CODESRIA

Thesis by
WOLASSA
LAWISSO
KUMO

UNIVERSITY OF SOUTH
AFRICA

THE EFFECTS OF RISK AND UNCERTAINTY ON INVESTMENT DECISIONS OF FIRMS WITH PARTICULAR EMPHASIS ON INVESTOR-LENDER FIRM BEHAVIOUR

June
2007

THE EFFECTS OF RISK AND UNCERTAINTY ON INVESTMENT DECISIONS OF FIRMS WITH PARTICULAR EMPHASIS ON INVESTOR-LENDER FIRM BEHAVIOUR

by

WOLASSA LAWISSO KUMO

Submitted in accordance with the requirements

for the degree of

DOCTOR OF LITERATURE AND PHILOSOPHY

in the subject

ECONOMICS

at the

UNIVERSITY OF SOUTH AFRICA

PROMOTER: PROFESSOR J A DU PISANIE

Date submitted: June 2007

ACKNOWLEDGEMENTS

I would like to thank the following people and institutions for their support during the study. First of all, I would like to thank Professor Johann A. Du Pisanie for his constant encouragement, valuable suggestions and provision of direction for the study. I would also like to thank Professor Oludele A. Akinboade for the many very useful informal discussions. Many thanks to Thabisa Tokwe for proof reading the entire first draft of the thesis. My wife, Addis O. Anata, deserves appreciation for her support and motivation. I thank the UNISA Economics librarian, Etrecia de Villiers, for her assistance in fast tracking the reference materials.

I am grateful to the School of Economic Sciences, College of Economic and Management Sciences, UNISA, for selecting me for the position of post graduate student assistant with in the college's "Grow Your Own Timber" programme from which I benefited immensely.

Furthermore, I would like to acknowledge and express my sincere appreciation to the following institutions each of which provided partial financial support for the present study: International Institution of Educations' Scholar Rescue Fund (IIE/SRF) based in New York, USA, and Commission for the Development of Social Science Research in Africa (CODESRIA) based in Dakar, Senegal. Their support was invaluable.

To the memory of my father,
Lawisso Kumo Fara

SUMMARY

Traditional theories of investment analyse firms' investment decisions based on the assumption of certainty. However, in real life economic agents face both risk and uncertainty. The present study investigated various theories of economic decisions under certainty, risk and uncertainty. Economic decisions are affected by the agent's risk preferences. If an agent's risk preference exhibits constant absolute risk aversion, economic decisions are independent of the level of wealth whether the individual is risk averse, risk seeking or risk neutral. But, for decreasing or increasing absolute risk aversion, economic decisions depend on the level of wealth. They are also affected by aversion to uncertainty (ambiguity) measured by nonadditive probability in the context of expected utility maximization. The nonexpected utility theories, developed after the Allais paradox, question the validity of the main axioms of EU theory by claiming that people often violate these assumptions. Accordingly, they offer alternative explanations to the problems of choice under uncertainty. However, various empirical tests of nonexpected utility models have shown that none of these theories provide comprehensive alternative model of choice under uncertainty.

The present study analyzed investment and lending choice problems of an investor-lender (I-L) firm under risk and uncertainty in the context of the maximization of the expected utility of profit. The analysis incorporated both return and cost-of-funds (deposit cost) uncertainty. In the presence of deposit cost uncertainty, the optimal investment decision of a profit maximizing I-L firm occurs when the expected return from loans is equal to the expected marginal deposit cost. This optimal investment involves lending instead of investment in riskfree assets. This decision is both rational and in line with the actual behaviour of the I-L firms or commercial banks.

The firm level problem of financial investment decision is extended to the problem of aggregate fixed investment of firms under uncertainty. The estimation results of empirical error correction model of conventional macroeconomic investment determinants controlling for the effects of uncertainty indicated that macroeconomic uncertainty has significant negative effects on aggregate fixed investment. This result is based on the assumption of linear relationship between uncertainty and investment. The investigation of possible nonlinear relationship requires further research.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS...ii
SUMMARY...iv
LIST OF FIGURES AND TABLES...xi

CHAPTER ONE: CONCEPTS AND APPROACHES...1

1.1 CONCEPTS...1
1.1.1 Risk and uncertainty...1
1.1.2 Is uncertainty deficient foresight?...6
1.2 INVESTOR-LENDER FIRM AND AGGREGATE FIXED INVESTMENT OF FIRMS...7
1.3 THE PROBLEM...9
1.3.1 Asset allocation under risk...10
1.3.2 Ambiguity aversion and asset allocation...11
1.3.3 Observed investment behaviour of investor-lender firms...12
1.4 METHODOLOGY AND OUTLINE OF THE STUDY...13

CHAPTER TWO: INVESTMENT DECISIONS OF FIRMS UNDER CERTAINTY AND RISK...19
2.1 INTRODUCTION...19
2.2 THEORIES OF INVESTMENT DECISIONS UNDER CERTAINTY...21
2.2.1 The Fisher theory of investment...21
2.2.2 The accelerator theory of investment...22
2.2.3 The neoclassical theories of investment...25
2.2.3.1 The cost-of-capital theory...25
2.2.3.2 The q theory of investment...26
2.3 ALTERNATIVE MODELS OF CHOICE UNDER CERTAINTY...29
2.4 THEORIES OF ECONOMIC DECISIONS UNDER RISK...30
2.4.1 The Von Neumann-Morgenstern expected utility theory as the

basic theory of decision under risk..30
2.4.2 Measures of risk..34
2.4.2.1 The Arrow-Pratt measures of risk version..........................34
2.4.2.1.1 The risk premium and risk preferences...............35
2.4.2.1.2 Alternative Arrow-Pratt measures of risk version...38
2.4.2.1.2.1 Constant absolute risk version..............40
2.4.2.1.2.2 Decreasing (increasing) absolute risk aversion...40
2.4.2.1.2.3 Relative risk version..........................41
2.4.2.2 The Ross characterisation of risk aversion 42
2.4.2.3 Rothschild-Stiglitz measures of riskiness............................46
2.4.2.4 Measures of country risk in developing and developed countries...50
2.5 ALTERNATIVE MODELS OF RISK BEHAVIOUR...................................52
2.5.1 The Friedman-Savage measure of risk aversion..........................53
2.5.2 Other alternative models of risk aversion....................................56
2.6 INVESTMENT DECISIONS UNDER RISK..57
2.7 SUMMARY AND CONCLUSION...59

CHAPTER THREE: UNCERTAINTY IN ECONOMIC DECISIONS..........63

3.1 INTRODUCTION...63
3.2 KNIGHT AND KEYNES ON THE IMPACT OF UNCERTAINTY ON ECONOMIC DECISIONS 64
3.3 AUSTRIAN ECONOMICS AND UNCERTAINTY.....................................67
3.3.1 Mises, praxeology and uncertainty...67
3.3.2 Hayek, knowledge and uncertainty...68
3.3.3 Kirzner, entrepreneurship and uncertainty..................................70
3.3.4 Austrian economics, time and ignorance....................................72
3.4 THE NEW CLASSICAL ECONOMICS AND RATIONAL EXPECTATIONS...74
3.5 THE ARROW-DEBREU THEORY OF COMPETITIVE EQUILIBRIUM UNDER UNCERTAINTY 77

3.6 THE COMPETENCE-DIFFICULTY (C-D) GAP APPROACH..................80
3.7 UNCERTAINTY, THE MARKET SYSTEM AND GOVERNMENT INTERVENTION 82
3.8 SUMMARY AND CONCLUSION..................84

CHAPTER FOUR: AMBIGUITY AVERSION AND OPTIMAL INVESTMENT DECISIONS..................87

4.1 INTRODUCTION..................87
4.2 AMBIGUITY AVERSION, INDEPENDENCE OF PREFERENCES AND THE ELLSBERG PARADOX..................87
4.3 THE CHOQUET EXPECTED UTILITY AND SCHMEIDLER'S THEORY OF UNCERTAINTY AVERSION..................90
4.4 THE EPSTEIN THEORY OF UNCERTAINTY AVERSION..................93
4.5 GUL'S THEORY OF DISAPPOINTMENT AVERSION..................95
4.6 AVERSION TO SEQUENTIAL RESOLUTION OF UNCERTAINTY..................96
4.7 DECISION WEIGHTS ON RISK AND UNCERTAINTY..................98
4.8 AMBIGUITY AVERSION AND OPTIMAL INVESTMENT DECISIONS..................100
4.9 SUMMARY AND CONCLUSION..................102

CHAPTER FIVE: UNCERTAINTY, IRREVERSIBLE INVESTMENT AND FINANCE..................104

5.1 INTRODUCTION..................104
5.2 UNCERTAINTY AND THE OPTION VALUE OF WAITING TO INVEST..................104
5.3 UNCERTAINTY AND INVESTMENT IRREVERSIBILITY..................107
5.4 FINANCE, UNCERTAINTY AND IRREVERSIBLE INVESTMENT..................114
5.5 SUMMARY AND CONCLUSION..................118

CHAPTER SIX: EXPECTED AND NONEXPECTED UTILITY THEORIES OF CHOICE UNDER UNCERTAINTY........120

6.1 INTRODUCTION..120
6.2 RATIONALITY IN ECONOMICS..122
6.2.1Instrumental rationality..123
6.2.2 Procedural rationality...124
6.2.3 Expressive (bounded) rationality..125
6.3 RATIONAL DECISIONS AND EXPECTED UTILITY MODELS.............128
6.3.1 Bernoullian expected utility theory..128
6.3.2 The von Neumann-Morgenstern expected utility theory as the theory of rational choice under uncertainty 130
6.3.3 The Savage expected utility theory..134
6.4 RATIONAL DECISIONS AND NONEXPECTED UTILITY MODELS......136
6.4.1Generalizations of expected utility theory.................................137
6.4.1.1 Machina's generalized expected utility analysis...............137
6.4.1.1.1 Empirical tests of the fanning-out hypothesis......140
6.4.1.2 Rank-dependent expected utility model 140
6.4.1.2.1 Empirical tests of RDEU model.....................144
6.4.2 Allais's nonlinear intensity theory...147
6.4.2.1 Empirical tests of the Allais Paradox..........................151
6.4.3 Prospect theories as alternative theories of choice under risk and uncertainty...152
6.4.3.1 Prospect theory as analysis of decision under risk...............152
6.4.3.1.1 Loss aversion..154
6.4.3.1.2 The reflection effect....................................157
6.4.3.1.3 Nonlinear weighting of probability.....................158
6.4.3.1.4 Empirical tests of prospect theory....................159
6.4.3.2 Cumulative prospect theory as analysis of decision

under uncertainty..160
6.4.3.2.1 Empirical tests of cumulative prospect theory......164
6.4.4 Nontransitive preference theories: regret theory.........................167
6.4.4.1 Empirical tests of regret theory...................................169
6.5 PREFERENCE REVERSAL..170
6.6 INEFFICIENT CHOICES AND INVESTOR IRRATIONALITY IN DECISION-MAKING...172
6.7 SUMMARY AND CONCLUSION...181

CHAPTER SEVEN: INVESTMENT DECISIONS OF BANKING FIRMS UNDER UNCERTAINTY..............................185

7.1 INTRODUCTION...185
7.2 THEORIES OF THE FIRM...186
7.3 BANKING FIRM MODELS..190
7.4 CHARACTERISTICS OF INVESTMENT AND LOAN MARKETS...........190
7.4.1 Investment markets 190
7.4.2 Loan markets...191
7.4.2.1 Determinants of lending...191
7.4.2.2 Types of loans...193
7.4.2.3 Pricing of business loans...194
7.5 BANK PORTFOLIO CHOICE UNDER CERTAINTY...............................196
7.6 THE RISK-RETURN PORTFOLIO CHOICE MODELS.............................197
7.7 BANK PORTFOLIO CHOICE MODELS UNDER UNCERTAINTY..........199
7.7.1 Porter bank portfolio model..200
7.7.2 Pyle's bank portfolio model..201
7.8 THE CHAVAS MODEL OF PORTFOLIO SELECTION UNDER UNCERTAINTY...204
7.9 YAARI'S COMPARATIVE STATICS MODEL OF PORTFOLIO CHOICE.206
7.10 ECONOMETRIC MODELS OF BANKING FIRM INVESTMENT DECISIONS UNDER UNCERTAINTY...209
7.11 BANK LENDING AND THE MANAGEMENT OF RISK.........................211
7.12 LENDING, MONITORING AND COSTLY STATE VERIFICATION-

THE GALE-HELLWIG MODEL....212
7.13 SUMMARY AND CONCLUSION....215

CHAPTER EIGHT: INVESTOR-LENDER FIRM MODEL....218

8.1 INTRODUCTION....218
8.2 THE MODEL....220
8.3 THE RESULT OF THE I-L MODEL....228
8.4 THE NONEXEPECTED UTILITY INTERPRETATION OF THE RESULT OF THE I-L MODEL....231
8.5 CONCLUSION....234

CHAPTER NINE: THE EFFECTS OF UNCERTAINTY ON AGGREGATE FIXED INVESTMENT OF FIRMS: AN EMPIRICAL ECONOMETRIC ANALYSIS....237
9.1 INTRODUCTION....237
9.2 MACRO LEVEL UNCERTAINTY AND AGGREGATE FIXED INVESTMENT OF FIRMS....240
9.3 SOURCES AND MEASURES OF MACROECONOMIC UNCERTAINTY: THE GARCH ESTIMATION OF VOLATILITY MODELS....241
9.4 ECONOMETRIC METHODOLOGY AND THE DATA....245
9.4.1 The model....245
9.4.2 Testing for the stationarity of the time series....247
9.4.3 Testing for cointegration....249
9.4.4 The empirical error correction equation....252
9.4.5 The data....254
9.5 THE ESTIMATION RESULTS....255
9.6 CONCLUSION....260

CHAPTER TEN: FINAL CONCLUSION....263

REFERENCES....266

LIST OF FIGURES

Figure 2.1 The Diamond-Stiglitz single crossing density functions 48

Figure 2.2 The Diamond-Stiglitz multiple crossing distribution functions..............49

Figure 2.3 The Friedman-Savage concave-convex utility functions....................54

Figure 2.4 The Friedman-Savage utility function with double inflexion................55

Figure 6.1 Indifference curves assumed by expected utility theory in a unit probability triangle..133

Figure 6.2 Indifference curves assumed by Machina's *fanning-out* hypothesis in a unit probability triangle..139

Figure 6.3 Indifference curves suggested by the prospect theory in a unit probability triangle..159

LIST OF TABLES

Table 6.1 The Gordon et al. outcomes for maximum investment on five gambles..174

Table 9.1 The estimation results of GARCH (1,1) volatility models for South Africa 1975(4)-2006(1)..243

Table 9.2 Augmented Dickey-Fuller unit root tests for the conventional macroeconomic variables in levels including constant and trend.........248

Table 9.3 Augmented Dickey-Fuller unit root tests for the conventional macroeconomic variables in first differences including constant but no trend...248

Table 9.4 The estimation results of E-G two step test for Cointegration, 1975(4)-2006(1)..251

Table 9.5 The estimated result of the ECM model for capital stock and macroeconomic determinants, 1978(3)-2006(1)...........................253

Table 9.6 Private fixed capital stock ECM equation for South Africa, 1979(2)- 2006(1)..257

Table 9.7 Correlation matrix for aggregate private fixed capital stock and macroeconomic uncertainty, i.e. GARCH (1,1) conditional variances....259

Table 9.8 Correlation matrix among measures of macroeconomic uncertainty, i.e. GARCH (1,1) conditional variances.......................................260

CHAPTER ONE

CONCEPTS AND APPROACHES

1.1 CONCEPTS

1.1.1 Risk and uncertainty

The concepts of risk and uncertainty have generated debate among economists for decades. However, Frank Knight (1921) was the first of these to indicate the importance of the difference between the two concepts in economic analysis. Knight (1921:214-216) defined risk as a situation where the outcome of events is known through either *a priori* probability or statistical probability, whereas uncertainty occurs when no probability distribution can be assigned to the outcome of events. Thus, for Knight, risk is a measurable uncertainty while uncertainty proper is an unmeasurable uncertainty.

Knight emphasized the difference between *a priori* probability and statistical probability where the former refers to a mathematical probability usually obtained in games of chance, and the latter to empirical generalization of risky or hazardous events. He stated, " …there is a fundamental difference between '*a priori*' probability, on the one hand, and 'statistical' on the other. In the former, the chances can be computed on general principles, while in the latter they can only be determined empirically" (Knight 1921:224). According to Runde (1995:198), Knight splits the situations of risk into two. These comprise those in which probabilities are determined purely on general principles (*a priori* probabilities) and those in which alternatives are not absolutely homogenous but are grouped together on an empirical basis (statistical probabilities). This implies that Knight's risk reflects reduction in uncertainty while true uncertainty can only show a tendency toward regularity if it can be grouped on the basis of nearly any similarity or common element. And hence, for Knight, true uncertainty is unmeasurable.

Although Knight made a seminal contribution to our understanding of the differences between risk and uncertainty and their importance in influencing the behaviour of economic agents, he did not provide any suggestions on how the impact of uncertainty on the decisions of economic agents can be measured.

Keynes (1921, 1936, 1937) provides similar definitions of the concepts of risk and uncertainty. He maintains that agents' knowledge cannot be characterized by probabilities. In other words, the knowledge agents have about the future is simply uncertain and this uncertainty cannot be measured by assigning probabilities. Driver and Moreton (1992:67) argue that, according to Keynes, even though probabilities can be numerically defined, probabilistic judgment can provide information only for one side of the two sided phenomenon. Moreover, the comparison between different probability distributions may be impossible because of different weights of confidence attached to these distributions.

Emphasizing the difference between probable events and uncertainty in the *General Theory*, Keynes (1936:148) stresses that by "very uncertain" he does not mean the same thing as "very improbable". In the *Quarterly Journal of Economics* he writes:

> By 'uncertain' knowledge, let me explain, I do not mean merely to distinguish what is known for certain from what is only probable. The game of roulette is not subject, in this sense, to uncertainty ... Even the weather is only relatively uncertain. The sense in which I am using the term is that in which the prospect of a European war is uncertain, or the price of copper and the rate of interest twenty years hence or the obsolescence of a new invention ... About these matters there is no scientific basis on which to form any calculable probability whatever. We simply do not know.
>
> (Keynes 1937:213-214)

Thus, for Keynes the future is unknown and hence unmeasurable. This concept is related to Knight's true uncertainty.

For Keynes, probabilistic risk, on the other hand, is a concept knowable on the basis of past and present market signals as opposed to uncertainty, where existing information is not a reliable guide to future performance. Probabilistic risk may characterize routine repeatable economic decisions where it is reasonable to presume an unchanging reality, whereas for an important economic decision involving investment, the accumulation of wealth etc., as Keynes declares, agents are dealing with an uncertain, nonprobabilistic creative economic reality in which today's human action can create a new and different future reality (Davidson 1995:111).

Furthermore, Shackle (1961:60) argues that when probability is objective, it consists of relative frequencies and while these represent knowledge, knowledge and uncertainty are mutually exclusive. Thus, for Shackle too, the concept of probabilistic risk is very different from the concept of uncertainty.

Although Keynes made a great contribution in emphasising uncertainty as a dominant feature of real life that confronts economic agents in their daily decision-making processes, he did not distinguish between different objects of uncertainty and different objects of knowledge. For instance, in the previous quotation Keynes considers almost everything that may or may not happen in the future as uncertain. He emphasizes the possibility of European war being uncertain but, in fact, it occurred two years after his publication in the *Quarterly Journal of Economics*. According to Lawson (1985:916), in the context of economic analysis, Keynes's use of the term uncertainty was exclusively reserved for the evaluation of future outcomes of all *currently* possible decisions or acts.

Although both Keynes and Knight agree on the sharp distinction between risk and uncertainty and the impossibility of measuring uncertainty using statistical calculations, there are some fundamental differences in their definitions of the concepts. Davidson (1996:23) states that Keynes's distinction between risk and uncertainty appears on the surface to be similar to Knight's but is in fact significantly different. While Knight associates risk with either statistical or Bayesian probabilities, uncertainty is associated with only unique events. Moreover, for Knight, uncertain future is the basis for the existence of business profit, while for Keynes, uncertainty involves the absence of any scientific basis on which to form probabilities, while the concept is particularly important when the topic of analysis is the accumulation of wealth.

Perlman and McCain (1996:17), in addition to distinguishing between risk and uncertainty, make a more detailed distinction between risk based on *a priori* probability and risk based on statistical probability, and factual uncertainty and uncertainty that arises from a lack of understanding. For them, probability is either an object of knowledge (it has an objective existence) or an epistemological concept, i.e. it is a form of knowledge. Thus, probability can be either an *aleatory* or an *epistemic* concept. The term *aleatory* derives from a seventeenth century game of dice; Cournot, Siméon Denis Poisson, Pierre Simon De Laplace and others of the French classical probability tradition referred to the concept as *la chance* or *la facilité*, whereas the term epistemic is employed to refer to consciousness or the product of the mind (Perlman and McCain 1996:19). Accordingly, Perlman and McCain define aleatory and epistemic risk and aleatory and epistemic uncertainty. Aleatory risk implies the existence of a known frequency defined by reference to a given empirical series as in rational expectation models, while epistemic risk implies that such a series is not necessarily objectively known but is nonetheless determinable, as in the Savage model. Aleatory uncertainty refers to factual uncertainty that refers to the uncertainty of Knight, while epistemic uncertainty is the type of uncertainty that exists owing to man's limited ability to understand beyond a limited sphere, and corresponds to Keynes's definition (Perlman and McCain 1996:17).

Many economists, however, dispute the distinction between risk and uncertainty made by Knight, arguing that there is no difference between Knight's uncertainty and his risk. Some argue that probabilities are only subjectively assigned expressions of belief and may not necessarily be connected to the true randomness of the real world. The expected utility theory, including Bayesian decision theory, and rational expectation theory do not support such sharp distinctions between risk and uncertainty. For instance, Runde (1995:197) states, "the widespread acceptance of the Bayesian decision theory has eroded Knight's (1921) famous distinction between risk and uncertainty". In Bayesian decision theory, a rational agent or a subjective utility maximizer chooses between two risky events on the basis of their subjective values and beliefs. The values are represented by subjective utility functions where a utility index can be assigned to each possible outcome and beliefs are represented by a probability function that assigns a probability index to each of the possible states of the world on which possible outcomes depend (Runde 1995:197).

Furthermore, Quiggin (1993:4) argues that it is not possible to strictly confine risk to situations where probabilities are known and uncertainty to situations where probabilities are unknown because decisions under risk may be made with or without objective information on probabilities. He uses the term ambiguity for situations where probabilities are unknown and uncertainty as a general term to include risk, instability and ambiguity. The present study follows an approach similar to Quiggin's (1993) in defining risk and uncertainty. The term ambiguity was first defined by Ellsberg (1961:656-657) who questioned why some uncertainties are not risks and provided an answer by defining ambiguity as a situation where an individual's confidence in assigning probability distribution to the occurrence of events is very low, reflecting neither complete ignorance nor risk. Accordingly, in the present study, the term ambiguity is used to represent a situation where probabilities are unknown while uncertainty is used as a general term to include both risk and ambiguity. The term ambiguity represents both Knight's aleatory and Keynes's epistemic uncertainties while risk, based on either *a priori* probabilities or statistical probabilities, is treated as one concept.

However, post Keynesians such as Davidson (1991) still consider Knight's distinction appropriate in economic analysis. They believe that the Knightian "uncertainty" may be the only relevant form of uncertainty for economics in the sense that the situations of Knightian "risk" arise only in very controlled situations when the alternatives are clear and experiments can be repeated, such as in the system of gambling. This concept of risk does not apply when a situation is unique and the alternatives are not all known or clearly understood. Davidson (1991:129-130) argues that although neoclassical economics equates probabilistic risk with uncertainty, probability distributions are not the basis for comprehending real world behaviour under uncertainty. He further states that "...there are many important situations where 'true' uncertainty exists regarding future consequences of today's choices. In these cases, today's decision makers believe that no expenditure of current resources on analyzing past data or current market signals can provide reliable statistical or intuitive clues regarding future prospects" (Davidson 1991:130). Other writers link the concept of uncertainty to the problem of deficient foresight.

1.1.2 Is uncertainty deficient foresight?

An important conceptual issue regarding uncertainty is whether it is possible to define it as a lack of correct foresight. This is based on the presumption that certainty reflects correct foresight and therefore, by implication, uncertainty reflects deficient foresight. Coddington (1982:483) questions whether certainty can be defined as a state of complete confidence in belief or as correct foresight and argues further that "perfect confidence in a belief is perhaps far better sustained by ignorance than by understanding. And one would not feel at all happy expounding a theory in which everyone could repeatedly be perfectly confident one moment and discover themselves to have been wrong the next. So it is by the second idea of certainty as correct foresight that we are driven."

Coddington does, however, admit that there are analytical difficulties in the concept of correct foresight. First, if certainty is identified with some unattainable ideal and if the foresight is considered to be correct when it occurs and resembles exactly what was foreseen, then its lack of attainment becomes a trivial matter. This assumes comprehensive exactness of foresight and makes all foresight "incorrect" if it does not meet this criterion. Thus, it does not matter whether a particular belief or expectation is uncertain when compared with some, unattainable, ideal state of omniscience. Second, at the other extreme, all foresights could be claimed to be correct by relaxing standards of approximation. In between these extremes there is a notion that correctness must consist in foresight's being within certain reasonable bounds of approximation and with the scatter of actual cases within these bounds not exhibiting any systematic error (Coddington 1982:482-484). Thus, assessing beliefs and expectations in terms of the best use of available information rather than in terms of omniscience provides one with framework thought that is consistent with being fallible (man makes mistakes) but avoids a systematic pattern of errors once they have become apparent (Coddington 1982:484).

Not all economists agree with this definition of uncertainty as deficient foresight. Some, especially post Keynesians, believe that agents can be uncertain about the future even though their forecasts are correct for a finite number of periods (Driver and Moreton 1992:78). According to these individuals, uncertainty cannot be equated with a mere deficiency in foresight. It is a more pervasive phenomenon in every real world life. Coddington's argument falls between the post Keynesians' views of nondeterministic future and the deterministic views of neoclassical economics. The present study will not follow the approach that defines uncertainty as deficient foresight because not only is the hypothesis of certainty as perfect foresight questionable, but neither can this notion be easily extended to uncertainty when the foresight is believed to be deficient.

1.2 INVESTOR-LENDER FIRM AND AGGREGATE FIXED INVESTMENT OF FIRMS

The investor-lender (I-L) firm is a term developed by the present study to define a financial firm involved in mobilisation of funds in the form of customer deposits and the use of these funds, either for investment in riskfree securities, particularly homogenous government bonds, and/or for lending to other firms or entrepreneurs which make direct investment in productive activities. In the real world, the concept of an I-L firm corresponds to the activities of commercial banks. Thus, the I-L firm can be considered as a commercial banking firm with only two investment alternatives, i.e. investment in government bonds or investment in loans, and with only one source of capital, i.e. customer deposits. There are different types of loans, but for the purpose of the present study the types of loans do not matter as long as they are used for business purposes. The present study is not concerned with nonbusiness loans.

At this juncture it becomes important to define the concept of investment. Investment can take two broad forms. The first is investment in fixed capital or fixed capital formation while the second refers to investment in financial assets. Investment in fixed capital stock in turn takes different forms. These involve investment in human resources in the form of training, investment in research and development activities, investment in finished goods and intermediate inputs, and investment in fixed capital stock in the form of machinery, equipment and nonresidential buildings. The present study is concerned with the aggregate investment in fixed capital stock by private business firms, i.e. the aggregate investment in machinery, equipment and non-residential buildings by private business firms.

Investment in financial assets can be of two types: direct financial investments and portfolio investments. Direct financial investment refers to investments in a company in order to gain control or ownership, while portfolio investment refers to financial investment for the purpose of interest or dividends (Online Dictionary of the Social Sciences 2002). That is, portfolio investment is financial investment in securities or shares or other interest traded in financial markets with the objective of obtaining interest or dividends.

Financial firms have the option of investing in financial securities or lending their funds to other firms. The two main concerns of the present study are the analysis of portfolio investment behaviour by investor-lender firms and the analysis of the effects of uncertainty on aggregate fixed investment. The I-L firm can invest either in riskfree financial assets, particularly government bonds, and obtain riskfree returns from the investment or it can lend its capital to other firms and face risk and uncertainty about its future returns. The first hypothesis of the thesis is that investor-lender firms tend to put more of their wealth into risky financial assets, i.e. lending their capital and obtaining higher but riskier returns rather than investing in riskfree securities which yield lower returns. The second hypothesis of the thesis is that the effects of uncertainty do not disappear at the aggregate level, i.e. aggregate uncertainty reduces aggregate private firm fixed investment.

1.3 THE PROBLEM

For over two centuries economists have assumed that economic agents make economic decisions with full knowledge of the future, i.e. with certainty. Certainty was an important characteristic of classical and neoclassical economics until and even after the concepts of risk, uncertainty and probability were incorporated into economic analysis in the 1920s and 1930s (Amariglio & Ruccio 1995:334-335). Accordingly, most theories of investment decisions of firms are based on this assumption.

This is true for both fixed investment decisions as well as for the investment decisions in financial assets. Financial firms face two broad choices when making investment decisions. The first is the possibility of investing in riskfree securities and earning riskfree returns from these investments while the second involves investment in risky financial assets and facing uncertainty about the future return from these investments. Given that the objective of these firms is the maximization of profit, how does the presence of risk and uncertainty affect their optimal decision? Is it possible to make rational choices under uncertainty in the same manner as rational choices under certainty? According to economic analysis of choices under risk and uncertainty this presents a problem. The optimal values of decisions under certainty will no longer be optimal under uncertainty, i.e. the maximization of profit under certainty is completely different from the maximization of profit under uncertainty. In terms of the utility analysis of choice, while agents maximize utility under certainty, they maximize the expected utility under uncertainty.

1.3.1 Asset allocation under risk

Based on the expected utility, the dominant theory of choice under risk and uncertainty, various attitudes towards risk have been developed in economic analysis. These attitudes involve risk aversion, risk seeking and risk neutrality, where the most prevalent behaviour is considered to be risk aversion.

The decisions to allocate one's wealth between risky and riskfree alternatives depend both on these attitudes towards risk and on whether the objective of the individual is the maximization of the expected utility of profit or not. For instance, when two individuals have certain levels of absolute wealth and both individuals are risk averse, and if their initial wealth is divided between a riskless asset and a risky asset whose expected return per dollar invested exceeds that of the riskless asset, then for all such wealth level and asset returns, the first expected utility maximizer would put at least as much wealth into the riskless asset as would the second expected utility maximizer (Fishburn 1988:19).

This comparative analysis provides a useful tool for comparing the behaviour of different risk averse individuals whose objectives are the maximization of expected utility. However, it does not explain why a single individual or an investor-lender firm may decide to put more of its wealth in risky assets with higher returns than in riskfree assets with lower returns. This choice problem seems to contradict the basic economic assumption that individuals prefer certainty to risk. This behaviour might be explained by alternative noncomparative optimization models which utilize the fundamental assumption of the maximization of the expected utility of profit under both price and cost-of-funds uncertainty. Chapter eight of the present study will investigate this issue in greater detail.

As an extension to the above firm level investment problem the present study investigates the effects of uncertainty on the aggregate fixed investment of private business firms using the measures of uncertainty generated by the generalized autoregressive conditional hetroscedasticity (GARCH) model. These measures of uncertainty are used to estimate the empirical investment model developed in chapter nine.

1.3.2 Ambiguity aversion and asset allocation

As in the case of risk aversion, attitudes towards ambiguity or uncertainty where probabilities cannot be assigned to occurrence of events are measured by ambiguity aversion. According to Ellsberg (1961:660-661), ambiguity is a subjective variable but it is possible to identify some situations likely to present high ambiguity by noting situations where available information is scanty, obviously unreliable or highly conflicting. Agents' attitudes towards ambiguity have been studied by other authors since Ellsberg. In particular, Epstein (1999:584) defines uncertainty aversion as preference for unambiguous events rather than ambiguous events or as the behaviour of ambiguity averse individuals preferring unambiguous events to ambiguous events.

For Schmeidler (1989:582), uncertainty or ambiguity aversion is related to the expected utility with respect to nonadditive probability. Does aversion to ambiguity have a similar impact on asset allocation of an investor-lender firm as aversion to risk? Dow and Werlang (1992) have investigated the impact of ambiguity aversion on the problems of portfolio choice using the analytical framework of expected utility. They argue that under the theory of expected utility, an agent who must allocate his wealth between a riskfree and a risky asset will invest in some of these assets if the price of the given asset is less than the expected value (Dow and Werlang 1992:197). When the expected utility maximization is analyzed in the context of nonadditive probability, i.e. where probabilities of occurrence of events do not add up to one, this situation reflects aversion to ambiguity. Under these circumstances, the agent will calculate the price of the asset as the expected value of the asset using the nonadditive probability measure and this expected price under nonadditive probability will be different from the conventional expected price. This is another problem raised in the study. Stated differently, does the agent's optimal investment decision reflect aversion to ambiguity or divergence from the conventional problem of the maximization of the expected utility?

1.3.3 Observed investment behaviour of investor-lender firms

The trouble with investor-lender firm behaviour is that these firms earn most of their income by investing in loans instead of in riskfree securities. But, at the same time, lending is the main cause of the collapse of such firms. Gup and Kolari (2005:271) argue that the problem with banking firm behaviour is that banks earn most of their income by lending instead of investing in riskfree securities while at the same time lending accounts for most of their risk and lending risk is the primary cause of bank failure.

Moreover, Cebenoyan and Strahan (2001:3) observe that banks with better risk management options, i.e. banks involved both in loan buying and selling activities, hold more risky loans as a percentage of their balance sheets than banks that are not involved in loan buying and selling or which are involved in only loan buying or selling activities but not both.

