premium as well as longer-term futures premiums. DeRoon et al. (1998) find that the spreads between futures and spot prices have power to explain term premiums for gold and soybean contracts.

We re-examine the forecast power of futures yields using not only the short-term yield but the entire term structure of futures yields. For each futures markets, we regress the semi-monthly return on the nearest-to-maturity contract on the current yield of the m-th nearby contract,

$$r_{f,t+1}^{(1)} = \alpha_{1m} + \beta_{1m} y_t^{(m)} + \varepsilon_{t+1}^{(1,m)}, \tag{5.9a}$$

for $m = 1, \ldots, 6$. Deviations of β_{1m} from zero imply that the spot-futures premium can be explained by the m-th nearby yield. Analogously, we regress the return on spreading strategies which go long in the n-th nearby contract and short in the nearest-to-maturity contract on the m-th nearby yield,

$$r_{f,t+1}^{(n)} - r_{f,t+1}^{(1)} = \alpha_{nm} + \beta_{nm} y_t^{(m)} + \varepsilon_{t+1}^{(n,m)}, \tag{5.9b}$$

for $n = 2, \ldots, 6$ and $m = 1, \ldots, 6$. Evidence of non-zero β_{nm} indicates that the m-th nearby yield has explanatory power for the n-th term premium in futures prices.

Equations (5.9a) and (5.9b) lead to 36 regressions for each of the 23 markets under scrutiny. Table 5.5 summarizes the results of these regressions. Panel A reports for all (n, m) combinations the p-value for a test that the slope coefficients are zero in all markets. Clearly, this hypothesis is rejected in many cases. In particular, the short yield has strong forecast power for short as well as most spreading returns, and the term structure of yields appears to contain information about both short returns and second, fourth, and sixth nearby spreading returns. Panel B of Table 5.5 shows for each (n, m) pair the number of markets for which predictability is found, i.e., the number of slope coefficients which differ significantly from zero at the 10 percent level. Predictability seems to be strongest for the short return using the short yield—significance is found in eight out of 23 markets.

Moreover, a clear pattern emerges from the signs of the slope coefficients, which are marked by a + or − in Panel B of Table 5.5. All markets in which predictability of the short return is found have negative yield coefficients, whereas virtually all markets with predictable spreading returns have positive yield coefficients. Hence, for a significant number of contracts, current yields tend to have a negative impact on the spot-futures premium and a positive impact on term premiums. The negative effect of yields on the spot-futures premium is also found by Fama (1984) and Fama and French (1987). Research on the relation between yields and term premiums is scarce; DeRoon et al. (1998) examine five markets using observation from March 1970 until December 1993, and detect a significant relation for gold and soybean contracts, which is negative rather than positive. We do not find reliable forecastability for these contracts in our sample period, which only partly overlaps with theirs.

Exploiting predictability from yields

The negative relation between futures yields and short-term returns suggests that a simple, active trading strategy, which goes long in the nearest-to-maturity contract if current yields are low, and short if current yields are high, would yield a positive expected return. Similarly, profits could be made using a trading strategy which takes a long position in a long-term contract combined with a short position in the nearest-to-maturity contract if current yields are high, and opposite positions if current yields are low. By constructing cross-market portfolios of offsetting positions in low-yield and high-yield markets, one may exploit the predictability of returns in futures markets as a whole.

Analogously to the work by Jegadeesh (1990) and others on momentum strategies, and the work by Fama and French (1992, 1995) on size and book-to-market factors in equity markets, we sort all 23 futures contracts on their short yield at every date in the sample into three groups of about equal size: a low-yield group, a high-yield group, and a group with intermediate yields.[4] We then form a simple spreading portfolio of equally-weighted long positions in the low-yield group combined with as many equally-weighted short positions in the high-yield group. Portfolios are updated in this way every period. Similar portfolios are constructed using yields of longer maturities.

The first column of Table 5.6 shows the averages, standard deviations, and Jensen's alphas (both CAPM and multi-factor based) for the nearest-to-maturity returns on these active trading strategies. Clearly, the average portfolio returns are positive and significantly different from zero. The average return increases from 7.4

[4] The low-yield group and the high-yield group each consist of the nearest integral value of $N_t/3$ contracts, where N_t is the number of markets for which price data is observed.

percent on an annual basis for the short yield-based strategy to 12.9 percent per year as the maturity of the yield goes up, but this also leads to a higher risk as measured by the standard deviation of the return. The performance of the yield-based strategies is little changed after correcting the returns for market risk. Moreover, the hypothesis that Jensen's alphas are zero is strongly rejected for all maturities. If, instead of the CAPM, we use a six-factor benchmark, the results are only slightly less powerful.

The next columns of Table 5.6 present the results for active trading strategies which combine two spreading strategies: one spreading strategy takes long positions in long-term contracts and short positions in short-term contracts in the high-yield group; the other takes short positions in the long-term contracts and long positions in the short-term contracts in the low-yield group. As expected, mean returns and standard deviations are lower than for the nearest-to-maturity returns. There is a clear upward sloping term structure in the expected returns and standard deviations of the trading strategies. All average returns differ significantly from zero at the 10 percent level, while many differ significantly from zero at the 5 and even the 1 percent level. Again, the size and significance of the results hardly changes if returns are corrected for market risk. A multi-factor correction dampens the results somewhat, but most alphas remain significant at the 10 percent level, with results being particularly strong for the longer-maturity term-spreading returns.

5.4.2 Strategies based on past hedging pressure

Next, we investigate the time variation of risk premiums through the hedging pressure effect. The hedging pressure effect implies that the net demand for futures contracts induces risk premiums in futures markets. Previous studies find that the empirical relevance of the effect is substantial. Carter et al. (1983) analyze the weekly returns on contracts of different delivery months in wheat, corn, soybean, cotton, and cattle markets, and provide strong statistical evidence that returns are a function of speculators' net positions. They are unable to distinguish between spot and term premiums, because they use returns on contracts of a fixed delivery month rather than a fixed time-to-maturity. Bessembinder (1992) analyzes the variation in the spot-future premium by using nearest-to-maturity returns in 22 futures markets including agricultural, metal, foreign currency, and (other) financial contracts. He finds that mean returns depend on net hedging, particularly in non-financial futures markets. DeRoon et al. (2000) use nearest-to-maturity as well as second nearby contracts and they also find significant and time-varying risk premiums. Moreover, they find evidence for spillover effects of hedging pressure from one market to another.

These studies do not make a distinction between the spot-futures premiums and term premiums in futures markets. It is not clear a priori if net hedge demand has the same influence on spot premiums as on term premiums. We examine the relevance of the hedging pressure effect for both spot and term premiums. Furthermore, previous studies have used current measures of net hedging to explain the variation in expected futures returns. However, data on hedge positions, which are published in the Commitment of Traders reports issued by the Commodity Futures Trading Commission (CFTC), only become available at a time lag of at least three days.[5] Hence, information on hedge positions is only observable to investors after a reporting lag, and therefore cannot be used as a conditioning variable in an active trading strategy. We examine whether the hedging pressure effect, which has been shown to have strong explanatory power for futures returns if no reporting lag is taken into account, also contains information about returns if net hedging is lagged one period. Moreover, we analyze whether predictability, if any, can be exploited using active trading.

Following previous works, we define the hedging pressure variable in a futures market as the difference between the number of short hedge positions and the number of long hedge positions by large traders, relative to the total number of hedge positions by large traders in that market,

$$q_t = \frac{\text{\# of short hedge positions} - \text{\# of long hedge positions}}{\text{total \# of hedge positions}}, \tag{5.10}$$

where positions are measured by the number of contracts in the futures market. Hedging pressures are calculated from the aforementioned Commitment of Traders reports, which were available semi-monthly (and every two weeks as of October 1992) in our sample period. The first two columns of Table 5.7 show the averages

[5] The Commission reports on her website that "[t]he Commitments of Traders reports are released at 3:30 pm Washington D.C. time. The [...] reports are usually released Friday. The release usually includes data from the previous Tuesday."

and standard deviations of the hedging pressure variables for all futures markets. Both net short and net long hedging are common, and variability is considerable. These figures are in line with results reported by, e.g., DeRoon et al. (2000). The next columns show the autocorrelation of the hedging pressure variables at the first four lags. Clearly, hedging pressure is strongly persistent for every market. The pattern resembles that of a first-order autoregressive model. One may expect that due to this strong persistence, return predictability is not much affected by lagging the hedging pressure measure.

