II. Commodity Risk Management Concepts

2.1 Measuring Risk and Exposure

A. Risk and Uncertainty
B. What is Value at Risk?
C. Commodity Basis Risk Characteristics

2.2 Optimal Risk Management Decisions

A. Risk Management Objectives
B. Transaction Hedging and Optimal Hedging
C. Optimal Hedge Ratios for *Ex Ante* Decisions

2.3 Strategic Risk Management

A. What is Strategic Risk Management?
B Shareholder Value and Corporate Risk Management
C. The Farmer's Risk Management Problem

> ...human decisions affecting the future, whether personal or political or economic, cannot depend on strict mathematical expectation, since the basis for making such expectations does not exist.
>
> *John Maynard Keynes,* (The General Theory, chap. 12)

> "It is sometimes said that hedging is the opposite of speculation. This is not so. They are different kinds of the same thing. The thing that is usually identified as speculation – that is, long or short positions in futures contracts – is speculation in price level. The thing that we identify as hedging – that is, long cash and short futures or vice versa – is speculation in price relationships. . . . Thus hedging and speculation are not opposite; in fact, they are conceptually similar. They are just different kinds of speculation."
>
> Thomas Hieronymous (1977)

2.1 Measuring Risk and Exposure

A. Risk and Uncertainty[1]

Defining Risk and Uncertainty

The concept of risk is fundamental to many fields of study. Risk appears in numerous guises, from theoretical modelling of financial decisions to determining the social consequences of expanded nuclear power usage. Despite this importance, the precise definition of risk depends on the context and application. Common usage is derived from insurance applications where risk represents the possibility of loss, injury or peril. This definition is reflected in various risk assessment and some risk management applications, ranging from social and psychological risk to environmental and bio-hazard risk. Units of measurement for risk vary with context. In contrast, risk in financial economics is associated with the possibility that the actual return for a real asset or security will differ from the expected return. There is a fundamental tradeoff between risk and return. This financial risk is typically measured using the variance or standard deviation of historical return from the mean return, a definition of risk that includes both positive and negative outcomes. Where only the possibility of financial loss is of concern, as in value-at-risk applications, measurements are evaluated using the left tail of the probability distribution for return or profit.

The evolution of methods for the identification, assessment and management of risk have played a central role in the progress of civilization. In ancient times, religious beliefs were important in reconciling the risks confronting a society. Appeals to the gods by the priesthood, prophecies from the oracle, chanting by the shaman were all methods of passively dealing with risks encountered. The development of scientific, mathematical and probabilistic methods during the Enlightenment permitted risk to be more actively identified and assessed. This advancement encountered a philosophical quandary concerning subjective and objective interpretations of probability. More precisely, the objective interpretation views probability as inherent in nature. Logic, scientific investigation and statistical analysis can be used to discover objective probabilities. In contrast, subjective probabilities quantify an individual's belief in the truth of a proposition or the occurrence of an event and are at least partially revealed in an individual's choice behavior. Such probabilities can vary across individuals due, say, to differing degrees of ignorance about the event of interest.

Debate over subjective versus objective probability reached a peak around the time that Frank Knight (1885-1972) introduced a distinction between risk — where the objective probability of an event is at least measurable — and uncertainty — where the probability is not knowable and has to be determined subjectively. This terminological distinction between risk and uncertainty has now faded from common usage as the subjectivist approach has gained prominence supported by seminal contributions from Frank Ramsey (1903-1930), Bruno di Finetti (1906-1985) and Leonard Savage (1917-1971). Attention has shifted to whether subjective beliefs derive from intuition or are only realized in choice behavior. The intuitive approach leads to a focus on the perception of risk, a concept often employed in psychometric and sociological research. Development of the choice-theoretic approach to subjective probability was facilitated by the expected utility function introduced by John von Neumann (1903-1957) and Oskar Morgenstern (1902-1976) in a classic work of social science, *The Theory of Games and Economic Behavior* (1944). The choice-theoretic approach has sustained the modelling of decision making under uncertainty that is a central component of modern economic theory, in general, and financial economics, in particular.

Prior to von Neumann and Morgenstern, neoclassical economic theory was based on certainty or perfect foresight, though consideration of risk in decision making could be found in the less formal approaches of Frank Knight, John Maynard Keynes (1883-1946) and Irving Fisher (1867-1947). These individual contributed to a range of future contributions and perspectives on the impacts of risk in economics. Precisely how to model predictions for random variable outcomes, e.g., using the conditional distribution, 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 Keynes and Knight 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. The essence of these contributions is often forgotten in the modern 'scientific' approach to risk management.

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 tradition of classical and neoclassical economics, seeking to explain how economic profits can arise from uncertainty in the process of production and distribution. Neoclassical 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 which 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.

In effect, Knight is saying that in many instances involving commercial risks there is insufficient empirical information to accurately form the subjective probabilities need to implement the expected utility approach. Risk is associated with objectively measured probabilities, while uncertainty requires subjective probability assessments that are difficult to determine in important practical situations. The economic rents to business ownership or, for that matter, commodity risk management arises from correctly anticipating uncertain outcomes.

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 among the different approaches for dealing with uncertainty. For example, increased specialization permits more firm resources to be devoted to data collection and analysis which 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 say very clearly is that the randomness associated with economic risks, such as strategic business risk, is composed of 'risk', which is measurable in an objective sense, and 'uncertainty', which is only measurable subjectively. It is in dealing correctly with uncertainty that "entrepreneurs" earn value.

Keynes and Fisher represent two alternative approaches to uncertainty. By explicitly recognizing the "caution coefficient" that measures the difference between the mathematical expectation and the price that will be paid for a gamble, Fisher laid the foundation for later contributions in mean-variance portfolio theory. This approach assumes sufficient empirical information for subjective probabilities to be specified. In contrast, the numerous contributions by Keynes on the impact of uncertainty range from the *Treatise on Probability* (1921) to the *General Theory of Employment, Interest and Money* (1936). As the chapter opening quote identifies, Keynes followed Knight in maintaining there was insufficient empirical information to determine the subjective probabilities needed to determine the mathematical expectations associated with means and variances. Disciples of Keynes, such as George L.S. Shackle (1903-1992), argue against the use of probability theory to model decision making under uncertainty. Poitras (2011, ch.5) demonstrates the failings of the ergodicity assumption in applications of probabilistic notions to model financial decisions.

Given the diverse approaches to risk generated by Knight, Keynes and Fisher, the application of expected utility in economic theory requires the use of preference orderings over state contingent commodities. Kenneth Arrow (born 1921) and Gerard Debreu (1921-2004) extended the neoclassical economics of Stanley Jevons (1835-1882), Leon Walras (1843-1910) and Alfred Marshall (1842-1924) to include decision making under uncertainty. This development follows naturally from using the choice-theoretic approach to subjective probability developed by von Neumann and Morgenstern. The utility of a certain outcome is replaced by the expected utility, calculated using known probabilities and the associated utilities for a set of random outcomes. The known probabilities can be notionally determined by direct observation of previous choice behavior. Using this approach, there is no formal distinction between risk and uncertainty, e.g., Penello (2009). Risk is associated with the variability of random outcomes and uncertainty with randomness. Sensitivity to risk is measured by comparing a certain outcome to a random outcome with the same expected value. Risky outcomes are measured in income, dollars or returns and can take both positive and negative values.

In financial economics, the expected utility framework has been applied to determining solutions to: the optimal combination of individual securities in a portfolio of securities; and, as discussed in sec. 2.2, the optimal hedge ratio to use in risk management decisions. The initial application involved using an expected utility function specified over the expected (portfolio) return and variance of (portfolio) return. Harry Markowitz (born 1927) and William Sharpe (born 1934) were able to demonstrate that the variability or risk of a portfolio can be further divided into two components: firm specific risk which is diversifiable and non-systematic; and, market related risk which is systematic and not diversifiable. Applying this to the tradeoff between risk and return, it is demonstrated that only increases in the systematic risk of an individual security will be rewarded with higher expected return. Hence, it is only that portion of the total variability of a security's return that cannot be diversified away that warrants higher expected return. A measure of systematic risk — the beta of a security — is provided. Beta can be calculated as the slope coefficient in a least squares regression of individual security return on market return: the ratio of the covariance between the individual security return and the market return divided by the variance of the market return.

More recently, a variety of risk measures have been developed to deal with perceived limitations of variance of return and beta in dealing with downside risk. These new measures include: value at risk; expected regret; conditional value at risk; and, expected shortfall. In many applications, losses have a different impact than gains, if only because losses may increase the risk of bankruptcy or lead to the delay of investment project. Risk measures such as variance, standard deviation or beta give similar weight to upside and downside movements. 'New' risk measures such as value at risk and expected shortfall are concerned only with the properties of downside risk. While this enhances the identification of potential impact of downside risk, the connection with overall firm profitability is obscured. The raises the well known quandary facing risk management: whether risk associated with losses or overall profitability is the ultimate objective.

Epistemology of Commodity Risk Management

Modern academic studies of commodity risk management face an enigma. The strategic importance of commodity risk management directly impacts the valuation process for the common stock of many non-financial firms. Yet, the vagaries of market pricing prevent precise identification of this impact. The commodity risk management enigma is situated within the more general epistemological difficulties that 'scientific studies' face when confronting the problem of valuing the common stock of any publicly traded company, e.g., Poitras (2011). Confronted with the practical difficulties of determining an ***ex ante*** market price of a commodity common stock, academics have found comfort in an analytical perspective based on investor rationality and market efficiency. Recognizing that market efficiency dictates against systematic abnormal gains to commodity price prediction or individual security selection, the upshot is an approach to commodity risk management which emphasizes a strategy based on empirical estimation. Analysis of the heterogeneous characteristics of individual commodity risk management situations is avoided in favor of the search for sources of homogeneity (factors) across firms. How the commodity risk management decisions of specific firms were determined have little role in this process.

In modern finance, various philosophical approaches compete to explain what constitutes knowledge and objective truth in determining a commodity price or valuing the common stock for a non-financial firm or determining an equity investment strategy for such firms. Finance is, at root, a human science, concerned with explaining and predicting that aspect of human behavior associated with financial activities. Much of interest has appeared in the epistemological debates about knowledge and objectivity in the human sciences since, say, Hayeks' The Counter-Revolution of Science (1955) or Gadamer's Truth and Method (1960). Unlike the natural sciences, what is required in the human sciences is recognition that there are differing approaches to what constitutes knowledge when human behavior is involved. It is naive and intellectually chauvinistic to believe the route to knowledge and truth in, say, valuing equity securities or determining a commodity risk management policy is unproblematic, provided that one adheres to the analytical and epistemological approach of modern Finance: it is inappropriate to conclude that deviations from the narrow parameters of the prevailing epistemology are 'unscientific' nonsense not worthy of academic consideration.

Knowledge appears in various guises: empirical observations, logical deductions and informed conjectures can all be part of the final picture. Making sense of the different facets requires that careful attention be given to the language being used. For example, a logical relationship derived from a theoretical model may have only limited empirical applicability. Yet, the logical relationship may be presented as though it has a strong 'factual' basis. This may confuse an uninitiated audience into concluding that the factual basis, which is logical, extends into the empirical realm. Academics, in general, and modern financial economics, in particular, are inherently attracted to logical facts, such as the capital asset pricing model or the mean-variance optimal hedge ratio. Whether logical facts have any *ex ante* empirical validity requires careful analysis that extends beyond the theoretical structure used to develop the model. Though this point may seem obvious, the resulting confusions are apparent even in introductory investments textbooks where logical relationships, such as the CAPM or optimal hedge ratio or Value at Risk measure, are presented as though there were an empirical validity which corresponds to the logical validity.

The term 'epistemology' comes from the Greek word for knowledge. Simply put, epistemology is the philosophy of knowledge. The central question of epistemology is how individuals come to know or, in slightly different terms, how knowledge is created. Methodology is concerned with the methods that are used in creating knowledge and, as such, is more practical in nature. Positivism is a philosophical movement, concerned with epistemology, characterized by an emphasis upon science and scientific method as the only sources of knowledge. Though the roots of positivism can be traced back to Francis Bacon (1561-1626), the beginnings of the movement are usually credited to Auguste Comte (1798-1857). Over time, positivism evolved substantively to the point where, in the

1920's, a new version, known as ‘logical positivism’ (also known as logical empiricism, logical neopositivism, neopositivism) emerged. Reflecting the German and Austrian roots of the so-called Vienna school, the leading founding figure is usually identified as Rudolf Carnap (1891-1970). However, the English philosopher A.J. Ayer (1910-1989) is usually credited with the most influential contribution Language, Truth and Logic (1936). The branch of positivism reflected in modern financial economics can be traced to Friedman (1953), e.g., Boland (1979).

Comte argued the search for knowledge had gone through three historical phases: the theological, that was concerned with obtaining knowledge about God and spirituality; the metaphysical, where the search was for philosophical truths; and, the positive or scientific phase, that involved the search for objective facts or ‘positive truths’. It was this last phase that Comte associated with positivism. As initially conceived by Comte, the positivist approach to knowledge made a sharp distinction between the realms of fact and value. There was also a strong hostility toward religion and traditional philosophy, in general, and metaphysics, in particular. The positivist philosophy maintained that all sciences rely upon the same methodology for determining facts about the physical and material world. As such, there are no important differences between, say, biology, physics or economics. This was referred to as the ‘unity of science project’. Facts are to be collected and summarized through a process of induction.

Echoes of positivism constantly resonate through modern financial economics. Elton and Gruber (1984, p.273) provide an excellent example: ***"As the physicist builds models of the movement of matter in a frictionless environment, the economist builds models where there are no institutional frictions to the movement of*** **[*commodity and*]** ***stock prices"*** (emphasis added). The epistemology of such an approach to pricing in commodity and stock markets can be traced to Friedman (1953) where the distinction between fact and value appears as a distinction between "positive economics" and "normative economics" (p.4):

> Positive economics is in principle independent of any particular ethical position or normative judgments ... it deals with "what is" not with "what ought to be". Its task is to provide a system of generalizations that can be used to make correct predictions about the consequences of any change in circumstances. Its performance is to be judged by the precision, scope, and conformity with experience of the predictions it yields. In short, positive economics is, or can be, an "objective" science, in precisely the same sense as any of the physical sciences.

Much of Friedman (1953) is concerned with the issue of whether a theory with unrealistic assumptions, even "wildly inaccurate descriptive representations of reality" can be "important and significant". For Friedman, the ultimate test of a theory was "whether it yields sufficiently accurate predictions", not whether the assumptions are realistic.

The concern of Friedman (1953) with the form of the theory being examined is consistent with the evolution of positivist epistemology. Initially, positivism placed heavy reliance on the inductive process of collecting facts. Spurred by the remarkable successes of the natural sciences during the late 19th and early 20th centuries, this view evolved into logical positivism, an epistemology that placed emphasis on theories and the logical deduction of hypotheses to test those theories as well as the collection of facts. The epistemology of logical positivism allows only two grounds for truth: there are deductive truths such as those in mathematics and formal logic, e.g., 12 - 3 = 9; and inductive statements that match reality precisely. As a consequence, truthful statements have to be verifiable to be meaningful. In logical positivism, statements have meaning relative to the conditions under which the statement can be verified. Friedman adapts this approach to where the test of verification for a hypothesis is the ability to predict. That is consistent with the tenet of logical positivism that a statement that does not describe an ‘experiential proposition’ carries no significance, i.e., it is not knowledge.

Friedman (1953, p.7) clearly reflects these tenets of logical positivism in what Boland (1979, 1991) has termed economic positivism: "theory has no substantive content; it is a set of tautologies ... Factual evidence alone can show whether the categories of the ‘analytical filing system’ have a meaningful empirical counterpart, that is, whether they are useful in analyzing a particular class of concrete problems." Statements that are verifiable provide a basis for building a science. Under positivism, science is the source of knowledge. As such, both positivism, in general, and economic positivism, in particular, share a fundamental commitment to empiricism, an epistemology where claims that have no empirical consequences are without meaning. Economic positivism extends empiricism by arguing that science can also seek to build theories to describe the regularities of cause and effect in order to explain the world. This requires theories to be expressed as a set of axioms or, less formally, basic assumptions. These theories have rules to systematically link the predictions with objective measurements of the real world. The connection to Friedman (1953), von Neumann and Morgenstern (1947) and innumerable other projects in positivist economics is apparent.

This discussion begs the question: so what is wrong with economic positivism? There are a number of different answers to this question. At this point, it is relevant to observe that positivism maintains that science is the only way to create knowledge and to allow individuals to understand the world well enough to predict and control outcomes. In the positivist framework, the objective world is viewed as deterministic, operated by laws of cause and effect that can be identified if the unique approach of the scientific method is correctly applied. Science is conceived as a mechanistic operation. It is possible to use deductive reasoning to postulate theories that can be empirically tested. Based on the results of these empirical tests, it is determined whether a theory ‘fits the facts’ or whether the theory needs to be revised in order to provide better predictions of reality. Ultimately, there is an objective reality that can be discovered if there is sufficient empirical information available to verify the ‘true’ deductive hypotheses.

Criticisms of economic positivism are numerous. One type of criticism focuses on the misunderstanding of the process by which science is conducted. Is there really a unity of science? Are the procedures used in physics and chemistry directly applicable to economics or psychology? Do scientists really develop deductive hypotheses that are then ‘verified’ on empirical data? Another related criticism observes that economic positivism says little or nothing about how axioms (or Friedman’s assumptions) are translated into possible testable hypotheses. In other words, positivism has no substantive insight into the process by which knowledge is created. Positivism is only interested in specifying the scientific process, without recommending criteria for selecting among permitted ideas. This leads to Friedman (1953) and the criteria of predictive ability. But, this leads to the problem of measuring predictive ability. The distinction between *ex ante* and *ex post* predictability is one key example of this type of problem in modern finance.

Positivism proposes that there is a unity of science. Certain developments in epistemology after positivism deny this proposition. As such, schools of thought have emerged that are concerned specifically with the epistemological problems arising in the human sciences. One such epistemology is ‘critical realism’, a school that observes all measurement is fallible is some way, e.g., Bhaskar (1978), Lawson (1997). For example, critical realists maintain that all observations are theory-laden and that individuals, in general, and scientists, in particular, are inherently biased by their cultural experiences, world views, and so on. Friedman (1953, p.4-5) recognizes this issue but does not view it as a basis for "a fundamental distinction" between economics and the natural sciences. For critical realists the challenge is how to move from a notion of objectivity that is inherently a social phenomenon to the identification of knowledge. If objectivity is not perfect, then how are these separate and imperfect individual interpretations of reality to be combined?

Friedman (1953) provides a window to the 20th century development of the philosophy of the social sciences. In this development, words like "hermeneutics" and "ontology" are essential to the discussion, though references to notions such as "the questionableness of romantic hermeneutics" require knowledge of the philosophical developments to be correctly interpreted. Hermeneutics has a long history in philosophy, starting with problems of biblical exegesis. During the 18th and early 19th century, hermeneutics evolved into a more general theory of textual interpretation, aiming to provide a set of rules for accurate interpretive practice applying to a wide range of subject matter. Taking hermeneutics as the relevant method for the recovery of meaning, Wilhelm Dilthey (1833-1911) broadened hermeneutics to represent a methodology for the recovery of meaning that is central to understanding knowledge within the 'human' or 'historical' sciences.

Strongly influenced by Martin Heidegger (1889-1976), Hans-Georg Gadamer (1900-2002) is "the decisive figure in the development of twentieth century hermeneutics" (Stanford Encyclopedia of Philosophy). Gadamer is part of a long line of thought that questions the ability to apply techniques of the natural sciences to the human sciences, e.g., (p.6): "the real problem that the human sciences present to thought is that one has not properly grasped the nature of the human sciences if one measures them by the yardstick of an increasing knowledge of regularity. The experience of the socio-historical world cannot be raised to a science by inductive procedure of the natural sciences." Though Gadamer's notion of the human sciences may seem to have more applicability to, say, political science or sociology, it is difficult to evade the observation that the prices of commodities and common stock are set in markets and are the outcome of a social interaction. In turn, commodity risk management practices form an integral part of the process by which commercial firms are operated and the securities of those firms are valued.

Unlike the natural sciences, the human sciences have to allow for prejudice derived from authority. In contrast, methodologically disciplined use of reason cannot accept arguments based on authority for that involves not using one's reason to reach conclusions. "If the prestige of authority takes the place of one's own judgment, then authority is in fact a source of prejudices". But the approach toward the human sciences proposed by Gadamer (1960, p.249) does not view prejudice either negatively or positively. As such, authority as a positive prejudice provides a basis for knowledge:

> ... the recognition of authority is always connected with the idea that what authority states is not irrational or arbitrary, but can be seen, in principle, to be true. This is the essence of the authority claimed by the teacher, the superior, the expert. The prejudices that they implant are legitimized by the person who presents them. But this makes them then, in a sense objective prejudices, for they bring about the same bias in favor of something that can come about through other means. e.g., through solid ground offered by reason.

The process of interpretation and understanding is fundamental to the human sciences. While knowledge about an object in the natural sciences gets progressive deeper over time, the same is not true about the human sciences where great achievements of the past "hardly ever grow old".

For Gadamer, the interpreter is an essential component of knowledge in the human sciences: "the object appears truly significant only in the light of him who is able to describe it to us properly. Thus it is certainly the subject that we are interested in, but the subject acquires its life only from the light in which it is presented to us." Subjects appear historically "under different aspects at different times or from a different standpoints" (p.252). Insightful interpretations require the past to be echoed in the present. As such, the human sciences are involved not only in the accumulation of empirical results but in the transmission of an important source of authority: tradition. "That which has been sanctioned by tradition and custom has an authority that is nameless, and our finite historical being is marked by the fact that always the authority of what has been transmitted – and not only what is clearly grounded – has power over our attitudes and behavior" (p.249).

Gadamer sees an essential role for tradition in the human sciences (p.251-2): "That there is an element of tradition active in the human sciences, despite the methodological nature of its procedures, an element that constitutes its real nature, and is its distinguishing mark, is immediately clear if we examine the history of research and note the difference between the human and natural sciences with regard to their history". For Gadamer: "the natural scientist writes the history of his subject in terms of the present stage of knowledge. For him errors and wrong turnings are of historical interest only, because the progress of research is the self-evident criterion of his study ... the human sciences cannot be described adequately in terms of this idea of research and progress." Knowledge in the human sciences does not proceed by distancing and freeing ourselves from what has been transmitted through tradition. Rather, the problem is to find the relationship of the present with the traditions of the past.

The positivist foundation of academic studies depends on the premise that knowledge in the subject is obtained solely from the methodology of the natural sciences. Somehow, increasingly greater knowledge is obtainable about the natural phenomena of stock or commodity markets, especially commodity prices, as increasingly larger amounts of data are examined or more precisely mathematical theories are derived. The historical evolution of markets is unimportant. The views of writers in the past, such as Holbrook Working or Nicolas Kaldor or J.M. Keynes, are only of historical interest, useful illustrations of how far knowledge has progressed since that time. Gadamer, and other philosophers of his ilk, would argue that this approach is predicated on the supposition that the subjects of finance, in general, and commodity risk management, in particular, can be a natural science. However, the objects of interest in commodity risk management, especially commodity prices, are the result of human interactions and, as such, belong in the realm of the human sciences. If correct, knowledge of the subject could be substantively increased by proceeding beyond the scientific search for sources of homogeneity in commodity risk management across firms to incorporate the notions of acceptable heterogeneity in commodity risk management decisions associated with the contributions of authorities from the past.

B. What is Value at Risk?

To say that the value at risk (VaR) methodology has produced a ‘scientific’ revolution in exposure measurement for financial risk management is, arguably, an understatement. Yet, the concept has not received as much traction in commodity risk management. 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 the US accounting standard FAS 133 has inspired financial firms to include VaR calculations in annual reports and other financial statements. As detailed by Emm et al. (2007), the SEC also imposes rules for reporting exposure:

> Under Item 305 of SEC Regulation S-K, "Quantitative and Qualitative Disclosures about Market Risk," companies must disclose in their 10-Ks information about their exposures to fluctuations in variables such as interest rates, foreign exchange, and commodity prices. While disclosure is mandatory, companies have the discretion to choose among three alternative methods: sensitivity analysis, value-at-risk (VaR), and the so-called "tabular" method.

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 which have been used for many years by more sophisticated firms. The VaR revolution is associated more with the system wide adoption of these techniques by depository institutions and other financial intermediaries. However, it is not clear this system wide introduction of VaR has resulted in a corresponding reduction in systemic risk in financial markets.

On balance, the VaR revolution has been profound for financial firms, e.g., Jorion (2006). 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. Yet, VaR can be of importance for non-financial firms, particularly multinational firms, seeking to assess and control the financial risk which 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 which could be done monthly or quarterly. VaR has the irresistible attraction of providing a single number which summarizes the 'risk' of a position. The risk exposure can arise from a number of different measurement 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 initial impetus to the spread of VaR was the widespread availability of software and technical material to support the implementation. In particular, the RiskMetrics group that was formerly associated with J.P. Morgan/Reuters was an important early promoter of the VaR methodology providing detailed technical publications, such as the RiskMetrics manual (JP Morgan 1996), and daily data sets for important financial variables that were free of charge on the J.P. Morgan web page. The RiskMetrics manual describes a set of methodologies outlining how risk managers can compute VaR on a portfolio of financial instruments where commodities are included as an 'asset class'. RiskMetrics paid close attention to modeling the VaR for positions containing options. Non-linear payoffs associated with options can pose problems for the VaR methodology. The RiskMetrics group was first established in 1989 when Sir Dennis Weatherstone was chairman of J.P. Morgan. In 1993, J.P. Morgan publicly launched the RiskMetrics methodology. In January 1998, the RiskMetrics Group was spun off from J.P. Morgan and the RiskMetrics Group was listed on the New York Stock Exchange (NYSE: RISK). Finally, in June 2010, RiskMetrics was acquired by Morgan Stanley Capital International.

VaR for the One Asset Case

Unlike financial risk management where diversification gains are an important element of VaR calculations, many commodity risk management applications can be structured for the case where the calculated VaR measures the impact of changes in a commodity price on commercial firm operations. Because VaR can be affected by the presence of non-linear payoffs arising from the presence of real options, a number of different methods can be used to arrive at VaR estimates. The simplest of these methods is the *variance-covariance* method, e.g., 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. For the one asset case, this method establishes an immediate connection between VaR and techniques of univariate 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 commodity price behaviour can be 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 to yearly. In some applications, VaR for hourly intervals could be calculated but such calculations would have limited information in a commodity risk management context. Large financial firms which 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 commodity price risk for Coca-Cola or the oil price risk for Canadian Oil Sands, could calculate a monthly VaR. In financial risk management, 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).