Why, then, do investor-lender firms fail to avoid this risk by increasing the size of their low or no risk but low earning equity investment? Stated differently, the question becomes: why do investor-lender firms lend their funds to entrepreneurs instead of investing them in riskfree financial assets and obtaining riskfree returns from the investment? This seems to negate the basic economic assumption that agents prefer certainty to risk.

This issue is compounded by state verification of the borrower. The problem of state verification (observation of the state of the borrower) under lending and investment decisions arises from the existence of asymmetric information. Gale and Hellwig (1985:648) have developed an optimal credit contract model and argue that the act of the observation of state can be considered as "bankruptcy", which means that the state is observed if and only if the firm cannot repay the loan in full or when the firm is insolvent. This argument implies that the optimal credit contract between lenders and borrowers is a standard debt contract under bankruptcy.

Thus, the observed behaviour of investor-lender or commercial banking firms seems to be at odds with the economic assumption that economic agents prefer certainty to risk and that they are averse to both risk and ambiguity. Preference for risky investment alternatives in the presence of a riskfree alternative is one of the central problems explored by the present study. The study will show that this behaviour is not a contradiction of the rational choice behaviour of economic agents under uncertainty.

Another central problem explored in the present study is the investigation of the effects of risk and uncertainty on the aggregate fixed investment of private business firms. This problem is explored by using time varying volatility measures of macroeconomic uncertainty which are used as some of the explanatory variables in the empirical investment equations. The aggregate fixed investment considered by the present study refers to the private business investment in fixed capital stock in South Africa.

1.4 METHODOLOGY AND OUTLINE OF THE STUDY

The objectives of the present study are twofold: (a) to investigate the investment decisions of the investor-lender firms by developing an investor-lender firm model under risk and uncertainty including an extension of the firm level analysis into empirical econometric analysis of the effects of uncertainty on aggregate fixed investment of private business firms; and (b) to analyze the global theoretical literature on the importance of risk and uncertainty on firm level economic decisions and the effect on the economy at large.

The methodology followed in the study involves the use of various techniques such as graphical analysis, mathematical and optimization models and empirical econometric and volatility (GARCH) models. The study investigates the behaviour of investor-lender (I-L) firms by developing an I-L firm optimization model controlling for the effects of cost-of-funds uncertainty. In real life, economic agents face many uncertainties. Typical sources of uncertainty faced by the investor-lender firm without production function constraint include price or return uncertainty and cost-of-funds uncertainty. The I-L firm model developed in the present study incorporates these sources of uncertainties in the analysis of the investment and lending decisions of these firms (see ch 8). Chapter 3 explores the impact of uncertainty on economic decisions.

Economic theories assume that economic agents prefer certainty to risk and risk to ambiguity. Accordingly, most neoclassical theories of investment base their analysis of investment decisions by firms on the assumption of perfect certainty. These theories include, among others, the cost-of-capital theory of Jorgensen (1963) and the q theory of Tobin (1969). However, alternative models of choice under certainty using the analytical tools of benefit and distance functions question the assumption of perfect certainty. According to these alternative models, the benefit and distance functions are used both to describe Pareto-efficient outcomes under certainty, and to derive equivalent measures of risk where compensating benefits are considered to represent the traditional measure of risk aversion (Quiggin and Chambers 1998:133). The theories of decisions under risk develop both premium as well as comparative absolute and relative measures of risk aversion. These measures are also known as the Arrow-Pratt measures of risk aversion and are based on the concavity of the utility function in Von Neumann-Morgenstern's expected utility theory. As discussed above, comparative absolute and relative measures of risk aversion are used in the analysis of the impact of attitudes towards risk on the allocation of wealth between risky and riskfree alternatives by two economic units. In special cases where risk preference exhibits constant absolute risk aversion (CARA), increase in initial wealth does not have any impact on the individual's attitudes towards risk. When the risk attitudes reflect increasing absolute risk aversion (IARA) or decreasing absolute risk aversion (DARA), however, increase in initial wealth increases or decreases an individual's willingness to purchase insurance against risk respectively. Relative risk measures have similar interpretations. Alternative measures of risk attitude, particularly the Friedman-Savage (1948) hypothesis, question smooth equating of risk aversion with the concavity of the utility function (see ch 2).

While attitudes towards risk are studied extensively, attitudes towards ambiguity are not as thoroughly explored. Recent theoretical developments, based on strict distinctions between risk and ambiguity, hypothesize that uncertainty (ambiguity) aversion can be measured in two ways. The first refers to nonadditivity of probabilities in the analysis of the expected utility maximization while the second refers to preferences for unambiguous events over ambiguous events. Chapter 4 provides further explanations of how these attitudes towards ambiguity affect optimal investment decisions of a financial firm. Explanations are also given of Palacios-Huerta's (1999) hypothesis where preferences reflect aversion to sequential resolution of ambiguity and preference for resolving ambiguity all at once.

Investment decisions of firms are not only constrained by the presence of risk and uncertainty. Some investments are irreversible in the sense that they cannot be partially or fully recovered if they are found to be unprofitable. Dixit and Pindyck (1994:3) argue that in the presence of uncertainty and investment irreversibility, the investor has an option to wait for more information about the future and as the value of this option increases, investment decreases. This problem can be worsened if a firm faces limited internal funds or is unable to borrow from financial institutions. Chapter 5 explores how the interaction between uncertainty, investment irreversibility and finance affects the investment decisions of firms. The empirical econometric analysis of the effects of uncertainty on aggregate fixed or aggregate irreversible investment of private business firms is carried out in chapter 9. Investment irreversibility is important not only for fixed capital formation but also for portfolio investment involving lending because loans involve irreversible investment depending on the degree of default.

The dominant theory of choice under uncertainty is the expected utility theory, particularly that of Von Neumann and Morgenstern (1944, 1947). This theory presents choice under uncertainty using various axioms. Among the five axioms of this theory are three main ones: transitivity, independence (linearity in probability) and continuity or invariance. Transitivity refers to preferring more to less, independence means that any linear combination of probability does not affect preference patterns while continuity means that a sufficiently small change in probability does not reverse strict preference (Chavas 2004:24). Both transitivity and independence are considered to be hypotheses of rationality of choices. However, these axioms, particularly the independence axiom, have been challenged by some authors. For instance, the Allais (1953) experiment questioned the validity of the independence axiom as the hypothesis of rationality of choices. Based on Allais's argument, several nonexpected utility theories have been developed, including the prospect theory of Kahneman and Tversky (1979), Tversky and Kahneman (1986), Machina's (1987) generalized expected utility theory or *fanning-out* hypothesis, Quiggin's (1993) rank-dependent expected utility model and Looms and Sugden's (1987) regret theory, all of which attempt to provide an alternative descriptive theory of choice under uncertainty. However, the empirical tests of these alternative nonexpected utility theories of choice have shown that none of them adequately organize choice under uncertainty. Chapter 6 provides empirical evidence from Battalio et al. (1990), Birnbaum and Navarrete (1998) and Birnbaum (2004a, 2005) that indicates that none of the main nonexpected utility theories can serve as adequate alternative descriptive theories of choice under uncertainty. Therefore, the investor-lender firm model developed in chapter 8 of the present study will be based on the approaches of the expected utility theory.

Portfolio investment theories have paid little attention to the study of the behaviour of financial firms that make decisions to lend their funds to other firms instead of investing them themselves as an important element in analyzing their investment decisions. In addition, most models of portfolio investment behaviour ignore the problems of risk and uncertainty. Those that do incorporate the latter focus only on price and revenue uncertainty. Little attention has been given to cost uncertainty because financial firms, particularly banks, are considered to be rational portfolio investors with no costs. Elyasiani (1983:1002) states that most banking models treat the bank as a rational investor with no resource cost constraint nor a production function. The present study does not intend to incorporate production function constraint in the analysis of the I-L firm investment and lending decisions but will incorporate cost-of-funds or deposit cost uncertainty in the analysis (see ch 7 and 8).

One of the key problems of the present study, that is, that investor-lender firms invest most of their wealth in risky financial assets instead of in riskfree government bonds, is investigated by developing an investor-lender firm optimization model. As stated earlier, this model incorporates both price and cost-of-funds uncertainty. Chapter 8 provides solutions to the problem of the optimal decision of profit maximizing I-L firms in the presence of these sources of uncertainty, where interpretation of the result is given in the context of both expected utility and nonexpected utility theories.

Another key problem of the study, that is, the investigation of the effects of uncertainty on aggregate fixed investment of private business firms is carried out in chapter 9. Five sources of macroeconomic uncertainty are used in the analysis. These are: real effective exchange rate uncertainty, output growth uncertainty, terms of trade uncertainty, uncertainty about changes in real interest rate and producer price uncertainty. The measures of these sources of uncertainty are obtained using GARCH (1,1) volatility model. The investment equation is estimated using restricted error correction (ECM) model that includes both the accelerator and neoclassical investment determinants and the uncertainty measures.

The remaining part of the present study is organized as follows: chapter 2 provides the analyses of the theories of investment decisions under certainty and risk. The impact of uncertainty on economic decisions is investigated in chapter 3. Chapter 4 deals with ambiguity aversion and its impact on optimal investment decisions of financial firms.

The impact of the interactions between uncertainty, investment irreversibility and financial constraints is analyzed in chapter 5. Chapter 6 concentrates on the expected utility and the alternative nonexpected utility theories of choice under uncertainty. Chapter 7 analyzes banking firm investment decisions under uncertainty while chapter 8 develops the investor-lender firm model. Chapter 9 is concerned with empirical econometric analysis of the effects of uncertainty on aggregate fixed (irreversible) investment of private business firms while chapter 10 provides the final conclusion.

CHAPTER TWO

THEORIES OF INVESTMENT DECISIONS UNDER CERTAINTY AND RISK

2.1 INTRODUCTION

Investment decisions by a firm are affected by the degree of risk and uncertainty it faces with regard to the future anticipated returns from these investments. However, traditional theories of investment ignored the latter aspects in their analysis of firms' investment behaviour because they based these analyses on the assumption of certainty.

Theories that incorporate risk in the economic decisions of firms state that economic decisions by agents depend on their attitude towards risk. Three factors contribute to the existence and prevalence of risk. These are: (a) inability to control and/or measure precisely some causal factors of events, (b) limited ability to process information and (c) the cost of information in the sense that obtaining and processing information is always costly, implying imperfect information and imperfect knowledge (Chavas 2004:6-8). Risky events are usually described by using probability (either *a priori* or statistical probabilities) of the occurrence of events. In terms of probability, an event z has a probability p(z) such that $0 \leq p(z) \leq 1$. The special case of this where $p(z) = 1$ refers to a situation of perfect certainty or sure events. And since risky events and sure events do occur at the same time, we can conclude that risky events are characterized by $p(z) < 1$.

An agent's attitude towards risk is classified into three categories. These are risk aversion, risk neutrality and risk seeking. The expected utility theory has been developed to explain two of these attitudes towards risk; namely, risk aversion and risk seeking. According to the expected utility theory, risk aversion is defined as a preference for a sure outcome over an uncertain prospect with equal or greater expected value, while risk seeking occurs when an uncertain prospect is preferred to a sure outcome with equal or greater expected value (Tversky and Fox 2000:93-94). The term prospect here refers to something expected, a possibility or chance. As far as attitude towards risk is concerned, of the two attitudes toward risk described above, the most prevalent behaviour in economic analysis is risk aversion as people are commonly assumed to be risk averse, i.e. they prefer certainty to risk.

One of the objectives of the present chapter is to identify the defects in traditional investment theories in terms of their failure to incorporate real life situations of uncertainty in their analysis of the economic decisions of firms. The chapter analyzes three groups of investment theories based on certainty assumptions. These are the Fisher theory, the accelerator theory, and the neoclassical theories of investment. A second objective of this chapter is to critically review various theories of risk attitudes and measures of risk, including the alternative models of risk attitude, and to investigate their implications for investment decisions of firms. The various theories of risk attitude are developed, based on the general Von Neumann-Morgenstern expected utility model. This chapter reviews this theory as the basic theory of decisions under risk.

The chapter is organized as follows: section 2.2 analyzes the various theories of investment decision under certainty with particular emphasis on the Fisher theory, the accelerator theory and the neoclassical theories of investment. Section 2.3 reviews alternative models of choice under certainty. Section 2.4 focuses on the theories of economic decisions under risk with particular emphasis on the Von Neumann-Morgenstern expected utility theory, the Arrow-Pratt measures of comparative absolute and relative risk aversion, the Ross characterization of risk aversion, and the Rothschild-Stiglitz measures of riskiness and measures of risk in developing and developed countries. Alternative models of risk behaviour are presented in section 2.5. Section 2.6 deals with investment decisions under risk, while section 2.7 presents a summary and conclusion.

2.2 THEORIES OF INVESTMENT DECISIONS UNDER CERTAINTY

2.2.1 The Fisher theory of investment

Traditional investment theories assumed that decision makers operated under full knowledge of the future, that is, with certainty. One such theory is the Fisher (1930) theory of investment. The Fisher theory states that investment is redistribution of consumption opportunities by an individual over time to maximize utility within his opportunity set that contains endowment, financial opportunity and productive opportunities that can be employed optimally with certainty (Hirshleifer 1965:509-510). Thus, the two important aspects of the Fisher theory of investment are that first, it uses an individual instead of a firm as the basis of its analysis, and, second, the choices these individuals make over time are considered to be riskless.

A criticism of Fisher's theory is that it does not make a clear distinction between investment and spending. He considers investment as nothing other than postponed spending and the value of goods, the purchase of which constitutes investment, is determined by the final or subjective income obtained after deduction of a discount (Haberler 1931:501). For instance, Fisher (1907:125-126) defines investment as purchasing the right to remote enjoyable income in preference to immediate enjoyable income, thereby equating investment with future spending.

In this theory, the objects of choice are present consumption and future consumption and individuals attempt to maximize their utility based on their opportunity set of endowment (or income), financial opportunities and productive opportunities. Thus, investment is considered to be the redistribution of consumption opportunities over time. The presence of financial opportunities for investment enables individuals to transform their endowment into alternative consumption bundles so as to maximize their utility. However, they can attain this only through exchange with other individuals.

Under pure exchange, the individuals will be in equilibrium when their current and future consumption equals their current and future endowment respectively, which implies that the social total of investment under pure exchange is equal to zero (Hirshleifer 1965:512). On the other hand, under productive opportunities the individual investor maximizes his utility when he attains the highest possible wealth level through productive investment. The above argument is based on an assumption of certainty about the future.

Under uncertainty, however, both the preference and objects of choice are different. Objects of choice become commodities instead of consumption decisions and preferences involve not the maximization of utility but rather the maximization of expected utility.

The assumption of certainty in investment decisions of individuals or firms ignores several issues that are important in investment decisions. These include, among others, the difference in returns from various forms of investment, liquidity preference, the willingness to insure against risk, and decisions involving debt and equity financing. Firms are concerned about these issues because of the uncertainty of the future outcomes of the decisions they make today. The Fisher theory is not the only theory based on the assumption of certainty. The following section presents the accelerator theory of investment that follows a similar line of argument in analyzing the investment behaviour of firms.

2.2.2 The accelerator theory of investment

The accelerator theory of investment, devised by Clark (1917), was one of the first theories of investment. However, due to its limited explanatory power, the model did not enjoy wide recognition until it was further transformed into a flexible accelerator model by Clark (1944). The accelerator model presents investment as a linear function of output with bidirectional causality where increase in output has a direct impact on investment, whereas investment affects output through the multiplier (Brown 1988:209).

The accelerator theory was made up of fixed and flexible accelerator models. In the former version, because the accelerator is assumed to be constant, there is always a fixed connection between current growth of capital stock and the current rise in final output. The validity of such a theory of investment depends upon whether the accelerator really is constant (Knox 1952:271-272). Similarly, Caballero (1999:816) argues that the fixed accelerator model was derived by inverting a simple fixed proportion production function and taking first differences and was thus unable to account for the serial correlation of investment beyond that of output.

In the flexible accelerator model it is assumed not only that capital output ratio is variable but also that firms close only part of the gap between the current and the desired capital stock. This makes investment a fraction, β, of the difference between the desired capital, K^*, and the actual capital, K_{-1}. Thus, in the flexible accelerator model:

$$I = \beta[K^* - K_{-1}]; \qquad 0 < \beta < 1 \qquad (2.1)$$

Alternatively, the change in capital stock is represented by

$$I_t = \sum_{i=0}^{n} \beta_i \Delta K^*_{t-i} \qquad (2.1')$$

where I_t = investment at time t, β_i = distributed lag parameters, K^*_{t-i} = the desired capital stock at time t-i.

In the case of the fixed capital output ratio the desired capital stock K^* is:

$$K^* = \alpha Q \qquad (2.2)$$

where Q is output.

Substitution of equation (2.2) for K^* in either (2.1) or (2.1′) demonstrates that in the accelerator model, investment is a linear function of output only. The model fails not only to account for the user cost of capital, which has an immediate and direct impact on the investment decisions of firms, but also to incorporate the greater uncertainty that firms face in the their day to day economic decisions.

The accelerator theory is based on a hypothesized relationship between sales and plant capacity utilization in which an increase in the former is expected to increase the latter. The implicit assumption of this model is that an increase in current sales leads to an increase in the current capacity utilization, which results in an increase in investment. However, the difficulty with this assumption is not only that the changes in one time sales alone cannot provide a good signal to induce firms to enter into a long term commitment of investment, but also that increased capacity utilization may not be linked to investment decisions (Brown 1988:213).

An attempt has been made by the proponents of the alternative econometric models of investment to improve the theoretical foundation of the accelerator model. For instance, Robert Eisner, who was one of the main advocates of econometric models of investment behaviour, hypothesized the existence of permanent acceleration and stated that, to induce investment, the change in output or sales should be permanent and that profits and capacity utilization have positive roles to play in investment (Eisner 1963:240).

In the accelerator model, the investment function is so simplified that it is only change in output that induces firms to acquire more fixed assets. However, in the real world investment decisions are affected by a host of other factors such as expected future returns on investment, costs of machinery, policy expectations and the like.

Moreover, in the face of an uncertain future, firms that have a past history of violent fluctuations in the demand for their output are most likely to be unwilling to expand their plant size in line with an increase in the current demand. Furthermore, firms may be concerned in their investment policy with distant possibilities that are different from current movements of markets (Knox 1952:292). The neoclassical theories try to correct these apparent deficiencies in the accelerator model by incorporating the user cost of capital into the equation of investment.

2.2.3 The neoclassical theories of investment

Neoclassical economics refers to the development in economic thinking following the marginal utility revolution of the 1870s when economic agents were believed to have made their decision based on the margins. Neoclassical economics is based on the fundamental assumptions of rational economic agents with utility maximizing individuals and profit maximizing firms with full information about the future, i.e. with certainty about the future. The assumption of a profit maximizing firm is behind the various theories of the firm. The neoclassical theories of investment are also derived from these fundamental assumptions of neoclassical economics. In the following sections we review the typical neoclassical theories of investment, i.e. the cost-of-capital theory and the q theories of investment.

2.2.3.1 The cost-of-capital theory

The cost-of-capital theory of Jorgenson (1963) emphasizes the central feature of the neoclassical investment theory which centres on the analysis of the response of change in demand for capital to changes in relative factor prices or ratio of factor prices to the price of output (Jorgenson 1963:247). The cost-of-capital theory is meant to remedy the weakness of the accelerator theory, which failed to incorporate the cost of capital in its investment function. In this theory, therefore, the main determinants of the level of optimal capital stock are output and the user cost of capital.

By utilizing the neoclassical optimization problem of a perfectly competitive firm with no adjustment costs and myopic expectations, and a constant returns Cobb-Douglas production function, Jorgenson derived the static first order condition relating investment to output and user cost of capital (Caballero 1997:4). Thus, for Jorgenson the desired amount of capital stock, K, becomes:

$$K = \alpha \frac{Y}{C_k} \quad (2.3)$$

where Y = output, C_k = user cost of capital and α = the share of capital in a Cobb-Douglas production function. In contrast to the accelerator model, in the cost-of-capital model, the level of investment is determined not only by the level of output or change in current demand but also by the level of current user cost of capital. The user cost-of-capital model is theoretically more plausible than its predecessors.

However, this theory also failed to provide an adequate empirical model for the analysis of investment behaviour. The main defect of the cost-of-capital model, in addition to its ignoring uncertainty about the future returns from investment, is that it used the same coefficient, α, for both the cost of capital and output and hence was unable to find a role for the cost of capital in the investment function. The failure of the cost-of-capital model led to the development of an almost equivalent but more dynamic q theory of investment.

2.2.3.2 The q theory of investment

This is a neoclassical version of a dynamic theory of investment, which is attributed to James Tobin (1969) and based on the idea that investment is a function of the ratio of the market value of a firm's assets to their replacement cost. This ratio, also known as the average q, is written as:

$$q = \frac{M_v}{R_c} \quad (2.4)$$

where M_v = the market value of the firm's assets and R_o = the cost of purchasing these assets in the market.

According to this theory, firms must invest when q >1 and disinvest when q < 1. For instance, if the value of the firm's outstanding shares of common stock is $2,000,000 and the replacement cost of the existing machinery and buildings is $1,000,000, the q ratio becomes 2. Since q is greater than 1, the firm can increase the value of its assets by issuing stock and buying more machinery and equipment. If we assume that the firm raises $500,000 by selling new stock and uses this to buy new equipment and machinery, the firm size increases by 50% while the value of outstanding stock increases only by 25%. Thus, the firm can pay the increased dividend claims on the outstanding shares of stock. However, if the opposite is true and q is less than 1, the firm will not be able to pay the increased dividend claims and therefore should not make any investment (Brown 1988:215-210).

When q =1, the market price of the firm's assets is equal to the value of the firm and hence the firm is in long run equilibrium. Tobin (1969:23) argues that when capital is valued at its reproduction cost, i.e. when q =1, this can be regarded as the condition of long run equilibrium. The q ratio summarizes all the necessary information relevant to future investment decisions and was considered to be a suitable statistic for investment in a wide variety of cases (Caballero 1997:5). Thus, the investment equation of the q theory became:

$$I = \eta q \tag{2.5}$$

where η is a strictly positive parameter since optimal investment is assumed to be an increasing function of the average q.

Abel (1983:230), however, has argued that the optimal rate of investment is the rate that equates the marginal adjustment cost with the marginal q, or the marginal value of the installed capital, instead of the average q. While the latter is potentially observable it is the marginal q that is relevant for investment decisions. Furthermore, Hayashi (1982:218) has shown that for a competitive firm, when the adjustment cost function and production functions are linearly homogeneous in factor inputs, marginal and average q are equal.

The q theory differs from the cost-of-capital theory in that it incorporates adjustment costs in the capital stock, making the firm's optimization problem dynamic as opposed to a static one time optimization problem assumed by the latter theory. This adjustment cost is assumed to be convex, i.e. marginal adjustment cost increases with an increase in investment.

Thus, for the neoclassical theories the optimal investment decision depends on profit maximization behaviour of firms subject to the constraint of the convex cost of adjustment. However, Caballero (1999:828-830) argues that the traditional q theory is not likely to work because the value function associated with it does not have the property of global concavity. Two points with the same value of marginal q lead to very different levels of investment and there is no function mapping marginal q to investment.

The q theory was thus also unable to provide an adequate theory of investment behaviour, not only because it failed to identify whether the firm's dynamic optimization problem could bring any degree of intertemporal uncertainty, but also because the model was unable to account for loss of global concavity and monotonicity. And so the neoclassical theories of investment were unable to provide adequate empirical models to investigate the investment behaviour of firms.

Most traditional theories of investment have implicitly assumed certainty mainly because of the ease of constructing an important investment decision rule of the net present value (NPV) decision criterion. The NPV rule is calculated using the rate of interest and the future stream of benefits from the investment. The problem presented by this rule is that it is difficult both to know about the future return from an investment project and to estimate the future rate of interest used in the calculation of the discount formula (Brown 1988:203). Because these variables are not known for certain, the NPV rule could lead to inappropriate investment decisions by firms. This is one of the serious weaknesses of the theories of choice under certainty.

The accelerator and neoclassical theories of investment focus on the analysis of the behaviour of firms when investing in fixed capital stock under the assumption of certainty. The investigation of these theories serves as a precursor to the analysis of various theories of investment in financial assets (which is the concern of the present study) that are based on similar assumptions of certainty. These theories are examined in detail in chapter seven of the present study.

2.3 ALTERNATIVE MODELS OF CHOICE UNDER CERTAINTY

The literature on choice under certainty has recently made some progress. It has developed new techniques for characterizing preferences and technologies using the concepts of distance functions (Shepard 1953; Malmquist 1953) and benefit functions (Luenberger 1992, 1994). Luenberger (1992, 1994) introduced the benefit function and demonstrated its usefulness in characterizing preferences and Pareto-efficient outcomes under certainty.

The Luenberger (1992) benefit function for the preference structure $B: R \times Y^S \to R$, defined for $g \in R^S$ by $B(w, y) = \max\{\beta \in R: W(y- \beta g) \geq w\}$ measures utility with respect to a given reference utility level and in terms of the willingness to trade for a given commodity bundle (Quiggin and Chambers 1998:122). In this sense, the benefit function is the natural precursor of the standard expenditure function, which means that the expenditure function is the dual of the benefit function (Luenberger 1992:462). Another related concept is the distance function of Shepard (1953:41) which is related to the distance function for homothetic production function, and Malmquist's (1953:233) distance function related to the price-quantity indifference curve. According to Quiggin and Chambers (1998:123), the distance function, D: $R^s_+ \times R \rightarrow R_+$, (which is only relevant when y is nonnegative), is defined by $D(\mathbf{y}, w) = \sup\{\lambda > 0: W(y/\lambda) \geq w\}$ $y \in R^s_+$. Both the distance function and the benefit functions are alternative representations of the preference structure. These functions fulfil the usual requirements of monotonicity, transitivity and continuity, but their most important characteristics are the translation property of the benefit function and the corresponding homogeneity property of the distance function. These latter properties are used in the analysis of the relationship between the certainty equivalent and various risk measures. Thus, benefit and distance functions are used not only to describe Pareto-efficient outcomes under certainty, but also to derive equivalent measures of risk.

Compensating and equivalent benefits derived from benefit functions lead therefore to the elimination of risk, which implies that the risk premium is a compensating benefit measure. On the other hand, the distance function yields analogous measures of compensating and equivalent relative benefits which leave the decision maker indifferent to both the original and the new prospects, and this implies that the relative risk premium is a compensating relative benefit (Quiggin and Chambers 1998:132-133). This recent development in the theory of choice under certainty is a major shift away from traditional artificial assumptions of perfect certainty. The next section presents a detailed analysis of economic decisions under risk.

2.4 THEORIES OF ECONOMIC DECISIONS UNDER RISK

2.4.1 The Von Neumann-Morgenstern expected utility theory as the basic theory of decisions under risk

Von Neumann and Morgenstern (1944, 1947) analyzed the decision-making behaviour of economic agents using the expected utility model, which became one of the pillars of economic theories under risk. According to Fishburn (1988:7), the theory begins with a binary relation $\succ$ on a convex set of P where the axioms of the theory imply the existence of a real valued function u on P that preserves the order of $\succ$ on P and is linear in the convexity operation so that for all p, q $\in$ P and all $0 \leq \lambda \leq 1$,

$$p \succ q \Leftrightarrow u(p) > u(q) \qquad (2.6)$$

$$u(\lambda p + (1-\lambda)q) = \lambda u(p) + (1-\lambda)u(q) \qquad (2.7)$$

where $\succ$ refers to preference, where $p \succ q$ is interpreted as "p is preferred to q"; P is the set of probability distributions or measures; p and q are random variables which are elements of P; u is real valued utility function and λ is any variable between 0 and 1.

The relation described in (2.6) refers to order preserving property, while (2.7) refers to the linearity property. These properties are further elaborated below under the main axioms of expected utility theory.

In this theory, P is interpreted as a set of probability distributions on X where P contains each measure that assigns probability 1 to some outcome, and defines u on X from u on P by u(x) = U(p) when p(x) =1. This theory derives the expected utility from this definition and from linearity described above, where for every simple measure of p in Px,

$$u(p) = \sum u(x)p(x) \qquad (2.8)$$

Equation (2.8) is the Von Neumann-Morgenstern measure of expected utility. The preference under Von Neumann-Morgenstern theory applies not to any outcomes but strictly to comparisons of risky alternatives (Fishburn 1988:7). This is ensured by the use of objective probability distribution P assigned to the outcomes of the states of nature.

Von Neumann and Morgenstern also developed a linear utility representation using axioms for simple preference comparison that refer to ordering, linearity and continuity (Fishburn 1988:24). There are five basic axioms of expected utility theory. These axioms are commonly identified as assumptions A1-A5 and are elaborated below:

Assumption A1 (Ordering and transitivity)

- For any random variable p and q exactly one of the following must hold:

  $p \succ q$ or $q \succ p$ or $p \sim q$

- If $p = \succ q$ and $q = \succ z$, then $p = \succ z$, where $= \succ$ refers to non strict preference. This refers to the assumption of transitivity or preferring more to less.

Assumption A2 (Independence)

For any random variables p, q and r and any λ $(0 < \lambda < 1)$

- $p \succ q \rightarrow \lambda p + (1-\lambda)r \succ \lambda q + (1-\lambda)r$,

this means that the preference between p and q is independent of r.

Assumption A3 (Continuity)

For any random variables p, q and r where $p \succ q \succ r$ there exist numbers α and β where $0 < \alpha < 1$ and $0 < \beta < 1$ such that

$$\alpha p + (1-\alpha)r \succ q \text{ and } q \succ \beta p + (1-\beta)r$$

This means that a sufficiently small change in probabilities will not reverse strict preference.

Assumption A4

For any risky prospects p and q satisfying $Pr[p \leq r: p \prec= r] = Pr[q \geq r: q = \succ r] = 1$ for some sure reward r, then $q = \succ p$.

Assumption A5

- For any number r, there exist two sequences of numbers $s_1 = \succ s_2, = \succ \ldots$ and $t_1 \prec= t_2, \prec= \ldots$ satisfying $s_m \prec= r$ and $r \prec= t_n$ for some m and n.

- For any risky prospects p and q, if there exists an integer m_0 [p|p $\geq s_m : p \succeq s_m$] $\succeq$ q for every $m \geq m_0$, then p $\succeq$ q. And if there exists an integer n_0 such that [p| p $\leq t_n$: p $\preceq t_n$] $\preceq$ q for every $n \geq n_0$, then p $\preceq$ q (Chavas 2004:24).

Under assumptions A1- A5, for any risky prospects p and q, there exists a utility function *U(x)* representing individual risk preference such that

$$p \succeq q \text{ if and only if } EU(p) \geq EU(q)$$

where *E* is the expectation operator.

This is the expected utility theorem. Under the above five axioms, the Von Neumann-Morgenstern expected utility model provides an accurate characterization of behaviour under risk. In other words, under the above assumptions, observing which decision an individual makes is equivalent to solving the maximization problem: max *EU(x)* and as such the expected utility model can be used in positive economic analysis to predict and explain human behaviour under risk (Chavas 2004:25). This argument serves as one of the analytical bases for the development of the investor-lender firm model in chapter 8 of the present study.

Most economic theories focus only on the first three axioms of expected utility model, i.e. the first axiom, A-1, as an axiom of economic rationality, the second axiom, A-2, as the linearity in probability axiom, and A-3, as the axiom of continuity in the sense that an act is prevented from being preferred indefinitely. In fact, the Von Neumann-Morgenstern utility theory is also known as the linear expected utility theory.

Various theories of risk were developed within the Von Neumann-Morgenstern framework of maximizing expected utility. The Arrow–Pratt measures of risk aversion that focus on the curvature of u on X and the Friedman-Savage double inflexion utility functions both lie within this framework. However, the Von Neumann-Morgenstern expected utility theory has been challenged by alternative nonexpected utility models.

One of the earliest challenges to this theory came from Allais (1953) and is known as the Allais paradox. The Allais paradox shows that when agents are presented with two pairs of choices, the first pair involving certainty with less payoff and a risky situation with more payoff and the second pair involving two risky options with higher probability of fewer payoffs and a lower probability of more payoffs, most agents choose fewer payoffs with certainty in the first case and more payoffs with lower probability in the second case. This challenged the assumption of the linear transformation of two pairs of choices as presented in the independence axiom (A-2) and was considered to be a violation of the independence axiom. This violation of the independence axiom was called the *certainty effect* after Allais who emphasized the security of certainty as the driving force behind the violation of the independence axiom illustrated above (Fishburn 1988:37). The Allais *certainty effect* was the first major challenge to the expected utility theory but it was not considered as an alternative to the theory. Instead, various authors provided explanations of it as an exception using various approaches within the EU theory. Later, however, the Allais paradox led to the development of alternative nonexpected utility theories. A detailed comparison of the expected utility models and the alternative nonexpected utility models will be provided in chapter 6 of the present study. The following sections continue the review of theories of risk aversion and riskiness (increased risk) based on the Von Neumann-Morgenstern expected utility theory.

2.4.2 Measures of risk

2.4.2.1 The Arrow-Pratt measures of risk aversion

In the expected utility theory, the behaviour of risk aversion is indicated by the concavity of utility functions, that is, the negativity of the second order derivative of the utility functions. However, since utility functions are not unique, their second derivatives are not unique either, making comparison between two utility functions difficult. The most famous measure of risk aversion was developed by John Pratt (1964) and Kenneth Arrow (1965) who used both risk premium and the concavity of the utility functions to measure risk aversion.

Let U_1 and U_2 be two elementary utility functions over wealth representing preferences $=\succ U_1$ and $=\succ U_2$ over M (the set of all random variables) respectively, and $\pi^{U1}(x)$ and $\pi^{U2}(x)$ be measures of risk premium, then risk aversion is defined as: if risk premium for utility function U_1, $\pi^{U1}(x)$, is greater than the risk premium for utility function U_2, $\pi^{U2}(x)$, for all $x \in M$, then U_1 has more risk-aversion than U_2 in the sense that the risk premium is greater for U_1 than for U_2. This is called the premium measure of risk aversion. Karni (1985:27) argues that the most appealing measure of risk aversion is the risk premium which represents the maximum a decision maker would be willing to pay to avoid risk.

