2. Risk Management Concepts

2.1 Profit Functions and Expected Utility Maximization

The profit function is an essential tool in the analysis of derivative security trading strategies, for both risk management and speculative purposes. The procedure for specifying the profit function proceeds by, first, writing down the relevant transactions in a trading schematic. The basic profit function is then specified from the schematic. For simple trading strategies, such as a naked speculation, the profit function and schematic are not too useful, as the basic insights can be obtained without much analysis. However, for more complicated trades, the profit function can be an invaluable aid. Once the basic profit function is specified, it is possible to do manipulations and substitutions that can be used to identify relevant features of the trading strategy. One important substitution that is often used is to replace the deferred contract prices with the cash and carry arbitrage conditions.

To complete the illustration of what determines the profitability of long and short positions, let Q be the number of units of the commodity purchased. At $t=0$, the two parties to the futures contract for delivery of Q units at $t=T$ agree to a price of $F(0,T)$. Consider what happens if $F(1,T) > F(0,T)$, i.e., that futures prices for the commodity rise. The short position who agreed to sell Q units at $F(0,T)$ now is faced with a situation where the value of the commodity to be delivered is $Q\ F(1,T)$ versus $Q\ F(0,T)$ the previous period. Pretending for the moment that $t=1$ was actually the delivery date T, then the short would have to cover by going into the cash market, purchasing the appropriate number of units of the commodity and delivering. This would require a larger outlay of funds, $Q\ F(1,1) = Q\ S(1)$, than would be received from the sale of the commodity through the futures markets, $Q\ F(0,T)$. The opposite type of example would hold for the long position. Hence, it can be concluded that: *short positions benefit from price falls while long positions benefit from price rises.*

While it is straight forward to illustrate the conditions under which long and short positions are profitable, the analysis is decidedly more complex when more involved positions are being considered. For this reason, it is instructive to illustrate how to convert the previous discussion into algebraic terms. Figure 2.1 demonstrates that the profit function for the long futures position is: $\pi(1,T) = Q\ \{F(1,T) - F(0,T)\}$. It follows that $\pi > 0$ when futures prices rise between $t=0$ and $t=1$. A similar profit function can be defined for ***short*** positions:

$$\pi(1,T) = Q\ \{F(0,T) - F(1,T)\}$$

In this case, $\pi > 0$ occurs when prices fall.

It is useful to recognize that the form of the profit function will depend on how the units, Q, are specified. Some presentations will vary the sign of Q such that when a long position is indicated then $Q > 0$ and, for a short position, $Q < 0$. In this case, what is referred to above as the long profit function will apply in both cases. If prices fall, and $Q < 0$ for a short, then profit will be positive. While this approach is somewhat tidier to use in presenting basic concepts, in the analysis of more complicated trading strategies this convention will lead to variations in the signs of some terms when compared to derivations based on $Q > 0$ everywhere. Assuming $Q > 0$ throughout facilitates use of the rule for calculating profit: "what you sold it for minus what you bought it for". While in many cases the difference is between using $Q > 0$ and Q of varying sign is transparent, there are instances where some care is required when comparing results given using the different specifications of Q.

Figure 2.1 **Profit Function for a Long Futures Position**

DATE	***Cash Position***	***Futures Position***
t=0	None	Long *Q* units @ *F(0,T)*
t=1	None	Close out position by going short *Q* units @ *F(1,T)*

The profit function, $\pi(1,T)$, can now be defined by observing that change in value of the futures position is calculated by subtracting the purchase price from the sales price:

$$\pi(1,T) = Q \{F(1,T) - F(0,T)\}$$

The position is profitable, $\pi > 0$, when prices rise.

The speculative profit function for a long or short position is relatively simple when compared to the profit function for a hedger, where account has to be taken of both the futures and cash positions. In addition, details for the hedge will vary depending on the specific hedge under consideration. To illustrate the hedger's profit function, consider the hedging problem confronting the stylized grain elevator operator of the 19th century. Grain would be hauled by land to the riverside where the grain elevator was situated. The elevator operator would pay the farmer the prevailing cash price and, in the case at hand, would store the grain through the winter until the river thaw in the spring. The elevator operator was concerned that grain prices would move adversely between the harvest and springtime. To offset this risk the farmer would engage in a futures hedge with the traders at Chicago Board of Trade. The relevant transactions are described in Figure 2.2. This profit function applies only to the hedge transaction, it does not take account of other profits associated with the actual business. For example, there may be spoilage or loss or grain to vermin such that $Q_A(0) \neq Q_A(1)$, or the grain may be processed and sold in a

Figure 2.2 **Profit Function for a Grain Elevator Hedge using Futures Contracts**

DATE	***Cash Position***	***Futures Position***
$t=0$	Buy Q_A units of grain at $S(0)$ for storage in grain elevator	Short Q_H units at $F(0,T)$
$t=1$	Q_A units are sold at S(1) and loaded for shipment	Close out position with Long Q_H units at $F(1,T)$

If costs associated with carrying the commodity are ignored, the profit function for this type of hedge can be specified:

$$\pi(1,T) = \{S(1) - S(0)\}\ Q_A + \{F(0,T) - F(1,T)\}\ Q_H$$

different form.

The intricacies of hedging can be illustrated by extending the grain elevator hedge example in Figure 2.2. The futures position in the example implicitly assumes that the elevator operator has grain for sale that is not deliverable against the futures position because it does not conform to the standarized grade. The possibility that the relationship between prices for different grades will change over the life of the hedge is a source of risk in the hedge. If the elevator operator has a deliverable grade of grain then the futures hedge can be completed by making delivery. In this case, the futures hedge profit function takes the form of a profit function for a forward sale, a hedge that is done using forward contracts, where the profit function is:

$$\pi(1) = Q_A\,\{F(0,1) - S(0)\}$$

In this case, the costs of financing, storing the commodity and making the delivery at maturity are ignored.

The profit function for the grain elevator hedge using futures explicitly recognizes that the size of the hedge position in futures may be different than the size of the cash position. As discussed in chapter 6, the precise relationship between the size of the cash and futures positions can be formulated as an optimization problem from which an ***optimal hedge ratio*** can be determined. However, it is still revealing to *assume that the hedge is one-to-one*, i.e., let $Q_A = Q_H = Q$, which gives after some manipulation:

$$\pi/Q = \{F(0,T) - S(0)\} - \{F(1,T) - S(1)\}$$

The profitability of the hedged position depends on the change in the difference between the spot and futures prices. If this difference narrows, the hedge will have $\pi > 0$.

The analytical usefulness of the profit function is also apparent when the futures spread trades are considered. For example, the basic *intra-commodity spread* trade, also called a *calendar spread*, involves offsetting, long and short, positions in different contract delivery months for the same commodity. If the spread is in different commodities, an *inter-commodity* spread, the delivery months involved less importance. Intra-commodity trades can be done for different reasons that will be discussed in later chapters. Recalling the use of N and T for the

Figure 2.3 **Profit Function for an Intra commodity Futures Spread Position**

Date	***Nearby Position***	***Deferred Position***
$t=0$	Short Q_N units at $F(0,N)$	Long Q_T units at $F(0,T)$
$t=1$	Close out position with Long Q_N units at $F(1,N)$	Close out position with Short Q_T units at $F(1,T)$

In this case, the profit function can be specified:

$$\pi(1,T) = \{F(0,N) - F(1,N)\}\, Q_N + \{F(1,T) - F(0,T)\}\, Q_T$$

nearby and deferred deliveries of amount Q_N and Q_T, this trade is depicted in Figure 2.3. One immediate interpretation of spread behavior from this profit function is to *assume that the spread is one-to-one and intra-commodity*, i.e., let $Q_N = Q_T = Q$, which gives after some manipulation:

$$\pi/Q = \{F(1,T) - F(1,N)\} - \{F(0,T) - F(0,N)\}$$

The one-to-one intracommodity spread that is short the nearby and long the deferred will be profitable if the difference between the deferred and nearby prices widens. The opposite would be true for the alternative spread, long the nearby and short the deferred.[1]

The Wealth Process

In order to obtain applicability to a range of decision-making situations, the approach taken is to specify the wealth process, admitting the possibility of two random variables, both price and yield uncertainty at the decision horizon date. The representative decision maker purchases an asset at time t and sells it at time $t+1$, and purchases derivative securities to provide protection against downside movements, either in price or yield or both. The price and the yield at $t+1$, the end of the investment horizon, can both be unknown at time t, the date the relevant risk management decision is initiated. In some types of decision problems, such as the typical problem of investment in domestic assets, this level of generality is more than is required because there is only one random variable in this problem, the yield on domestic assets. However, where the problem involves investment in foreign assets, there are two random variables involved, the exchange rate and the yield on foreign assets (denominated in foreign currency terms). In other problems, such as a farmer subject to random crop yield or a mine subject to random ore quality, both price and quantity are uncertain. Given that price and yield can be uncertain, the optimization problem does not permit the amount of initial wealth to invest in the asset to vary.[2] Starting from the given initial level of wealth, the investor's objective is to maximize a moment preference function for the value of terminal wealth assuming that the balance (possibly negative) of initial wealth that is not allocated to the risky asset will earn (pay) the risk free rate of interest.

Initially, consider the wealth process for a decision maker not having access to any derivative securities. Once the initial structure of the terminal wealth function is specified, usage of derivative securities will be introduced. Following Poitras (1993) and others, allowing for ***both*** the quantity and the price to be random leads to the underlying wealth process:

$$W_{t+1} = A\, Y_{t+1} P_{t+1} + [W_t - C(A)]\,(1+r)$$

where: W_{t+1} is wealth at time $t+1$ and W_t is the known level of initial wealth; A is the fixed initial size of the asset, e.g., acres planted for a farmer; Y_{t+1} is the possibly random quantity per unit or yield per unit of the asset observed at $t+1$; P_{t+1} is the random spot price at $t+1$; $C(A)$ is the given cost function associated with purchasing A; r is the risk-free interest rate.[3] Manipulation of the underlying wealth process gives the more conventional form of the wealth process for a single risky asset:

$$W_{t+1} = W_t\,(x(1+R) + (1-x)(1+r)) = W_t\,((1+r) + x(R-r)) = W_t + \pi_{t+1}$$

where: π_{t+1} is the profit defined by the wealth process realised at time $t+1$, x is $(C(A)/W_t)$ the given fraction of initial wealth invested in the risky asset, and $(1+R)$ is $[(A\, Y_{t+1}P_{t+1})/C(A)]$ one plus the rate of return on the risky asset. For simplicity of exposition, it will be assumed that $x > 0$ in what follows.[4]

The basic specification for the decision maker's terminal wealth function requires some additional terms if there is a need to capture the payoffs associated with, say, introducing a put option. While the terminal wealth function derived from the wealth process, with put options included, follows appropriately, some motivation is required. In particular, in the absence of some form of put option to provide asset insurance, there is a natural minimum on R, the rate of return on the investment. Either a complete catastrophic loss occurs where $Y_{t+1}=0$, or a spot price of zero occurs at time $t+1$, both cases corresponding to the result $(1+R)=0$. Significantly, three possible variants of put option pay-out are possible, each aimed at dealing with the different types of risks faced by the risk manager. More precisely, put option pay outs can depend on the deviation of ***price***, ***yield*** or ***revenue*** from a stated exercise value. Pay-outs based on revenue provide protection against $Y_{t+1}P_{t+1}$ falling below a given floor. In contrast, pay outs based on yield or price cannot guarantee a minimum return higher than $(1+R) = 0$. For the farmer example of Section 6.1, while put pay-outs based on revenue set a lower level for farm income, put option pay-outs guaranteeing a price of $\$K$ per bushel cannot prevent a 100% loss due to crop failure, nor can a put payout based on yield providing for, say, $\underline{Y}$ bushels an acre prevent the future spot price falling to zero. However, put pay outs based on either price and yield do reduce the probability of the total return attaining low values and, as a result, do alter the distribution for terminal wealth.

In practice, conventional exchange traded put options are structured with pay outs based on price. Other types of put options, such as multiple peril crop insurance schemes, are a type of yield insurance. Still other types of put options, such as some types of real options, provide revenue or income protection. The case where the put payout is based on revenue insurance produces a wealth process similar to the yield insurance case. Introduction of a put option based on price produces:

$$W^z_{t+1} = AY_{t+1}P_{t+1} + (W_t - C(A))(1+r) + Q_z(\max[0,\ K - P_{t+1}] - z)$$

$$= W_t\,\{x(1+R) + (1-x)(1+r) + \frac{Q_z P_t}{W_t}\,[\max[0,\ \frac{K - P_{t+1}}{P_t}] - \frac{z}{P_t}]\}$$

$$= W_t\,\{(1+r) + x(R-r) + \gamma(\max[0,\ -R_P] - \frac{z}{P_t})\}$$

where K is exercise price on the put option *that is assumed to be "at the money" (where $K = P_t$)*, z is the price per unit of output of the put, Q_z is the number (in output units) of puts purchased, with the ratio γ being the asset value covered by the option position divided by initial wealth.[5]

This specification can be contrasted with that for put option payouts based on yield where, instead of the number of options to purchase, it is the fraction of A to insure that is the decision variable:

$$W^{y}_{t+1} = AY_{t+1}P_{t+1} + (W_t - C(A))(1+r) + Q_y\,(P_{t+1}\ \max[0,\ \underline{Y} - Y_{t+1}] - L)$$

where L is the price (put premium) per unit of A for the yield put option, Q_y is the number of units covered by the yield put option and $\underline{Y}$ is the yield floor provided by the put option or insurance plan. Defining the optimization problem by allowing the risk manager to choose the fraction of A to insure leads to:

$$W^{y}_{t+1} = W_t\{(1+r) + x[R-r] + x\lambda[\max[0,\underline{RR}-R] - l]\}$$

where l equals $(LA/C(A))$, $\lambda = (Q_y/A)$ is the fraction of A, e.g., the total planted acreage, covered or insured with the physical yield put option and $\underline{RR} = \{P_{t+1}\ \underline{Y}\ A\}/C(A)$.[6] Assuming actuarially fair pricing requires insurance to impact the decision problem through its effect on downside risk and skewness.

This basic structure can be readily adjusted to account for other derivatives, such as futures or forward contracts. For example, if it is assumed that the only hedging instrument available is futures contracts then the underlying wealth dynamics can be specified:

$$W_{t+1} = A\ Y_{t+1}P_{t+1} + [W_t - C(A)]\ (1+r) + Q_f\,(f_{t+1} - f_t)$$

where: Q_f is the quantity of futures contracts sold (-) or bought (+); and f_{t+1} and f_t are the futures prices observed at $t+1$ and t respectively. Manipulation gives:

$$\begin{aligned} W_{t+1} &= W_t\,(x(1+R) + (1-x)(1+r) + HR_f) \\ &= W_t\,((1+r) + x(R-r) + HR_f) \\ &= W_t + \pi_{t+1} \end{aligned}$$

where: H is the value (f_t times Q_f) of the hedge position divided by initial wealth (*not the value of the spot position*), R_f is $(f_{t+1} - f_t)/f_t$ and $(1+R)$ is, again, $[(A\ Y_{t+1}P_{t+1})/C(A)]$ one plus the rate of return on planting for a farmer.

The Expected Utility Function

The study of decision making under uncertainty is a vast subject. Financial applications almost invariably proceed under the guise of the expected utility hypothesis: people rank random prospects according to the expected utility of those prospects. Analytically, this involves solving problems requiring selecting choice variables to maximize an expected utility function. In some cases, such as the basic optimal hedging problem, the associated budget constraint is embedded in the argument of the utility function. In other cases, such as in optimal portfolio diversification models, the budget constraint appears as the restriction that the sum of the value weights equals one. In either event, the central concern is expected utility, an essentially subjective construct that cannot be directly observed. A key step in the optimization problem is to specify an expected utility function that captures the true expected utility mapping.

The central tool in analyzing preferences over random outcomes is the expected utility function. Expected utility calculations involve taking expectations, which are conventionally modelled using statistical properties of random variables. This may involve the explicit introduction of probability densities. There is a profound connection between the choice of a specific probability distribution and the risk aversion properties required of the expected utility function, e.g., Heaney and Poitras (1994). As for the utility component of the expected utility function, even before von Neumann and Morgenstern (1947), it has been recognized that choosing over risky prospects is decidedly different than the textbook model of economic choice. As is well known, von Neumann and Morgenstern made a seminal contribution by proposing a set of axioms governing choice under uncertainty. Observing that the axioms are difficult to reject lends strong support to the von Neumann and Morgenstern approach.[7]

A key construct of the axiomatic approach is the linear choice function over risky prospects, better known as the expected utility function:

$$EU[x] = \sum_{j=1}^{S} \theta_j \ U_j[x]$$

where: $EU[x]$ is the expected utility of x; S is the number of possible futures states of the world; θ_j is the probability that state j will occur; and, $U[x_j]$ is the utility associated with the amount of x received in state j. The EU function ranks risky prospects with an ordering that is unique up to a linear transformation. While there are a number of possible selections for x, in what follows either terminal wealth or terminal profit will typically be used.

Beyond this foundation, developing various arguments can be tricky. For example, it is not apparent how to determine the θ_j. The axiomatic foundation cannot say much more than that the probabilities are subjective. General equilibrium models often proceed by assuming that expectations are homogeneous or that individual agents are homogeneous. Such assumptions permit the derivation of market equilibrium conditions, such as the CAPM. However, general equilibrium concerns are of little use here. The decision problems encountered are partial equilibrium. The theoretical results apply to speculators and hedgers confronted with a parametric world of atomistic competition where their activities will not impact prices. In this process, the expected utility function can be an invaluable analytical tools. This can be readily demonstrated by applying an essential tool from functional analysis: the Taylor series expansion (see Appendix I).