An essential input to the VaR calculation for commodity risk management is the data for the commodity prices of interest. Where commodities are traded on exchanges, market prices can be used. If the commodity is not exchange traded, e.g., tungsten or uranium, then an estimate for the price has to be obtained from a survey of dealers and brokers specializing in the commodity or from an appropriate pricing model. For financial institutions subject to BIS-style rules, the relevant pricing models have to conform to certain requirements, , e.g., Hendricks and Hirtle (1997). 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. However, the situation is less

> Originally introduced by the Riskmetrics group, Andren et al. (2005) give the following description of Cash Flow at Risk (CFaR):
>
> CFaR ... is the cash flow equivalent of Value-at-Risk, or VaR, which is widely used as the basis for risk management systems within financial institutions. Whereas VaR-based systems specify the maximum amount of total value a firm is expected to lose under most foreseeable conditions (for example, with a 95% confidence level), CFaR-based systems determine the maximum shortfall of cash the firm is willing to tolerate (again, with a specified level of statistical confidence}. CFaR is gaining in popularity among industrial companies for much the same reasons VaR has succeeded with financial firms: it sums up all the company's risk exposures in a single number that can be used to guide corporate risk management decisions. It is this number—the maximum shortfall given the targeted probability level—and the fact that it can be directly compared to the firm's risk tolerance that are the uniquely attractive features of both VaR and CFaR.

transparent for non-financial firms where the physical 'assets' involved are typically non-traded and the impact of commodity price change on the value of the asset is difficult to determine.

Extending the VaR methodology to commodity risk management situations is complicated, e.g., Al Janabi (2009), Manfredo and Leuthold (1999). In financial risk management, the profit or loss due to a change in the price of traded securities can be calculated directly. For non-financial firms, the impact of the change in commodity prices is less direct. Following Stein et al. (2001) and Andren et al. (2005), the change in firm cash flow due to the commodity price change and other variables can be used to construct a cash flow at risk measure (CFaR). As Andren et al. (2005, p.77) observe: "while VaR has met the demand for measures of downside risk in the context of portfolios of financial assets, it is clearly inappropriate for most non-financial firms. If applied to a non-financial firm's portfolio of financial instruments (debt instruments, swaps, FX contracts, and so on), VaR will capture only a small part of the company's overall exposure since it ignores the risk of its underlying commercial cash flows." While conceptually appealing, implementation difficulties arise because: "The derivation of CFaR requires a forecast of the probability distribution of cash flow at some future point in time".

The potential difficulties of determining the cash flow distribution for the CFaR methodology are apparent upon inspection of the one-asset VaR case. More precisely, for the single asset VaR case, the mean (μ) and volatility (σ) of the asset value distribution have to be calculated. For CFaR, these parameters are associated with the distribution of future cash flow. While the change in asset value is the appropriate random variable for VaR, the random variable for CFaR -- the change in a selected cash flow measure -- depends on the impact of a change in commodity price or other applicable random variable, e.g., exchange rate. For the VaR portfolio case, in addition to the individual asset volatilities, the asset return covariances are also calculated and the VaR of the portfolio determined. For CFaR, the connection between the covariance of two commodity prices and the associated impact on cash flow is not apparent leading to 'top down' and exposure-based CFaR calculations that seek to model the "*drivers*" of total variability of cash flow.

The CFaR calculation can be compared with practical 'bottom up' applications common in the financial reports on numerous non-financial firms where cash flow exposure associated with specific market risks is typically identified using sensitivity analysis. Figure 1.B.a provides an example of such point estimate calculations from CNQ. No attempt is made to model the cash flow distribution or to determine the maximum shortfall at a given probability level. This approach can be contrasted with the exposure-based CFaR approach suggested by Andren et al. (2005, p.77):

> Exposure-Based CFaR involves the estimation of a set of exposure coefficients (deltas) that provide information about how various macroeconomic and market variables are expected to affect the company's cash flow. We argue that these coefficients can be estimated using a multivariate regression framework for analyzing corporate exposures to macroeconomic and market risks that recognizes the interdependence of such exposures.

In effect, CFaR is a sophisticated form of optimal hedge ratio calculation. Can such a 'scientific' approach to risk management can be successful?

INSERT Figure 1.B.a Canadian Natural Cash Flow Sensitivity

To illustrate the CFaR methodology, consider the cash flow generated by a real asset producing a single commodity being sold at a random price. Let $r_t = \ln(1 + R_t)$ where $R_t = (CF_t - CF_{t-1})/ CF_{t-1}$ and CF is the cash flow at a given time which changes due to fluctuations in 'market risk' associated with the commodity price. This transformation to a rate of return calculation is done for statistical reasons which will be discussed shortly. Due to commercial factors, such as the level of fixed costs, the impact of commodity price changes on cash flow may not be linear. 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, though other densities could be used, e.g., Fuss (2010).

INSERT Figure 1.B.b Distributions and VaR

Having assumed that r is normally distributed, parameter estimates for the volatility and mean of $\Phi[r]$, 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:

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

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) can be used when this equation is inverted to solve for r^*:

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

Cash flow at risk can now be determined by evaluating $CF_t\ r$ and $(CF_t\ r^*)$.

For short time horizons, such as weekly CFaR, it can often be assumed that $\mu = 0$ to produce the result:

$$CFaR = -2.33\ \sigma\ (S)$$

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

$$CFaR = (S)\ (\mu - 2.33\ \sigma)$$

In VaR calculations, some presentations of the calculation of the return form 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, e.g., Hamidieh and Ensor (2010). 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.

Cash Flow at Risk, Normality, and Options

There are definite pitfalls which can arise in extending the VaR model to CFaR. In general, excessive reliance on quantitative risk measures can produce a chimera for individual firms. Quantitative measures can provide false confidence that commodity risk and other risks have been effectively identified and appropriate actions have been taken to deal with quantitative risk which has been identified. Even Alan Greenspan (1996) has observed 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." Given the limitations of VaR for financial firms, e.g., Duffie and Pan (1997), Ju and Pearson (1999), Fresard et al. (2011), there are also serious problems which can arise in determining the CFaR for a non-financial firm, e.g., Culp et al. (1998). Some problems with VaR and CFaR modeling may be theoretically rectifiable. Other problems can be rectified only by significantly increasing the complexity of the modeling process. Some problems may not be rectifiable at all and heuristic adjustment will be required. Because the set of problems facing a CFaR 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. This assumption is typically 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 and predict exposure to financial prices, commodity prices and cash flows are widely recognized, e.g., Brooks et al. (2005), Wood (2011). It is well known that the probability distribution for changes or returns in many market determined variables, such as commodity prices, are not normal, being typically fat tailed (leptokurtic) and, often, skewed, e.g., Lechner and Ovaert (2010). 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. A range of potential solutions have been proposed to deal with limitations of the normality assumption. The basic idea of many approaches is, somehow, to adjust the normal distribution to accurately reflect the true tail density.

The problem of empirically fitting a distribution to a time series of past data that is useful for predicting future values of the time series is not without difficulties but may be manageable in many situations. A common statistical approach to fitting non-normal data is to use a series approximation, such as an Cornish-Fisher or Edgeworth expansion, to the model the true distribution. This would typically result in higher moments, such as skewness and kurtosis, being estimated and used to adjust the tail densities, e.g., Erb and Harvey (2006). Regarding the use of a Cornish-Fisher expansion, Fuss et al. (2010, p.263) observe:

> Although the Cornish-Fisher (CF) VaR adjusts the critical values of the standard normal distribution, neither normal VaR nor CF-VaR react sufficiently to changes in the return process, which can be problematic for forward-looking investment decisions.

As a consequence of observed empirical limitations, such approaches are less popular than approaches which make appropriate adjustments to the parameters which are needed for testing, the volatility and, possibly, the drift.

Calculating Cash Flow at Risk (CFaR)

Fuss et al. (2010, p.76) make the following observations about different methods of determining CFar:

The calculation of a risk statistic such as CFaR requires an estimate of the probability distribution of cash flow at some future point in time. RiskMetrics, the firm that originally developed CFaR, generally relies on a "bottom-up" approach that attempts to identify cash flow components that are exposed to market risk. Their definition of CfaR thus targets cash flow volatility *conditional on specified levels of market risk.* This approach is useful when management has confidence in its estimates of risk and in its understanding of how changes in market prices affect corporate cash flows. But when there is considerable uncertainty about both the risks and their expected effects on cash flows, management will want to calculate the firm's overall (as opposed to its "conditional") CFaR—and for this purpose the bottom-up approach will not work. Moreover, in cases where it is impossible to identify all sources of exposure to market risk, a firm's total exposure is more accurately measured by its cash flow "delta" (the sensitivity of its cash flow to a small change in the underlying market price). And here again the bottom-up approach cannot provide such a measure.

[Stein et al. (2001)] instead apply a "top-down" approach in which the focus is on overall cash flow volatility. In place of bottom-up estimates based on a company's historical data and line managers' projections, the top-down method pools cash flow data for a large number of comparable companies to estimate a pooled cash flow distribution. The advantage of such an approach is its ability to provide a historical average exposure estimate that reflects the collective experience of many firms under a variety of market conditions. Bur this approach also has an obvious limitation in that the firm in question could be very different from the "average" company in the sample. Moreover, the top-down approach does not provide an estimate of CFaR conditional on market risk, nor can it be easily adapted to do so.

In particular, Fuss et al. (2010) demonstrate that unconditional VaR estimates based on normal distributions for commodity price changes are improved using estimators that admit non-normality such as the GARCH-VaR, e.g, Giot and Laurent (2003), and the conditional auto-regressive value at risk, e.g., Engle and Manganelli (2004). Kuester et al. (2006) provide a survey and comparison of many of the available method of adjustment. Yet, while the problems associated with the normality assumption in determining VaR estimates for financial applications are in the realm of the theoretically rectifiable, determination of CFaR for non-financial firms is decidedly more complicated because the CFaR approach is concerned with cash flows being generated by the firm and not with value changes resulting from changes in financial prices. Even if the sources of ‘market risk’ could be meaningfully identified, the time series of past cash flows for an individual firm are insufficient to generate a meaningful ‘bottom-up’ distribution for use in CFaR measurement. Similarly, the heterogeneity of firms prevents a meaningful ‘top-down’ distribution from being identified.

A final theoretical concern with CFaR concerns a well-known problem for calculating a VaR for a portfolio is that many securities traded in financial markets, such as options, have non-linear payoffs, e.g., Jorion (2000). Similarly, many real assets also contain various types of options which can have a non-linear impact on firm valuation. Non-linear payoffs substantially complicate the problem of dealing with non-normality, e.g., Hull and White (1998). To address these problems in VaR models, the alternative *delta* and *delta-gamma* approaches have been proposed, e.g., Riskmetrics (1996), Duffie and Pan (1997), Byers (2005). Much like the use of duration and convexity in fixed income analysis, these approaches use a Taylor series expansion (see Poitras 2005) to approximate the non-linear payoff function for an instantaneous change in the random variable. In fixed income analysis, the payoff function is usually assumed to be convex, to reflect the inverse relationship between price and yield. Such a convexity assumption may not be valid for situations involving real options impacting the cash flows of a non-financial firm.

VaR is not Sub-additive

VaR is one of a number of possible risk measures. Being concerned only with the measuring the size of possible losses, VaR is a 'one-sided' or 'downside' risk measure. Alternative one-sided measures include 'expected shortfall' and 'extreme tail loss'. Other risk measures, such as the standard deviation, are two sided taking account of both positive and negative outcomes. Grootveld and Hallerbach (1999) discuss implications of different risk measures for decision making under 'uncertainty'. To differentiate among risk measures, the notion of a coherent risk measure was introduced, e.g., Arzner (1999). A coherent risk measure satisfies apparently innocuous properties of monotonicity, sub-additivity, homogeneity, and translational invariance. Monotonicity requires that if one asset has a higher return in all future states than another asset then the risk measure of that asset will be lower. Homogeneity requires that if the size of the asset increases by a factor of α, then the risk measure will increase by a factor of α. Translational invariance requires that if a riskless amount λ is added to an asset position then the risk of the position is unchanged. While these three properties are satisfied by VaR, sub-additivity is not. Hence, VaR is not a coherent risk measure. A similar comment applies to CFaR.

The sub-additive property for a risk measure $\rho[\cdot]$ requires that:

$$\rho[A + B] \leq \rho[A] + \rho[B]$$

$$\rho[A + B + C] \leq \rho[A] + \rho[B] + \rho[C]$$

and so on for larger and larger combinations ...

In words, sub-additivity requires that the measured risk of a combination of real assets be no greater than the sum of the measured risk for the assets individually. To see this, let there be cash flows from three real assets A, B and C – gold mines or oil wells in different locations or acres planted in different crops are possible examples – and for simplicity assume the cash flow from each asset is initially \$100. For each real asset there are three possible future outcomes for the cash flow from the asset {X = 110, Y = 100, Z = 10}with probabilities: Prob[X] = 0.9 (90%); Prob[Y] = .095 (9.5%); and, Prob[Z] = .005 (0.5%). The 99% VaR[A] = 100 - 100 (Initial Price) = 0 = 99%

VaR[B] = 99% VaR[C]. It follows that, because the probability of Z is less than 1%:

$$\text{VaR[A]} + \text{VaR[B]} + \text{VaR[C]} = 0$$

Consider all the possible non-Z outcomes for a 'portfolio' company that combines the three real assets:

Prob[X/A] * Prob[X/B] * Prob[X/C] = (.9)*(.9)*(.9) = .729
Prob[X/A] * Prob[X/B] * Prob[Y/C] = Prob[X/A] * Prob[Y/B] * Prob[X/C] =
Prob[Y/A] * Prob[X/B] * Prob[X/C] = (.9) * (.9) * (.095) = .07695
==> .23085 (3 cases of 2X with Y)
Prob[Y/A] * Prob[Y/B] * Prob[X/C] = Prob[Y/A] * Prob[X/B] * Prob[Y/C] =
Prob[X/A] * Prob[Y/B] * Prob[Y/C] = (.9) * (.095) * (.095) = .0081225
==> .02436 (3 cases of 2Y with X)
Prob[Y/A] * Prob[Y/B] * Prob[Y/C] = VaR[0] = .095 * .095 * .095 = .0008574

Observing that whenever Z occurs, the change from the initial value will be negative, from these calculations it follows that the .729 + .23085 + .02436 + .0008574 = .9805 is the total probability associated with non-Z events. All value changes below a .9805 probability have a VaR > 0 (Remember VaR is a positive number when there is a loss). Therefore, VaR is not sub-additive at the 99% level.

Standard Deviation is Sub-Additive

While a sub-additive risk measure ρ[·] satisfies the following:

$$\rho[A + B] \leq \rho[A] + \rho[B]$$

A super-additive risk measure satisfies:

$$\rho[A + B] \geq \rho[A] + \rho[B]$$

and so on for larger collections of asset cash flows. When only equality holds, the risk measure is additive. The variance (*var*[·]) of two random variables with **positive covariance** (*cov*[A, B] > 0) is a super-additive risk measure:

$$var[A+B] = var[A] + var[B] + 2\ cov[A,B] \geq var[A] + var[B]$$

However the standard deviation (*sd*[·]) is subadditive:

$$sd[A + B] \leq sd[A] + sd[B]$$

To see this, consider the case where *sd*[*A*] = *sd*[B] = 1 and *cov*[*A,B*] = .5. In this case:

$$\sqrt{1 + 1 + 2(.5)} = \sqrt{3} \leq 1 + 1$$

Only in the perfectly, positively correlated case will equality hold. Unlike VaR, the standard deviation is a two-sided risk measure.

C. Commodity Basis Risk Characteristics

Types of Basis

In market terminology, a ***basis*** refers to the difference between two prices. The study of basis relationships is fundamental to understanding cash markets as well as futures and forward markets. Various types of basis relationships are of interest. One basis relationship which is of theoretical interest is the ***maturity basis***, the difference between the delivery date price of a futures contract and the corresponding spot price. It is often theoretically convenient to assume that the maturity basis is zero, implying that $F(T,T)=S(T)$. Though the maturity basis is almost always zero for forward contracts which are written with delivery in mind, for futures the maturity basis to be zero it is necessary but not sufficient that the spot and futures prices both refer to the deliverable commodity. When the maturity basis for a deliverable spot commodity deviates from zero, a profit opportunity is provided for delivery arbitragers operating on the futures exchange. Lack of convergence in the futures basis in the wheat futures market was a central concern in US Senate (2009) (see Fig. 1.C.bb).

INSERT Figure 1.C.bb Wheat Futures Basis

Evaluation of the maturity basis when the spot commodity is not the same as that deliverable on the futures contract is often complicated by the grade and location characteristics of the spot commodity. Even for futures contracts, the deliverable standard grade specified in the futures contract often permits multiple delivery grades or locations. There are also embedded options associated with time of delivery within the delivery month, deferral of delivery to future months and so on that are commodity and exchange specific. For example, non-ferrous London Metals Exchange contracts can allow for delivery in ports such as Bristol or Hamburg. In addition to the complications this presents to delivery arbitragers, the presence of ***multiple delivery specifications*** in futures or forward contract requires the ***cheapest deliverable*** commodity to be identified in order to determine the precise commodity grade and location which is being traded for future delivery.[2] Increasingly, futures exchanges are moving to financial settlement as a method of dealing with problems of multiple delivery specification.

The ***quality basis*** is a development on the maturity basis. The quality basis relates to the difference between prices for different grades of a commodity. Consumers encounter various types of quality basis decisions on a daily basis. For example, automobile drivers have to decide whether to buy regular or premium gas when tanking up. The difference between these two prices is a quality basis. When the maturity basis is not zero, the difference between spot and futures prices will be a quality basis. Numerous examples of the quality basis are provided in the cash market price quotes provided in the financial press. For example, on Aug. 8, 1994 the Wall Street Journal reported that for New York delivery, Brazilian coffee was selling for \$1.74/lb. while Columbian coffee was selling for \$1.84/lb (Poitras 2002, p.193). Engelhard refined industrial and fabrication quality platinum was selling for \$408 and \$508 per troy ounce. Copper cathodes and copper scrap were selling for \$1.11 and \$.87 per pound. On the same date, Arabian heavy and light oil in Rotterdam were selling for \$15.05 and \$16.25 per barrel, respectively.

Price differences due to variations in quality are determined by market considerations and are not typically constant. This is apparent from the time series behavior of the Brent- WTI (West Texas Intermediate) crude oil spread. This variation will be of concern to hedgers, both for determining the hedge position for the spot commodity and identifying the appropriate deliverable commodity to use for the calculating a cash-futures basis. Because of possible variation in the quality basis, futures contracts with multiple delivery specifications must provide an acceptable method to account for changes in the quality basis of deliverable commodities. In determining the quality basis, considerations of ***location basis*** can often be relevant. For example, the Brent-WTI quality basis also incorporates a location basis differential because WTI prices are for Cushing, OK delivery and Brent is for Rotterdam delivery (see Fig. 1.C.a).

INSERT Figure 1.C.a Brent-WTI crude oil spread

More precisely, the ***location basis*** refers to the difference between the price of more-or-less the same commodity at two different locations, e.g., live hogs in Des Moines and Chicago, winter wheat in Topeka and Kansas City. An example of a location basis on Aug. 8, 1994 is provided in Poitras (2002, p.195) where No. 1 Canola in Vancouver is \$481.30/tonne and in Thunder Bay \$473.60/tonne. For many of the physical commodities, including canola, transportation costs have an important impact on the location basis. In addition to transport costs, a number of other factors can impact the location basis, such as local supply and demand considerations. Information on the behavior of the location basis is important for hedgers to determine hedging strategies. For futures hedgers, the relevant location basis is the difference between the local price of the relevant spot commodity being hedged and the price of the commodity deliverable against the futures contract. Where multiple delivery locations are permitted, the location basis can also be important for determining which location is the cheapest for purposes of making delivery. For many commodities, grade and location basis are combined, as is the case for WTI crude oil in Cushing, Okla. and Brent Sea crude in Rotterdam which feature somewhat different API gravity and sulphur content.

When referring to ***the basis***, without further adjectives, reference is usually being made to the difference between an appropriate forward or futures price and the cash price: $F(0,T) - S(0)$, e.g., Leuthold and Peterson (1983). This form of basis is also referred to as the cash basis or the cash-futures/cash-forward basis. Certain markets have specialized terminology for the basis, e.g., in FX markets the basis is referred to as the swap rate or swap points. Comparing the conventionally quoted cash prices with the futures prices for grain and oil reveals that, when the price of the deliverable spot commodity is correctly identified, the nearby futures contract price is often almost identical with the spot price of the appropriate deliverable commodity. As noted previously, failure of the basis for wheat to convergence to zero at maturity was a major element in US Senate (2009). To make the relevant comparisons needed to evaluate the basis using futures contracts, it is necessary to identify the deliverable commodity associated with the futures contract of interest by referencing the contract specifications from the relevant exchange website. This exercise will reveal that, even though it is not always possible to precisely reconcile cash market quotes with futures markets quotes, various commodities such as gold, silver, crude oil, and soybean meal do typically exhibit a near zero maturity basis.

INSERT Fig. 1.C.b Soybean Futures Prices
INSERT 1.C.c Natural Gas Prices

Having established the correspondence between the prices of the cash and nearby futures contract, it is important to determine the behavior of prices for futures or forward contracts as delivery dates get progressively more deferred. Significant deviations between spot and futures prices can be observed. For example, soybean futures prices exhibit a pattern related to the soybean harvest cycle (see Fig. 1.C.b). In Figure 1.C.c, heating oil exhibits futures prices which are seasonal, highest in the inventory buildup going into winter than in summer. The precise pattern of the futures price deviations for different delivery dates varies considerably across commodities. A classical 'contango' relationship, where prices increase with the time to delivery, is observed in the gold market (see Fig. 1.C.d below) where futures prices which increase monotonically with interest carrying charges. The basis relationship between the prices of futures contracts for different delivery dates is the ***futures basis***. A 'backwardation' in the futures basis occurs when prices decrease montonically with time to delivery. In general, the basis and the futures basis are theoretically determined by cash and carry arbitrage considerations.

Spot, Forward, Futures and Option Prices: **Notation and Definitions**

At any time, there are many prices available for a given commodity. Prices can vary by location, by quality, by delivery date and so on. The study of 'basis risk' is concerned with identifying and explaining differences in prices for the same commodity. Among other contributions, understanding of basis risk is essential for the construction of risk mitigation strategies employing derivative securities and insurance products. The use of analytical concepts such as profit functions to discuss such strategies requires some notation to be introduced:

$F(t,T)$: the forward or futures price observed at time t for delivery at time T.

$S(t) \equiv S_t$: the *cash* or *spot* or *physical* price of the deliverable commodity observed at time t.

Typical subscripts that will be used are $t=0$ and $t=1$ with N for contracts which are *nearby* or closer to delivery and T for contracts which are *deferred* or farther from delivery. For consistency, it has to be that $T \geq N \geq t$. In much of what follows, the assumption, $F(T,T) = S(T)$ is made in order for the price of a futures contract observed on the delivery date $t=T$ to be equal to the price of the deliverable commodity. Unless otherwise stated, the spot commodity is taken to be the deliverable commodity on the derivative security contract. This condition is often satisfied for forward contracts, but requires assuming away the possibility of cross hedging if futures contracts are involved.

As the quote from Tim Horton's Inc. in sec. 1.1 illustrates, contracting methods are an essential component of practical commodity risk management activities. In turn, contracting methods employed depend on the economic characteristics of the commodity industry involved. For example, the iron ore industry – where about 2/3 of seaborne trade originates from three large suppliers – does not feature exchange traded derivative securities to manage price risk. Instead, by setting prices on a quarterly basis, the industry employs short dated forward contracts tailored to the requirements of suppliers and the end consumers – primarily large steel mills and iron foundries. A similar situation can be found in the market for metallurgical coal. In contrast, the copper industry features a larger number of ore suppliers – many of which do employ forward contracting methods to manage price risk – and a consumption chain that involves a sufficient number of refiners and end users to sustain active exchange trading in derivative securities for copper. The agricultural commodities feature different variations on forward contracting and exchange traded derivative securities.

Much of the traditional academic literature on risk management is concerned with applications of exchange traded derivative securities, e.g, Chance and Brooks (2010). To capture the essence of commodity risk management decisions in practice, it is helpful to determine the role, if any, for exchange traded derivative securities. This requires some basic background on the futures and option contracts that are traded on derivatives exchanges. Using a strictly legal definition, it is possible to be reasonably precise about what constitutes a futures contract. Subtle differences in the legal definition of futures contracts across jurisdictions does not create substantive problems, for present purposes. A brief summary of a commodity futures contract is all that is required:

> A ***futures contract*** is an exchange traded agreement between two parties, guaranteed by the clearinghouse, that commits one party to sell a standardized grade and standardized quantity of a commodity to the other party at a given price and specified location at a future point in time.

While useful, this brief summary disguises important features of futures trading. For example, one of the significant limitations of forward contracting is the requirement of precise specification of the grade and quantity of the commodity be determined by the parties to the contract. This procedure raises the problem of how to ascertain whether the commodity delivered meets the grade and quantity requirements. Because forward contracts typically require delivery of the commodity, this procedure is an essential feature of forward contracting. Because futures markets deal in a standardized commodity for which delivery can be avoided by taking an offsetting position prior to delivery, futures contracting avoids this problem.

Like futures, options have a specialized nomenclature. To understand this jargon, the essential notions of an option contract for a commodity needs to be identified:

> An ***option contract*** is an agreement between two parties in which one party, the writer, grants the other party, the purchaser, the *right*, but not the obligation, to either buy or sell a given commodity at a future date under stated conditions.

Options almost always involve the purchaser making some type premium payment to the writer. The timing and form of the premium payment depends on the specifics of the contract. For exchange traded options and many OTC options, the premium is paid up front, when the option agreement is initiated. It is essential to recognize that an option does not represent an ownership claim. Rather, an option is a claim against ownership under prespecified conditions. While it is not necessary that options be exchange traded, many option contracts do originate on exchanges.

Given this, two types of commodity options can identified:

> A ***call*** option gives the option buyer *the right to purchase* the underlying commodity from the option seller at a given price.
>
> A ***put*** option gives the option buyer *the right to sell* the underlying commodity to the option seller at a given price.

The seller of the option is often referred to as the option ***writer***. An option purchaser makes a payment to the option writer referred to as the option ***premium***. Once the premium has been paid, the purchaser has no further liability.