2.4.2.1.1The risk premium and risk preferences

The premium measure of risk aversion is useful in studying individual risk preferences. Consider the case of a decision maker facing uncertain monetary returns represented by random variable "α". Let w denote the initial wealth of the decision maker which is certain. Thus, the terminal wealth of the individual becomes: $(w + \alpha)$. Here, $\alpha > 0$ because if $\alpha = 0$, this means that individuals keep all their wealth in riskfree assets. If the preference pattern of the individual is represented by the utility function $U(w + \alpha)$, in line with the expected utility model where the utility function is an increasing function of the terminal wealth, the individual is made better off by an increase in his wealth (Chavas 2004:33).

The main question associated with the above analysis is how to measure the monetary value of risk in relation to the uncertain monetary reward expected by this decision maker. According to Chavas (2004:33), this can be done by *income compensation tests* which involve finding the change in sure income that would make the decision maker indifferent to changes in risk exposure. He suggests the use of three monetary valuations of risk. These are: (a) the selling price of risk, (b) the asking price (bid price) of risk and (c) the risk premium.

The selling price of risk, Z_s, is defined as the sure amount of money a decision maker would be willing to receive to eliminate or sell the risk "α" if he had it. In particular, Z_s refers to the sure amount of money that makes the decision maker indifferent to facing the risky prospect $\{w + \alpha\}$ versus facing the sure prospect $\{w + Z_s\}$. That is, Z_s is the monetary amount that satisfies the indifference relationship (Chavas 2004:33):

$$\{w + Z_s\} \sim \{w + \alpha\} \tag{2.9}$$

where ~ refers to indifference.

Under the expected utility hypothesis, this implies that Z_s is the solution to the implicit equation that equates the utility of the prospect with the expected utility of the risky prospect, i.e.

$$U(w + Z_s) = EU(w + \alpha) \tag{2.10}$$

According to Pratt (1964:24), Z_s is the cash equivalent or the smallest amount for which the decision maker would willingly sell the risk "α", if he had it.

The need for monetary valuation of risk arises because, given the assumption of diminishing marginal utility and the maximization of expected utility, individuals will always have to be paid to induce them to bear risk (Friedman and Savage 1948:280).

Another monetary valuation of risk is the asking price (bid price) of risk. The bid price of risk, Z_b, is defined as the sure amount of money a decision maker would be willing to pay to obtain or buy the risk "α". More specifically, Z_b is the sure amount of money that makes the decision maker indifferent to facing either the sure prospect $\{w\}$ or the risky prospect $\{w + \alpha - Z_b\}$. In other words, if Z_b is the monetary amount which satisfies the indifference relationship presented in equation (2.11) below (Chavas 2004:34), then,

$$\{w\} \sim \{w + \alpha - Z_b\} \tag{2.11}$$

under the expected utility model, the above implies that Z_b is the solution to the implicit equation that equates the utility of the sure prospect w and the excepted utility of the risky prospect, $w + \alpha - Z_b$, i.e.

$$U(w) = EU(w + \alpha - Z_b) \tag{2.12}$$

Pratt (1964:124) defines Z_b, the bid price of risk, as the largest amount the decision maker is willing to pay to buy risk "α". Pratt further argues that if the risk "α" is unfavourable, it is natural to consider the insurance premium π such that the decision maker is indifferent either to facing the risk "α" or paying the nonrandom amount π but if "α" is actuarially neutral, the risk premium and insurance premium will be the same (Pratt 1964:125).

Furthermore, in the absence of income effects, and where preferences satisfy $U(w + \alpha) = w + V(\alpha)$, the selling price of risk Z_s and bid price of risk Z_b are equal, i.e. $Z_s = Z_b$ while on the other hand, if preferences $U(w + \alpha) \neq w + V(\alpha)$, then preferences exhibit an income or wealth effect and hence the bid price and the selling price of risk differ from each other , i.e. $Z_s \neq Z_b$ (Chavas 2004: 34).

The third monetary valuation of risk is the risk premium. As discussed earlier, the risk premium π is defined as the sure amount of money the decision maker would be willing to receive to become indifferent to receiving either the risky return "α" or the sure amount [E(α)- π], where E(α) is the expected value of α. In other words, π is the monetary amount which satisfies the following indifference relationship:

$$\{w + \alpha\} \sim \{w + E(\alpha) - \pi\} \tag{2.13}$$

Under the expected utility model, this implies that π is the solution to the implicit equation that equates the expected utility of the risky prospect $w + \alpha$ with the utility of the sure amount $w + E(\alpha) - \pi$, i.e.

$$EU(w + \alpha) = U[w + E(\alpha) - \pi] \qquad (2.14)$$

Chavas (2004:34) shows that given that $U(w + a)$ is a strictly increasing function its inverse function U^{-1} exists and $U(w) = u$ and $w = U^{-1}(u)$ are equivalent, implying $U^{-1}(EU(w + \alpha)) = w + E(\alpha) - \pi$. Thus, the risk premium π can always be written as

$$\pi = w + E(\alpha) - U^{-1}(EU(w + \alpha)) \qquad (2.15)$$

In the literature of economics, $w + E(\alpha) - \pi$ is commonly referred to as the *certainty equivalent* of $EU(\bullet)$ or a sure money metric measure of utility while the risk premium π is defined as the shadow price of private risk bearing or an individual's willingness to ensure against risk (Chavas 2004:35).

2.4.2.1.2 Alternative Arrow-Pratt measures of risk aversion

The second Arrow-Pratt measure of risk aversion is based on the curvature of the utility function, u on X, where X is an interval of monetary amount defined as either a wealth level or gains or losses around a given present wealth level. This approach defines various types of economic behaviour in risky situations in terms of the curvature of u on X within the Von Neumann-Morgenstern framework of maximizing expected utility. These measures of risk aversion are implied by the Pratt (1964:128) theorem. According to this theorem:

Let u_1, u_2 be elementary utility functions over wealth, x, which are continuous, monotonically increasing and twice differentiable. Then, the following are equivalent:

(1) $-u_1''(x)/u_1'(x) > -u_2''(x)/u_2'(x)$ for every $x \in R$.

(2) $u_1(x) = T(u_2(x))$ where T is a concave function.

(3) $\pi^{u1}(z) \geq \pi^{u2}(z)$ for all $z \in M$.

The term $-u_1''(x)/u_1'(x)$ is the measure of risk aversion related to the concavity of the utility function (Pratt 1964:125). Equating the former with $r_{u1}(x)$, that is $r_{u1}(x) = -u_1''(x)/u_1'(x)$, is known as the *Arrow-Pratt Measure of Absolute Risk-Aversion* (ARA).

Thus, for a risk averse individual, $r_{u1}(x) > 0$ if u_1 is monotonically increasing and strictly concave, while for a risk neutral individual with a linear utility function, $r_{u1}(x) = 0$ and for a risk loving individual with strictly convex utility function, $r_{u1}(x) < 0$. In other words, if u is twice differentiable and increasing on X so that $u^{(1)}(x) >1$, then u is risk averse if $u^{(2)}(x) < 0$, risk seeking if $u^{(2)}(x) > 0$ and risk neutral if $u^{(2)}(x) = 0$, where the superscripts 1 and 2 indicate the first and second derivatives of the utility function respectively.

The Arrow-Pratt measure of risk aversion is useful in analyzing investment decisions by agents. Suppose that there is a possibility of choice of dividing a normalized unit amount of wealth between a riskless asset with gross return x and a risky asset with return $x + \check{z}$ where $E(\check{z}) \geq 0$, then the individual with utility function u_1 would demand at least as much of the risky asset as would the individual with the utility function u_2. This refers to the Arrow-Pratt characterization of asset demand under risky situations. Thus, two risk averse expected utility maximizers with certain initial wealth levels divided into risky assets with high return and riskfree assets with lower returns will put at least the same amount of their wealth in the riskfree asset.

The Arrow-Pratt measure of absolute risk aversion (ARA) can be constant, decreasing or increasing. The next three subsections provide a detailed investigation of the characteristics of ARA.

2.4.2.1.2.1 Constant absolute risk aversion

When the Arrow-Pratt ratio -U′′/ U′ is constant for all levels of wealth *w*, the special case of risk preference called constant absolute risk aversion (CARA) exists. Economists define different utility functions based on a preference pattern that follows CARA. The following are typical CARA utility functions for r ≡ -U′′/ U′ (Chavas 2004: 38):

a) r > 0 (risk aversion) corresponds to the utility function $U= -e^{-r.(w+a)}$

b) r =0 (risk neutrality) corresponds to the utility function U = w + a, and

c) r < 0 (risk seeking) corresponds to the utility function $U = e^{-r.(w+a)}$

The most important economic implication of the CARA risk preference pattern is the absence of an income or wealth effect, because changing the initial wealth level does not affect economic decisions. This applies whether the individual is risk averse, risk neutral or risk seeking. Thus, the CARA risk preference pattern implies that the risk premium π is independent of initial wealth *w*. If we define the risk premium as the willingness to purchase insurance, under CARA, the change in initial wealth does not affect the individual's willingness to purchase insurance (Chavas 2004:39).

2.4.2.1.2.2 Decreasing (increasing) absolute risk aversion

For decreasing absolute risk aversion (DARA), the risk premium π decreases with initial wealth and for increasing absolute risk aversion (IARA), it increases with initial wealth. If we define risk premium as the individual's willingness to insure, under DARA an increase in initial wealth tends to reduce the individual's willingness to purchase insurance, while under IARA it tends to increase the willingness to purchase insurance. Thus, under DARA, private wealth accumulation and the need for insurance act as substitutes while under IARA, they complement each other. Among the three types of risk preferences, empirical evidence suggests that besides being risk averse, most individuals exhibit DARA (Chavas 2004:40). However, this preference pattern may not apply to all levels of wealth. In section 2.5, we will review the alternative models of risk preference, those that argue that individuals do not exhibit a uniform pattern of risk preference for all levels of wealth and those that argue that risk preferences depend on changes of wealth rather than initial or terminal wealth levels.

2.4.2.1.2.3 Relative risk aversion

Some economists argue that absolute risk aversion does not explain risk preferences for proportional changes in the levels of terminal wealth and hence is a less appealing measure of risk attitude than relative risk aversion (RRA). RRA is another Arrow-Pratt measure of risk aversion related to the curvature of the utility function. RRA is obtained by weighting the absolute measure of risk aversion by the level of wealth, x. Thus, the RRA is defined as $Ru_1(x) = xru_1(x) = -xu_1''(x)/u_1'(x)$. The RRA is the most suitable measure of risk aversion, since it incorporates individuals' change in attitude towards risk in relation to changes in the level of their wealth.

While absolute risk premium is a monetary measure of the cost of private risk bearing, relative risk aversion will use the relative risk premium, which is expressed as a proportion of the individual's wealth and hence does not depend on units of measurement (Chavas 2004:44). As in the case of the absolute risk premium, the use of the relative risk premium implies three types of risk behaviour. These are: constant relative risk aversion (CRRA), where the relative risk premium is independent of a proportional change in terminal wealth; increasing relative risk aversion (IRRA) where the relative risk premium increases with the increase in terminal wealth; and decreasing relative risk aversion (DRRA) where the relative risk premium decreases with increases in terminal wealth. However, there is no empirical evidence that suggests which of the three relative risk aversion attitudes is prevalent among decision makers (Chavas 2004:45).

Although the Arrow-Pratt measure of risk aversion is very useful in the analysis of individuals' risk attitudes, when this type of measure of risk is applied to changes in the level of wealth, it has been observed that some people show preference reversal between gains and losses in that they tend to be risk averse in gains but risk seeking in losses (Fishburn 1988:18). This argument is related to the hypothesis of nonexpected utility theories, particularly the prospect theory. However, this assumption is not supported by empirical evidence. A detailed analysis of this and other hypotheses of the alternative nonexpected utility models is conducted in chapter six of the present study.

2.4.2.2 The Ross characterization of risk aversion

According to the Arrow-Pratt characterization of comparative risk aversion for expected utility maximizers and for a pair of twice differentiable Von Neumann-Morgenstern utility functions $U(\bullet)$ and $U^*(\bullet)$, the following two conditions are equivalent (Machina and Neilson 1987:1139-1140):

(AP.1) $U^*(x) \equiv \rho(U(x))$ for some increasing concave functions $\rho\ (\bullet)$,

(AP.2) $\frac{U^*_{11}(x)}{U_1^*(x)} \geq \frac{U_{11}(x)}{U_1(x)}$ for all x, and

(AP.3) if π^* and π solve $U^*(x-\pi^*) = \int U^*(\omega)dF_{x+\tilde{\varepsilon}}(\omega)$ and

$U(x-\pi) = \int U(\omega)dF_{x+\tilde{\varepsilon}}(\omega)$

where $E(\tilde{\varepsilon}) = 0$, then $\pi \geq \pi^*$,

and if $U(\bullet)$ and $U^*(\bullet)$ are both concave, these are equivalent to:

(AP.4) if $E(\tilde{z}) \geq 0$ and $\overline{\alpha}^*$ and $\overline{\alpha}$ respectively, maximize $\int U^*(\omega)dF_{x+\alpha\tilde{z}}(\omega)$ *and* $\int U(\omega)dF_{x+\alpha\tilde{z}}(\omega)$, then $\overline{\alpha}^* \leq \overline{\alpha}$.

The first condition states that the function $U^*(\bullet)$ is a concave transformation of $U(\bullet)$ while the second condition states that the Arrow-Pratt measure of absolute risk aversion $U_{11}(x)/ U_1(x)$ is everywhere as great for $U^*(\bullet)$ as for $U(\bullet)$. The third condition states that if $U^*(\bullet)$ is willing just to pay a risk premium of π to avoid actuarially neutral risk $\tilde{\varepsilon}$, about the wealth level x, then $U^*(\bullet)$ would be willing to pay at least this amount to avoid risk. The fourth condition states that, given the choice of dividing a normalized unit amount of wealth between a riskless asset with gross return x and a risky asset with return $x+\tilde{z}$ where $E(\tilde{z}) \geq 0$, $U(\bullet)$ would demand at least as much as the risky asset, as would $U^{*}(\bullet)$ (Machina and Neilson 1987:1140). Pratt (1964:124-125) argues that for unfavourable risk, $\tilde{\varepsilon}$, it is natural to consider the insurance premium which is the function of wealth and risk such that the decision maker is indifferent between facing the risk or paying the fixed amount to ensure. However, as discussed earlier, if the risk parameter $\tilde{\varepsilon}$ is actuarially neutral, the risk premium and the insurance premium coincide.

Using the Arrow-Pratt ratio of absolute risk aversion, the individual's marginal rate of substitution between risk and premium payments about an initial situation of certainty can be analyzed. Accordingly, for actuarially neutral risk $\sqrt{t} \bullet \tilde{\varepsilon}$ with var $(\tilde{\varepsilon}) = \sigma^2$ and with var $(\sqrt{t} \bullet \tilde{\varepsilon}) = t.\sigma^2$, where t, $\tilde{\varepsilon}$ and σ^2 are scale of risk parameter, risk measure and variance of risk measure respectively. Then the premium π that the individual would be just willing to pay to avoid this risk is obtained by solving the following equation (Machina and Neilson 1987:1140):

$$U(x-\pi) = \int U(\omega)dF_{x+\sqrt{t}.\tilde{\varepsilon}}(\omega). \qquad (2.16)$$

The Taylor expansion of the above equation in π and t at $t = 0$ yields the following equation:

$$-U_{11}(x)\, d\pi = \tfrac{1}{2} . \sigma^2 . U_{11}(x).dt \qquad (2.17)$$

so that

$$\frac{d\pi}{dt}\Big|_{t=0,} = \tfrac{1}{2} . \sigma^2 . U_{11}(x)/ U_1(x) \qquad (2.17')$$

which is equivalent to the Pratt (1964:125) measure of risk premium for a small, actuarially neutral risk. Therefore, for any level of actuarially neutral risk $\tilde{\varepsilon}$, the greater an individual's Arrow-Pratt ratio, or absolute risk aversion, the greater is the marginal rate of substitution between the scale of risk parameter t and the premium level π about an initial situation of certainty (Machina and Neilson 1987:1140).

Ross (1981:621-622) argues, however, that while the Arrow-Pratt measure of risk has been very useful in the analysis of individual behaviour towards risk, the risk premia and asset demand conditions, i.e. AP.3 and AP.4, are both formulated with reference to situations of complete certainty which implies the existence of premia for complete insurance against risk and the allocation of wealth between risky and completely riskfree assets. But Ross does note that in real world situations there is no complete security and that most insurance typically covers only some types of risk and not others: in the world of price level uncertainty and bankruptcy no asset can be completely riskfree (Machina and Neilson 1987:1141).

Accordingly, Ross (1981) has further generalized the Arrow-Pratt theory of risk to address questions of economic concern and risk attitudes not easily dealt with by the original theory. Fishburn (1988:18) argues that Ross was especially interested in the nonavailability of riskfree alternatives and the impact of this on comparative economic analysis.

Ross (1981:628-629) has developed an alternative characterization of comparative risk aversion according to which the following conditions for the pair of twice differentiable risk averse and concave utility functions $U(\bullet)$ and $U^*(\bullet)$ are equivalent:

(R.1) $U^*(x) \equiv \lambda.\, U.(x) + G(x)$ for some positive λ and nonincreasing concave

function $G(\bullet)$,

(R.2) - $\frac{U^{*}_{11}(x)}{U^{*}_{1}(y)} \geq \frac{U_{11}(x)}{U_{1}(y)}$, for all x, y, and

(R.3) if π^* and π solve $\int U^*(\omega)dF_{\tilde{x}-\pi^*}(\omega) = \int U^*(\omega)dF_{\tilde{x}+\tilde{\varepsilon}}(\omega)$ and

$\int U(\omega)dF_{\tilde{x}-\pi}(\omega) = \int U(\omega)dF_{\tilde{x}+\tilde{z}}(\omega)$ where $E(\tilde{\varepsilon}/x) \equiv 0$, then $\pi^* \geq \pi$ and

that this in turn implies

(R.4) if $E(\tilde{z}/x) \geq 0$ and $\overline{\alpha}^*$ and $\overline{\alpha}$ respectively maximize $\int U^*(\omega)dF_{\tilde{x}+\alpha\tilde{z}}(\omega)$

and $\int U(\omega)dF_{\tilde{x}+\alpha\tilde{z}}(\omega)$, then $\overline{\alpha}^* \leq \overline{\alpha}$

The main difference between the Arrow-Pratt conditions 3 and 4 and the Ross conditions 3 and 4 are that in the Ross conditions, an individual does not generally have any opportunity for complete certainty. Moreover, the risk premium condition (R.3) assumes that the initial wealth $\tilde{x}$ is itself random and that the individual is, at most, able to ensure against a conditional actuarially neutral risk $\tilde{\varepsilon}$ while the asset demand condition (R.4) involves an asset with random return $\tilde{x}$ and with return $\tilde{x}$ plus $\tilde{z}$ with a higher mean, but possibly with greater risk (Machina and Nielson 1987:1141).

The Ross characterization of risk aversion is useful in building a model of an investor-lender firm faced with the choice of whether to lend its assets to investors who make direct investments and face a risk of future loan default, or to invest its assets in riskfree securities and obtain riskfree returns. Moreover, Ross's characterization of risk attitude is interesting not only because it provides stronger measures of risk behaviour, but also because it indicates that riskfree options are limited in economic decisions. Chapter eight of the present study will explore the implications of this result on the decision problems of investor-lender firms. However, to avoid complications, in economic analysis, riskfree situations are considered as benchmark scenarios in the analysis of choice under risk and uncertainty.

2.4.2.3 Rothschild-Stiglitz measures of riskiness

One of the measures of higher risk suggested by Rothschild and Stiglitz (1970) was the idea of Mean Preserving Spread (MPS) or Mean Preserving Increase in Spread (MPIS) which simply means the movement of the tails of a probability distribution without any change to the mean of the distribution. In other words, given a distribution *S*, we construct another distribution *T* by increasing the spread or making the tails fatter in such a way that the means of both distributions remain the same, i.e. $\mu_S = \mu_T$.

Using the MPIS approach, Rothschild and Stiglitz (1970:227-229) provided the following definition of riskiness. The random variable *s* is less risky than *f* if *f* is generated by an MPIS on *s* if the following two conditions hold:

(a) The difference between the means of distributions s(x) and *f*(x) is equal to 0 (mean-preservation), and

(b) The difference between the area of distributions *f*(x) and *s*(x) ≥ 0 (increase in spread).

Alternatively, if *f* is a density function and if $g = f + s$, then

$$\int_0^1 g(x)dx = \int_0^1 f(x)dx + \int_0^1 s(x)dx = 1 \qquad (2.18)$$

and

$$\int_0^1 xg(x)dx = \int_0^1 x(f(x) + s(x))dx = \int_0^1 xf(x)dx \qquad (2.19)$$

The above equations imply that if $g(x) \geq 0$ for all *x*, then g is a density function with the same mean as *f*. Hence, adding a function such as *s* to *f* shifts probability weight from the centre to the tails. The function that satisfies this condition is referred to as mean preserving spread (MPS) and if *f* and *g* are densities and *g* - *f* is an MPS, then *g* differs from *f* by a single MPS (Rothschild and Stiglitz 1970:229). Thus, if Y corresponds to the density function *f* and X corresponds to the density function *g*, then Y is preferred to X by every decision maker with a nondecreasing concave utility function, u.

Muller (1998:189-194) introduced a concept of mean preserving dispersals. He tried to unify this concept, which is another notion of increasing risk, with Rothschild-Stiglitz's MPIS and other notions of increasing risk in the analysis of comparative risk by restricting the set of possible spreads.

Rothschild and Stiglitz (1970:225-226) also define riskiness by linear transformation of random variables. Thus, the random variable y is riskier than the random variable x if there is a random variable ε such that $y = x + \varepsilon$ and $E[\varepsilon|x] = 0$ for all x, i.e. the expected value of the disturbance term conditional on x is zero for all x. This means that y is equal to x plus a disturbance term (noise). The addition of the disturbance term ε, which implies an increased riskiness, ensures that y is definitely riskier than x. However, both definitions of riskiness are essentially equivalent. An increase in risk always decreases demand for risky assets if the relative risk aversion is less than or equal to unity and the absolute risk aversion is nonincreasing (Rothschild and Stiglitz 1971:71-72).

Diamond and Stiglitz (1974:337) extend the concept of mean preserving increase in spread of Rothschild and Stiglitz (1970) by providing an alternative definition of increase in risk and risk aversion which keeps the expectation of utility (instead of the mean of the random variable) constant. They further argue that increases in risk should affect more risk averse individuals more than they do less risk averse individuals and that the appropriate definitions of increases in risk and in risk aversion must be closely linked.

Their alternative definition refers to the concept of mean preserving increase in risk (MPIR) instead of the mean preserving increase in spread (MPIS), which depends on the constant mean of the random variable. Consider the following family of distribution functions F(θ, r), where F is assumed to be twice continuously differentiable with respect to θ and r, and θ is a random variable defined over a finite range which for convenience is limited to the unit interval. Consider also two distributions in the family $F(\theta, r_1)$ and $F(\theta, r_2)$. Diamond and Stiglitz (1974:338) argue that if $F(\theta, r_2)$ and $F(\theta, r_1)$ are generated by taking weight from the centre of the probability distribution and shifting it to the tails, while keeping the mean of the distribution constant, it is acceptable to say that $F(\theta, r_2)$ represents a riskier situation than $F(\theta, r_1)$ and that the difference between these two variables is a *mean preserving increase in risk.*

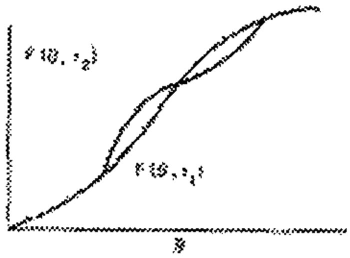

Figure 2.1 The Diamond-Stiglitz single crossing density functions

Figure 2.1 shows that the distribution crosses only once and hence $F(\theta, r_2)$ has more weight in both tails. Diamond and Stiglitz (1974:338) state that when the above situation holds, the difference between the distributions has the single crossing property and that this difference represents a *simple mean preserving spread.*

Such a spread can be characterized by the following two conditions:

$$\int_0^1 [F(\theta, r_2) - F(\theta, r_1)] d\theta = 0 \qquad (2.20)$$

and there exists θ such that

$$F(\theta, r_2) - F(\theta, r_1) \leq (\geq)\ 0 \text{ where } \theta, \geq (\leq)\ \bar{\theta}\,. \qquad (2.21)$$

The first condition ensures that the two distributions have the same mean while the second condition shows that there is a single crossing. Thus, the indefinite integral of the difference in the distributions is nonnegative, i.e.

$$\int_0^y [F(\theta,r_2)-F(\theta,r_1)]d\theta \geq 0 \qquad 0\leq y\leq 1 \qquad (2.22)$$

Consider another distribution, $F(\theta, r3)$, generated from $F(\theta, r_2)$ by a simple mean preserving increase in spread. $F(\theta, r_3) - F(\theta, r_1)$ does not , in general, have the single crossing property, as shown in figure 2.2 below.

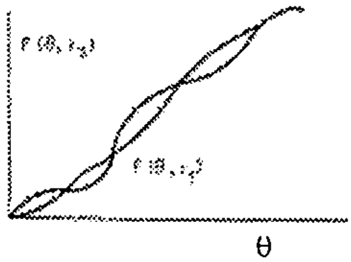

Figure 2.2 The Diamond-Stiglitz multiple crossing distribution functions

Since $F(\theta, r_3)$ is riskier than $F(\theta, r_2)$ and $F(\theta, r_2)$ is riskier than $F(\theta, r_1)$, it can be said that $F(\theta, r_3)$ is riskier than $F(\theta, r_1)$. Based on this argument, Diamond and Stiglitz (1974:339-340) state that the first two equations do not provide an adequate basis for the definition of riskier, but since $F(\theta, r_3) - F(\theta, r_1)$ does satisfy conditions (2.20) and (2.22), and since $F(\theta, r_3)$ can be generated from $F(\theta, r_1)$ as a limit of a sequence of simple mean preserving spreads, (2.20) and (2.22) provide a natural definition of increased risk. Thus, an increase in r (the "shift parameter") represents a mean preserving increase in risk if

$$\int_0^1 F_r(\theta,r)d\theta = 0 \qquad (2.23)$$

and

$$T(y,r)=\int_0^y F_r(\theta,r)d\theta \geq 0 \qquad 0\leq y\leq 1 \qquad (2.24)$$

Based on the above formulation, Diamond and Stiglitz (1974:342) define mean utility preserving increase in risk as

$$\hat{T}(y,r) = \int_0^y U_\theta F_r(\theta,r)d\theta \geq 0 \qquad \text{for all } y, \qquad (2.25)$$

and

$$\hat{T}(1,r) = \int_0^y U_\theta F_r(\theta,r)d\theta = 0 \qquad (2.26)$$

The interpretation of the above result is that if the first individual finds his utility decreasing from a given change for any mean preserving increase in risk for the second individual, the first individual is more risk averse than the second and the first individual will pay more for perfect insurance against any risk (Diamond and Stiglitz 1974:345).

Diamond and Stiglitz clearly link the two concepts of increase in risk and risk aversion. Thus, given their earlier definition of increase in risk, there is increased risk aversion if there is a monotone concave function g such that

$$U^1(\theta, \alpha) = g(U^2(\theta, \alpha) \qquad (2.27)$$

For a family of utility functions indexed by ρ, the requirement for identical indifference curves implies that we can write the family of functions $Z(\theta, \alpha, \rho)$ in the separable form $Z(U(\theta, \alpha), \rho)$, where the second order or concavity condition gives a derivative property

$$\partial^2 \log Zu/\partial u \partial \rho < 0 \qquad (2.28)$$

which indicates that risk aversion increases with ρ (Diamond and Stiglitz 1974:349).

2.4.2.4 Measures of country risk in developed and developing countries

In developed countries where asset markets are well developed, country systematic risk is measured using the beta measure as a ratio of covariance of equity return with a market return to the variance of market return developed by capital asset pricing models (CAPMs). However, while this approach is suitable to measure risk in developed markets, it is not appropriate for markets in developing countries. This is because we do not have betas for many developing economies' markets, as equity markets do not yet exist in these countries (Erb et al. 2000:126). Thus, since risk measures based on equity markets are not relevant for developing economies, alternative risk measure approaches are required to measure country risk in these countries.

The alternative risk measure for developing countries, which can also be used for developed countries, is the country risk rating. There are various services that provide country risk ratings. The most prominent ones are *International Country Risk Guide (ICRG)* and *Institutional Investor* (Erb et al. 2000:127). Four risk indexes have been developed by the ICRG. These are: political risk, economic risk, financial risk and composite risk indexes. Institutional Investor's risk index is called country credit ratings (CCR). These risk rating providers combine both quantitative and qualitative information into a single index to measure risk. For instance, CCR is calculated based on a survey of leading international bankers who are asked to rate each country on a scale from 0 to 100, 100% representing maximum creditworthiness. On the other hand, political risk scores of ICRG are based on subjective staff analysis of the available information while economic risk scores are based on objective analysis of quantitative data. Likewise, financial risk scores are based on an analysis of a combination of quantitative and qualitative data while the composite risk index is a linear combination of the three indexes' point scores. In this process, political risk, which represents willingness to pay, is given twice the weight of each of economic and financial risk, which reflect ability to pay (Erb et al. 1996:29-30).

Some authors, like Ferson and Harvey (1998:1635), link the fundamental valuation attributes in an economy, such as book to price, earning to price, dividend to price and price to cash ratios, to the risk exposure of national markets. These authors suggest an asset pricing model in which the fundamental attributes are linked to dynamic country risk models based on beta measures for both developed and emerging market economies.

There are also other approaches to measuring risk. These involve the use of macroeconomic variables and social development indicators to develop expected return, volatility measures and correlation estimates in developing countries. Some important variables include per capita GNP, output growth, the size of the trade sector, change in the exchange rate versus a benchmark, volatility of exchange rate, the size of the government sector, external debt level, the number of years of schooling, life expectancy and quality of life (Erb et al. 2000:133).

The above country risk measures assume that risk and uncertainty are one and the same thing. Therefore, these international risk measures can be considered as empirical counterparts of the neoclassical, new classical and rational expectation theories that reject a sharp distinction between risk and uncertainty. However, it is not always possible to assign numerical values to all types of uncertainty about the future faced by economic agents. As Knight and Keynes argued almost eight decades ago and as post Keynesians emphasize today, the separation between measurable risk and unmeasurable uncertainty continues to be an important aspect in the analysis of the behaviour of economic agents. However, in the present study uncertainty covers both risk and ambiguity while ambiguity is considered to be what Knight and Keynes define as unmeasurable uncertainty.

2.5 ALTERNATIVE MODELS OF RISK BEHAVIOUR

The expected utility model provides the basis for the study of decision making under risk. Under the expected utility model, individuals make decisions on alternative wealth levels x by maximizing the expected utility of x, i.e. maximizing $EU(x)$ where the utility function is defined up to a positive linear transformation. One of the main advantages of the expected utility model is its empirical tractability (Chavas 2004:79). However, some theories have challenged the ability of the expected utility model to fully explain individual behaviour under risk. The first challenge to this model regarding its capacity to fully explain individual risk behaviour came from Friedman and Savage (1948). We will review Friedman and Savage's alternative characterization of risk attitudes in the next section.

2. 5.1 Friedman-Savage measure of risk aversion

The Friedman-Savage challenge to the predictions of the expected utility model came in the form of the following question: is the expected utility consistent with the fact that some individuals purchase insurance and gamble at the same time? Or does the expected utility model explain the behaviour of simultaneous gambling and the act of insuring against risk? Friedman and Savage (1948:279) argue that in practice individuals exhibit the behaviour of simultaneous acts of insuring and gambling and the choice of different degrees of risk so prominent in such acts of insuring and gambling is clearly present and important in a much broader range of economic choices. They argue further that the empirical evidence of the willingness of individuals of all income classes to buy insurance is extensive as is that of the willingness of individuals to purchase lottery tickets or engage in similar forms of gambling (Friedman and Savage 1948: 285-286).

Friedman and Savage (1948:287-289) hypothesize that in choosing from alternatives open to it, whether or not these alternatives involve risk, an economic unit behaves as if (a) it had a consistent set of preferences; (b) these preferences could be completely described by a function attaching a numerical value or "utility" to alternatives, each of which is regarded as certain; and (c) its objectives were to maximize the expected utility of profit.

They argue, however, that the above hypotheses cannot rationalize the major assumption representing the behaviour of economic agents without imposing further restrictions on the utility function. Some of these statements which describe the behaviour of economic agents, include (i) consumers prefer larger to smaller certain incomes; (ii) low income consumers buy, or are willing to buy, insurance; (iii) low income consumers buy, or are willing to buy, lottery tickets; (iv) many low income households are willing to buy both insurance and lottery tickets (Friedman and Savage 1948:293-294). One of the restrictions they imposed is that total utility increases with the size of money income. They further argue that statement (ii) can be rationalized only if the utility functions of the corresponding units are not everywhere convex; statement (iii) can be rationalized only if the utility functions of the corresponding units are not everywhere concave and statement (iv) can be rationalized only if the utility functions of the corresponding units are neither everywhere convex nor everywhere concave (Friedman and Savage 1948:294). Thus, they conclude that the utility function consistent with the willingness of the low income consumer unit both to purchase insurance and to gamble is a concave-convex utility function which has different curvatures for different levels of wealth, i.e. a utility function which is initially concave but which becomes convex for higher levels of income. Figure 2.3 shows the Friedman-Savage concave-convex utility function.