Documenting predictability from past hedging pressure

Panel A of Table 5.8 documents the predictability of nearest-to-maturity and term-spreading returns using lagged hedging pressure. It summarizes the results of six regressions for every market, i.e., one regression of the nearest-to-maturity return on past hedging pressure, and five regressions of spreading returns on past hedging pressure:

$$r^{(1)}_{f,t+1} = \alpha_1 + \beta_1 q_{t-1} + \varepsilon^{(1)}_{t+1} \tag{5.11a}$$

$$r^{(n)}_{f,t+1} - r^{(1)}_{f,t+1} = \alpha_n + \beta_n q_{t-1} + \varepsilon^{(n)}_{t+1}, \tag{5.11b}$$

for $n = 2, \ldots, 6$. The first line of Panel A shows the p-value for a joint test that all slope coefficients are zero. For the short returns, the hypothesis that slopes are zero is rejected at the 10 percent level; evidence of non-zero slopes is much stronger for the spreading returns, indicating that past hedging pressure explains the variation in term premiums much better than the variation in spot premiums. The weak forecast power found for the short-term returns is striking in the light of the strong, positive effect of hedging pressure found by studies which do not take into account a reporting lag, particularly given the high level of persistence in the hedging pressure variables. One possible explanation for this result is that there is a negative relation between past hedging pressure and the forecast errors of the regressions which use current instead of lagged hedging pressure. Indeed, we find a large negative covariance between these variables for each market which cancels out (and in some cases dominates) the effect of persistence in the hedging pressure variable.

Exploiting predictability from past hedging pressure

Panel A of Table 5.8 also reports the number of markets in which we find significant slope coefficients, broken down according to sign. In most cases where predictability of the short returns is found, past hedging pressure has a negative effect on returns. This is opposite to the effect documented for current hedging pressure for the reasons mentioned above. In contrast, the much stronger effect of past hedging pressure on the variation in the term premium is in almost all cases negative. This suggests that trading strategies which go long in contracts which have had high hedging pressure in the past and short in contracts which have had low hedging pressure in the past, could be profitable. Panel B of Table 5.8 shows that this is indeed the case. Significant abnormal returns can be achieved by trading according to a strategy which sorts futures markets every period into three equal-sized groups according to lagged hedging pressure, and takes (equally weighted) long positions in long-term contracts combined with short positions in the nearest-to-maturity contract in markets with high lagged hedging pressure, and opposite positions in markets with low lagged hedging pressure. The higher the maturity of the long-term contract, the higher the expected return (but also the risk) of the trading strategy. As expected, the strategy which is designed to exploit the (weak) predictability in the variation of the spot premium does not outperform the market or the six-factor benchmark.

5.4.3 Momentum strategies

Finally, we examine the presence of the momentum effect in futures markets and we analyze the performance of momentum strategies. From equity markets we know that stocks that have performed well in the past are likely to perform well in the future, while stocks that have performed poorly in the past are likely to perform poorly in the future. Such predictability in equity markets has been shown to be exploitable using spreading strategies. Early works include Jegadeesh (1990), Lehmann (1990), and Jegadeesh and Titman (1993) who show that strategies which buy past winners and sell past losers generate significant positive abnormal returns. In a recent paper, Jegadeesh and Titman (2001) show that, contrary to other stock market anomalies, momentum profits have continued in the 1990s. Whether momentum is also present and profitable in futures markets is an open question.

Documenting momentum

If futures returns are autocorrelated over time, then this is evidence of momentum in futures markets. Historically, little evidence has been found of autocorrelation in futures returns. For instance, Dusak (1973) reports semi-monthly serial correlations at up to 10 lags for the returns on wheat, corn, and soybean contracts during the 1950s and 1960s, and finds that hardly any differ significantly from zero. In a sample of 12 agricultural, foreign currency, and metal futures markets in the 1970s and 1980s analyzed by Bessembinder and Chan (1992), no appreciable autocorrelation is found either. We also find little evidence of autocorrelation in the subsequent years among an even more extensive cross-section; only silver (−.15) and sugar (.16) contracts show significant semi-monthly serial correlation in the nearest-to-maturity returns at the 1 percent level, while soybeans (−.12) and crude oil (.11) also show significant autocorrelations at the 5 percent level. Equally weak (or even weaker) results are found for the serial correlations at further lags, largely confirming the evidence documented in previous studies.

The autocorrelation coefficient of futures returns coincides with the slope coefficient in a forecast regression of current on past returns. Panel A of Table 5.9 summarizes the results of such regressions for the nearest-to-maturity returns and the returns on term-spreading strategies using semi-monthly lags. As noted before, hardly any momentum is found at this horizon for the nearest-to-maturity contracts. The hypothesis that all slope coefficients, i.e., all autocorrelation coefficients, are zero cannot be rejected at all conventional confidence levels. This implies that the variation in spot premiums cannot be explained by its own history.

We do, however, find significant results for the term premiums. A substantial number of futures markets shows momentum in the spreading returns at all delivery horizons, while the hypothesis that autocorrelations are zero is rejected in all cases (at the 5 percent level) and strongly rejected (i.e., at the 1 percent level) in most. Momentum therefore induces term premiums in futures markets, a result that has to our knowledge not been documented previously. There is, however, no clear pattern in the direction of the predictability. In some markets, past returns have a positive effect on future returns, while in others a negative effect is found. This finding suggests that the active trading strategies discussed earlier have a lesser chance of performing abnormally.

Exploiting momentum

Indeed, Panel B of Table 5.9 shows that there are no significant (abnormal) returns to be made from taking long positions in futures markets with low past returns and short positions in markets with high past returns. Therefore, the momentum effect, which is clearly present in the term structure of futures prices, does not appear to be exploitable using such simple trading rules. Hence, the profitability of momentum strategies in equity markets does not translate to futures markets. However, we do retain the pattern of average returns, standard deviations, and alphas increasing with maturity observed for yields and past hedging pressure.

5.5 Conclusions and robustness of the results

For convenience, we briefly recapitulate the conclusions of the previous sections here. Firstly, we find zero-mean unconditional spot-futures premiums in virtually all markets, while unconditional term premiums are non-zero for some markets. Basically the same results are obtained after correcting for market or multi-factor risk, except for the financial index spot-futures premiums, which appear to be largely due to market risk. Both premium components can be explained by futures yields and past hedging pressure, while the momentum effect appears to have explanatory power for the term premiums only. Momentum is not found for the spot premiums. Finally, predictability in both spot and term premiums is found to be exploitable using yield-based strategies, while strategies based on past hedging pressure are only profitable using the term premiums. Momentum strategies do not yield (abnormal) returns.

To test the robustness of these conclusions, we perform a number of sensitivity tests. First, we investigate the possibility that the size and predictability of futures premiums changes over the sample period by splitting the sample period in half and redoing the entire analysis for each subperiod. Second, we examine whether our findings stay the same if, instead of semi-monthly returns, we use returns with longer horizons. Finally, we test whether the active trading strategies which were found to outperform benchmarks, are still profitable if transaction costs are taken into account.

5.5.1 Subperiod results

The sample period is split up into two intervals of about equal size. The first subperiod consists of semi-monthly data from January 1986 to December 1993, and the second subperiod is from January 1994 to December 2000.

The average returns and volatilities of the nearest-to-maturity contracts and the spreading strategies are about the same size in each of the subperiods as in the entire sample. We find about the same number of non-zero mean returns, albeit that some of markets in which they are obtained differ across subperiods. Similar results are obtained for the unconditional Jensen's alphas.

Furthermore, we find that futures yields have strong forecast power in both periods, albeit slightly less than in the entire sample. However, while the first subperiod shows a clear pattern of negative regression coefficients for the spot premium and positive coefficients for the term premiums, the results are mixed in the second subperiod. Nevertheless, the yield-based trading strategies produce significant and positive returns in both periods, with constant volatilities across time. Moreover, the strategies outperform the market to a similar degree in both periods. However, they do not outperform the six-factor benchmark in the first subperiod, whereas they do in the second.

The hedging-pressure forecast regressions show similar results across the two subperiods. As in the entire sample, we find strong predictability for the term premiums, while only weak predictability is found for the spot premium. Again, the predictability in the term premiums turns out to be exploitable using active trading strategies. In fact, in the first subperiod positive abnormal returns are obtained for the spreading strategies which use the third, fourth, and fifth nearby contracts, while the strategies which use the second and third nearby contracts yield positive abnormal returns in the second subperiod.

Finally, there is a momentum effect in both periods which is comparable to the momentum effect in the entire sample. We find that term premiums are predictable by past term premiums, with positive and negative coefficients in each subperiod.

As in the large sample, no predictability is found for the spot premiums. Also, the momentum effect does not appear to be exploitable in either subperiod.

To sum up, although there are differences between mostly individual premiums across time, the forecast power of yields, hedging pressure, and past returns, and the extent to which active trading strategies can exploit forecastability, are obtained consistently through time.