To see this, consider the problem of determining the cost of risk. The solution to this problem would be useful in analysing whether to buy insurance or to invest in a risky capital project. While there are a number of possible methods to extract the cost of risk, consider the following solution. Let the expected value of terminal wealth be: $E[W_{t+1}] = \Omega$. Observe that Ω is a parameter that permits the *certainty equivalent* income of a risky prospect to be defined as Ω - C, where C is the cost of risk. It follows from the expected utility axioms that the cost of risk, C, can be calculated as the difference between the expected value of the risky prospect and the associated certainty equivalent income:

$$U[\Omega - C] = \sum_{i=1}^{S} \theta_i \ U[W_i] = EU[W_{t+1}]$$

It is now possible to expand $U[\Omega - C]$ in a Taylor series and estimate the cost of risk by manipulating the first and second order approximations.

More precisely, expanding the function $U\Omega - C]$ around Ω the first order approximation is:

$$U[\Omega - C] = U[\Omega] + U'[\Omega] \ (\Omega - C - \Omega) = U[\Omega] - U'[\Omega] \ C$$

Similarly, a second order approximation for the function $U[W_{t+1}]$ can provide:

$$U[W_{t+1}] = U[\Omega] - U'[\Omega] \ (W_{t+1} - \Omega) + \frac{1}{2} U''[\Omega] \ (W_{t+1} - \Omega)^2$$

$$\rightarrow \quad EU[W_{t+1}] = U[\Omega] + \frac{1}{2} \ U''[\Omega] \ var[W_{t+1}]$$

Using $U[\Omega - C] = EU[W_{t+1}]$ and manipulating gives:

$$C = -\frac{U''[\Omega]}{2 \ U'[\Omega]} \ var[W_{t+1}] \quad \rightarrow \quad \frac{C}{W_t} = -\frac{U''[\Omega] \ W_t}{U'[\Omega]} \ var[1 + R]$$

This demonstrates theoretically that the cost of risk will vary across utility functions. This result also provides theoretical measures of the cost of risk. The measures of absolute risk aversion, $-\{U''/U'\}$, and relative risk aversion, $-\{U'' \ W_t\}/U'$ are now textbook concepts, e.g., Elton and Gruber (1995).

Expected Utility and Moment Preference

The relationship between moment preference and expected utility has received considerable attention. Important topics have included: the conditions under which mean-variance analysis is consistent with maximizing expected utility, e.g., Kroll, Levy and Markowitz (1984), Ormiston and Quiggin (1994), Bell (1995); and, the implications of introducing skewness preference into the mean-variance framework, e.g., Kraus and Litzenberger (1976), Hassett et al. (1985), Lim (1989), Diacogiannis (1994), Poitras and Heaney (1999). Brockett and Kahane (1992) among others, have shown that there is not a direct correspondence between the derivatives of the expected utility function and moments of the return distribution. The implication is that maximization of a function defined over moments, such as mean-variance or mean-variance-skewness, may not give the same solution as directly maximizing expected utility. Yet, Meyer (1987), Ormiston and Quiggin (1994) and others demonstrate that the conditions on the random variables sufficient for mean-variance rankings to provide solutions consistent with expected utility rankings are relatively weak. Extensions providing the conditions on random variables required for mean-variance-skewness ranking to be consistent with expected utility ranking are currently unavailable.

As discussed in numerous sources, e.g., Loistl (1976), Levy and Markowitz (1979), Poitras and Heaney (1999), the relationship between expected utility and moment preference objective functions can be motivated using a Taylor series expansion of $U[W]$, the decision maker's utility function (U) for wealth (W), evaluated at the expected value for terminal wealth $[\Omega]$ $(E[W_{t+1}] = \Omega)$:

$$U[W_{t+1}] = U[\Omega] + U'[\Omega](W_{t+1} - \Omega) + \frac{U''[\Omega]}{2!}(W_{t+1} - \Omega)^2 + \frac{U'''[\Omega]}{3!}(W_{t+1} - \Omega)^3 + ...$$

Exploiting this type of expansion requires certain technical conditions be satisfied. For example, convergence of the power series within the interval of interest is needed.[8] In addition, desirable properties for utility functions require: $U'[W] > 0$, non-satiation; $U''[W] < 0$, risk aversion; and, $U'''[W] > 0$, preference for positive skewness.

With relatively weak distributional restrictions, e.g., Hassett et al. (1985), the Taylor series representation of $U[W]$ can be transformed into an approximation for a general expected utility function based on the moments of the conditional distribution for W_{t+1}. The relevant approximation is derived by taking conditional expectations at time t and ignoring terms associated with moments higher than the second, for a mean-variance approximation, and moments higher than the third, for a mean-variance-skewness approximation. The general notation $EU[\cdot]$ will be used to denote such a moment preference functional. Taking expectations for the mean-variance-skewness case gives:

$$EU_{MVS}[W_{t+1}] \equiv EU_{MVS} = U[\Omega] + 0 + \frac{U''[\Omega]}{2!} var[W_{t+1}] + \frac{U'''[\Omega]}{3!} skew[W_{t+1}]$$

$$= U[\Omega] - b \ var[W_{t+1}] + c \ skew[W_{t+1}]$$

where $var[W_{t+1}]$ is the variance of terminal wealth, $skew[W_{t+1}]$ is the skewness or centralized third moment for terminal wealth. Restrictions imposed by assuming risk aversion and positive skewness preference permit the coefficients in EU_{MVS} to be immediately signed as $b,c > 0$. Further restrictions on b and c, as well as the admissible range of W, can be derived by taking further derivatives of the Taylor series expansion and invoking Jensen's inequality. Setting $c=0$ permits the mean-variance-skewness moment preference function to be reduced to the mean-variance function, EU_{MV}.

What are the implications of introducing this additional skewness term into the moment preference objective function? Currently, little information is available comparing solutions from mean-variance and mean-variance-skewness approximations. Information about such comparisons would be relevant for a range of decision making situations, especially those involving skewness altering securities such as options and insurance. The few studies that do compare the mean-variance and mean-variance-skewness objective functions illustrate some confusion as to the implications of introducing skewness. In particular, Prakash et al. (1996, p.240) claim to "show how a risk-averse manager with sufficient preference for positive skewness may undertake projects with skewed payoff distributions that appear to be unfair gambles." Horowitz (1998) correctly takes exception to the Prakash et al. claim, arguing that there is no underlying utility function that is consistent with the central theoretical condition that Prakash et al. use, i.e., $3U''/U''' > skew[W_1]/var[W_1]$. Horowitz refutes the Prakash et al. claim by demonstrating that it is not possible for an expected utility function to conform to the Prakash et al. restrictions.

Studies examining the impact of skewness have been largely concerned with asset pricing and portfolio theory, e.g., Kraus and Litzenberger (1976), Sears and Trennepohl (1983), Lim (1989), Simaan (1993). However, combining of securities into portfolios almost certainly reduces the skewness of the portfolio relative to the value weighted sum of the individual asset skewness values. The structure of the decision problem under consideration here is typically more stylized, being concerned only with transforming the return distribution for an exogenously determined amount of a single asset into a “more desirable” distribution using a derivative security. This type of problem is typical of many risk management situations, e.g., farming or mining, where the size of the spot position is predetermined by production considerations and the decision problem is to solve for the size of the hedge position. Allowing both the quantity of the risky asset and the size of the derivative security position to be endogenous substantively complicates the analysis without adding significantly to the usefulness of the solutions, e.g., Poitras (1993). More importantly, practical situations where derivative securities would be used to manage risk often involve having the level of the risky asset fixed prior to assessing the size of the derivative security position.

Finally, while one obvious potential benefit of introducing skewness preference into the objective function is enhanced ability to model certain types of decisions problems, this gain is not without some costs. Compared to the mean-variance approach, the introduction of skewness significantly increases complexity of the solutions, permitting only complicated preference dependent closed form solutions to be derived. This is due to the presence of quadratic terms in the first order conditions arising from $skew[W]$. Though intuitive results can still be obtained by fully solving for the mean-variance part of the solution, this leaves an additional unresolved coskewness term that is associated with the quadratic terms in the first order conditions. This unresolved term will be utility function dependent, as it will contain the parameters b and c. Hence, solving for the mean-variance-skewness optimal demand requires b/c to be specified before an optimal solution can be obtained. In this process, the mean-variance optimal solution acts a control variate against which the mean-variance-skewness optimal solution can be compared. Properties of this comparison can be developed in detail, e.g., Poitras and Heaney (1999)..

A Stylized Risk Management Decision Problem

Both firms and individuals face an array of risks that have to be managed. Some of these risks, such as those associated with loss from fire or theft, can be safely unbundled from decisions involving other types of risk. In formulating a general risk management strategy for business enterprises, it is conventional to ignore these unrelated risks and restrict attention to the following categories of risk (Dowd 1998): *general business risks*, that are risks specific to the industry or market of interest, e.g., yield uncertainty in farming, sales uncertainty for a retailer, production uncertainty for a mine, also referred to as *commercial risk*; financial and commodity *market risks* associated with changes in prices for equities, exchange rates, interest rates and commodities; *credit and liquidity risks*, associated with factors such as counter-party failure, costs associated with having to unwind a position and the possibility that credit lines may be restricted; *operational risks*, that can include inadequate management control systems or incorrect pricing models; and, *legal risks*, that are associated with contract enforcement and variation.

The presence of these various types of risks begs an obvious question: under what conditions can each of these risks be managed independently of the other types of risks? Some attention has been given to identifying *theoretical* conditions under which it is possible to separate the production decision from the risk management decision, e.g., Danthine (1978), Feder et al. (1981). If there is separability, this implies that it is possible to use a risk management process that considers the problem of managing, say, market risks independently of general business risks. A common theme running through various so-called derivatives debacles is the apparent inability to understand the speculative component of the risk management strategies that were being used. Under reasonable conditions, optimal risk management decisions can be decomposed into the sum of two parts: the solution to a risk minimizing problem; and, the solution to a speculative profit maximization problem. In effect, as demonstrated in various derivative debacles, implementing optimal risk management requires understanding of speculation.

Textbook presentations often portray risk management activities, such as hedging, as eliminating the risk of price fluctuations, leaving firm profit to be dependent solely on underlying productive activities. For example, by hedging the farmer is able to lock in the price that is received for the crop at harvest. This leaves profit to be determined by factors influencing the yield per acre. In practice, the hedging problem is much more complicated. Because hedge positions can lose money as well as make money, the hedging decision has speculative features. If the hedge loses money, then profits will be increased by not hedging. In the farmer example, this would occur when the price at harvest was higher than the price that was locked in using the hedge. If prices move adversely, then hedges will make money. Hence, unless the hedger is completely risk averse, the optimal hedge must take expected future prices into account. Is it optimal to hedge when prices are expected to move favourably?

In order to understand the optimal hedging problem, it is expedient to first consider the optimization problem for a speculator. Because speculators, by definition, do not have any cash market position, these traders are concerned solely with making profits from changing prices. Analytically, this problem can be structured as a question about ***optimizing*** some appropriately specified objective function. The convention in modern finance is to model the optimization problem using the maximization of expected utility. While a number of slightly different variables could be selected to determine expected utility, for present purposes an appropriately defined π will be used as the sole argument. Because it is difficult to interpret the optimal solutions when a general form is used for the expected utility function, this leaves the specific functional form to be selected. Following the convention in much of the modern finance, (e.g., Elton and Gruber 1995, Alexander and Sharpe 1990, Ingersoll 1987), the ***mean-variance expected utility*** function *(EU)* will be used:

$$EU[\pi] = E[\pi] - b\, var[\pi]$$

where $b\ (> 0)$ measures the sensitivity of expected utility to changes in risk. The optimal speculative position for this objective function can now identified.

To construct a mean-variance solution for the optimal speculative position, consider the profit function for a futures speculator who is either long the actual $(Q > 0)$ or short the actual $(Q < 0)$:

$$\pi(1) = Q\,\{F(1,T) - F(0,T)\}$$

The resulting expectation and variance of profit lead to the following (see Appendix 2):

$$\max_{Q} \quad Q\ \{E[F(1,T)] - F(0,T)\} \ - \ b\ Q^2\ \sigma_f^2$$

$$\frac{\partial EU}{\partial Q} = \{E[F(1,T)] - F(0,T)\} \ - \ 2b\ Q\sigma_f^2 \ = \ 0$$

$$\Rightarrow \qquad Q^* = \frac{E[F(1,T)] - F(0,T)}{2b\ \sigma_f^2}$$

where $b\ (> 0)$ is a parameter that measures the sensitivity of expected utility to changes in risk. The optimal speculative position size is seen to depend on three elements: the expected change in the futures price; the conditional variance of futures prices; and, the speculator's sensitivity to risk. It is instructive to consider the different solutions that are associated with varying these elements.

The solution depends on a combination of the trader's attributes: subjective probability assessments of the trader about future states of the world; the trader's degree of risk aversion; and, the trader's ability to forecast. Under analytically restrictive conditions (that possibly could be loosened), the solution to the speculative trader's optimization problem can be aggregated to get implicit indications about the nature of market equilibrium. Given this, if, in aggregate, speculators behave as though they were risk neutral, then the speculators' offer curve for Q in terms of $\{F(1,T) - F(0,T)\}$ would be (theoretically) infinite at $E[F(1,T)]=F(0,T)$. To see this, observe what conditions the numerator of the optimal speculative position must satisfy for Q to be finite (required for markets to clear) and b goes to zero. If this were the case, hedgers would not pay a risk premium to speculators in the form of a systematic bias in the forecasting accuracy of the futures price. This is because speculators are already willing to participate when the futures price is an unbiased forecast. Given that the futures markets are designed to facilitate the participation of a wide range of traders, it is possible that b may vary across market environments, from risk loving to risk neutral to risk averse. Analysis of this situation could be explored by appropriate differentiation of the optimal speculative solution.

Given the solution to the optimal speculative position, it is now possible to develop a solution to the optimal hedging problem. In most practical situations, the hedger is faced with the question of what ratio of cash to futures positions should be selected. This can be translated into questions about ***optimizing behavior***. As for the speculator, optimality has to be defined using the maximization of expected utility. This objective includes minimizing the variance (risk) of the hedged position as an important special case of the slightly more general *mean-variance expected utility* function *(EU)*: $EU[\pi] = a(E[\pi]) - c\, var[\pi]$ where π is defined for the appropriate hedger profit function and a and c (> 0) are appropriately defined parameters. If risk is taken to be variance, then the objective of minimizing risk can be reformulated in expected utility form as:

$$EU[\pi] = - var[\pi]$$

which is the general mean-variance objective function with $a = 0$ and $c = 1$. These variations of the mean-variance objective function can be used to address the issue of whether hedgers are minimizers of risk or maximizers of expected utility (or both).

Figure 2.4 **Stylized Short (Long) Hedge Profit Function**

DATE	***Cash Position***	***Futures Position***
$t=0$	Buy (Sell) Q_S at $S(0)$	Short (Long) Q_F at $F(0,T)$
$t=1$	Sell (Repurchase) Q_S at $S(1)$	Close out with Long (Short) at $F(1,T)$

This leads to the associated profit function for the short (long) hedger:

$$\pi(1) = Q_S \{S(1) - S(0)\} + Q_H \{F(0,T) - F(1,T)\}$$

$$(= Q_S \{S(0) - S(1)\} + Q_H \{F(1,T) - F(0,T)\})$$

The profit function can now be used to derive $var[\pi]$, that is the same for both the long and short profit functions:

$$var[\pi] = Q_S^2\, \sigma_S^2 + Q_H^2\, \sigma_f^2 - 2\, Q_S Q_H\, \sigma_{Sf}$$

where $\sigma_{Sf} = cov[S(1), F(1,T)]$, the conditional covariance of spot and futures prices with σ_f^2 and σ_S^2 being

Over time, considerable academic attention has been given to the solution of the fundamental question: what ratio of spot to futures positions is most appropriate to maximize the expected utility of end-of-period profit? (e.g., Johnson 1960, Ederington 1979, Hill and Schneeweis 1982, Stulz 1984, Toevs and Jacob 1986, Herbst, et al. 1989, Heaney and Poitras 1991). Much of this research has focused on estimating hedge ratios using an ordinary least squares *(OLS*) regression of spot prices on futures prices; the "optimal" hedge ratio being the estimated slope coefficient. This result can be derived from the stylized short (long) hedger trading profile in Figure 2.4, where short refers to the hedger's position in futures.

Given the variance of trade profit, the optimal hedge ratio follows by solving the first order conditions for max *EU* (with $EU = - var[\pi]$) using Q_H as the choice variable:

$$\frac{\partial EU}{\partial Q_H} = 2Q_H \, \sigma_f^2 - 2Q_S \, \sigma_{Sf} = 0 \qquad \Rightarrow \qquad \left(\frac{Q_H}{Q_S}\right)^* = \frac{\sigma_{Sf}}{\sigma_f^2} = \frac{\sigma_S}{\sigma_f} \, \rho_{Sf}$$

where ρ is the correlation coefficient between *S* and *F*. A number of observations can be made about this solution. Most importantly, it identifies the minimum variance hedge ratio as the *OLS* slope coefficient in a bivariate regression of spot on futures prices.[9] This is the operational result that makes the minimum variance hedge ratio empirically attractive and accounts for its widespread use among practitioners. However, despite the widespread popularity of the approach, there are significant analytical limitations on its use and unanswered questions about its validity. For example, one important limitation is the dependence of the optimal solution on *one choice variable*, the size of the futures position. The size of the cash position is taken as fixed and certain. No allowance is made for leveraging to purchase the spot commodity or for hedging situations where the size of the cash position is uncertain, e.g., the farmer who faces stochastic output. Before addressing these issues, it is important to address the unanswered questions about its validity, i.e., does $EU = -var$ correspond to optimal solutions for other, more theoretically plausible, expected utility functions?