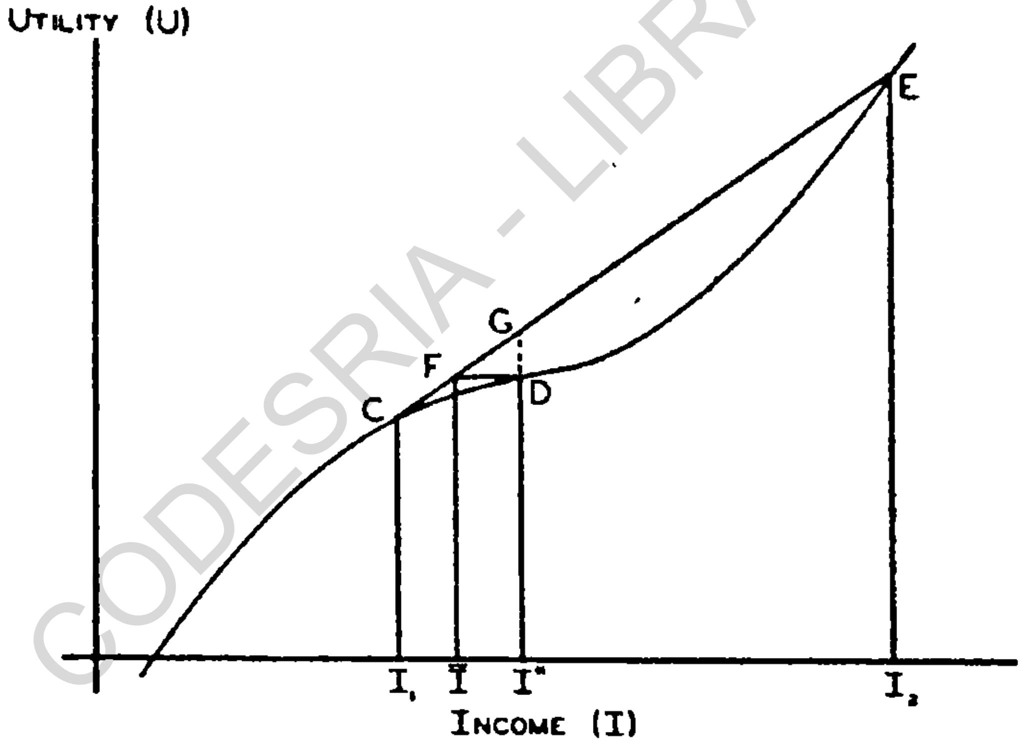

Figure 2.3 The Friedman-Savage concave-convex utility function

For levels of income below I^*, a consumer exhibits risk averse behaviour while for levels of income above I^* it exhibits risk seeking behaviour. This utility curve reflects an observed consumer's preference for both gambling and insuring. In other words, figure 2.3 indicates the utility function that is consistent with the willingness of a low-income consumer both to purchase insurance and to gamble. Friedman and Savage (1948:296) state that $I^{*} - \bar{I}$ represents the maximum amount in excess of the actuarial value that the corresponding consumer would pay for a gamble involving some fixed chance of winning $I_2 - I^{*}$ and some fixed chance of losing $I^{*} - I_1$.

The Friedman-Savage analysis implies that it is not plausible to smoothly equate conditions of risk aversion, risk neutrality and risk seeking with the concavity and convexity of the shapes of utility functions. The individual's utility function may not have the same curvature everywhere. The curvature may change from concave to convex, indicating different levels of socioeconomic class a consumer can attain at different times. At some levels of socioeconomic class (or wealth) an individual may be risk averse and at other levels of wealth he may be risk seeking (Freidman and Savage 1948:298).

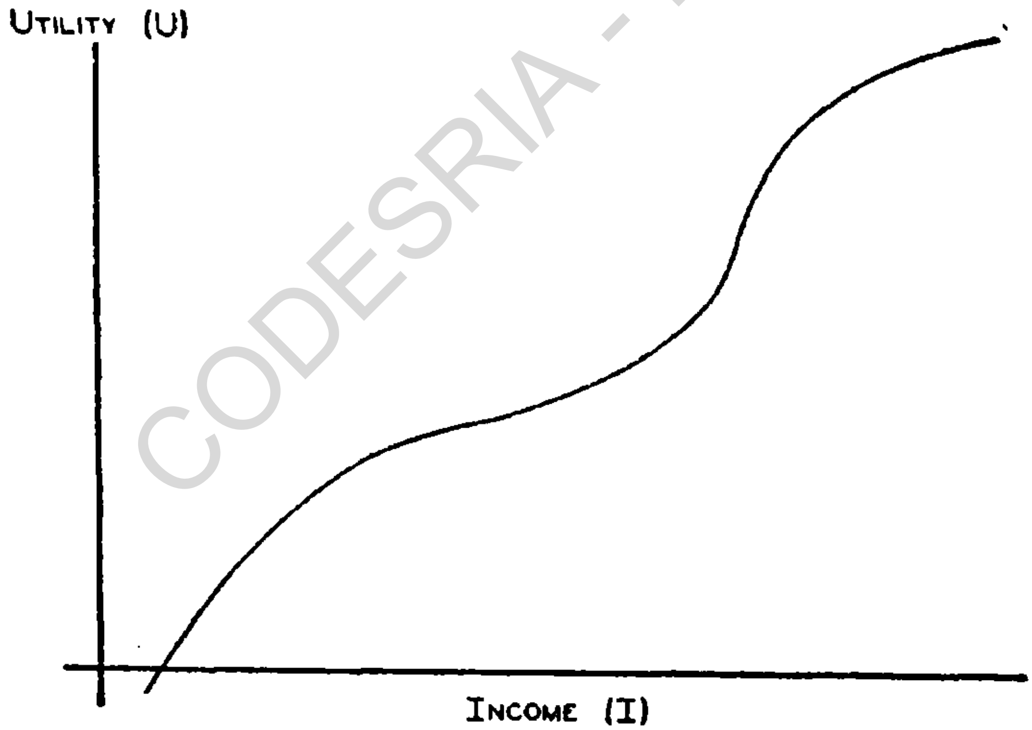

Figure 2.4 The Friedman-Savage utility function with double inflexion

Freidman and Savage (1948:297) generalize the utility analysis of choice involving risk by extending the curvature of the utility function into a doubly inflected one or into a concave-convex-concave utility function as depicted in figure 2.4. This indicates that an individual may exhibit different types of risk attitude at different levels of wealth. At low income levels (where the utility curve is concave) agents exhibit risk averse behaviour, risk loving behaviour at middle levels of income (where the utility curve is convex) and then again risk verse behaviour at higher levels of income (where the utility curve is once again concave) (Friedman and Savage 1948:297). Assuming that u on X is twice differentiable and increasing in x then, following Friedman and Savage, we can say that an individual with preference function u is risk averse in an interval of X if $u^{(2)}(x) < 0$, in that interval, he is risk seeking if $u^{(2)}(x) > 0$, in that interval. In other words, Friedman and Savage (1948:297) suggest the use of concave-convex-concave utility functions to explain economic behaviour such as simultaneous acts of purchasing insurance and gambling in actuarially unfair lotteries, which seems to go against the assumption of the expected utility theory in the sense that the stated behaviour is at variance with aversion to all unfair risks.

However, Friedman and Savage agree, as do other economists, that the above behaviour is not inconsistent with the expected utility model. Chavas (2004:79) state that, according to Friedman and Savage, for most individuals, the utility function U(x) is probably concave (corresponds to risk aversion and willingness to ensure) for low and moderate monetary rewards, but convex (corresponding to risk seeking and a positive willingness to gamble) for high monetary rewards, and in this context a particular individual may ensure against "downside risk" while at the same time gambling on "upside risk" and still be consistent with the expected utility model.

2.5.2 Other alternative models of risk attitude

The preceding theories of risk derive from the basic assumptions of dominance, invariance and transitivity of the expected utility theory. However, some of the alternative nonexpected utility theories, particularly the prospect theory, indicate that individuals' behaviour towards risk may differ depending on their preference patterns. These are presented in detail in chapter six of the present study.

2.6 INVESTMENT DECISIONS UNDER RISK

There are two types of risk that are important to investment decisions. These are systematic risk and non-systematic risk. Systematic risk refers to risk which cannot be diversified by investors, while non-systematic risk is potentially diversifiable. However, these concepts must be analyzed in relation to the types of investment the agents are planning to undertake. In financial investment, risk is important for the shareholders only in so far as it is not diversifiable. In other words, only systematic risk is important in their investment decisions because holding a balanced, diversified portfolio of shares eliminates non-systematic risk, i.e. risk incurred by returns uncorrelated with the return on a general share index (Driver and Moreton 1992:37).

Other authors distinguish between unique risk that is peculiar to a given stock and market risk that is associated with market wide variations where investors are believed to eliminate unique risk by holding a fully diversified portfolio, whereas they cannot eliminate market risk by diversifying the portfolio (Brealey and Myers 2003: 159). The two concepts of systematic and non-systematic risk and market and unique risk are two sides of the same coin in the sense that systematic risk refers to market risk which cannot be eliminated by diversifying a portfolio whereas non-systematic risk is unique risk because both can be eliminated by holding diversified portfolios.

Portfolio risk is calculated using portfolio variance (*PV)* defined as:

$$PV = \sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \sigma_{ij} \qquad (2.29)$$

which when expanded becomes:

$$PV = x_1^{\,2}\sigma_1^{\,2} + x^2{}_2\sigma_2^{\,2} + 2(x_1 x_2 \rho_{12}\sigma_1\sigma_2) \qquad (2.29')$$

where x_i's measure proportion invested in stock *i*, σ^2_i, variance of return on stock *i*, σ_{ij} = covariance of return on stock *i* and *j*, and ρ_{ij} correlation between returns on stock *i* and *j*.

On the other hand, the market risk is measured by beta which is a measure of sensitivity to market change. It is measured using the formula:

$$\beta_i = \frac{\sigma_{im}}{\sigma_m^{\ 2}} \qquad (2.30)$$

where σ_{im} is the covariance between stock *i*'s return and the market return while σ^2_m is the variance of the market return (Brealey and Myers 2003:149-153). Expected returns are calculated using objective probability distributions. For this reason managers take risk factors into consideration whenever they make investment decisions.

Using the Arrow-Pratt measure of relative risk aversion, Nickell (1978:79) shows that the risk averse firm has not only a lower optimal capacity compared to risk neutral firms, but its optimal capacity also decreases with the actual degree of risk aversion. The limitation of the Arrow-Pratt measure of risk aversion is that it is based on a comparative analysis of the attitudes towards risk of two economic entities. It does not deal directly with the problems of economic decisions made by a single economic unit such as the decision to allocate wealth between risky and riskfree assets. However, the Arrow-Pratt measure of risk aversion has been extremely valuable in economic analysis and the intuition behind it is useful in analyzing decision problems involving a single economic entity. Moreover, it is argued that risk measures do not explain the entire behaviour of economic agents when making investment decisions because in real world situations economic agents not only face measurable risk but also unmeasurable uncertainty.

It is important to distinguish between measures of risk and measures of attitudes towards risk. The variance and beta measures presented above are measures of risk but not attitudes towards risk. The attitudes towards risk are measured by the individual's absolute or relative aversion to risk, risk seeking or risk neutrality.

2.7 SUMMARY AND CONCLUSION

Although the concepts of risk, uncertainty and probability have been incorporated into economic analysis since the early 20th century, traditional theories of investment have continued to analyze firms' investment decisions based on the assumption of certainty. One of the earliest theories of investment, the Fisher theory, considered investment as the redistribution of consumption over time by individual economic agents so as to maximize their utility based on their endowment, financial opportunities and productive opportunities. Financial opportunities allow the individual to maximize his utility through transformation of his consumption bundle by entering into exchange with other individuals, while productive opportunities lead to productive investment and increased wealth, which ultimately increases utility with certainty.

The accelerator model attempted to formulate the investment theory in a more formal manner than the Fisher approach. However, the model was based on a simple transformation of a simple fixed proportion production function taking first differences and hence was unable to account for determinants of investment other than output and changes in demand.

The neoclassical theories of investment, on the other hand, assumed that the optimal investment decision depends on profit maximization behaviour of firms subject to the constraint of the convex cost of adjustment. Thus, both the cost-of-capital theory and the q theory of investment focused on the profit maximization behaviour of firms without any consideration of how this behaviour might be influenced by the presence of risk and uncertainty about the future.

There is ample evidence that the assumption of certainty in investment decisions of individuals or firms leaves out several issues that are important in investment decisions. These include return differentials between various forms of investment, liquidity preference, the willingness to gamble yet also to insure against risk, decisions involving types and sources of investment financing and so on. These issues are important in firms' investment decisions because they have limited or no knowledge about the future outcomes of the decisions they make today.

Alternative mathematical models of individual decisions under certainty have developed new techniques for characterizing preferences. This has involved the use of benefit functions and distance functions. The benefit function allows decision makers to derive compensating and equivalent benefits which equate the risk premium with the compensating benefit, which implies the elimination of risk. The analogous measures of compensating and equivalent relative benefits derived from the distance function leave the decision maker indifferent to the original and new prospects respectively, thereby equating the relative risk premium with the compensating relative benefit. Therefore, the alternative theories of decisions under certainty represent a real advance in terms of creating a link between certainty and risk.

Three conditions reflect prevalence of risk. These are: inability to control and/or measure precisely some causal factors of events, limited ability to process information and the cost of information, which implies imperfect information knowledge.

The expected utility model, based on the axiomatic characterization of individual preferences, provided a theoretical basis for the analysis of risk attitudes. There are two important attitudes toward risk. These are risk aversion and risk seeking. However, since it is assumed that people always prefer certainty to risk, risk aversion has been the dominant phenomenon in economic analysis. Accordingly, economists have developed various measures of risk aversion. Arrow and Pratt developed two measures of risk attitude. These are the premium measure of risk aversion and measures of comparative absolute and relative risk aversion which are based on the concavity of the utility function in line with the expected utility theory.

The risk premium is one method of monetary valuation of risk in that it represents the sure amount of money the decision maker would be willing to receive to become indifferent to receiving the risky return or the sure amount. Other monetary valuations of risk involve the selling price and the asking price (bid price) of risk.

The Arrow-Pratt comparative absolute and relative risk aversion may be constant, decreasing or increasing. Under constant absolute risk aversion (CARA), changing the initial wealth of an individual does not affect economic decisions, whether the individual is risk averse, risk seeking or risk neutral. Thus, CARA implies that the risk premium is independent of initial wealth, and that if the risk premium is defined as the willingness to purchase insurance, under CARA, the change in initial wealth does not affect the individual's willingness to purchase such insurance. However, for decreasing absolute risk aversion (DARA) and increasing absolute risk aversion (IARA), an increase in initial wealth decreases and increases willingness to purchase insurance respectively. Moreover, empirical evidence suggests that DARA reflects the most prevalent risk behaviour. Relative risk aversion (RRA) is obtained by weighting the absolute measure of risk aversion by the level of wealth and this also exhibits constant, increasing or decreasing properties. However, there is no empirical evidence to prove which of these properties reflects the dominant risk attitude.

Ross, Machina and Neilson have extended the Arrow-Pratt measures of risk aversion by incorporating issues such as the nonavailability of riskfree alternatives and the impact of this on the comparative economic analysis. In particular, Ross has further generalized the Arrow-Pratt theory of risk to address questions of economic concern and risk attitudes not easily dealt with by the original work of Arrow and Pratt. Ross not only provided stronger measures of risk aversion but also indicated that, in economic decisions, riskfree alternatives are limited. This argument is particularly relevant for the problems of choice confronting an investor-lender firm: when the objective of the firm is the maximization of the expected utility of profit, there may not be any riskfree alternatives that can maximize this objective.

Rothschild and Stiglitz and Diamond and Stiglitz focused on measures of riskiness. For the former, increase in risk is represented by the mean preserving increase in spread (MPIS) while the latter linked the definition of increase in risk and increase in risk aversion by employing similar concepts of mean preserving increase in risk and mean utility preserving increase in risk, both of which are used to measure increase in risk aversion.

Other authors provide completely different alternative definitions of risk aversion to the standard Arrow-Pratt and related measures. The first of these is the Friedman-Savage hypothesis regarding the empirical problems of two opposing decisions by economic agents in gambling and purchasing insurance. Based on this observed behaviour, they reject the smooth equating of risk preference with concavity or convexity of a utility function. They suggest that an individual can exhibit different risk attitudes at different levels of wealth and this can be represented by doubly inflected or concave-convex-concave utility functions where individuals are risk averse at lower levels of wealth, risk seeking at medium levels and again risk averse at higher levels of wealth, while and at the same time conforming to the behaviour described by the expected utility hypothesis.

In firms' investment decisions, the most significant forms of risk are systematic or market risk that cannot be eliminated by holding diversified portfolios, and non-systematic or unique risk which can be eliminated by diversifying the portfolio. These types of risks are usually measured by beta based on CAPM models. However, since capital markets do not exist in developing countries alternative country risk measures have been developed. These country risk measures are widely used in empirical economics. Risk measures are to be distinguished from measures of attitudes towards risk, which reflect individual aversion to risk, risk seeking or risk neutrality.

CHAPTER THREE

UNCERTAINTY IN ECONOMIC DECISIONS

3.1 INTRODUCTION

The analysis of the behaviour of investment in relation to changes in risk captures only part of the uncertain future faced by economic agents. Most real economic activities involve situations where simple mathematical calculations of the distribution of future outcomes may not be possible. Therefore, most economic decisions are made in an environment where the probability distribution of the outcomes of the decisions is unknown. Such an environment is characterized by the concept of ambiguity or uncertainty without known probability distribution.

This chapter reviews five groups of theories concerned with the impact of uncertainty on economic decisions. These are: Knight and Keynes's theories, Austrian economics, the new classical economics and rational expectations, the Arrow-Debreu theory of competitive equilibrium under uncertainty and Heiner's competence difficulty (C-D) gap theory. Knight and Keynes emphasize that in real life, economic agents face unmeasurable uncertainty and this affects their economic decisions negatively. On the other hand, the Austrian school of economic thought emphasizes the importance of uncertainty in the economic process and do not view it as a negative factor. For the Austrian school, uncertainty is necessary for the very existence of the market economic system. The new classical economics and rational expectations hypothesize that economic agents can meet uncertainty with the correct use of available information. Unlike other theories, the Arrow-Debreu theory of competitive equilibrium under uncertainty incorporates uncertainty into a general equilibrium setting through the introduction of contingent commodity markets, i.e. markets for trading of goods contingent on the state of the world. This is the Arrow-Debreu state preference approach. Heiner's competence-difficulty gap analysis attempts to rationalize Keynesian epistemic uncertainty by using reliability ratio measures. The remaining sections in this chapter will explore these theories.

The chapter is organized as follows: section 3.2 presents Knight and Keynes's theories on the impact of uncertainty on economic decisions. The views of the Austrian school of economic thought are reviewed in section 3.3. Section 3.4 concentrates on the new classical economics and rational expectations while section 3.5 analyzes the Arrow-Debreu theory of competitive equilibrium under uncertainty. Section 3.6 presents the analysis of Heiner's C-D gap theory while section 3.7 investigates the link between uncertainty, the market system and the government intervention. Finally, section 3.8 presents a summary and conclusion.

3.2 KNIGHT AND KEYNES ON THE IMPACT OF UNCERTAINTY ON ECONOMIC DECISIONS

Knight (1921) emphasizes the difference between true uncertainty and risk and their impact on firms' economic decisions and profit. Using the framework of perfect competition in economics, he shows how the presence of risk and uncertainty can create a wedge between actual and theoretical competition. It is true uncertainty, not risk, that "forms the basis of a valid theory of profit and accounts for the divergence between actual and theoretical competition" (Knight 1921:20). However, other than showing us how uncertainty creates this wedge between theoretical and actual competition, Knight did not provide us with any alternative means by which the magnitude of these divergences could be measured.

Instead, Knight focuses on the mechanisms with which economic agents cope with uncertainty. According to him, the possibility of reducing uncertainty depends on two fundamental sets of conditions. These are: (a) Uncertainties are fewer in groups of cases than in single instances and hence in the case of *a priori* probability, the uncertainty tends to disappear altogether as the group increases in inclusiveness while with statistical probabilities, the same tendency is manifest to a lesser degree. (b) The second mechanism for reducing uncertainty is reflected in the differences among human individuals with regard to uncertainty itself. There may be differences in individuals or their position with regard to uncertainty. Knight calls these two methods of reducing uncertainty by grouping and selection of individuals who bear it, "consolidation" and "specialization", to which he adds "control of the future" and "increased power of prediction" as two additional methods for coping with uncertainty (Knight 1921:238-239). However, there are two problems inherent in these methods of coping with uncertainty. First, Knight himself argues that consolidation of the true uncertainty is almost impossible because of the high degree of uniqueness of the instances being dealt with. This makes consolidation irrelevant as a method of reducing true uncertainty. Second, in a situation where the future is unknown, the control of that future as a mechanism to reduce the unknown outcome seems conceptually incoherent.

Knight is a neoclassical economist and a strong adherent of *laissez faire*. His main concern, therefore, is to analyze how the presence of uncertainty affects any future actions of economic agents and hence interferes with the perfect working of competition in accordance with the laws of pure theory (Knight 1921:230). However, for Knight, the market mechanism itself can correct uncertainty. That is why he emphasizes that the entrepreneurs and hence the market system can meet uncertainty by using the methods of consolidation, organization, control of the future, and increased power of prediction

In a similar vein, Keynes emphasizes the impact of uncertainty on economic decisions. In his analysis of asset market behaviour in *General Theory*, Keynes (1936:162-163) states that human decisions affecting the future cannot depend on strict mathematical expectations since there is no basis for making such calculations; instead, it is an innate urge (animal spirit) for activity that makes individuals choose the best possible alternative. For Keynes, the market system cannot meet uncertainty and the presence of uncertainty can force investors to behave in a nonrational manner. Gerrard (1995:190-191) remarks that, for Keynes, instead of calculating most probable forecasts which require probability distributions and the use of the marginal efficiency of capital involving the calculation of expected future returns from investment, a large portion of investment decisions depend on nonrational factors such as animal spirit and a spontaneous urge to action rather than inaction.

In addition to the innate urge to action, Keynes's argues further that in spite of being uncertain about the future, people have knowledge about certain aspects of the existing situation and have ways of coping if they do not have definite calculable knowledge of the possible outcomes of the current actions. One way of managing is to fall back on the conventional judgement used by the rest of the world (Lawson 1985:916). According to Keynes, there is both direct knowledge and indirect knowledge and people can obtain direct knowledge about these conventions or other general situations through direct acquaintances: experience, understanding and perception are the three main forms of direct acquaintance (Keynes 1921:12).

In other words, Keynes (1936:114) argues that although economic agents face uncertainty in the sense of being unable to determine precisely the future outcome of all current actions, they possess extensive knowledge of current societal practices which can provide the basis for determining how to get by. It is this action, based on such existing practices and conventions, that "saves the faces of men as rational economic agents". Thus, under Keynes's analysis, economic agents cope with the unknown future by using the existing practices and conventions. However, as uncertainty increases, non-rational, spontaneous decisions cannot be ruled out. Fitzgibbons (1995:217) argues that as society becomes dominated by commercial values and as a concept of extended reason diminishes, a high degree of uncertainty will tend to overcome any scope for rational action. He further states that if the cost of shifting between assets is low and if confidence has been lost in conventions that support investment behaviour, then investment decisions may become volatile. Although Keynes and post Keynesians emphasize the importance of uncertainty in investment decisions, they do not offer any alternative for analyzing the impact of this uncertainty on investment decisions as they preclude any numerical analysis in the uncertainty-investment relationship. Fitzgibbons (1995) describes Keynesian uncertainty as uncertainty without judgment.

Coddington (1982:482) also believes that, for Keynes, it is not the fact of uncertainty that is important but rather how individuals are supposed to respond to uncertainty. Therefore, in Keynes's analysis it is not the uncertainty surrounding private sector investment decisions that is the central issue but rather the wayward and unruly behaviour of the aggregates resulting from the decisions taken in the face of uncertainty.

3.3 AUSTRIAN ECONOMICS AND UNCERTAINTY

3.3.1 Mises, praxeology and uncertainty

The Austrian school of economic thought considers uncertainty an important issue in economic decisions. While the oldest generation of Austrian writers did not discuss uncertainty, its importance was emphasized by the second and third generation. In particular, Ludwig von Mises, who is considered by many to be the intellectual father of the neo-Austrian school, linked uncertainty directly to his key epistemological concept of human action (Wubben 1995:106-7). "The uncertainty of the future is already implied in the very notion of action. That man acts and that the future is uncertain are ... only two different modes of establishing one thing" (Mises 1949:105).

According to Mises (1949:118), economic agents cope with the uncertain future with the help of: (a) the faculty of understanding and (b) the experience of past events. Thus, for him, while the experience of past events serves as a starting point for a planning process, the only appropriate method of dealing with the uncertainty of future conditions is understanding because it provides knowledge about what is going on in the minds of other people.

Mises presents another interesting argument in the way he links human action, uncertainty and equilibrium. For him, the presence of uncertainty or human action means that there cannot be equilibrium in an economic system. "According to Mises, it is useless to consider an equilibrium, because in such a situation there are no economic problems. Using its praxeological meaning, no one would act any more. And in the absence of feelings of dissatisfaction, there is no Misesian action, and hence the situation of equilibrium. Thus, Mises's conception of human action implies the existence of disequilibrium" (Wubben 1995:116). The above argument implies that in general feelings of dissatisfaction reflect the existence of disequilibrium, uncertainty and the need for action to improve ones current position.

Mises claims that the equilibrating tendency in an economy is brought about by the price system. Understanding, calculations and subsequent human actions are made possible by the existence of the price system which provides the framework for guiding competitors under disequilibrium conditions. Profits are the result of noticing the shortcomings in the markets, and hence they direct the activities of entrepreneurs (Wubben 1995:116). However, within the context of praxeology (the science of human action), profit opportunities only exist when individuals understand them as such. Even if this was not the case, the aim to better one's position does not necessarily bring about a perfect mutual or socially coordinated outcome. This is because Mises' analysis broke up the relationship between the past and the future and introduced uncertainty with regard to the future without introducing any mechanism for learning from past experiences. This is also associated with his use of the concept of praxeology. Other Austrian economists question the empirical meaningfulness of the idea of praxeology and emphasize the necessity for introducing time and place elements into economic theorizing (Wubben 1995:116-117). Hayek, Kirzner and later O'Driscoll and Rizzo have attempted to bridge these gaps in Mises's economic analysis by including the issues of incomplete subjective knowledge, entrepreneurial alertness as well as time and place elements in their theories. The following sections will review their theories.

3.3.2 Hayek, knowledge and uncertainty

Hayek's analysis of the problem of rational economic order focused primarily on incomplete subjective knowledge and its consequences instead of uncertainty as such. He states that the knowledge we use in society never exists in concentrated and integrated form, but only as the dispersed bits of incomplete and frequently contradictory knowledge which all separate individuals possess (Hayek 1945:519). Hayek emphasizes the role of learning and making economics not only a science of human action but also one of human interaction and hence of coordination (Wubben 1995:117).

Similarly, Schmidt (1996:1) argues that if we frame the economy as a system, uncertainty is inseparable from real crisis in economics, whereas if human actions are assumed to constitute the real background of economics, uncertainty immediately emerges as a necessary dimension for every decision maker and, in that sense, uncertainty is a normal component of economic knowledge.

Hayek distinguishes between well organized scientific knowledge and disorganized, but very important, knowledge of particular circumstances of time and place. He argues that when such knowledge of the relevant facts is dispersed among many people, prices can serve as coordinators of separate actions by individuals in the same way that subjective values can help the individual to coordinate the parts of his plan (Hayek 1945:521-526). Hayek emphasizes that prices perform two important functions: (a) they communicate information about the relative scarcity of resources, and (b) they improve coordination of transactors' plans under definite assumptions. These functions serve to solve the equilibrating and the coordination process (Wubben 1995:123-124). Thus, the market system is the best possible method of allocating resources because it is a relatively cheap communication network and works as a system of signals (Hayek 1945:526). It is for this reason that the Austrian school of economic thought provides a very strong defence of the market system.

Hayek brings together problems of incomplete information, disequilibrium and expectation in his analysis of the knowledge problem (Wubben 1995:117). In the same manner as Mises, Hayek dismisses the importance of equilibrium analysis in economics. Hayek (1945:530) argues that while it is a useful preliminary to the study of the main problem, the situation equilibrium analysis describes has no direct relevance to the solution of practical problems and does not deal in any way with the social process.

Hayek shows the importance of the consequences of the assumption of dispersed information, implying that each one of us is guided in our actions by the expectation of an uncertain future where market prices are supposed to communicate scattered information and improve the coordination of plans (Wubben 1995:141-142). However, he does not show how and where the new price information originates or how agents discover unknown future information to keep the economic process moving. Kirzner focuses on these aspects of the Austrian economic theory.

3.3.3 Kirzner, entrepreneurship and uncertainty

Among the recent writers of the Austrian school, Israel Kirzner provides more elaborate discussion on uncertainty. He emphasizes the purposefulness of human action and defines uncertainty as "the essential freedom with which the envisaged future may diverge from the realized future" (Kirzner 1985:58). Kirzner's analysis of uncertainty focuses on two issues: (a) the relationship between uncertainty and past errors, and (b) the relationship between uncertainty and entrepreneurship.

Regarding the link between uncertainty and past errors, Kirzner argues that people do not intentionally fail to achieve the most preferred available outcome: rather, they fail to notice relevant facts. That is, people make errors because they fail to obtain the relevant information before they make decisions. Kirzner calls such an error a genuine error and defines it as "a decision being made in unwitting ignorance of pertinent information" (Kirzner 1979:207). He further argues that "we cannot rule out the possibility that market decisions have been made, not out of deliberately accepted ignorance, but out of genuine error" (Kirzner 1979:207). Thus, for Kirzner, genuine error not only is not inconsistent with fundamental postulates of economics but economics in fact depends on the presence of this kind of error for its most elementary and far reaching theorems

(Kirzner 1979:131). Thus, unlike Mises, for Kirzner, errors do not constitute irrationality. Equilibrium exists if there is no error. However, he argues that past errors might be attributed to the degree of uncertainty prevailing during that period and that present uncertainty and past errors can be considered as two sides of the same coin, called entrepreneurship (Wubben 1995:127).

Another important issue in Kirzner's analysis of uncertainty is the link between uncertainty and entrepreneurship. He links uncertainty to entrepreneurship through the concept of entrepreneurial alertness. He states that "entrepreneurial alertness consists in the ability to notice without search opportunities that have hitherto been overlooked" (Kirzner 1979:148).

In addition, Kirzner (1973:78-79) argues that "entrepreneurial activity undoubtedly involves uncertainty and the bearing of risk. However, it in no way depends on any specific attitude toward uncertainty bearing on the part of decision makers. Even if agents are neutral to uncertainty or fail to perceive profit opportunities, we have to find a place in our theory of the market process for entrepreneurial alertness and for its effect upon the continued availability of perceived opportunities for pure profit." Thus, entrepreneurial activity enables the economic process to move forward in an uncertain world with "the process of spontaneous discovery of information about other participants in the market and enabling the participants to exploit this opportunity in a way that is beneficial to all participants" (Kirzner 1979:148-150). This process of spontaneous discovery links the present uncertainty to the past error.

Thus, the Austrian school of economic thought does not regard uncertainty as a negative factor. Instead, it considers uncertainty as an important factor in propelling the economy forward. It is considered to be important for the very evolution of the economic process. Kirzner's analysis of uncertainty and entrepreneurship bears some relation to Knight's view in the sense that the very existence of entrepreneurship and the market economic system is attributed to the existence of uncertainty about the future. However, while

Knight believes that uncertainty poses threats to the competitive market system if the latter does not develop a proper mechanism to cope with the former, for Austrian economics uncertainty provides opportunity by increasing entrepreneurial alertness that leads to spontaneous discovery of profit opportunities.

For the Austrian school, an economic agent has the mechanism to cope with economic reality, however great uncertainty about the future may be. However, Austrian theories do not believe in agents' complete knowledge about the future. In fact, they assume that in the short run an agent's probabilistic knowledge regarding economic reality is incomplete or even completely lacking. Littlechild (1986:29) argues that, according to the Austrian theories, agents are assumed to know only some components but not others. He or she knows there will be other components but not what they will be; as a result, the agent cannot form a probability judgement as to the likelihood of their occurring. Thus, the basic feature of Austrian theories on uncertainty is ignorance about the future.

3.3.4 Austrian economics, time and ignorance

More recent neo-Austrian writers such as O'Driscoll and Rizzo (1985) incorporate the idea of genuine uncertainty into that of real time and equilibrium. They argue that genuine uncertainty is characterized by (a) inherent unlistability of all possible outcomes resulting from a course of action, i.e. ignorance, and (b) a complete endogeneity of uncertainty (O'Driscoll and Rizzo 1985:71). For these authors, the neoclassical method of modelling uncertainty denies the basic tenet of subjectivism or the autonomy of the human mind by portraying the uncertain future as an objective probability distribution defined over an exhaustive set of events. Genuine uncertainty differs fundamentally from neoclassical uncertainty and is incompatible with the rational expectations equilibrium (O'Driscoll and Rizzo 1985:4). Moreover, real time implies a characterization of uncertainty that is fundamentally different from neoclassical uncertainty (O'Driscoll and Rizzo 1985: 66).

Genuine uncertainty is characterized by ignorance because individual decision-making is partly autonomous and partly dependent on what others choose to do. In a world in which there is autonomous or creative decision-making, the future is not merely unknown, but is unknowable (O'Driscoll and Rizzo 1985:2). This reflects the Austrian interpretation of subjectivism. For these authors, the fundamental aspect of ignorance is the perceived unlistability of all possible outcomes. It is not merely lack of knowledge about which possibility out of a given set will occur; the set itself is unbounded, which reflects the dynamic aspects of subjectivism as opposed to the static subjective probability of neoclassical economics (O'Driscoll and Rizzo 1985: 4).