5.5.2 Multi-period returns

The results so far are based on semi-monthly returns as in, for example, DeRoon et al. (2000). However, other authors have used different horizons to analyze futures premiums. Fama (1984), Fama and French (1987), Chang (1985), and Bessembinder and Chan (1992) examine monthly returns, while Carter et al. (1983) use weekly returns, and Bessembinder (1992) and Bessembinder et al. (1995) analyze daily returns. DeRoon et al. (1998) analyze both daily returns and returns over longer, contract-specific holding periods. To test the resiliency of our results in this dimension, we repeat all analysis for different return horizons. The semi-monthly frequency at which we observe the hedging pressure data dictates the minimum return horizon we can use. Hence, the basic holding period is half a month. We construct multi-period returns by adding semi-monthly (log) returns over multiple periods. We consider two-period (monthly), three-period (semi-quarterly), and four-period (bi-monthly) returns.[6]

In the interest of conciseness, rather than repeat all empirical results for the multi-period returns, which are available from the authors on request, we briefly summarize the main conclusions here. The multi-period short and spreading returns from passive trading show little difference from the one-period results; we obtain unconditional futures premiums which are similar in size and in statistical significance. The same is true for the CAPM and multi-factor alphas. Interestingly, the forecast power of yields found for the semi-monthly returns is even stronger

[6]To minimize loss of data, we use overlapping series of multi-period returns. As a consequence, the innovations in the forecast regressions will be autocorrelated. We use the method of Newey and West (1987) to correct the covariance matrix of the innovations for heteroskedasticity and autocorrelation.

for returns over longer horizons. A re-examination of yield-based predictability for monthly returns as in Table 5.5 shows that nearly all joint tests for zero slopes result in p-values smaller than 5 percent, with most being well below 1 percent. On average, we find that the number of predictable markets is increased by half. The results for the semi-quarterly and bi-monthly returns are only slightly weaker but still considerably stronger than for the semi-monthly returns. As for the multi-period returns on the yield-based trading strategies, we find similar averages, standard deviations, and alphas compared to the one-period case. Only the strategies using semi-quarterly returns seem to produce even stronger statistical significance.

Multi-period versions of the forecast regressions and the trading strategies based on hedging pressure and momentum also confirm the qualitative results found for the one-period returns. The forecast power of hedging pressure is found to be consistent over all holding periods, with the exception of the one-month horizon, in which case predictability is somewhat stronger. Predictability from past returns—the momentum effect—is about equally strong for all horizons. Finally, the trading strategies which aim to exploit predictability from hedging pressure or momentum produce similar results for all horizons.

5.5.3 Transaction costs

As a final robustness test, we investigate whether the active trading strategies which we found to outperform the benchmark portfolios, still yield abnormal returns after correcting for transaction costs. Active trading, contrary to passive trading, involves regular updating of long and short positions, and such updating is costly. These transaction costs, which comprise brokerage commissions, exchange and clearing fees, taxes, the bid-ask spread, etc., vary by type of trader, type of transaction, type of market, as well as through time. Hence, it is not easy to estimate transaction costs and incorporate them in the returns on trading strategies. Instead, we compute for each active strategy a critical transaction cost, defined as the average transaction cost per contract, expressed as a percentage of the futures price, for which the (abnormal) return on the strategy is just significantly different from zero at a given confidence level.[7] Thus, the critical transaction cost is the maximum transaction cost for which the strategy is still profitable.

The total transaction costs of an active trading strategy depend on the proportion of futures positions which need to be replaced with new positions each period. Panel A of Table 5.10 shows the average replacement rates for the long positions and the short positions of the yield-based strategies. The average replacement rate of 25 percent for the long positions of the semi-monthly updated short-yield strategy means that, on average, one in four long positions is substituted with a new long position every half month. Each substitution involves one "round turn," i.e., closing an existing long position by selling a contract, and taking a new long position by buying a new contract. Likewise, 23 percent of all short positions is replaced with new ones every half month. Average replacement rates for the semi-monthly updated yield-based strategies vary between 12 and 25 percent. More contracts need to be replaced as the return horizon increases.

Table 5.11 shows the critical transaction costs for the semi-monthly mean returns and alphas on the yield-based strategies corresponding to a 95 percent confidence level. Clearly, the short-yield strategies and the long-term spreading strategies permit the largest per-contract transaction costs for the strategies to remain profitable. For instance, the strategy based on the second nearby yields will still be profitable when transaction costs are under 54 basis points per trade. After correcting for market risk, the critical transaction costs are only slightly lower at 51 basis points, while correcting for multiple benchmarks reduces the critical cost to 33 basis points. More generally, critical transaction costs hardly change from average returns to CAPM alphas, but they go down quickly for the multi-factor alphas.

An individual trading small quantities is likely to pay more than these critical transaction costs in brokerage fees alone. Hence, the abnormal returns on yield-based trading may vanish. Large traders, on the other hand, may be able to mitigate this cost; however, it may be difficult to trade large quantities at once without moving the price a few basis points. Hence, the outperformance of benchmark

[7] Note that for our simple, equally-weighted trading strategies, the costs of updating futures positions only involves the costs of replacing one position by an other, and not the additional costs of changing portfolio weights. On the other hand, we abstract from the trading costs resulting from rolling over contracts when they approach maturity (or when the order of maturity changes).

portfolios by yield-based trading strategies with a semi-monthly return horizon may disappear once we include transaction costs, whether they be due to commissions and fees or market impact.

The critical transaction costs found for the other semi-monthly updated spreading strategies are even lower, and they go down as maturity decreases. The average returns on the strategies using the shortest maturities are far too small to tolerate transaction costs. Again, critical transaction costs are about the same for the average returns as for CAPM alphas, and considerably lower for the multi-factor alphas.

We also computed the critical transaction costs for the yield-based strategies with longer return periods. Again, we find the same pattern: low critical costs for the short-term spreading returns, and higher critical costs for the short-term returns and the long-term spreading returns. While critical transaction costs differ across return horizons, they seldom exceed one hundred basis points. The yield-based strategies which use semi-quarterly returns allow for the highest transaction costs, even though average replacement rates are relatively high. This is explained by the fact that these are also the strategies which produce the strongest evidence for outperformance of benchmark returns, as noted before.

Panel B of Table 5.10 shows the average replacement rates for the trading strategies based on past hedging pressure and on momentum. Average replacement rates for the hedging-pressure strategies are of the same order of magnitude as for the yield-based strategies, showing the same pattern of rates going up with the return horizon. The momentum strategies, on the other hand, require considerably higher replacement rates (implying larger total transaction costs), and they remain constant across return periods. Critical transaction costs for these active trading strategies (not shown here) suggest that positive abnormal returns disappear for reasonable values of the transaction costs.

5.6 Conclusion

This paper has analyzed trading strategies which capture the various risk premiums that have been distinguished in futures markets. On the basis of a simple decomposition of futures returns, we showed that the return on a short-term futures contract measures the spot-futures premium, while spreading strategies isolate the term premiums. Using a broad cross-section of futures markets and delivery horizons, we examined the components of futures risk premiums by means of passive trading strategies and active trading strategies which intend to exploit the predictable variation in futures returns.

We find that passive strategies which capture the spot-futures premium do not yield abnormal returns, in contrast to passive spreading strategies which capture the term premiums. The term structure of futures yields has strong explanatory power for both spot and term premiums, which can be exploited using active trading strategies that go long in low-yield markets and short in high-yield markets. The profitability of these yield-based trading strategies is not due to systematic risk. However, transaction costs may eliminate these gains, in particular for the strategies which capture short-term premiums.

Furthermore, we find that spreading returns are predictable by net hedge demand observed in the past, which can be exploited by active trading, but only if transaction costs are relatively low. Finally, there is momentum in futures markets, but momentum strategies do not outperform benchmark portfolios.

Table 5.2: **Summary statistics for short and spreading returns**

Returns are calculated from semi-monthly data for the period January 1986 to December 2000. Average returns and standard deviations are annualized and in percentages. The short return is defined as the return on the nearest-to-maturity contact. The n-th spreading return is the return on a strategy which takes a long position in the n-th nearby contract and a short position in the nearest-to-maturity contract.