Various features for exchange traded commodity option contracts can be identified. Some or all of these features may apply to other types of option transactions. In order to be accurately specified, option contracts require an ***exercise*** or ***strike*** price as well as an ***expiration date***, on which the right is terminated. The exercise price is the contractually specified price at which the purchaser is allowed to buy (for a call) or sell (for a put) the underlying commodity. When the exercise price is below (above) the current underlying price, the call (put) option is said to be ***in the money***. When the exercise price is above (below) the underlying price, the call (put) option is ***out of the money***. An ***at the money*** option has the exercise price and underlying asset or commodity price approximately equal. Exercising an option involves completion of the relevant transaction specified in the option contract. Options which can be exercised prior to the stated expiration date are referred to as ***American options***; to be contrasted with options which can only be exercised on the expiration date commonly referred to as ***European options***.[3] Depending on the type of option, either a spot delivery (physical settlement) or a net dollar value transaction (cash settlement) may be required to satisfy the conditions of exercise. Finally, the option contract will typically contain other adjustment provisions, e.g., handling of variation in quality and location of delivery.

Arbitrage and Convenience Yield

The distinction between price risk and basis risk is fundamental to commodity risk management. For storable commodities, the behavior of the near-term basis and the near-term futures basis is determined by the execution of ***cash-and-carry arbitrage*** trades. The use of the term arbitrage in this context has a technical meaning. What is colloquially called an "arbitrage" by market practitioners may, under a more technical definition, be little more than a potentially profitable trading strategy with limited risk of losses. An ***arbitrage*** opportunity is defined here as: ***a riskless trading strategy which generates a positive profit with no net investment of funds***. By construction, for any commodity there will be two associated arbitrage trades: ***the long arbitrage***, where the arbitrage transactions involve holding a long position in the cash commodity; and, ***the short arbitrage***, where the arbitrage involves holding a short position in the cash commodity. In some presentations, the long arbitrage is referred to as the cash and carry arbitrage, with the short arbitrage being the reverse cash and carry.

By design, arbitrages are defined only in terms of ***current*** prices and interest rates. If dependence on expected future values were admitted, the trading strategies would not be riskless. For much the same reason, arbitrages cannot involve a net investment of funds. Where the purchase of a commodity is involved, the long cash-and-carry arbitrage is executed by borrowing funds at the current interest rate and using those funds to purchase the spot commodity, simultaneously contracting to sell the commodity at some future date at the current futures price. Hence, the cash-and-carry arbitrage provides a relationship between the futures and spot prices which depends on the net price of carrying the commodity until delivery. Similarly, other derivative securities such as options on futures contracts also provide instances of arbitrage trades, which may or may not involve purchase of the physical commodity. In this fashion, arbitrages determine relationships among current prices of spot commodities and derivative securities.

In practice, the execution of cash-and-carry arbitrages for commodities and the associated behavior of the basis depends on the type of commodity under consideration. Various questions must be addressed to identify the details involved in a specific arbitrage. Is the commodity storable, storable with loss, or not storable? What costs are associated with storage? Is the commodity harvestable? What is the appropriate borrowing rate? Is there an off-setting carry return? To illustrate how arbitrage trading determines the basis for a specific commodity, it is convenient to develop the trades assuming that markets are perfect. In other words, there are no transactions costs, either in the form of commission or bid/offer spreads. Other components of perfect markets include no taxes or other regulatory restrictions and equal lending and borrowing at riskless rate of interest. Relaxing these assumptions converts the cash and carry equality conditions to be derived into upper and lower bounds on forward prices associated with the long and short arbitrage conditions.

Cash-and-Carry Arbitrage: The Case of Gold

Perhaps the simplest cash-and-carry arbitrage to consider is the case of gold, a commodity which is storable at low cost, earns no pecuniary carry return and is not affected by harvests. Assuming for simplicity that there are no other storage costs other than financing charges, i.e., ignoring insurance, vault charges and administrative expenses, a 'long-the-spot' cash-and-carry arbitrage for deliverable gold is described in Table 1.1.[4] Recalling that absence-of-arbitrage requires that the arbitrage profit be non-positive, $\pi \leq 0$, and observing that $Q > 0$ gives a ***long*** *cash-and-carry arbitrage* upper bound restriction on the gold futures price: $F(0,T) \leq S(0)\{1 + r(0,T)\}$. If this condition is violated, the long arbitrage can be profitably executed. When the equality is binding the futures price is said to be at *full carry*, i.e., to reflect the full carrying charges.

INSERT Figure 1.C.d   Gold Futures Prices

Table 1.1: **Profit Function for a Long Gold Cash-and-Carry Arbitrage**

DATE	***Cash Position***	***Futures/Forward Position***
$t=0$	Borrow $\$[Q_G\, S(0)]$ at interest rate $r(0,T)$ and buy Q_G ounces of gold at $S(0)$ for storage until $t=T$	Short Q_G units at $F(0,T)$

-- The cash gold position provides no pecuniary return between $t=0$ and $t=T$

$t=T$	Deliver the Q_G units against the maturing futures/forward contract and use the proceeds to repay the maturity value of the loan, $\$[Q_G\, S(0)]\{1 + r(0,T)\}$	

Ignoring the minimal storage costs, the profit function can be specified:

$$\pi(0) = \{F(0,T) - S(0)(1 + r(0,T))\}\, Q_G \;\leq 0$$

Consider the gold futures price data for Feb. 25, 2011 given in Figure 1.C.d. Examining the June 2011/June 2016 price relationship, for the associated closing prices of \$1410.7 and \$1626.7, the five year true yield interest rate implied in the gold futures prices is 2.89%:

$$\left(\frac{1626.7}{1410.7}\right) = 1.1531 \quad => \quad 1.1531^{\frac{1}{5}} = 1.0289$$

while the June 2011/Jun 2012 contracts (= (1421.3/1410.7) give an implied interest rate of 0.751%. This is roughly consistent with the rates offered on 3 month Euro-US deposit futures which, on Feb. 25, 2011, were trading at 99.625 (= 37.5 basis points) for June 2011 and 98.735 (= 1.265%) for June 2012 delivery. Examination of the futures price structures on Feb. 25, 2011 for other commodities reveals a range of different relationships. Silver, for example, is in backwardation on this date with July 2011 at \$32.91, July 2012 at \$32.713 and July 2015 at \$31.928. Soybeans exhibit a backwardation for the July 2011 and July 2012 contracts at 1382 ¾ ¢ and 1310 ½ ¢. The price structure of copper is also inverted, with prices for deferred June 2012 delivery at \$4.4275 being lower than the nearby June 2011contracts at \$4.46. Due to the impact of the harvest cycle on supply available for storage, most of the agricultural commodities exhibit some form of kinking or reversal in the direction of futures prices as delivery dates get more distant, e.g., the CME/CBT wheat contract. The diversity of futures price spread behavior should be apparent. Full carry relationships are the exception and situations can vary over time. For example, until the recent emergence of backwardation in silver, all the precious metals futures complex prices exhibited contango.

Because each futures contract involves both a long and a short position, the arbitrage relationship between cash and futures prices involves two trading strategies. In addition to the long-the-cash strategy described above which involves combining a fully leveraged purchase of the spot commodity with a short futures position, it is also possible to combine a short position in the cash commodity with a long futures position, a ***short cash-and-carry arbitrage*** trade. Hence, cash-and-carry arbitrage strategies for futures contracts are *two-sided*, having both a long and a short arbitrage trade to be satisfied. While there are differences across commodities, in practice, due to restrictions on the ability to short the cash commodity, execution of the short *or reverse* cash-and-carry arbitrage can be substantially more difficult than the long arbitrage, which only involves leveraged purchase of the spot commodity and costs associated with carrying the commodity to delivery. In cases where sufficient commodity supply is unavailable for lending to short sellers, the cash-futures basis cannot be determined by the execution of cash and carry arbitrages.

Consider the extreme case of gold, where central bank stocks have traditionally provided a significant supply of gold for short selling. Payments from short sellers increase the central bank return on required holdings of gold inventory. In addition, there are funds that hold physical gold stocks to earn the carry return by selling for future delivery. Deviations from full carry in gold induce dishoarding which puts pressure on the spot price. As a consequence of limited restrictions on execution of the reverse cash and carry arbitrage and the activities of funds that hold physical stocks of gold, gold futures prices are at full carry.[5] Another less extreme example would be a grain storage company holding inventory over the crop cycle . At harvest, prices for future delivery dates determine the return to storage. If the futures prices at harvest fall "too far" below a level consistent with an adequate return to storage, then grain inventories would not be purchased for storage and the required grain inventory needed to maintain physical delivery commitments would be acquired using long futures positions. This acts to restore an adequate return to storage until the next harvest when a new storage cycle begins. Fig. 1.C.e illustrates the return to storage for corn from Dec 2011 until July 2012 is 5.35% (annualized).

INSERT 1.C.e Corn Futures Prices

Table 1.2: **Profit Function for a Short Gold Cash-and-Carry Arbitrage**

DATE	***Cash Position***	***Futures/Forward Position***
$t=0$	Borrow Q_G ounces and sell at $S(0)$. Invest the funds received at interest rate $i(0,T)$	Long Q_G ounces at $F(0,T)$
$t=T$	Take delivery of the Q_G units against the maturing futures/forward contract, pay with the proceeds of the investment, $\$[Q_G\, S(0)]\ \{1 + i(0,T)\}$, returning the Q_G units to settle the short position	

Ignoring the gold short sell borrowing fee, the profit function can be specified:

$$\pi(0) = \{S(0)(1 + i(0,T)) - F(0,T)\}\ Q_G \ \leq 0$$

To execute the *short cash-and-carry arbitrage* for gold, the funds received from the sale of the borrowed gold will be invested at a different, probably lower, rate of interest than the long arbitrage. Taking $i(0,T)$ to be the all-in lending rate and again ignoring incidental costs, the short arbitrage is described in Table 1.2. Recalling that absence of arbitrage implies $\pi \leq 0$ and observing that $Q > 0$ gives the short cash-and-carry arbitrage restriction on the gold futures price: $F(0,T) \geq S(0)\ \{1 + i(0,T)\}$. If this lower bound inequality is violated the short arbitrage can be profitably executed.

The combination of the long and short arbitrage conditions imposes upper and lower boundaries on the gold futures price:

$$S(0)\ \{1 + r(0,T)\} \geq F(0,T) \geq S(0)\ \{1 + i(0,T)\}$$

In an idealized world where $r(0,T) = i(0,T) = r$, the equality condition is binding and $F(0,T) = S(0)\ \{1 + r\}$: the futures price will be fully determined at $t=0$ by the current spot price and the cost of financing. This idealized result requires that: lending and borrowing rates be equal; there are no short sale costs or restrictions; the commodity earns no pecuniary or non-pecuniary carry return and is costless to store; and, transactions costs such as those associated with the bid/offer spread are ignored. When the futures prices for gold obey this condition, the commodity is at full carry.

Given the idealized full carry model for gold, it is useful to evaluate what happens to basis behaviour as maturity of the gold contract approaches. Defining $\Delta X \equiv X(1) - X(0)$ and substituting $X(1) \equiv X(0) + \Delta X$ permits the profit function for a one-to-one long spot/short futures gold position to be expressed as:

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

This result is useful for interpreting the profitability of inventory hedges. Recalling that $r(i,j)$ is not annualized but, rather, reflects the actual interest cost over the i to j holding period, $r(0,T)$ will typically be larger than $r(1,T)$. This is due to the one period reduction in the number of days until maturity for both the loan and the futures contract. Hence, even if $\Delta S = 0$ and the level of annualized interest rates is unchanged, there is a time decay in the basis associated with $\Delta r < 0$ which is fundamental to understanding hedge profit determination. There is a 'time clock' at work in a hedge, acting to reduce the difference between the spot and futures price. The time clock continues to wind down to the point where $r = r(T,T) = 0$ and $F(T,T) = S(T)$. Significantly, this time decay is ***not*** present in the futures basis.

Factors Impacting the Basis:
Generalized Cash-and-Carry Arbitrage Conditions

Gold is the commodity that provides cash and carry arbitrage conditions that are closest to the idealized, perfect market conditions associated with the cash and carry arbitrages for financials such as exchange rates and Eurodollars. In general, conditions required for the long and short arbitrages to be combined to approximate an equality condition do not occur in practice. Allowing for the conditions associated with actual commodity markets converts the equality conditions into much weaker upper and lower boundaries on the derivative security price provided by the long and short arbitrage conditions. In addition, the characteristics of various commodities often impose severe restrictions on the ability to execute the short arbitrage trade. As a consequence, the lower bound on the forward or futures price associated with the short arbitrage condition will often provide a much weaker boundary than the upper bound which is determined by the long arbitrage. In cases where the short arbitrage cannot be executed, the cash-and-carry is said to be *one-sided*. The result is that basis determination for various commodities can be complicated in practice.

Problems of basis determination were a central concern in the US Senate investigation into speculation by commodity index funds in the wheat futures market from 2003-2008 (US Senate 2009, p.63):

> Because both the cash price and the futures prices of a commodity vary with time, the basis varies with time as well. In addition, at any particular time, there may be multiple different cash prices for a commodity, depending upon the precise grade of the commodity, the location of the commodity, and the costs of transporting the commodity to market. The basis will vary, therefore, according to the commodity's attributes and location. At any given time, the same crop could give rise to a variety of basis calculations, depending upon specific variables in the crop and the market where it is to be sold. A buyer or seller of a crop must take into account each of those variables when computing the basis for a particular sale or purchase.

An important feature of this Senate investigation was the failure of the ***delivery basis*** for wheat futures prices to converge to the spot price of deliverable wheat, substantially undermining the price discovery and hedging function of wheat futures markets.

Non-storability can be associated with rapid spoilage or prohibitive storage costs. There is a tradeoff between 'non-storability' and length of time to delivery. For non-storable commodities, neither the short or long cash-and-carry arbitrage can be executed for contracts more than a few days or weeks in the future. Commodities in this group have included potatoes, eggs and onions. Poultry and sugar beets also have elements of non-storability, e.g., "Chickens cannot be shipped far before losing value, due to both the direct costs of transport or extra feed and the indirect costs from the birds losing value due to stress, weight loss, or death during transport, or aging during additional feeding. Similarly, sugar beets lose value quickly, and transport costs are still quite high" (Macdonald and Korb 2011, p.3-4).

When available, the price performance of longer dated futures contracts for 'non-storable' commodities is decidedly more erratic when compared to such contracts for storable commodities. Without the ability to do cash-and-carry arbitrages, price determination for futures and forward contracts relies on expectations about future spot prices. At times, this can create market clearing problems, due to lack of liquidity on one side of the market. Empirically, while a number of different types of contracts for non-storables have been offered over the years, there are currently only a few thinly traded futures contracts for truly non-storable commodities. Where contracting for future delivery occurs, short dated forward contracting is the norm, e.g., Macdonald and Korb (2011, p.4). A number of the older studies on non-storables can be found in A. Peck (1977).

In addition to the truly non-storables, certain other commodities may not satisfy strict requirements of storability. For example, feeder cattle associated with delivery on the CME futures are required to be 650 to 849 pound steers. Given weight gain over time, heavier steers eligible for delivery in one contract month cannot be carried for delivery against the next contract as would be the case for, say, gold or silver. However, in these cases, non-storability does not significantly affect cash-futures price determination because there are feeder cattle available which can be used to do the cash-and-carry arbitrage. There are opportunities to purchase feed stock which is at a point in the growth cycle that the stock will qualify for future delivery. Similarly, agricultural crops are only profitably storable from harvest time until the beginning of the next harvest. As illustrated in Fig. 1.C.e, there is a break in the corn basis from the Sept. to Dec. contracts associated with the new supply created by the corn harvest. On balance, it is safe to say there is considerable diversity across commodities in the execution of the long cash-and-carry arbitrages.

Table 1.3 **Profit Function for a Long Grain Elevator Cash and Carry Arbitrage**

DATE	***Cash Position***	***Futures/Forward Position***
$t=0$	Borrow $\{Q_A * S(0)\}$ @ $r[0,T]$ to buy Q_A units of deliverable grain at $S(0)$ for storage in grain elevator @ $cc[0,T]$	Short Q_A units at $F(0,T)$
$t=T$	Deliver the Q_A units of grain (warehouse receipts) against the short	

Assuming all costs of storage are variable, the arbitrage profit function is:

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

Examining the long grain elevator arbitrage in Table 1.3 reveals the importance of storage companies in arbitrage trade execution. Grain elevators and similar storage operators are able to provide storage facilities at cheapest cost and, as a consequence, are appropriately situated to exploit long arbitrage opportunities. Henderson and Fitzgerald (2008) describe the grain marketing process:

> Since their emergence in the mid-1800s, grain elevators have earned income by collecting, storing, and readying grain for transportation. Smaller, country grain elevators collect grain from farmers, hold it in storage, and coordinate transportation to final end users or larger terminal elevators, which coordinate larger shipments to other domestic or international users. The grain held in storage is either owned by the elevator or by the farmers, who pay storage fees.

Grain storage companies also control the bulk of supply available for short selling. In contrast to gold where short sales involve borrowing of physical gold, short arbitrage trade execution in grains involves selling activities of storage operators. This process is decidedly more complicated than for gold (see Table 1.4). Because the bulk of grain in storage at a given time has already been contracted for future delivery, the supply available in storage for sale in the spot market between harvests is restricted. Costly movement of grain from other storage locations can be difficult due to transport costs and uncertain supply conditions in other locations. Committed grain for deferred delivery that is sold needs to be repurchased for a nearby delivery date that will accommodate the completion of previously initiated deferred delivery contracts.

Table 1.4 **Profit Function for a Short Grain Elevator Cash and Carry Arbitrage**

DATE	***Cash Position***	***Futures/Forward Position***
$t=0$	Sell $\{Q_A\}$ units of grain @ $S(0)\}$ and invest proceeds at @ $r[0,N]$ saving grain storage costs @ $sc[0,N]$	Long Q_A units at $F(0,N)$
$t=N$	Use maturing investment funds to take delivery of the Q_A units of grain (warehouse receipts) against the long position.	

Assuming all costs of storage are variable, the arbitrage profit function is:

$$\pi(0) = \{S(0)(1 + r(0,N)) + sc(0,N) - F(0,N)\} \; Q_A \; \leq 0$$

Consideration of the short cash and carry grain arbitrage exposes a fundamental distinction between financial 'commodities' and the physical 'commodities'. Shorting a financial product such as a common stock index or a foreign exchange rate is straight-forward. The long and short arbitrage bounds on forward and futures prices for financial 'commodities' are tight. Due to limitations of physical supply, short selling of commodities is problematic. As a consequence, the short arbitrage bound on the basis for commodities is not binding and ***convenience yield*** is introduced to explain deviations from the cash and carry arbitrage bound. More precisely, ignoring fixed costs the general cash and carry condition for forward prices is:

$$F(0,T) = S(0)\,\{1 + cc(0,T) - cr(0,T)\}$$

where $cc(0,T)$ is variable carry costs, expressed as a rate and $cr(0,T)$ is the variable pecuniary and non-pecuniary returns to holding the commodity, also expressed as a rate. Decomposing cc and cr into components gives:

$$F(0,T) = S(0)\,\{1 + r(0,T) + oc(0,T) - d(0,T) - cy(0,T)\}$$

where $r(0,T)$ is the interest rate portion of carry costs, $cc(0,T) - r(0,T) \equiv oc(0,T)$ is that portion of variable carry costs not attributable to interest carrying costs, $d(0,T)$ is the pecuniary interest or dividend return associated with carrying the commodity, $cy(0,T) \equiv cr(0,T) - d(0,T)$ is that portion of cr not associated with pecuniary interest or dividend return.

Unlike financial products where attention centres on $r(0,T)$ and $d(0,T)$, physical commodities do not earn a pecuniary carry return – $d(0,T) = 0$ – and elements in $oc(0,T)$, such as storage costs and spoilage, can have greater importance than $r(0,T)$. Recognizing the essential role of storage operators in determining commodity forward prices, the diversity of arbitrage execution across commodities can be captured by generalizing the profit function for the cash-and-carry arbitrage to include both fixed and variable carry costs and carry returns. More precisely, for a transaction starting at $t=0$ and ending at $t=T$, the profit function for the marginal storage operator would produce:

$$F(0,T) = S(0)\{1 + cc(0,T) - cr(0,T)\} + [CC(0,T) - CR(0,T)\}]$$
$$= S(0)\{1 + r(0,T) + oc(0,T) - cy(0,T)\} + [CC(0,T) - CR(0,T)]$$

where $CC(0,T)$ is the fixed cost component of carry costs, e.g., regular maintenance and capital costs of storage facilities; and, $CR(0,T)$ is any fixed return earned for being able to hold the commodity, e.g., payments associated with 'take or pay' storage contracts with grain consumers. Though this approach allows for all elements of cost and return to be introduced, the specification of CC and CR can be problematic. In many cases, $CR = 0$ and CC will vary inversely with the amount of the commodity in storage.

For notational convenience, it is sometimes expedient to ***define*** the 'implied carry' as $ic(0,T)$ where $ic(0,T) \equiv cc(0,T) - cr(0,T)$ combining the carry cost and carry return elements. It is also typical to ignore fixed costs and to work with the non-pecuniary variable, $cy(0,T)$, to explicitly allow for potential returns the commodity may provide during the 'arbitrage' period. Recalling that the pecuniary return $d(0,T) = 0$, it follows:

$$F(0,T) = S(0)\{1 + cc(0,T) - cy(0,T)\}$$

Following the discussion of the long spot/short futures inventory hedge for gold, this formulation can be directly applied to an inventory storage hedge profit function involving a general commodity. In this case, the profit function can be written:

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

Again, defining $\Delta X \equiv X(1) - X(0)$ and using $X(1) \equiv X(0) + \Delta X$ permits this form of the profit function to be expressed as:

$$\pi/Q = S(0)\{\Delta cy - \Delta cc\} + \{cy(1,T) - cc(1,T)\}\Delta S$$
$$= -[S(0)\Delta ic + ic(1)\Delta S]$$

This provides a general framework for examining the profitability of an inventory hedge as well as numerous other hedge trades. Application to specific cases requires correct interpretation of cy. For example, while there are no direct financial returns involved in carrying commodities such as wheat and other grains, there are other benefits. This is based on the notion that stocks of a commodity provide some net benefit to the owner. When stocks are low, the convenience yield is high; when stocks are plentiful, the convenience yield is low. The cy may, at times, be an important element of basis behavior, especially for commodities with high absolute carry cost such as the grains and industrial metals.

In practice, the role of cy depends intimately on the importance of absolute carry costs. Commodities with a cash-futures basis determined by relative carry costs, such as the financial futures and gold, usually adhere closely to the cash-and-carry arbitrage condition. When the cash-futures basis depends predominately on absolute carry costs, the arbitrage conditions provide only wide boundaries on the cash-futures basis, and convenience yield can act to off set the costs of carrying the physical. Numerous instances of this are provided in the daily commodity futures price quotes. For example, this explains how harvestability affects the basis. Stocks of grain are most plentiful after the harvest. The value of stocks carried from one crop year to another will fall by the amount of the associated loss in convenience yield. This will be reflected in futures or forward prices, typically resulting in a discontinuity in the futures price structure occurring with the harvest delivery contract, when stocks of grain will be high and the convenience yield low. This type of discontinuity can be seen in the futures prices for corn, soybeans and wheat. A similar situation often occurs in the oil complex contracts and for copper.

Convenience Yield and the Supply of Storage

The notion of convenience yield, and the closely related concept of the ***supply of storage***, were subjects of central interest in the early research on futures and forward markets and continue to attract attention . Seminal contributions include: Kaldor (1939) where the concept of 'convenience yield' was introduced; and, Working (1949) which builds on empirical evidence for the supply of storage in wheat futures provided in Working (1933) (see Fig. 1.C.g). Working (1949) advanced the notion that convenience yield can explain storage under backwardation, i.e., when the price for future delivery is below the current spot price (see Figures 1.C.f). Further study of these concepts continued with Brennan (1958), Muth (1961), Weymar (1968), Danthine (1978) and Turnovsky (1983) making notable contributions. Building on this earlier work, more recent contributions have pursued a number of different approaches to: empirical estimation of the supply of storage function; and, theoretical specification of the convenience yield. Carter and Ghia (2007), Zulauf et al. (2006) and Gao and Wang (2005) provide useful overviews of more recent contributions.

INSERT Fig. 1.C.f Supply of Storage
INSERT Fig. 1.C.g Working (1933) Supply of Storage Curve

Analytically, the notions of convenience yield and supply of storage have direct implications for explaining the behavior of a key variable in the speculator and hedger profit functions: $F(0,T) - E[S(T)]$.[6] However, the connection between the forecasting accuracy of the futures or forward price and the notions of convenience yield and supply of storage is not immediately obvious. Building on empirical work in Working (1949), modern studies model the supply of storage as a relationship between the basis ($F(0,T) - S(0)$) and commodity inventory levels. Because convenience yield and the supply of storage are concerned with properties of the physical commodity, some early approaches ignore the role of the futures market and examine $E[S(T)] - S(0)$. The two period, two agent equilibrium model of the supply of storage of Brennan (1958) is a case in point. Supply and demand functions are derived for a consumer-merchant market. Brennan describes the market this way:

> During any period there will be firms carrying stocks of a commodity from that period into the next. Producers, wholesalers, etc. carry finished inventories from the periods of seasonally high production to the periods of low production. Processors carry stocks of raw materials. Speculators possess title to stocks held in warehouses.

> These firms may be considered as supplying inventory stocks or, briefly, supplying storage....On the other hand, there will be groups who want to have stocks carried for them from one period...to another period....These consumers may be regarded as demanding storage.

In this case, the supply and demand functions for storage are behavioral, dependent on both the spread between the expected future spot price and the current spot price as well as on the levels of stocks being held. The upshot is an identified supply of storage function which provides a (potentially nonlinear) montonically increasing relationship between physical inventory levels and $E[S(T)] - S(0)$.

The development of the partial equilibrium supply of storage model to include futures markets was provided initially by Weymar (1966, 1968) and extended by Turnovsky (1983). In Weymar's model, three agents are identified: merchants, manufacturers and speculators. Futures markets provide cash market participants with an additional method of carrying inventories. Equilibrium in the futures market is directly specified and a supply of storage function is derived. Much as in Brennan's case, there is a monotonically increasing relationship between physical inventory levels and $E[S(T)] - S(0)$. Using a more sophisticated, but similar model, Turnovsky is able to show:

> ...with risk averse behaviour, the current futures price is a weighted average (with weights summing to less than unity) of the current spot price and the expected future spot price. Only if ... producers and speculators are risk neutral...does $F(0,T)= E[S(T)]$ and the futures price become an unbiased predictor of the future spot price. Otherwise, the futures price is a biased predictor, with the direction of the bias depending on the magnitude of the (relevant) cost parameters.

A final implication of Turnovsky's model is that 'under normal conditions', hedgers should be net short and speculators net long.