The second aspect of genuine uncertainty is real time. Genuine uncertainty is endogenous and ineradicable in nature and hence activities directed towards overcoming future uncertainty in a world of real time cannot fully succeed (O'Driscoll and Rizzo 1985:72). O'Driscoll and Rizzo identify two types of time. These are: (a) Newtonian, static time, and (b) real time. The use of Newtonian or static time ignores the processes of economic change and hence leads to static, deterministic theories similar to those of neoclassical economics.

Real or subjective time is not static; it is a dynamically continuous flow of novel experiences (O'Driscoll and Rizzo 1985:59-60). There are two consequences of real time. First, real time is irreversible and, second, real time involves creative evolution, i.e. it produces creative change (O'Driscoll and Rizzo 1985: 62). Therefore, the critical contrast involved in uncertainty is not that between Knight's measurable and unmeasurable uncertainty, nor between Savage's subjective and objective interpretations of probability; instead, it is between purely time dependent and Newtonian types of uncertainty (O'Driscoll and Rizzo 1985:75). Thus, purely time dependent uncertainty is endogenously generated and is inconsistent with the concept of equilibrium, while Newtonian time is static and hence consistent with stochastic models of equilibrium. In the world of time and genuine uncertainty there is continuous endogenous change which contradicts the concept of equilibrium, which assumes the absence of such changes (O'Driscoll and Rizzo 1985:79).

Genuine uncertainty is ineradicable because actions dealing with genuine uncertainty may only transform uncertainty (O'Driscoll and Rizzo 1985:66). A further reduction in uncertainty may also be inhibited by certain characteristics of markets, particularly those which involve independent decision-making. "As long as independent decision-making remains independent, the search for more knowledge does not reduce uncertainty. It merely increases the level of guessing and counter guessing. Endogenous uncertainty can be reduced or eliminated only if actors agree to follow arbitrary conventions which involve exchange of price information among rival firms." (O'Driscoll and Rizzo 1985:234) On this point, O'Driscoll and Rizzo (1985) seem to agree with Keynes's view of falling back on conventions to get by in the face of uncertainty. However, their main argument provides a more radical view regarding uncertainty than the arguments of their predecessors in the Austrian school. Nevertheless, all writers in the neo-Austrian school of economic thought have two important points in common. These are (a) that the existence of uncertainty implies that the economic system cannot be in equilibrium at any point in time, and (b) that agents know only part of what is going on in an economic system. Thus, they are characterized by partial or complete ignorance. For this school of economic thought this is not necessarily bad news for the economic system.

3.4 THE NEW CLASSICAL ECONOMICS AND RATIONAL EXPECTATIONS

Classical economics dominated economic thinking until the emergence of neoclassical economics in the 1870s. Classical economics is based on the assumption that economic agents operate in a world of perfect certainty by presuming that these agents have full knowledge of an unchanging external economic reality that governs all past, present and future economic outcomes. However, their current counterparts reject the perfect certainty models, while still accepting as a universal truth the existence of an unchanging reality that can be fully described by objective conditional probability functions (Davidson 1995:107).

The use of objective probability does not necessarily imply that the economic reality is unchangeable. Instead, it reflects the fact that the information provided by the market system can explain most of the events that occur in an economic environment. In the light of this, Friedman (1968) and Phelps (1967) provided a new explanation regarding the relationship between the rate of inflation and the rate of change in unemployment. This gave rise to a new school called New Classical Economics which adopted the Natural Rate Hypothesis (NRH) and the resulting full information general equilibrium called Natural Rate Equilibrium (NRE) (Van Zijp and Visser 1995:184). The explanations of Friedman (1968) and Phelps (1967) both carried the clear implication that "excess demand" was neither necessary nor sufficient for price or wage inflation. They also showed that any average inflation was consistent, theoretically, with any level of unemployment, contradicting the Philips curve, which was at the centre of all models based on neoclassical synthesis (Lucas 1980:705). This school adopted four crucial assumptions. These are: (a) economic agents are considered to be price takers with perfect price flexibility and continuous market clearing, (b) aggregate real output is formulated as the "Lucas supply function", where deviation from the NRE can only result from unanticipated nominal price changes and hence from expectational errors which are rational in the sense that they differ only randomly from actual realizations of economic variables, (c) the Rational Expectations Hypothesis, and (d) incomplete information about the future (Van Zijp and Visser 1995:185).

The rational expectations (RE) theory (certainty equivalent model) does not claim that the agents in their model possess a complete knowledge of reality. Rational expectations only require that agents obtain reliable probabilistic knowledge from their processing of market signals. Past and present price signals will serve as the basis for obtaining future information. Those who fail to estimate their objective probabilities correctly in the short-run will possess faulty views about risk and uncertainty and will fail in the long-run while those who do will survive in the long-run (Davidson 1995:107).

For rational expectations and other related theories in the short-run, agents have at least partial knowledge of external economic reality. Therefore, for rational expectations, uncertainty is inherent in reality and hence can be measured by stochastic terms in economic relationships. Suppose we specify the following monetary demand equation in line with Dow (1995):

$$M/P = \alpha + \beta y + \delta r + \varepsilon \qquad (3.1)$$

where P is the general price level, M is the money supply, y is real income and r is the rate of interest, and α, β, δ are parameters, and ε is the stochastic term. In the above equation the degree of uncertainty is measured by the variance (σ^2) or the standard deviation σ of the stochastic term ε.

The rational expectations approach, which is part of New Classical business cycle theory, explains economic fluctuations in terms of expectational errors and maintains that these errors are caused by discrepancies between changes in the relative prices and in the general price level. Economic agents must determine how much of a given change in their local prices can be attributed to a change in general price level, and how much of a change in local prices reflects changes in real factors. Thus, agents must form expectations about the difference between their local price level and the general price level (Van Zijp and Visser 1995:187).

Another crucial move of the rational expectations school is to make beliefs rational because, according to this hypothesis, subjective beliefs coincide with objective ones. "In effect the circularity of a purely subjective set of beliefs is broken by the supposition that there are objective beliefs and a rational agent will seek to close the gap between the objective and subjective" (Heap 1989:57).

For RE models, statistical aggregate time series data used in the analysis are "assumed to be random variables and the parameters of a regression model describe the movement of the process that is stochastic and assumed to be stationary. Deviations around expected values have a mean of zero and are assumed to be normally distributed; a necessary condition for assuming that linear regression techniques will yield rationally expected equilibrium values. Thus, in RE models the only kind of expectations that matter are those that correspond to real-valued, explicit data where individuals are believed to adopt the mean value of the observed distribution as their subjective expectation of future values" (Butos 1997:223-4).

Robert Lucas (1980:707), however, argues that the rational expectations hypothesis can make use of the Arrow-Debreu approach to measuring uncertainty. This incorporates uncertainty into a static general equilibrium framework where goods are indexed both by the date of exchange and state of nature, where state of nature refers to different states of occurrence of events. This can be conveniently achieved if the equilibrium is assumed to be determined via a sequence of spot market transactions in which current prices are set, given certain expectations about the future prices. This is because the Arrow-Debreu contingent-claims interpretation of a competitive equilibrium model takes all information to be simultaneously and freely available to all traders. The following section analyzes the Arrow-Debreu theory of competitive equilibrium under uncertainty.

3.5 THE ARROW-DEBREU THEORY OF COMPETITIVE EQUILIBRIUM UNDER UNCERTAINTY

Arrow and Debreu (1954) developed a theory for an economic system where unique competitive equilibrium always exists and where any competitive equilibrium of convex "Arrow-Debreu" model economies is Pareto optimal and, conversely, every Pareto optimal allocation of resources can be realized by a competitive equilibrium (Arrow and Debreu 1954:265). Arrow and Debreu extended the concept of competitive equilibrium to an analysis of consumption, production and exchange decisions under uncertainty involving state and time indexed commodities. Heller et al. (1986:1) stated that "Arrow recognized early on that uncertainty could be incorporated in a general equilibrium setting through the introduction of contingent commodity markets (markets for trading of goods contingent on the state of the world); the presence of such markets, together with other standard assumptions of the competitive model, would assure the first best allocation of risk for the same reasons that the competitive equilibrium is Pareto efficient in the certainty world."

Arrow's discussion focused on the attainment of equilibrium in a competitive securities market where different states of nature and alternative realizations of uncertainty occur at a given point in time. Geanakoplos and Polemarchakis (1986:66) argue that for Arrow, "if for each time period and realization of uncertainty, a pure security exists that yields one unit of 'revenue' or of *numéraire* commodity at that date event wise and zero otherwise, any allocation obtained as competitive equilibrium with a complete market in contingent commodities at the initial period can be alternatively obtained as competitive equilibrium with a complete market in pure securities in the initial period, and subsequent spot markets". Under the state preference approach commodities are identified, not only by their physical location, but also by their location in different states and individuals maximize their utility over bundles of these state contingent commodities. With commodities and preferences defined in relation to states, the theory developed for a world of certainty is formally applicable to a world of uncertainty without modification, and in particular, the concepts of price equilibrium and Pareto optimum, the existence and the equivalence theorems, will be applied under uncertainty as well (Drèze 1987:262).

Moreover, Arrow (1964) and Debreu (1959) observed that uncertainty can be incorporated into "static" general equilibrium theory in exactly the same way that Hicks proposed to incorporate the passage of time, by indexing goods both by the time in which they are going to be exchanged and by the stochastically selected "state of nature" contingent on which the exchange is to occur (Lucas 1980:707). Thus, a contingent claims equilibrium can be interpreted as a description of an economy in which all state contingent prices are determined in advance in the clearing of a single grand futures market, where individual traders may assess the probabilities of the occurrence of the future states of nature, but with prices determined in advance, which excludes the issue of price expectations (Lucas 1980:707). Lucas further argues that the Arrow-Debreu approach in measuring uncertainty rapidly and easily absorbed and clarified a variety of special results in the economics of uncertainty and facilitated their unification and extension, as a result of which it is now in standard use in virtually every applied field of economics, including business cycle theory (Lucas 1980:707).

Among such empirical applications of the Arrow-Debreu state-preference approach, Hirschleifer (1966) is the most common example. He empirically applied the Arrow-Debreu state-preference approach, or time-state-preference approach to use his terminology, to the three states and one commodity conditions in choice processes. There is one current state (c_0) that is certain and two mutually exclusive future states (c_{1a} and c_{1b}) that are uncertain. Each individual who has an endowment of such claims has preference relations ordering combinations of his claims, and has certain opportunities to transform this endowment into alternative combinations. The possible transformation can take the form of financial trading in the market involving exchange with other individuals or productive transformation involving transactions with nature (Hirschleifer 1966:253-254). Using equal state probabilities of $\pi_{1a} = \pi_{1b} = 1/2$, he argues that individuals exhibit the usual convex indifference curve between the two commodities c_{1a} and c_{1b}, implying risk aversion. The convex indifference map is equivalent to a concave Von Neumann-Morgenstern utility function indicating diminishing marginal utility of income and hence risk aversion. This happens not only because of the general principle of diminishing marginal rate of substitution between any ordinary commodities, but also because individuals usually prefer to hold diversified contingencies rather than placing all their eggs in a single basket (Hirshleifer 1966:255).

Thus, although the Arrow-Debreu state-preference approach does not use any probability distribution, the theory incorporated both risky and uncertain choices and allowed extended application of the economics of uncertainty. Moreover, the Arrow-Debreu incorporation of the concept of equilibrium into the analysis of choice processes under uncertainty differs from the Austrian approach that considers equilibrium analysis as irrelevant under uncertainty because, for Austrians, the very existence of the latter indicates that the economy is always in disequilibrium.

3.6 THE COMPETENCE-DIFFICULTY (C-D) GAP APPROACH

In his analysis of the *Origin of Predictable Behaviour*, Heiner (1983, 1985a, 1985b) outlines a general theory for investigating the gap between an agent's decision-making competence and the difficulty of a decision problem which he calls a (C-D) gap (Heiner 1985a:391). This theory argues that the "current pattern of behaviour arises due to decision-making uncertainty resulting from a C-D gap; so that uncertainty becomes the basic source of predictable behaviour" (Heiner 1985a:391). This theory uses two ratios in analyzing the use of available information in decision-making by economic agents. These are (a) a reliability ratio defined as the ratio of the chance of correctly responding under the right circumstances to the chance of mistakenly responding under the wrong circumstances, and (b) a tolerance level defined as the ratio of posterior expected gain to expected loss. The argument is that if the first ratio is greater than the second, the agent will not use even the freely available information.

Accordingly, using the models of behavioural entropy, Heiner (1985b:581) argues that uncertainty arises when the behavioural probabilities r_a is less than and w_a is greater than one (i.e. $r_a < 1$ and $w_a > 0$). The probability ratio r_a is the chance of correctly responding under the right circumstances while w_a is the chance of mistakenly responding under the wrong circumstances. In other words, defining the ratio $r_a/w_a = \rho$ as the set of possible actions that describe an agent's reliability in reacting to information about when to select a specific action, Heiner claims that the term uncertainty arises when $r_a < 1$ and $w_a > 0$, i.e. when the solution is bounded. But $w_a = 0$

implies ρ^{∞} which refers to a special case where agents are perfectly reliable. In this case, agents still face risk from imperfect information but no additional uncertainty is involved. The term uncertainty thus refers to a case where the agent's reliability, ρ is bounded (Heiner 1985b:581).

If, on the other hand, the reliability ratio $r_a/w_a = 1$, an agent does not know whether to choose the wrong or the right action. This is another extreme case of complete ignorance. Between the two extremes of complete ignorance and perfect knowledge are ranges of uncertainty possibilities. Heiner argues that, "The reliability sets describe the range of uncertainty possibilities beginning with $\rho = 1$ at one extreme, and proceeding through intermediate cases where ρ is still bounded, finally limiting on $\rho = \infty$ where only imperfect information remains" (Heiner 1985b:582). Thus, Heiner's argument provides some rational basis for measuring Keynesian uncertainty with some proxy variables.

The reliability condition provides a method of modelling behaviour using separable value functions and behaviour probabilities r_a and w_a which do not require the maximizing postulate of conventional decision theories (Heiner 1985a:393). However, Heiner's approach fails to go beyond the provision of rational explanation and actual assignment of probabilities to measure uncertainty, because the reliability ratios he uses in his analysis are themselves not measurable.

In spite of this, Heiner advanced Keynes's thought by providing "a rational explanation of how induction and analogy are used to guide decisions by judging the associated reliability. Action will be avoided unless reliability exceeds a given threshold. His approach also explains how convention (conservatism) arises as a product of excluding complex considerations, and how variation about convention is possible due to different C-D gaps" (Driver and Moreton 1992:72). This complements Keynes's view that, as uncertainty increases, agents will depend on conventions to get by.

In his analysis of valuation of stock exchanges, Keynes himself emphasizes the importance of convention in investment decisions. He argues that "the essence of this convention lies in assuming that the existing state of affairs will continue indefinitely, although we do not really believe that the existing state of affairs will continue indefinitely. The actual results of an investment over the long-term of years very seldom agree with the initial expectation" (Keynes 1936:152). Thus, for Keynes the conventional method of calculating future expectations is compatible with the prevailing situation only as long as this convention is maintained.

3.7 UNCERTAINTY, THE MARKET SYSTEM AND GOVERNMENT INTERVENTION

In analyzing the impact of uncertainty in economic decisions, the five theories reviewed in this chapter differ in their approaches regarding the role of the market system and the government. Among these, Keynes and post Keynesians strongly believe that the market system cannot meet uncertainty and, therefore, there is a permanent role for the government to take corrective measures in the economic system.

In other words, Keynes's approach to uncertainty implies the possibility that there is a permanent positive role for government in designing policies and institutions that might provide results preferable to those generated by the competitive market system. Moreover, government policies designed to eliminate the deficiency of effective demand produces a Keynes-productive-efficiency gain that supposedly outweighs any allocative inefficiency that might arise as a result of government interferences with relative prices in the market place (Davidson 1995:111).

Coddington (1982:486) argues that Keynes and the post Keynesian school of economic thought is "interested in uncertainty insofar as it helps to show that, under capitalist institutions, the decentralization of production and investment decisions leads inevitably to chaos and waste. What this group needs to establish, then is: (i) that the allegedly wayward and unruly nature of production and investment decisions under capitalist institutions is in some way the product of these institutions; and (ii) that under some alternative institutional arrangements things would be different and better". He further argues that post Keynesians seem to support the idea of having some kind of tripartite institution to oversee investment programmes at the national level, but they do not explain why this may not increase the degree of uncertainty surrounding aggregate investment instead of reducing it (Coddington 1982: 486).

The Keynes and post Keynesian camp continue to argue that there is nothing in a *laissez faire* system that guarantees that there are endogenous forces in the economy to move the system automatically always to full employment. Reliance on the price system alone may mean that Keynes-productive-efficiency and Schumpeterian-"creative destruction"-efficiency are impossible to attain and hence neither can the Keynesian demand policy produce output in excess of the predetermined long run output level nor can entrepreneurs creatively destroy the predetermined production process (Davidson 1995:112).

The other economic theories reviewed in this chapter reflect a belief that the market signals provide complete information in the long run and hence economic laws cannot be changed by any human policy actions. The Austrian school is one such strong adherent to the free market system. This school of economic thought considers uncertainty as an opportunity that propels the economic system forward through deliberate human action, use of incomplete subjective knowledge and spontaneous discovery of profit opportunities by alert entrepreneurs. For Knight and the Austrian school, although uncertainty is unmeasurable and the future is unknown, this does not imply a failure of the market mechanism to correct deficiencies in the economic system. Thus, uncertainty does not necessarily imply the need for government intervention.

Neither new classical nor rational expectations models differ much from Austrian economics on the role of the market system in ensuring results that are superior to those that might be achieved under government intervention. The three theories, i.e. Knight's theory, Austrian economics, and new classical economics and rational expectations can be considered as the advocates of the free market system (*laissez faire*).

This chapter critically reviewed different schools of economic thought regarding their analysis of the impact of uncertainty on economic decisions. However, the chapter did not cover theories on agents' attitudes towards uncertainty (ambiguity). This is an equally important concept but requires separate treatment. A review of the theories of uncertainty (ambiguity) aversion and their implications for optimal investment decisions is presented in the following chapter.

3.8 SUMMARY AND CONCLUSION

Knight and Keynes both emphasized the negative impact of unmeasurable uncertainty on economic decisions. For Knight, the true uncertainty forms the basis of a theory of profit and accounts for the divergence between actual and theoretical competition. For Keynes, in the face of uncertainty, economic agents are dictated by innate urges or animal sprits or fall back on conventional judgement in making the best possible economic decisions.

Austrian economics does not consider uncertainty as a negative factor. In this theory, uncertainty is a necessary condition for the very evolution of the economic processes. That is, entrepreneurship and the market system exist because of the presence of uncertainty. Another important aspect of the Austrian analysis is that, under uncertainty, the economic system cannot be in equilibrium. The very existence of uncertainty implies that the economy is always in disequilibrium. Austrian economics emphasizes two aspects of uncertainty: ignorance and time. Time is in turn categorized as real time or Newtonian (static) time. The latter ignores the processes of economic change and hence leads to static, deterministic theories while real or subjective time is a dynamically continuous flow of novel experiences. In the world of time and genuine uncertainty, there is continuous endogenous change which contradicts the concept of equilibrium that assumes the absence of such changes. Moreover, for the Austrian school, genuine uncertainty is ineradicable because actions dealing with genuine uncertainty may only transform uncertainty.

In contrast to the Austrian theories, new classical economics and rational expectations models adopt the Natural Rate Hypothesis and believe in the existence of the full information general equilibrium called Natural Rate Equilibrium. Another crucial assumption of the rational expectations school is that beliefs must be made rational because, according to this hypothesis, subjective beliefs coincide with objective ones. For new classical economics and rational expectations, uncertainty is inherent in reality and agents are required to obtain reliable probabilistic knowledge by processing the market signals using the past and present price signals which serve as the basis for obtaining future information.

In a similar manner, Arrow and Debreu hypothesized the existence of competitive equilibrium under uncertainty without assigning any probability distributions. They used the concept of competitive equilibrium to analyze consumption, production and exchange decisions under uncertainty involving state and time indexed commodities. They conclude that in the same way that competitive equilibrium is Pareto optimal in the certainty world, general equilibrium based on contingent commodity markets ensures the best allocation uncertainty.

Theories of uncertainty in economic decisions differ regarding the roles of the market system and government in the presence of uncertainty. Keynes and post Keynesians emphasize the negative impact of uncertainty on economic decisions and advocate a permanent role for the government in the economic system. Knight stresses the negative impact of uncertainty on economic decisions but believes that the market mechanism itself can cope with uncertainty. The Austrian school focuses on the role of uncertainty in economic decisions but also believes that uncertainty is a necessary condition for the very existence of entrepreneurship and the free market system and that government intervention is not necessary. The new classical and rational expectations models believe that economic agents cope with uncertainty by fully utilising the available information provided by the market system, which is capable of correcting its own deficiencies, implying that there is no need for government intervention.

CHAPTER FOUR

AMBIGUITY AVERSION AND OPTIMAL INVESTMENT DECISIONS

4.1 INTRODUCTION

While greater attention has been given to the study of risk and risk aversion in economics, little attention has been given to the broader concept of ambiguity aversion. Recently, however, some theories of uncertainty/ambiguity aversion have been developed, primarily in the context of nonadditive probability. This chapter reviews these theories of ambiguity aversion and investigates the link between the latter and optimal investment decisions.

The rest of the chapter is organized as follows: section 4.2 analyzes uncertainty aversion, the independence of preferences and the Ellsberg paradox. Section 4.3 reviews Schmeidler's theory of uncertainty aversion developed in the context of Choquet's expected utility while section 4.4 deals with Epstein's theory of uncertainty aversion. Gul's theory of disappointment aversion is presented in section 4.5. Section 4.6 investigates the essence of aversion to sequential resolution of uncertainty. Section 4.7 presents decision weights on risk and uncertainty. Section 4.8 investigates the link between uncertainty aversion and optimal investment decisions, while section 4.9 presents a summary and conclusion.

4.2 AMBIGUITY AVERSION, INDEPENDENCE OF PREFERENCES AND THE ELLSBERG PARADOX

Ellsberg (1961:650-655) analyses state-indexed preference relations under uncertainty and emphasises the nonadditivity of probability and the weakening of the strong independence assumption of the expected utility theory. When preferences under uncertainty are state indexed, these preference relations can be weakly or strongly independent with respect to states. "The preference relation over state indexed commodities is strongly independent with respect to states if the preference relation among commodities in any group of states is independent of the commodity levels in the other states. Under this assumption, if preferences are described by a continuous, strongly monotonic utility function and if $S > 2$ then the utility function can be written in strongly separable form" (Luenberger 1995:382).

Strong independence is more commonly used in state preference approaches than weak independence, where utility functions are weakly separable. However, Ellsberg showed that strong independence might not hold when individuals are confronted with uncertain choices. The example given by Ellsberg (1961:653-654) illustrates the attitude towards ambiguity and the nonadditive nature of probability. Imagine an urn containing 30 red balls and 60 black and white balls, the latter in unknown proportion. One ball is to be drawn at random from the urn. If we think of four benefit options, b_1 (one red and zero others), b2 (one black and zero others), b_3 (one white, one red and zero black) and, b_4 (one black and one white and zero red). The action b_1 is the bet on red while the benefit b_2 is the bet on black. The action b_1 is objectively determined and is 1/3 since 1/3 of the balls are red. But b_2 is not objectively determined. Thus, if an individual is asked to indicate his preference between b_1 and b_2 he will clearly prefer b_1. This is because the uncertainty of the reward is objectively known in b_1 but not in b_2. Further, if an individual is asked to indicate his preference between b_3 and b_4 he will prefer b_4 because uncertainty is objectively known in b_4.

In some cases, however, individuals may prefer b_2 to b_1 and b_3 to b_4. If these preferences occur, they violate the Sure-thing principle which requires that the ordering of b_1 and b_2 be preserved in b_3 and b_4. In this case, the first pattern of choice implies that the agent prefers red to black and at the same time prefers black to red which contradicts the qualitative probability relationship (i.e. if a is more probable than b, then "not-a" is less probable than "not-b"), since it would indicate that he regarded red as more likely than black, but also "not-red" as more likely than "not-black". Similarly, the second preference pattern would indicate that the agent would be acting as though he regarded "red or white" as less likely than "black or white" although red were more likely than black, and red, white and black were mutually exclusive (Ellsberg 1961:654-655). This violates another qualitative probability relationship which states that if a and b as well as a and c are mutually exclusive (i.e. $a \cap b = b \cap c = 0$), and if a is more probable than b, then the union $(a \cup b)$ is more probable than $(b \cup c)$ (Ellsberg 1961:648). On the basis of such choices, it is impossible to infer even qualitative probabilities for the events in question, specifically for events that include white or black but not both. Thus, choices involving white or black balls involve ambiguity.

Since preferences are separable under strong independence assumptions, under the above three state situations, the preference can be written as follows:

$$U(X_R, X_B, X_W) = U(X_R) + U(X_B) + U(X_W), \tag{4.1}$$

where X_R, X_B and X_W are benefits if red, black and white occurs respectively (Luenberger 1995:383). The preference relation $b_1 \succ b_2$ means that $U_R(1) > U_B(1)$ and the preference relation $b_4 \succ b_3$ means that $U_B(1) + U_W(1) > U_R(1) + U_W(1)$. If we cancel $U_W(1)$ from both sides, we obtain the inequality $U_B(1) > U_R(1)$. This contradicts the first inequality indicating that preferences under the above case are not strongly independent. This is known as the Ellsberg paradox which challenges Savage's substitution principle and additivity axiom (Fishburn 1998:190). Sarin and Wakker (1998:223) have argued, though, that Ellsberg's concerns about the inadequacy of probability have been satisfactorily addressed by the nonadditive probability model developed by Schmeidler (1989) and Gilboa (1987). Gilboa (1987:67) states that no additive probability measure may explain the preference pattern characterized by the Ellsberg paradox; it can only be explained using the nonadditive probability measures.

Heath and Tversky (1991:6) argue that "Ellsberg's example, and most of the subsequent experimental research on the response to ambiguity and vagueness, were confined to chance processes, such as drawing a ball from a box. The potential significance of ambiguity, however, stems from its relevance to the evaluation of evidence in the real world". They further argue that, at the same time, the current literature in this area did not provide any answer as to whether ambiguity (uncertainty) aversion is limited to games of chance and stated probabilities or whether it also holds for judgemental probabilities.

Moreover, the Ellsberg paradox does not imply that the independence axiom of the expected utility theory should be discarded. Some theories argue that models that use a strict distinction between risk and uncertainty, as in the case of Ellsberg (1961), imply maximizing the expected utility with a nonadditive probability with some weakening of the independence axiom. For these theories, the nonadditive probability reflects both the presence of uncertainty and the agent's aversion to it.

4.3 THE CHOQUET EXPECTED UTILITY AND SCHMEIDLER'S THEORY OF UNCERTAINTY AVERSION

Schmeidler's (1989) theory of uncertainty aversion begins with a critical assessment of certain aspects of prior probability. He argues that the probability attached to an uncertain event does not reflect the heuristic amount of information that led to the assignment of that probability. For instance, when information on the occurrence of two events is symmetric they are assigned equal prior probabilities and if the events are complementary the probabilities will be ½, independently of whether the symmetric information is meagre or abundant (Schmeidler 1989:571). However, while Schmeidler accepts the rule that symmetric information with respect to the occurrence of events results in equal probabilities, he does not agree with the rule that states that if the space is partitioned into k symmetric (equiprobable) events, then the probability of each event is 1/k. That is, he does not agree with additivity of probabilities. He argues that if each of the symmetric and complementary uncertain events is assigned the index 3/7, the number 1/7, 1/7 = 1-(3/7 + 3/7) would indicate the decision maker's confidence in the probability assessment. Thus, allowing nonadditive probabilities enables transmission or recording of information that additive probabilities cannot represent (Schmeidler 1989:572). Thus, according to Schmeidler, not all decision makers can represent their beliefs with respect to the occurrence of uncertain events through an additive probability.

Furthermore, Schmeidler (1989:573-574) believes that, in cases where preferences between acts do not admit additive subjective probability (as in the case of, for instance, the Ellsberg paradox) it is possible to define in some consistent way a unique nonadditive subjective probability as well as to define the expected utility maximization criterion for the nonadditive case. Schmeidler's (1989) model rationalizes nonadditive (personal) probabilities and admits the computation of expected utility with respect to these probabilities. The model formally extends the additive model and makes the expected utility criterion applicable to cases where additive expected utility is not applicable, thereby translating the concepts of risk aversion, risk premium and certainty equivalence used for the analysis of decision under risk into uncertainty aversion, uncertainty premium and risk equivalence in the case of decisions under uncertainty (Schmeidler 1989:574).

By replacing the common axioms of independence with comonotonic (common monotonic) independence, and the finite additive probability p with nonadditive probability v in the context of the Choquet-expected utility (CEU), Schmeidler (1989) provides the first definition of uncertainty aversion. Uncertainty aversion in the sense of Schmeidler is equivalent to the convexity of capacity, v, (or nonadditive probability) in the Choquet utility function (Epstein 1999:581). Equations (4.2) and (4.3) below define this function. In Choquet's expected utility, a "capacity" v is used instead of the additive probability measure p used in subjective expected utility measures. According to Sarin and Wakker (1998:227), it is assumed that v assigns value 0 to the impossible event and value 1 to the universal event, S, and that $A \supset B$ implies $v(A) \geq v(B)$. Accordingly, the CEU value of an act $(A_1, X_1, \ldots, A_n, X_n)$, with $X_1 \geq, \ldots, \geq X_n$ is given by

$$\sum_{i=1}^{n} \pi_i U(x_i) \qquad (4.2)$$

where U is the utility function as in subjective expected utility and π_i denotes decision weights defined by

$$\pi_i = v(A_1 \cup \ldots A_i) - v(A_1 \cup \ldots A_{i-1}) \qquad (4.3)$$

Unlike the theory of risk aversion where aversion to risk is associated with the concavity of utility functions, the theory of uncertainty aversion does not link this phenomenon to the concavity of the utility function. Instead, it is the convexity of the capacity in CEU that explains aversion to uncertainty. That is, given the Choquet expected utility in (4.2) and (4.3), Schmeidler's uncertainty aversion is equivalent to the property:

$$v(A \cup B) + v(A \cap B) \geq v(A) + v(B) \qquad (4.4)$$

for all measurable events A and B.

According to Schmeidler, a binary relation $= \succ$ (indicating not strict preference) on L is said to reveal uncertainty aversion if for any three acts, f, g and h in L and any α in [0,1], if $f = \succ h$ and $g = \succ h$, then $\alpha f + (1-\alpha)g = \succ h$. Similarly, if $f = \succ g$, then $\alpha f + (1-\alpha)g = \succ g$. The definition of strict uncertainty aversion requires a strict preference $\succ$. However, some restrictions have to be imposed on f and g for the above relations to hold; and one such restriction is that f and g are not comonotonic (Schmeidler 1989:582).

The intuitive definition of uncertainty aversion, according to Schmeidler (1989:582), is that "smoothing" or averaging utility distributions results in the decision maker being better off. Another way of defining this is to say that substituting objective mixing for subjective mixing makes the decision maker better off. Schmeidler's definition of uncertainty aversion accommodates the widely studied theories of preferences such as the Ellsberg type behaviour. However, other authors such as Epstein (1999) do not agree with this definition of uncertainty aversion as it appears to be less appealing and less useful for applications, such as in the Savage domain, that are more widely used in descriptive modelling. Consequently, Epstein (1999) provides an alternative definition of uncertainty aversion. The following section reviews this definition of uncertainty aversion.

4.4 THE EPSTEIN THEORY OF UNCERTAINTY AVERSION

Epstein (1999) objects to the widely used behavioural interpretation of convexity of the capacity in Choquet's expected utility as uncertainty aversion. He argues that "the notion of uncertainty aversion leads to concern with 'local probabilistic beliefs' implicit in an arbitrary preference order or utility function and these beliefs represent the decision maker's underlying 'mean' or 'ambiguous free' likelihood assessment for events which may not be unique but can be so if utility is eventwise differentiable" (Epstein 1999:580). Thus, using eventwise differentiability of utility, Epstein provides an alternative definition of uncertainty aversion that is suitable for applications in the Savage domain of acts.

Epstein (1999:583) argues that the theory of risk aversion can provide some perspective on analysis of uncertainty aversion, and that if a distinction between risk and uncertainty is to be made then the theory of risk aversion must be modified. Accordingly, he emphasizes the need to extend the notion of comparative and absolute risk aversion to comparative and absolute uncertainty aversion.

By making a distinction between risk as an unambiguous event where probability can be assigned to events, and uncertainty as an ambiguous event where it is not possible to assign precise probabilities, Epstein (1999:584) defines uncertainty aversion in relation to an individual's preference for unambiguous events. He begins his model of uncertainty aversion by proving an *a priori* specification of the "absence of uncertainty" which takes the form of an exogenous family $\mathcal{A} \subset \Sigma$ of "unambiguous" events and with an intuitive requirement for $\mathcal{A}$ which also contains S being

$A \in \mathcal{A}$ implies that $A^c \in \mathcal{A}$; where A^c is A complement

$A_1, A_2 \in \mathcal{A}$ and $A_1 \cap A_2 = \emptyset$ imply that $A_1 \cup A_2 \in \mathcal{A}$.

Epstein (1999:584) argues that, intuitively, if an event being unambiguous means that it can be assigned a probability by the decision maker, the sum of the individual probabilities is naturally assigned to a disjoint union, while the complementary probability is naturally assigned to the complementary event. Moreover, intuition does not require that $\mathcal{A}$ be closed with respect to nondisjoint unions or intersections or that $\mathcal{A}$ be an algebra. By denoting f^{ua} the set of $\mathcal{A}$-measurable acts, also called unambiguous acts, and given two orderings $\succeq^2$ *is more uncertainty averse than* $\succeq^1$ if for every unambiguous act h and every act e in f

$$h \succeq^1 (\succ^1) e \text{ implies } h \succeq^2 (\succ^2) e. \tag{4.5}$$

Thus (4.5) represents Epstein's (1999) definition of comparative uncertainty aversion. He also argues that there is no loss of generality in assuming that the acts h and e deliver identical outcomes. The difference between the two acts lies in the nature of the events, where these outcomes are delivered. For h, the typical outcome x is delivered in the unambiguous event $h^{-1}(x)$, while it occurs in an ambiguous event given e. In plain language, Epstein's definition of comparative uncertainty aversion implies that decision makers prefer risk (events with certain probability distributions) to uncertainty which entails an ambiguous event.