	Averages						Standard deviations					
	Short return $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$					Short return $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
		$n=2$	$n=3$	$n=4$	$n=5$	$n=6$		$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
Wheat	−5.4	2.7**	5.0**	5.1**	5.9**	9.5**	21.2	5.2	8.1	9.7	11.0	13.7
Corn	−8.4	1.4	3.1*	4.3**	6.6***	8.4***	21.7	4.7	7.2	8.6	10.0	12.2
Oats	−15.1**	3.6*	5.4*	7.7**	6.6	.	30.4	8.2	11.4	13.9	16.8	.
Soybeans	−2.3	−1.4*	−1.3	−1.0	−0.1	1.5	19.8	3.0	5.0	5.9	6.5	7.7
Soy oil	−9.2*	0.4	1.8***	3.1***	4.1***	5.0***	21.8	1.7	2.7	3.8	5.0	5.9
Soy meal	5.7	−3.4***	−5.4***	−5.8***	−6.1***	−5.5**	21.4	3.9	6.5	7.9	9.1	10.3
Live cattle	5.3*	−0.8	−3.0	−2.7	−3.9*	−3.8	12.5	5.7	7.3	8.2	8.6	9.2
Feeder cattle	3.9	−0.7	−0.6	−0.5	−0.3	−0.7	12.0	2.9	4.1	5.1	5.7	6.3
Live hogs	6.4	0.6	−2.1	−3.1	−4.5	−2.9	23.9	9.6	14.4	16.7	18.5	19.8
Crude oil	7.8	0.2	−1.0	−1.9	−2.6	−3.0	34.5	7.1	10.2	12.5	14.3	15.8
Heating oil	13.1	−8.1***	−9.3**	−8.7**	−8.6*	−9.2*	33.5	10.3	14.6	16.4	17.9	18.8
Gold	−6.2*	−0.1	0.0	−0.1	−0.1	−0.2	13.0	0.4	0.6	0.8	1.0	1.2
Silver	−8.8	0.4	0.4	1.0*	0.9	1.1	22.9	1.3	1.8	2.3	2.8	3.1
Platinum	1.9	0.4	−0.5	−1.0	.	.	19.7	2.1	3.0	3.2	.	.
Coffee	−8.4	−0.6	−1.5	−1.5	−0.3	−1.4	37.7	7.2	10.2	12.4	14.6	16.2
Sugar	1.8	2.6	1.2	0.4	0.8	−2.7	39.4	19.6	21.4	22.7	23.9	26.2
Eurodollar	0.5*	0.2**	0.4***	0.5***	0.5**	0.4**	1.1	0.4	0.6	0.7	0.8	0.8
Swiss franc	−0.1	−0.1	−0.2	.	.	.	11.9	0.3	0.8	.	.	.
British pound	2.6	−0.1	−0.2	.	.	.	10.0	0.4	0.8	.	.	.
Japanese yen	0.8	0.0	−0.1	0.1	.	.	12.4	0.3	0.6	1.0	.	.
Can. dollar	0.6	−0.1	0.0	0.0	−0.1	.	4.7	0.4	0.8	1.1	1.4	.
S&P 500	8.6**	0.0	0.0	−0.2	.	.	14.4	0.3	0.6	0.8	.	.
NYSE	7.8**	0.1	0.1	.	.	.	13.9	0.3	0.7	.	.	.

*/**/*** indicates significance at the 10/5/1 percent level. No result is reported if more than one-third of the data is missing.

Table 5.3: **Unconditional Jensen's alphas**

The unconditional Jensen's alpha in the Capital Asset Pricing Model is the intercept in a regression of the short return (or a spreading return) on the return of the market portfolio in excess of the risk-free rate. The market portfolio is measured by the MSCI U.S. equity index, and the risk-free asset is the one-month Eurodollar deposit. The multi-factor alphas are implied by a six-factor model including U.S. and non-U.S. equities, U.S. and non-U.S. government bonds, emerging market stocks, and the U.S. dollar. All alphas are annualized and in percentages.

	Capital Asset Pricing Model						Multi-factor model					
	Short return	Spreading returns $r_f^{(n)} - r_f^{(1)}$					Short return	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
	$r_f^{(1)}$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$r_f^{(1)}$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
Wheat	−5.7	2.8**	5.2***	5.2**	6.1**	9.7**	−8.5	2.8**	5.6***	6.1**	7.2**	11.6***
Corn	−8.1	1.3	3.2*	4.2**	6.4**	8.0***	−6.5	1.6	3.5*	4.4**	6.2**	8.3**
Oats	−15.0**	3.3	5.1*	7.4**	7.0	.	−20.3***	3.3*	6.6**	9.4***	11.2**	.
Soybeans	−2.1	−1.1	−0.8	−0.7	0.2	1.6	−3.3	−0.9	−0.4	−0.4	0.3	1.9
Soy oil	−8.8	0.4	1.7**	3.0***	3.9***	4.8***	−8.9	0.4	1.6**	2.9***	3.8***	4.8***
Soy meal	6.3	−3.2***	−5.1***	−5.5***	−6.0**	−5.5**	3.8	−2.3**	−3.7**	−3.9**	−4.5**	−4.0
Live cattle	4.5	−1.0	−2.8	−2.4	−3.6*	−3.6	3.3	−0.2	−1.8	−1.2	−2.1	−2.2
Feeder cattle	3.4	−0.7	−0.6	−0.5	−0.2	−0.6	3.5	−1.1	−1.5	−1.1	−0.7	−1.1
Live hogs	6.5	1.0	−1.3	−2.3	−3.7	−2.3	3.8	1.3	0.8	−0.3	−2.0	0.3
Crude oil	9.9	0.2	−1.2	−2.3	−3.1	−3.6	15.1*	0.3	−1.1	−2.3	−3.3	−4.0
Heating oil	14.4*	−7.7***	−8.9**	−8.4**	−8.4*	−9.1*	19.1**	−7.0**	−8.4**	−7.9*	−8.2*	−8.9*
Gold	−4.9	−0.1	−0.1	−0.1	−0.1	−0.2	−6.8*	−0.1	0.0	0.0	0.0	−0.1
Silver	−8.9	0.5	0.4	1.1*	1.1	1.2	−8.9	0.5	0.5	1.3*	1.3*	1.6*
Platinum	1.1	0.4	−0.4	−0.8	.	.	2.3	0.3	−0.7	−1.3	.	.
Coffee	−7.4	−0.8	−2.0	−2.1	−1.0	−1.7	−2.7	−0.7	−2.2	−2.7	−2.0	−2.6
Sugar	1.8	2.7	1.7	1.0	1.2	−1.9	4.9	−0.3	−1.2	−2.7	−2.3	−7.1
Eurodollar	0.4	0.2*	0.3**	0.4**	0.4*	0.3	0.1	0.1	0.3**	0.3**	0.3*	0.2
Swiss franc	1.1	−0.1	−0.2	.	.	.	−2.1*	−0.1	−0.1	.	.	.
British pound	3.1	−0.1	−0.1	.	.	.	1.2	−0.1	−0.1	.	.	.
Japanese yen	1.2	0.0	−0.1	0.1	.	.	−1.2	0.0	0.1	0.2	.	.
Canadian dollar	0.2	−0.1	0.0	0.0	−0.1	.	−0.7	−0.1	−0.1	0.0	−0.1	.
S&P 500	−0.2	0.0	−0.1	−0.2	.	.	−0.3	0.1**	0.1	0.0	.	.
NYSE	−0.5	0.0	0.1	.	.	.	−0.3	0.1**	0.3**	.	.	.

*/**/*** indicates significance at the 10/5/1 percent level. No result is reported if more than one-third of the data is missing.

Table 5.4: **Summary statistics for futures yields**

The yield on the m-th nearby futures contract is defined as the difference between the log price of the m-th nearby futures contract and the log spot price divided by the estimated time to maturity. Settlement is assumed to take place on the 15-th of the delivery month on average. Yields are calculated from semi-monthly data for the period January 1986 to December 2000. Averages and standard deviations are annualized and in percentages.