The transition from expected spot prices to futures prices in empirical work is captured by Brennan (1958, p.58):

> For stocks which are hedged on an active futures market the price spread relevant to a decision about storage levels is the difference between a futures and a spot price. Arbitrage between cash and futures markets will insure that the cash price expected to exist in a future period is accurately reflected in the current quotation of the futures price for delivery in that period.

The transition from unobserved expected prices to observed futures prices leads to more recent empirical studies of the supply of storage where the observed basis ($F(t,T) - S(t)$) at a given point in time is attributed to three factors: the costs of storage; the convenience yield; and, a risk premium to account for the substitution of the futures price for the expected spot price in empirical estimations. From this point, empirical studies have examined the implications of different definitions of 'inventory' and alternative estimation methods that do not directly rely on inventory levels. In addition, theoretical studies of convenience yield have developed using option valuation methodology. This approach inspired the influential Gibson and Schwartz (1990) and Schwartz

(1997).

Carter and Ghia (2007, p.864-5) give the following helpful overview of studies evolving from the seminal Working (1949).

There are two main explanations for the particular shape of the Working curve. The first is the “supply of storage” explanation (Working 1949), which includes the concept of *convenience yield* to explain storage under backwardation ... However, the supply of storage approach has been criticized for taking *convenience yield* as given, and then incorporating it into modern commodity models through an *ad hoc* functional form (see Deaton and Laroque 1996, and Brennan,Williams, and Wright 1997). The second explanation views the Working curve as an artifact of an aggregation problem (Wright and Williams 1989; Brennan, Williams, and Wright 1997). Holbrook Working aggregated U.S. wheat stocks from different domestic locations and for different wheat grades, and then plotted these stocks against intertemporal Chicago price spreads. According to this second explanation, if the Working curve had been drawn using single location prices and stocks, storage under backwardation would not have been found, and the intertemporal theory of prices would be complete without any need to appeal to the *convenience yield* idea.

A graphical presentation of the modern supply of storage function is presented in Figure 1.C.f. This Figure illustrates the notional theoretical behavior of the convenience yield as the supply of inventory varies and the physical storage capacity is held fixed. Heuristically, Figure 1.C.f indicates that when inventory levels are 'normal', fully hedged holding of stocks will earn storage operators a return to compensate for the costs of maintaining storage capacity. This return will be reflected in a forward or futures price which is higher than the current stock by the relevant cost of providing storage and, typically, a small convenience yield to holding stocks. The differential between forward and spot prices is relatively constant across a wide range of inventory levels. However, when inventory levels approach the physical limit set by storage capacity, the costs of providing storage increase and, as a consequence, the convenience yield goes to zero or becomes negative. Conversely, at very low inventory levels stocks are in short supply relative to demand and inventories have a high convenience yield.

As the quote from Carter and Ghia (2007) illustrates, various difficulties have emerged in empirical studies of the supply of storage function due to the problem of accurately specifying inventories. At least since Weymar (1966) it has been recognized that current inventory levels are less relevant for distant futures prices than expected future inventory levels. This point is particularly relevant for seasonally harvested commodities. While Gray and Peck (1981) demonstrate that storage costs for Chicago wheat futures are determined by the supply of immediately deliverable wheat in exchange approved warehouses rather than more general inventory that is not yet immediately available for delivery, Thompson (1986) finds the opposite situation for coffee. Brennan et al. (1997) claim that the aggregation of inventory from different locations gives the misleading appearance that convenience yield is sufficient for inventory to be held when the basis is negative. In effect, when the basis is negative the unobserved ‘convenience yield’ is a statistical artifact created by subtracting storage costs from the basis. However, on balance, available empirical results for direct tests of the supply of storage function do provide support for a non-linear relationship between the basis and inventory levels.

Gao and Wang (2005, p.401) describe the indirect method of testing for the supply of storage proposed by Fama and French:

To avoid the difficulties in determining a suitable measurement for inventory level, Fama and French (1988) proposed alternative indirect tests of the theory of storage to investigate the validity of implications derived from the theory of storage. They used the sign of interest-adjusted basis as the proxies for high and low inventory periods. Fama and French (1988) applied indirect tests on forward and spot prices of five metal contracts traded on the LME with sample daily data from 1972 to 1983. They found that the variability of spot prices is higher than the variability of forward prices during those periods when the interest-adjusted basis is negative (indicating a low inventory period) and the variability of spot and forward prices tends to be equal when the basis is positive (indicating a high inventory period). Furthermore, they found that the variability of basis is greater when the basis is negative than when the basis is positive.

To avoid problems associated with estimating the supply of storage function directly using inventory levels, Fama and French (1987, 1988) build on Samuelson (1965) to employ an indirect method to test for a supply of storage relationship (see Guo and Wang quote). The indirect approach is based on the hypothesis that, when compared to the volatility of forward prices, spot prices will be more volatile than forward prices when inventory levels are low. Both Fama and French and Cho and McDougall (1990) found evidence in favor of the supply of storage hypothesis. Using related methodologies, Ng and Pirrong (1994) and Dutt et al. (1997) found further support for the supply of storage relationship using the indirect approach. On balance, the indirect test is univariate in specification, ignores other elements that impact the basis, and does not take account of inventory data that might be available. To address these difficulties, Gao and Wang (2005) employ “an ARMAX-asymmetric GARCH model framework” to combine direct and indirect tests. The empirical results obtained indicate: “when modeling dynamic behavior of the daily basis of storable (industrial use) commodities, researchers should consider incorporating an asymmetric volatility term in their conditional variance function even if the inventory data is available.” More recently, Volmer (2011) has identified difficulties with the Fama and French approach in the natural gas market.

In addition to studies on the supply of storage relationship, considerable effort has been given to conceptualizing and testing models of convenience yields. There are various threads in these studies.

Some studies extend the direct approach to testing supply of storage by examining the relationship to convenience yield, measured by subtracting storage costs from the basis, e.g., Sorensen (2002), Wei and Zhu (2006). Building on Bresnahan and Spiller (1986) and Heinkel et al. (1990), a number of studies have modelled the convenience yield as a call option on storage, e.g., Milonas and Thomadakis (1997). As inventory levels fall to 'low' levels, the value of this option will increase because of the increased possibility of increasing prices for stocks available for immediate sale. When viewed as a call option, a number of additional factors impacting the option price are suggested, e.g., volatility of the spot price. Building on Longstaff (1995) and Heaney (2002), Zulaf et al. (2006) is a recent example of this approach where implied volatility is used as a measure of spot price variability.

The final substantive group of studies incorporates convenience yield within the more general problem of 'capturing the dynamics of futures prices'. Building on Gibson and Schwartz (1990) and Schwartz (1997), a stochastic process for the convenience yield is assumed in these studies and combined with assumptions about the stochastic process for spot prices for a two factor model plus interest rates for a three factor model. The two factor specification of Gibson and Schwartz (1990) and Hilliard and Reis (1998) takes the stochastic differential equation form:

$$dS(t) = (\mu - \delta(t))\ S(t)\ dt + \sigma_s\ S(t)\ dW_s(t) \qquad d\delta(t) = \kappa\ (\alpha_\delta - \delta(t))\ dt + \sigma_\delta\ dW_\delta(t)$$

where δ is the convenience yield, κ is a speed of adjustment coefficient, μ is the drift of the spot price process after adjustment for the current level of convenience yield, α is the long run average for the convenience yield, σ is the instantaneous volatility for each process, and W_S and W_δ are the Weiner processes for the spot price and convenience yield. Casassus and Collin-Dufresne (2005) extend to a three factor specification which also includes an interest rate process, allowing both spot prices and interest rates to impact the convenience yield.

The Term Structure of Futures and Forward Contract Prices

At any time t, there are a range of futures prices for a given commodity associated with different delivery dates. In liquid markets where standardized forward contracts are traded, there will also be price quotes for different delivery dates. This is apparent from a casual inspection of the futures prices in Figures 1.C.b to 1.C.e. One unresolved question concerns the relationship between $F(t,T)$ and $F(t,N)$, $N \leq T$. Returning to the perfect markets assumption and ignoring fixed costs and returns, the cash and arbitrage conditions for any two futures prices can be specified:

$$F(0,N) = S(0)\ (1 + ic(0,N)) \qquad F(0,T) = S(0)\ (1 + ic(0,T))$$

Taking the ratio of $F(0,T)$ to $F(0,N)$ provides:

$$F(0,T)/F(0,N) = (1 + ic(0,T))/(1 + ic(0,N)) = \{(1 + ic(0,N))(1 + ic(0,T\text{-}N))\}/(1 + ic(0,N))$$

It follows that:

$$F(0,T) = F(0,N)\ (1 + ic(0,T\text{-}N))$$

This relationship holds between any two delivery dates.

The equation $F(0,T) = F(0,N)\ (1 + ic(0,T\text{-}N))$ can be conceptualized as a deferred cash and carry transaction. Consider a trade which established at $t=0$ and is long $F(0,N)$ and short $F(0,T)$. At time $t=N$, the long position is settled by borrowing $Q\ F(0,N)$ at $r(N,T)$ and incurring other possible carry costs $oc(N,T)$ where $cc(N,T) = r(N,T) + oc(N,T)$. The borrowed funds are used to settle the long position by buying the spot commodity at the agreed price $F(0,N)$. This spot position is carried for $T\text{-}N$ periods, where applicable earning a carry return of $cr(N,T)$. At time $t=T$, the spot commodity is then used to settle the short position at $F(0,T)$. This sequence of transactions can be used to specify $ic(0,T\text{-}N)$ as the implied carry cost, reflected in $F(0,T)$ and $F(0,N)$ at time $t=0$, for a cash and carry arbitrage which will begin at $t=N$ and end at $t=T$. Observe that the actual implied carry earned on the cash and carry transaction between N and T, $cc(N,T)$ - $cr(N,T)$, will not be the same as that reflected in futures prices at time $t=0$.

2.2 Optimal Risk Management Decisions

A. Risk Management Objectives

Risk Management Decision Problems

Developing a framework for adequately identifying and managing the range of commodity risks confronting the non-financial corporation is not possible. There is too much variation across the types of commodity risk encountered by the various commodity producing and consuming firms that a general framework is unhelpful at best, and could be misleading. Some method of simplifying the process is needed. 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" (Chance 1998, p.672). In this approach, risk management is closely identified with the types of instruments which 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 financial risk management typically proceeds by classifying risks into categories: 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. For both financial and non-financial firms, the financial risk management approach requires each of these risks to be assessed for the specific firm 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 (2006), emphasize the measurement of market risks using VaR; others, e.g., Oxelheim and Wihlborg (1997), Andren et al. (2005), emphasize the use of CFaR to provide an integrated treatment of business risks.

Still others, e.g., Smithson, Wilford and Smithson (1994), examine the methods for handling commodity or 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 for non-financial firms 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 (see sec.1.3). In all of this, the traditional distinction between hedging and speculation seems misplaced. This unfortunate, as there are useful lessons contained in the earlier discussions of risk management, which typically structure the discussion as a partial equilibrium problem in hedging specific transactions.

The traditional approach to commodity 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 risks, such as the volatility in market prices associated with the introduction of commodity ETF trading or variation in cash flows due to the impact of flexible exchange rates, were still on the horizon. In this simpler world, commodity 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 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.

Profit Functions and Expected Utility Maximization

The profit function is an essential tool in the theoretical analysis of derivative security trading strategies, for both risk management and speculation 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 which can be used to identify relevant features of the trading strategy. One important substitution which 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. This produces the elementary result: *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. Table 2.2 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 $t=1$. Observing $Q > 0$ by assumption, 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.

Table 2.2 **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, π(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.

Table 2.3 **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:

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

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 the Chicago Board of Trade. The relevant transactions are described in Table 2.3. 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 different form.

The intricacies of hedging can be illustrated by extending the grain elevator hedge example in Table 2.3.[7] The futures position in the example implicitly assumes that the elevator operator has grain for sale which 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 which is done using forward contracts. Because $T = 1$ and $F(1,1) = S(1)$ in the case, the forward sale profit function where $Q_A = Q_H$ 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 for simplicity.

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. 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 ($Q_A = Q_H$) 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 hedge 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 have less importance. Intra-commodity trades can be done for different reasons

Table 2.4 **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:

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

which are discussed in Poitras (2002, ch.3). Recalling the use of N and T for the nearby and deferred deliveries of amount Q_N and Q_T, this trade is depicted in Table 2.4. 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 which is short the nearby and long the deferred will be profitable if the difference between the deferred and nearby prices widens, i.e., becomes more positive or less negative. The opposite would be true for the alternative spread, long the nearby and short the deferred.[8]

The Wealth Process

In order to obtain applicability to a range of decision-making situations, the theoretical 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. In commodity risk management applications, the 'asset' could be a crop that is planted or a producing base metal mine investment or an oil sands lease purchase and so on. 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 commodity prices denominated in a foreign currency or investment in a foreign commodity producing asset, there are two random variables involved, the exchange rate and the yield on assets. In other commodity risk management 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.[9] Starting from the given initial level of wealth, the risk manager's objective is to maximize a moment preference function for the value of terminal wealth assuming that the balance (possibly negative) of initial wealth which 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), Poitras and Heaney (1999), Poitras (2002) 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 producing or purchasing A; r is the risk-free interest rate.[10] Manipulation 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.[11]

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, 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 or embedded options in forward contracts, 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:

$$\begin{aligned} W_{t+1}^{z} &= AY_{t+1}P_{t+1} + (W_t - C(A))(1+r) + Q_z(\max[0,\ K - P_{t+1}] - z) \\ &= W_t\ \{x(1+R) + (1-x)(1+r) + \frac{Q_z P_t}{W_t}\ [\max[0,\ \frac{K - P_{t+1}}{P_t}] - \frac{z}{P_t}]\} \\ &= W_t\ \{(1+r) + x(R-r) + \gamma(\max[0,\ -R_P] - \frac{z}{P_t})\} \end{aligned}$$

where K is exercise price on the put option <u>which is assumed to be "at the money" (where $K = P_t$)</u>, 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.[12]

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 which is the decision variable:

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

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 $\underline{RR} = \{P_{t+1}\ \underline{Y}\ A\}/C(A)$.[13] 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 for a farmer 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: π^*_{t+1} is the profit for the futures hedge realized at time $t+1$, x is, again, $(C(A)/W_t)$ the fraction of initial wealth invested in the crop production, 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 typically proceed by employing the expected utility hypothesis where agents rank random prospects according to the expected utility of those prospects, e.g., Elton and Gruber (1995). Analytically, this involves solving problems determining choice variables that maximize an expected utility function. In the basic optimal hedging problem, the associated budget constraint is embedded in the argument of the utility function, e.g., in the specification of the profit function. In contrast, 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 which cannot be directly observed. A key step such optimization problems is to specify a moment preference function which captures the true expected utility mapping for the decision maker.

Given that the central tool in analyzing preferences over random outcomes is the expected utility function, calculation of an optimal solution involves taking conditional expectations, which are conventionally modeled using statistical properties of random variables. This involves 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 (1991). 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 deterministic 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. Asserting that the axioms are difficult to reject lends strong support to the von Neumann and Morgenstern approach.[14] Are such assertions valid?

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 which 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. One of the key difficulties with this approach -- the non-additivity of subjective probability -- was identified during the 'years of high theory' from 1926-39 (Shackle 1967) that produced a number of insightful approaches to resolving the difficult quandaries surrounding the distinction between risk and uncertainty. Knight (1921) identified the ability of entrepreneurs to resolve uncertainty as a source of economic profit. Keynes (1921) used of subjective probability by to motivate the distinction. Shackle (1952, 1958) observed that non-additivity corresponds to an inability to accurate define the sample space.

Confronted with the difficulties raised by uncertainty in forming expectations, George Shackle proposed "non-additive" measures of uncertainty. "After Shackle wrote on this subject, mainly in the 1950s, Probability Theory has come to consider classical probabilities as a special case of monotone measures of uncertainty, of which Choquet capacities are the best known representatives" (Fioretti 2009, p.284; Klir 2006). Following Shafer (1976), evidence theory uses Choquet capacities of infinite order to specify the monotone uncertainty measures. The context is cognitive, "it does not take a gambler as its prototypical subject, but a judge or a detective ... managers making investments [are] ... more akin to a judge or a detective looking for cues than to a gambler looking for luck." A gambler is able to fully enumerate the possibilities that can occur. "On the contrary, judges and detectives know that unexpected proves and testimonies may open up unexpected possibilities" (Fioretti 2009). Beyond this, a number of alternative, not completely compatible, approaches to the handling of uncertainty are possible.

Confronted with the shortcomings of probability theory as a method of handling uncertainty, Shackle proposed "potential surprise", and the associated concepts of 'focus-loss' and 'focus-gain', as a replacement . This concept was a non-additive, cardinal measure that depended on the "certainty of wrongness". As Ford (1993, p.696) observes: "Shackle raised fundamental issues concerned with the individual's perception and measurement of uncertainty: some of those basic issues await resolution." Though the analytical methods Shackle proposed have not gained much traction, "in spite of its conceptual problems and the criticisms it has received, the theory can be utilised to explain many types of economic behaviour under uncertainty" (Ford 1993). Consider the following from Shackle (1952, p.76):

> Investment can be depressed by either an increase in focus-losses or a decrease in focus-gains, and the effect in *both* these cases will seem to be immediate, though in fact the downward movement following an event which is going to increase focus-losses will be due at first to the mere 'standstill' effect of the surprise event. This asymmetry seems at least partly to explain why the downturn of investment and employment after a boom is usually more abrupt and rapid than their upturn after a slump.

Altering the context slightly from aggregate investment to commodity risk management, the Shackle 'potential surprise' approach appears capable of explaining the empirical observation of a more violent downdraft of prices with a bear market compared to the gradual upswing of a bull market. This implies value for hedging programs that protect against downside risk

Expected Utility and the Cost of Risk

Even if the assertion of additive probabilities is accepted, i.e., the sample space for future outcomes is subjectively known, developing various arguments can still be tricky. For example, it is not apparent how to determine the probabilities θ_j in the expected utility function. 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 'the speculative efficiency hypotheis' or the CAPM. However, general equilibrium concerns are of little use in commodity risk management applications. 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 tool. This can be readily demonstrated by applying an essential tool from functional analysis: the Taylor series expansion.

To see this, consider the problem of determining the cost of risk. The solution to this problem would be useful in analyzing 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 which 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 deterministic 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 Taylor series approximation for the deterministic function $U[W_{t+1}]$, expanded again about Ω provides:

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

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

Using $U[\Omega\text{-}C]=EU[W_{t+1}]$, observing $var[W_{t+1}] = var[(1+R)\ W_t] = W_t^2\ var[(1+R)]$ and manipulating gives:

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

This demonstrates theoretically that the cost of risk will vary across utility functions, providing a foundation for widely used 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).

As with other results that rely on a Taylor series expansion to derive a theoretical result, e.g., Poitras (2005, ch.5), subtle changes occur when a higher order of the expansion is used. First order expansions only provide a linear approximation to the function. Such approximations are helpful when the functions are approximately linear or the change in the variable – in this case the cost of risk – is small. For convex and other non-linear functions, sizable changes in the variable will produce inaccurate approximations indicating a second order approximation is in order. Consider what happens when a second order approximation is used:

$$U[\Omega - C] = U[\Omega] + U'[\Omega]\ (\Omega - C - \Omega) + \frac{1}{2}\ U''[\Omega]\ (\Omega - C - \Omega)^2$$

$$= U[\Omega] - U'[\Omega]\ C + U''[\Omega]\ C^2$$

Using $U[\Omega\text{-}C]=EU[W_{t+1}]$ no longer provides a convenient solution. The solution to the resulting quadratic equation now has two solutions, as opposed to one for the linear approximation. Neither solution has the appealing form of the commonly used measures of absolute and relative risk aversion.

Expected Utility and Moment Preference

The relationship between moment preference and expected utility has received considerable attention dating back to the rise in popularity of the mean-variance expected utility to model portfolio optimization problems. 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), Zakamouline and Koekebakker (2009); and, the implications of introducing skewness preference into the mean-variance framework, e.g., Kraus and Litzenberger (1976), Hassett et al. (1985), Lim (1989), Poitras and Heaney (1999). Various sources, e.g., Brockett and Kahane (1992), 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, 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 decidedly more complicated.

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 $[\Omega]$ $(E[W_{t+1}] = \Omega)$:

$$U[W_{t+1}] \;=\; U[\Omega] \;+\; U^{\prime}[\Omega](W_{t+1}-\Omega) \;+\; \frac{U^{\prime\prime}[\Omega]}{2!}(W_{t+1}-\Omega)^2 \;+\; \frac{U^{\prime\prime\prime}[\Omega]}{3!}(W_{t+1}-\Omega)^3 \;+.....$$

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.[15] 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.

Employing what are typically viewed as 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:

$$EU_{MVS}[W_{t+1}] \;\cong\; EU_{MVS} \;=\; U[\Omega] \;+\; 0 \;+\; \frac{U^{\prime\prime}[\Omega]}{2!}var[W_{t+1}] \;+\; \frac{U^{\prime\prime\prime}[\Omega]}{3!}skew[W_{t+1}]$$

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

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 and "≅" means 'approximately equal to' because the higher order terms in the expansion have been ignored. 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? Poitras and Heaney (1999) compare solutions from mean-variance and mean-variance-skewness approximations for the optimal demand for a put option – a negative skewness reducing security. 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. It is demonstrated that under reasonable conditions on the random variable the introduction of skewness into the moment preference function can ***reduce*** the amount of a fairly priced put option demanded compared to the mean-variance solution. This seemingly contrary result follows because the point of maximum positive skewness is interior to the point of minimum variance. This enigmatic result is not related to the Prakash et al. (1996, p.240) claim that "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 which is consistent with the central theoretical condition which 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.

Theoretical exploration of conditions required for mean-variance solutions led Ross (1978) to propose the family of separating distributions. Chamberlain (1983) further characterizes the set of distributions from which the mean-variance expected utility function may be derived as distributions with spherical symmetry, a special case of elliptical distributions. Wei et al. (1999) extend these results to a "linear conditional expectation" (LCE) relation. A well-known distribution that admits LCE is the elliptical distribution class, of which the normal is a special case. Significantly, LCE is also consistent with the Pearson system of distributions. In turn, the skew elliptic distributions, e.g., Genton et al. (2004), can be used to motivate solutions using the mean-variance-skewness expected utility function. Chiu (2010) demonstrates such a generalization is justified. In effect, restrictions on the subjective distributions of the ***ex ante*** random variable(s) can be used to justify the selection of a particular expected utility function for use in determining the optimal solution to problems of decision making under 'uncertainty'. As with other arguments based on Taylor series, further properties of the expansion can raise significant complications.

More generally, continuing with the procedure used to determine the Taylor series approximation for the mean-var-skewness case to include more terms produces:

$$EU[W_{t+1}] \;=\; U[\Omega] \;-\; b\ var[W] \;+\; c\ skew[W] \;+\; d\ kurt[W] \;+\; e\ \sigma^5 \;+\; f\ \sigma^6$$

where: σ^5 and σ^6 are the fifth and sixth central moments of the distribution for terminal wealth; *kurt*[*W*] is the kurtosis, or fourth central moment; and, *d, e* and *f* are coefficients depending on the fourth, fifth and sixth utility function derivatives with respect to wealth.[16] Though seemingly arcane, introducing these additional terms into the expansion raises substantive issues about using moment preference in approximating expected utility. Recalling that an ***ex ante*** subjective distribution is used to determine conditional expected utility, Poitras (2011) examines the limitations of the ergodicity hypothesis to conclude that as the *ex ante* decision time interval increases, the ensemble of future time paths for the random variables determining future wealth could become more bi-modal than unimodal. The fifth and sixth central moments of the distribution assume significance when the density is bimodal.

The immediate theoretical implication of bi-modality in the *ex ante* probabilities is non-linear dynamics for the mean of the process, e.g., Poitras (2011, ch.5) gives numerous references. The inability to reject bimodality of the *ex ante* time path is a consequence of the fundamental quandary arising from the use of non-experimental data to make decisions about the future, as is the case with expected utility modelling: there is only one *ex post* time path available to estimate the theoretically infinite ensemble of possible *ex ante* future time paths for the random variable(s) of interest, such as the spot price at harvest for a farmer or metal price over the life of a mine for a miner. The implications of non-linear dynamics in the mean is complicated for conventional academics studies, e.g., Wang and Tomek (2007). Various theoretical and empirical estimation approaches have been introduced to deal with the failure of the normality assumption while still retaining essential features of unimodality. For example, Poisson jump processes are added to mean-reverting processes to capture manageable random changes in the mean, e.g., Bertus et al. (2009), Bernard et al. (2008). Markov switching models and long memory models with structural change can also be employed, e.g., Coakley et al. (2011). Some approaches, such as the use of GARCH estimation methods, model changes in the variance rather than the mean. The limitations of non-experimental data imply that problems of identifying a 'best methods' 'scientific' approach may not be resolvable on empirical grounds.

A Stylized Risk Management Decision Problem

Both firms and individuals face an array of risks which 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*, which 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 commodities, equities, exchange rates, and interest rates; *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*, which can include inadequate management control systems or incorrect pricing models; and, *legal risks*, which are associated with contract enforcement and variation, regulatory approval and the like.

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. (1980). If there is separability, this implies that it is possible to use a risk management process which considers the problem of managing, say, market risks independently of general business risks. A common theme in some derivatives debacles is the apparent inability to understand the speculative component of the risk management strategies which were being used, e.g., Metallgesellschaft AG. Under seemingly reasonable theoretical 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 some derivative debacles, implementing optimal risk management requires understanding of speculation.

Textbook presentations often portray risk management activities, such as hedging, as eliminating or mitigating 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 which 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 which 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 for a producer to hedge when spot position prices are expected to increase or for a commodity consumer to hedge when prices are expected to fall?

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 involvement, 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, e.g., Elton and Gruber (1995, Alexander, Sharpe and Bailey (2000), Ingersoll 1987). To this end, a ***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 or forward contract speculator who is either long the actual $(Q > 0)$ or short the actual $(Q < 0)$:

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

Assuming that EU is depends only on the choice of Q, the resulting expectation and variance of profit lead to the following:[17]

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

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

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

where $b\ (> 0)$ is a parameter which 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 *ex ante* variance of futures prices; and, the speculator's subjective sensitivity to risk. It is instructive to consider the different solutions which 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 certain conditions, 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 with respect to b.

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. 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 $E[\pi] = 0$ and $b = 1$. This variant 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).