Thus, whenever the greater ambiguity inherent in e leads $\succeq^1$ to prefer h, the more ambiguity averse $\succeq^2$ also prefers h, an unambiguous act, to e, an act with ambiguity. This definition of comparative uncertainty aversion is parallel to the definition of comparative risk aversion, given two orderings of preferences, but it does not imply the concavity or convexity of utility function considered in the theory of risk aversion. Moreover, "there appears to be no logical connection between uncertainty aversion and convexity. Convexity does not imply uncertainty aversion unless additional conditions are imposed. Furthermore, convexity is not necessary even for the stricter notion 'more uncertainty averse than some expected utility order' that seems close to Schmeidler's notion" (Epstein 1999:589).

4.5 GUL'S THEORY OF DISAPPOINTMENT AVERSION

Gul (1991:667) claims that his theory of disappointment aversion is an alternative model of decision-making under uncertainty that not only includes expected utility theory as a special case but is also consistent with the Allais paradox. According to Gul's formulation, the preference of a disappointment averse individual can be represented by U(x), β where U(x) is a conventional utility function, describing the utility of consuming good x, with the usual properties of risk aversion described by the concavity of the utility function, i.e. $U' > 0$ and $U'' < 0$. Moreover, U is considered to be unique up to an affine transformation (i.e. linear transformation between two vectors followed by translation), and β is unique. When $\beta = 0$, this corresponds to the case of the expected utility theory where U is the Von Neumann-Morgenstern utility function. Preferences with $\beta \geq 0$ are described as disappointment averse while strict disappointment aversion is described by $\beta > 0$ (Gul 1991:667-668).

By using four basic assumptions regarding preferences, i.e. the preferences relation is complete and transitive; it is continuous; it exhibits weak independence; and it is symmetrical, Gul (1991:672-673) argues that there exists a preference relation β $\in (-1, \infty)$ such that

$$\tau(\alpha) = \frac{\alpha}{1+(1-\alpha)\beta} \qquad \text{for all } \alpha \in [0,1] \qquad (4.6)$$

and shows that the expected utility corresponds to the special case $\tau(\alpha) = \alpha$; that is, $\beta = 0$ and that if $\beta > 0$, then $\tau(\alpha) < \alpha$ for all $\alpha \in [0,1]$ and $\tau(\alpha)$ is convex. On the other hand, if $-1 < \beta < 0$, then $\tau(\alpha) > \alpha$ for all $\alpha \in [0,1]$ and $\tau(\alpha)$. Thus, the preference relation $\sim\succ$ reflects disappointment aversion if $\beta \geq 0$ and preference relation $\sim\succ$ is elation loving if $\beta \in (-1,0)$ and that unlike risk aversion, disappointment aversion is, by definition, a global property (Gul 1991:673).

The theory of disappointment aversion is parsimonious in the sense that it is only one variable (β) richer than the expected utility and yet it is believed by its author to be able to include the expected utility as a special case and be consistent with models that predict violations of the main assumption of the latter theory. This claim is not empirically substantiated as there are to our knowledge no empirical tests of the theory of disappointment aversion. However, the basic concept of this theory, disappointment aversion, is used by Palacios-Huerta in his theory of aversion to sequential resolution of uncertainty. We shall review this theory in the next section.

4.6 AVERSION TO SEQUENTIAL RESOLUTION OF UNCERTAINTY

In a dynamic choice setting, when individuals face uncertain prospects, they often have a preference for the time of the resolution of uncertainty. A growing body of both theoretical and empirical literature focuses on the role and implications of preferences for the resolution of uncertainty. Palacios-Huerta (1999:249) states that the implications of these preferences range from different areas in asset pricing and finance, such as models of security prices and asset trading, to models of demand for information (in terms of both quantity and quality), theories of price formation, the formation of preferences, models of information aggregation, investment, and several other types of individual and social choices under uncertainty.

Palacios-Huerta's (1999) formulation of sequential resolution of uncertainty differs from the previous studies in that it focuses not only on the *timing* but also on the *form of* *the timing* of the resolution of uncertainty and as such its main focus is the aversion attached to the sequential resolution of uncertainty instead of a complete resolution all at one time. In other words, this theory is concerned with an individual's preferences for the resolution of uncertainty all at one time rather than sequentially (Palacios-Huerta 1999:150). The intuition for the aversion to the sequential resolution of uncertainty is that when a disappointment averse individual is taken through the sequential process of the resolution of uncertainty, some disappointment may occur each time the event takes place sequentially, and this leads to the loss of more utility compared to a situation where the disappointment averse individual is able to resolve the uncertainty all at one time (Palacios-Huerta 1999:254). In other words, a disappointment averse individual strictly prefers the resolution of uncertainty all at one time rather than sequentially.

To explain this behaviour, Palacios-Huerta uses the example of a Brazilian professor of economics watching the penalty shootout of the football World Cup final between Brazil and Italy in 1994. He preferred to switch off his TV set during the shootout to resolve his uncertainty all at once instead of watching the shootout process. He further argues that this behaviour is related to the preference pattern described by the Allais paradox and that it can be explained by theories such as Gul's (1991) theory of disappointment aversion.

People do not always prefer to resolve uncertainty all at on time, however. This is because the utilities attached to the sequential processes and the final outcome are different and in some cases the utility gain of the sequential process can be greater to the individual than the utility loss associated with not being able to resolve the uncertainty all at one time. Such situations refer, for example, to processes where individuals prefer to participate in sequentially uncertain events and by so doing wish to influence the final outcomes in their favour. These two sources of utility mean that it will not be straightforward to document empirically the extent of aversion to the sequential resolution of uncertainty (Palacios-Huerta 1999:257).

Palacios-Huerta (1999:259) further emphasizes that his theory of sequential resolution of uncertainty can contribute to the literature on endogenous timing of actions and the differential between clustering of decisions and herding, in the sense that individuals decide to take action after all the uncertainties have been resolved. This may be true in some cases where, for instance, economic decisions involve investment in irreversible capital stock and, in the face of uncertainty about the future return, investors prefer to delay their investment until some future time when they have full information. Palacious-Huerta's hypothesis implies that if, for instance, a firm faces a series of macroeconomic uncertainties such as uncertainty about the interest rate, exchange rate or inflation, the firm prefers to see all three uncertainties resolved at once instead of one by one before it commits itself to investing in irreversible capital stock. This particular issue will be discussed in detail in chapter 5 of the present study. However, the empirical measure of the extent of the aversion to such sequential resolution of uncertainty is complex because of the implied psychological degree of disappointment associated with various sequences of the occurrence of uncertain events.

4.7 DECISION WEIGHTS ON RISK AND UNCERTAINTY

In line with Knight (1921), Keynes (1921) and Ellsberg (1961), Tversky and Fox (2000:116) distinguish uncertain prospects based on the degree to which the uncertainty can be quantified. At one extreme, when uncertainty is characterized by a known probability distribution, this refers to a situation of risk. At the other extreme, when decision makers are unable to assign numerical probability to prospects, this refers to a situation of ignorance. Most decisions under uncertainty lie between these two extremes. Thus, decision under uncertainty can be segmented into the three elements of risk, uncertainty and ignorance (or ambiguity, or vagueness).

The studies of choice between risky prospects, principally by the prospect theory, have suggested the nonlinear transformation of the probability scale that overweights lower probabilities and underweights moderate and higher probabilities. Tversky and Fox

(2000:93) extend the above notion from risk to uncertainty by using the principle of bounded subadditivity and show that an event has a greater impact when it turns impossibility into possibility, or possibility into certainty, than when it makes a possibility more or less likely. This is similar to Epstein's definition of uncertainty aversion where agents are believed to prefer risk (uncertainty with known probability) to uncertainty (ambiguity or ignorance).

The expected utility theory which was developed to explain attitudes towards risk, in terms of risk aversion and risk seeking, defined the former as a preference for a sure outcome over a prospect with equal or greater expected value. Risk seeking, on the other hand, refers to the situation where a prospect with equal or greater expected value is preferred to a sure outcome. However, in expected utility theory individuals are commonly assumed to be risk averse – the behaviour explained by the concave utility function.

Tversky and Fox (2000:116) argue that there is subadditivity for both risk and uncertainty and that this effect is more pronounced for uncertainty than for risk. Moreover, they further suggest that subadditivity is amplified by vagueness or ambiguity, implying that studies of decision under risk can underestimate the degree of subadditivity that characterizes decisions involving real world uncertainty. Therefore, subadditivity emerges as a unifying principle of choice that occurs to varying degrees in decision under risk, uncertainty and ignorance.

However, empirical tests have shown that the nonlinear transformation of probability that satisfies bounded subadditivity, as well as the fourfold characterisation of risk and the implied reflection effects are sometimes observed but do not reflect universal behaviour of economic agents.

4.8 AMBIGUITY AVERSION AND OPTIMAL INVESTMENT DECISIONS

Dow and Werlang (1992) provide an alternative definition of increase in perceived uncertainty and its impact on investment decisions. Using a version of the expected utility models that recognises the distinction between risk and uncertainty in line with Frank Knight (1921), they argue that such models entail maximizing expected utility with a nonadditive probability measure when the independence axiom is weakened (Dow and Werlang 1992:197). In other words, the model of expected utility with nonadditive probability reflects uncertainty aversion when the independence axiom is weakened, instead of being totally eliminated. Nonadditive probability measures refer to a situation where the probability that either of two mutually exclusive events occurring will not add up to one as opposed to conventional belief. Moreover, Schmeidler (1989:573) argues that preferences between acts do not necessarily admit additive subjective probabilities. The same holds true for objective probabilities. Thus, Dow and Werlang (1992:197) argue that if the sum of the probabilities is less than 1, then expected utility calculations using this probability measure will reflect uncertainty aversion as well as (possibly) risk aversion and that it should be stressed that the probabilities, together with the utility function, provide a representation of behaviour and are not objective probabilities.

Suppose the probability that two mutually exclusive outcomes are p_1 and p_2. Dow and Werlang (1992:199) argue that if the sum $p_1 + p_2 < 1$, the agent's decision reflects uncertainty aversion. But if both outcomes are equally likely, i.e. $p_1 = p_2 = 1/2$, this shows that the agent believes that the two outcomes are equally likely and he is not averse to uncertainty. It does not mean that the agent knows the risk with certainty. They further argue that nonadditive prior probability represents both the presence of uncertainty and the agent's aversion to it.

The above arguments can be directly applied to the investment decisions of economic agents. An agent who starts from the position of certainty will invest in an asset if and only if the expected value of the asset exceeds the price. The amount of the asset purchased will depend on the agent's attitude to risk. On the other hand, if the expected value of the asset is lower than the price, the agent will sell the asset. "The highest price at which the agent will buy the asset is the expected value of the asset under the nonadditive probability measure, while the lowest price at which the agent would sell the asset is the expected value of the asset. These prices depend only on the beliefs and aversion to uncertainty incorporated in the agent's prior probability, and not on attitude to risk. This result is the nonadditive analog of the local risk neutrality result" (Dow and Werlang 1992:198).

Furthermore, Dow and Werlang (1992:198-201) argue that with the nonadditive probability measure, the expectation of the random variable is less than the negative of the expectation of the negative of the random variable, i.e. $E(X) < -E(-X)$. Suppose the investor is faced with the problem of choosing the sum of money N he will invest in an asset. The present value of one unit of the asset next period is a random amount X with nonadditive probability distribution P. The demand for the asset is a function of the price. Under these circumstances, a risk averse or risk neutral investor with certain wealth W, who is faced with an asset which yields X per unit, whose price is $p > 0$ per unit, will purchase or invest in the asset if $p < E(X)$ and only if $p \leq E(X)$. He will sell or disinvest the asset if $p > -E(-X)$ and only if $p \geq -E(-X)$ (Dow and Werlang 1992:2002).

The Dow-Werlang analysis is based on Schmeidler (1989) and Gilboa's (1987) representation of choice problems under uncertainty using nonadditive probabilities. However, Dow and Werlang differ from the latter authors in that they use the nonadditive probability measures and the resulting measure of uncertainty aversion in the context of the maximization of expected utility.

Uncertainty and its overall impact on investment decisions of firms is also emphasized by the theory of investment under uncertainty. In addition to agents' attitudes towards uncertainty, the theory of investment under uncertainty includes investment irreversibility and financial constraint as important factors that determine the optimal level of firm investment. This theory is reviewed in the next chapter.

4.9 SUMMARY AND CONCLUSION

Most theories of uncertainty aversion are based primarily on the notion of nonadditivity of probability. When preferences under uncertainty are state indexed preferences may be weakly or strongly independent with respect to the state. The Ellsberg paradox showed that preferences are not always state independent thereby challenging the additivity of probability and the Sure-thing principle.

Schmeidler pioneered the definition of uncertainty aversion. By using the concept of nonadditive probability in the context of the Choquet expected utility, Schmeidler provided the first definition of uncertainty aversion as the convexity of capacity, or nonadditive probability in the Choquet utility function.

However, more recent authors, such as Epstein, have questioned the link between uncertainty aversion and convexity of the capacity in the Choquet utility function and have provided an alternative definition of uncertainty or ambiguity aversion based on eventwise differentiability of utility. Epstein developed a definition of comparative uncertainty aversion which indicates agents' preferences for unambiguous events rather than ambiguous ones. According to him, agents' attitudes towards uncertainty are measured by their preference for unambiguous events to which prior probabilities can be assigned, as opposed to events which are ambiguous.

Based on the theory of disappointment aversion, Palacios-Huerta developed the theory of aversion to sequential resolution of uncertainty. This theory focuses on both the timing and the nature of the timing of the resolution of uncertainty and concludes that individuals are averse to the sequential resolution of uncertainty. The economic implication of this is that if a firm faces a series of macroeconomic uncertainties, such as uncertainty about the interest rate, exchange rate and inflation, then it will not be willing to invest in irreversible capital stock until all of these uncertainties have been fully resolved. The aversion to the sequential resolution of uncertainty and a preference for complete resolution of uncertainty all at once can be a useful tool in economic analysis. However, the empirical measure of the extent of the aversion to such sequential resolution of uncertainty is complex because of the implied psychological degree of disappointment associated with various sequences of the occurrence of uncertain events.

The existence of nonadditive probability (the probability that does not add up to 1) does not undermine the expected utility theory. Instead, nonadditivity of probability under the expected utility theory implies the prevalence of uncertainty aversion and this has clear implications for optimal investment decisions of firms. Empirical studies have shown that individual beliefs and aversion to uncertainty (presence of nonadditive probability under the expected utility theory) affect the expected asset prices and subsequent decisions to sell or purchase assets.

CHAPTER FIVE

UNCERTAINTY, IRREVERSIBLE INVESTMENT AND FINANCE

5.1 INTRODUCTION

The theory of investment under uncertainty, developed in the period from the early 1980s and presented more formally by Dixit and Pindyck (1994), reemphasized the importance of uncertainty in the investment decisions of firms. According to this new theory, the optimal decisions of investors are affected by the interaction of investment irreversibility, the decision to wait for more information and the presence of uncertainty about the future. The most recent literature has included firms' financing problems as an additional factor that affects investment decisions of firms that plan to undertake irreversible investments under uncertainty.

The present chapter reviews the theory of investment under uncertainty in view of establishing the link between these factors and the investment and lending decisions of the investor-lender firm. The chapter is organized as follows: section 5.2 investigates the link between uncertainty and the option of waiting to obtain more information about the future. Uncertainty and the problem of investment irreversibility is the concern of section 5.3. Section 5.4 assesses the link between uncertainty, investment irreversibility and finance while section 5.5 presents a summary and conclusion.

5.2 UNCERTAINTY AND THE OPTION VALUE OF WAITING TO INVEST

The theory of irreversible investment under uncertainty states that when investors face uncertainty about the future returns from their planned investment, they have an option to wait, that is, to defer the investment. Thus, the theory of irreversible investment under uncertainty is sometimes referred to as the new investment theory of real options.

Trigeorgis (1999:6) argues that the option to wait is particularly valuable in resource extraction industries, farming, paper manufacturing and real estate development due to high uncertainties and long investment horizons. There are various real options but, for the theory of investment under uncertainty, the most important real option is the option to defer investment.

According to Dixit and Pindyck (1994:3-4), the option to defer investment, partial or complete irreversibility of most investment expenditures and uncertainty over the future returns from that investment are the three most important factors that determine the optimal decision of an investor. These conditions are the central feature of the new theory, which departs from the traditional neoclassical theory that emphasizes the importance of the net present value (NPV) rule in investment decisions.

The NPV rule states that if the present value of the expected stream of benefits from a given investment is greater than the present value of the stream of the expenses of the investment then the investor should go ahead with the investment. However, this rule is questioned by the new theory because of the presence of a sunk cost, that is, the investment expenditure which may not be partially or fully recovered. In the presence of a sunk cost, once a decision is made to invest, the option to wait is lost and this lost option value is an opportunity cost that must be included in the calculation of the cost of the investment. Thus, "the NPV rule 'invest when the value of a unit of capital is at least as large as its purchase and installation cost' must be modified. The value of the unit must exceed the purchase and installation cost, by an amount equal to keeping the investment option alive" (Dixit and Pindyck 1994:6).

Trigeorgis (1999:3-4) argues, however, that although the NPV and other discounted cash flow (DCF) approaches to capital budgeting do not provide adequate information, given the real world situations of change, uncertainty and competitive interactions, it does not imply that these traditional approaches should be discarded. Instead, managerial adaptability is required to adjust to these real world situations. This calls for an expanded NPV rule that combines the traditional (static or passive) NPV of expected cash flows with the option value of operating and strategic adaptability. This expanded (strategic) NPV = static (passive) NPV of expected cash flows + value of options from active management.

Such real options can be properly valued using contingent claims analysis (CCA) within a backward risk neutral valuation process using risk neutral probabilities, p, and a riskless discount rate, or interest rate, r. In this case, risk refers to a measurable uncertainty. According to Trigeorgis (1999:12), in such a risk neutral world, the current value of the project (or equity holders' claim), *E*, is obtained as:

$$E = [pE^{+} + (1-p)E^{-}]/(1+r), \tag{5.1}$$

$$P = [(1+r)S - S^{-})/(S^{+} - S^{-}), \tag{5.2}$$

where S^{+}, S^{-} are the highest and lowest prices of the output of the next period. The probability, p, is estimated from the price dynamics. The *p* value can be used to determine the certainty equivalent values (expected cash flows) which can be properly discounted at riskfree rate.

Accordingly, the option to delay (defer) investment can be calculated as the difference between expanded NPV and passive NPV, i.e. option to wait (defer) = expanded NPV - passive NPV. Suppose the gross value of a project is $105 million dollars and its NPV is -5, implying that the investment should not go ahead, while the project's total value (i.e. the *expanded* NPV, which includes the value of the option to defer (or real options) obtained using contingent claims analysis (CCA)) is 26, then the option to wait (defer) = 26 - (-5) = 31. This is almost one-third of the gross value of the investment. The above calculation is based on Trigeorgis (1999:12-13). The option value of waiting caused by uncertainty about the future prices or values of the project is found to be high and this could be higher in the case of investment irreversibility.

The implication of the increase in the option values is that there will be further delay in investment decisions. In other words, as the option to wait increases, investment will be delayed. Thus, in the presence of investment irreversibility and uncertainty, investors delay their investment because they have the option to wait and obtain more information about the future. Bernanke (1983:86) argues that individual investment projects are economically irreversible in the sense that once constructed, they cannot be undone or changed to radically different projects without a high cost. New information for assessing long run project returns arrives over time, so that, by waiting, the potential investor can improve his chance of making a correct decision. However, when the option value of waiting is high, then it can lead to a decrease in investment by delaying the timing of the investment decisions. If the investors wait until all the uncertainties confronting them are fully resolved before they make any investment, this implies aversion to sequential resolution of uncertainty discussed in the previous chapter of the present study.

5.3 UNCERTAINTY AND INVESTMENT IRREVERSIBILITY

The new theory argues that most investment expenditures are sunk costs in that the firms or even industries cannot disinvest without cost should their investment prove to be unprofitable. In the presence of uncertainty, therefore, firms should postpone their investment decisions in order to get more information in the future, since they cannot fully recover their capital once they have made an irreversible investment. As uncertainty increases, then, firms become more reluctant to invest and hence investment decreases. In other words, as stated earlier, as uncertainty increases the option value of waiting becomes high leading to a fall in investment. Therefore, in the presence of irreversibility, uncertainty can be a major deterrent to investment.

When investments are irreversible and can be delayed, they become very sensitive to future returns. Thus, if the goal of macroeconomic policy is to stimulate investment in the short to medium term, stability and credibility may be much more important than particular levels of tax rates or interest rates (Pindyck and Solimano 1993:1-2). Moreover, McDonald and Siegel (1986:714) show that in the presence of irreversibility friction a small rise in uncertainty can lead to a high level of the return required to trigger investment. In other words, the presence of investment irreversibility makes investment adjustment costs asymmetric and larger for downward than upward adjustment. Under appropriate conditions, this creates a range of inaction in that investment takes place only when expected profitability is greater than a certain threshold (Serven 1998:2). The central theme of the new theory is that uncertainty affects the investment decisions of a firm through its impact on this trigger or threshold rate of return.

According to the new theory of investment, when investments are irreversible firms face two choices. These are either spending their resources with no possibility of reversing their decisions, or exercising an option to wait for more information that implies a period of inaction. One way to analyze this problem is by using the basic model of MacDonald and Siegel (1986) which presents irreversible investment as a sunk cost, *I*, needed to undertake an investment project whose value, *V*, follows a geometric Brownian motion:

$$dV = \alpha V dt + \sigma V dz \qquad (5.3)$$

$$dz = \varepsilon_t \sqrt{dt}\,\mathrm{dz} \qquad (5.4)$$

where dz is the increment of the Wiener process, α is a drift parameter and σ is the variance parameter or the standard deviation of the change in value of the project for a given period.

Given the properties of the Brownian motion, percentage changes in V $(\Delta V/V)$ or change in the natural logarithm of V $(\ln V)$ are normally distributed with mean $(\alpha - 1/2\sigma^2)dt$ and variance $\sigma^2 t$, while absolute changes in V, ΔV, are log normally distributed. If V_t follows a geometric Brownian motion then $F(V) = Log\ V$ becomes the following

Brownian motion: $dF = (\alpha - 1/2\sigma^2)dt + \sigma dz$. Over a specified time period the change in the logarithm of V is normally distributed with mean $(\alpha - 1/2\sigma^2)t$ and variance $\sigma^2 t$ (Dixit and Pindyck 1994:71).The application of a geometric Brownian motion to the value of the investment project implies that the current value of the project is known, while future values are always uncertain and this uncertainty increases linearly with the time horizon.

The problem the firm faces is to maximize the expected present value of the investment opportunity. Pindyck and Solimano (1993:6) denote this investment opportunity as $F(V)$. Since the benefit from the investment opportunity at time t is $V_t - I$, the firm needs to maximize

$$F(V) = \max E[(V_t - I)e^{-\eta t}] \qquad (5.5)$$

where E is the expectation operator, t is unknown future time of investment and η is the discount rate. Thus, the firm must maximize this value subject to the preceding constraint of the value equation. Moreover, it is assumed that the trend parameter α is less than the discount rate, η, so that the difference, $\eta - \alpha = \omega > 0$.

Thus, according to the theory of irreversible investment under uncertainty, the optimal investment decision depends on the critical value V^* and hence firms should invest if $V \geq V^*$. That is, firms will not invest until the expected return from the project is at least as large as the threshold return and in the presence of uncertainty this threshold can be many times larger than the cost of investment, I. If the firm invests only when the expected future value of the project reaches the threshold V^*, then the value of the investment opportunity becomes:

$$F(V) = bV^{\rho} \qquad (5.6)$$

Where b is a constant and ρ is a parameter whose value depends on the level of uncertainty σ and the discount rate η.

In the above equation ρ is the root of a quadratic equation and hence is given by

$$\rho = \frac{1}{2} - \frac{(\eta - \omega)}{\sigma^2} + \sqrt{[(\eta - \frac{\omega}{\sigma^2} - \frac{1}{2})]^2} + \sqrt{\frac{2\eta}{\sigma^2}} > 1 \qquad (5.7)$$

The threshold value V^* and the constant b are given by:

$$V^* = \frac{\rho}{(\rho - 1)I} \qquad (5.8)$$

and

$$b = \frac{(V^* - I)}{(V^*)^\rho} = \frac{(\rho - 1)^{\rho - 1}}{\rho^\rho I^{\rho - 1}} \qquad (5.9)$$

The above result is in line with Pindyck and Solimano's (1993:6) solution to the investment problem. The most important argument of the theory of irreversible investment under uncertainty is therefore the following: since ρ > 1, $\frac{\rho}{\rho - 1} > 1$ which ensures that the threshold value V^* is greater than the cost of investment, I. It is because of this that the net present value argument, invest when the future return is greater than or equal to I, becomes incorrect. Thus, according to the new theory of investment, the presence of uncertainty and irreversibility creates a wedge which is equal to $[\frac{\rho}{\rho - 1}]$ between the threshold V^* and I, thereby rendering the simple NPV rule incorrect. This disparity increases with increase in uncertainty about the future return of the project. This is where the new theory of investment differs fundamentally from the neoclassical and other orthodox theories.

Abel and Eberly (1999:340) state that the chief result of the literature on irreversible investment under uncertainty is that irreversibility increases the hurdle that projects must clear in order to be profitably undertaken and that this hurdle increases as uncertainty increases. They further state that, although these results are enormously helpful to a manager making a capital budgeting decision, they do not specifically indicate the optimal amount of investment in the long run behaviour of capital accumulation.

In a similar manner, other authors argue that while the growing theoretical literature has shown the importance of uncertainty in the presence of irreversibility in the short run, its long run impact is not clear at all. Caballero (1997:21) argues that the rise of uncertainty can reduce investment if it leads to increased reluctance to invest, but if not, average investment may increase in the presence of an irreversibility constraint. Another concern about the impact of uncertainty emanates from its relationship to the threshold required to trigger investment. While uncertainty raises the threshold, it can also raise the volatility of the threshold; in other instances, increased uncertainty may raise marginal expected present value and hence lower the threshold and increase investment (Fedderke 2000:14).

Bertola and Caballero (1991:2) do not agree with the preceding analyses and state that even risk neutral firms are reluctant to invest when the projects are irreversible and the future is uncertain, because when the project is adopted, the option to wait for some of the uncertainty to be resolved at some future date is lost and such options are valuable even to risk neutral firms. These results are crucial for the analysis of the investment decisions of firms in the face of risk and uncertainty.

Furthermore, other authors argue that the impact of uncertainty on investment decisions depends on the presence of investment lags. By analyzing the investment decisions of a single firm, Bar-Ilan and Strange (1996) found that the presence of long lags between the time of decision and the beginning of receiving a return can reduce the deterrent effects of uncertainty on irreversible investment decisions. They believe that, while uncertainty can act as a serious deterrent to investment with a short lag, it may encourage investment with a longer lag (Bar-Ilan and Strange 1996:619).

The theoretical debate regarding the long run impact of uncertainty on irreversible investments is wider and deeper in scope than the above statements would imply. According to the standard neoclassical model of investment reversibility, optimal capital budgeting occurs when the marginal revenue product of capital equals the Jorgenson (1963) user cost of capital. Abel and Eberly (1999:340) argue, however, that when the investment is irreversible, the optimal investment policy involves the purchase of capital only as needed to prevent the marginal revenue product of capital from rising above an optimally derived hurdle. This hurdle, which is the user cost of capital appropriately defined to take account of irreversibility and uncertainty, is higher than the Jorgensonian user cost. Thus, if the firm is currently planning to obtain capital and faces a given marginal revenue product schedule, as a decreasing function of capital stock, the optimal capital stock under irreversibility is smaller than the optimal capital stock under reversibility. Abel and Eberly (1999:340) call this the "user cost" effect.

Opposed to the user cost effect is Abel and Eberly's (1999) "hangover" effect which indicates the dependence of the current capital stock on past behaviour. This occurs, under investment irreversibility, when the firm facing a low demand for its products is unable to sell its capital stock to start afresh at some future date. In this case, the firm is currently constrained by its own past investment behaviour. According to Abel and Eberly (1999:341), this dissonance between the firm's actual capital stock and the level that it would choose to hold does not reflect any failure of rationality. Instead, it reflects the presence of the hangover effect that may lead to a higher capital stock under irreversibility than under reversibility. Thus, in the presence of investment irreversibility, the user cost effect tends to reduce the expected capital stock while the hangover effect tends to increase the expected capital stock.

Abel and Eberly (1999:341) make a further point that the two effects also have opposing implications regarding the effects of increased uncertainty on the expected long run capital stock. The user cost effect implies that increased uncertainty tends to reduce the expected long run capital stock under irreversibility while the hangover effect implies that increased uncertainty tends to increase the expected long run capital stock.

Guiso and Parigi (1999:186), on the other hand, state that in an uncertain environment, irreversibility increases the value of waiting for the uncertainty to be at least partly dispelled and naturally leads to postponing investment. In other words, the decision to make an irreversible investment now precludes the option of investing the same resources in the future when more is known. This entails an opportunity cost and this cost increases with uncertainty and depresses current investment. They do state, however, that the effects of uncertainty on investment depend not only on irreversibility or access to second hand markets for capital goods, but also on technology, and the elasticity of the product demand (Guiso and Parigi 1999:187). Similarly, Bo and Lensink (2001:4) believe that the impact of uncertainty on investment also depends on the competitiveness of the firms in the product markets and the technology the firms employ. In the presence of perfect competition and constant returns to scale, the marginal product of capital is the convex function of the uncertainty variable. Hence, by means of Jensen's inequality, an increase in uncertainty will positively affect investment.

Serven (1997:8) disagrees with the preceding view and states that if a firm faces a decreasing returns technology and a downward sloping demand curve, successive marginal increments to the capital stock can be regarded as distinct projects and hence the profitability threshold that must be reached for investment to take place exceeds the user cost of capital and rises with the degree of uncertainty faced by the firm.

The theoretical debates on the impact of uncertainty on investment are not conclusive. These debates imply that the impact of uncertainty on investment is ambiguous and can only be determined with empirical investigation. Nevertheless, despite this concern, a growing empirical literature has proved that the introduction of irreversibility and uncertainty in investment theory has been of paramount importance and any investment project that ignores these issues risks making wrong decisions (Kumo 2006:193). For instance, having carried out empirical investigation of the impact of uncertainty on manufacturing investment, Price (1995:153) concludes that uncertainty has a significant negative effect on investment and the consequences of misspecifying the investment equation by excluding uncertainty may be large. Moreover, using the risk premium as a proxy measure of uncertainty, Ferderer (1993:31) conducted empirical assessment of the impact of uncertainty on investment. He concluded that, in the presence of investment irreversibility, uncertainty, proxied by the risk premium, not only has considerable negative impact on investment, but its impacts are also far greater than those predicted by the cost-of-capital and the q-theories of investment. This is one of the reasons why the study of the impact of uncertainty on investment behaviour of firms facing an irreversibility constraint has become increasingly important since the beginning of the 1990s. The present study carries out empirical investigation of the effects of uncertainty on fixed investment using econometric methodology and the GARCH generated measures of uncertainty. This is done in chapter 9 of the present study.

The problem of investment irreversibility is relevant for the investment and lending decisions of investor-lender (I-L) firms. Investment in loans is irreversible in the sense that once the borrower firm is bankrupt, the lender may fail to recover the whole or part of the capital it invested in loans. The investor-lender firm may face partial or full irreversibility of loans depending on the degree of loan default. Thus, in the presence of uncertainty, loan irreversibility can affect the lending decisions of the I-L firms.

5.4 FINANCE, UNCERTAINTY AND IRREVERSIBLE INVESTMENT

This section investigates the relationship between uncertainty, investment irreversibility and investment financing of the firm. Firms can finance their investment either by borrowing (debt) or by equity. Irreversibility of most investment expenditures and the uncertainty about the future returns from these investments can have a profound impact on the financing policy of the firm and hence on its capital structure.

Some authors have argued that firms' financing policies can be independent of their investment decisions. Modigliani and Miller (1958) studied the investment financing mechanism of firms and concluded that the value of the firm is not affected by the leverage ratio, that is, the ratio of the proportion of debt to the value of the firm, and that investment decisions of the firm can be independent of the financing decisions. More specifically, Modigliani and Miller (1958:268-269) state that the market value of any firm is independent of its capital structure, or alternatively, the average cost of capital to any firm is completely independent of its capital structure. Nickell (1978:150) reemphasized the Modigliani-Miller theorem and stated that "according to this theorem, under certain circumstances, the owners of the firm will be quite indifferent between the outcomes under different financial policies, and [that] the investment decisions of the firm can be taken quite independently of its financing decisions". However, various other authors have questioned this theorem.

Hirshleifer (1966) reviewed the Modigliani-Miller theorem that the market value of any firm is independent of its capital structure by using state preference analysis for choices under uncertainty. He concluded that solving the problem of optimal balance of debt and equity financing requires the integration of personal tax and corporate tax effects, as well as the consideration of other factors such as the magnitude of bankruptcy penalties. This implies crucial limitations to the much debated Modigliani-Miller theorem (Hirshleifer 1966:268).