	Averages						Standard deviations					
	Short yield	2nd nearby yield	3rd nearby yield	4th nearby yield	5th nearby yield	6th nearby yield	Short yield	2nd nearby yield	3rd nearby yield	4th nearby yield	5th nearby yield	6th nearby yield
Wheat	24.3	9.0	5.2	4.1	4.0	3.7	13.5	6.1	4.8	3.9	3.1	2.8
Corn	44.2	20.3	15.6	12.6	10.5	9.1	10.6	5.1	4.0	3.4	2.9	2.6
Oats	−117.8	−28.8	−13.9	−8.1	−4.5	.	22.2	4.9	3.4	3.0	2.6	.
Soybeans	21.0	9.2	6.8	5.6	4.9	4.4	5.3	3.1	2.6	2.1	1.8	1.5
Soy oil	14.3	9.7	8.4	7.4	6.6	6.0	8.2	3.7	2.7	2.3	2.1	2.0
Soy meal	12.4	3.5	2.1	1.9	1.9	1.9	14.3	7.2	5.2	4.1	3.5	3.0
Live cattle	11.9	0.9	−0.7	−1.1	−1.2	−1.0	9.5	4.7	3.2	2.4	1.9	1.5
Feed. cattle	−117.0	−49.1	−31.6	−22.2	−16.9	−14.4	17.4	6.0	3.4	2.4	2.0	1.7
Live hogs	102.5	41.5	26.9	18.8	15.0	11.9	39.1	17.1	12.1	9.7	7.9	6.5
Crude oil	0.1	−3.3	−4.4	−4.8	−4.9	−4.8	1.6	2.9	3.2	3.1	3.0	2.9
Heating oil	−7.9	−8.4	−6.6	−5.6	−5.2	−4.9	11.4	9.6	7.9	6.7	5.8	5.1
Gold	3.1	4.3	4.5	4.6	4.6	4.7	1.7	0.6	0.4	0.4	0.3	0.3
Silver	−6.4	1.7	3.8	4.7	5.1	5.3	3.6	1.4	0.9	0.6	0.6	0.5
Platinum	0.5	0.8	1.2	1.9	.	.	2.0	1.0	0.8	0.5	.	.
Coffee	43.7	14.5	10.4	8.6	7.6	6.9	26.8	8.9	6.3	5.2	4.4	3.9
Sugar	−22.0	−5.1	−3.6	−2.8	−2.1	−1.6	14.7	4.7	3.3	2.8	2.4	2.3
Eurodollar	−0.1	−0.3	−0.4	−0.5	−0.6	−0.6	0.3	0.3	0.2	0.2	0.2	0.1
Swiss franc	1.9	1.9	1.9	.	.	.	0.8	0.6	0.5	.	.	.
Br. pound	−2.3	−2.3	−2.1	.	.	.	0.6	0.5	0.4	.	.	.
Jap. yen	3.0	3.1	3.2	3.4	.	.	0.8	0.5	0.5	0.5	.	.
Can. dollar	−1.1	−1.0	−1.0	−1.0	−0.9	.	0.5	0.4	0.3	0.3	0.3	.
S&P 500	3.0	3.2	3.2	3.3	.	.	0.7	0.3	0.3	0.3	.	.
NYSE	2.7	2.9	3.0	2.9	.	.	0.7	0.3	0.3	0.3	.	.

No result is reported if more than one-third of the data is missing.

Table 5.5: **Yield-based forecast regression scoreboard**

For each futures market, semi-monthly short and spreading returns are regressed on the short yield. The analysis is repeated using yields of other maturities, i.e, the yield on the second nearby contract, the yield on the third nearby contract, et cetera. If, for a particular regression, more than one-third of the sample days has missing observations, the market is excluded from the analysis. Panel A gives the p-value for a test that the slope coefficients are equal to zero for all markets. Panel B shows the number of markets with slope coefficients which differ significantly from zero at the 10 percent level. The sign (+/−) indicates whether these coefficients are positive or negative. Between parentheses is the total number of analyzed markets, i.e., markets with sufficient data.

	Short return $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
		$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
A. Test: all slope coefficients zero (p-value)						
Short yield	0.000	0.012	0.006	0.001	0.263	0.005
2nd yield	0.000	0.124	0.147	0.005	0.485	0.004
3rd yield	0.000	0.457	0.600	0.054	0.750	0.054
4th yield	0.000	0.011	0.140	0.151	0.671	0.135
5th yield	0.001	0.015	0.249	0.418	0.646	0.088
6th yield	0.127	0.051	0.102	0.320	0.357	0.155
B. Number of predictable markets and sign of predictability (total number of markets)						
Short yield	8−(23)	4+(23)	6+(23)	3+(21)	2+(17)	2+(15)
2nd yield	6−(23)	3+(23)	3+(23)	2+(21)	2+(17)	3+(15)
3rd yield	6−(23)	2+(23)	2+(23)	0 (21)	0 (17)	0 (15)
4th yield	6−(21)	1+(21)	1+(21)	0 (20)	0 (17)	0 (15)
5th yield	2−(17)	2[a](17)	2[a](17)	1[b]−(17)	0 (17)	1[b]−(15)
6th yield	1−(15)	2[a](15)	2[a](15)	1[b]−(15)	0 (15)	1[b]−(15)

[a]One negative sign (silver) and one positive sign (soybean meal).

[b]Silver.

Table 5.6: **Yield-based trading strategies**

At each date, futures markets are sorted on the short yield into three groups of about the same size. Averages, standard deviations, and alphas (all annualized and in percentages) are reported for the short returns on trading strategies which take long positions in the low-yield group and as many short positions in the high-yield group. The analysis is repeated for yields of other maturities, as well as for the (term-)spreading returns on trading strategies which go long in high-yield markets and short in low-yield markets.

	Short returns $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
		$n=2$	$n=3$	$n=4$	$n=5$	$n=6$
	Averages					
Short yield	7.4**	1.9**	3.3***	3.6**	4.9**	7.5***
2nd nearby yield	10.4***	1.9**	3.5***	4.0***	5.5***	8.4***
3rd nearby yield	11.0***	2.1**	3.2***	3.8**	5.2***	8.0***
4th nearby yield	11.7***	2.3**	3.1**	3.6**	5.2***	7.8***
5th nearby yield	11.8***	2.2*	3.3**	3.7**	4.4**	6.5**
6th nearby yield	12.9***	2.3*	3.1*	3.8*	3.7*	4.9*
	Standard deviations					
Short yield	12.0	3.8	5.0	6.1	7.5	9.4
2nd nearby yield	12.4	3.7	5.0	6.0	7.3	9.1
3rd nearby yield	12.9	3.7	4.9	6.0	7.3	9.2
4th nearby yield	14.0	4.2	5.5	6.4	7.8	9.3
5th nearby yield	16.7	4.8	6.3	7.2	8.3	10.1
6th nearby yield	18.5	5.4	6.9	7.9	8.9	10.3
	CAPM alphas					
Short yield	7.5**	1.9**	3.3***	3.6**	4.7**	7.4***
2nd nearby yield	10.3***	1.9**	3.5***	4.0***	5.3***	8.3***
3rd nearby yield	10.8***	2.1**	3.2***	3.8**	5.1***	7.9***
4th nearby yield	11.6***	2.3**	3.0**	3.5**	5.2***	7.7***
5th nearby yield	11.6***	2.2*	3.3**	3.7**	4.4**	6.4**
6th nearby yield	13.1***	2.3*	3.1*	3.8*	3.8*	4.9*
	Multi-factor alphas					
Short yield	5.2	1.8*	2.8**	2.6	3.3	6.4**
2nd nearby yield	9.0***	1.8*	3.0**	3.3**	4.3**	7.7***
3rd nearby yield	10.0***	1.9*	2.7**	3.2**	4.0**	7.1***
4th nearby yield	10.3***	2.0*	2.5*	3.0*	4.0*	7.0***
5th nearby yield	9.9**	1.7	2.3	2.6	3.0	5.7**
6th nearby yield	11.6**	1.6	1.9	2.6	2.4	3.5

*/**/*** indicates significance at the 10/5/1 percent level.

Table 5.7: **Summary Statistics for Hedging Pressures**

The hedging pressure variable is defined as the number of short hedge positions minus the number of long hedge positions divided by the total number of hedge positions. Hedging pressures are calculated from semimonthly data for the period January 1986 to December 2000. Averages and standard deviations are in percentages. ρ_τ denotes the autocorrelation at lag τ.

	Avg.	Std.	ρ_1	ρ_2	ρ_3	ρ_4
Wheat	17.1	20.3	0.78	0.63	0.54	0.44
Corn	−0.9	15.6	0.88	0.76	0.67	0.58
Oats	38.4	15.9	0.87	0.71	0.56	0.44
Soybeans	17.0	20.6	0.89	0.82	0.78	0.71
Soy oil	12.7	19.8	0.84	0.72	0.62	0.55
Soy meal	13.2	15.5	0.79	0.64	0.56	0.50
Live cattle	14.6	17.4	0.94	0.88	0.82	0.77
Feeder cattle	−11.5	24.3	0.91	0.78	0.66	0.55
Live hogs	2.3	24.0	0.79	0.61	0.51	0.45
Crude oil	0.2	6.8	0.79	0.64	0.49	0.42
Heating oil	9.1	9.4	0.79	0.56	0.39	0.26
Gold	−3.1	21.7	0.80	0.65	0.53	0.46
Silver	43.0	15.9	0.87	0.76	0.69	0.66
Platinum	35.7	23.7	0.84	0.68	0.57	0.51
Coffee	17.9	14.5	0.75	0.53	0.36	0.28
Sugar	20.0	19.8	0.87	0.74	0.62	0.51
Eurodollar	−3.2	5.4	0.92	0.86	0.80	0.75
Swiss franc	−7.8	43.5	0.74	0.49	0.35	0.22
British pound	0.5	41.3	0.66	0.37	0.26	0.13
Japanese yen	−10.1	37.3	0.79	0.62	0.52	0.46
Canadian dollar	14.2	39.2	0.75	0.57	0.44	0.36
S&P 500	−5.1	6.7	0.84	0.73	0.63	0.56
NYSE	−14.4	45.0	0.76	0.65	0.55	0.46

Table 5.8: **Hedging pressure-based forecast regression scoreboard and trading strategies**

Caption as in Tables 5.5 and 5.6.