To see this, consider the optimal hedge ratio that is associated with max $EU = E[\pi] - b\ var[\pi]$. Observing that for the short hedge $E[\pi] = Q_S\{E[S(1)] - S(0)\} + Q_H\{F(0,T) - E[F(1,T)]\}$, the following problem and solution can be posed:

$$\max_{Q_H} \quad EU[\pi] = E[\pi] - b \ var[\pi]$$

$$\frac{\partial EU}{\partial Q_H} = (F(0,T) - E[F(1,T)) - b \ (2Q_H \, \sigma_f^2 - 2\overline{Q_S} \, \sigma_{Sf}) = 0$$

$$\Rightarrow \qquad \frac{Q_H^*}{\overline{Q_S}} = \frac{\sigma_{Sf}}{\sigma_f^2} + \frac{F(0,T) - E[F(1,T)]}{2 \ b \ \overline{Q_S} \, \sigma_f^2}$$

The mean-variance solution is composed of two parts: the minimum variance hedge ratio and the optimal speculative position. While the minimum variance component depends on the ratio of statistical parameters, the speculative component depends on the hedger's risk attitudes as reflected in *b*. Hedgers who are "less risk averse" will have lower b (ceteris paribus) and, as a result, will be more willing to take speculative positions in the form of over or under hedges. In addition, because the futures price variance enters in the numerator of the "speculative" term, as the *perceived* volatility increases the hedger will be less willing to take positions over or under the minimum variance hedge. More precisely, variances as well as expectations are conditional on the information available on the hedge date. These values are derived from the subjective probability assessments of the hedger. Hence, the less capable or willing the hedger is to make forecasts, the less important is the speculative component of the hedge.

2.2 Measuring Risk and Exposure

Risk and Uncertainty

To be practical, the moment preference approach to modeling the optimal risk management problem requires values for the relevant statistical parameters. The optimal solution to the stylized risk management problem of Sec. 2.1 requires the variance and expected value for the distribution of futures prices as well as the covariance between the futures price and spot price. These parameters are all from the conditional distribution. Precisely how the conditional distribution is to be modeled raises deep philosophical questions, variants of which have been debated for centuries. For example, Thomas Bayes (1701-1761) suggested that the conditional (posterior) distribution is determined by combining prior beliefs with available empirical evidence. In the 20th century, both J.M. Keynes (1883-1946) and Frank Knight (1885-1972) advanced the notion that the variation in future outcomes is a combination of a measurable component, risk, and an unmeasurable component, uncertainty. At the time, this was an intellectual step forward, a reaction to the 19th century beliefs of Stanley Jevons, Francis Galton and others that future outcomes were ultimately measurable.

Knight and Keynes were both struggling with different facets of the impact of randomness on economic activity. When put within the context of the problems at hand, their seemingly arcane ideas still have considerable relevance. Knight worked within the traditional of classical economics, seeking to explain how economic profits can arise from uncertainty in the process of production and distribution. Classical economic theory depends on the assumption that outcomes are certain, if there is randomness then the probabilities of the possible outcomes are known with certainty. In the absence of market imperfections, such as monopoly, classical economic theory argues that economic profits will dissipate to zero and each of the factors of production will earn their value of marginal product. Knight questioned this view, arguing that economic profits could still arise from the ability of entrepreneurs to resolve the uncertainty facing factors of production.

Frank Knight still has relevance, not because of his theoretical musings, but because of his interpretation of the randomness arising from commercial risks. Part Three of *Risk, Uncertainty and Profit* (1921), especially the chapters on "The Meaning of Risk and Uncertainty" and "Structures and Methods for Meeting Uncertainty", contain many insights. For example, Knight discusses the application of "the principle of insurance" to "business hazards". After recognizing the wide divergence of insurable risks, from life to fire to marine to theft and burglary, Knight concludes (p.252): "The possibility of ... reducing uncertainty by transforming it into a measurable risk ... constitutes a strong incentive to extend the scale of operations of a business establishment. This fact must constitute one of the important causes of the phenomenal growth in the average size of industrial establishments which is a familiar characteristic of modern life". Knight also clearly recognizes "specialization" in activities that isolate the "true uncertainty" in business risk including "organized speculation as carried on in connection with produce and security exchanges" (p.257).

Perhaps the most important point involves Knight's interpretation of commercial risks, for example (p.226):

> A manufacturer is considering the advisability of making a large commitment in increasing the capacity of his works. He "figures" more or less on the proposition, taking account as well as possible of the various factors more or less susceptible of measurement, but the final result is an "estimate" of the probable outcome of any proposed course of action. What is the "probability" or error (strictly, of any assigned degree of error) in the judgment? It is manifestly meaningless to speak of either calculating such a probability *a priori* or of determining it empirically by studying a large number of instances. The essential and outstanding fact is that the "instance" in question is so entirely unique that there are no others or not a sufficient number to make it possible to tabulate enough like it to form a basis for any inference of value about any real probability in the case we are interested in.

Risk is associated with objectively measured probabilities, while uncertainty requires subjective probability assessments. The economic rents to business ownership arise from correctly anticipating uncertain outcomes. While recognizing that "real uncertainty" can be reduced through consolidation, Knight (p.234) recognizes "that it *is possible* does not necessarily mean that it *will be done*".

As for methods of dealing with uncertainty, Knight (p.239) recognizes four general approaches:

> We may call the two fundamental methods of dealing with uncertainty, based respectively upon reduction by grouping and upon selection of men to "bear" it, "consolidation" and "specialization", respectively. To these two methods we must add two others ... (3) control of the future, (4) increased power of prediction.

Knight recognizes the complementarity of the different approaches of dealing with uncertainty. For example, increased specialization permits more firm resources to be devoted to data collection and analysis that increases power of prediction. Writing in 1921, Knight has little to say about the use of derivative securities to "control the future". Other than occasional references, Knight also does not deal with specific aspects of financial risk and uncertainty. What Knight does says very clearly is that the randomness associated with economic risks, such as business risk, is composed of 'risk', that is measurable in an objective sense, and 'uncertainty', that is only measurable subjectively. It is in dealing correctly with uncertainty that "entrepreneurs" earn value. In terms of the stylized risk management solution given in Sec. 2.1, risk can be associated with variance minimization and uncertainty with speculation.

In contrast to Knight, Keynes provides little guidance on general methods of managing risks. Whereas Knight's *Risk, Uncertainty and Profit* wanders toward an endpoint, in *The General Theory of Employment, Interest and Money* (1936) Keynes proposes "not one, or two, but three or four 'models' of the workings of a modern economy" (Blaug 1978, p.682). Chapter 12 of *The General Theory* is a largely self-contained essay on "The State of Long Term Expectation". In this chapter, Keynes is concerned with the social consequences of instability in stock markets, arguing for government intervention to offset inherent deficiencies. The core of the argument revolves around an examination of the process by which expectations are formed in financial markets. Due to an excess bias towards maintaining liquidity, expectations in financial markets are focused on near-term prospects (p.157): "Investment based on genuine long-term expectation is so difficult today as to be scarcely practicable".

The General Theory is a difficult book to read, quite untidy and poorly written. The importance of the book lies in the substance of certain arguments, who was making those arguments and when the book was presented, i.e., during the stagnation following the economic collapse of the early 1930's. Many ideas are presented, some seemingly off-the-cuff. Such is the case with Chapter 12. Some of the observations are insightful, for example (p.154-5):

> It might be supposed that competition between expert professionals, possessing judgment and knowledge beyond that of the average private investor, would correct the vagaries of the ignorant individual left to himself. It happens, however, that the energies and skill of the professional investor and speculator are mainly occupied otherwise. For most of those persons are, in fact, largely concerned, not with making superior long-term forecasts of the probable yield of an investment over its whole life, but with forecasting changes in the conventional basis of valuation a short time ahead of the general public. They are concerned, not with what an investment is really worth to a man who buys it "for keeps", but with what the market will value it at, under the influence of mass psychology, three months or a year hence. Moreover, this behaviour is not the outcome of a wrong-headed propensity. It is an inevitable result of an investment market organized (to concentrate resources upon the holding of "liquid" securities). For it is not sensible to pay 25 for an investment of which you believe the prospective yield to justify a value of 30, if you also believe that the market will value it at 20 three months hence.

In true Keynesian fashion, this is shortly followed with the unerudite statement (p.155): "The social objective of skilled investment should be to defeat the dark forces of time and ignorance which envelop our future". The modern reader is left glancing about for a Wall Street investment banker dressed as Luke Skywalker or Hans Solo.

What Keynes develops in Chapter 12 is a model where the heterogenous, subjective expectations of market participants leads to a financial market equilibrium in which prices are "subject to waves of optimistic and pessimistic sentiment, which are unreasoning and yet in a sense legitimate where no solid basis exists for a reasonable calculation" (p.154). The implication is that prices can change "violently as the result of a sudden fluctuation of opinion due to factors which do not really make much difference to the prospective yield"(p.154). Not only will prices be considerably more volatile than is justified by the long term expectation, prices will typically depend more on "what average opinion expects average opinion to be" rather than on valuations that capture "the prospective yield of an investment over a long term of years" (p.155). Prices are determined more by "*speculation* ... the activity of forecasting the psychology of the market" than by "*enterprise* ... the activity of forecasting the prospective yield of assets over their whole life" (p.158).

Keynes was concerned about the potentially negative impact that price formation in capital markets can have on the macroeconomy: “When the capital development of a country becomes a by-product of the activities of a casino, the job is likely to be ill-done” (p.159). As evidenced by the technology stock price bubble of 1999-2000, such observations still have relevance. However, this is not a book on macroeconomics. While Keynes has little to say about risk management practices and the use of derivative securities, there is considerable insight about the stochastic properties of financial prices and the role of speculation in determining financial prices. Modern corporate risk management is largely concerned with managing commercial (enterprise) risks and financial risks to achieve the objective of shareholder wealth maximization. Keynes warns about the possibility of that financial prices may not reflect long term expectations of prospective yields and will likely be subject to inexplicable volatility. If so, this substantially complicates the problem of formulating optimal risk management strategies.

Types of Risks to be Managed

Following Merton (1993), Tufano (1996), Scholes (2001) and others, risk management objectives of the type considered in this book can be achieved through diversification, hedging and insurance. This classification is somewhat misleading, as it disguises the problem of specifying the hedging situation and gives the appearance that the risk being managed can, somehow, be managed in a systematic and unambiguous fashion. In emphasizing pro-active management techniques, methods needed for risk avoidance and risk absorption are not examined. The various risk management approaches also differ in applicability to specific cases. In some situations, such as Tufano's gold mining sample, the firms involved have little opportunity to exploit diversification opportunities to manage risk. In other cases, such as globally diversified investment funds, diversification is an integral part of risk management. Hedging situations vary and the identification of an optimal risk management strategy depends on the objective function specified. Similarly, the management of “risks” disguises the importance of the “uncertainty” contained in random future events. It is difficult to formulate general rules. Even if general rules can be derived, such rules rely crucially on information about the properties of the relevant random variables, i.e., the types of risks being managed

Continuing with the “insurance principle” approach provided by Frank Knight, actuarial science can provide an excellent source of general insights into the risk management problems of interest, e.g., Vaughan (1982). By design, actuarial science examines situations where only the chance of loss or no loss is considered. This is a restriction on the properties of the random variables being modeled. As such, there is only partial overlap with the situations of current interest, where the random variables, such as profits or wealth, can take both positive and negative values. Whereas insurance problems seek to reduce risk, it may be desirable to increase certain commercial risks if the potential gains significantly outweigh the possibility and size of loss. Given these qualifications, Vaughan (1982) suggests the following methods for handling the risks faced in actuarial science: *risk can be avoided*, e.g., by foregoing the writing of a policy; *risk may be retained*, e.g., by self-insuring; *risk may be transferred*, e.g., through hedging; *risk may be shared*, e.g., through the purchase of reinsurance; and, *risk may be reduced*, e.g., by increasing audit surveillance.

The contrast between the actuarial science approach and that suggested by Merton (1993) is revealing. There is some close correspondences. Risk reduction can be equated to diversification, risk transference to hedging and risk sharing to insurance. This leaves risk avoidance and risk retention not counted. These omissions are significant. Finance tends to approach risk management by emphasizing applications of the various risk management products that are available in the financial marketplace. Limited attention is given to identifying methods of self-insuring or risk avoidance. Yet, these methods do receive attention in studies outside the financial risk management arena. For example, a number of strategic management studies, e.g., Oxelheim and Wihlborg (1997), propose techniques for strategic hedging, that can lead to self-insurance as an outcome of active risk management. Similarly, strategic risk management preaches risk avoidance through natural hedging. These different potential approaches to risk management tend to take different views on the randomness that faces decision makers.

Assessing the relevance of these different views becomes more complicated when it is recognized that actuarial science, the mathematical science of insurance, is not concerned with the range of risks that are conventionally encountered in financial and commodity markets. Actuaries are concerned with the probability of loss versus no loss. Many of the risks encountered in financial and commercial markets are *speculative risks*, where there is a possibility of loss, as well as a possibility of gain. Such risks can be distinguished from *pure risks* that involve situations with only the chance of loss or no loss (Vaughan 1982, p.8):

> The distinction between pure and speculative risks is an important one, because normally only pure risks are insurable. Insurance is not concerned with the protection of individuals against those losses arising out of speculative risks. Speculative risk is voluntarily accepted because of its two-dimensional nature, which includes the possibility of gain.

It is apparent that the study of risk management requires a careful and detailed discussion of the definition and classification of the types of risks that are going to be managed.

This observation extends to the comparison between financial risk management and strategic risk management. Financial risk management typically treats risks in isolation. The measurable component of randomness is modeled and estimates of possible gains and losses are assessed. In some cases, risks are considered in a portfolio context, by taking into account the relevant correlations between the measurable components. Extreme cases are approached by stress testing the modeling process using extreme observations. Strategic risk management focuses more on evaluating the uncertain component of randomness. The business operation is considered as a whole, and an attempt is made to provide a coordinated approach to the corpus of risk and uncertainty facing the firm. Compared to financial risk management, this process is considerably more difficult to implement in a quantitative fashion. In some situations, it may not be possible or desirable to engage in meaningful strategic risk management, outside of the realm of financial risk management. This argues for examining the techniques of financial risk management before considering strategic risk management.

What is Value at Risk?

To say that the value at risk (VaR) methodology has revolutionized financial risk management is, arguably, an understatement. The importance of value at risk extends well beyond the implementation of the Bank of International Settlements (BIS) capital adequacy standards (BIS 1996, Danielson et al. 1998). For example, the introduction of SFAS 133 has inspired firms, both financial and non-financial, to include VaR calculations in annual reports and other financial statements. The increased attention to risk management has led many firms to reform the process by which risk management is integrated into the hierarchy of managerial control. The VaR technique is, in and of itself, not much different than risk management techniques that have been used for many years by more sophisticated firms. The VaR revolution is associated more with the system wide adoption of these techniques, particularly by depository institutions and other financial intermediaries. Presumably, the system wide introduction of VaR has resulted in a corresponding reduction in systemic risk in financial markets.

On balance, the VaR revolution has been more profound for financial firms, e.g., Jorion (2001). Non-financial firms pose a somewhat different risk management problem (Oxelheim and Wihlborg 1997, p.21):

> For a non-financial firm the primary risk would be its commercial risk – i.e. its uncertainty about the value of cash flows that can be generated by its physical assets producing output. Its liquidity risks are secondary in the sense that they merely enhance or modify the primary risk. The importance of a specific kind of risk can shift depending upon the situation.

As such, the value at risk (VaR) revolution is somewhat narrowly confined to financial firms, especially firms making markets in derivative securities and other leveraged instruments such as bond portfolios financed using repurchase agreements. VaR is also of importance for non-financial firms, particularly multinational firms, seeking to assess and control the financial risk that is associated with activities such as currency and interest rate risk management.

Wilmott (1998, p.547) provides a useful definition for value at risk (VaR):

> *Value at risk is an estimate, with a given degree of confidence, of how much one can lose from one's portfolio over a given time horizon.*

The reliance on degree of confidence immediately suggests a connection to probability theory and the specific topic of hypothesis testing. The time horizon selected will vary according to the specifics of the situation. For example, a financial trading firm will do daily VaR calculations, while a portfolio manager will examine VaR when the portfolio is being rebalanced, a task that could be done monthly or quarterly. VaR has the irresistible attraction of providing a single number that summarizes the total risk of a position. The total risk can arise from a number of different situations, e.g., the equity value of a financial firm, a derivatives trading portfolio, an internationally diversified portfolio of stocks and so on.

A useful starting point for an introductory treatment of VaR is Hull (2000, p.342):

> The VaR calculation is aimed at making a statement of the following form: "We are X percent certain that we will not lose more than V dollars in the next N days." The variable V is the VaR of the portfolio. It is a function of two parameters: N, the time horizon, and X, the confidence level. One attractive feature of VaR is that it is easy to understand. In essence, it asks the simple question: "How bad can things get?" In calculating a bank's capital, regulators use N = 10 and X = 99. They are, therefore, considering losses over a 10-day period that are expected to happen only one percent of the time. The required capital for market risk is, at the time of writing, three times the 10-day 99% VaR.