Other authors agree that the tax effect, the magnitude of bankruptcy penalties and financial distress are considered to be the three most important factors determining the optimal capital structure, i.e. the ratio between equity and debt. However, they do not agree that the investment decisions can be independent of financing decisions by the firm. Marsh (1982:122) argues that the traditional view is that debt has a tax advantage, but that this can be counterbalanced, after a certain level, by the costs of bankruptcy and that financial distress and debt level will be influenced by the probability of financial distress, which in turn is a function of operating and financial risk. He argues that firms with high operating risks are more likely to use less debt.

The most recent literature on investment under uncertainty reemphasizes the impact of financial constraints in the presence of investment irreversibility. Where financial markets are well developed, firms may find it easier to sell used capital goods that may reduce the degree of the irreversibility friction, thereby increasing the chance of the firm to borrow from banks to finance its investments.

This, of course, requires the existence of an active secondhand capital market. If such a market exists for capital goods, investment in fixed capital can be easily reversible and banks will be more willing to lend to these firms (Pattillo 2000:109). Thus, firms with more irreversible investment would rely more on internally generated funds and hence for such firms investment should be more sensitive to internally generated funds (Worthington 1995:52). Moreover, Carruth et al. (1998:2) argue that if capital markets are perfect and the tax treatment of different sources of finance is the same, then investment spending should never be limited by shortage of internal finance; however, they go on to argue that in reality this is not the case. Therefore, the presence of investment irreversibility and uncertainty about the future returns from these investments directly affects the financing decisions of these firms and hence their capital structure.

Moreover, Ghosal and Loungani (2000:342) state that according to the financing constraint hypothesis, firms that borrow in external capital markets can face a premium because of asymmetric information between borrowers and lenders. The increase in uncertainty about the future returns from the investment exacerbates such information asymmetries and can have adverse impact on investment by increasing the premium charged on external funds.

Furthermore, if the firm decides to use debt finance instead of equity finance, it increases its chances of bankruptcy arising from the increased level of debt. In this case the firm may decide not to borrow and not to invest in the absence of internal finance. Vercammen (2000) studied the impact of debt financing on firm investment behaviour and concluded that if the risk of bankruptcy arising from new investment increases with debt, the firm delays investment.

Thus, in the presence of an irreversibility constraint, the financial constraint becomes the most important factor in determining the level of investment under uncertainty. Although the two constraints arise independently of each other, they may complement each other. The financial constraint can exacerbate the irreversibility constraint by making the firm more reluctant to undertake investment, while irreversibility can exacerbate any tendency of the firm to face the financial constraint (Holt 2000:8-9). Both constraints, therefore, raise the value of options to wait and hence delay investment decisions by firms.

The financing decisions of firms under uncertainty have direct bearing on the decisions of investor-lender (I-L) firms. The lending decisions of these firms depend on the financing policy of the firms that decide to make irreversible investments. If the I-L firms are able to verify the state of the borrowing firms they will be able to avoid loan default and their own investment irreversibility problem. If a financially constrained firm decides to embark on irreversible investment which may lead to bankruptcy, the investment of the I-L firm also faces investment irreversibility in the form of unrecoverable loans.

Several theories of economic decisions under uncertainty question the rationality of such economic decisions. The theories of rational decisions under uncertainty are divided into two broad categories. These are expected utility and nonexpected utility theories of choice under uncertainty. The next chapter examines these theories in greater detail.

5.5 SUMMARY AND CONCLUSION

The theory of investment under uncertainty reemphasizes the importance of uncertainty in investment decisions. According to this theory, investment decisions of firms are affected by the interaction of three factors. These are: partial or full irreversibility of most investment expenditures, the option value of waiting to obtain more information about the future and uncertainty about the future returns to an investment.

The option to defer investment is particularly important in resource extraction industries, farming, paper manufacturing and real estate development due to high uncertainties and long investment horizons. Recent studies attempted to include the real option of deferring investment in the traditional NPV rule of investment decisions, using the contingent claims analysis (CCA). Using the expanded NPV rule and static NPV it is possible to calculate the value of the option to defer investment. In the presence of irreversibility and uncertainty, a high option value makes investors more reluctant to invest, leading to a drop in investment. In other words, the high value of the option to wait can lead to a decrease in investment by delaying the investment decisions.

The new theory of investment under uncertainty emphasizes the importance of investment irreversibility as the main friction in investment decisions. In the presence of uncertainty, when investors believe that they will not be able to fully recover the cost of their investment when such investments prove to be unprofitable, they prefer to delay investment. The presence of irreversibility and uncertainty delays investment by creating a wedge between the required threshold rate of return and the cost of investment, and by causing the value of the threshold to exceed the cost of investment. This disparity increases with an increase in uncertainty about the future returns to the project. This also contradicts the usual net preset value rule. Thus, in the presence of investment irreversibility, uncertainty can be a major deterrent to investment.

However, the theoretical literature does not agree on the sign of the impact of uncertainty on investment in the presence of irreversibility. While one strand of the literature emphasizes the negative impact of uncertainty on irreversible investment, another strand suggests that uncertainty can have a negative impact on investment in the short run, while its long run impact is unclear or may even be positive. Others suggest that the negative impact of investment depends on the duration of the investment lags. The most persuasive of these debates is the Abel-Eberly user cost and hangover effects where, in the presence of investment irreversibility, the user cost effect tends to reduce the expected capital stock, while the hangover effect tends to increase the expected capital stock. These two effects also have opposing implications regarding the effects of increased uncertainty on the expected long run capital stock. While the user cost effect implies that increased uncertainty tends to reduce the expected long run capital stock under irreversibility, the hangover effect suggests that increased uncertainty tends to increase the expected long run capital stock.

In addition to the irreversibility constraint, the financial constraint can be an additional deterrent to investment in the presence of uncertainty. Although earlier theories stipulated that investment decisions of firms can be independent of financing decisions, the new theory of investment under uncertainty stresses that, in the presence of investment irreversibility and uncertainty, investment decisions may not be independent of financing decisions. Banks will be reluctant to lend to firms whose investments are fully irreversible. Moreover, if the firm decides to finance its investment by debt and the risk of bankruptcy increases with the level of debt, it will not go ahead with the investment. Thus, in the presence of uncertainty, investment irreversibility and financial constraints reinforce each other, increase the option value of waiting, and reduce investment.

CHAPTER SIX

EXPECTED AND NONEXPECTED UTILITY THEORIES OF CHOICE UNDER UNCERTAINTY

6.1 INTRODUCTION

The expected utility theories have been the most dominant theories in the analysis of choice under uncertainty. These theories, particularly the Von Neumann-Morgenstern and the Savage expected utility theories follow axiomatic presentations of agents' preferences. Some of these axioms are hypotheses of rational choices, that is, economic agents make rational choices.

However, some empirical economic studies have suggested that economic agents do not behave in the manner explained by the expected utility theories. These studies usually focus on the violation of the strong independence assumption of the Von Neumann-Morgenstern expected utility (EU) theory. The earliest challenge to the independence or linearity in probability axiom of the EU theory came from Allais (1953) who investigated the linearity in probability assumption of the EU theory and concluded that economic agents follow nonlinear weighting of probability in their decision processes. This was called the Allais paradox. Allais (1979b) further investigated the violation of the independence axiom of the EU theory using nonlinear intensity theory, which will be reviewed in detail in this chapter. Ellsberg's (1961) paradox, discussed in chapter 4 of the present study, was another theory that questioned the independence axiom of the EU theory.

Among other alternative nonexpected utility theories, the Prospect Theory (PT) of Kahneman and Tversky (1979) and Tversky and Kahneman (1986) provided more detailed alternative explanations of the choice behaviour of economic agents under uncertainty. Using the analytical tool of framing of decisions, these authors state that the

EU theory is a normative analysis used to predict and explain actual behaviour, but that it cannot serve as an adequate descriptive model of human behaviour. They argue that when individuals are presented with choices in different frames, they make decisions that violate the main axiom of the EU theory, i.e. the independence axiom or linearity in probability axiom. In prospect theory this refers to violation of the dominance axiom. They also argue that individuals not only violate the dominance axiom but also the invariance axiom. There are two types of invariance: procedure invariance and description invariance. The failure of procedure invariance is known as the phenomenon of preference reversal first observed by Lichtenstein and Slovic (1971) and Lindman (1971) and later used by Tversky, Slovic and Kahneman (1990) and Tversky and Thaler (1990) in support of the prospect theory. The framing effect reflects the violation of description invariance. Recently, based on such systematic violations of the basic axioms of the EU theory, some authors adopted the explicit definition of investor irrationality.

The conclusion of the violation of the independence axiom of the EU theory is based on the fundamental assumption used by the nonexpected utility models. This assumption is that the objective of economic agents is not the maximization of the expected utility of profit, but the maximization of some value. The fundamental nature of such an assumption is that, while it incorporates price and revenue uncertainty, it excludes cost uncertainty in the economic analysis of the decisions of economic agents. In real life situations, profit maximization is one of the most important objectives of economic agents and hence analysis of economic decisions should incorporate all sources of uncertainty that affect the attainment of this objective. When all sources of uncertainty are included and the objective of an economic agent is the maximization of the expected utility of profit, the EU theory is capable of serving as an adequate descriptive theory of choice under uncertainty.

Moreover, Battalio et al. (1990), Birnbaum and Navarrete (1998) and Birnbaum (2004a,2005) have carried out empirical tests of the alternative nonexpected utility theories. Battalio et al. (1990) conducted empirical tests of the four main nonexpected utility models of choice under uncertainty, viz. generalized expected utility analysis, prospect theory, rank-dependent expected utility models and regret theory, all of which are reviewed in this chapter, in order to assess their adequacy as alternatives to EU theory. These authors conclude that none of the four theories considered consistently organize choices and further research is needed to develop alternatives to EU theory as a descriptive theory of choice under uncertainty. Similarly, Birnbaum and Navarrete (1998) and Birnbaum (2004a, 2005) have empirically tested rank-dependent utility models and the cumulative prospect theory and conclude that these theories violate their own basic assumptions, such as stochastic dominance and cumulative independence, and hence cannot serve as alterative descriptive theories of choice under uncertainty. Therefore, the I-L model developed in Chapter 8 of the present study will be based on the assumption of the maximization of the expected utility of profit in line with the EU theory. The objective of this chapter, however, is to review both expected utility and nonexpected utility theories in terms of their capacity to provide an alternative descriptive theory of choice under uncertainty.

The chapter is organized as follows: section 6.2 reviews various concepts of rationality used in economic analysis. Section 6.3 investigates the expected utility models as the models of rational choice under uncertainty, beginning with the Bernoullian expected utility theory. Section 6.4 analyzes alternative nonexpected utility theories including prospect theories as the alternative theories of rational choice under uncertainty. Section 6.5 reviews the phenomenon of preference reversal. Section 6.6 reviews the Levy et al. (2000) hypothesis of investor inefficiency and irrationality. Finally, section 6.7 presents a summary and conclusion.

6.2 RATIONALITY IN ECONOMICS

The rationality postulate has been one of the cornerstones of economic analysis. Its premises are based on the belief that economic agents always make well reasoned decisions when confronted with various choices. There are three major postulates of rationality used in economic analysis. These are: (a) instrumental (global) rationality, the usual neoclassical utility maximizing rational economic agent model, (b) procedural rationality, i.e. behavioural rationality based on rules and norms developed by society, and (c) expressive rationality (bounded rationality) involving decisions based on limited knowledge due to uncertainty about the future (Simon 1978, 1982; Heap 1989).

6.2.1 Instrumental rationality

The hypothesis of instrumental (global) rationality is used predominantly in the neoclassical model of the utility maximizing rational economic agent. Under this hypothesis the rational action is equated with the identification of the end or objective to be achieved and the means to achieve this end. This rationality assumption provides an explanation that is normally referred to as intentional. However, there are situations where individual actions can be causally connected in such a way that they produce unintended consequences producing a supra-intentional causality (Heap 1989:39). This kind of causality refers, for example, to the famous Adam Smith (1776) assertion that "the acts of butchers, brewers and bakers arising from their own interest, provide us with our dinner" which reflects the workings of the invisible hand in the free market system.

According to the instrumental rationality assumption, a rational action is defined with respect to a given set of objectives and it is the action that best satisfies this objective. In terms of the consumer theory, the action of an instrumentally rational agent ensures the maximization of a well behaved utility function in the sense of reflexivity, completeness, transitivity and continuity, as well as convexity or quasi-concavity of the utility function subject to non-satiation or budget constraint. An action is instrumentally rational if it satisfies these well ordered preferences.

The most important application of the instrumental rationality assumption is the neoclassical general equilibrium theory. This theory is based on the causal explanations of interactions of individual actions leading to unintended consequences which are beneficial to society. Arrow and Hahn (1971) studied an economy with several agents with preferences as described above and with tradable initial endowments, where the production function was differentiable and quasi-concave as in the case of the utility function, and the objective of firms was to maximize profit subject to the production constraint under competitive conditions. They concluded that there exists a price vector that ensures general equilibrium or that equates supply and demand in both commodity and factor markets.

The major weakness of instrumental rationality is its limited scope in terms of its capacity to explain the institutional and informational structures that orchestrate instrumentally driven explanations and hence the implied human freedom and uncertainty in the social world (Heap 1989:10). However, theories of decision under uncertainty extended the assumption of instrumental rationality to choice under uncertainty by assuming that uncertainty can be measured by probability distributions, in which case utility maximization becomes expected utility maximization. Typical of these theories is the expected utility theory of choice under risk and uncertainty.

6.2.2 Procedural rationality

Procedural rationality refers to behavioural rationality, which is based on rules and norms developed by society. These rules cannot be reduced to, or explained by, instrumental rationality alone and are often shared by several agents. These rules and norms are sources of coordination and communication in the social world constituting a building block for our shared institutions and culture. Thus, when instrumental rationality faces problems of limited scope in some intentional explanations, procedural rationality can be invoked (Heap 1989:116).

Simon (1978:8-9) distinguishes between substantive rationality that refers to the extent to which agents choose an appropriate course of action, and procedural rationality, which refers to the effectiveness of the procedure used to choose actions in the face of limited human cognitive and computational powers. Thus, according to Simon, economic agents fall back on procedures and norms when they face a limited ability to correctly calculate future outcomes. By contrast, Heap (1989:118) argues that, while limited computational capacity is one reason for agents relying on procedural rationality, it is not the only reason. Economic agents use shared procedures to communicate and coordinate their actions with each other. In this way procedural rationality fills the gaps in the scope of instrumental rationality. It becomes crucial in our explanation of institutions and information and the implied concepts of uncertainty and freedom. Keynes's (1936:114) statement that agents fall back on conventions when faced with uncertainty reflects reliance on procedural rationality instead of instrumental rationality.

6.2.3 Expressive (bounded) rationality

Expressive rationality is related to the analysis of various approaches regarding the issues of what are good reasons for holding some preferences or objectives rather than others, and is associated with the human behaviour of self respect. It is complementary to both instrumental and procedural rationality and, in particular, can be reduced to the former with the introduction of the objective or goal of self respect (Heap 1989:148-151).

Expressive rationality involves a process of making decisions under risk and uncertainty. Heap (1989:164) states that economic theories that observed violations of some assumptions of the expected utility theory, such as the regret theory, could be regarded as an example of expressive rationality, because "uncertainty in the souls" has been added to the risk present in the decision environment. However, this definition of expressive rationality contradicts the notion of the complementarity between the three postulates of rationality. Expressive rationality, unlike the other two, captures the process in which incomplete information about the future faced by agents forces them to make decisions which are less than rational. This notion is related more to Simon's bounded rationality rather than to instrumental or global rationality and is useful in the analysis of decisions under uncertainty. Simon (1982:405-411) defines *bounded rationality* as *subjective rationality* caused by the presence of risk and uncertainty, incomplete information about alternatives and complexity in the environmental constraint facing agents, all of which limit their power to calculate the best course of action.

Kreps (1990:151) distinguishes between bounded rationality and retrospection. While he agrees with Simon's definition of boundedly rational behaviour as the behaviour that is *intendedly rational but limitedly so*, in the sense that the individual strives consciously to achieve some goals but does so in a way that reflects cognitive and computational limitations, Kreps questions whether the line dividing bounded rationality and irrationality is drawn somewhere between behaviour that follows some coherent procedure and behaviour that is simply chaotic (Kreps 1990:151). For Kreps, it is here that the notion of retrospective behaviour becomes important. Retrospective behaviour is loosely defined as behaviour in which the past influences current decisions. When behaviour is governed by social convention, retrospection can be usefully applied (Kreps 1990:151-152). Thus, retrospective behaviour complements boundedly rational behaviour where economic agents are considered to be not only boundedly rational but also retrospectively rational. However, Farmer (1995:70) interprets bounded rationality within the neoclassical rational agent hypothesis and argues that the notion of bounded rationality, or of actors as rule followers, runs the risk of "putting the blame" on individual humans' limited cognitive abilities, but fails to recognise the role of the changing nature of the social world in explaining this behaviour. Her argument does not make any distinction between bounded and instrumental rationality and emphasizes the limitation of the entire concept of the rational actor using the hypothesis of bounded rationality.

The assumption of rationality is so powerful because no other theory of judgment and decision matches it in scope, power or simplicity. Thus, it is commonly assumed that the substantial violations of this model of the rational agent are: (a) restricted to insignificant choice problems, (b) quickly eliminated by learning, and (c) irrelevant to economics because of the corrective function of the market forces (Tversky and Kahneman 1986:S273). Tversky and Kahneman argue that there are cases where these forces fail to elicit the required rational behaviour and economic agents may make significant decisions that are less than rational. However, to date no economic theory has produced a forceful alternative assumption that replaces this powerful assumption of rationality.

Sen (1985:121-122) splits rational choice into two elements; namely, correspondence rationality and reflection rationality. Correspondence irrationality refers to the inadequate correspondence between the person's reasoned reflection and his actual choices. This can be caused by any of the following: acting "without thinking", "lazy reflection" and "weakness of will". On the other hand, reflection irrationality is a lack of careful reflection or the phenomenon that, despite reflecting carefully, connections may be missed and relevant considerations ignored because of intellectual limitations, possibly due to a lack of knowledge of decision problems. He links the problem of violations of strong independence with the problem of reflection rationality.

Rationality should not be considered as perfectly coherent decision-making. It can be defined as the behaviour that strives to avoid mistakes, but admits that, in the face of an unknown future, mistakes are unavoidable and hence decisions are made with this in mind. In the following sections, we investigate various expected and nonexepected utility theories that analyze the rationality of individual decision-making processes under uncertainty.

6.3 RATIONAL DECISIONS AND THE EXPECTED UTILITY MODELS

6.3.1 Bernoullian expected utility theory

During the early years of the development of probability theory, the widely held view was that risky monetary ventures ought to be evaluated by their expected returns. Suppose $(p_1,\ldots,\ p_n)$ and $(q_1,\ldots,\ q_n)$ are probability distributions on a set of X of monetary gains $(x \geq 0)$ and losses $(x < 0)$ that correspond to two risky ventures. Then the expected return of the two ventures is given by:

$$E(x, p) = \sum xp(x) \text{ and } E(x, q) = \sum xq(x)\,, \quad \text{for } x \in X. \qquad (6.1)$$

In the above case q is preferred to p when $\sum xq(x) > \sum xp(x)$.

This principle of expected value maximization was challenged for the fist time by Daniel Bernoulli in 1738 (Fishburn 1988:1). He states that, given an initial wealth w_0, a return from a risky prospect should be evaluated by its expected subjective value or expected utility:

$$E(v, p) = \sum v(w_0 + x)p(x) \text{ and } E(v, q) = \sum v(w_0 + x)q(x), \quad \text{for } x \in X. \ (6.2)$$

Bernoulli's work on evaluating risky ventures was motivated by the famous *St Petersburg paradox* devised by his cousin Nicolas Bernoulli in 1713, which violated the principle of maximum expected return. According to this game, suppose a fair coin is tossed until a head appears. If the first head appears during the first toss, the payoff is \$1, if it appears at the second toss the payoff is \$2, \$4 if it takes three tosses, etc. Suppose a person is entitled to one play of the game without cost. What is the least amount one would sell his entitlement for? Since this game offers ½ chance of winning \$1, ¼ chance of winning \$2, 1/8 chance of winning \$4, etc., its expected value is (1/2)\$1 + (¼)\$2 + (1/8)\$4 +… = \$1/2 +\$1/2 +\$1/2 +… = infinite, implying that a person would want to sell his right for an infinite amount of money. However, since in real life individuals would agree to sell their right for a relatively small amount, the above shows a fundamental violation of the principle of expected value maximization. Daniel Bernoulli devised a solution to the above problem by providing a unique solution s to the equation

$$\sum v(w_0 + 2^n)2^{-n} = v(w_0 + s) \qquad (6.3)$$

for any finite w_0, where s is the minimum selling price or equivalent monetary value.

Thus, in terms of the expected utility approach, the sure gain g which would yield the same utility as the St. Petersburg gamble, i.e. the certainty equivalent of this gamble, is determined by the equation

$$U(w_0 + g) = (\tfrac{1}{2})U(w_0 + 1) + (\tfrac{1}{4})U(w_0 + 2) + (1/8)U(w_0 + 4) + \ldots$$

Assuming that the utility function takes a logarithmic form and that the initial wealth is \$50,000, Machina (1987:123) has shown that the individual's certainty equivalent gain g would only be about \$9 even though the gamble has an infinite expected value.

The Bernoullian expected utility theory is the theory of choice in risky decisions when they consist of the following (Fishburn 1988:6)

1) A set X of outcomes and a set P of probability distribution and or measures of X.
2) A utility function v on X based on a notion of riskless comparable preference differences, usually presumed unique up to positive linear transformations.
3) The principle of choice, which says that the most desirable distributions or their corresponding risky alternatives are those that maximize expected utility, $\sum v(x)p(x)$.

Of the three conditions, it is the second condition that makes the Bernoullian theory fundamentally different from the Von Neumann-Morgenstern expected utility theory, which we shall review in the following section.

6.3.2 The Von Neumann-Morgenstern expected utility theory as the theory of rational choice under uncertainty

The Von Neumann-Morgenstern expected utility (EU) theory (1944, 1947) has identical forms of expected utility to the Bernoullian expected utility theory. Thus, while the Bernoullian expected utility model is described as $\sum v(x)p(x)$, the Von Neumann-Morgenstern expected utility model is described as $\sum u(x)p(x)$. However, the two models are radically different from each other. While u and v represent the same individual preference ordering and both u and v are unique up to positive linear transformations, their interpretation and assessment are quite different (Fishburn 1988:7).

The Von Neumann-Morgenstern theory follows an axiomatic approach in analyzing a binary preference relation $\succ$ on a convex set of P which implies the existence of the real valued function u on P that preserves the order of $\succ$ on P and is linear in the convexity operation, i.e. *it is an order preserving linear functional.* The axioms of the Von Neumann-Morgenstern theory apply solely to $\succ$ on P. Unlike the Bernoullian theory, preference under the Von Neumann-Morgenstern theory applies to comparisons of risky alternatives, not just outcomes. Furthermore, the Von Neumann-Morgenstern axioms involve no notion of comparable preference differences or strengths of preference, since they use only "ordinal" preference comparisons (Fishburn 1988:7). However, Machina (1987:123) argues that the utility concept of Von Neumann-Morgenstern is quite distinct from the ordinal utility function of the standard consumer theory, because, while the latter can be subjected to any monotonic transformation, the Von Neumann-Morgenstern utility function is cardinal in that it can only be subjected to transformations of the form aU(x) + b(a > 0), i.e. transformations which change the origin and/or scale of the vertical axis, but do not affect the "shape" of the function. Machina further argues that the monotonicity of the utility functions reflect the property of stochastic dominance preference, where one venture is said to stochastically dominate another if it can be obtained from it by shifting probability from lower to higher outcome levels. For instance, a 2/3:1/3 chance of \$100 or \$20 and a ½ : ½ chance of \$100 or \$30 stochastically dominates a ½ : ½ chance of \$100 or \$20. Stochastic dominance is the probabilistic equivalent of the attitude that more is preferred to less.

Stochastic dominance is assumed in the first three most important axioms of the five axioms, i.e. axioms A1 to A5 of the Von Neumann-Morgenstern expected utility model discussed in chapter 2 of the present study. The first axiom, A1, of ordering and transitivity, is related to the main economic conception of rationality. The assumption of transitivity means that if an individual prefers one object to a second one and prefers this second object to a third one, then he prefers the first object to the third one. That is, if $A \succ B$ and $B \succ C$, then $A \succ C$. Therefore, violations of A1 are seen as a serious weakening of the assumptions of the expected utility theory. According to Camerer (1989:63), in indifference curves in a unit probability triangle, transitivity implies that indifference curves do not cross inside the triangle, as indicated in figure 6.1.

In expected utility theory, the indifference curve in the triangle diagram is a set of gambles with the same expected utility E(U). Consider three gambles, X_L, X_M and X_H. The expected utility from these gambles can be expressed as:

$$E(U) = P_L U(X_L) + P_M U(X_M) + P_H U(X_H) \tag{6.4}$$

where P_L, P_M, and P_H are probabilities corresponding to the three gambles and subscripts *H, M* and *L* refer to highly valued, middle valued and lowest valued gambles. The sum of the three probabilities is one, i.e. $P_L + P_M + P_H = 1$. The slope of the tangent line to an indifference curve at a point is given by

$$\frac{dP_H}{dPL} = \frac{U(X_M) - U(X_L)}{U(X_H) - U(X_M)} \tag{6.5}$$

Equation (6.5) can be interpreted as a discrete marginal rate of substitution of P_H for P_L, or as the shadow price of probabilistic units of the highly valued gamble in terms of probabilistic units of the lowest valued gamble (Camerer 1989:64).

However, the most challenged axiom of the expected utility theory is the second axiom A2, the independence axiom. This axiom is also known as the linearity assumption and is closely associated with similar axioms, such as substitution principles, cancellation conditions, additivity axioms and Savage's sure-thing principle. It simply says that if p is preferred to q then a convex combination of p and r is preferred to the similar combination of q and r and reflects the consistency of preferences (Fishburn 1988:11). This is another important notion of rational decisions. We will present theories that challenge the independence or linearity assumption of the EU theory in the next section.

The third assumption, A3, guarantees that one object is not preferred to another infinitely. It provides probability limits α and β in (0, 1) where if p is preferred to q and q is preferred to r, then with a probability, $\alpha < 1$, $\alpha p + (1-\alpha)r$ is preferred to q and with a probability, $\beta > 0$, q is preferred to $\beta p + (1-\beta)r$ (Fishburn 1988:11).

The axiom of continuity of EU theory ensures that every gamble lies on some indifference curve and that there are no holes in the indifference map in a unit probability triangle. The independence axiom implies that indifference curves are parallel straight lines (Camerer 1989:63). Figure 6.1 shows the indifference map in a unit probability triangle that satisfies the assumptions of the EU theory.

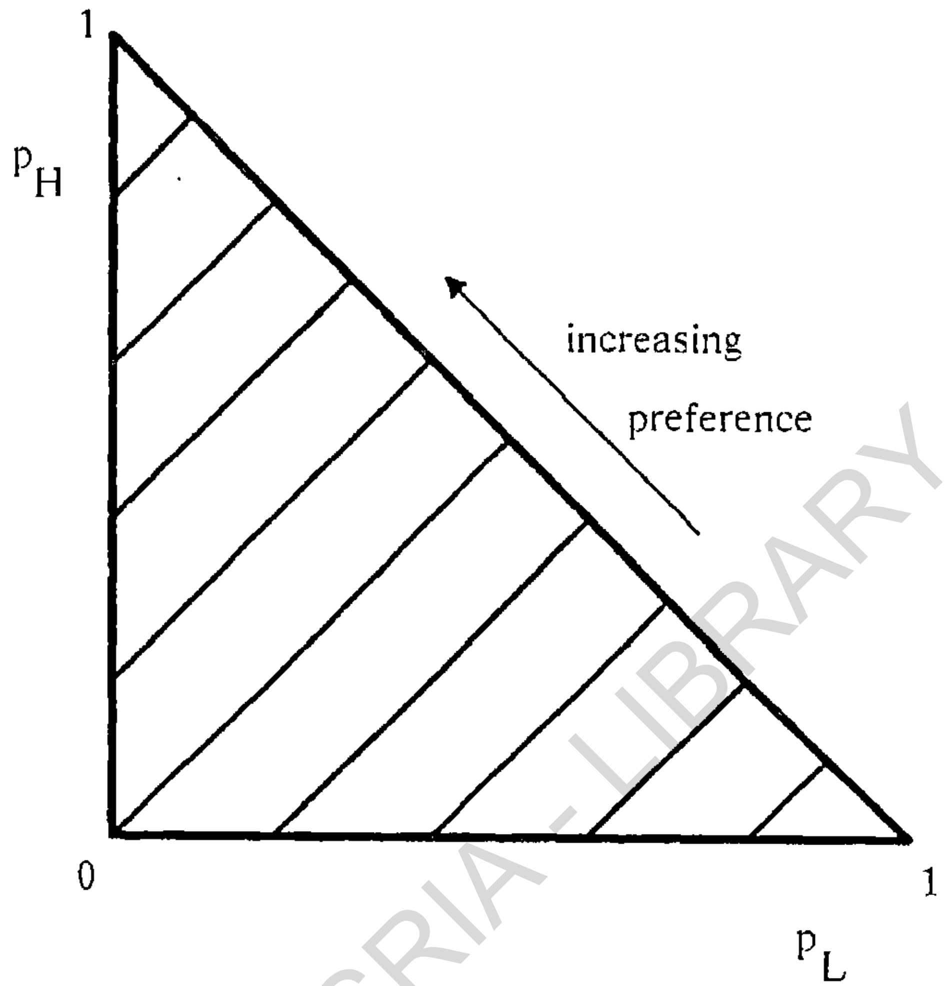

Figure 6.1 Indifference curves assumed by expected utility theory in a unit probability triangle

The expected utility theory has long been a dominant tool in economic analysis. However, some theories based on empirical experiments have challenged its accuracy as a descriptive tool in economic analysis. As stated earlier, the first theory to challenge the basic assumption of independence or linearity of EU theory was Allias's (1953, 1979b) nonlinear intensity theory, also known as the Allais paradox. Another theory that emphasizes systematic violation of the independence axiom of the EU theory is the prospect theory of Kahneman and Tversky (1979) and Tversky and Kahneman (1986).

Some authors argue that the challenge to the EU theory may run deeper than violations of the independence axiom. Two assumptions implicit in any conventional theory are *procedure invariance* (preferences over prospects are independent of the method used to elicit them) and *description invariance* (preferences over prospects are purely a function of the probability distributions of consequences implied by prospects and *do not depend on how those given distributions are described*) (Starmer 2004:110). The phenomenon that describes the failure of procedure invariance is known as *preference reversal* and the phenomenon that describes the description invariance is known as the *framing effect*. However, before we review theories that challenge the main axioms of the Von Neumann-Morgenstern expected utility theory, we present the Savage expected utility theory.

6.3.3 The Savage expected utility theory

Savage's (1954, 1972) theory of expected utility is considered to be a subjective revolution compared to the Von Neumann and Morgenstern objectivistic approach. The theory focuses on personal probabilistic views of a "rational" economic agent faced with a problem of decision-making under uncertainty (Savage 1972:6-7). The term subjective also refers to the additive nature of the subjective probabilistic view of an individual in Savage's theory.

Savage formulates his theory by using a set S of states of the world and the set X of consequences or outcomes, where states are the carriers of uncertainty, and outcomes the carriers of value. He defines the world as the object about which the person is concerned, state as the description of the world including all relevant aspects, and true state as the state that obtains (Savage 1972:9). In other words, the states in S are considered to be mutually exclusive and collectively exhaustive so that exactly one state, the state that obtains, is the true state. The decision maker is uncertain about which state is the true state and the identification of the true state will not be known before the decision is taken and the decision is presumed not to affect the state that obtains (Fishburn 1988:159). Thus, Savage's subjective expected utility and subjective probability theory was developed with the assumption of state independent preferences.

In the Savage (1954) model, the canonical version of the Bayesian model, the agent is assumed to choose one form of the set of acts, taking into account the consequences of each act for each possible state of the world. In terms of Savage's first postulate, a preference relation is defined for any pair of acts constructed in this way (Runde 1995:206). A decision alternative called an act is a function from a set of states S into a set of outcomes X. The outcome assigned by act f to state s is $f(s)$ and this act f is constant if $f(S) = \{x\}$ for some $x \in X$ and is simple if $f(S) = \{f(s): s \in S\}$ is finite (Fishburn 1998:160). Applying the preference relation $\succ$ to a set of F of acts in his axioms, Savage states that an agent prefers a set of constant acts $x \succ y$ if $f \succ g$ when $f(S) = \{x\}$ and $g(S) = \{y\}$ and depending on the additive subjective probability for the occurrence of the state of nature. The three important axioms used in Savage's theory are additivity, independence and substitution, and these are related also to Savage's sure-thing principle.

Savage's additive probability theory has been challenged by several authors. The first of these challenges came from Allais's additive nonexpected intensive utility theory for decisions under uncertainty, which analyzed the violations of the independence and transitivity axioms of the Von Neumann-Morgenstern expected utility theory and of Savage's independence and substitution axioms. Since then various alternative models have been developed to accommodate violations of the independence, substitution, reduction and transitivity axioms of expected utility theory (Fishburn 1988:187).

Ellsberg (1961) used the notion of event ambiguity to construct examples that challenge Savage's substitution principle and the closely related additivity axiom. A variety of alternatives have been proposed to accommodate noncomparability of incommensurable events, imprecise or vague judgements, ambiguity, failure of additivity and intransitivities represented by the following theories: nonadditive subjective probability, additive nonexpected intensive utility, expected regret theory, nonadditive linear utility, nonadditive expected utility and nonadditive, nontransitive theories (Fishburn 1988:190-204). The following section investigates alternative nonexpected utility theories of choice under uncertainty.