	Short return $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
		$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
A. Forecast regression scoreboard						
Test: all zero (p-value)	0.066	0.000	0.000	0.000	0.000	0.000
Predictable markets (+)	1	8	8	8	6	5
Predictable markets (−)	4	1	0	1	0	0
Number of markets	(23)	(23)	(23)	(20)	(17)	(15)
B. Trading strategies						
Average	4.2	2.1***	3.0***	3.8***	4.9***	5.7**
Standard deviation	12.6	3.1	3.9	5.2	7.0	10.9
CAPM alpha	3.5	2.1***	3.0***	3.8***	4.9***	5.7**
Multi-factor alpha	4.7	1.9**	3.0***	3.6**	4.5**	5.8*

*/**/*** indicates significance at the 10/5/1 percent level.

Table 5.9: **Momentum-based forecast regression scoreboard and trading strategies**

Caption as in Tables 5.5 and 5.6.

	Short return $r_f^{(1)}$	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
		$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
A. Forecast regression scoreboard						
Test: all zero (p-value)	0.211	0.008	0.000	0.003	0.032	0.000
Predictable markets (+)	1	3	2	0	1	1
Predictable markets (−)	3	2	6	4	3	2
Number of markets	(23)	(23)	(23)	(19)	(17)	(15)
B. Trading strategies						
Average	2.7	−0.3	−0.5	0.2	1.0	1.7
Standard deviation	15.0	3.6	4.8	6.5	8.4	10.3
CAPM alpha	2.9	−0.3	−0.4	0.3	1.3	2.0
Multi-factor alpha	1.8	−0.8	0.0	1.0	1.3	2.9

*/**/*** indicates significance at the 10/5/1 percent level.

Table 5.10: **Average replacement rates**

The average replacement rate is the average proportion of long or short contracts replaced by new contracts every period (semi-monthly, monthly, etc.) for a given trading strategy.

A. Yield-based trading strategies

	Semi-monthly		Monthly		Semi-quarterly		Bi-monthly	
	long	short	long	short	long	short	long	short
Short	25%	23%	26%	26%	31%	28%	31%	31%
2nd	15%	17%	18%	21%	22%	25%	25%	27%
3rd	12%	16%	15%	19%	19%	23%	21%	25%
4th	13%	25%	18%	36%	22%	33%	26%	42%
5th	15%	21%	20%	26%	23%	30%	27%	31%
6th	17%	21%	22%	29%	26%	33%	31%	34%

B. Trading strategies based on past hedging pressure

Semi-monthly		Monthly		Semi-quarterly		Bi-monthly	
long	short	long	short	long	short	long	short
19%	19%	26%	26%	31%	33%	31%	37%

C. Momentum strategies

Semi-monthly		Monthly		Semi-quarterly		Bi-monthly	
long	short	long	short	long	short	long	short
64%	63%	64%	64%	62%	63%	64%	65%

Table 5.11: **Critical transaction costs for the yield-based strategies**

The critical transaction cost is the average replacement cost for which the hypothesis that the mean return or alpha on an active trading strategy is zero is just not rejected at the 5 percent level. The table displays critical transaction costs for the yield-based strategies with a semi-monthly return horizon. Transaction costs are measured in basis points of the futures price.

	Critical transaction costs					
	Short return	Spreading returns $r_f^{(n)} - r_f^{(1)}$				
	$r_f^{(1)}$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$
Yield-based trading strategies						
	Averages					
Short yield	13	1	7	5	10	25
2nd nearby yield	54	1	14	14	24	50
3rd nearby yield	69	3	12	12	24	52
4th nearby yield	52	2	4	4	14	34
5th nearby yield	41	.	3	1	4	18
6th nearby yield	41	.	.	.	.	.
	CAPM alphas					
Short yield	11	1	7	5	8	24
2nd nearby yield	51	1	14	13	22	49
3rd nearby yield	64	4	12	12	22	50
4th nearby yield	49	2	4	4	14	34
5th nearby yield	39	.	2	1	3	17
6th nearby yield	42	.	.	.	.	.
	Multi-factor alphas					
Short yield	.	.	1	.	.	12
2nd nearby yield	33	.	5	1	6	36
3rd nearby yield	47	.	1	0	2	33
4th nearby yield	33	.	.	.	.	22
5th nearby yield	14	.	.	.	.	2
6th nearby yield	20	.	.	.	.	.

Bibliography

Ait-Sahalia, Y. and Lo, A. W.: 1998, Nonparametric estimation of state-price densities implicit in financial asset prices, *Journal of Finance* **53**(2), 499–547.

Anderson, R. W. and Danthine, J.-P.: 1980, Hedging and joint production: Theory and illustrations, *Journal of Finance* **35**(5), 487–498.

Anderson, R. W. and Danthine, J.-P.: 1981, Cross hedging, *Journal of Political Economy* **89**(6), 1182–1196.

Arrow, K. J.: 1970, *Essays in the Theory of Risk-Bearing*, North-Holland Press, Amsterdam.

Bailey, W. and Chan, K. C.: 1993, Macroeconomic influences and the variability of the commodity futures basis, *Journal of Finance* **48**(2), 555–573.

Balduzzi, P. and Kallal, H.: 1997, Risk premia and variance bounds, *Journal of Finance* **52**(5), 1913–1949.

Balduzzi, P. and Robotti, C.: 2001, Minimum-variance kernels, economic risk premia, and tests of multi-beta models, *Technical report*, Federal Reserve Bank of Atlanta, WP 2001-24.

Bessembinder, H.: 1992, Systematic risk, hedging pressure, and risk premiums in futures markets, *Review of Financial Studies* **5**(4), 637–667.

Bessembinder, H. and Chan, K.: 1992, Time-varying risk premia and forecastable returns in futures markets, *Journal of Financial Economics* **32**, 169–193.

Bessembinder, H., Coughenour, J. F., Seguin, P. J. and Smoller, M. M.: 1995, Mean reversion in equilibrium asset prices: Evidence from the futures term structure, *Journal of Finance* **50**(1), 361–375.

Black, F. and Scholes, M. S.: 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* **81**, 637–654.

Bollerslev, T.: 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* **31**, 307–327.

Boyer, B. H., Gibson, M. S. and Loretan, M.: 1999, Pitfalls in tests for changes in correlations, *Technical report*, International Finance Discussion Papers, No. 597, Board of Governors of the Federal Reserve System.

Breeden, D. T., Gibbons, M. R. and Litzenberger, R. H.: 1989, Empirical tests of the consumption-oriented capm, *Journal of Finance* **44**(2), 231–261.

Brennan, M.: 1958, The supply of storage, *American Economic Review* **48**, 50–72.

Burmeister, E. and McElroy, M. B.: 1988, Joint estimation of factor sensitivities and risk premia for the arbitrage pricing theory, *Journal of Finance* **43**(3), 721–733.

Campbell, J. Y.: 1996, Understanding risk and return, *Journal of Political Economy* **104**(2), 298–345.

Carhart, M. M.: 1997, On persistence in mutual fund performance, *Journal of Finance* **52**(1), 57–82.

Carter, C. A., Rausser, G. C. and Schmitz, A.: 1983, Efficient asset portfolios and the theory of normal backwardation, *Journal of Political Economy* **91**(2), 319–331.

Chang, E. C.: 1985, Return to speculators and the theory of normal backwardation, *Journal of Finance* **40**(1), 193–208.

Chen, N.-F., Roll, R. and Ross, S. A.: 1986, Economic forces and the stock market, *Journal of Business* **59**(3), 383–403.

Chen, Z. and Knez, P. J.: 1996, Portfolio performance measurement: Theory and applications, *Review of Financial Studies* **9**(2), 511–555.

Cherubini, U. and Luciano, E.: 2002, Multivariate option pricing with copulas, *ICER Applied Mathematics Working Paper Series, WP no. 5* .

Cochrane, J. H.: 2001, *Asset Pricing*, Princeton University Press.

Cox, J. C., Ross, S. A. and Rubinstein, M.: 1979, Option pricing: A simplified approach, *Journal of Financial Economics* **7**, 229–263.

DeRoon, F. A., Nijman, T. E. and Veld, C. H.: 1998, Pricing term structure risk in futures markets, *Journal of Financial and Quantitative Analysis* **33**(1), 139–157.

DeRoon, F. A., Nijman, T. E. and Veld, C. H.: 2000, Hedging pressure effects in futures markets, *Journal of Finance* **55**(3), 1437–1456.