Though the VaR methodology can conceptually be applied to a wide range of situations, applications have focused on situations involving market risk: the potential for changes in the value of a position resulting from changes in market prices.

An important impetus to the spread of VaR has been the widespread availability of software and technical material to support the implementation. For example, the RiskMetrics group at J.P. Morgan/Reuters has been an important promoter of the VaR methodology by providing detailed technical publications, such as the RiskMetrics manual (JP Morgan 1996), and daily data sets for important financial variables free of charge on the J.P. Morgan web page at http://www.jpmorgan.com/RiskMangement/RiskMetrics/RiskMetrics.html. The RiskMetrics manual describes a set of methodologies outlining how risk managers can compute VaR on a portfolio of financial instruments. RiskMetrics pays close attention to modeling the VaR for positions containing options. The non-linear payoffs associated with options pose definite, if solvable, problems for the VaR methodology.

VaR for the One Asset Case

VaR can be calculated for a single component of the firm’s operations, e.g., to assess the activities of a single trader, or for the full portfolio of a firm’s activities. Because VaR can be affected by the presence of non-linear payoffs arising from the presence of securities such as options, a number of different methods have been proposed to arrive at VaR estimates. The simplest of these methods is the *variance-covariance* method, e.g., Alexander and Leigh (1997), Duffie and Pan (1997). Though this method may produce inaccurate estimates if non-linear payoffs are present, it is the easiest to understand and implement. As a consequence, the variance-covariance method is also the basis of the most widely used VaR applications. This method establishes an immediate connection between VaR and techniques of probability and statistics. As with all the VaR methodologies, risk is treated as a measurable quantity, ignoring the implications of uncertainty. Extreme deviations from previously observed financial price behavior are handled by stress testing, i.e., further assessing the impact that extreme observations can have on the VaR estimates.

VaR calculations require a number of exogenous inputs. Before starting, the level of confidence and the time horizon for the VaR estimate are needed. In turn, these exogenous values depend on the degree of aversion to losses. Conceptually, as the aversion to loss increases the level of confidence in the estimate will increase from, say, 95% to 99%. A similar comment applies to the selection of a time horizon,which can vary from daily to weekly to monthly. In theory, VaR for hourly intervals could be calculated. Large financial firms that face considerable market risk typically calculate a daily VaR. Non-financial firms which face a more limited range of market risk, e.g., the currency and interest rate risk for Coca-Cola, could calculate weekly or monthly VaR. It is sometimes maintained that the time horizon "is supposed to be the timescale associated with the orderly liquidation of the portfolio, meaning the sale of assets at a sufficiently low rate for the sale to have little effect on the market" (Wilmott 1998, p.548).

The other essential input to the VaR calculation is the data for the prices of the assets of interest. If the asset is traded, then market prices can be used. If the asset is not traded, then an estimate for the price has to be obtained from an appropriate pricing model, e.g., Hendricks and Hirtle (1997). For financial institutions subject to BIS-style rules, the relevant pricing models have to conform to certain requirements. In-house models are acceptable, perhaps preferable, as long as the resulting prices are, on average, accurate. From this data, the relevant statistical parameters can be calculated using conventional formulas appropriate for the probability distribution selected. For the single asset case, the mean (μ) and volatility (σ) of the asset value have to be calculated. For statistical reasons, the parameters are calculated for the rate of return distribution, as opposed to working directly with price levels or price changes. For the portfolio case, in addition to the individual asset return volatilities, the asset return covariances are also calculated and the VaR of the asset portfolio determined using the familiar formula for the portfolio variance that is covered in introductory investment analysis (see Sec. 6.3).

To illustrate the VaR methodology, consider the value of a portfolio containing a single non-dividend paying asset. Let $r_t = \ln(1 + R_t)$ where $R_t = (S_t - S_{t-1})/S_{t-1}$ and S is the asset price. Define the probability density associated with r as $\Phi[r]$. With this density it is possible to obtain the probability that a future value of r will take a value less than r^*:

$$Prob[\, r_{t+1} < r^*] = \int_{-\infty}^{r^*} \Phi[r]\, dr = c$$

In this calculation, $c = (1 - \alpha)$ where α is the desired level of confidence, e.g., 5% or 1%, and r^* is defined by the level of confidence. Parameter estimation and calculation of confidence levels proceeds by assuming that $\Phi[r]$ is a normal (Gaussian) probability density.

Having assumed that r is normally distributed, parameter estimates for the volatility and mean of asset returns, σ and μ, are obtained from the available historical data and the probability equation for, say, a 99% degree of confidence can now be determined by using the standard normal form:

$$Prob[\, r_{t+1} < r^* \,] = Prob[\, Z < \frac{r^* - \mu}{\sigma} \,] = c = .01$$

Using the appropriate value from the standard normal distribution tables (2.33 for a one-tailed test at the 99% level, 1.645 at the 95% level and so on), this equation can inverted to solve for r^*:

$$r^* = (\mu - 2.33\, \sigma)$$

Value at risk can now be determined by evaluating $QS\, r$ and $(QS_t\, r^*)$, where Q is the number of units of the asset held.

For short time horizons, such as daily VaR, it is usually assumed that $\mu = 0$ to produce the result:

$$VaR = -2.33\, \sigma\, (Q\, S)$$

For longer time horizons, where $\mu \neq 0$, the solution is:

$$VaR = (QS)\, (\mu - 2.33\, \sigma)$$

Some presentations of the calculation of the return form of VaR have an additional time scale factor to account for differences between the time scale used to estimate the volatility and the time horizon for the VaR, e.g., if a weekly VaR is desired and the volatility estimate is for annual returns, scaling by $(1/52)^{1/2}$ is required. This adjustment is unnecessary if the sampling frequency of the data used to estimate the parameters is the same as the horizon for the VaR forecast.

Value at Risk, Normality, and Options

As evidenced by the numerous articles detailing potential problems associated with the variance-covariance VaR model, there are numerous pitfalls that can arise in VaR modeling and application, e.g., Duffie and Pan (1997), Ju and Pearson (1999). For example, Alan Greenspan (1996) observes that "disclosure of quantitative measures of market risk, such as value at risk, is enlightening only when accompanied by a thorough discussion of how the risk measures were calculated and how they relate to actual performance." It seems that even the regulators who proposed and promoted these quantitative risk measures have real reservations about the practical implications of naively using the techniques. To be sure, the VaR revolution has led to the institutionalization of quantitative risk management. Insofar as such techniques were not in place at various financial institutions, the revolution has produced a substantive reduction in system wide exposure to market risks. Financial firms are required to identify and measure the set of risks that confront them. Yet, excessive reliance on quantitative risk measures can produce a chimera for individual firms and, possibly, for the financial system as a whole. Quantitative measures can provide false confidence that financial risk has been effectively identified and appropriate actions have been taken to deal with quantitative risk that has been identified.

As it turns out, there are a number of serious problems that can arise in determining the VaR for a given firm., e.g., Culp et al. (1998), Marshall and Siegel (1996). Some problems with VaR modeling are theoretically rectifiable. Other problems can be rectified only by significantly increasing the complexity of the VaR modeling process. Some problems may not be rectifiable at all and heuristic adjustment will be required. Because the set of problems facing a VaR modeler will vary from firm to firm, there is value added to considering the various limitations and extensions of the VaR model. Consider the assumption of normality, e.g., . This assumption is made for ease of implementation. It allows immediate application of techniques of hypothesis testing using the standard normal distribution inherited from elementary probability and statistics. Familiar estimators for μ and σ can be employed to determine the parametric inputs. The limitations of using normality to model financial prices are widely recognized. It is well known that the probability distribution for changes in financial prices are not normally, being typically fat tailed (leptokurtic) and, often, skewed. If the deviation from normality is significant, this will impact the critical (α) values, e.g., testing at the 1% level may actually be testing at, say, the 12% level.

The problems associated with the normality assumption are in the realm of the theoretically rectifiable. A range of potential solutions have been proposed to deal with limitations of the normality assumption. The basic idea is to adjust the normal distribution to accurately reflect the true tail density. This can be done by empirically fitting a distribution to the past data, a task that is not without difficulties. Once the functional form for the distribution has been determined, tail densities can be accurately identified, with the VaR formulas given above being appropriately adjusted. As the whole distribution has to be modeled, this is a process that can be more difficult than required. A somewhat less complicated approach would using a series approximation, such as an Edgeworth expansion, to the true distribution. This would result in high moments, such as skewness and kurtosis, being estimated and used to adjust the tail densities. Such approaches are less popular than approaches that make appropriate adjustments to the parameters that are needed for testing, the volatility and, possibly, the drift.

The most popular of these solutions revolve around providing improved volatility estimates. Important approaches include GARCH based estimators (Bollerslev et al. 1992) recommended by the Riskmetrics group or, where available, implied volatility estimates backed out of option pricing models (see Chapter 8). In effect, the difficult problem of directly fitting the complete distribution is replaced by the more tractable problem of determining the parameter(s) of interest. However, which particular method will work best is not clear. The issues involved are similar to those in the debate surrounding whether advanced econometric estimators of volatility, such as GARCH, will provide more accurate forecasts of spot price volatility than estimates obtained using implied volatilities. Evidence on this issue is mixed, being complicated by various issues, such as determining precisely what the variable be forecasted is and how the implied volatility estimate is determined, e.g., Takezawa (1995).

To assess whether the volatility estimate is reliable, it is recommended that techniques such as "back testing" the model be used, e.g., Kupiec (1995), Jackson et al. (1997), SBC Warburg (1998), Lopez (1999).

A more serious problem for the variance-covariance approach to VaR can arise with the types of cash flows being generated by the firm. In particular, many securities traded in financial markets, such as options, have non-linear payoffs. Many real assets also contain various types of options that can have a non-linear impact on firm valuation (see Sec. 7.4). Non-linear payoffs substantially complicate the problem of dealing with non-normality, e.g., Hull and White (1998). To address these problems, the alternative *delta* and *delta-gamma* approaches to VaR have been proposed, e.g., Riskmetrics (1996), Duffie and Pan (1997). Much like the use of duration and convexity in fixed income analysis, these approaches use a Taylor series expansion (see Appendix I) to approximate the non-linear payoff function. In fixed income analysis, the payoff function is usually assumed to be convex, to reflect the inverse relationship between price and yield. This assumption may not be valid for situations involving options.

To see this consider the VaR for a one security portfolio containing a European call option, with price C, on a non-dividend paying common stock with price S. The delta VaR approach uses the first order Taylor series approximation:

$$C(S) - C(S_0) = (\partial C / \partial S)\,(S - S_0)$$

where the partial derivative, $(\partial C / \partial S)$, is evaluated at S_0. Sec. 9.1 derives a functional form for $(\partial C / \partial S)$ that could be used in this case, with the change in S being handled, say, in the same fashion as in the variance-covariance approach. For larger changes in the value of S, a second order approximation will be needed. This can be done by using the second order Taylor series expansion, that involves the second derivative or gamma term. However, to be consistent with option valuation theory, the most appropriate form for the expansion is second order in the state variable S and first order in time. Bookstaber (1997) discusses difficulties with applying VaR to "globally managing" the range of risks facing a financial firm.

2.3 Risk Management and Speculation

What is Risk Management?

Risk management is an immense subject, ranging from medicine to engineering to finance to political science.[10] Risk is a pervasive phenomenon. As such, methods for managing risk are a natural adjunct to everyday life. Risk management decision problems range from the relatively straight forward, such as those involved in quality control on an assembly line, to the ethically and morally challenging, such as those involving treatment selection for a terminally ill patient. Though there are some general principles that apply to most risk management situations, it is practical to narrow the focus to the specific types of risks that are of interest. Yet, this narrowing of focus is not without difficulty. Treating risks individually can over simplify the problem, ignoring the complementarity that arises between different types of risk. Examining groups of risks together can also suggest different types of solutions. This problem of how best to structure risk management decisions is relevant to financial and corporate situations.

Consider the financial risk management situation of a depository institution. These firms are subject to at least three types of risk: market risk, credit risk and operational risk. Market risk can be further subdivided into interest rate risk, currency risk, equity risk and commodity risk. Liquidity risk could be added to this list or treated separately. Credit risk can be subdivided into liquidity risk, sovereign risk, corporate risk and individual/personal risk. Operational risk can be subdivided into system and control risk, management failure risk and the risk of human error. Though much discussion of financial risk management focuses on market risk, examination of the recent derivative debacles reveals the importance of operational risk and credit risk (Kuprianov 1995, p.2). Within this environment, the risk management process is decidedly complex. It is the goal of senior management to have in place "a risk management system that links capital, risk and profit in a way that enhances profitability whilst satisfying the ... demands of regulators and the marketplace" (Arthur Anderson 1998).

From a corporate perspective, the need for an integrated approach to risk management has been understood for many years. Dowd (1998, p.9) captures the general approach required:

> The first point to appreciate is that all sensible approaches (to risk management) have the same first step, i.e., we formulate a corporate risk management philosophy to impose some guidelines on risk management decision-making. This tells us what kinds of risks we wish to bear, what risks we want to avoid, what sort of options we will consider to manage our risks, and so forth. Usually, we will readily bear those risks that we have some particular expertise in handling (e.g., risk unique to our particular line of business), but there will also be other risks that we will usually wish to avoid (e.g., the risk of our factory burning down). This philosophy should also give us some indication of what attitude we should take towards the many other types of risks that we might face -- when we should bear them, when we should not bear them and the like.

Strategic risk management is the term often used to describe the process of formulating and implementing a corporate risk management philosophy. This vital step in the risk management process receives relatively little attention in conventional presentations of financial risk management that tend to focus on risk measurement and techniques of hedging, diversifying and insuring.

The importance of the risk management function has led to the creation of risk management offices within larger corporations, run by risk management officers. Such offices engage in a whole range of activities that are of little or no relevance to financial risk management. Monitoring of Occupational, Health and Safety rules, internal security, handling of various insurance plans for fire, theft and the like, these types of activities could be localized in the risk management office. Depending upon the specific corporation, it is possible that the risk management officer may have little or no financial expertise. By design, narrowing the focus to corporate financial risk management abstracts from the integrated risk management process. This runs the risk of failing to make an adequate connection between the implementation of a financial risk management program and the overall risk management philosophy of the firm. Arguably, the failure to make such a connection has contributed to a number of recent and not so recent debacles, e.g., the Metallgesellschaft losses.

Developing a framework for adequately identifying and managing the range of risks confronting the corporation is not possible. There is too much variation across the types of risk encountered by the various firms that a general framework is unhelpful at best, and could be misleading. Some method of simplifying the process is needed. Following Chance (1998, p.672) and others, one possible method is to restrict the types of risks encountered: “Risk management is the practice of defining the risk level a firm desires, identifying the risk level a firm currently has, and using derivatives or other financial instruments to adjust the actual level of risk to the desired level of risk.” In this approach, risk management is closely identified with the types of instruments that can be used to manage risk. Jorion (2000, p.3) takes a similar tack: “*Risk management* is the process by which various risk exposures are identified, measured and controlled. Our understanding of risk has been much improved by the development of derivatives markets”.

The modern approach to risk management typically proceeds by classifying risks into the categories encountered: business or commercial risks; market risks; credit risks; liquidity risks; operational risks; and legal risks. These importance of these classifications in practical risk management is reflected in the annual reports of major banks.[11] For both financial and non-financial firms, the risk management process requires each of these risks to be assessed for the specific corporation involved. Decisions are then made on which of these risks will be assumed and which will be managed. Beyond this general intuition, things get more difficult as it is not possible to deal with all the various aspects in detail. Further focus is required. As a consequence, there is a myriad of different possible approaches to corporate risk management. Some treatments, e.g., Dowd (1998), Jorion (2000), emphasize the measurement of market risks using VaR; others, e.g., Oxelheim and Wihlborg (1997), emphasize the integrated treatment of business risks. Still others, e.g., Smith, Smithson and Wihlford (1995), examine the methods for handling financial risk in specific situations.

The modern, integrated approach to corporate risk management is a utopian ideal. It is conventional, if not essential, to treat risks in isolation in order to better conceptualize the methods of managing the risk. In certain situations, such as for financial firms making markets in securities, the risks being managed are primarily market risks and utopian models, such as those derived from VaR, can be used effectively. These situations stand in stark contrast to cases where the risks are less amenable, e.g., Proctor and Gamble or Coca-Cola seeking to manage firm wide business risks across different geographical markets. Even where the risks are perceived to be primarily market risks, the complexity of the risks being faced can defy adequate treatment, e.g., LTCM. In all of this, the traditional distinction between hedging and speculation seems misplaced. The is unfortunate, as there are useful lessons contained in the earlier discussions of risk management, which typically structure the discussion as a problem in hedging specific transactions..

The traditional approach to risk management predates the modern Renaissance of derivative securities. Products for managing risks were limited, both by legislation and market practices. Financial derivatives, so important in modern financial risk management, were largely traded OTC and were not widely used. Various types of financial risks, such as the volatility associated with flexible exchange rates, were still on the horizon. In this simpler world, risk management was typically associated with hedging using agricultural forward or futures contracts, where risks were treated in isolation and transactions involved in the hedge were emphasized. Considerable attention was dedicated to clarifying the distinction between hedging (risk management) and speculation. This distinction between risk management and speculation seems to have been lost in the modern approach to risk management. There is a modern belief that the engineering of risk is a precise enough science to make this distinction irrelevant.