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? Early studies include Johnson 1960, Ederington 1979, Hill and Schneeweis

Table 2.1 **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[π]*, which is the same for both the long and short profit functions:

1982, Toevs and Jacob 1986, Herbst, et al. 1989 and 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 Table 2.1, where short refers to the hedger's position in futures. In addition to the Rahgozar and Najafi (2003) and Chen et al. (2003) surveys, there are a number of recent approaches which employ more advanced estimation methods, e.g., Moschini and Myers (2002), Jin (2007), Bertus et al. (2009), Park and Jei (2010)

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:

$$\frac{dEU}{dQ_H} = 2Q_H\,\sigma_f^2 - 2Q_S\,\sigma_{Sf} = 0 \qquad \Rightarrow \qquad (\frac{Q_H}{Q_S})^* = \frac{\sigma_{Sf}}{\sigma_f^2} = \frac{\sigma_S}{\sigma_f}\,\rho_{Sf}$$

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.[18] This is the operational result which makes the minimum variance hedge ratio empirically attractive and accounts for its widespread use among practitioners. However, despite the 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 unanswered questions about its validity: does $EU = -var$ correspond to optimal solutions for other, more theoretically plausible, expected utility functions?

To see this, consider the optimal hedge ratio which 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:

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

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

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

The mean-variance optimal 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. Similarly, if $F(0,T) = E[F(0,T)]$ due to the willingness of speculators to provide sufficient liquidity at such prices, the incentive to engage in 'active' hedging – where speculation about future commodity price behaviour impacts the risk management decision – is severely dampened.

B. Transactions Hedging and Optimal Hedging

Merton (1993), Tufano (1996) and others conventionally state that corporate risk management can be achieved through diversification, hedging and insurance. In many commodity risk management situations, such as the Tufano (1996) gold mining sample or the Mackay and Moeller (2007) sample of oil refineries, the firms involved have little opportunity to exploit diversification opportunities to manage business risk. In contrast, financial risk management situations, such as globally diversified investment funds or the trading book of an investment bank, diversification is an integral part of risk management. Situations vary and the identification of an optimal risk management strategy depends on both the objective function specified and the characteristics of the commercial situation. It is difficult to formulate general rules, and it is unclear whether solutions to expected utility maximization problems can be specific enough to provide useful guidance. Difficult does not mean impossible and this theme has guided an immense academic literature on developing general rules for possible approaches to corporate risk management involving hedging and diversification.[19] Do such theoretically based general rules provide essential background to aid the firm in developing a risk management philosophy?

As discussed in Part III, observed hedging behaviour often deviates substantively from theoretically prescribed solutions. It is possible to decompose the optimal solution for the hedging problems into a speculative component and a risk minimizing component. Factors determining the speculative part can be significantly different than the elements of the risk minimization problem. Faced with an inability or unwillingness to forecast key price variables, the firm is reduced to seeking risk management solutions aimed at minimizing the variability of a target variable, such as direct transactions cash flow, cash flow from operations, EBIT and so on. As illustrated in Part III, this observation is consistent with the rationales expressed by commodity producing and consuming firms for using derivatives. Yet, derivatives can also be used as an essential component of a business plan, as numerous firms have demonstrated. Integrating a commodity risk management program into the optimal strategic management of the business is the ultimate goal.

Hedging techniques can be illustrated using two approaches. The transactions hedging approach emphasizes the trading mechanics involved in *fully hedging* a specific transaction. A cash position is identified and the appropriate derivative security position is described and determined. It is conventional to have spot and derivative positions that have little or no basis risk, though this does not have to be the case. Various examples of transactions hedges are provided in textbook discussions. The transactions hedging approach does not address what the optimal size of the hedge position needs to be for a specific cash position. This is addressed in the optimal hedging approach. While transactions hedging naively takes the hedge ratio to be one for the cash flow component being hedged, optimal hedging requires estimating the hedge ratio for the firm from empirical data. The transactions hedging approach is the basis for arguments involving the relative benefits of hedging vs. no hedging, such as appear in the free lunch argument for currency hedging. The optimal hedging approach avoids these questions, starting from the premise that full hedging and no hedging are only two of a theoretically infinite number of possible hedge ratios.

The relevance of these two different approaches extends well beyond academic pedagogy, as illustrated by the MGRM saga. This firm was in the business of intermediating the long term forward market for oil. Faced with liquidity constraints on available contract maturities, MGRM employed a rolling stack hedge which featured a concentration of nearby contracts being used to hedge a deferred spot commitment. The hedging decision was modelled using the transaction hedging approach. Mello and Parsons (1995, p.19) draw the following conclusion about the MGRM strategy: MGRM had to

> incorporate the fact of cash flow variability across months into a decision about whether and how much to hedge. When the time pattern of cash flows matters — as it typically does for corporations looking to hedge — a smaller size hedge may be preferable to a one-for-one

rolling stack or other strategy that actually increases initial cash flow variability.

In effect, MGRM used a transaction hedging approach when the situation was theoretically better suited to using an optimal hedge due to the substantial amount of basis risk facing MGRM. However, how an optimal hedge would have been estimated in the MGRM case is not obvious.

The airline industry provides another illustration of the contrast between an optimal hedge and a transactions hedge. Based on information contained in annual reports, this industry is characterized by a range of risk management activities using both exchange traded and OTC derivative securities to manage jet fuel price risk. Some firms, e.g., Thai Airways, engage in only limited use of derivatives, e.g., for hedging specific FX transactions. Other firms, e.g., Singapore Airlines, have a sophisticated program that covers the range of market risks, especially the most volatile cost component – jet fuel prices. A transactions hedge of this commodity price risk would fully hedge the quantity of jet fuel to be purchased over the planning horizon. For example, based on estimated passenger volume and other factors, the airlines will quote ticket prices for flights up to one year in advance. An optimal hedge would attempt to estimate the impact of changes in jet fuel prices on changes in firm net cash flows. In Part III, industry practice among airlines using derivative security strategies is determined by examining airline company financial statements to find what fraction of jet fuel purchases are hedged (see Part III.3).

Price Risk vs. Basis Risk

For present purposes, hedgers are traders using the derivatives market to cover a cash position. While there are some legal and regulatory interpretation problems with this broad definition of hedgers, analytically there is little problem using this approach, though analytical difficulties do arise in handling interaction of the two basic components of hedge design: risk management, which is inherent in hedging, and speculation, which is required in order to design optimal trades. To see how speculation affects the problem, consider the unhedged grain elevator case (see Table 2.6). As a random variable, π will have a conditional mean and variance which can be used to assess the risk (volatility) of the unhedged position (see Poitras 2002, Appendix I). The grain elevator industry is of considerable economic significance. “According to the 2002 Economic Census, grain elevators operated in almost 6,000 locations and employed over 61,000 workers. Grain elevators generated almost $90 billion in sales and revenue.”

Table 2.2 **Profit Function for an Unhedged Grain Elevator**

DATE	**Cash Position**	**Futures Position**
$t=0$	Buy units of grain at $S(0)$ for storage in grain elevator	None
$t=1$	Q_A units are sold at $S(1)$ and loaded for shipment	None

Operations in the grain elevator business are described in US Senate (2009):

> Grain elevators usually purchase grain from farmers with cash purchases or forward contracts which set a specified date in the future for the delivery of the commodity ... Many grain elevators, particularly co-operatives owned by farmers, also sell seed, fertilizer and other items that farmers need. In order to protect themselves from the risk of falling crop prices, elevators usually hedge their cash or forward purchases by entering into futures contracts on the futures exchanges to sell the grain at a price they expect will cover their expenses. Grain elevators that possess grain in storage are said to be "long" in the cash market; when they enter into futures contracts to sell that grain in a future month, they are said to be "short" in the futures market. Once the purchase of a cash crop is hedged with a futures contract, any decline in the value of the crop in the cash market should be offset with a gain in the futures market.

In contrast to the profit function described in Table 2.2, grain elevators are averse to taking unhedged positions for grain in storage.

Application of the definition for conditional variance (where the conditioning notation has been dropped for convenience), $var[\,\cdot\,] = E\{(\pi - E[\pi])^2\}$ provides the result that the conditional variance for an unhedged spot position is :

$$var[\pi] = Q_A{}^2\, var[S(1)]$$

In the case of unhedged cash positions, risk depends on the size of the position and the volatility of the cash price. Despite some stylized textbook treatments to the contrary, hedging does not typically eliminate all the risk of cash price fluctuations. To see this, recall the profit function for the ***one-to-one*** grain elevator hedge from Sec. 2.2.A:

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

It was remarked that the profitability of the hedged position depends on the change in the basis. This discussion can now be extended to observe that the conditional variance for the hedged position is:

$$var[\pi] = Q^2 \, var[F(1,T) - S(1)]$$

In other words, ***a transaction hedge substitutes basis risk for price risk.***

A Variety of Hedge Types

Working (1962) and Leuthold et al. (1989, p.145-6) provide the following taxonomy of traditional hedges. Two important types, optimal hedges and natural hedges, are not identified.

Carrying-charge Hedge: A carrying-charge hedge is associated with the storage of a commodity. A merchant purchases and stores the commodity, and hedges to profit from storage. The merchant seeks to profit from changes in the basis relationship, rather than price level changes.

Operational Hedge: An operational hedge facilitates merchandising or processing operations. A merchant hedges to establish the price of an input or output, often ignoring changes in the basis. The hedge protects the merchant against rapid change in price while a product is being processed or transported. Typical examples of an operational hedge include a flour miller, buying wheat futures to offset a forward sales contract of flour to a baker, or a shipper, exporter or importer selling futures against a cash purchase. These hedges act as temporary substitutes and are liquidated as soon as the trader takes a corresponding cash position.

Active or Selective Hedge adjusts the hedge according to price expectations. The holder of the commodity adjusts the hedge if prices are expected to fall or rise. Selective hedging introduces an additional speculative element to hedging as traders hedge position changes with price expectations. This common hedging procedure is often done to prevent large losses, or variation margin difficulties. Such hedging may be related to an optimal hedging strategy.

Anticipatory Hedge: An anticipatory hedge is usually not matched or offset by a contemporaneous goods or merchandising commitment. The anticipatory hedge serves as a temporary substitute for merchandising to be done later, that is, an expected future cash purchase or sale. The anticipatory hedge involves either the purchase of futures contracts against raw material requirements, or the sale of contracts by producers in advance of the completion of production. For example, flour millers and soybean processors may buy futures contracts in anticipation of subsequent purchases of wheat and soybeans, respectively. Livestock feeders may sell live cattle and live hog futures long before the animals are ready for market. Similarly, grain farmers forward sell their crops before harvest ...

Cross Hedge involves a derivative security position opposite an existing cash position, but in a different commodity. Typically, there is no active futures contract in the commodity corresponding to the cash position, so the trader must select a related commodity for hedging. To be effective, the prices and commodity values of the cash commodity and futures contract must have a fairly high positive correlation. Examples include hedging corporate bonds in the Treasury bond market, grain sorghum in corn, and boneless beef in live cattle.

The importance of understanding basis behaviour for designing effective hedges is apparent from the variance of the one-to-one hedge profit function. Basis information can be used to make various adjustments to improve hedging performance. For example, a refiner seeking to hedge the price of purchasing copper scrap using the high grade copper contract will want to know the basis relationship between the grade of refined copper being produced from the scrap as well as the approximate amount of refined copper which can be produced per unit of scrap. With this information, appropriate adjustments can be made to Q. When will the hedge completely eliminate price risk? This will occur when $F(1,T) = S(1)$. Accomplishing this result requires a delivery hedge where the commodity being hedged is the same as the commodity specified in the futures contract and $T = 1$. In this case, $F(1,1) = S(1)$ and the futures position is satisfied by delivery of the commodity. This result is typically easier to achieve with the use of ***forward*** contracts which can be tailored to match the size, grade, location and other factors which can impact the basis relationship.

C. Optimal Hedge Ratios for Ex Ante Decisions [20]

Multivariate Optimal Hedge Ratio Estimation

The mean-variance optimal univariate hedge ratio is determined by optimizing the expected utility function: $EU = E[\pi] - b\ var[\pi]$. The problem is univariate because a single cash position is being hedged using a single derivative security. The optimal mean-variance solution was shown to be 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* or *ex ante* volatility increases the hedger will be less willing to take positions over or under the minimum variance hedge. Because variances as well as expectations are conditional on the information available on the hedge date, i.e., on the subjective probability assessments of the hedger, then the less capable or willing the hedger is to make forecasts, the less important is the speculative component of the hedge.

In some hedging situations, it is possible to reformulate the hedging problem to have not one but a number of different derivative securities involved in the hedge. One example would be a farmer hedging a crop using both forward contracts and crop insurance, e.g., Poitras (1993). Another example would be an airline with expenses and revenues denominated in a number of different currencies. In this case, currency futures for relevant currencies could be used to construct a hedge that protects the cash flow denominated in the 'domestic' currency from fluctuations in exchange rates. Another example is a commodity, such as tungsten or titanium, where there is no traded hedging contract available. A hedge could be constructed by using an regression analysis to determine an appropriately weighted combination of a number of different futures contracts. The 'costless collars' commonly used by oil and gas producers are another example where a hedge using related commodities that are available can be constructed with combinations of option contracts.

The objective of increasing the number of different derivative security contracts used in the hedge is, ultimately, to improve hedge performance. At some point, this will be a fruitless exercise because there would be so many hedging contracts to monitor and transactions costs would increase accordingly. The obvious question is: how to optimally construct 'multivariate hedges' using combinations of derivative securities? Theoretically, it would be most appropriate to develop an equilibrium model explaining the relationship between the commodity being hedged and the commodities underlying the contracts being used to construct the hedge. The approach used here is to assume that such a model has been specified and to work with the general profit function for a multivariate hedge position. For simplicity it is assumed that only futures contracts are being used to construct the hedge of a commodity position.

Two problems will be considered, the minimum variance hedge and the mean-variance optimal speculative solution. These two solutions will then be combined to produce the mean-variance optimal multivariate hedge solution. As with the univariate hedge solutions, it is assumed that there is one random variable to be hedged, the price risk of the cash commodity position. The size of the cash position is fixed and the relevant components of perfect markets are adopted. Given this, consider the general profit function for a hedge involving k futures contracts and a fixed cash position:

$$\pi(1) = \overline{Q_s}\,(S_1 - S_0)$$

$$- Q_1\,(F_1(1,T) - F_1(0,T)) - Q_2\,(F_2(1,T) - F_2(0,T)) - \ldots\ldots - Q_k\,(F_k(1,T) - F_k(0,T))$$

$$\equiv \overline{Q_s}\,\Delta S - Q_1\,\Delta F_1 - Q_2\,\Delta F_2 - \ldots - Q_k\,\Delta\,F_k$$

where the bar on Q_s indicates that the size of the cash position is not a choice variable. Like the previous univariate hedge where it was assumed that the cash position was long and that the futures position was short when $Q > 0$, this formulation also permits futures to be short or long with the $Q_i > 0$ solution denoting a short futures position matched with a long cash position and $Q_i < 0$ denoting a long futures position matched with a long cash position. It is possible for some of the ΔF to represent option prices.

The key step in transforming the general profit function into a form suitable for ***estimating*** a minimum variance solution is to reformulate the problem in vector space form:

$$y = X\,\beta + u$$

where $y = Q_s\,\Delta S$, $X = (1, \Delta F_1, \Delta F_2, \ldots, \Delta F_k)$, $\beta = (\alpha, Q_1, Q_2, \ldots, Q_k)$ and u is an equation error which captures the unexplained variation in y not accounted for by $X\beta$ and α is the equation constant term. It is possible to normalize this formulation by dividing through by Q_s, in which case β would refer to the hedge ratios Q_i/Q_s. The general formulation with Q_s on the lhs permits the possibility of using income as the random variable to be hedged. It is also useful for deriving the mean-variance optimal solution.

Using $y = X\beta + u$ to specify the problem permits the OLS solution ($\hat{\beta}$) to be immediately specified as:

$$\hat{\beta} = (X^T X)^{-1}\, X^T y$$

In a time series estimation framework, this is an OLS regression estimator where y is a $Tx1$ vector containing ΔS_t and X is a $T\,x\,(k+1)$ matrix containing k columns of $Tx1$ vectors of ΔF_{it} together with a $Tx1$ column of ones to represent the constant term. (This specification for y assumes that Q_s has been used to normalize and that β represent hedge ratios.) For a long cash position, $\beta_i > 0$ indicates the fraction of the cash position which will be hedged with a short position in commodity future i. If $\beta_i < 0$, then this indicates the fraction of the cash position which will be hedged with a long position in commodity future i. Extensions of this solution to other forms of estimation such as GLS, ARCH or 2SLS estimation methods follow naturally, e.g., Lien and Tse (2002), Lien (2009a).

Transforming the solution for the multivariate minimum variance hedge ratio estimator to an equilibrium framework involves taking expectations at the appropriate point to produce the result that:

$$(X^T X)^{-1} = \Sigma^{-1} \qquad X^T y = \{\sigma_{1,s}\ ,\ \sigma_{2,s}\ ,\ ....\ ,\sigma_{k,s}\}^T\ \overline{Q_s} = cov[F,s]\ \overline{Q_s}$$

where Σ is the variance-covariance matrix of the ΔF_i:

$$\Sigma = \begin{bmatrix} \sigma_1^{\ 2} & \sigma_{1,2} & \dots & \sigma_{1,k} \\ \sigma_{2,1} & \sigma_2^{\ 2} & \dots & \dots \\ \dots & \dots & \dots & \sigma_{k-1,k} \\ \sigma_{k,1} & \dots & \sigma_{k,k-1} & \sigma_k^{\ 2} \end{bmatrix}$$

and the individual σ_{ij} refer to variances and covariances for changes in the relevant futures prices. It follows that:

$$Q_{mv} = \frac{Q^*}{\overline{Q_s}} = \Sigma^{-1}\ cov[F,s]$$

This result is identical to the estimation solution, with the proviso that the equilibrium solution involves *ex ante* conditional population parameters while the estimated solution involves *ex post* estimates of the population parameters determined using *T ex post* observations of data on the ΔF_i and ΔS.

As an example of the multivariate solution, consider the case of a minimum variance hedge using two futures contracts. In this case:

$$\Sigma^{-1} = \frac{1}{\det}\begin{bmatrix} \sigma_2^{\ 2} & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^{\ 2} \end{bmatrix} \qquad X^T y = \begin{bmatrix} \sigma_{s,1} \\ \sigma_{s,2} \end{bmatrix} \overline{Q_s}$$

$$\det = \sigma_1^{\ 2}\sigma_2^{\ 2} - \sigma_{12}^{\ 2}$$

The conventional result follows:

$$\frac{\hat{Q}_1}{\overline{Q_s}} = \frac{\sigma_2^{\ 2}\sigma_{s,1} - \sigma_{1,2}\sigma_{s,2}}{\sigma_1^{\ 2}\sigma_2^{\ 2} - \sigma_{1,2}^{\ 2}} \qquad \frac{\hat{Q}_2}{\overline{Q_s}} = \frac{\sigma_1^{\ 2}\sigma_{s,2} - \sigma_{1,2}\sigma_{s,1}}{\sigma_1^{\ 2}\sigma_2^{\ 2} - \sigma_{1,2}^{\ 2}}$$

These results are intuitively the same as those given previously, with the proviso that the sign of β has positive coefficients reflecting short hedge positions combined with long cash positions and negative coefficients representing long futures combined with a long cash position.

As in the univariate hedge case, the optimal mean-variance solution is a combination of the minimum variance solution and the mean-variance optimal speculative solution. To derive the optimal speculative solution, observe that the objective is to maximize expected utility which, in this case, is specified using the mean and variance of speculative profit:

$$L = E[\pi] - b\ var[\pi] = Q_1\ E[\Delta F_1] + Q_2\ E[\Delta F_2] + ... + Q_k\ E[\Delta F_k] - b\{Q^T \Sigma Q]$$

$$= Q^T \begin{bmatrix} E[\Delta F_1] \\ E[\Delta F_2] \\ . \\ . \\ . \end{bmatrix} - b\{Q^T \Sigma Q\}$$

Differentiating the objective with respect to the choice variables Q^T produces the solution:

$$Q^* = \begin{bmatrix} Q_1^* \\ Q_2^* \\ . \\ . \\ . \end{bmatrix} = \frac{1}{2b}\ \Sigma^{-1} \begin{bmatrix} E[\Delta F_1] \\ E[\Delta F_2] \\ . \\ . \\ . \end{bmatrix}$$

where * denotes an optimum value.

Solving the general mean-variance optimal speculative solution for the case of two futures positions gives:

$$\begin{bmatrix} Q_1^* \\ Q_2^* \end{bmatrix} = \frac{1}{2b\ \det} \begin{bmatrix} \sigma_2^2 & -\sigma_{1,2} \\ -\sigma_{1,2} & \sigma_1^2 \end{bmatrix} \begin{bmatrix} E[\Delta F_1] \\ E[\Delta F_2] \end{bmatrix}$$

where det is the same as that given in the minimum variance hedge example given above. Solving for the specific case of Q_1*:

$$Q_1^* = \frac{\left[\sigma_2^2\ E[\Delta F_1] - \sigma_{1,2}\ E[\Delta F_2]\right]}{2b\ (\sigma_1^2 \sigma_2^2 - \sigma_{1,2}^2)}$$

The result for Q_2* follows appropriately.

This result has much the same implications as in the univariate case. As $b \rightarrow \infty$ the speculator's risk aversion becomes so high that the optimal speculative solution goes to zero and as $b \rightarrow 0$ the speculator's risk aversion approaches risk neutrality and the size of speculative position is very sensitive to even small expected changes in futures prices. In this case, the speculator's offer curve sufficient for markets to clear at a futures price that is an unbiased predictor of the future price is somewhat more complicated than in the univariate case. Similarly, the speculator's ability to forecast price changes, as reflected in the size of the elements in Σ, also affects the solution. In the univariate case the result was that if the speculator has limited ability to forecast the next price change then the variance of the forecast (σ_f^2) will be large and the difference between current and expected prices will have to be large in order to induce a significant speculative position. The multivariate case has similar intuition but adjusts this result to account for the covariances between the various contracts involved in the position as well as the variances of the forecasts for the prices of other contracts involved in the position.

Having derived the minimum variance and optimal speculative solutions, it is now possible to derive the mean-variance optimal solution for a multivariate hedge. The objective function for this problem is:

$$EU[\pi_{mh}] = \overline{Q_s}\ E[\Delta S] - Q_1\ E[\Delta F_1] - Q_2\ E[\Delta F_2] - \ldots - Q_k\ E[\Delta F_k] - b\{var[\pi_{mh}]\}$$

where, using the definitions from the previous derivation of the equilibrium minimum variance solution:

$$var[\pi_{mh}] = \overline{Q_s^2}\ \sigma_s^2 + Q^T\ \Sigma\ Q - 2\{\overline{Q_s}\ Q_1\ \sigma_{s,1} + \overline{Q_s}\ Q_2\ \sigma_{s,2} + \ldots + \overline{Q_s}\ Q_k\ \sigma_{s,k}\}$$

The first order condition for futures position 1 gives:

$$\frac{\partial EU}{\partial Q_1} = -\{E[\Delta F_1]\} - 2b\ \{Q_1^*\ \sigma_1^2 + Q_2\ \sigma_{1,2} + \ldots + Q_k\ \sigma_{1,k} - \overline{Q_s}\ \sigma_{s,1}\} = 0$$

where, as before, the * indicates an optimum value. Solving Q_1* involves determining the solution of the other *k-1* positions.

Proceeding to determine all *k* derivatives and expressing the solution in matrix form:

$$\Sigma\ \frac{Q^*}{\overline{Q_s}} = -\frac{E[\Delta F]}{2b\ \overline{Q_s}} + cov[F,s]$$

Inverting Σ gives the solution:

$$Q^* = -\Sigma^{-1}\ \frac{E[\Delta F]}{2b\ \overline{Q_s}} + \Sigma^{-1}\ cov[F,s] = Q_{mv} - \frac{Q_{os}}{\overline{Q_s}}$$

where Q_{mv} is the minimum variance solution and Q_{os} is the mean-variance optimal speculative solution. This verifies the generalization of the univariate result: the mean-variance optimal *ex ante* hedge ratio is determined by combining the minimum variance hedge ratio and the mean-variance optimal speculative solution.

Again considering the example where two futures contracts are being used to hedge a fixed cash position, the optimal mean-variance solution can now be specified as:

$$\frac{Q_1^*}{\overline{Q_s}} = \frac{\sigma_2^2\ \sigma_{s,1} - \sigma_{1,2}\ \sigma_{s,2}}{\sigma_2^2\ \sigma_1^2 - \sigma_{1,2}^2} - \frac{\sigma_2^2\ E[\Delta F_1] - \sigma_{1,2}\ E[\Delta F_2]}{2b\ \overline{Q_s}\ (\sigma_1^2\ \sigma_2^2 - \sigma_{1,2}^2)}$$

From this it follows that if the different futures contracts involved in the hedge have price changes which are uncorrelated ($\sigma_{1,2} = 0$), then the individual hedge ratios will be equal to the univariate solutions. The dependence of the optimal solution on the utility parameter *b* casts doubt on the ***general*** validity of the numerous *ex post* empirical studies which have taken the minimum variance solution to be the optimal solution.

The two period structure of the profit function leads to a number of other qualifications about the generality of the optimal mean-variance solution. Extending to a multi-period framework raises the possibility of readjusting the hedge position over time, producing a different hedge ratio at each hedge readjustment date. In addition, the solution to an inter-temporal, multi-period optimization problem will not necessarily produce the decomposition of the hedge ratio into the minimum variance and optimal speculative solution (Heaney and Poitras 1991). Among other reasons, this is because as prices evolve over the life of the hedge, risk propensities will affect the desire to adjust to observed data.

Optimal Hedge Ratios for Different Utility Functions

Regarding the minimum variance component of the optimal hedge, it has long been recognized that OLS is optimal for only a restricted set of expected utility functions. Various studies have demonstrated that"optimal" futures hedge ratio estimation does depend on the objective function selected, e.g., Chen et al. (2003), Turvey and Baker (1989). With this in mind, it is possible to formulate solutions to the general expected utility problem underlying the hedge ratio optimization. Two types of solutions are considered. Firstly, under the assumption of bivariate normality of the spot and futures returns, a general relationship between the OLS estimate and the hedge ratio implied by a general expected utility function will be specified, and a number of specific expected utility functions examined as illustrations. Secondly, admitting riskless lending and borrowing in the wealth dynamics, the hedge ratio can be shown to be independent of preferences, i.e., depending solely on parameters of the joint distribution of the return generating processes. These results are derived for the traditional, single period "myopic" objective function.