DeRoon, F. A., Nijman, T. E. and Werker, B. J. M.: 2003, Currency hedging for international stock portfolios: The usefulness of mean-variance analysis, *Journal of Banking and Finance* **27**, 327–349.

Dixit, A. K. and Pindyck, R. S.: 1994, *Investment under uncertainty*, Princeton University Press, Princeton, New Jersey.

Dixit, A. K. and Pindyck, R. S.: 1995, The options approach to capital investment, *Harvard Business Review* pp. 105–115.

Duan, J.-C.: 1995, The garch option pricing model, *Mathematical Finance* **5**(1), 13–32.

Duffie, D.: 1989, *Futures Markets*, Englewood Cliffs, New Jersey: Prentice-Hall.

Dusak, K.: 1973, Futures trading and investor returns: An investigation of commodity market risk premiums, *Journal of Political Economy* **81**(6), 1387–1406.

Embrechts, P., McNeil, A. J. and Straumann, D.: 2002, Correlation and dependence in risk management: properties and pitfalls, *in* M. A. H. Dempster (ed.), *Risk Management: Value at Risk and Beyond*, Cambridge University Press, Cambridge, England, pp. 176–223.

Engle, R. F. and Kroner, K. F.: 1995, Multivariate simultaneous generalized arch, *Econometric Theory* pp. 122–150.

Fama, E. F.: 1984, Forward and spot exchange rates, *Journal of Monetary Economics* **14**, 319–338.

Fama, E. F. and French, K. R.: 1987, Commodity futures prices: Some evidence on forecast power, premiums, and the theory of storage, *Journal of Business* **60**(1), 55–73.

Fama, E. F. and French, K. R.: 1992, The cross-section of expected stock returns, *Journal of Finance* **47**(2), 427–465.

Fama, E. F. and French, K. R.: 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* **33**(1), 3–56.

Fama, E. F. and French, K. R.: 1995, Size and book-to-market factors in earnings and returns, *Journal of Finance* **50**(1), 131–155.

Ferson, W. E. and Harvey, C. R.: 1991, The variation of economic risk premiums, *Journal of Political Economy* **99**(2), 385–415.

Fung, W. and Hsieh, D. A.: 1997, Survivorship bias and investment style in the returns of ctas, *Journal of Portfolio Management* **24**(1), 30–41.

Fung, W. and Hsieh, D. A.: 2000, Performance characteristics of hedge funds and commodity funds: Natural vs. spurious biases, *Journal of Financial and Quantitative Analysis* **35**(3), 291–307.

Genest, C.: 1987, Frank's family of bivariate distributions, *Biometrika* **74**, 549–555.

Genest, C. and MacKay, R. J.: 1986, Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données, *The Canadian Journal of Statistics* **14**(2), 145–159.

Ghoudi, K., Khoudraji, A. and Rivest, L.-P.: 1998, Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles, *The Canadian Journal of Statistics* **26**(1), 187–197.

Glosten, L. R. and Jagannathan, R.: 1994, A contingent claim approach to performance evaluation, *Journal of Empirical Finance* **1**, 133–160.

Hansen, L. P.: 1982, Large sample properties of generalized method of moment estimators, *Econometrica* **50**(4), 1029–1054.

Hansen, L. P. and Jagannathan, R.: 1991, Implications of security market data for models of dynamic economies, *Journal of Political Economy* **99**(2), 225–262.

Hicks, J. R.: 1939, *Value and Capital*, Oxford University Press.

Hirshleifer, D.: 1989, Determinants of hedging and risk premia in commodity futures markets, *Journal of Financial and Quantitative Analysis* **24**(3), 313–331.

Jagannathan, R.: 1985, An investigation of commodity futures prices using the consumption-based intertemporal capital asset pricing model, *Journal of Finance* **40**(1), 175–191.

Jegadeesh, N.: 1990, Evidence of predictable behavior of security returns, *Journal of Finance* **45**(3), 881–898.

Jegadeesh, N. and Titman, S.: 1993, Return to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* **48**(1), 65–91.

Jegadeesh, N. and Titman, S.: 2001, Profitability of momentum strategies: An evaluation of alternative explanations, *Journal of Finance* **56**(2), 699–720.

Johnson, H.: 1987, Options on the maximum or the minimum of several assets, *Journal of Financial and Quantitative Analysis* **22**(3), 277–283.

Keynes, J. M.: 1930, *A Treatise on Money, Vol.II*, McMillan, London.

Lehmann, B. N.: 1990, Fads, martingales, and market efficiency, *Quarterly Journal of Economics* **105**(1), 1–28.

Magill, M. and Quinzii, M.: 1995, *Theory of Incomplete Markets*, MIT Press, Cambridge, Massachusetts.

Margrabe, W.: 1978, The value of an option to exchange one asset for another, *Journal of Finance* **33**(1), 177–186.

Mas-Colell, A., Whinston, M. D. and Green, J. R.: 1995, *Microeconomic Theory*, Oxford University Press, Oxford, New York, New York.

Mayers, D.: 1972, Non-marketable assets and capital market equilibrium under uncertainty, *in* M. C. Jensen (ed.), *Studies in the theory of capital markets*, Praeger Publishers, New York, pp. 223–247.

McDonald, R. L. and Siegel, D. R.: 1985, Investment and the valuation of firms when there is an option to shut down, *International Economic Review* **26**(2), 331–349.

McDonald, R. L. and Siegel, D. R.: 1986, The value of waiting to invest, *Quarterly Journal of Economics* **101**(4), 707–728.

Nelsen, R. B.: 1999, *An Introduction to Copulas*, Lecture Notes in Statistics 139, Springer-Verlag, New York.

Newey, W. K. and West, K. D.: 1987, A simple, positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**, 703–708.

Patton, A. J.: 2002a, Modelling time-varying exchange rate dependence using the conditional copula, *Working Paper* .

Patton, A. J.: 2002b, Skewness, asymmetric dependence, and portfolios, *Working Paper* .

Reiner, E.: 1992, Quanto mechanics, *From Black-Scholes to Black Holes: New Frontiers in Options*, RISK Books, London, pp. 147–154.

Rosenberg, J. V.: 1998, Pricing multivariate contingent claims using estimated risk-neutral density functions, *Journal of International Money and Finance* **17**, 229–247.

Rosenberg, J. V.: 1999, Semiparametric pricing of multivariate contingent claims, *Stern School of Business, Working Paper S-99-35* .

Rosenberg, J. V.: 2003, Nonparametric pricing of multivariate contingent claims, *Journal of Derivatives* **10**(3).

Shimko, D. C.: 1994, Options on futures spreads: Hedging, speculation, and valuation, *Journal of Futures Markets* **14**, 183–213.

Sklar, A.: 1959, Fonctions de répartition à n dimensions et leurs marges, *Publ. Inst. Statist. Univ. Paris* **8**, 229–231.

Stoll, H. R.: 1979, Commodity futures and spot price determination and hedging in capital market equilibrium, *Journal of Financial and Quantitative Analysis* **14**(4), 873–894.

Stulz, R. M.: 1982, Options on the minimum or the maximum of two risky assets: Analysis and applications, *Journal of Financial Economics* **10**, 161–185.

Van der Vaart, A. W. and Wellner, J. A.: 1996, *Weak Convergence and Empirical Processes*, Springer-Verlag, New York.

Vassalou, M.: 2003, News related to future gdp growth as a risk factor in equity returns, *Journal of Financial Economics* **68**, 47–73.

White, H. L.: 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* **48**, 817–838.

Working, H.: 1949, The theory of the price of storage, *American Economic Review* **12**, 1254–1262.

Samenvatting (Summary)

Dit proefschrift beslaat, zoals de titel reeds suggereert, een breed scala aan onderwerpen binnen het vakgebied der financiële economie. Het behelst een verzameling studies over investeringsbeslissingen en waarderingsvraagstukken op diverse deelterreinen van de financiële economie, uiteenlopend van de reële optietheorie en de analyse van macro-economische risico's, tot termijnmarktmodellen en de waardering van multivariate financiële derivaten.

In hoofdstuk 2 van het proefschrift wordt de theorie van dynamisch investeringsgedrag onder onzekerheid bestudeerd. Deze theorie, die ook wel de *reële optietheorie* wordt genoemd, stelt dat het van waarde kan zijn een investering met een positieve netto contante waarde, maar met onzekere opbrengsten, uit te stellen in situaties waarin de investering niet kosteloos ongedaan kan worden gemaakt. Tot dusverre is deze theorie alleen ontwikkeld voor de bijzondere omstandigheid dat de investeerder risiconeutraal is, en voor het speciale geval dat alle onzekere toekomstige opbrengsten van de investering exact kunnen worden *gehedged* (nagebootst) op de ter beschikking staande vermogensmarkten. In dit hoofdstuk wordt onderzocht hoe de optimale investeringsbeslissing wordt beïnvloed wanneer niet aan deze veronderstellingen is voldaan.