Hedgers vs. Speculators: The Cargill Corn Case

The traditional approach to risk management divides derivative market participants, particularly those operating in forward and futures markets, into the two general groups of *hedgers* and *speculators*. This distinction has both economic and legal implications. Cummins et al. (1998, p.31) describe the economic distinction: “The terminology typically used is that if managers are attempting to reduce risk through their actions, they are said to be hedging; if managers are trying to increase the firm’s risk exposure because they believe such a strategy will yield abnormal profits, they are said to be speculating.” A number of the legal implications of the distinction between hedgers and speculators can be found in the Commodity Exchange Act, where a legal separation of traders in futures markets is identified and differential reporting requirements and position limits are specified. Commissions and margin requirements can and do vary between the two types of traders. The basic textbook distinction between hedgers and speculators has hedgers trading to reduce the risk associated with a cash market position while speculators are trading solely on the basis of expected price changes. While useful analytically, there is much more to this dichotomy that needs to be explained.

Historically, some legal distinction between the two types of traders was essential. The anti-speculative sentiment surrounding the passage of the CEA (1934) had to be tempered to accommodate commercial interests with a real need to trade derivative securities in order to manage risk. Initial definitions of hedgers were quite restrictive. By the 1970s, the easing of anti-speculative sentiment spurred on by the increasing needs of commercial enterprise was sufficient to produce a rethinking of the legal treatment of hedging and speculation. Of particular interest, as part of the process surrounding the revisions to the Commodity Exchange Act (1974), the CFTC was required to provide a definition of hedging in order to determine which traders would be subject to position limits on speculative trade, and which traders would be considered as hedgers, and not subjected to limits on trading positions. This definition for a hedger was released in 1977. Instead of attempting a precise legal definition, the CFTC opted for a long and involved definition, derived from the economic motivations for hedging.

Though there is considerable scope in the CFTC definition for consideration on a case-by-case basis, the CFTC definition generally requires (Leuthold et al. 1989, p.71):

> (a) it must be economically appropriate to reduce risks; (b) risk must arise from operating the commercial enterprise; (c) the futures positions normally represent a substitute for transactions to be made later in the physical market; and (d) price fluctuations in futures markets must closely relate to fluctuations in the cash market value of assets, liabilities or services hedged. Thus, hedging is more than just enumeration of specific transactions and positions -- it is a process of risk reduction. Prior to this definition, several legitimate hedging operations, especially cross and anticipatory hedges, were not recognized as hedges because futures positions did not meet the approximate equal and opposite requirement to the cash position.

While this explicit attempt to incorporate economic motivations into the definition of hedging is definitely an improvement over a strict legal approach, the underlying issues may be unresolvable. As discussed in Section 2.1, there is a speculative element in the risk management decision, optimal hedging involves an element of speculation.

The Cargill corn case (Falloon 1998, ch.8) provides a classic instance of the difficulties, and associated implications, of distinguishing hedging from speculation. The case originated from actions of Cargill, Inc., a grain marketing company, primarily in the trading of corn futures contracts on the CBT during 1936 and 1937. Cargill was a major player in the grain industry, at that time handling approximately 12% of American grain being marketed (Broehl 1992, p.467), arguably the largest single private grain distributor. Also at the time, Cargill was the largest user of grain futures contracts on the CBT. Despite being a major player in the markets, Cargill was something of an outsider in the CBT hierarchy, having only grudgingly been admitted as a clearinghouse member in 1935. Cargill did not have a seat on any of the primary CBT governing committees. As a result, Cargill reacted negatively to a series of adverse decisions from the CBT that were implemented to prevent what was perceived as blatant market manipulation by Cargill. The resulting court cases that originated were the first major test of the market manipulation provisions of the CEA.

As a major player in the corn market, Cargill was able to forecast tight conditions in US corn supplies in 1936 and again in 1937. In response, during July 1936 Cargill bought corn offshore, in Argentina, for import and placed a large long position in Sept 1936 corn futures on the CBT. Such activities are consistent with the role of Cargill as a grain distributor seeking to match orders with purchases. What was questioned by both the CBT and the regulators was whether the size of the position was fully consistent with hedging activity. Did the position contain a significant speculative component? The size of Cargill's corn futures position constituted about 1/4 of open interest and can be compared with the 22.5 million bushels, which was the four year historical average for the "visible supply of corn in the US" at that time (Falloon 1998, p.190). As the end of the delivery month approached, the size of this position attracted the attention of the business conduct committee of the CBT, which conducted meetings on the matter involving, on Sept 25, the president of Cargill. Despite assurances from Cargill that the position would be unwound in an orderly fashion and no evidence to the contrary, the CBT board of directors took action on Sept. 29 allowing an extension of the deadline for notice of physical delivery. This action precipitated a substantial price drop in the cash corn price, adversely impacting Cargill.

Not surprisingly, the Cargill management was incensed. The atmosphere was again poisoned on December 7, 1936 when the business conduct committee took action regarding Cargill's position in the Dec 1936 corn contract, which was deemed to be too large to be justified by legitimate hedging activity. The committee took the unprecedented action of ordering Cargill to reduce positions. Even though Cargill complied, at an estimated cost of 15¢ per bushel per contract, positions were only further hardened setting the stage for the events of Sept. 1937. Again confronted with supply shortages, during the summer of 1937 Cargill placed large long positions in CBT corn futures, first in the July contract and then for the Sept contract. The Sept position was about double that of the previous year, being as large as 9.4 million bushels, about one half of the contract open interest. Based on past experience, Cargill surmised that the size of this position would come under intense scrutiny by the CBT. To counteract such scrutiny Cargill entered into temporary futures-for-cash exchange contracts with the Continental Grain and Uhlmann Grain companies. The result was a reduction in Cargill's reported position to 2.2 million bushels.

The grain business being a closely knit community, it was not possible for Cargill to disguise the actual activities from the CBT business conduct committee which was only too aware of Cargill's controlling position in the Sept contract open interest. As the delivery month progressed, Cargill did little to reduce the size of its position, resulting in Cargill having an increasingly larger share of open interest in the contract. By Sept. 22, the pressure on the price of deliverable stocks was evident. Despite apparent assurances to liquidate in an orderly fashion, Cargill did not move promptly to reduce its position. In response, on Sept. 23 a cease and desist order was sent to Cargill and on Sept. 24 a trading halt was ordered and a settlement price for all outstanding contracts was set at $1.10½, 2 cents below the close of Sept. 22. The impact on the cash market was predictable, the cash price fell resulting in significant losses for Cargill on the cash grain it held in company stocks, as well as on the cash corn it had acquired in the cash-for-futures swap it had done with Continental and Uhlmann. Cargill was incredulous and a long series of CBT committees as well as two court actions under the CEA were initiated. Hearings of the CBT board of directors in March 1938 resulted in the expulsion of Cargill from the CBT. The resolution of the CEA cases also went against Cargill, though only specific managers at Cargill were sanctioned. In the end, Cargill, Inc. was still able to use the futures markets to facilitate grain marketing.

Was Cargill engaged in legitimate hedging actitivites? Surely, there was a significant hedging element behind some of Cargill's futures activities. Cargill maintained that insiders at the CBT acted to undermine the legitimate activities of a grain distributor, with the members of the CBT business conduct committee directly benefitting from the negative decisions made against Cargill. This accusation that the futures exchanges act as monopolies seeking to further the interests of the members has been replayed in other cases. For example, a similar comment applies to the Hunt silver manipulation. The apparent evidence of manipulative activities on the part of Cargill was confirmed in various forums, from CBT hearings to court actions under the CEA. Yet, somehow, the arguments are not clear cut. The boundary between legitimate hedging, speculation and speculation supported by manipulative intent is not clear cut. The hedging decision involves a speculative component. Combined with sufficient impact in the cash market, it is possible to rig the game. Precisely when rigging the game is happening is not as easy to discern as might appear.

Speculation and Manipulation[12]

What is manipulation? The answer to this question is important, if only because manipulation is an activity that is considered illegal under a number of US statutes. For example, the Commodity Exchange Act (1936) makes it a felony "to manipulate or attempt to manipulate the price of any commodity in interstate commerce". The CFTC licences futures exchanges with guidelines requiring rules being in place that prevent manipulation. Various other statutes dealing with price fixing and monopoly also make manipulation a criminal activity. As illustrated in a large number of civil cases, e.g., *Minpeco vs. Hunt*, there are also severe civil sanctions associated with attempts to manipulate markets. Yet, despite all this legal foundation: "The law governing manipulations has become an embarrassment -- confusing, contradictory, complex, and unsophisticated" (McDermott 1979, p.205).

The difficulties in the law surrounding manipulation speak to the difficulty in defining manipulation. Following Gray (1981), it is useful to make a distinction between *manipulation*, which is an economic concept, and *illegal manipulation* which is a legal notion. There are two essential elements required for an illegal manipulation: intent and "the creation of an artificial price by planned action".[13] The issue of intent is primarily a legal concept that is difficult to capture in a legal sense. It often speaks to the specifics of the case at hand. Similarly, price artificiality is an economic concept, that is also difficult to capture. Presumably, an artificial price is not a market clearing price. The underlying forces of supply and demand have been circumvented for personal or corporate gain. But this puts too much pressure on economic theory, a science that can usually provide only a vague estimate of what the 'true' market price ought to be in a given situation. Various measures of artificiality have been proposed, e.g., Leuthold et al. (1989, p.383), with some of these measure being adopted in specific legal cases.

In approaching manipulation, the courts have chosen to proceed piecemeal. The Congress has stated quite clearly that manipulation is not a desirable economic activity. Yet, a precise definition of manipulation is not available in the relevant statutes. The number of legal cases dealing with manipulation in the US is small, probably not more than thirty such cases have gone to trial since the 1950s, e.g., Johnson (1981), Gray (1981). Most of these cases have originated from trading on futures exchanges. From these cases, certain actions have been identified that are essential features of a manipulation: a controlling position in the appropriate derivative contracts; a dominant position in the deliverable commodity; and the undertaking of specific actions that would produce an artificial price. The first activity is associated with a squeeze. The first two activities together constitute a corner. The last activity encompasses what Williams (1995, p.6) refers to as:

> A "rumor" manipulation, in which someone with a previously established position in the physical commodity or in futures convinces other traders through false reports that a shortage in that commodity will occur, for example, through a rumor of a freeze. The rumor must be believed by others only long enough for the manipulator to close out his position at top prices ... (or) ... An "investor-interest" manipulation, in which a series of trades and statements made by the manipulator convinces others of a broadly based desire to hold the commodity, thereby increasing its price. Until others realize that

the underlying interest is merely temporary, the manipulator can sell her holdings at a high price.

The resulting confusion associated with applying all these standards in a specific legal situations is understandable.

The confusion surrounding manipulation extends to the jargon used to describe the possible strategies. Some sources, e.g., Williams (1995, p.6), use the terms squeeze and corner interchangeably. Others, e.g., Leuthold et al. (1989), require a controlling position in both the derivative market and the cash market for there to be a corner. For a squeeze, the trader only takes advantage of cash market shortages (oversupply) by establishing long (short) positions in derivative contracts. Corners and squeezes can occur from both the short and long side of the market, though most attempts in practice are from the long side. For a long corner, the trader establishes long derivative positions, typically well in excess of available deliverable supplies. At the same time, the trader has attempted to obtain a controlling position in the available supply. The process of standing for delivery on the contracts forces the short side to pay high prices to bring available supplies onto the market. Yet, those available supplies are controlled by the holder of the long derivative positions. The shorts are forced to go "hat in hand" to the longs to cover their positions.

The textbook description of a long side corner is usually more involved in practice. The Hunt silver operations leading up to the silver price peak in 1980 had elements of a corner, but there are real questions about the presence of whether the Hunts and their confederates had controlling positions in both the spot and futures markets. The Sumitomo copper operations that ended in 1995 is a much better example of a corner. This particular operation was spread over a long period of time, with the position in the deliverable spot commodity growing gradually, starting around 1986 when Yasuo Hamanaka assumed control of Sumitomo's team of copper futures traders. In the Sumitomo case, a plausible explanation was given for the buildup in stocks: even before Hamanaka began his trading activities for Sumitomo, the firm was an important player in the international copper market.

The Hunt silver operation had many of the earmarks of a traditional corner. Swashbuckling entrepreneurs making big bets on rigged games. The evolution of the Sumitomo copper corner has a decidedly more modern flavour. Hamanaka was a career man at Sumitomo, with a 20 year history in the company division. Whether more senior Sumitomo executives were aware of his activities is not clear. However, in any event, Hamanaka was legitimately able to assume huge positions in cash and futures on Sumitomo's behalf. He also had signing authority over various corporate bank accounts and access to corporate lines of credit. Hamanaka also was able to geographically disperse his positions around the globe and to exploit the laxness of regulators in specific jurisdictions. For example, Hamanaka did a considerable amount of trading on the London Metal Exchange which, together with the Comex are the most important markets for forward and futures trading of copper. Despite being vigorously warned about possible wrong doing by Hamanaka as early as November 1991, the LME did not get actively involved in serious investigations of Hamanaka until the CFTC became involved in October 1995.[14]

Due to filing requirements and other regulatory oversight, cornering activities in modern markets require considerable effort to avoid detection. The elaborate networks of traders involved in the Hunt silver operations was needed to avoid the appearance that large positions were being accumulated on one side of the market by one group of traders. Despite the laxness of the LME, Hamanaka had to enter into arrangements to hide the total size of Sumitomo's position in deliverable supplies of copper. A combination of LME regulatory laxness and careful planning permitted Hamanaka to successfully deny involvement in a market manipulation. This despite a number of instances of market turbulence, such as that in the extreme cash-futures price backwardation of September of 1993, where evidence of a cornering operation was difficult to deny. Eventually, it was the vigilance of the CFTC that in April 1996 announced that it had uncovered sufficient irregularities in the Sumitomo accounts to proceed with regulatory actions.

Both the Hunt and Sumitomo operations were manipulations aimed at forcing up prices, to the detriment of the short side of the market. Operations aimed at driving down prices, to the detriment of the long side of the market, are much less common. Take the case of a corner aimed at squeezing the longs. The trader aims to acquire a controlling position in the cash commodity without significant impact on cash prices. Once this is done, the trader establishes a controlling short derivative position across a range of delivery dates and aims to delivery a large amount of the commodity against the nearby contracts. The trader may simultaneously engage in cash market sales in order to further depress the price. The combination of selling pressure from the short derivative positions and the weakness in the cash market permits the trader to profit from the excess short future positions that are being held. The investment in the cash commodity that had been accumulated has now been recouped through deliveries on the futures contracts.

Manipulation is legally difficult to prove. When prosecuted, those involved in activities aimed at manipulating are often convicted of crimes associated with covering up the manipulation, e.g., forging documents, lying to regulators, and not for violation of statutes directly concerned with manipulation. The lines between manipulation and legitimate speculation are difficult to define. Consider the case where an astute trader identifies a trading opportunity associated with the lack of deliverable supplies for a nearby contract delivery. The possibility of a squeeze causes the trader to take much larger nearby long positions than would be customary. Is this trader involved in manipulating markets? Say this trader was also a major player in the cash market and had what would be reasonably considered a potentially controlling position in the cash market. Would the unusually large long, nearby speculative position be manipulative in this case? Preventing the trader from taking positions aimed at profiting from the potential squeeze would be unreasonable, as other smaller traders without cash market influence would not be similarly restricted.

2.4 Strategic Risk Management

What is Strategic Risk Management?

Strategic risk management is becoming a popular buzz word in discussions of financial risk management, e.g., Dowd (1998). Unfortunately, this terminology can cover a range of possible notions, creating some semantic confusions. Recognizing this lack of precision in the concept, various notions of strategic risk management have been adopted enthusiastically by numerous financial and non-financial corporations, e.g., Arthur Anderson (1998). A rough description of strategic risk management applicable to all the various notions can be formulated as:

> *Strategic risk management is the process of identifying, implementing and monitoring systems for managing the range of risks confronting the firm.*

The goal of strategic risk management is to deal with the risks facing the firm in a systematic and enterprise-wide fashion, instead relying on the ad hoc and independent risk management functions that often characterize traditional firm activities surrounding risk. As such, strategic risk management is squarely aligned with risk management for corporations.

The importance of the drive to strategically manage risks has not been lost on the management consulting industry. The perceived need for such a strategic management function has led to the major players in that industry to develop programs for implementing the appropriate "business organization and management structures, geographic, regulatory and reporting matrices, and the mandates which underwrite these." Arthur Anderson (1998, p.9) gives the following description of strategic risk management:

> An organization's risk management profile must reflect current business complexity as well as business dynamics, so that risk controls and risk management structures can be adjusted to changing business flows and regulatory requirements. The difficulty for many organizations is that risk management structures have evolved on an ad hoc, rather than organization-wide model. As a result, these structures are disjointed, with risk controls that are not aligned or comparable, and communications processes which do not yield the type or quality of management information required to meet both internal and regulatory requirements.

Upon closer inspection, it seems that the management consulting industry identifies strategic risk management with enterprise wide risk management. This approach focuses on operational systems, such as reporting channels, methods of identifying risks, and solutions to information technology requirements.