While there are a number of roughly equivalent specifications of the hedger's optimization problem, following Heaney and Poitras (1991) the relevant hedger's problem can be expressed as maximizing the expected utility of terminal wealth for a hedged portfolio with *basic wealth dynamics* determined by:

$$W_{t+1} = W_t\,(1 + R_s(t+1) - h_t\,R_f(t+1))$$

where: h_t is defined to be the ratio of the values (price times quantity) of the spot and futures positions at time t,[21] $R_s(t+1)$ and $R_f(t+1)$ are the t to $t+1$ returns to holding spot and futures and $t\ \varepsilon\ [0....T]$, i.e., $R_J(t+1) = (J_{t+1} - J_t)/J_t$ where J is either the spot or futures price at t and $t+1$. By construction, this specification for the wealth dynamics restricts the problem in order to derive implementable solutions. In particular, the wealth dynamics assumes a single period decision framework with no potential for variation in the quantity of the spot commodity being held.

Another significant feature of this approach to wealth dynamics is the absence of portfolio theoretic considerations. In particular, by not incorporating lending and borrowing into the specification of the profit function, the size of the cash position has been fixed. In effect, the resulting hedging optimization assumes away the portfolio decision by having the hedger fully invested in the spot commodity. A number of sources argue that “farmers can manage their risk exposure through adjusting their leverage, obviating the need for hedging instruments”, e.g., Pannell et al (2008), Simmons (2002). If lending and borrowing is admitted, considerations of leveraging to buy the spot commodity and short-selling the spot to invest in the riskless asset enter the hedger's decision process. Theoretically, this is translated to the *leveraged wealth dynamics*:

$$W_{t+1} = W_t\,(1 + x_t R_s(t+1) + (1\text{-}x_t)\ r(t+1) - H_t\,R_f(t+1))$$

where: x is the fraction of total wealth invested in the spot commodity, H is the value (price times quantity) of the hedge position divided by initial wealth (*not the value of the spot position*) and r is the riskless rate. In turn, the leveraged wealth dynamics can be used as the argument in the hedger's optimization problem. In practice, the primary advantage of using the basic wealth dynamics specification over the leveraged dynamics specification is analytical simplification: the optimal hedge ratio requires only a joint probability distribution and a utility function. The addition of lending and borrowing results in the introduction of an additional choice variable.

Given this, the conventional hedge ratio optimization problem can be generalized to admit any type of well-behaved utility function, Using basic wealth dynamics, the ***optimization problem*** can be specified:

$$\max_{h_t} \quad E\{\ U[\pi(t+1)\ ]\ \} \qquad \Rightarrow \qquad \max_{h_t} \int_{R_f}\int_{R_s} U[\ W_t(1\ + R_s(t+1)\ -\ h_t\ R_f(t+1))\ ]\ d\Psi[F(t),\ S(t)]$$

where the conditional expectation $E[\cdot]$ has been formally defined using an appropriately specified conditioning information set. In this form, the joint conditional probability distribution for $S(t)$ and $F(t)$ associated with the expectation is $\Psi[S, F]$ and the profit function is either $\pi(t+1) = W_t\,(R_s(t+1) - h_t R_f(t+1))$ or, allowing leveraging, $\pi(t+1) = W_t\,(x_t\,R_s(t+1) + (1-x_t)\,r(t+1) - H_t\,R_f(t+1))$.[22] In practice, there are restrictions on the types of commodities for which the *EU* optimization using the basic wealth dynamics is the appropriate hedging problem. For example, because no allowance has been taken of unexpected variation in quantity of the spot commodity, the hedged portfolio would not fully capture the wealth dynamics associated with many harvestable crops.

In terms of solutions, under some strong distributional assumptions, the minimum variance solution leads to an optimal hedge ratio which equals the slope coefficient in an ordinary least squares (OLS) regression of spot on futures prices, e.g., Lien (2005). By construction, OLS depends fundamentally on the selection of joint probability distributions which are constant over time. This assumption results in equality of conditional and unconditional parameters. When time variation in the joint probability distributions is permitted, e.g., due to ARCH errors, the decision problem can be more complicated. In this case, specification of the optimal hedging problem typically takes on a more complicated form and has to be solved using some dynamic optimization procedure, e.g., dynamic programming, which takes account of the state variable time paths. The resulting solutions are potentially intractable and difficult to interpret. However, in the special case of *log utility*, the dynamic solution will reduce to a sequence of one-period solutions (Samuelson 1969). This important simplification permits the introduction of certain types of temporal variation in the conditional variances and covariances without significantly complicating the solution.

In addition to complications arising from non-constant distributional parameters, when the structure of the optimal hedging problem is altered by the introduction of riskless lending and borrowing, variation in the size of the spot position means that the hedge ratio cannot be determined by choosing the relative size of the futures position. There are now two choice variables, the fraction of initial wealth invested in (borrowed using) the riskless asset and the size of the hedged position. Again, while there has been explicit recognition of riskless lending and borrowing, analysis has been restricted to special cases, particularly mean-variance (e.g., Bond and Thompson 1986, Turvey and Baker 1989). In certain special cases (e.g., Poitras 1989), the resulting optimal hedge ratio has been shown to be independent of hedger risk preferences, depending solely on the parameters of the (un)conditional joint distribution of returns.

Given this background, it is possible to derive two Propositions corresponding to the two different formulations of the "myopic" optimal hedge ratio problem, where myopia is a direct consequence of the single period specification of the optimization problem. The first formulation is based on the conventional approach which omits lending and borrowing from the portfolio decision, where the basic wealth dynamics are the basis of the objective function. The second approach uses leveraged wealth dynamics thereby admitting riskless lending and borrowing. In this analysis, "myopia" dictates that future time paths of the conditioning variables are ignored; the trade is initiated at time t and profits are taken at $t+1$. This permits use of the unconditional distributions. Given this, Proposition 2.C.1 extends the conventional constant distributional parameter solution to include a general expected utility function. It is shown that the optimal hedge ratio can be decomposed into the OLS-based hedge ratio (h_{OLS}) and a utility function dependent term. Proposition 2.C.2 incorporates riskless lending and borrowing to determine a market equilibrium hedge ratio which is shown to be independent of the utility function selected.

Conventionally, h_{OLS} has been the foundation of empirical estimation of hedge ratios. Hence, it is important to know the relationship between specific solutions to the expected utility optimization problem and the associated OLS estimate. More precisely, the linkage between the optimization problem and the minimum variance hedge ratio is given by the following Propositions (see Heaney and Poitras 1991 for proofs):

Proposition 2.C.1: **Optimal Myopic Hedge Ratio**

Under the assumption of constant parameter bivariate normality of $R_s(t)$ and $R_f(t)$, the generalized optimal hedge ratio can be specified as:

$$h^* = h_{OLS} + \frac{E[R_f]}{var[R_f]} \frac{E[U'(\cdot)]}{W_t \; E[U''(\cdot)]}$$

where: $E[\cdot]$ is the (un)conditional expectation taken with respect to the joint density; U' and U'' are the first and second derivatives of the selected utility function with respect to π; $var[R_f]$ is the (un)conditional variance of R_f ; h_{OLS} is from an OLS regression of spot on futures prices.

In words, Proposition 2.C.1 demonstrates that, for myopic agents, the optimal hedge ratio can always be decomposed into a sum of the OLS-based hedge ratio and an additional term which is fully determined by statistical parameters and the risk aversion propensity of the selected utility function.

The primary upshot of Proposition 2.C.1 is that in addition to h_{OLS} consideration must be given to the variance-deflated expected return on the futures position. When the expected return is non-zero, the properties of the particular expected utility function assumed, i.e., the inverse of the coefficient of relative risk aversion, takes on importance. Examining the affect of the statistical parameters, an important *general* corollary follows: when the current futures price is an unbiased predictor of the distant futures price ($E[R_f] = 0$) or when $var[R_f] \to \infty$, h_{OLS} is optimal. Hence, results which apply for specific utility functions (e.g., Benninga, et al. 1984, Poitras 1989) can be generalized to any type of admissible utility function, albeit under the restriction of bivariate normality. However, for many commodities, $E[R_f] = 0$ or $var[R_f] \rightarrow \infty$ does not always hold *ex ante* in which case the issue of selecting an appropriate utility function is raised.

To better illustrate, consider some specific examples. Because of the normality assumption, if utility is taken to be negative exponential, $U = -exp\{-\alpha W\}$, this is equivalent to assuming mean-variance expected utility, a case for which a solution has already been provided using a slightly different approach. The resulting optimal solution reflects the different approaches:

$$h^*_{MV} = h_{OLS} - \frac{E[R_f]}{\alpha \; W_t \; var[R_f]}$$

This form of solution also emerges for other methods of generating mean-variance expected utility, such as quadratic utility where $U = \pi - ½\, b\, (\pi - E[\pi])^2$. In order to contrast constant absolute and relative aversion utility functions, consider the power utility function, $U = (\pi^p / p)$, where $p < 1$. In this case:

$$h^*_{pow} = h_{OLS} - \frac{E[R_f]}{var[R_f]} \frac{E[\pi^{p-1}]}{W_t \; (1 - p) \; E[\pi^{p-2}]}$$

For the important specific power utility case of log utility, $U = ln(\pi)$, the solution reduces to:

$$h^*_{\ln} = h_{OLS} - \frac{E[R_f]}{var[R_f]} \frac{E[\pi^{-1}]}{W_t \; E[\pi^{-2}]}$$

From these results it follows that a given optimal hedge ratio depends on parameters of both the conditional distribution and the expected utility function.

Significantly, Proposition 2.C.1 demonstrates that, when $E[R_f] \neq 0$ and $var[R_f] < \infty$, it is not "optimal" to use OLS hedge ratios without making further assumptions about the return and profit generating processes and the form of expected utility. In practice, given specific distributions for the relevant processes, the h^* of interest can be approximated with quasi-maximum likelihood estimation or other numerical methods, e.g., Moschini and Myers 2002; Haigh and Holt (2002). Following Baillie and Myers (1991), it is also possible to empirically model and estimate the conditional hedge ratio, e.g., Kroner and Sultan (1993); Haigh and Holt (2000); Park and Jei (2010); Chang et al. (2011). However, such applications typically result in a substantive increase in the complexity of the estimation problem. Unfortunately, it is not possible to establish, theoretically, whether there will be corresponding increases in the value of the resulting hedge ratio estimates. For example, Alexander et al. (2012) observe:

> Although it is true theoretically that GARCH-based minimum-variance hedging should provide greater variance reduction than naive hedging, due to uncertainty in parameter estimates, and in the correct specification of the GARCH process, this may not be the case in practise. Moreover, typically, the hedge ratios derived from GARCH-type models are extremely volatile, suggesting unrealistic re-balancing of the hedged portfolio. Doing so would amount to large transaction costs

The value-added of more direct specification of the both the expected utility function and return generating processes is still largely an unresolved empirical issue.

Proposition 2.C.2: **Quasi-Separation of the Hedge Ratio**

In the presence of riskless lending and borrowing, the myopic optimal hedge ratio depends solely on expectations and other statistical parameters and is not affected by risk attitudes or initial wealth. Specifically, assuming bivariate normality of S and F the optimal hedge ratio for the leveraged wealth dynamics process is given by:

$$H^* = \frac{var[R_s]}{var[R_f]} \frac{(\beta[F,S]\,(E[R_s - r]\,) - E[R_f])}{(\,(E[R_s - r]) - \beta[S,F]\; E[R_f]\,)}$$

where: $\beta[S,F] = (cov[R_s,R_f] \,/\, var[R_f])$; $\beta[F,S] = (cov[R_s,R_f] \,/\, var[R_s])$.

Within the myopic model the introduction of riskless lending and borrowing alters the objective function such that the leveraged wealth dynamics process is used to specify the optimization problem. The general importance of leveraging and financial constraints in the risk management decision is gradually being recognized in academic studies, e.g., Haushalter et al. (2007); Adam et al. (2007). Solution of the stylized optimization problem with leveraging leads to Proposition 2.C.2 which holds for the case of bivariate normality of the spot and futures price. Because this result depends only on parameters of the joint (un)conditional distribution, the optimal hedge ratio with riskless borrowing and lending is independent of both the specification of the expected utility function and initial wealth. Such separation results are relatively rare and require investigation.

While not immediately apparent, the relationship between Propositions 2.C..1 and 2.C.2 can be seen by evaluating H^* in Proposition 2.C.2 where $E[R_f] = 0$. In this case, $H^* = \beta[S,F] = h_{OLS}$. Significantly, as is the case without lending and borrowing, h_{OLS} is optimal when the current futures price is an unbiased predictor of the distant futures price. Hence, the introduction of lending and borrowing into the hedger's optimization problem does not alter the general result that the optimal hedge ratio is decomposable into h_{OLS} and another term which depends on statistical parameters. However, admitting the ability to short-sell and leverage eliminates the need to consider the risk aversion properties of the selected utility function. The upshot is that, in practice, optimization problems based on (2.C.2) may produce more implementable solutions than those based on (2.C.1).

On a more critical level, as demonstrated by Poitras (1988), Kroner and Sultan (1993), Lien (2009a) and many others, hedge ratio estimation is complicated because both the means and the variance-covariance matrix of the relevant variables are not typically constant through time as required for ordinary least squares (OLS) regression. In addition, because the minimum variance hedge ratio does not take the mean return on the hedge portfolio into account, regression-based hedge ratios will not necessarily be optimal for all types of hedger expected utility functions. Improved *ex post* estimation of *ex ante* hedge ratios suggests the use of sophisticated estimation procedures such as GARCH, e.g., Lien (2009a). However, straight forward implement of ARCH methods assumes that only minimum variance solutions are relevant to the hedger. In an intertemporal model which permits changing means and variances, Heaney and Poitras (1991) demonstrate that minimum variance techniques will only be valid for log utility. Lee and Yorder (2007) rationalize minimum variance by assuming that the prices of futures contracts used in the hedge follow a martingale. Empirical evidence indicating that introducing more advanced estimation methods, such as the Markov switching techniques used by Lee (2009), substantially improves the practical hedge performance over the use of OLS-based hedge ratio estimates is mixed. Similarly, while it is possible to incorporate information on the mean returns of the hedge portfolio into the estimation procedure, this also leads to considerably more complicated estimation procedures.

INSERT Fig. 1.C.h Conditional and Unconditional Corn Hedge Ratio

Figures 1.C.h provides an example from Moschini and Myers (2002) comparing OLS and ARCH hedge ratios for corn hedges and currency hedges. The conventional unconditional hedge ratio is the OLS estimate, which is constant over the sample. The conditional hedge is a GARCH estimate of the hedge ratio. The relevance of conditional volatility models such as ARCH and GARCH for hedge ratio estimation has been examined in numerous sources going back at least to Hsieh (1988). Because a bivariate regression hedge ratio is determined as the ratio of a covariance between the dependent and independent variables divided by the variance of the independent variable, if these parameters change conditionally as time evolves, the associated hedge ratio will also change. This is the basis for the conditional hedge which is reported in Figure 1.C.h. Unfortunately, evaluation of the relative performance of the conditional volatility and OLS hedge ratios is difficult. More importantly, such a performance evaluation is largely irrelevant to the actual risk management practices of non-financial firms.

The numerous academic studies that recommend the use of more sophisticated estimation methods than OLS to determine hedge ratios typically employ statistical criteria such as minimum mean squared error in assessing performance. Following Leitch and Tanner (1991) comparing different hedge ratios using statistical criteria, e.g., by examining the in-sample or out-of-sample minimum mean square forecast error, fails to account for the profitability of the hedge. Significantly, for both corn and currencies, estimates for the OLS hedge ratio is close to one and conditional hedge ratios fluctuate around one providing some support for the use of full hedges, as opposed to no hedging. While attention focuses almost exclusively on the 'fit' of the estimate, the level of hedge ratio estimates goes largely unmentioned. In particular, while hedge ratio estimates – OLS and otherwise – often indicated close to full hedging is optimal, such estimates do not correspond to observed hedging behaviour. In many cases, risk management directives from the Board prevent hedging beyond some fraction of total exposure, presumably to reduce the possibility of corporate speculation.

Determining the Dynamic Hedge Ratio

Heaney and Poitras (1991) and others demonstrate that extending the static framework to the dynamic case involves a reformulation of the underlying the optimization problem. While of considerable practical interest, Propositions 2.C.1 and 2.C.2 are theoretically imprecise due to the assumption of myopia which allows expected utility of terminal wealth to be optimized without considering the entirety of lifetime consumption. This is an important theoretical development because it permits the future time paths of the conditioning variables to affect the optimization problem. As before, the results feature a general expected utility function. In addition, conditional probability distributions with state-dependent parameters are admitted. Useful results can be obtained using the assumption of conditional joint normality of the underlying spot and futures returns. While it is possible to further generalize to include other types of joint distributions, this typically reduces the sharpness of the results, e.g., Bertus (2009).

Ingersoll (1987, p.235-48) demonstrates the problem of maximizing the expected utility of lifetime consumption (not terminal wealth) can be achieved expediently by assuming additive separability of the utility function. Given this, at any time, *t,* the hedger's more general "intertemporal" optimization problem is modelled using the "derived utility of wealth function" (Ingersoll 1987, p.235):

$$J[W_t \; ; \; X_t] = \max_{C_i \, , \; i=\{t,\ldots,T\}} E\left[\sum_{i\,=\,t}^{T} U[C_i] + D_T \mid X_t \right]$$

where C_i is consumption in period i, X_t is the conditioning information and D_T is the bequest function for the terminal date T. The introduction of consumption into the intertemporal problem is dependent on using a different specification for the wealth dynamics. More precisely, ***W for the intertemporal case*** is:

$$W_{t+1} = (W_t - C_t) \left[1 + R_s(t+1) - h \; R_f(t+1) \right]$$

In addition to incorporating consumption, the intertemporal problem involves conditional distributions that require the relevant state variables to be identified. While it is possible to be more general, for present purposes, the rate process (R_s and R_f) for spot and futures prices are the only state variables considered. Given this, the general solution to the optimization problem must now incorporate compensation for the hedger's "nervousness" about future changes in the state variables.

Applying Bellman's dynamic programming principle (e.g., Malliaris and Brock, 1982), the dynamic generalization of Proposition 2.C.1 reveals the corresponding complications (see Heaney and Poitras 1991 for proofs)

Proposition 2.C.3: **Optimal Intertemporal Hedge Ratio**

Under the assumption of conditional bivariate normality of $R_s(t)$ and $R_f(t)$, the generalized optimal hedge ratio H^* can be obtained by solving the "intertemporal" optimization problem using $J[W\,;\,X]$ and W for the intertemporal case as:

$$H_t^* = (1 + \gamma) \; h_{OLS}\big|_{x(t)=c} + \frac{E[R_f]}{var[R_f]} \, \frac{E[J_W(\cdot)}{E[J_{WW}(\cdot)]'} + \frac{E[J_{Wf}]}{E[J_{WW}(\cdot)]'}$$

where:

$$\gamma = \frac{E[J_{Ws}]}{E[J_{WW}(\cdot)]'} \qquad E[J_{WW}(\cdot)]' = (W_t - C_t)\, E[J_{WW}(\cdot)]$$

and where c is the observed values of the state variables *(X)* up to and including *t,* and all expectations are taken conditionally on *X(t)* with:

$$E\left[\frac{\partial J}{\partial W_{t+1}} \mid X(t)\right] \equiv J_W \qquad E\left[\frac{\partial^2 J}{\partial W_{t+1}{}^2} \mid X(t)\right] \equiv J_{WW}$$

$$E\left[\frac{\partial^2 J}{\partial W_{t+1}\ \partial R_{s,t+1}} \mid X(t)\right] \equiv J_{Ws} \qquad E\left[\frac{\partial^2 J}{\partial W_{t+1}\ \partial R_{f,t+1}} \mid X(t)\right] \equiv J_{Wf}$$

In addition to h_{ols} now being a conditional estimate, the role of preferences in the intertemporal optimal hedge ratio is more complicated than in the myopic case. More significantly, h_{ols} is no longer generally optimal when $E[R_f] = 0$.

Specifically, the additional preference-dependent terms arise from expected changes in the state variables affecting the marginal utility of wealth. In this situation, utility function selection takes on added importance. For example, log utility posses the important simplifying property that $J_{ws} = J_{wf} = 0$, which allows the intertemporal solution to correspond directly to the myopic case, with the caveat of potential inequality of conditional and unconditional statistical parameters. Empirical estimation proceeds by assuming a specific form of conditional distribution, e.g., GARCH (Moschini and Myers 2002; see Fig. 1.C.h). Significantly, other important types of utility functions such as quadratic and power are not so well behaved. Estimation of intertemporal optimal hedge ratios for these types of functions can be problematic.

Similar complications arise when riskless lending and borrowing is admitted. In this case, the wealth specification is

$$W_{t+1} = (W_t - C_t)\left[1 + (1 - x_t)\ r(t+1) + x_t\ R_s(t+1) - H_t\ R_f(t+1)\right]$$

In an intertemporal context this leads to the introduction of an additional conditioning (state) variable, *r*, which has to be taken into account. As in Proposition 2.C.3, the risk associated with the potential changes in the state variables must be compensated. This leads to an equilibrium condition for the generalized hedge ratio that has terms that involve $J_{ws,}$ $J_{wf,}$ and J_{wr} i.e., 'indirect utility function' *(J)* terms appears. However, J_{ww} terms do not appear and, as a result, the hedge ratio does not depend on risk attitudes of the hedger. The exact expression for the optimal hedge ratio is quite complicated and not revealing. However, as before, log utility provides an important simplification: because the $J_{ws,}$ $J_{wf,}$ and J_{wr}, terms are zero, the myopic result of Proposition 2.C.2 applies.

2.3 Strategic Risk Management

A. *What is Strategic Risk Management?*

Though much discussion of commodity risk management focuses on market risk, examination of the recent derivative debacles also reveals the importance of operational risk and liquidity risk (see sec. I.3). 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 and *enterprise risk management* are terms often used to describe the process of formulating and implementing a corporate risk management philosophy.[23] Though these terms are often used synonymously, it is possible to define distinct notions. Making this distinction identifies a vital step in the commodity risk management process that receives relatively little attention in conventional academic studies that tend to focus on risk measurement and techniques of hedging, diversifying and insuring.

The multifaceted importance of the risk management function has led to the creation of corporate risk management offices run by 'risk management officers'. Mikes (2008) examines the role of chief risk offices within the corporation. Such offices engage in a whole range of activities which are of little or no direct relevance to commodity 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. Commodity risk management decisions could be made by marketing, purchasing and treasury departments. By design, narrowing the focus to corporate commodity 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 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.

'Enterprise risk management' is now a popular, arguably essential, concept in discussions of corporate risk management practice, e.g., Hampton (2009); Moeller (2007). Unfortunately, this terminology can cover a range of possible notions, creating some semantic confusions. Recognizing lack of precision in these concepts, various notions of 'strategic risk management' and 'enterprise risk management' have been adopted enthusiastically by numerous non-financial corporations. For present purposes, a rough description of enterprise risk management applicable to all the various notions can be formulated as:

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

This definition of enterprise risk management is similar to that given by the Casualty Actuarial Society: "the discipline by which an organization in any industry assesses, controls, exploits, finances and monitors risk from all sources for the purposes of increasing the organization's short- and long-term value to its stakeholders." The goal of enterprise risk management is to deal with the risks facing the firm in a systematic and enterprise-wide fashion, instead of relying on the ad hoc and independent risk management functions which often characterize traditional firm activities surrounding risk.

Enterprise risk management is a multi-faceted concept. For airline companies, a leading concern of enterprise risk management is catastrophic risk of plane crashes and other repair and maintenance issues. Similarly, enterprise risk management for open pit mining companies will be concerned with adherence to environmental standards. While commodity risk management is only one part of the larger task of enterprise risk management, the resulting decision have fundamental strategic implications for actual firm performance. To recognize this key aspect of enterprise risk management, the concept of 'strategic risk management' can be introduced to narrow the focus to only include market risk that impacts on the business strategy of the firm such as the ability to: finance future projects and acquisitions; maintain or increase the level of dividend payments; improve firm profitability; and, reduce the possibility of bankruptcy. As such, strategic risk management is squarely aligned with management of business risks facing corporations and other firms operating in commodity producing and consuming industries.

The importance of the drive to strategically manage risks has not been lost on the management consulting industry. Yet, it is the perceived need for an enterprise risk management function that 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." The now defunct accounting firm Arthur Anderson (1998, p.9) gives the following description of enterprise 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.

The management consulting enterprise risk management approach focuses on operational systems, such as reporting channels, methods of identifying and handling risks, and solutions to information technology requirements. A common recommendation relevant to commodity risk management is that such decisions need to be allocated to the treasury department and not left to the purchasing department.

The notion of 'strategic risk management' is appealing in commodity risk management applications. 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 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 in corporate situations. It is likely that those senior managers responsible for risk management will be an essential cog in the process, due to potentially limited knowledge about specific 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., Rawls and Smithson (1990), Ahn and Falloon (1991), Oxelheim and Wihlborg (1997), Andersen (2006), Damodaran (2008), Andersen and Schroder (2010), identify a specific type of risk management philosophy with 'strategic risk management'. This essence of this approach can be illustrated by examples. Consider Gallo Wines, a company which 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? From a transactions hedging perspective, 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 (2010) 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 measuring and evaluating the risks. There is feedback between the conceptual and empirical aspects. In a corporate context, 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 commodity risk management problems at various firms, e.g., Metallgesellschaft, China Aviation Oil, the Chinese State Reserve Bureau, Amaranth Advisors, 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. 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.

Arguments Concerning the Use of Derivatives in Corporate Risk Management

> "…we find that there is generally no difference in firm values between firms that hedge and firms that do not hedge." Jin and Jorion (2004)

> " ... risk management can add value when revenues and costs are nonlinearly related to prices ... For a sample of 34 oil refiners, we find that hedging concave revenues and leaving concave costs exposed each represent between 2% and 3% of firm value." Mackay and Moeller (2007)

The problem of risk management for the corporation has been well studied using techniques adapted from traditional corporate finance where managers act as agents for the owners of the firm, the common stockholders. Following conventions from traditional corporate finance, 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 leading to potential agency problems. Yet, under reasonable conditions, the primary corporate objective for a non-dividend paying firm 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). In addition, in ignoring the relevance of dividend payout restrictions, a potentially important motive for risk management is not captured. 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. Observing the variables in the fundamental NPV equation, it follows that 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 managers facing exposure to commodity risks must address a natural question: when are risk management actions such as hedging commodity price risk using derivative securities consistent with the primary corporate objective? Consistent with the observed behavior of non-financial firms that do not hedge commodity price risk, a number of persuasive arguments have been made ***against*** hedging commodity price risk and other such market risks, e.g., corporate foreign exchange (FX) risk. Academic arguments against hedging commodity price risk are specific applications of more general arguments which claim, in general, that using derivative securities for corporate risk management activities will not be value enhancing for common stockholders. When such risk management activities are too costly to implement, monitor and execute, firms are generally recommended to forego the direct use of derivative securities to manage commodity risk. In such situations, risk can be managed by indirect methods such as: holding significant cash balances; allowing other entities to forward contract, e.g., small farmers selling to cooperatives; or, embedding forward contracting provisions in marketing contracts, e.g., metallurgical coal producers contracting annually with steel mills to set forward delivery prices.