Met gebruikmaking van een model waarin ook met risico-aversie rekening wordt gehouden, wordt het bestaan van een drempelwaarde voor de investeringsopbrengsten aangetoond. Evenals in de reeds ontwikkelde theorie, komt de optimale investeringsbeslissing erop neer dat men dient te investeren wanneer de huidige waarde van de opbrengsten de drempelwaarde overstijgt, en anders voorlopig dient af te zien van de investering. Met behulp van een nutsfunctie die gekenmerkt wordt door afnemende risico-aversie bij grotere bedragen, wordt vervolgens onderzocht wat de invloed is van veranderingen in risico-aversie, de onzekerheid van de opbrengsten en andere parameters van het model op de optimale investeringsbeslissing.

In termen van de comparatieve statica van het model, zijn de belangrijkste bevindingen dat een hogere mate van risico-aversie leidt tot een hogere investeringsdrempel, vooral als de omvang van de investering aanzienlijk is. Ook een toename van de onzekerheid van de investeringsopbrengsten heeft tot gevolg dat het aantrekkelijker wordt de investering uit te stellen. Dit effect blijkt sterker te zijn naarmate de risico-aversie van de investeerder groter is.

Verder worden in hoofdstuk 2 niet eerder gepubliceerde analytische formules afgeleid voor de comparatieve statica van het model met risiconeutraliteit. Een vermeldenswaardige conclusie is dat een renteverlaging voor de risiconeutrale investeerder altijd leidt tot een lagere investeringsdrempel, omdat het positieve effect van een renteverlaging op investeringen (ten gevolge van een hogere contante waarde van de opbrengsten) altijd groter is dan het tegengestelde effect dat veroorzaakt wordt door een hogere optiewaarde van uitstel (de mogelijkheid toekomstige verliezen te verminderen of te vermijden).

Hoofdstuk 3 biedt een analyse van zogeheten *economic hedging portfolios*. Dit zijn beleggingsportefeuilles die investeerders aanhouden om macro-economische risico's waaraan zij zijn blootgesteld af te dekken. Zo kan bijvoorbeeld onvoorziene inflatie de koopkracht van een belegger in gevaar brengen. Door een hedgeportefeuille aan te houden waarvan het verwachte rendement zo veel mogelijk samenhangt met de beweging van de inflatie, kan dit inflatierisico verminderd worden. Zo'n hedgeportefeuille bestaat bijvoorbeeld uit beleggingen in aandelen en obligaties. In dit hoofdstuk wordt aangetoond hoe de optimale samenstelling van een hedgeportefeuille verkregen kan worden. De optimale portefeuillegewichten blijken een functie van de mate van risico-aversie van de investeerder. Bovendien is de hedge preciezer bij tegenvallende aandelen- en obligatiekoersen dan in betere tijden. Deze benadering vormt daarmee een uitbreiding van de optimale hedgeportefeuille die volgt uit het gangbare *mean-variance* model.

Een empirische analyse van optimale hedgeportefeuilles ter bescherming tegen inflatie-, rente-, krediet- en andere macro-economische onzekerheden voor Amerikaanse beleggers met toegang tot de aandelen- en obligatiemarkten laat zien dat nutsmaximaliserende beleggers die blootgesteld zijn aan macro-economische risico's bereid zijn significante compensaties te betalen voor het afdekken van die risico's. Afhankelijk van de mate van risico-aversie van de belegger verandert de samenstelling van de hedgeportefeuille, waarmee de eerder beschreven theoretische samenhang empirisch wordt ondersteund. Verder kan hedgen tegen macro-economische risico's, zo blijkt uit dit onderzoek, wellicht ook een verklaring bieden voor het feit dat feitelijk waargenomen rendementen niet stroken met het Capital Asset Pricing Model (CAPM), dat voorspelt dat beleggers alleen compensatie kunnen verwachten voor marktrisico en derhalve kunnen volstaan met beleggingen in een marktindex.

In hoofdstuk 4 wordt de prijsbepaling van multivariate opties onder de loep genomen. Multivariate opties zijn financiële derivaten waarvan de uitbetaling een functie is van twee of meer onderliggende activa. Gewoonlijk betreft het call- of put-opties die recht geven op de toekomstige aan- respectievelijk verkoop, tegen een vooraf vastgestelde prijs, van het onderliggende activum met het hoogste (of juist laagste) gerealiseerde rendement. Voor de premie van een dergelijk derivaat is de samenhang tussen de onderliggende activa van belang, die in dit hoofdstuk modelmatig beschreven wordt door middel van parametrische copula's. Dit zijn functies die de afhankelijkheid tussen stochastische variabelen in kaart brengen en een alternatief bieden voor de veelal veronderstelde Gaussische of normale afhankelijkheidsstructuur.

In tegenstelling tot andere onderzoeken op dit terrein wordt in deze studie niet uitgegaan van een constante afhankelijkheidsstructuur, maar is een in de tijd veranderende samenhang ook toegelaten. Daarbij wordt aangenomen dat de mate van samenhang tussen de onderliggende activa verband houdt met de volatiliteit in de afzonderlijke markten, hetgeen aansluit bij het verschijnsel dat in roerige tijden hogere correlaties tussen de rendementen op financiële activa worden gemeten dan in rustiger tijden. Deze dynamische copulamodellen zijn toegepast op call- en put-opties op twee Amerikaanse beursindices, de S&P 500 en de Nasdaq. De bevindingen van dit onderzoek zijn dat de optiepremie die volgt uit dynamische copulamodellen sterk afwijkt van de optiepremie die berekend wordt op basis van modellen die een constante afhankelijkheid veronderstellen, met name in geval van grote fluctuaties op de onderliggende markten. Verder blijkt dat de in ogenschouw genomen niet-

Gaussische copula's significant verschillende premies opleveren ten opzichte van de Gaussische copula, ongeacht de volatiliteit op de onderliggende markten, terwijl er geen noemenswaardig verschil is tussen de optiepremies die volgen uit de niet-Gaussische copula's.

Onderwerp van hoofdstuk 5 zijn de risicopremies op termijnmarkten. Het verwachte rendement op termijncontracten (*futures*) kan ontleed worden in een *spot premium* en een *term premium*. De spot premium meet het verschil tussen het verwachte eenperioderendement op de onderliggende waarde en de kortetermijnyield. De kortetermijnyield is het procentuele prijsverschil tussen het kortetermijncontract en de onderliggende waarde. De *term premium* is de wig tussen de som van verwachte toekomstige kortetermijnyields en de langetermijnyield. Aangetoond kan worden dat de spot premium gelijk is aan het verwachte rendement op een long-positie in een kortetermijncontract, terwijl de term premium gelijk is aan het verwachte rendement op een *spreadingstrategie*, waarbij een long-positie in een langetermijncontract gecombineerd wordt met een short-positie in een kortetermijncontract. Op basis van deze decompositie worden in dit hoofdstuk de genoemde premiecomponenten geschat, en wordt de voorspelbaarheid van de bijbehorende rendementen onderzocht om vervolgens de winstgevendheid van zogeheten *sortingstrategieën* te analyseren.

De empirische analyse in dit hoofdstuk concentreert zich op een groot aantal Amerikaanse futuresmarkten voor goederen en financiële waarden. Gemiddeld genomen blijken de rendementen op vrijwel alle kortetermijnfutures nihil te bedragen, wat duidt op verwaarloosbare onconditionele spot premiums. Echter, de gemiddelde rendementen op de eerdergenoemde spreadingstrategie, en daarmee de schattingen van de onconditionele term premiums, blijken voor veel markten wel van nul af te wijken, zelfs na correcties voor systematisch risico.

De voorspelbaarheid van de kortetermijnrendementen en de rendementen op de spreadingstrategie is onderzocht aan de hand van drie variabelen: de termijnstructuur van futuresyields, de zogenoemde *hedging pressure* (d.i. het saldo van short- en long-posities van hedgers in een futuresmarkt) en het momentumeffect. Yields hebben een dusdanig sterke verklarende waarde voor zowel spot als term premiums, dat men met dynamisch handelen door steeds long te gaan in de markten met de laagste yields en short in de markten met de hoogste yields, de rendementen op benchmarkportefeuilles significant overtreft. Een dergelijke sortingstrategie gebaseerd op hedging pressure levert eveneens hogere rendementen op dan die men op basis van het CAPM of een meerfactorenmodel zou verwachten. Vergelijkbare momentumstrategieën blijken echter niet winstgevender dan deze benchmarks. Alleen wanneer de transactiekosten, die met deze dynamische strategieën gepaard gaan, aanzienlijk zijn verdwijnen de gevonden anomalieën.