The basic idea underlying the more general notion of strategic risk management is appealing. The process starts with the formulation of a risk management philosophy for the firm. This requires an initial evaluation of the range of risks facing the firm. Decisions are then made about the exposures the firm wants to manage and what types of systems will be used to manage those risks. This step in the strategic risk management process is referred to as developing a philosophy because there is much that is subjective and intuitive, especially for non-financial firms. The correct method of identification and handling of risks is not obvious. Loosely put, philosophy has to do with ways of looking at the world. What risks are relevant and how these risks are to be handled depends on the managers' view of the world. This stage in the process is top-down, with senior management being an integral part of the decision making process. It is likely that those senior managers responsible for risk management will be an essential cog in the process, due to the potentially limited knowledge about risk management matters by those at the most senior levels of the firm.

A number of academic studies, many originating from the strategic management area, e.g., Ahn and Falloon (1991), Oxelheim and Wihlborg (1997), identify a specific type of risk management philosophy as strategic risk managment. This essence of this approach can be illustrated by examples. Consider Gallo Wines, a company that produces and sells the bulk of its outputs in the US. Cash flows and assets are denominated primarily in US dollars. Does this firm need to manage FX risks arising from changes in the US dollar? Seemingly no, but if it is recognized that the major competitors for Gallo are situated offshore with cost structures denominated in foreign currency, then Gallo's exposure to FX changes becomes apparent. What about the range of other macroeconomic risks? Oxelheim and Wihlborg (1997) examine the issues surrounding the management of macroeconomic risks, and do a detailed analysis of Volvo Cars. Shapiro (2000) gives numerous other examples, from US ski resorts to Monsanto. These 'soft' risks can be contrasted with the 'hard' risks arising from pure financial decisions, such as funding of debt or investing in marketable securities. These types of issues can be considered as the *conceptual* aspect of the identification phase of strategic risk management.

The process of formulating a risk management philosophy also involves an *empirical* aspect. The conceptual aspect requires detailed empirical data about the various risks facing the firm. These data have to be collected, processed and evaluated. Decisions have to be made about which variables to include, the relevant sample periods to examine and the types of techniques to use in evaluating the risks. There is feedback between the conceptual and empirical aspects. Whereas senior management is primarily involved in the conceptual aspect, the empirical aspect has to have wider involvement, with data inputs being collected and processed in the various risk management units within the firm. Once the basic empirical results have been obtained, decisions have to be made about the appropriate risk management techniques to use for managing the risks. The data may require a fresh look at the firm's approach to risk management, a rethinking of the conceptual aspect, and a retooling of the empirical aspect.

Judging from the risk management problems at various firms, e.g., Metallgesellschaft, Barings, Proctor and Gamble, Orange County, the costs of ignoring the implementation phase of strategic risk management can be considerable. The first step of the implementation process is to determine the relevant chain of command, ensuring that each level in the chain understands the risk management philosophy and subscribes to it. Implementation also requires putting decision making systems in place to adequately manage risk. In this vein, Arthur Anderson (1998, p.10) focuses on the importance of risk controls and management information systems: "Risk control is ... the independent identification measurement, monitoring and reporting of risk, returns and capital utilization ... The quality of management information systems is central to management's ability to assess both business performance and risk management effectiveness."

In financial firms, monitoring of risk management is intimately connected with value at risk (VaR) calculations. The practical experience of a number of financial firms indicates the importance of adequate monitoring. For example, Jorion (1997) argues forcefully that the Orange County bankruptcy was due to inadequate monitoring. For non-financial firms, there have been some efforts to apply value at risk techniques, e.g., Godfrey and Espinosa (1998). Others, such as Culp, Miller and Neves (1998, p.34), suggest that non-financial firms are more concerned with cash flow volatility than financial firms. In such situations, firms "are better off eschewing VaR altogether in favor of a measure of cash flow volatility." With all these competing, ad hoc approaches to risk management on the landscape, conceptual guidance is needed. To this end, it is time to take a step back and undertake a careful discussion of the rationales for corporate risk management.

Arguments Against the Use of Derivatives in Corporate Risk Management

The problem of risk management for the corporation has been well studied using techniques adapted from traditional corporate finance. In traditional corporate financial management, because managers act as agents for the owners of the firm, the common stockholders, the appropriate primary objective is to ***maximize the expected utility of the end-of-period wealth of stockholders.*** Achieving this objective is complicated by the inability of managers to observe the expected utility functions of individual shareholders. Yet, under reasonable conditions, the primary corporate objective can be reformulated as ***maximizing the long run value of the firm's common stock***. Given Keynes's observations about the formation of prices in stock markets, this objective is not without difficulties, e.g., (Poitras 1994). Proceeding on the assumption that long run common stock prices will correctly reflect firm value, the market value of the firm can be determined as the sum of the net present values (NPV) of the firm's ventures. Given this, NPV increases due to corporate risk management activities can arise from: reductions in discount rates; increases in net cash flows; and, increases in the real option value of projects.

Corporate financial managers facing exposure to, say, currency risks must address a natural question: is hedging of currency risk consistent with the primary corporate objective? A number of persuasive arguments have been made ***against*** hedging corporate foreign exchange (FX) risk and other such market risks. The arguments against hedging are specific applications of more general arguments that claim that using derivative securities for corporate financial risk management activities will not be value enhancing for common stockholders. Because such risk management activities are costly to implement, monitor and execute, firms are generally recommended to forego the use of derivative securities to manage risk. The general content of these arguments is that, in perfect markets, the role of derivative securities in the risk management policy of the firm is irrelevant to the market valuation of the common stock, e.g., Siegel and Siegel (1990, p.146-9).

There are a variety of arguments that have been advanced to attempt to demonstrate that the hedging policy or, more generally, the use of derivative securities to manage firm risks, is irrelevant, e.g., Dufey and Srinivasulu (1983), Levi and Sercu (1991). Are the arguments that belie the importance of financial risk management correct? To determine the answer to this question, it is helpful to classify the important arguments against corporate currency hedging into the following groups: Modigliani-Miller (MM) arguments; CAPM (capital asset pricing model) arguments; and, market efficiency (expected value) arguments.[15] There is a complementarity among the various irrelevance arguments. Irrelevance is demonstrated using perfect markets assumptions (see Sec. 4.1). Because, in practice, the use of derivative securities involves an expenditure of firm resources that would not be required if derivatives were not used, it is argued that the use of derivatives is impractical. The firm is better off not using derivative securities at all.

In summarizing the irrelevance arguments, it is conventional to start with the MM arguments. The gist of the MM argument is captured by Levi and Sercu (1991): "It is a well-accepted principle of finance that managers of a firm will not increase the firm's value by doing anything the shareholders of the firm can do themselves at the same or lower cost." This argument is an extension of the MM arguments from traditional corporate finance that propose the financial policies of the firm are irrelevant in determining the market value of the firm. The original MM arguments focused on demonstrating that the capital structure and dividend policies of the firm have no implications (***are irrelevant***) for the market value of the firm.[16] Value is determined by the asset side of the balance sheet. The extension to the international arena is that, as a financial decision of the firm, the use of derivative securities to implement corporate FX risk management decisions is irrelevant for the same reasons as outlined by Levi and Sercu: the market will not increase the value of the firm for engaging in practices that can be done directly by investors.

The MM irrelevance argument relies on the perfect market assumptions. Within the MM framework, violations of key assumptions can dramatically change the results. For example, when corporate taxes are admitted and tax deductibility of interest payments on the debt is allowed, then instead of debt irrelevance, the simple MM model indicates that all debt financing is the optimal method to maximize the market value of the firm. Though introducing taxes can also provide a rationale for the use of derivative securities, this type of motivation does not appear to be widespread in practice and will not be examined here (see, for example, Siegel and Siegel, (1990, p.150-1) for an illustration of when taxes could provide a motive for hedging). More importantly, the MM argument is not exempt from the implications of relaxing perfect market assumptions such as ***no bankruptcy costs***. If the market value of the firm is affected by bankruptcy risk, then by reducing the total variability of cash flow, hedging and other risk management activities can increase the market value of the firm by lowering the default premium and, thereby, lowering the discount rate in the long run NPV calculation.

Though it is possible to demonstrate selected weaknesses in the MM argument by making appeals to the differences between perfect financial markets and actual financial markets, it is not as simple to completely dismiss the MM arguments as it would appear. For example, Levi and Sercu (1991) maintain that it is not possible to argue that there are economies of scale, of various kinds, that are available when the firm, not the individual shareholder, engages in risk management activities such as hedging. Shareholders may not want the firm to engage in the types of risk management policies being pursued. In spite of the scale economies, the hedge that the firm puts on may be considerably different than the individual investor desires. While appealing, this view is subject to the same criticism, that it fails to directly account for the primary goal of management, the maximization of the (long run) market value of the firm.

Following Frank Knight, firms earn economic rents from correctly handling uncertainty. Measurable risks, that can be handled by conventional risk management techniques such as purchasing insurance, are part of the cost structure, not a source of economic value added. In many cases, firms that do not accurately handle measurable risks will, in the long run, suffer the consequences of the market place. This is all too apparent from the MGRM debacle. Yet, if risk management activities are aimed at increasing expected net cash flows then it does not follow that the firm will also be able to reduce discount rates by reducing the variability of future cash flows. Using risk management to increase expected cash flow will likely increase the firm's cash flow variability. As demonstrated by LTCM, risk management activities can be a source of economic profit, i.e., increased net cash flow, only by moving out of the realm of measurable risks and into the grayer area of uncertainty. As demonstrated in Sec. 2.1, the optimal risk management solution is composed of a risk minimizing component and a speculative component. Optimal risk management may increase bankruptcy risk to achieve a speculative gain.

Being derived using perfect market assumptions, CAPM arguments have many similarities with the MM arguments. A version of the CAPM argument can be found in Levi and Sercu's discussion of FX hedging:

> It is surprisingly common to hear it argued that hedging is a good idea because it reduces the variance of the value of an asset or liability when translated into a reference currency... Of course this rationale for hedging can be quickly dismissed when it is recognized that investors do not care about the variance of the value of an individual asset or liability, *but rather the risk the asset or liability contributes to an efficiently diversified portfolio.* That is, it is only the undiversifiable or systematic part of risk that matters, and this can be defined only in the context of an investment portfolio. (emphasis added)

The CAPM argument is based on an analysis of variance argument. Total risk is decomposed into systematic (non-diversifiable) and unsystematic (diversifiable) risk. The argument is that in an efficiently diversified portfolio, the unsystematic component will be unimportant. Because FX risk is primarily unsystematic, there are no stock price implications to hedging unsystematic risk. FX risk is not likely to be priced and, if it is, any systematic risk would be incorporated in forward exchange rates and, hence, all that hedging would do is to move the firm's stock along the security market line. Again, no benefit is obtained from risk management activities such as hedging.

To better appreciate the CAPM argument, examine the discussion of the future basis from Sec. 3.1. Does the futures price provide an unbiased prediction of the future spot price? If not, then what factors determine the difference between $F(0,T)$ and $E[S(T)]$? The CAPM provides a solution to these questions by providing an elegant solution to the relationship between $S(0)$ and $E[S(T)]$. More precisely, the CAPM requires that all assets earn a return consistent with the level of systematic risk for that asset: $E[R_{i,t}] = R_f + \beta_i (E[R_m] - R_f)$, where $E[R_{i,t}]$ is the conditional expected return on asset i, R_f is the return on the riskless asset, β_i is the measure of systematic risk and $E[R_m]$ is the expected return on the market portfolio. If the CAPM holds, it follows that if S is the price of asset i then $E[S(T)] = S(0)(1 + E[R_{it}])$. Because $F(0,T) = S(0)(1 + ic(0,T))$, the CAPM can be used to solve the future basis: $E[S(T)] - F(0,T) = S(0)(E[R_{it}] - ic(0,T))$ or $F(0,T) = E[S(T)] - S(0)(E[R_{it}] - ic(0,T))$. (Where relevant, it is possible to reduce this result by netting out the pecuniary carry returns.) As a consequence, the CAPM implies that $E[S(T)]$ will typically be higher than $F(0,T)$.

To extend this result to firm valuation, make the conventional assumption that the value of the firm is determined by the discounted value of the expected net cash flows generated by the firm. Now, for simplicity, consider the case of the market value of an all-equity financed silver mining firm. The firm has not yet started production, but exploratory drilling results estimate an ore body of 10 million ounces (with no other economically recoverable by-product ores). If it takes one year to recover this ore and the firm does ***not hedge***, then the market value of the firm's output will be the 10 million ounces times the expected price of the silver in one year's time. Assuming without loss of generality that there are also 10 million shares of stock outstanding, then the share price of the company will be given by: $S_U(0) = E[S(1)]/(1 + E[R_{s,1}])$, where the discount rate is determined by the CAPM.

To demonstrate that hedging is irrelevant to the market value of the firm, consider the value of a share if the firm decides to *fully hedge* and uncertainties in production are ignored: $S_H(0) = F(0,1)/(1 + R_f)$. Because the output price has been locked in by hedging, the discount rate will be lower than for the unhedged firm. The CAPM argument involves making sufficient assumptions to ensure $S_u(0) = S_H(0)$. This requires $E[S(T)]$ to be higher that than $F(0,T)$ by precisely the amount needed to offset the difference in the discount rates. Using the notation from Sec. 4.1 and observing that silver is typically near full carry, then it is possible to assume that $R_f = r(0,1)$. Similarly, because this company is a pure silver mining play, the discount rate for this company can be assumed to be the same as the CAPM discount rate for holding spot silver, $E[R_{s,1}] = E[R_{i,1}]$. Given this, $S_u(0) = S_H(0) = S(0)$ and the CAPM argument is validated. The market price of the firm's stock will be equal to the current spot price of silver. This is consistent with the company being a pure play on the deterministic stock of silver that will be marketed in one years time.

Survey evidence on motivations for firm hedging presented in Sec. 1.4 reveals that the most important determinant of firm hedging activity was the desire to reduce the volatility of firm cash flows. By exhibiting less volatile cash flows firms can potentially lower the cost of capital. However, the CAPM argument maintains that this motivation is fictional. ***Any decrease in the cost of capital from hedging is exactly offset by a decrease in the expected cash flows of the firm***. This follows from the equilibrium underlying the determination of forward prices. The forward price will differ from the expected spot price by just the amount needed to offset the gain associated with the reduction in the cost of capital. Yet, as the discussion in Sec. 3.3 illustrates, the CAPM argument imposes unrealistic empirical conditions on the futures basis. In addition, the full hedge assumption implies that the motivation for hedging is to only to reduce the volatility; no attempt is made to identify the optimal hedge and to examine the associated valuation implications of employing such a hedge.

Much like the MM argument, the CAPM argument can be criticized by demonstrating that relaxation in the underlying assumptions substantively changes the results. Under the perfect markets assumptions required for the CAPM to hold, the CAPM argument could have considerable validity. However, the CAPM assumptions are relatively severe. Of particular interest are the assumptions that would make total instead of systematic variability a concern. ***No bankruptcy costs*** is, again, a key CAPM assumption. The basic hedging framework explicitly identifies risk afford ability as an essential element in establishing a hedging program. While there are numerous possible examples, two that could be used are: the B.C. Hydro case where the combination of US$ borrowing and C$ revenues meant that a significant exchange rate change could more than eliminate firm capital; and, the (now defunct) Vancouver Grizzlies who earned the bulk of revenues in C$ but were obliged to incur expenses, including but not limited to salaries, in US$. Among other things, the presence of bankruptcy costs can affect a firm's cost of and accessibility to capital.

Another difficulty with the CAPM argument is the significant restrictions on diversification associated with real assets. The CAPM framework was developed to explain optimal portfolio selection, where the assets involved are highly liquid and divisible. However, real assets may be "lumpy" and not easily divisible. The alternative assets needed to adequately diversify may not be available for purchase. Where such assets are available, capital constraints may prevent their acquisition. In short, it may not be possible to construct "an efficiently diversified portfolio". Again, factors such as the lumpiness of assets and the inability to adequately diversify, means there is an element of uncertainty in business decisions that cannot be reduced to the type of measurable risk argument that underlies the CAPM.

Another group of arguments against the use of derivatives in risk management can be classified as ‘expected return’ arguments. These arguments make the empirical observation that risk management activities involving derivatives will be, on average, a negative expected value operation, once the fixed costs of the risk management program are considered. To see this, consider the case of a futures hedge. Assume that forward prices are unbiased predictors of future spot rates, then a policy of continuous hedging will just reflect back the price changes. Sometimes hedges will make money, sometimes hedges will lose money. On balance, the gains and the losses will net out and the hedged position will have the same expected value as the unhedged position. In effect, the expected values of the returns on the hedged and unhedged positions will be equal.

As it turns out, this argument may have validity in both practical and theoretical settings. For example, due to the costs associated with initiating, monitoring and executing a derivatives program, practical concerns dictate that such activities are unnecessary for many firms. As Cummins et al. (1998, p.34) conclude: "whatever the underlying value-related motive for risk management, the existence of fixed costs associated with using derivative instruments may make it more likely that only larger firms, with the resources to pay these large up-front costs, will manage risk through derivatives trading." Theoretically, in perfect markets with risk neutral participants, traders are indifferent between risky prospects with the same expected values, independent of the variance of the prospect. However, the theoretical assumptions required to generate some results are associated with models that do not have a well defined general equilibrium. In a world with risk averse participants, the dependence solely on expected value omits one of the primary reasons for hedging: controlling the *total variability* of the firm's cash flows due to changes in financial prices such as exchange rates, interest rates, commodities and equities. Following Perold and Schulman (1988), by not hedging the firm is foregoing a "free lunch" opportunity. A practical objective of heding could be to cause fluctuations in the firm's market value to be due solely to changes in the firm's business activities, not (random) changes in financial prices. As in the MM and CAPM arguments, this could enhance long run share prices by reducing bankruptcy cost thereby reducing the cost of capital.