The general content of these academic 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). Such early theoretical arguments provide the foundation for subsequent empirical studies that explore implications of specific perfect market assumptions, e.g., bankruptcy risk, agency costs, taxes, financial constraints. More recent studies generally support the use of derivative securities to manage commodity price risk, e.g., Aretz et al. (2007), Smithson and Simpkins (2005), Servaes et al. (2009). In addition, some recent studies have progressed beyond questions related to whether risk management is relevant to examine the interaction between corporate strategy, which is impacted by product market dynamics, and risk management decisions, e.g., Haushalter et al. (2007); Adam et al. (2007). In the process, the connection between risk management and the level of cash holdings has been recognized.

In earlier studies, a variety of theoretical arguments were proposed that attempted 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 which belie the importance of such corporate risk management activities correct? To determine the answer to this question, it is helpful to classify the important arguments against such corporate risk management into the following groups: perceived shareholder preference; Modigliani-Miller (MM); CAPM (capital asset pricing model); and, market efficiency (expected value).[24] There is a complementarity among irrelevance arguments that depend on perfect markets assumptions. Because, in practice, the use of derivative securities involves an expenditure of firm resources which would not be required if derivatives were not used, it is argued that the use of derivatives is impractical if that use is not value enhancing. The firm is better off not using derivative securities at all.

Shareholder preference arguments against the use of commodity price risk management arise in many practical situations. An excellent example is Barrick Gold (see sec. 3.1.B), once a leading advocate of gold price hedging, now states (Barrick Annual Report, 2010):

> "Barrick's revenues are primarily derived from the sale of gold and the market price of gold can fluctuate widely due to macroeconomic factors that are beyond our control. Consequently, the market price of gold is one of the most significant factors in determining the profitability of our operations. All of our future gold production is unhedged, providing full leverage to changes in the market gold price. To maximize our realized gold price, we have a corporate treasury function which monitors the gold market and is responsible for our gold sales".

Significantly, Barrick is involved in managing interest rate and FX risk, but not commodity price risk. Canadian Oil Sands (see sec. 3.2.A) which receives a relatively predictable amount of synthetic crude oil at the Syncrude plant gate reflects a similar view, but does identify a possible situation where hedging of commodity price risk would be considered (Annual Report, 2010):

> "Canadian Oil Sands prefers to remain un-hedged on crude oil prices; however, during periods of significant capital spending and financing requirements, management has in the past, and may again, hedge prices and exchange rates to reduce revenue and cash flow volatility to the Corporation. Canadian Oil Sands did not have any crude oil price hedges in place for 2010 or 2009. Instead, a strong balance sheet was used to mitigate the risk around crude oil price movements. As at February 23, 2011, and based on current expectations, the Corporation remains un-hedged on its crude oil price exposure."

COS also does hedge FX risk. The often stated rationale for such risk management inaction is that shareholders want exposure to commodity price risk. Hedging such exposure would frustrate the objectives of shareholders. Presumably such perceptions are filtering through from the Board of Directors where significant shareholdings are usually represented.

In summarizing the remaining 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.[25] 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 which can be done directly by investors.

The MM irrelevance argument relies on 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, e.g., Graham and Smith (1999), this type of motivation does not appear to be widespread in practice, e.g., Graham and Rogers (2002). Siegel and Siegel, (1990, p.150-1) provide an illustration of when taxes could provide a motive for hedging. More importantly, the MM argument is not exempt from the implications of relaxing other 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, e.g., Purnanandam (2008).

Following Frank Knight, firms can earn economic rents from correctly handling uncertainty. Measurable risks, which 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 may also increase the firm's cash flow variability. As demonstrated in the debacles of sec. 1.3, 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. Ultimately, 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 an expected 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 the Levi and Sercu (1991):

> 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 commodity price and FX risk are primarily unsystematic, there are no stock price implications to hedging unsystematic risk. Commodity price risk is not likely to be priced and, if it is, any systematic risk would be incorporated in forward or futures contract prices 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 commodity risk management activities such as hedging.

To better appreciate the CAPM argument, recall the discussion of the future basis. 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 $E[S(T)] = S(0)(1 + E[R_{it}])$, where the discount rate is determined by the CAPM. 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 $F(0,T) = E[S(T)] - S(0)(E[R_{it}] - ic(0,T))$. 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 drilling results estimate an ore reserve 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. Observing that if silver is typically near full carry, then it is possible to assume that $R_f = r(0,1) = ic(0,1)$. 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 which will be marketed in one years time.

Evidence on motivations for non-financial firm hedging from surveys and other sources reveals that the important determinant of firm hedging activity was the desire to reduce the volatility of firm cash flows, e.g., Bodnar et al. 1996, 1998; Servaes et al. 2009. 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, among other problems, the CAPM argument imposes unrealistic empirical conditions on the basis and 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 which 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 case of AMR just prior to filing for bankruptcy in 2011 where the exposure to jet fuel cost meant that a significant change in could more than eliminate firm cash resources needed to avoid bankruptcy; and, the FX exposure of 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$ at a time when the US$/C$ was falling. 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, the non-financial firms involved in the production, transportation and consumption of commodities use real assets that are typically "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 and other factors may prevent their acquisition. In short, in commodity risk management situations it is not usually possible to construct "an efficiently diversified portfolio" of real assets sufficient to achieve a CAPM solution. Again, factors such as the lumpiness of assets and the inability to adequately diversify, means there is an element of uncertainty in business decisions which cannot be reduced to the type of measurable risk argument which 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 zero expected value operation. Due to the costs associated with initiating, monitoring and executing a derivatives program, practical concerns dictate that such activities are unnecessary. To see this, consider the case of a futures hedge. Assume that futures 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 specific theoretical settings. For example, 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 risk neutrality 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 commodity prices, exchange rates, interest rates, and equity prices. A practical objective of this 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 market 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 using futures or forward contracts 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 market prices. The important question is whether, in this situation, firms that successfully pursue active commodity 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 number of largely theoretical arguments against corporate risk management using derivative securities, a considerable number of theoretical and empirical arguments in favor have accumulated. Some of the arguments supporting corporate risk management using derivative securities have already been mentioned in passing. 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." In a survey of available studies, Aretz et al. (2007, p.434) makes the conclusion:

> While there is empirical support for these rationales of hedging at the firm level, the evidence is only modestly supportive, suggesting alternative explanations.

Another recent survey article , Smithson and Simpkins (2005) makes a similar conclusion:

> Is there any direct evidence that risk management increases firm value? The answer is yes, but the evidence is fairly limited as yet. A number of more recent studies show a clearly positive correlation between higher share values and the use of derivatives to manage foreign exchange rate risk and interest rate risk. And one study provides fairly compelling evidence that the use of commodity price derivatives by commodity *users* increases share values. But studies of hedging by commodity producers provide no clear support for the argument that risk management adds value. At a minimum, whether hedging adds value appears to depend on the types of risk to which a firm is exposed.

As hinted at in the last sentence in the quote, a possible explanation for such tepid results is the heterogeneity of commodity risk management situations which defies the search for 'one-size-fits-all' explanations.

Employing Tufano's method of classifying studies extends the conventional corporate finance classifications which are based on the elements of the NPV calculation. Aretz et al. (2007, p.434) identify the various theories supporting risk management activities associated with reductions in the discount rate, increases in expected net cash flow or increases in real option value of projects:

> When there are imperfections in capital markets, corporate hedging can enhance shareholder value through its impact on agency costs, costly external financing, direct and indirect costs of bankruptcy, as well as taxes. More specifically, corporate hedging can alleviate underinvestment and asset substitution problems by reducing the volatility of cash flows, and it can accommodate the risk aversion of undiversified managers and increase the effectiveness of managerial incentive structures through eliminating unsystematic risk. Lower volatility of cash flows also leads to lower bankruptcy costs.

Tufano extends the universe of theories justifying commodity risk management to include characteristics of managers, e.g., the amount of firm ownership, degree of risk aversion.

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. By reducing cash flow variability corporate risk management could lessen the probability of default. *Ceteris parabus,* the incentive to hedge for this reason would be stronger in firms with higher degrees of leverage and lower levels of profitability compared to other firms. Yet, there is not strong empirical evidence to suggest that firms with a higher probability of financial distress engage in higher levels of risk management compared to firms with lower levels. For example, Geczy et al. (1997) and Mian (1996) are unable to find a link between derivatives usage and the capital structure of non-financial firms while Haushalter (2000) and Graham and Rogers (2002) find some support. The heterogeneous 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 that 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. (1994), 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 are 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. Various empirical studies provide mostly favorable evidence on the under-investment hypothesis, e.g., Gay and Nam (1998), Adam (2002), Carter et al. (2006). For example, using a sample of 325 firms using derivatives combined with 161 firms not using derivatives, Copeland and Copeland (1999, p.74) 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" .

Other promising explanations for corporate risk management have been advanced in academic studies. 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 within 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. As in the argument against the CAPM approach, real assets can 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.

B. Designing a Commodity Risk Management Philosophy

Types of Risks to be Managed

Risk management is a diverse subject, with applications ranging from medicine to engineering to finance to political science.[26] Because risk is a pervasive phenomenon, methods for managing risk are a necessary 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 which apply to most risk management situations, in practice it is sensible to narrow the focus to the specific types of risks most relevant to the decision problem at hand. This narrowing of focus is not without difficulty. Treating risks individually can over simplify the problem, ignoring the complementarity which 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 directly relevant to commodity risk management problems facing firms and individuals.

The theoretical rationales for corporate risk management using derivatives provide some foundation for the discussion of practical issues involved in commodity risk management decisions and activities. 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 (1981), Poitras (2002), Fisher and Kumar (2010). It is in formulating answers to the various questions that essential elements of a risk management policy become apparent. To this end, consider the following sequential list of questions which are of particular relevance for a firm considering the implementation of derivatives trading program for commodity risk management purposes:

A Simple Guide to Designing a Risk Management Philosophy

What are the firm's aggregate and specific risk exposures? How is risk to be measured? This initial step requires some data and analysis. It is essential to make detailed calculations of the possible losses if no risk management actions are taken. Adjustment for expectations about future movement in prices may be incorporated, producing a range of scenarios. Unfortunately, in many cases, the risk measurement calculations required are not obvious. For example, Gallo Wines was for many years a US wine producer which 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, if difficult to measure, economic 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 that generated a variation margin liquidity crisis associated with exchange traded derivative hedges. Without a traded market price for the long term forward contracts, it is difficult to provide a precise objective estimate of a change in value of the spot position when the futures market 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), which underpin the optimal hedging strategy and the transactions hedging strategies discussed previously. Accounting exposure measures on a transaction by transaction basis. This approach is reflected in conventional textbook presentations of risk management involving a derivative security hedge which assumes that there is only one transaction of interest being hedged. For example, a US meat exporting 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 which will take place in three months. Similarly, MGRM identified market risk using an accounting exposure. As a consequence, the size of MGRM rolling stack hedge position was set to be approximately equal to the number of barrels in the long term delivery contracts.

Where individual transactions need be considered, accounting exposure is a useful measure. Yet, in many situations there are numerous transactions which contribution to the risk exposure. These transactions can involve a large number of commodity and financial prices. Economic exposure measures attempt to assess the impact of a specific commodity or financial price on the firm's net cash flow. An important example is the economic exposure of an airline to change in the price of jet fuel. While the *ex post* accounting exposure would equal the amount of fuel consumed multiplied by the unanticipated change in jet fuel prices, the 'economic exposure' would also depend on the impact that jet fuel prices had on competitors. Airlines with more fuel efficient planes that are also able to keep other costs down could gain competitive advantage if other airlines were forced by rising jet fuel cost pressures to cut routes and raise fares. In such cases, a less than full transactions hedge would be appropriate.

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. Profitable firms may use derivative security trading as part of a liquidity management program. 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 coal prices are an important component of cost variability in the steel industry. Firms that hedge metallurgical coal prices through forward contracts with metallurgical coal producers may be able to gain market share from those that do not hedge due, say, to being able to deliver steel at prices where high spot coal prices have squeezed profits at competitors to the point where it is not profitable to bring steel to market.

In some situations the decision to trade derivatives may be imposed by lenders. An example of this occurs with intermediate Australian and Canadian mining companies which are required by the banks making them loans to implement a hedging program to cover future mine output as a condition of being granted credit. As banks are reluctant to lend unless there is actual mining operations in place producing identifiable output, it is unusual for non-producing mining exploration ventures to obtain significant bank financing. A detailed example of this situation, Capstone Mining Corp., is provided in sec. 3.1. Such bank dictated 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 metal prices. Affordable for management is not necessarily consistent with affordable for shareholders which is not necessarily consistent with affordable for bond holders. Similarly, firms may want to use hedging to protect the ability to make future dividend payments to common and preferred stockholders.

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, B.C. Ferries, the government monopoly ferry service in British Columbia is able to arbitrarily add surcharges to regulated ferry fares when fuel costs are higher than expected, instead of hedging the full expected amount of fuel usage over the period the regulated fares are mandated. Some government-linked Canadian electric utilities, such as BC Hydro, have a substantial portion of debt denominated in US dollars while the bulk of cash flows are in Canadian dollars. Large, adverse changes in the exchange rate which would be sufficient to eliminate the net asset value of a private company, would not have the same impact on the government-owned utility. This situation is further complicated for some Canadian utilities, such as BC Hydro, that do considerable business selling surplus power to US consumers.

Can the risks be hedged? Are other methods of risk management applicable? This is the problem of hedge design, a topic which is the central concern of Poitras (2002, ch. 6). There may be a variety of possible hedging techniques which have to be considered. An important practical concern is whether there are derivative contracts available which qualify as feasible hedging instruments. In many commodity markets, forward contracts are available which allow the cash position to be matched with the commodity underlying the forward contract. In some cases, no forward contracting method is available and a ***cross hedge*** can be executed using exchange traded derivative securities. Even in cases where forward contracting is available, 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 which is written for a commodity which differs from the cash commodity, e.g., Borger (2009). For example, a copper scrap dealer can cross hedge using copper futures contracts which feature copper cathodes as the deliverable commodity or a wheat farmer in S. Dakota can cross hedge with Chicago-deliverable wheat futures. 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 have been discussed in sec. 2.2. Information on basis relationships is an essential element in determining the size and type of the hedge position to be initiated. The time series behavior of the basis can also be used to make inferences about the current and near-term future state of cash markets.

What are the costs of hedging? Costs will vary depending on the intellectual capital available within the firm available for risk management activities. In large corporations, there needs to be sufficient investment in internal control systems and corporate governance structures to ensure all relevant costs are made available to senior management. Given this, cost efficiency requires 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 and possible variation margin costs which need 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 oversee hedging operations. There will typically be considerable variation in the specific costs associated with various potential hedging instruments and programs. Firms that use OTC traded contracting solutions need to account for potential costs associated with liquidity and credit risks.

What are the tax and accounting implications of hedging with derivative securities and other risk management activities? The relevant issues involved here are discussed in other sources, e.g., Okochi (2008), Zhang (2009). These issues are not incidental and will have to be determined in order to precisely calculate the costs and feasibility of risk management activities such as hedging with derivative securities. In particular, the introduction of FAS 133 and IAS 139 raises a host of questions and queries which lie outside the confines of the present inquiry. Small to medium sized non-financial firms often rely on outside consultants, e.g., Accenture or Deloitte, to handle the preparation of relevant derivative related tax and accounting items. Emm et al. (2007) provide a valuable discussion of accounting disclosure issues relevant for risk management. Pincus and Rajagopal (2002) examine the impact of hedging on accrual management.

By design, this general framework for designing a risk management philosophy cannot deal with all important issues which 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.

C. Shareholder Value and Corporate Currency Risk Management

Measuring Corporate Economic Currency Exposure

There are various reasons to extend the commodity risk management umbrella to include FX risk, even though FX risk is technically in the financial realm. In particular, many commodity prices are determined in global markets where the US dollar is the unit of account. Gold, oil, grains and other commodities could be increasing in US dollar terms but stagnant in Canadian dollar terms. The increasing globalization of commodity markets has put increasing pressure on management of non-financial firms to determine the appropriate method of handling exposure to currency fluctuations. Some non-financial firms that do not hedge commodity risk do hedge FX risk. 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 often concerned with the implications of accounting rules, contained primarily in FAS and IASB standards, which deal with the handling of accounting items which are denominated in foreign currency. While useful within an accounting context, there are many forms of corporate currency exposure which are not captured using this measure.[27]

Economic currency exposure is a broader concept which 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 Shapiro (1992, p.225) suggests, "the focus must be not on nominal exchange rate changes, but instead on changes in the purchasing power of one currency relative to another." 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 purchasing power parity (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 which 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 which 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 which 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.[28] 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 which 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. Where there is flexibility in forward contract terms, hedging outcomes can be achieved by, say, embedding forward FX rates in contract language. The other general type of strategy for managing corporate currency exposure involves natural hedges which are dependent on strategic multinational firm management decisions. These long horizon techniques apply to corporate net cash flows that are relatively indeterminate, consistent with cash flows that originate from many real assets of commodity producing and consuming firms.

Natural hedging strategies for managing corporate currency exposure involve assessment of the competitive exposures which 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 to employ natural hedging techniques. This is an essential, if not always well understood, point. Many commodity risk management situations where non-financial institutions deal with FX 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 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 which matches the foreign currency exposure.

Measuring Corporate Exposure to Currencies and Commodity Prices

Against this heuristic background, attempts have been made to provide a more formal 'scientific' approach to measuring corporate exposure to currencies to currencies and commodity prices. At least since Adler and Dumas (1984), the use of regression analysis has been proposed to identify the impact of changes in the nominal exchange rate on the domestic currency value of the firm's net cash flows or stock prices. For example, Patnaik and Shah (2010) use changes in firm stock price (measured as a rate of return in domestic currency, $R_{i,t}$) regressed against changes in a stock market index ($R_{M,t}$) and nominal exchange rates ($R_{FX,t}$), both measured as rates:

$$R_{i,\,t} = \alpha + \beta_1 R_{M,\,t} + \beta_2 R_{FX,\,t} + u_{i,\,t}$$

The resulting estimated slope coefficient on $R_{FX,t}$ is proposed as a proxy for currency exposure. This methodology for measuring corporate FX risk exposure is also used to measure exposure to commodity price risk. For example, substituting in the regression equation the percentage change in Gulf Coast jet fuel prices for the rate of change in FX rates, Carter et al. (2006b) use the same methodology to measure the exposure of US airlines to jet fuel costs. While 'scientifically' appealing, this regression approach to measure corporate risk exposures suffers from a number of difficulties and potential shortcomings associated with the heterogeneity of firms.

In addition to well known difficulties associated with 'market model' estimations of CAPM-based relationships, measurement of exchange rate changes can be challenging. While some studies take the view that the change in the US$ exchange rate is the appropriate measure, e.g., Patnaik and Shah (2010); Parsley and Popper (2006), other studies recommend trade weighted measures, e.g., Dominguez and Tesar (2006). For US firms, where some decision about the construction of a trade weighted 'exchange rate' measure would be necessary, specification of such a measure is problematic. 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. Such changes could be due to mergers and acquisitions, or significant changes in unhedged commodity positions. 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 which 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.

PPP Arguments

Though the roots of 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 are the foreign and domestic prices of the good with S being the spot exchange rate. Norman (2010, p.919) observes: “The concept of purchasing power parity (PPP) is one of the most empirically well studied theories in international economics, perhaps because evidence of its existence has been so elusive.”

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 in (Poitras 2002, Table 4.2.1). Casual examination of this Table reveals that real exchange rates for many currencies does deviate significantly from the PPP requirement that the real exchange rate is relatively constant over time.

The basic approach of the PPP arguments against hedging currency risk 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.[29] 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. 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. The persistence of deviations from PPP in the face of a volatile real exchange rate led Rogoff (1996) and others to identify the "Purchasing Power Parity Puzzle". Rogoff (1996) estimates the half lives of shocks that generate deviations from PPP to be on the order of three to five years which "seem to be extremely long even when PPP is viewed as a long run concept" (Norman 2010, p.919)

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

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

In the context of an international firm, an early study by 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. Similarly, Chowdhry and Howe (1999, p.229) find: "For firms with plants in both a domestic and foreign location, the foreign currency cash flow generally will not be independent of the exchange rate and consequently the optimal financial hedging policy cannot be implemented with forward contracts alone ... the optimal financial hedging policy can be implemented using foreign currency call and put
options and forward contracts."

C. The Farmer's Commodity Risk Management Problem

Farming and Risk Management

The risk management problem for producers of agricultural commodities is unlike that for non-financial firms involved in metals and energy production. Barnett and Coble (2009, p.36) describe the general situation in the US farming sector:

> Agricultural production is still quite concentrated with less than 6% of the farms in the United States producing 75% of the value of production ... Most U.S. farms still produce undifferentiated commodities for markets where production is characterized by relative ease of entry and exit. And farming is still a risky business.

For the metals and energy complex commodities, information about commodity risk management practices is relatively thin, available risk management tools are limited, and capital costs and other factors can pose significant barriers to entry, e.g., ICME (2001). In contrast, agricultural risk management features a variety of commodity risk management tools, including crop insurance, exchange traded derivatives, forward contracts with elevators, disaster and other income assistance, production contracts, crop diversification and cash on hand, e.g., Harwood et al. (1999). The vast scope of studies on agricultural risk management extends into the Third World where various NGOs and international agencies have produced numerous studies on the topic, e.g., UNCTAD (1997), World Bank (1999).

INSERT Figure 2.3.m Price Trends for major field crops 2000-09

INSERT Figure 2.3.n Monthly percentage price changes for corn, wheat and soybeans

On the CME, commodities within the agricultural complex are divided into: grains and oil seeds; livestock and dairy; forestry; and, the soft commodities, coffee, sugar, cocoa and cotton. Of these, grain and oils seeds exhibit the most seasonality due to the harvesting cycle. As illustrated in Figures 2.3.m and 2.3.n, the volatile price behaviour of agricultural commodities observed from 2006-2011 poses significant risk management issues. Giot (2003) discusses the possibility of using implied volatility from agricultural options to predict near-term price changes. Despite the considerable volatility in prices, there are various sources that identify limited use of risk management products by US farmers. For example, Niemeyer (2008) stated:

> Several recent estimates indicate that a relatively small percentage of farmers directly use the futures markets for hedging. Instead, most farmers rely upon grain elevators and other grain buyers to obtain forward contracts for their crops. The National Corn Growers Association recently stated: "By one estimate, probably less than 10% of farmers are directly using the futures market for risk management."

Based on numerous studies by the USDA, such views are somewhat misleading as it is the limited amount of interest in using futures contracts that is being observed, not the overall use of risk management tools. USDA sources such as the Agricultural Contracting Update (e.g., Macdonald and Korb 2008, 2011) and Macdonald et al. (2004), combined with results from some academic studies provide a much clearer picture of the changing risk management landscape.

INSERT Fig. 2.3.p Characteristics of US Farms and Use of Marketing Contracts
INSERT Figure 2.3.q Farmer Usage of Risk Management Products

The important and useful report by the US Senate Permanent Subcommittee on Investigations on excessive speculation in the wheat market (US Senate 2009) is a gold mine of information about the workings of that market. For example, consider the following recognition of the impact that increased variation margin costs resulting from price volatility have on the incentive for grain elevators to hedge:

Even if an elevator is completely hedged – so that the elevator will have "locked in" a gain regardless of the direction of the market – a steeply rising market can impose significant additional costs upon the elevator operator. In a rising market, grain elevators and merchants that have hedged by selling futures may be subject to margin calls from the exchanges to cover the loss in value of their "short" positions. These margin calls, which are made at the end of each trading day, require payments by the grain elevator or other party to the futures exchanges into a margin account. The amounts in the margin account are not recovered by the elevator until the short position is closed out – in this case, until the elevator sells its grain and terminates the hedge. If a grain elevator cannot make the requisite margin payments, the exchange will close out its position at the current market price, possibly causing further losses. In 2008, rising grain prices in the cash markets, together with rising margin calls, required many grain elevators to make much larger cash outlays than normal. The National Grain and Feed Association estimated that a typical grain elevator faced a 300% increase in hedging costs in 2008, compared to 2006. It stated that "recent commodity price increases have led to unprecedented borrowing by elevators – and unprecedented lending by their bankers – to finance inventory and maintain hedge margins."

The characteristics of farming differ by crop and geographical region, e.g., Parkinson (2000). Despite this, Harwood et al. (1999) examine the comprehensive 1996 US Agricultural Resource Management Study (ARMS) data to discover:

> Keeping cash on hand for emergencies and good buys was the number one strategy for every size farm, for every commodity specialty, and in every region.

Figure 2.3.p provides further examination of results from the ARMS data, identifying both farm characteristics and use of marketing contracts for risk management. Examining the results for the use of marketing contracts, raises some issues when the following from Mark et al. (2008, p.22) is considered:

> Goodwin and Schroeder (1994) found in a sample of Kansas producers that only 11% hedged any of their grain using futures. Schroeder [et al.] (1998) summarized several studies that consistently showed that more producers used forward contracts than used futures hedges. These studies showed that 42–74% of producers used forward contracts to price any of their grain.

Examining Figure 2.3.p with Figure 2.3.r which was calculated from a survey of a smaller number of farms reveals that the ARMS data for 'marketing contracts' in wheat at just under 10% likely relates to the use of futures contracts. However, sugar beets, a relatively non-storable commodity for which there is no actively traded futures contract available has 87% usage of marketing contracts. As a consequence, some care is needed when using ARMS data and the related Census of Agriculture data (www.nass.usda.gove/census/) to compare usage of different types of derivative security contracting across field crop producers, e.g., agreements covering the sale of harvested commodities in storage are not defined as agricultural contracts in ARMS.

INSERT Figure 2.3.r US Agricultural Contracting

For grains and oilseeds, one common method for hedging to be implemented is for grain elevators to provide forward price quotes to farmers. Henderson and Fitzgerald (2008) describe the process:

> In a forward contract, an elevator agrees to purchase a quantity of grain from a farmer at a specified quality or grade to be delivered on a future date at an agreed on price. Forward contracts are typically consummated pre-harvest, allowing farmers to guarantee a crop price and eliminate the risk of falling crop prices as harvest approaches

As the various USDA studies on agricultural contracting detail, this seemingly simple process that has characterized commodity risk management in this sector since the 19th century has been changing as the landscape of agricultural production has changed:

> Farm production is shifting from smaller to larger family farms and from spot (or cash) markets to contracts. Technological developments may underlie much of the shift to larger farms, but expanded use of production and marketing contracts supports that shift by reducing financial risks for farm operators. For farm operators, contracts provide benefits from reduced risks, but also result in loss of managerial control and reduced autonomy.