It could be further argued that the expected return arguments against using derivative securities for risk management do not fully develop the implications of observing that such activities have zero expected value. For example, consider the statement that 'it is just as likely to be surprised on a foreign exchange hedge as on the cash position'. Given the additional costs associated with having a derivatives program, an expected return argument would conclude this is a reason for not hedging. However, situations when the hedge loses money, i.e., provides an unanticipated 'surprise', will likely be situations where a windfall gain would have occurred as a result of favorable exchange rate changes. In order to reduce volatility, the hedge trades off both downside and upside changes in the cash prices being hedged. This allows the firm to concentrate on production problems without having to worry about complications related to unexpected changes in financial prices. The important question is whether, in this situation, firms that successfully pursue active financial risk management will have a substantive competitive advantage over firms that are continuously fully hedged (or do not hedge at all).

Arguments in Favor of Using Derivatives in Corporate Risk Management

Compared to the limited number of arguments against currency hedging, there is a daunting number of arguments in favor of using derivative securities for corporate risk management. Some of these arguments have already been mentioned in passing, e.g., in Sec. 1.3. There are various ways these theories could be organized. For example, based on empirical evidence from risk management practices in the gold mining industry, Tufano (1996, p.1099) distinguishes between theories which focus on managerial characteristics such as stock option ownership by management, e.g., Smith and Stulz (1985), and "theories that explain risk management as a means to reduce the costs of financial distress, to break the firm's dependence on external financing, or to reduce expected taxes." This is a substantial break from the conventional corporate finance classifications which, based on the elements of the NPV calculation, identify theories supporting derivative use with reductions in the discount rate, increases in expected net cash flow or increases in real option value of projects. If Tufano is correct, explanations for risk management which sustain share value-maximization may be a blind for managers engaging in activities which may be value reducing for shareholders. In opposition, Cummins et al. (1998) suggest that Tufano's empirical results may not apply in other industries.

In examining MM and CAPM theories regarding the use of derivatives in corporate risk management, it was pointed out that the presence of bankruptcy costs would undermine both arguments. Yet, there is not strong empirical evidence to suggest that firms with a high probability of financial distress engage in higher levels of risk management compared to firms with lower levels. For example, while Wall and Pringle (1989) find the use of swaps is more likely for firms with lower credit ratings than for higher rated on-financial firms, more recently, Geczy et al. (1997), Mian (1996) and Nance, Smith and Smithson (1993) all are unable to find a link between derivatives usage and the capital structure of non-financial firms. Financial firms provide similar mixed results (Cummins et al. 1998, p. 38). The conclusion to be drawn from this is that, while financial distress may be a factor for some firms to use derivatives, this explanation does not seem to stand out in the data.

It is difficult, if not unwise, to abandon the notion the managers act to maximize long run share value, in favor of model of derivative usage which focuses on managerial incentives. Perhaps, as Copeland and Copeland (1999) argue, the financial distress hypothesis needs to be reformulated. Yet, it does seem that, if the MM and CAPM arguments against corporate use of derivatives are to be voided, another perfect markets assumption will have to be altered. Along this line, Froot et al.(1994, p.98) maintain that by stabilizing cash flows, firms can use derivative securities to align the internal supply and demand of funds. By stabilizing cash flows, corporate risk management permits the firm to participate in investment opportunities that may arise at inopportune times: “Managers who adopt our approach should ask themselves two questions: How sensitive are cash flows to risk variables such as exchange rates, commodity prices, and interest rates? and How sensitive are investment opportunities to those risk variables? The answers will help managers understand whether the supply of funds or the demand for funds are naturally aligned or whether they can be better aligned through risk management.”

According to Froot et al., by stabilizing cash flow, risk management permits firms to undertake some positive NPV projects that would be avoided in the absence of such activities. This hypothesis could be targeted at any of the elements in the NPV calculation. By stabilizing cash flows, the firm is better able to access sources of internal financing, which is cheaper to use than external financing. This will lower the discount rate. By using derivative securities to avoid under investment, risk management increases expected future cash flow by increasing the number of positive NPV projects. Finally, stabilizing cash flows can permit the firm to exercise real options, such as the development option, thereby increasing the value of these options to shareholders. Gay and Nam (1998) provide evidence on the under-investment hypothesis. Using a sample of 325 firms using derivatives combined with 161 firms not using derivatives, it is estimated that “firms with enhanced investment opportunities, lower liquidity, and low correlation between investment expenditures and internally generated cash flows tend to be more likely users of derivatives” (Copeland and Copeland 1999, p.74).

Other promising explanations for corporate risk management have been advanced. Key factors in these explanations include: the ownership structure of the firm (Smith 1995); resolving conflict between firms by enhancing the contracting relationship between firms (Pennings and Leuthold 2000); risk shifting with the firm (Smith 1995); and, lowering expected tax costs (Smith and Stulz 1985). Ownership structure can be related to both managerial incentives and shareholder wealth maximization. “Managers whose human capital and wealth are poorly diversified strongly prefer to reduce the risk to which they are exposed. If managers judge that it will be less costly (to them) for the firm to manage this risk than to manage it on their own account, they will direct their firms to engage in risk management” (Tufano 1996, p. 1109). Similarly, concentrated ownership, whether in the hands of management or not, likely means that owners do not have well-diversified portfolios, again providing an incentive for the firm to engage in risk management. One argument against the CAPM case against risk management was that assets could be lumpy, it was not easy to hold an efficiently diversified portfolio.

Tufano on Risk Management by Gold Mining Companies

Tufano (1996) is a useful source on risk management practices of gold mining companies:

Most of the 48 North American gold mines studied ... are not well diversified. Risk management strategies can be implemented using explicit derivative transactions, such as the forward sale of gold, or they can be combined with financing activities. For example, in borrowing via a gold or bullion loan, a mining firm combines dollar-base financing with a forward sale of gold.

Hedging instruments include over-the-counter forward sales of gold, exchange-traded futures contracts, gold or bullion loans, gold swaps, and spot deferred contracts (which are economically similar to rolling forward contracts.) Firms wishing to establish *insurance* strategies can use either exchange-traded or over-the-counter gold put options, or can dynamically replicate puts by trading forwards and futures...

The rich menu of risk management instruments gives firms an ability to customize their gold price exposure, and firms have embraced risk management. For example, over four years American Barrick Resources Corporation used put and call options, gold warrants, bullion loans, forward sales, spot deferred contracts, and customized gold-linked equity financing as part of its risk management program.

A Simple Guide to Designing a Risk Management Philosophy

The theoretical rationales for corporate risk management using derivatives provide foundation for the discussion of practical issues involved in engaging in such risk management decisions. Translating academic discussion into practice is facilitated by detailing some heuristic guidelines. The development of a risk management philosophy for a specific firm is an essential step in developing an effective risk management program. This process can be motivated by identifying a number of basic considerations to be addressed in order to determine the type of derivatives trading program to be undertaken, e.g., Powers and Vogel (1983). It is in formulating answers to the various questions that the elements of a risk management policy become apparent. To this end, consider the following sequential list of questions that are of particular relevance for a firm considering the implementation of derivatives trading program:

What are the firm's aggregate and specific risk exposures? This step requires data and analysis. It is essential to make detailed calculations of the possible losses if no derivatives trades are made. Adjustment for expectations about future movement in prices may be incorporated, producing a range of scenarios. Unfortunately, in many cases, the calculations required are not obvious. For example, Gallo Wines was for many years a US wine producer that produced and sold almost all of its output in the US. However, even though almost all the cash flows for Gallo were denominated in US dollars, the profitability of Gallo depended fundamentally on the price of competing wines from other countries. Hence, Gallo had a considerable risk exposure to changes in the value of US dollar. Another example is MGRM. In this case, the risk exposure was to changes in the value of the long term forward oil byproduct delivery contracts. Without a traded market price for these contracts, it is not possible to provide an object estimate of a change in value when the spot price for the byproducts change, say, one dollar.

The problem of determining risk exposures leads to the fundamental notions of ***economic exposure*** and ***accounting exposure,*** e.g., O'Brien (1997), that underpin the optimal hedging strategy and the transactions hedging strategies discussed in Chapter 6. Accounting exposure measures on a transaction by transaction basis. This approach is reflected in conventional textbook presentations of a risk management involving derivatives that assumes that there is only one transaction of interest. For example, a US company books a sale in yen to be settled in three months. There is now an accounting exposure equal in size to the anticipated yen to dollar spot transaction that will take place in three months. The method of determining risk exposures can have severe consequences. For example, MGRM identified market risk using an accounting exposure. As a consequence, the size of MGRM hedge position was set to be approximately equal to the number of barrels in the long term delivery contracts.

In many situations, accounting exposure is a useful measure. For example, when basis risk is low, the number of derivative transactions is small and the objective is to minimize the variance of the firm's net cash flow, then transactions hedges have desirable properties (see Chapter 6). Yet, in many cases there are numerous transactions that contribute to the risk exposure. These transactions can involve a large number of financial and commodity prices. Some process of aggregating the risk exposures is required. Economic exposure measures attempt to assess the impact of a specific financial or commodity price on the firm's net cash flow. An important example is the economic exposure of a financial institution to changes in the level and term structure of interest rates. Recognition of the difficulty in measuring risk exposure for financial institutions led to the development of measures such as the ***duration gap*** which is used in asset and liability management. Derivative securities can such as interest rate swaps play a crucial role in allowing financial institutions to significantly reduce the duration gap. A substantial portion of interest rate swap market activity is derived from duration gap managing institutions, both financial and non-financial.

Are the risks affordable? This involves comparing the calculated risk exposure with various measures of the capital invested in the business, also taking into account various possible remedies already in place, such as insurance policies and natural hedges. If the risk is affordable, the arguments in favor of implementing a derivatives trading program are substantively different than if the risk is sufficient to cause financial distress or bankruptcy. If the risk is not affordable, there can be real gains associated with implementing a derivatives trading program, such as a lower achievable cost of capital due to a lower probability of bankruptcy. It is also possible that competitive factors may impact whether a risk is affordable. For example, changes in jet fuel prices are an important component of cost variability in the airline industry. Firms that hedge jet fuel prices may be able to gain market share from those that do not hedge due, say, to being able to quote lower prices when jet fuel prices are high and profits are being squeezed.

In some situations the decision to trade derivatives may be imposed by lenders. An example of this occurs with some Australian gold mines that have been required by the banks making them loans to implement a hedging program as a condition of being granted credit. Such hedging programs would dramatically alter the risk management philosophy of a firm which maintained that shareholders wanted the share price to be fully exposed to changes in the gold price. Affordable for management is not necessarily consistent with affordable for shareholders which is not necessarily consistent with affordable for bond holders.

The risk afford ability issue is also difficult to determine for government enterprises where, ultimately, afford ability is determined by the ability to raise general tax revenue, borrow against future tax revenue or levy increased user rates. For example, some Canadian electric utilities, such as BC Hydro, have a large portion of debt denominated in US dollars while the bulk of cash flows are in Canadian dollars. Large, adverse changes in the exchange rate that would be sufficient to eliminate the net asset value of a private company, would not have the same impact on the government-owned utility. It may be that the government, in some other enterprise, has a corresponding amount of US dollar denominated assets. As such, the government's aggregate balance sheet would be fully hedged, without the need to manage the specific exposure at BC Hydro.

Can the risks be hedged? This is the problem of hedge design, a topic that is the central concern of chapter 6. There may be a variety of possible hedging techniques that have to be considered. An important practical concern is whether there are derivative contracts available that qualify as feasible hedging instruments. In certain cases, forward contracts will be available that allow the cash position to be matched with the commodity underlying the forward contract. For example, money center banks can use forward exchange contracts and foreign exchange swap contracts to directly manage currency risk. In other cases, no forward contracting method is available and a ***cross hedge*** is required. In other cases, the pricing on the forward contract may be considered to be expensive relative to doing a cross hedge. For example, an airline may undertake a cross hedge using NYMEX heating oil futures in lieu of doing a short dated jet fuel swap in the OTC market because the cost of the swap is deemed to be too expensive relative to doing the NYMEX hedge and absorbing the basis risk.

Cross hedging involves managing a specific commodity risk using a derivative that is written for a commodity which differs from the cash commodity. For example, a copper scrap dealer can cross hedge using copper futures contracts that feature copper cathodes as the deliverable commodity or a corporate bond issuer can hedge the issue cost using the Tnote futures contract. Cross hedges can sometimes involve quite different commodities, such as hedging brass scrap with a combination of aluminum, copper and lead futures/forward contracts. Cross hedging raises questions about the appropriate size to derivative position relative to the cash position (an optimal hedging problem) and whether the hedge will be effective.

What are the basis relationships? Examination of basis risk is a key part of the process in deciding whether a specific risk can be hedged. This is often a situation specific problem. Some fundamentals regarding basis variation are discussed later in this chapter 3. Information on basis relationships is an essential element in determining the size and type of the hedge position to be initiated. Basis variation is an important impediment to implementing the transactions hedge described in Chapter 6.

What are the costs of hedging? This could start with calculation of execution and transaction costs: bid/ask spreads, commissions, possible interest losses on margin and administrative expenses to initiate and monitor trades. Except where the contracting process permits sufficiently precise specification, there will be an element of basis changes that needs to be calculated. Such changes appear to have come as a surprise to MGRM, for example. For substantial hedging programs, there can also be significant managerial costs in terms of the time required to monitor hedging operations. There will typically be considerable variation in the specific costs associated with various potential hedging instruments and programs. For example, firms aiming to use complicated risk management products, such as exotic derivatives, need to provide internal checks for prices derived from proprietary pricing models.

What are the tax and accounting implications of the hedge? The relevant issues involved here are discussed in other sources. These issues are not incidental and will have to be determined in order to precisely calculate the costs of hedging. In particular, the introduction of FAS 133 raises a host of questions and queries that lie outside the confines of the present inquiry.

By design, this general framework for design a risk management philosophy cannot deal with all important issues that may arise in specifying the appropriate risk management/hedging program. It is only a guide to the appropriate mind set required to structure the risk management process.

Measuring Corporate Economic Currency Exposure

The increasing globalization of markets has put increasing pressure on management to determine the appropriate method of handling exposure to currency fluctuations. As discussed previously, there are two general approaches to measuring corporate currency exposure that can be identified: ***accounting*** *exposure* and ***economic*** *exposure*. The first of these is concerned with the implications of accounting rules, contained primarily in FASB-5 and FASB-52, that deal with the handling of accounting items which are denominated in foreign currency. When basis relationships permit, accounting exposure management typically leads to the use of transactions hedges (see Sec. 6.1). While useful within an accounting context, there are many forms of corporate currency exposure that are not captured using this measure.[17] For handling economic exposure, optimal hedging methods are needed.

Economic exposure is a broader concept that measures "the extent to which the value of the firm-- as measured by the present value of its expected future cash flows-- will change when exchange rates change." (Shapiro 1992, p.224) This occurs not only because components of firm cash flow's are directly denominated in foreign currency, but also because the relative competitiveness of the firm can be affected. In order to identify how this happens, "the focus must be not on nominal exchange rate changes, but instead on changes in the purchasing power of one currency relative to another." (p.225) This leads to the notion of "real" as opposed to nominal exchange rate change. It is changes in *real* exchange rates that produce the conventional economic result that exchange rate increases (decreases) will increase (decrease) imports and decrease (increase) exports. However, in the case of the multinational firm, a number of further complications have to be introduced.

The real exchange rate is an implication of PPP: "...if changes in the nominal exchange rate are fully offset by changes in the relative price levels between the two countries, then the real exchange rate remains unchanged. Alternatively, a change in the real exchange rate is equivalent to a deviation from PPP." (Shapiro 1992, p.155) Being based on PPP, the real exchange rate can be used to identify substantive changes in foreign currency values. In other words, if the economic implications of nominal exchange rate changes are offset by corresponding changes in price levels, then the real exchange rate is unchanged and, presumably, there is no incentive to change economic behaviour.

This simplified model ignores various complications such as financial obligations that are fixed in nominal terms, this will include unhedged fixed rate debt, sales and labor contracts and other types of receipts and disbursements denominated in foreign currency. In the absence of indexing, these factors cannot be readjusted when unanticipated changes in the *nominal* exchange rate occur. Hence, it is possible for the real exchange rate to be unchanged and still have substantive changes in economic behaviour. Similarly, it is possible for the nominal exchange rate to be unchanged and for changes in relative inflation rates to occur that will have substantive economic implications. Shapiro (1992, p.228-9) provides an illustration of this happening 1979-82 in Chile where a government attempt to fix the value of the Chilean peso led to a significant erosion in international competitiveness that had a disastrous impact on the Chilean economy.

A useful Canadian example of how economic currency exposure can affect firm profitability is the hotels and related businesses at the Whistler/Blackcomb ski resort in B.C.[18] Even though virtually all revenues and costs are in Canadian dollars, revenues are indirectly dependent on competition from overseas ski resorts. In effect, Whistler/Blackcomb is operating in a global market for skiing and other vacation services. Changes in the Canadian dollar will change the relative value of overseas ski vacations, for both domestic and foreign vacationers. More generally, even though a firm does not have any direct foreign currency exposure, the presence of foreign competition in either the input or output market means that there could be substantial economic currency exposure.

Another Canadian example of corporate currency exposure is provided by the Canadian mining industry. Because the price of metals is set in global markets in US dollars, mining company US dollar revenues will not be affected by changes in the Canadian dollar, assuming the price of oil sales in Canadian dollars is allowed to change to reflect the US dollar price. While US dollar revenues will not change, changes in the value of the Canadian dollar will alter the US dollar cost of Canadian labor and supplies used in the production of metals. This type of situation occurs in many other Canadian cases, where the product being produced is being priced on international market in terms of US dollars. This is the case with the grains such as wheat and energy products such as oil, natural gas and hydro electricity.

As a final example of corporate foreign exchange exposure consider Toyota, an automobile manufacturer where both revenues and costs are affected by exchange rate changes. On the revenue side, Toyota sells the bulk of its production overseas, concentrating on the US. Changes in the value of the yen will force a pricing policy decision. For example, in the face of an appreciation of the yen, to maintain market share the US dollar price has to be held constant, reducing yen revenues because the yen price per unit has fallen. If the yen price is held constant, market share will be reduced because of higher US dollar prices. On the cost side, Toyota is a purchaser of commodities required in car production that are priced on international markets. Changes in costs will tend to offset changes in revenues, though not one-for-one. In the case of Toyota, because such a large component of revenues is in US dollars while only a relatively small portion of costs is not yen determined, the impact of appreciation in the yen is negative. Hence, there are numerous ways in which currency exposure can impact a given firm.

Natural Hedging of Corporate Currency Exposure

While there are various strategies available for managing corporate currency exposure, it is possible to distinguish between two general types of strategies. One type is associated with traditional derivative security hedging techniques, suitable for nominal contracts stated in a foreign currency. Applications of these techniques include the important area of international asset/liability management, where relatively predictable cash flows originate from foreign financial assets. The techniques of swaps, futures and options are well developed in this area. The other general type of strategy for managing corporate currency exposure involves natural hedges that are dependent on multinational firm management decisions. These techniques apply to corporate cash flows that are relatively indeterminate, consistent with cash flows that originate from many real assets.

The other group of strategies for managing corporate currency exposure involve assessment of the ***competitive exposures*** that originate from inherent differences in firm competitiveness due to costs and revenues being denominated in different currencies. Currency exposure management in these cases will typically involve adjustments to be made to operating procedures, encompassing marketing, production and capital structure decisions. By design, this will require integrated, long term decision making. Natural hedging techniques are inherent in these types of strategies. This is an essential, if not always well understood, point. Many risk management situations, such as those faced by financial institutions in dealing with interest rate risk, can be most effectively managed using natural hedges.

Because competitive conditions can be altered by a real exchange rate change (Luehrman 1990), the firm must attempt to anticipate such changes and decide whether a given change will be transitory or persistent (permanent). For example, a Japanese car maker faced with an increase in the nominal $/Yen exchange rate, not matched by corresponding price level increases, must decide whether to increase dollar prices, attempting to sustain the yen price, or to hold dollar prices constant, thereby reducing the yen price. If the nominal exchange rate change was anticipated to be matched by price level adjustments in the near term, then the car manufacturer may be willing to hold the dollar price constant in order to maintain market share. This loss of income would have to be balanced against the cost of recovering market share when the real exchange rate is restored. On the other hand, if the real exchange rate change was anticipated to be persistent, then competitive conditions would require undertaking various adjustments such as lowering the yen cost of production. For example, this could be done by sourcing production of automobiles to the US and other countries where real production costs will be lower. There are various other possibilities.

Product strategy provides one potential method for adjusting to changes in currency exposure. Faced with long term appreciations in their real domestic currencies relative to the dollar, both VW and the major Japanese car producers have had to adjust the nature of the product being sold in the US. In effect, real exchange rate changes made competition at the low end of the market unprofitable. As a result, these companies have made long term product adjustments by offering higher-priced automobiles targeted at middle to upper middle income consumers. In contrast to persistent real exchange rate changes, temporary exchange rate changes will usually not require substantial adjustments to product offerings. However, temporary depreciations may provide timing opportunities for firms seeking to penetrate foreign markets. This is an important point because the high fixed startup costs associated with overseas expansion are often incurred in the initial stages of establishing a market presence. Favourable, if temporary, exchange rate changes can partially offset these costs.

Perhaps the most widely recognized method for multinational firms to manage currency risks is to create natural hedges through appropriate ***plant location*** and input purchase decisions. Firms that have similar production facilities in areas with different currencies can, potentially, shift production to plants where production is least expensive. Where it is not possible to establish production facilities in the appropriate locales, then a similar result can be achieved by spreading sources for inputs across countries in different currency areas. In practice, the benefits associated with multinational sourcing and production facilities have to be balanced against the costs associated with plant redundancy and loss of economies of scale. In a corporate context, this requires managerial decision makers to incorporate forecasts of exchange rate changes into company strategies. Hence, there is an element of active management in adjusting to currency exposure (in keeping with the results in Sec. 2.1). In addition to the natural hedges provided by plant shifting and alternative sourcing, it is also possible to react to currency related changes in competitive conditions in a more traditional fashion, i.e., by raising domestic productivity.

The final important method for corporations to use natural hedges to manage currency exposure is by adjusting the ***capital structure*** of the firm. Conventionally, this involves taking advantage of the natural hedge provided by financing real assets with foreign debt. Where the cash flows of the real assets have an identifiable currency exposure, either because of foreign competition or dependence on foreign markets for inputs or sales, changes in operating cash flows arising from exchange rate changes will be met by offsetting changes in debt service costs. As with any type of hedging situation, there will be situations when the hedge position is unprofitable, i.e., where the domestic currency value of the foreign borrowing increases. In these cases, it would be desirable to finance real assets with domestic debt. Because it may not be possible to adjust borrowing programs to keep pace with the numerous exchange rate changes, once again the natural hedge decision depends on active management of currency exposure to achieve the highest return. If management is not able to forecast or has a high degree of risk aversion, the optimal solution will be to establish a natural hedge that matches the foreign currency exposure.

Figure 1.5 **Questions Relevant to Formulating Natural Hedging Strategies**

1. What is the foreign/domestic breakdown of sales?
2. Are the company's key competitors foreign or domestic?
3. What is the short and long run price elasticity of demand for firm output?
4. What is the foreign/domestic breakdown of production activities?
5. What is the foreign/domestic breakdown of input sources?
6. What currency is used to determine the firm's inputs and outputs?

In the face of the various complications, Figure1.5 provides a heuristic framework of key questions designed to identify exchange risk from Shapiro (1992). The answers provided to these questions can be used to guide the implementation of natural hedging strategies. Luehrman (1990) and, more recently, Oxelheim and Wihlborg (1997) provides a theoretical development of these issues. Against this heuristic background, attempts have been made to provide a more formal approach to measuring currency exposure. Adler and Dumas (1984), Shapiro (1975) and others propose the use of regression analysis to identify the correlation between changes in the nominal exchange rate and the domestic currency value of the firm's cash flows. In effect, changes in firm cash flow (in domestic currency) are regressed on changes in nominal exchange rates.[19] The resulting estimated slope coefficient can be used as a proxy for currency exposure. While appealing, this approach suffers from a number of potential shortcomings. For example, the use of historical data for the regression requires that the nature of the firm has not changed substantively over the sample period, e.g., due to mergers or significant changes in unhedged foreign debt issues. Similarly, there should be no anticipated changes in the nature of the firm over the decision-making period for which the regression information will be used. In addition to these practical problems, the lead-lag relationship that is often associated with a currency change affecting firm cash flows may complicate identification of the appropriate regression equation. Various other problems also have to be addressed for this approach to be correctly implemented.

Purchasing Power Parity Arguments

Purchasing power parity plays a key role in decisions for naturally hedging currency risk. Though the roots of Purchasing Power Parity (PPP) can be found in Adam Smith and early 19th century classical political economy, the PPP theory is usually credited to Gustav Cassels, writing in the 1920's. The earliest versions of PPP took the form of the Law of One Price: assume a one good world with no transactions or transportation costs, then the price of that good denominated in different currencies will sell at the same price:

$$P_t^* \, S_t = P_t \qquad \Rightarrow \qquad S_t = \frac{P_t}{P_t^*}$$

where P* and P a

re the foreign and domestic prices of the good with S being the spot exchange rate.

Extending the Law of One Price using price indices instead of individual prices is known as ***Absolute Purchasing Power Parity*** (APPP). Even in the unlikely event that the Law of One Price holds for each good individually, the APPP extension may be invalid if the index weights are not the same for both economies. The problem of traded and untraded goods also creates significant difficulties. Nevertheless, ignoring the various possible problems with APPP, substituting price levels p* and p into the Law of One Price and taking logs produces:

$$\ln S = \ln p - \ln p^* \quad \Rightarrow \quad \frac{1}{S}\frac{dS}{dt} = \frac{1}{p}\frac{dp}{dt} - \frac{1}{p^*}\frac{dp^*}{dt} \quad \Rightarrow \quad \dot{S} = \dot{p} - \dot{p}^*$$

Hence, APPP holds that foreign exchange rate changes are determined by the difference between foreign and domestic inflation rates. One implication of this appealing interpretation of exchange rate changes is that predicting domestic and foreign inflation rates will permit exchange rate changes to be forecasted accurately.

A more popular form for PPP to take is ***Relative Purchasing Power Parity*** (RPPP). This is the version used to define the real exchange rate as the nominal exchange rate adjusted for changes in the relative purchasing power of each currency since some base period. In a one period framework, the relative form of the PPP condition can be expressed:

$$\frac{S_{t+1}}{S_t} = \frac{p_{t+1} \,/\, p^*_{t+1}}{p_t \,/\, p^*_t} = \frac{1 + \dot{p}_{t,\,t+1}}{1 + \dot{p}^*_{t,\,t+1}} \qquad \Rightarrow \qquad S_t = S_{t+1} \frac{1 + \dot{p}^*_{t,\,t+1}}{1 + \dot{p}_{t,\,t+1}}$$

where p is the appropriate price level index, $\dot{p}$ the inflation rate and * denotes a foreign value. The real exchange rate (s_t) notion is an attempt to convert observed exchange rates back to some base period. Starting from some base year where $S_0 = s_0$, then:

$$s_1 = S_1 \frac{1 + \dot{p}^*_{0,\,1}}{1 + \dot{p}_{0,\,1}} \qquad \Rightarrow \qquad s_n = S_n \frac{1 + \dot{p}^*_{0,\,n}}{1 + \dot{p}_{0,\,n}}$$

The multiperiod form of s_t involves compounding the inflation term over the time between the selected base year and the desired date. Some evidence on the historical behavior of nominal and real foreign exchange rates is given Shapiro (1999,p.217). Casual examination of the empirical evidence reveals that real exchange rates for many currencies do deviate significantly from the PPP requirement that the real exchange rate is relatively constant over time.

The basic approach of the PPP arguments is to attack the notion of exchange risk. This follows from the PPP implication that, in the ***long run,*** exchange rate changes will offset price level changes.[20] Take the example of a Canadian sugar refiner selling output in C$ but purchasing sugar in US$. The PPP argument indicates that a deterioration in the FX rate will be compensated for in price level increases. If, say, the US$/C$ increased by 50%, (C$/US$ falls) causing the cost of raw sugar inputs to increase proportionately, then PPP dictates that the Canadian inflation rate will be such that the price of refined sugar in Canada will increase to completely offset the Canadian dollar increase in input costs. When appropriate assumptions are satisfied, PPP holds and the real foreign exchange rate is unchanged. In this case, there are no real implications to nominal foreign exchange rate changes.

The argument that PPP holds and, hence, corporate hedging is unnecessary has a number of obvious and not-so-obvious shortcomings. A list of these would include:

a) ***Empirical applicability of PPP***: There is a sizable literature on the empirical validity of PPP, e.g., Corbae and Ouliaris (1988). The long lead-lag time period for the relationship to hold makes PPP inconsistent with the typical types of business decision time frames; the applicability of PPP to tradeables more so than non-tradeables creates complications if the hedger is interested in non-tradeables.

b) The slippage created between the price index that underlies PPP and the specific prices that are of interest to the hedger. It is relative, not aggregate, prices that are of interest.

c) The presence of financial and operating contracts that are fixed in nominal terms, i.e., cash flows that do not adjust when the aggregate price level changes.

In the context of an international firm, Shapiro (1984) has demonstrated that in the face of deviations from PPP (changes in real foreign exchange rates) a combination of forward exchange contracts, nominal debt and fixed price sales are required in order to hedge against currency risk (composed of inflation and real exchange rate risk) and relative price risk.

NOTES

1. There are a number of pitfalls in the practical interpretation of the spread trade profit function. For example, if the t=0 difference between the deferred and the nearby prices were negative, then profitability for the short nearby/long deferred spread would require that the absolute difference between the prices narrow.

2 In addition to the practical situations already listed, there are numerous other situations where the size of the risky asset position is fixed. Insurance decisions provide many cases, such as those involving fire or earthquake insurance on a house or how much crop insurance to purchase for an apple orchard. Other examples include the purchase of currency put options to protect against changes in exchange rates by a company bidding on a contract denominated in a foreign currency or a metals refinery concerned about declining prices for scrap already in inventory. In most practical situations, the decision about how many put options to purchase is unbundled from the real asset decision. In other words, the hedging decision is separated from the production decision, e.g., Feder et al. (1980).

3 In the domestic asset investment problem, it is typical to assume that $AP_t = AP_{t+1}$ is the initial value of asset units, e.g., shares of stock in the initial investment, making the problem somewhat simpler.

4 Extending the analysis to situations where $x < 0$ does changes the underlying conditions of the decision problem somewhat. For example, optimal solutions would involve the sale of put options. In many practical situations, e.g., crop insurance, this would not be possible. In some situations, the purchase of call options could be a feasible alternative. In addition, when the shape of the return distribution is negatively skewed and $x > 0$, this leads immediately to a negatively skewed distribution for terminal wealth. This situation changes when $x < 0$.

5 It is also possible to specify the put option using futures prices. However, because this involves the introduction of basis considerations, this complicates the analysis. Because exchange traded options are often written using futures prices, construction using futures prices is in some cases potentially more realistic. The assumption that the option is at-the-money is not restrictive and is used only for notational convenience.

6 It is simple to extend the profit function for the yield insurance case to cover revenue insurance. For revenue insurance, instead of two random variables associated with price and yield interacting to determine revenue (PY), there is only one random variable for revenue (R). The put option decision problem involves determining (Q_R/A), the fraction of A covered or insured with the revenue put option. Substitution of $\underline{R} = \{\underline{R}\, A\}/C(A)$

motivates the relevant profit function.

7 The axiomatic approach to choice under uncertainty has produced a considerable number of studies. Accessible and brief overviews are available in various sources, e.g., Henderson and Quandt (1980, Sec. 3.8), Layard and Walters (1979, ch.13).

8 Further discussion of issues related to the general properties of a Taylor series expansion for approximating a general expected utility function can be found in Loistl (1976). Hassett et al. (1985) examine specific types of problems with the Taylor series that arise where skewness is involved. Brockett and Kahane (1992) discuss the connection between preference for moments and expected utility rankings of risky prospects, arguing that " $U'' < 0$ and $U''' > 0$ are not related to variance avoidance or skewness preference".

9. This follows from the equivalence of the OLS estimator and the *sample* estimators for the minimum variance hedge ratio. In terms of population parameters, the minimum variance hedge ratio is equivalent to the slope coefficient in a bivariate normal regression of spot on futures prices. Much of the recent discussion of the OLS result focuses on whether the price variables should be expressed in levels, changes or rates of return, e.g., Myers and Thompson (1989); Toevs and Jacob (1986); Witt, et.al. (1987).

10 There are a number of societies dedicated to various aspects of risk management such as the Society of Actuaries, the Risk and Insurance Managers Asso.,. Risk Management Asso., and the Association of Financial Engineers.

11 Various examples of the risk management sections of recent annual reports can be found at the website, www.sfu.ca/~poitras. Follow the links to the book website.

12 Williams (1995) is an excellent source for more in depth discussion of the issues surrounding the definition and legal application of manipulation. The following discussion draws liberally from that source.

13 *General Foods vs. Brannan*, 170 F.2d 220 (7th Circuit, 1948), p.231, quoted in Williams (1995, p.5).

14 In November of 1991, David Threlkeld, a US copper broker operating on the LME, received a letter requesting him to backdate trade confirmation dates for a fake deal worth $425 million. This letter was apparently from Hamanaka. Recognizing the illegality of the request, Threlkeld passed the letter along to the head of the LME. The LMEs view on the letter was, more or less, that Threlkeld was well advised to keep quiet over the matter to avoid getting sued. At this point, it is not clear whether the LME did anything to follow up on the Threlkeld complaint.

15. A standard reference on the basics of the CAPM is Alexander and Sharpe. A more detailed discussion of the CAPM in an international context can be found in Adler and Dumas (1983) which is also a useful reference on PPP and other issues.

16. The MM theorems and subsequent literature are discussed in numerous sources, including Brealey and Myers, Principles of Corporate Finance (1992).

17 Accounting exposure identifies specific accounting items that are subject to risk of exchange rate changes. Hedging accounting items proceeds much as in the discussion in Section 6.1.

18. An American example would be hotels and related businesses at Aspen, Colorado.

19. The specification given in Shapiro (p.243) is in terms of levels, not changes, the relevant variables. This is not satisfactory on statistical grounds. Similarly, the discussion provided on the bottom of p.243 about R^2 and significant beta coefficients also appears to be lacking.

20. For a discussion of PPP, there are a number of useful sources. In addition to the Shapiro (1999), see also Officer (1982), Roll (1979), and Shapiro (1983).