As illustrated in Fig. 2.3.r, though the percentage of agricultural producers using contracting methods to manage commodity price risk is small at 10-12%, the fraction of total production is much higher at around 40%. This is a result of the small percentage of farms that comprise the largest agricultural producers being the most reliant on contracting methods. "Nearly 70% of the largest farms (those with at least $1 million in annual sales) used contracts in 2008, compared with 7 percent of small farms. Contracts covered 49 percent of production among the largest farms, compared with 16 percent among small farms (those with less than $250,000 in annual sales)" (Macdonald and Korb 2011, p.9). Nevertheless, it cannot be overlooked that: "Most transactions for US agricultural products are conducted through spot market exchanges in which commodities are bought and sold for immediate delivery" (Macdonald and Korb 2008, p.iii).

In addition to farm size, there considerable diversity in commodity risk management practices across sectors in the agricultural industry, e.g., Musser and Patrick (2002). Certain agricultural segments – especially hogs and poultry – use production contracts as a mechanism for dealing with price and production risks. For example, Dimitri et al. (2009) examine the transition from cash trading to contracts in the broiler industry. Such contracting methods (Macdonald and Korb 2011 p.1-2):

> specify services provided by a farmer for a contractor who owns the commodity while it is being produced. The contract covers (1) the services provided by the farmer, (2) the manner in which the farmer is to be compensated for the services, and (3) the specific contractor responsibilities for provision of inputs. For example, farmers provide labor, housing, and equipment under livestock and poultry production contracts, while contractors provide such other inputs as feed, veterinary and livestock transportation services, and young animals. The farmer's payment resembles a fee paid for the specific services provided by the farmer, instead of a payment for the market value of the product. Since contractor-provided inputs may account for a large share of production costs, the fee paid to the farmer may be a small fraction of the commodity's value.

In situations where the farmer also has obtained financing for capital investments, and especially if the financing was obtained from the contractor, such arrangements can lead to significant problems. For example, the length of time required to repay the financing on the capital investment can be longer than the term of the pricing agreement leading to 'holdup' costs where the contractor is able to extract a much lower fee to the farmer when the pricing arrangement is renegotiated at the end of the term. In general, farms with more debt also tend to be greater users of contracts (Key 2004).

The Farmer's Optimal Hedging Problem

One of the limitations of applying the conventional optimal hedging model to farming is that there is only one source of uncertainty which is hedged, e.g., Moschini and Hennessy (2001). For example, this is reflected in the assumption of a non-stochastic cash position when there is uncertainty about both the price at harvest and crop yield. Another related example occurs when hedging the domestic currency value of a commodity being priced in foreign currency, where both the exchange rate and the foreign commodity price interact to determine the variable to be hedged, e.g, Batterman and Broll (2004). Hence, in a number of hedging situations, it is not practical to assume there is only source of uncertainty. In the case where both price and quantity to be hedged are random, it is ***income*** and not price that is the variable to be hedged. The relevance of hedging farm revenue instead of examining price and quantity risk separately has received considerable attention, e.g., Acs et al. (2009), Guinvarc'h et al. (2004), Hart et al. (2001), Poitras (1993). Various sources have also identified the amount of farm leverage as an important element in hedging decision, e.g., Shapiro and Brorsen (1988).

The decision problem is motivated by the stylized farming situation, at planting time the farmer must estimate the size of the future crop in order to determine the size of the hedge. More generally, this problem also applies to hedging situations where future inputs to the production process have to be estimated. For the farmer's problem, because Q_S is not fixed at $t = 0$, it must be estimated. At $t = 0$ the farmer plants with the objective of reaping $\hat{q}$ where $\hat{q}$ is generated by some production function, $Q[\, K , L , Land \,]$, and offered for sale at some expected price, $E[S(1)]$, where K is the capital stock used in production and L is the labour input. Even at this basic level, the complexities of the farmer's problem are apparent. To make the problem more manageable, a number of simplifying assumptions can be used.

In the analysis of the minimum variance hedge ratio, it was shown that $EU = -var$ was the associated objective function. This functional form can be compared to the mean-variance EU. Given the similarities in the mean-variance and minimum variance optimal hedge ratios, it is possible to rationalize the minimum variance approach by assuming that hedgers, in this case farmers, do not forecast spot prices. In effect, the price in the future is expected to be the same as it is today. This form of myopia permits use of the minimum variance hedge ratio by eliminating the need to consider expected price change. Another simplifying assumption is the requirement that the hedge position put on at $t = 0$ is held and not changed until harvest at $t = 1$. More formally, no dynamic hedge adjustment is permitted. Given these two assumptions, two situations can be considered: where $\hat{q} = Q_H = Q$, i.e., there is no cross hedging of estimated output; and, $\hat{q}$ and Q_H are allowed to differ, i.e., there is cross hedging. If farm output realized at $t = 1$ is taken to be $\tilde{q}$, then it is possible to define the profit function for the farmer doing a delivery hedge (where no cross hedge is involved).

Creating a new random variable $\tilde{q}(1)\, S(1) = Y(1)$, which effectively represents future farm income, it is possible to define an associated variance of profit function which can be used to solve for the minimum variance hedge ratio:

$$var[\pi] \;=\; \sigma_Y^{\,2} \;+\; Q^2\, \sigma_F^{\,2} \;-\; 2Q\, \sigma_{YF}$$

$$\frac{\partial var[\pi]}{\partial Q} \;=\; 2[Q\, \sigma_F^{\,2} \;-\; \sigma_{YF}] \;=\; 0$$

$$Q^* \;=\; \frac{\sigma_{YF}}{\sigma_F^{\,2}}$$

While interesting, this form of the optimal hedge ratio has little practical value, if only because it is difficult to determine σ_{YF}. Expanding the solution to allow for Q_H to differ from the expected size of the crop position does not substantially change the practicality of the solution. The primary analytical difficulty is the presence of a random variable, farm income, which is the product of two random variables. Only one of these variables, spot prices, is hedgable. With this in mind, it is possible to reconstruct the optimal problem into the form used in determining the optimal myopic hedge ratio and the quasi-separation of the hedge ratio.

The basic model is discrete. Farmers have access to a variety of possible risk management instruments to hedge production decisions. The representative grain farmer plants a crop at time t and harvests it at time $t+1$. Both the price at harvest and the quantity harvested are unknown at time t, the date the relevant risk management and planting decisions are initiated. As conceived here, in addition to choosing the usage of hedging instrument(s), the farmer's optimization problem also involves choosing the amount of initial wealth to invest in crop production. Hence, the production decision is treated in a portfolio context. As a result, the costs associated with planting the given acreage are also determined. Starting from a given initial level of wealth, the farmer's objective is to maximize the value of terminal wealth assuming that the balance (possibly negative) of initial wealth which is not allocated to planting costs will earn (pay) the riskfree rate of interest.

Given this basic structure, it will initially be assumed that the only hedging instrument available is forward contracts with a grain elevator. In this case, 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: W_{t+1} is wealth at time $t+1$ and W_t is the known level of initial wealth; A is the number of acres planted; Y_{t+1} is the random yield per acre observed when the crop is harvested at $t+1$; P_{t+1} is the random spot price at $t+1$; $C(A)$ is the known cost function associated with planting the A acres; r is the riskfree interest rate; 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 terminal wealth used in Proposition 2.3.C:

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

where: π_{t+1} is the profit realized at time $t+1$, x is $(C(A)/W_t)$ the fraction of initial wealth invested in the crop production, 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 $[(A\ Y_{t+1}P_{t+1})/C(A)]$ one plus the rate of return on planting.

Given this, the farmer's optimal risk management decision problem is to choose x and H such that the expected utility of terminal wealth is maximized. As previously, the decision problem is modelled with a general expected utility function. However, in order to achieve analytically concise results, joint normality of R and R_f is invoked. This leads to the following:

Proposition 2.3C: **The Crop Investment and Hedging Decision**

Assuming that the returns R and R_f are jointly normal random variables, and that the farmer chooses x and H so as to maximize the expected utility of terminal wealth then:

$$\frac{H^*}{x} = \frac{\dfrac{E[R_f]}{\sigma^2_f} - \rho\dfrac{E[R]-r}{\sigma_R\sigma_f}}{\dfrac{E[R]-r}{\sigma^2_R} - \rho\dfrac{E[R_f]}{\sigma_R\sigma_f}}$$

where:

$$\rho = \frac{\sigma_{Rf}}{\sigma_R \sigma_f} \qquad \sigma_{Rf} = Cov(R,R_f) \qquad \sigma_f^2 = Var(R_f) \qquad \sigma_R^2 = Var(R)$$

Significantly, while the derivation of Proposition 2.3C reveals that the individual optimal solutions (denoted by *) to the farmer's risk management problem *(x*, H*)* depend on preferences, the ratio *(H*/x*) only involves ex ante statistical parameters*.

The portfolio-theoretic intuition behind the Proposition is as follows: the farmer faces two problems, one involving hedging, the other involving the scale of production. To determine the fraction of the crop to hedge, the farmer must solve a portfolio problem involving two risky "assets" with returns $(R - r)$ and R_f. From mean-variance portfolio theory, it is well known that if asset returns are jointly normal and riskless borrowing and lending is permitted then all investors, regardless of preferences, hold the same portfolio of risky assets. In addition, the ratio of any two assets in an optimal portfolio will be independent of risk preferences. Since the farmer's choice of the fraction of initial wealth to invest in crop production (x) is unconstrained, as long as returns are independent of the scale of production-- and the other assumptions relevant to Proposition 2.3.C are satisfied-- *(H*/x*)* will not involve preferences.

When used to analyze *(H*/x*)*, the practical implication of Proposition 2.3.C is that the fraction of the investment in crop production to be hedged $(Q_f f_t / C(A))$ is independent of the size of the crop. As a consequence, observed differences in the use of hedges and other contracting methods by large and small farmers revolve around the determination of R. Further, when the futures or forward price is unbiased $(E[R_f] = 0)$, only joint normality of R and R_f is required to motivate ordinary least squares as the optimal hedge ratio estimation technique. Though similar types of conditions have been derived for related problems, e.g., Benninga, et al. (1984), Heaney and Poitras (1991), this result has not been recognized as applying to the farmer's hedging problem with both price and production uncertainty. On balance, Proposition 2.3.C is of theoretical significance because it establishes a connection between the results of portfolio theory and the farmer's hedging problem further clarifying a possible motivation for the use of regression analysis to estimate the farmer's optimal hedge ratio.

Applications of Crop Insurance

Agriculture is of central importance to the economies of many countries. As a consequence, there are a myriad of government sponsored and private sector programs to support this sector. Many government programs, including disaster relief and crop insurance, were introduced to address the failure of private markets to “provide risk management tools for farmers to deal appropriately with production risk” (Agriculture Canada 1998, p.3). Crop insurance programs in the US and Canada originate in the agricultural reforms introduced following the Great Depression. In the US, the Federal Crop Insurance Corporation (FCIC) – currently responsible for administering a variety of federal crop insurance programs – was created in 1938 to administer a crop insurance program that only covered major crops in major producing areas. In Canada, the introduction of crop insurance begins with the passage of the federal Prairie Farm Assistance Act in 1939 aimed at providing

USDA Risk Management Agency, 'History of the Crop Insurance Program'

Crop insurance remained an experiment until passage of the Federal Crop Insurance Act of 1980. The 1980 Act expanded the crop insurance program to many more crops and regions of the country. It encouraged expansion to replace the free disaster coverage (compensation to farmers for prevented planting losses and yield losses) offered under Farm Bills created in the 1960s and 1970s, because the free coverage competed with the experimental crop insurance program. To encourage participation in the expanded crop insurance program, the 1980 Act authorized a subsidy equal to 30 percent of the crop insurance premium limited to the dollar amount at 65-percent coverage.

Although more farmers took part in the program after passage of the 1980 Act, it did not achieve the level of participation that Congress had hoped for. Therefore, after a major drought in 1988, ad hoc disaster assistance was authorized to provide relief to needy farmers. Another ad hoc disaster bill was passed in 1989. A third one enacted in 1992 gave farmers the option of claiming disaster losses on a farm-by-farm basis for any year between 1990 and 1992. An extremely wet and cool growing season in 1993 caused more losses, and Congress passed yet another ad hoc disaster bill. However, dissatisfaction with the annual ad hoc disaster bills that were competing with the crop insurance program led to enactment of the Federal Crop Insurance Reform Act of 1994.

The 1994 Act made participation in the crop insurance program mandatory for farmers to be eligible for deficiency payments under price support programs, certain loans, and other benefits. Because participation was mandatory, catastrophic (CAT) coverage was created. CAT coverage compensated farmers for losses exceeding 50 percent of an average yield paid at 60 percent of the price established for the crop for that year. The premium for CAT coverage was completely subsidized. Participants paid $50 per crop per county subject to maximum amounts for multiple crops and counties insured by the same individual. Subsidies for higher coverage levels were increased.

In 1996, Congress repealed the mandatory participation requirement. However, farmers who accepted other benefits were required to purchase crop insurance or otherwise waive their eligibility for any disaster benefits that might be made available for the crop year. These provisions are still in effect. In the same year, the Risk Management Agency (RMA) was created to administer FCIC programs and other non-insurance-related risk management and education programs that help support U.S. agriculture.

Participation in the crop insurance program increased significantly following enactment of the 1994 Act. For example, in 1998, more than 180 million acres of farmland were insured under the program. This is more than three times the acreage insured in 1988, and more than twice the acreage insured in 1993. According to estimates by the USDA National Agricultural Statistics Service, in 1998, about two-thirds of the country's total planted acreage of field crops (except for hay) was insured under the program. The liability (or value of the insurance in force) in 1998 was $28 billion, the largest amount since the inception of the program. The total premium, which includes subsidy, and the premium paid by insured persons (nearly $950 million) were also record figures.

In 2000, Congress enacted legislation that expanded the role of the private sector allowing entities to participate in conducting research and development of new insurance products and features. With the expansion of the contracting and partnering authority, RMA can enter into contracts or create partnerships for research and development of new and innovative insurance products. Private entities may also submit unsolicited proposals for

disaster assistance to western wheat farmers.[30] Initially these programs were small in scope in Canada and considered 'experimental' in the US.

From these early beginnings, crop insurance in the US and Canada has expanded to where crop insurance is a key federal agricultural support program. In Canada, the Crop Insurance Act (1959) was intended as a major step to stabilize farm incomes against production related risks. In contrast to other regions such as Europe and Australia where private markets write actuarially priced insurance, mostly for hail related damage, crop insurance in Canada and the US is subsidized and covers a wide range of crops and perils.[31] Recognizing that agriculture is a joint federal-provincial responsibility in Canada, the extent and evolution of the subsidy in Canada is given in Agriculture

Canada (1998, p.3):

> The CI Act of 1959 enabled the federal government to assist provinces in making CI available to producers at a 60% coverage level. Originally the federal government's share of total premiums was 20%, with a 50% share of administrative expenses. In 1964, the Act was amended to incorporate general provisions for a reinsurance agreement between the provinces and the federal government ... amendment to the Act, in 1973, provided two options for the federal-provincial-producer cost-sharing arrangements. In one option, the federal and provincial governments each contributed 25% of total premiums and 50% of administrative costs. In the other option, the federal government contributed a total of 50% of premiums and the provinces paid all administrative costs. In the 1990 amendment, the maximum coverage was increased to 90% for low risk crops. Furthermore, the single cost-sharing formula was adopted, where the federal government and provinces each pay 25% of total premiums and 50% of administration costs. Other changes included waterfowl crop damage compensation, and regulations concerning self-sustainability and actuarial soundness requirements.

Approximately a third of total federal agricultural support payments are contributed to the subsidy associated with crop insurance premiums offered through the AgriInsurance and delivered by the provinces. Each province tailors the insurance program to the crops grown in-province, e.g., Yeh and Yeh (1995). For example, in B.C. the program is delivered by the Ministry of Agriculture which refers to crop insurance as ‘production insurance’. Policies are offered that are tailored to specific production perils for: berries - blueberries, blueberry plants, cranberries, raspberries, strawberries and strawberry plants; flower bulbs - daffodil bulbs, tulip bulbs; forage - grass, legumes, greenfeed and silage corn; grain - canola, wheat, barley, oats, field peas, rye, forage seed (spot loss only); grapes - wine grapes, table grapes; tree fruits - apples, apricots, peaches, pears, plums, cherries; and, vegetables - carrots, onions, potatoes, cabbage, lettuce, broccoli, cauliflower, brussels sprouts, beans, peas, corn.

Though there is a rough similarity to the crop insurance programs featured in Canada and those provided in the US since 1996 by the Risk Management Agency (RMA), the US programs have considerably greater scope. Being a larger agricultural producer, the US has a greater number of crops requiring coverage: “RMA provides policies for more than 100 crops. Policies typically consist of general crop insurance provisions, specific crop provisions, policy endorsements and special provisions ... Policies are available for most commodities. In contrast to Canadian programs that focus on production insurance, the US features policies with a variety of payout features. In particular, a partial list of the policies the RMA offers involve the following: actual production history; actual revenue history; adjusted gross revenue; “dollar plan” crop yield; specified revenue protection; and, specified yield protection. For example, a crop insurance policy based on actual production history can be described as:

> Actual Production History policies insure producers against yield losses due to natural causes such as drought, excessive moisture, hail, wind, frost, insects, and disease. The producer selects the amount of average yield to insure; from 50-75 percent (in some areas to 85 percent). The producer also selects the percent of the predicted price to insure; between 55

> and 100 percent of the crop price established annually by RMA. If the harvested plus any appraised production is less than the yield insured, the producer is paid an indemnity based on the difference. Indemnities are calculated by multiplying this difference by the insured percentage of the price selected when crop insurance was purchased and by the insured share.

Specific revenue protection policies have a different procedure for determining the price used in calculating the indemnity. It is possible for farms to insure using individual or area wide variables to determine the indemnity.

Analysis of commodity risk management in agriculture is complicated by the array of government programs and subsidies aimed at mitigating significant production and revenue risks. These various programs have a complementarity that make it difficult to ascertain the commodity risk management activities of farmers. While the introduction of the Federal Crop Insurance Act of 1980 generated academic discussion surrounding the failure of private markets to provide a risk pooling solution, e.g., Myers and Oehmke (1988); (Knight and Coble (1997), the period leading up to passage of the Crop Insurance Reform Act (CIRA) (1994) was characterized by a reluctance by many farmers to participate in the crop insurance program. This reluctance coincided with Congress authorizing disaster relief payments in eight of the 10 years from 1985-1994. The CIRA required farmers to purchase a minimum amount of catastrophe insurance to quality for disaster relief programs. The CIRA was successful in doubling the number of acres insured under crop insurance plans from 1993 to 1998.

Changes to US crop insurance since CIRA have focussed on increasing the diversity of policy types and increasing private sector participation. These changes have coincided with the appearance of academic studies concerned with empirically identifying farmer attitudes to crop insurance, e.g., Mishra and El-Osta (2002); Sherrick et al. (2003); Ginder et al. (2009), and to identifying sound insurance policy designs, e.g., Miranda and Glauber (1997); Mahul (1999); Mahul and Wright (2003);Deng et al. (2007). Since 2000, when the government increased the subsidy from about 42 percent of the premium to about 60 percent, farmers not only increased purchases of crop insurance, but also tended to choose more expensive revenue insurance policies and at higher coverage levels. As a consequence, US crop insurers paid out a record $9.1 billion in indemnities in 2011, a total that could top $10 billion when all claims are settled. These claims originated primarily from damage due to drought, flooding, and freezing weather. The previous record was $8.7 billion in 2008. This has led to pressure to introduce cost reduction in the farm bill due in 2012.

Risk pooling mechanisms in agriculture can be traced back to the rural, agrarian economies of antiquity. At least since Evans-Pritchard (1940) it has been recognized that concerns surrounding economic and social security played an important role in the institutions and practices traditional rural societies (Platteau 1997, p.764). In turn, this has stimulated interest in ‘insurance’ arrangements in the village economies of the Third World, e.g., Townsend (1995); Jalan and Ravallion (1999); Little et al. (2001); Ligon et al. (20020. Because these primitive ‘mutual insurance’ schemes are second best compared to a market solution, considerable attention has been given to developing crop insurance programs that would be feasible for low-income, rural village economies, e.g., Sakurai and Reardon (1997), Skees and Enkh-Amgalan (2002); Skees et al. (2002); Skees et al. (1999). As with the US experience, the success of such programs depends on the level and type of subsidization and the presence of competing government programs also aimed at farm income support.

NOTES

1. This section follows Poitras (2008), Poitras (2005) and Poitras (2011, sec. 1.3.A).

2 A number of useful studies, both theoretical and empirical are available on multiple delivery specifications and the cheapest deliverable application, e.g., Chance and Hemler (1993), Kamara and Siegel (1987), Lien (1989), Cornell (1997), Adam-Muller and Wong (2003).

3. This terminology can create confusions. For example, the bulk of options traded in Europe are actually American options. While European options are not as commonly traded, this form is often used for the analytical simplifications provided. Another confusion is the use of "cash settlement" to refer to satisfaction of the option exercise requirements with a net dollar value transaction. In effect, the use of "cash" here does not refer to the spot commodity, but rather to actual cash.

4. The interest rate $r(i,j)$ is assumed to have been adjusted for the period over which interest is paid. This means that the interest rate has *not* been annualized.

5. Anand (2000) discusses difficulties with futures and forward markets achieving the full carry solution.

6. As before, $E[\ \cdot\ ]$ is the conditional mathematical expectation of $S(T)$ given the information available at time $t=0$. For notational simplicity the conditioning information is dropped because, in virtually ever case encountered in the analysis of derivative securities, expectations are conditional.

7. This is a stylized example. In practical applications, grain handling operations are often large, integrated firms that use forward contracting both with farmers supplying grain and consumers purchasing grain. The forward contracts involved are tailored to the specifics of the commerical operations involved. The example being given is more descriptive of the grain elevator operations in the latter half of the 19th century when the Chicago Board of Trade was available place futures contracts.

8. 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.

9 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).

10. 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.

11 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 $x < 0$.

12. 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.

13. 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.

14. 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 (1978, ch.13).

15. 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 which 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".

16. There is complementarity between the higher and lower moments that can be used under specific conditions. For example, Matz (1978) demonstrates that in the case of the quartic exponential density a change of variables can be used to eliminate the skewness term.

17. See Poitras (2002, Appendix 2) for derivation of σ_f^2, the variance of the futures price. In Poitras (2002, p. 113-6) derivation of the optimal solution is expressed using the partial derivative notation '∂' instead of the total derivative 'd' used in the derivation given here. The use of partial derivative notation is intended to recognize the reliance of terminal wealth on variables other than

the position in futures contracts, even though these variables are not directly specified. For example, either b and σ^2 could be variables instead of parameters. However, that approach is not used here.

18. This follows from the equivalence of the OLS estimator and the *ex post* (*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. Early 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). More recent applications, e.g., Godbey and Hilliard (2007), Bertus et al. (2009), have used alternative estimation methods to OLS to determine the minimum variance solution. Following Schwartz (1997), using stochastic differential equations to specify factors for the convenience yield, the spot price and, possibly, an interest rate process, the Kalman filtering technique can be used to determine the minimum variance solution. As illustrated in Lien (2005, 2012), which particularly technique is 'optimal' to determine the hedge ratio in a practical sense is still an unresolved issue.

19. Outside of agricultural applications, the insurance implications to risk management receive limited attention. Available studies typically focus on the relationship between insurance and firm financing decisions, e.g., Mayers and Smith (1982), MacMinn (1987) and Grace and Rebello (1993), Rebello (1995). Paulson et al. (2008) discuss the insurance approach to risk management in ethanol production.

20. This section follows Poitras (2002, sec. 6.H to 6.J).

21, Because h_t enters with a minus sign, this defines a short futures position to be a positive quantity.

22. The transformation from terminal wealth, W_{t+1}, to terminal profit, π_{t+1}, follows because the expectation of $U[W_t]$ reduces to a constant which does not affect the optimization. In addition, it will always be assumed that the only state (conditioning) variables of interest are R_s and R_f. However, in general, this need not be true.

23. When employed in common usage, the terms strategic risk management and enterprise risk management are often used interchangeably. However, enterprise risk management has become more commonly used since the actuarial societies in the US and the UK adopted 'enterprise risk management' as a recognized area of specialization. For example, the Society of Actuaries in the US now offers a Chartered Enterprise Risk Analyst designation. Enterprise risk management is also somewhat broader description involved in integrating risk into 'business strategy' and, perhaps more importantly, dealing with legal requirements for internal control mechanisms. In this vein, Section 404 of the Sarbanes-Oxley Act requires U.S. publicly-traded corporations to utilize a 'control framework' for internal control assessments. In addition, since 2007 the SEC has, in conjunction with the Committee of Sponsoring Organizations of the Treadway Commission (COSO), placed increasing scrutiny on top-down risk assessment, especially fraud risk assessment. In this context, strategic risk management focuses on the 'business strategy' component of risk assessment while enterprise risk management encompasses both strategic risk management and internal control

assessment, e.g., Moeller (2007). As issues of internal control receive little attention in the following discussion, reference is made to strategic, as opposed to enterprise, risk management.

24. 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.

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

26. 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.

27. Accounting exposure identifies specific accounting items which are subject to risk of exchange rate changes. Such exposures can usually be treated using hedging accounting.

28. An American example would be hotels and related businesses at Aspen.

29. For a further discussion of PPP, there are a number of useful sources. Early studies include L. Officer (1982), Roll (1979) and Shapiro (1983). Rogoff (1996) is an important review article where the purchasing power parity puzzle is identified. Sarno and Valente (2006) and Norman (2010) are recent studies dealing with the non-linear dynamics of PPP deviations.

30. The wheat farm risk management situation in Canada differs significantly from that in the US due to the presence of the Canadian Wheat Board (CWB) marketing monopoly that commenced in 1935 and has continued until the Conservative federal government passed legislation in Dec. 2011 to end the requirement that all non-feed wheat be sold through the CWB. In other regions, such as Australia and Europe, crop insurance is provided by the private sector.

31. Despite the government subsides, there is still private sector participation in crop insurance in the US where insurance companies market policies to farmers and then reinsure with the RMA. Private plans can also be structured to exceed maximum coverages associated with RMA policies. Circa 2011 there were approximately 15 private sector firms involved in providing crop insurance policies include John Deere Insurance Co. and Agrinational Insurance Co., a branch of Archer-Daniels-Midland. In 2009, the top five writers of multi-peril crop insurance in the United States were: Wells Fargo Insurance Group (16.5% market share); Ace INA Group (16.2%); NAU Country Insurance Co. (10.6%); Great American Property & Casualty Insurance Group (9.2%); and, Allianz of America (8.1%).