6.4 RATIONAL DECISIONS AND NONEXPECTED UTILITY MODELS

The expected utility model dominated the analysis of individual decision-making under risk and uncertainty for several decades. However, various attempts have been made to better understand the determinants of individual choice behaviour since as early as the 1950s. This involved both empirical and theoretical developments. On the empirical front, the discovery by Allais (1953) of the individual choice behaviour that violates the independence axiom of the expected utility theory has stimulated theoretical developments in nonexpected utility theory. These developments have taken both conventional and nonconventional approaches. The conventional approaches include the *"fanning-out"* hypothesis of Machina (1982) which proposes an analytical extension of EU theory in the form of generalized expected utility and the decision weighting models with more sophisticated probability transformation designed to ensure the monotonicity of the value function. The most common of these is the *rank-dependent expected utility* theory. There are two theories that take the nonconventional route. These are: (a) the prospect theory and (b) the nontransitive preference theories, the most prominent of which is Loomes and Sugden's (1982) *regret theory* (Starmer 2004:116-132). This chapter reviews all five main nonexpected utility theories, including the Allais nonlinear intensity theory. The following section analyzes the theories that generalize the expected utility theory.

6.4.1 Generalizations of expected utility theory

6.4.1.1 Machina's generalized expected utility analysis

In light of some empirical evidence of the violations of the independence axiom of the expected utility theory, Machina (1982) proposes an analytical extension of EU theory in the form of generalized expected utility, instead of discarding the theory for its descriptive deficiency. He believes that while the independence axiom, the key behavioural assumption of the EU theory, is systematically violated in practice, the basic concepts, tools and results of this theory do not depend on this assumption. Instead, they may be derived by merely assuming smoothness of preferences or differentiability of the preference functional over alternative probability distributions (Machina 1982:277-79). He argues further that the generalization of the expected utility analysis to include some hypotheses about the shape of a fixed nonlinear preference functional over probability distributions can generate predictions consistent with cases that show the violations of the independence axiom. This refers to Machina's *fanning-out* hypothesis: indifference curves that fan out can yield results that are consistent with Allais as well as St Petersburg paradoxes, the Friedman-Savage hypothesis on individual behaviour towards risk, and Markowitz's observation that preferences over risky outcomes are independent of the current level of wealth (Machina 1982:279-80). Machina's attempt to generalize the expected utility theory is intended to improve its descriptive validity as a theory of decision under uncertainty.

As in the expected utility theory, Machina (1982:293-294) uses axioms similar to those of EU theory, such as complete weak order, continuity and monotonicity, but replaces the axiom of independence with Fréchet differentiability. Fréchet differentiability is the natural notion of differentiability on spaces such as D[0, M]. Suppose that $F(\bullet)$ is a Fréchet differentiable preference function on D[0, M] then $F(Q^*)$ is strictly less than or equal to $F(Q)$ whenever $Q^*(\bullet)$ differs from $Q(\bullet)$ by a mean-preserving increase in risk and if and only if $U(x:Q)$ is a concave function of x for all $Q(\bullet) \in D[0,M]$. The fact that any Fréchet differentiable preference function may be thought of as "locally expected utility maximizing" follows from the fact that differentiable functions are "locally linear," and that for preference functionals over probability distributions, linearity is equivalent to expected utility maximization (Machina 1982:294-295).

According to Machina (1982:282), the generalized expected utility analysis allows us to model behaviours inconsistent with the EU theory, such as the purchase of lottery tickets at all wealth levels, and yet at the same time avoid the adverse behavioural implications of unbounded Von Neumann-Morgenstern utility functions. Therefore, according to the generalized expected utility (GEU) analysis, the assumption of linearity or the independence axiom is related to the "boundedness of utility" debate. According to this debate, the assumption of terminally convex local utility functions merely implies that the linear approximations to the preference functional are unbounded linear functionals, whereas the assignment of infinite certainty equivalents by an expected utility maximizer with unbounded utility, as in the St Petersburg paradox, follows from the fact that for such an individual, the preference functional itself is unbounded linear functional. However, once the assumption of linearity or the independence axiom is dropped, these two conditions become quite distinct (Machina 1982:282-284).

Under Machina's fanning-out hypothesis, preferences are represented by nonparallel indifference curves that fan out within a unit probability triangle, as shown in figure 6.2. The steepness of the indifference curves in the triangle diagram is a measure of risk aversion and hence *hypothesis II* predicts that curves will be steeper for gambles to the northwest, i.e. with lower P_L or higher P_H.

In Machina's fanning-out hypothesis indifference curves fan out or grow steeper from the lower right to the upper left of the unit probability triangle in figure 6.2. Machina's main concern is to demonstrate theoretically that tools of the EU theory may be used in economic analysis in spite of some violations of the assumptions of the theory. By using general preference functions of the form V(X), Machina argues that many properties of the local utility functions of V(X), such as risk aversion, imply global properties of V(X), and hence many of the standard assumptions of the EU theory hold for V(X) even though V(X) does not satisfy EU (Camerer 1989:71).

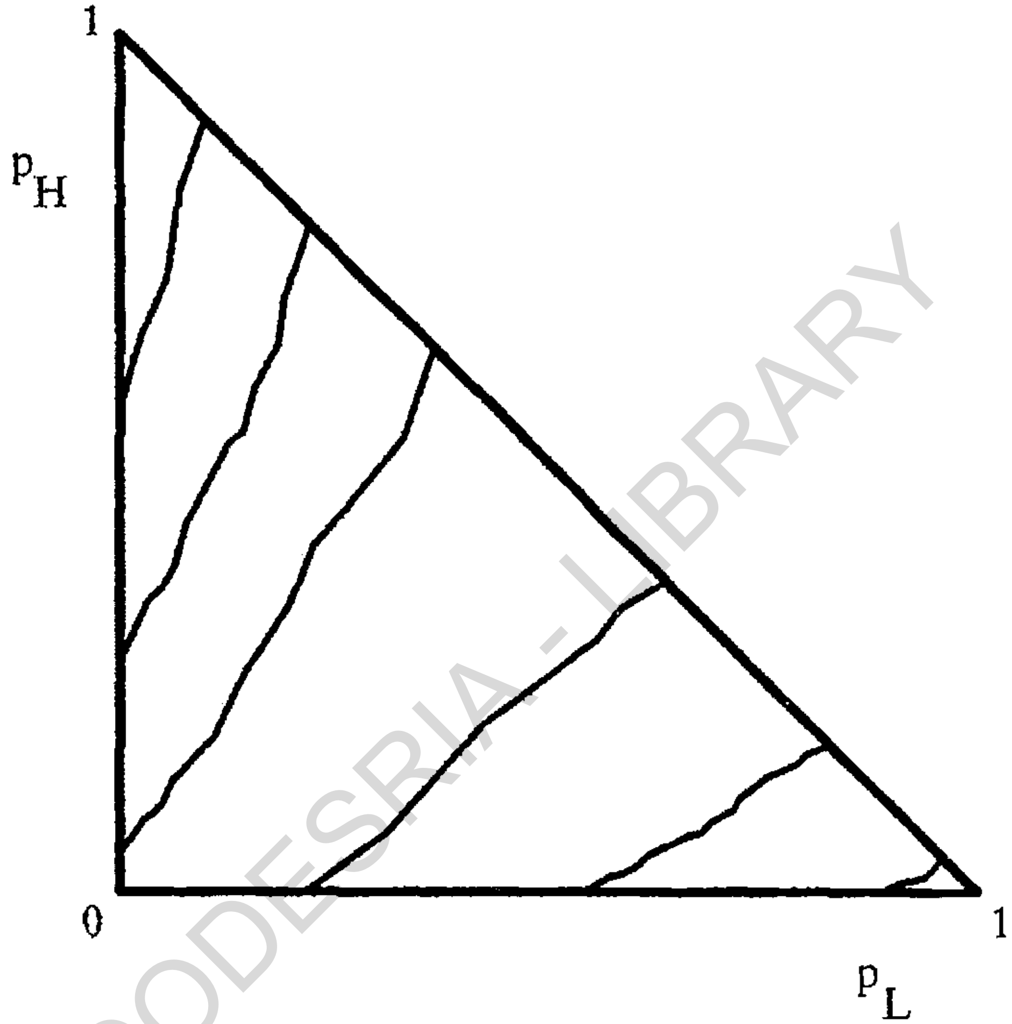

Figure 6.2 Indifference curves assumed by Machina's fanning-out hypothesis in a unit probability triangle.

6.4.1.1.1 Empirical tests of the fanning-out hypothesis

Battalio et al. (1990) tested Machina's type II hypothesis empirically together with other nonexpected utility theories. The type II hypothesis is an assumption about the shape of the individual preference functional used by Machina to explain four classes of systematic violations of the independence axiom of the EU theory. These are: (i) the common consequence effect, the most common example of which is the Allais paradox, (ii) the common ratio effect, (iii) oversensitivity to changes to small probability outlying events, and (iv) the utility evaluation effect (Battalio et al. 1990:41 and 48). According to hypothesis II, in moving from one probability distribution to another that stochastically dominates it, the local (linear in probabilities) utility function retains the same degree of concavity, or becomes more concave at each point. In terms of the unit probability triangle, this yields indifference curves that are parallel, as expected utility theory predicts, or that "fan out" relative to the origin as the type II hypothesis predicts (Battalio et al. 1990:41). However, the empirical test of the "fanning-out" effect by Battalio et al. (1990:42) has revealed that in all cases of the test, fanning out explained less than 50% of the deviation from the expected utility theory while in some specific choice problems the expected utility organized 44.4% of the data, compared to only 20% by the fanning-out hypothesis. In their tests of this hypothesis, Battalio et al. (1990:46) found conditions leading to large systematic violations of hypothesis II in the areas of the unit probability triangle and hence concluded that this limits the capacity of the theory to serve as an adequate alternative descriptive theory of choice under uncertainty.

Machina's fanning-out or type II hypothesis was also tested, together with other alternative theories, by Camerer (1989). He concludes that "indifference curves seem to fan out in most portions of the triangle diagrams, and fan in in some portions; the fanning-out hypothesis is sometimes violated" (Camerer 1989:94). This and other evidence presented above, limits the adequacy of Machina's type II hypothesis in serving as an alternative general theory of choice under uncertainty.

6.4.1.2 Rank-dependent expected utility model

The most common generalized expected utility theory is the rank-dependent expected utility model (RDEU). This model is one of the decision weighting models with more sophisticated probability transformation designed to ensure the monotonicity of the value function. The most prominent rank-dependent utility model is Quiggin's (1982) anticipated utility model which he later referred to as the rank-dependent expected utility model (RDEU) (Quiggin 1993).

The main feature of this model is the modification of the expected utility model by dropping the independence axiom or the linearity in probabilities assumption. The RDEU is a generalization of EU theory that preserves the standard properties of continuity, transitivity and first order stochastic dominance. It is considered to be the only natural generalization of the EU theory that can incorporate notions such as probability weighting (Quiggin 1993:53). Kahneman and Tversky (1979) and authors of other related studies maintain that any approach that uses probability weighting cannot avoid the violation of the dominance axiom of the EU theory. Quiggin (1993:53) argues, however, that the rank-dependent expected utility approach offers the only extension of two outcome probability weights consistent with the first order stochastic dominance. He further argues that the RDEU model offers an alternative notion of risk aversion and risk preference that depends on the existence of risk instead of the nature of the outcome space. This concept is related to Allais's (1979b) approach, according to which the cardinal utility of wealth is assumed to be known under certainty, while risk aversion refers to risk in utilities. Quiggin claims that one of the main reasons for the development of the RDEU theory is the desire to resolve the paradox in EU theory of simultaneous gambling and insurance, also a concern of Friedman and Savage (1948). The main interest in this regard is a situation where investor risk aversion does not apply. An example of this is "a speculative asset that will probably yield a zero return, but which will yield a high return if the associated venture is successful. An RDEU-maximizing investor may purchase such an asset even if expected profits are negative" (Quiggin 1993:96). It is also contended that the RDEU model is based on the modification of Savage's sure-thing argument. In this regard, it is argued that it is natural to code prospects in a rank-ordered fashion and that the sure-thing principle can be applied only to states which are both equally probable and equally ranked, which gives rise to the notion of ordinal independence (Quiggin 1993:55).

The development of the RDEU model was motivated by the evolution of models that use probability weighting approaches. Like these models, RDEU is based on the probability weighting approaches with the weighting function q: $[0, 1] \rightarrow [0, 1]$, where the weighting function is applied not to the probabilities of individual events, but to the cumulative distribution function (Quiggin 1993:59). The RDEU model is given by:

$$V(\{x; p\}) = \sum_{i=1}^{n} U(x_i)h_i(p), \qquad (6.6)$$

where

$$h_i(p) = q(\sum_{j=1}^{i} p_j) - q(\sum_{j=1}^{i-1} p_j) \qquad (6.6')$$

is the weighting function.

For the two outcome case the weighting vector is

$$h(p) = (q(p_1), q^*(p_2) \qquad (6.6'')$$

where q is the weight placed on a worse outcome and q^* is the weight on the better outcome and the weighting function is said to be symmetric when $q(p) = q^*(p)$ for all p. Moreover, since q is applied to the cumulative distribution function, the weight $h_i(p)$ depends not merely on p_i but on all the elements of p. This means that it is possible to have a different weight for two events with the same objective probability. It is this basic feature which distinguishes RDEU from other related theories such as the prospect theory of Kahneman and Tversky (1979) and Tversky and Kahneman (1986), whose application of probability weights leads to the violation of the dominance axiom (Quiggin 1993:58). However, Kahneman and Tversky's (1992) cumulative prospect theory recognises this fact and uses a cumulative functional in the weighting process, solving any problems that give rise to violations of dominance.

In a related development, Yaari (1987) analyzed the relationship between the EU theory and the special case of RDEU in which the utility function is linear. In Yaari's dual model, dual to expected utility theory, the linearity of the utility function (utility is linear in wealth) means that risk attitudes are completely separated from the declining marginal utility of wealth. Yaari (1987:95) argues that under expected utility theory, risk aversion and diminishing marginal utility of wealth are considered to be synonymous, while in actual fact they are horses of different colours. In contrast to the RDEU, Yaari's dual model uses the decumulative distribution function $G(x) = (1-F(x))$ which leads to a function that is linear in outcomes and nonlinear in probabilities. This function is presented as: $V(G) = \int_x x dq^*(G(x))$. This property of linearity in outcomes implies that Yaari's dual model is unlikely to serve as an adequate description of individual choice under uncertainty. The model is inconsistent with decreasing absolute risk aversion since it implies that risk attitudes are not affected by changes in levels of wealth. This model displays both constant absolute risk aversion and constant relative risk aversion, neither of which is possible in EU theory or consistent in practice (Quiggin 1993:66).

Nevertheless, Yaari's dual interpretation and the use of the linear utility approach may be useful in analyzing agents' wealth allocation behaviour as presented, for instance, in Yaari's example of portfolio choice. Considering the simple model with one risky and one riskfree asset, where the return on the riskfree asset is assumed to be zero and that on the risky asset to be positive, according to EU theory all agents, whether risk averse or not, will allocate some portion of their wealth to the risky asset. This is because agents are essentially risk neutral in a neighbourhood of certainty. As the rate of return on the risky asset rises, the amount of wealth allocated to it increases until all the wealth is allocated to the risky asset. However, in Yaari's model the agent is presented as allocating wealth only in an all or nothing fashion (Quiggin 1993:66).

The general RDEU model incorporates both the EU theory and the dual model of Yaari as special cases because it is capable of accommodating the behaviour found in the usual EU theory paradoxes as well as in Yaari's dual paradoxes (Quiggin 1993:67). Another work related to RDEU is that of Schmeidler (1989) and Gilboa (1987) which focuses on the analysis of decision problems under ambiguity in that the probabilities of the different states of nature are unknown. Their analysis is based on the notion of nonadditive measures or capacities developed originally by Choquet (1953). Their work is motivated by the discovery of the Ellsberg paradox, where individuals are faced in their decision-making with subjective nonadditive probabilities. The main difference between the Schmeidler-Gilboa model and the RDEU is that of interpretation. RDEU is based on the assumption of known objective probabilities that are transformed to yield nonadditive decision weights, whereas in the Schiemdler-Gilboa model, probabilities are initially unknown, and the decision weights are interpreted as nonadditive subjective probabilities (Quiggin 1993:72). Segal (1987:145), however, argues that for RDEU theory to explain the empirical paradoxes of the EU theory, in particular the Allais paradox or common consequence effects and common ratio effects, it needs to assume that the decision weighting function is concave.

6.4.1.2.1 Empirical tests of the RDEU

Battalio et al. (1990), Birnbaum and Navarrete (1998) and Birnbaum (2005) have carried out empirical test of the rank-dependent models. Battalio et al. (1990:46) state that, in the empirical test of the RDEU model, they observed Allais type common ratio violation of the expected utility theory under conditions where RDEU argues that they should not be observed.

The objective of Birnbaum and Navarrete (1998) was to test the alternative nonexpected utility theories for their capacity to explain the violations of stochastic dominance and cumulative independence, which they assume in place of the linearity in probability assumption of expected utility theory. To test for the violations of stochastic dominance, Birnbaum and Navarrete (1998:49-51) first investigated the recipe for the violation of the stochastic dominance by using the concepts of outcome monotonicity, transitivity, coalescing and probability branch independence. The first three concepts are assumed or implied by the RDEU.

Transitivity is explained in chapter two of the present study. The other three concepts are explained below. Outcome monotonicity implies that increasing the value of one outcome while holding everything else constant in the gamble, should improve the gamble. Accordingly, Birnbaum and Navarrete (1998:50) state that for three outcome gambles with outcomes selected such that $x^+ \succ$ x, $y^+ \succ$ y and $z^+ \succ$ z, outcome monotonicity requires:

$$(x^+, p; y, q; z, r) \succ \; G = (x, p; y, q; z, r) \qquad (6.7)$$

$$(x, p; y^+, q; z, r) \succ \text{G.} \qquad (6.8)$$

$$(x, p; y, q; z^+, r) \succ \text{G} \qquad (6.9)$$

The above three relations indicate that improving the gamble by the value of one outcome without changing any of the other elements in the gamble makes that gamble preferable to a similar one whose outcomes are all kept unchanged.

The third cause for the violation of stochastic dominance is coalescing. According to Birnbaum and Navarrete (1998:51), coalescing implies that if two or more outcomes within a gamble have the same value, it is possible to combine them by adding their probabilities such that the individual becomes indifferent between the two options. Thus, for two outcome gambles coalescing requires:

$$(x, p; x, q; z, r) \sim (x, p+q; z, r) \qquad (6.10)$$

$$(y, p; y, q; z, r) \sim (y, p+q; z, r) \qquad (6.11)$$

In other words, coalescing refers to the assumption that if the gamble has two (probability-consequence) branches yielding an identical consequence, those branches can be combined by adding their probabilities without affecting the utility. For example, if G = ($100, 0.2; $100, 0.2; $0, 0.6) and G′ = ($100, 0.4; $0, 0.6) then the decision maker would be indifferent between G and G′ (Birnbaum 2005:265). Coalescing is implied by rank-and sign-dependent utility models.

Another assumption used by Birnbaum and Navarrete (1998) and Birnbaum (2004a, 2005) in their tests of the adequacy of alternative nonexpected utility models is branch independence. This implies that if two gambles have a common outcome for a given state of the world with known probability, then the value of the outcome on that common branch should not affect the preference order induced by other components of the gamble. Branch independence is a weaker assumption that requires that outcomes be distinct and their probabilities be known, nonzero and add up to 1. In a special case, in which all elements in all gambles have the same rank, this refers to comonotonic branch independence (Birnbaum and Navarrete 1998:51).

Birnbaum (2005:265-266) argues that if people satisfy transitivity, coalescing and consequence monotonicity, then they will not violate first order stochastic dominance. Since RDEU models assume these three principles, they cannot explain systematic violations of stochastic dominance. Moreover, Birnbaum and Navarrete (1998:62) state that people systematically violate stochastic dominance in the special recipe, and rarely violate dominance in transparent tests of outcome or probability monotonicity.

Furthermore, any theory that satisfies comonotonic independence, monotonicity, transitivity and coalescing must also satisfy cumulative independence conditions (Birnbaum and Navarrete 1998:53). However, the empirical evidence suggests that RDEU violates stochastic dominance and cumulative independence. Thus, these findings represent similar paradoxes to the RDEU as the Allais paradox does to the expected utility model. These findings place in serious doubt the RDEU's claim that it is a natural generalization of the EU theory .

6.4.2. Allais's nonlinear intensity theory

The Allais (1953, 1979b) nonlinear intensity theory provides the first and most common example of systematic violation of the basic assumption of the EU theory, i.e. linearity in probability or the independence axiom (A2). Allais's original example was presented as two pairs of gamble as situations A and B versus situations C and D (Allais 1953:527). It involves the analysis of an individual's preferences for two pairs of gambles presented as follows:

a_1: {1.00 chance of $1,000,000} versus

a_2: {0.10 chance of $ 5,000,000, 0.89 chance of $1,000,000, 0.01 chance of $0}

and

a_3: {0.10 chance of $ 5000,000, 0.90 chance of $0} versus

a_4: {0.11 chance of $1000,000, 0.89 chance of $0}

There are three prospects to be chosen in this problem. These are defined, following Machina (1987), as: $\{x_1, x_2, x_3\}$ = {$0, $1,000,000, $5,000,000}. Under the expected utility hypothesis, any risk averse individual would prefer a_1 in the first case and a_4 in the second. But if the gambler is risk seeking, he would prefer a_2 in the first pair and a_3 in the second pair.

However, Allais (1953:527) found that the majority of individuals chose a_1 in the first pair and a_3 in the second pair, indicating the nonlinearity of the probabilities or violation of the independence axiom of the expected utility theory. According to the independence axiom, if $a_1 \succ a_2$ then $a_4 \succ a_3$. But individuals violate this in their decisions because the difference between the payoff probabilities between a_3 and a_4 of 0.01 is outweighed by the larger payoff for a_3. The independence axiom is considered to be the most important postulate of rationality and hence its violation is assumed to be a major challenge to the expected utility theory. The violation of independence indicates that individuals prefer certainty to higher probabilities but prefer taking risk for smaller probabilities. Therefore, the main reason for the violation of the independence axiom as described above is considered today to be the *certainty effect*, and not the framing effect. In line with the reduction principle, choices are presented in a straightforward manner rather than in special frames. On the other hand, when certainty is not involved, the failure of the independence axiom is observed in the case of a proportional rise in the probabilities of two positive returns such as 0.8/0.6 and 0.4/0.3. This phenomenon describing the failure of the independence axiom is called the *common ratio effect*.

A second phenomenon that explains the violation of the independence axiom is known as the *common consequence effect*. Machina (1987:129) argues that the Allais Paradox, initially considered to be an isolated phenomenon, is now known as a special case of a general empirical pattern called the *common consequence effect*.

Suppose individuals are asked to choose between the following prospects:

b_1 : {1.00 chance of \$1,000,000 } versus

b_2 : {0.10 chance of \$5,000,000; 0.89 chance of \$1,000,000; 0.01 chance of \$0}

and to compare

b'_1 : {0.11 chance of \$1,000,000; 0.89 chance of \$0}

b'_2 : {0.10 chance of \$5,000,000; 0.90 chance of \$0}

In line with the certainty effect described earlier, many individuals will have $b_1 \succ b_2$ and $b'_2 \succ b'_1$. If t(\$0) =1 and s = \$5,000,000 with P = 0.11 , \$0 with P = 0.89, we have

$$b_1 = (0.11)\, b_1 + (0.89)\, b_1, \qquad b_2 = (0.11)\, s + (0.89)\, b_1$$

and

$$b'_1 = (0.11)\, b_1 + (0.89)t, \qquad b'_2 = (0.11)\, s + (0.89)\, t$$

The "common consequence" in b_1 versus b_2 is \$1,000,000 while in b'_1 versus b'_2 it is \$0 or t. According to the independence axiom, the preference between b_1 and b_2, and b'_1 and

b'_2 should depend on the b_1-versus s-preference, independent of the common consequence in each case. Thus, independence requires $b_1 \succ b_2$ and $b'_1 \succ b'_2$ if $b_1 \succ s$ or $b_2 \succ b_1$ and $b'_2 \succ b'_1$ if $s \succ b_1$.

The nonlinear intensity theory assumes that the main defence of independence as a postulate of rationality involves the two stage choice process in which the final preference between b'_1 and b'_2 ought to depend on the preference $b_1 \succ b_2$. That is, for independent prospects, it is the illusion created by two stage interdependent framing of choices that makes the independence axiom seem a compelling normative principle. However, when prospects are presented in holistic form, implying that their underlying events are independent, any reasonable choice rejects the independence axiom as a normative principle. Thus, the Allais nonlinear intensity theory rejects independence as a guide to rationality since this two stage approach destroys the holistic nature of the prospects under consideration and is based on a specialised framing effect (Fishburn 1988:38-39). Allais (1953:504) argues that rationality must be defined in either of two ways. First, it may be defined in the abstract by referring to the general criterion of internal consistency which implies the coherence of desired ends and the use of appropriate means to attain them. Second, rationality can be defined experimentally by observing the actions of people who can be regarded as behaving in a rational manner. His experiment was designed to provide the second definition of rationality. However, Allais accepts the first degree stochastic dominance (which reflects the probabilistic extension of the "greater than" relation between sure outcomes) and reduction principles as the normative criteria (Fishburn 1988:39).

The assessment of the violations of the EU axioms focuses primarily on the independence axiom. According to Camerer (19989:63), the reason for the focus on the independence axiom as a source of many of the violations of the EU theory is that the variants of the other two most important axioms, i.e. ordering and continuity, are required in virtually all axiomatized theories of choice.

Allais's (1953, 1979b) nonlinear intensity theory presents preferences in risky situations using the reduction principle and weak first degree stochastic dominance, replacing the Von Neumann-Morgenstern independence axiom. In short, the three basic axioms of Allais nonlinear intensity theory are:

Assumption A1: Ordering

- A1: $\succ$ on P is a weak order

Assumption A2: Stochastic dominance

- Weak first degree stochastic dominance: if p $\succ$ q or p = q, then p $=\succ$ q.

Assumption A3

- An axiom sufficient to ensure the existence of V: P $\rightarrow$ R such that, for all p, q $\in$ P, p$\succ$ q $\Leftrightarrow$ V(p) $\succ$ V(q).

Allais accepts reduction and weak order stochastic dominance, but rejects independence or linearity in probability. He does, however, subscribe to additive subjective probability for decisions under uncertainty with a very different interpretation (Allais 1979b:469-473). Both the Allais and Ellsberg paradoxes challenge Savage's substitution principle and additivity axiom (Fishburn 1998:190).

Allais's rejection of the independence axiom of expected utility theory as a guide to rationality is opposed by several authors. Sen (1985:115) argues that rational decision models such as the Von Neumann-Morgenstern expected utility model have been successful, both in raising important questions about rational behaviour under uncertainty and in explaining real life behaviour, although with some limitations. Harsanyi (1977:16) agrees with this and adds that the Von Neumann-Morgenstern expected utility model has been successful in explaining or predicting real life human behaviour.

Moreover, Dow and Werlang (1992:197-198) argue that although the expected utility model has been questioned, there is one factor that is strongly in its favour. This is that, while the theory of consumer behaviour under certainty has only elementary empirical implications (such as homogeneity of degree zero and continuity of the demand functions, and symmetry and negative semi-definiteness of the Slutsky matrix where demand is differentiable), expected utility theory yields some stronger predictions, in particular the results on local risk neutrality and on complete insurance on actuarially fair policies. Thus, the generalization of the theory, which completely eliminates the independence axiom, would also lead to the loss of these useful predictions.

6.4.2.1 Empirical tests of the Allais paradox

The Allais paradox has been tested empirically by Birnbaum (2004a). The test investigates how significant the most famous Allais common consequence effect is in explaining the violation of the independence assumption of the expected utility model. Using empirical data, Birnbaum (2004a:96-97) tested the Allais common consequence effect and found that the empirical behaviour systematically violated the Allais independence and that this violation was so statistically significant that it forced him to reverse the model choice.

Birnbaum (2004a:105) states that, according to his empirical test, Allais paradoxes are found when games involve small sums of money, with real prices at stake and in comparison with gambles having an equal number of branches, without the use of zero consequences, and without the need for sure things. However, there is an exception to the above findings: when the assumption of coalescing is used it not only eliminates the Allais paradox, but actually reverses it. Coalescing or splitting of branches appears to provide the best explanation for common consequence paradoxes (Birnbaum 2004a:105). Therefore, at present, the Allais paradox is no longer considered to be a serious weakening of the expected utility model.

The above conclusions of the Allais nonlinear intensity theory regarding the violations of the independence axiom of the EU theory are based on the fundamental assumption that the objective of economic agents is not maximization of the expected utility. While this assumption incorporates price and revenue (benefit) uncertainty present in the gamble or in the analysis of the decision process, it excludes cost uncertainty. In the above presentation of choices in two groups of gambles, the uncertainty about the benefits the gambler expects when he chooses a certain outcome is represented by some subjective probabilities, but the associated cost uncertainties of the decisions taken are absent from the analysis. This is because the objective of the economic agent in the nonexpected utility models is not the maximization of the expected utility of profit. When the objective of the economic agent *is* the maximization of the expected utility of profit and cost uncertainty is incorporated into the analysis, the conclusions of nonexpected utility models can be fundamentally altered and the EU theory still provides the description of preferences people actually have. The I-L model developed in chapter 8 of the present study provides evidence in support of this argument.

6.4.3 Prospect theories as alternative theories of choice under risk and uncertainty

6.4.3.1 Prospect theory as analysis of decision under risk

As stated earlier, the EU theory is one of the dominant economic theories based on the assumption of rational choices. The prospect theory (PT) of Kahneman and Tversky (1979) and Tversky and Kahneman (1986) is one of the alternatives to the EU theory that provide alternative explanations to the choice behaviour of economic agents under risk and uncertainty.

The PT investigates the strength and weaknesses of some core assumptions of the expected utility theory that are developed from the principles of rational choice within the framework of the theories of decision under risk. These authors argue that the most important aspects of the descriptive limitations of the EU theory have been reflected in the failures of the most important principle of the theory which reflects rationality of choices, i.e. the failure of the independence or dominance principle. They add the failure of the invariance principle to the former. According to Tversky and Kahneman (1986: S253), the *invariance principle* states that different representations of the same choice problem should yield the same preference, i.e. the preference between options should be independent of their description. This refers to *description invariance*. Thus, if p $\succ$ q or q $\succ$ p, then depending on the framing of choices, this type of preference implies the violation of description invariance. The violation of description invariance as explained here is due to the framing effect. Another form of invariance is called *procedure invariance*. Procedure invariance states that preferences over prospects should be independent of the method used to elicit them. A well known phenomenon that reflects the failure of procedure invariance is *preference reversal* (Starmer 2004:110). Tversky and Kahneman consider the invariance principle an essential condition for normative choice theory although many writers do not openly state this principle. The principle of invariance is also closely associated with the reduction principle.

The reduction principle states that the degree of causal or stochastic dependence among the events that give rise to the probabilities across different alternatives should not affect preferences in risky choices. That is, each alternative is characterized by its probability distribution over potential outcomes and hence probability interdependence should not affect preferences (Fishburn 1988:27). Invariance can be presented in two forms. These are: (a) using different wordings of the frames but no difference between probability distributions in the two frames or in the way probabilities arise, and (b) using the differences between the way probabilities arise to induce violations of asymmetry under the reduction principle. In this case the violations of invariance can be considered as violations of reduction (Fishburn 1988:28).

According to PT, variations in the *framing* of decision problems produce systematic violations of invariance and dominance that cannot be defended on normative grounds. Framing involves posing two equivalent versions of a given problem which predictably yields different choices by different people. By doing so, framing violates the basic requirement of rationality which is called invariance. There are only two conditions that could ensure invariance of preferences. These are standard canonical representation of the same prospects such as a cumulative probability distribution of the same random variable, and the use of expected actuarial value (Tversky and Kahneman 1986: S256).

The concept of *framing* originated from an analysis of Allais's paradox by Savage (1954:101-4) and Raiffa (1968:80-86) to check whether agents apply the cancellation axiom. As stated earlier, the Allais paradox explains the situation where agents reverse their choices when the same two prospects are transformed by adding the same amount of another prospect to both of the previous ones. This is considered to be inconsistent with the expected utility theory.

The fundamental difficulty inherent in the PT and other nonexpected utility theories is that they frame decision problems in isolation from the costs those decisions entail, as the objective of the economic agents in these theories is not the maximization of profit. However, profit maximization is an important element in the analysis of the behaviour of economic agents and cannot be ignored, particularly when the economic decision in question involves investment of resources.

However, the prospect theory continues to defend its hypothesis using three basic elements that are meant to provide alternative explanations for the behaviour of economic agents that are considered to be abnormal under the assumptions of the expected utility theory. These are: (a) loss aversion, (b) reflection effects and, (c) nonlinear weighting of probability.

6.4.3.1.1 Loss Aversion

Loss aversion refers to a situation where losses generally loom larger than the corresponding gains. That is, the loss of utility associated with giving up valued goods looms greater than the utility gain associated with receiving the goods (Kahneman and Tversky 2000:481; Tversky and Kahneman 2000:145). In terms of the prospect theory

$\sum \pi(p_i)v(x_i - r)$, the value function v(x-r) exhibits *loss aversion* if the value of loss -x is greater in magnitude than the value of an equal sized gain (i.e. -v(-x) > v(x) for x>0). This implies that the value function is steeper in the negative domain than in the positive domain.

The implications of loss aversion are that investors make decisions based on change in wealth, x, rather than on total terminal wealth, w + x. In PT preferences are affected by gains and losses or changes in wealth rather than the state of wealth as implied by the rational agent model. Loss aversion is the main driving factor that leads to choices different from those presumed by the expected utility model. For problems framed in certain (gain) and risky or uncertain (loss) formats, PT suggests that agents always choose outcomes with certainty (gain) rather than risky outcomes or those that lead to loss even if the decision problems are identical. Agents are always risk averse when decisions involve gains and risk seeking when decisions involve losses (Tversky and Kahneman 1986:S255).

In decision under risk, loss aversion leads to reluctance to accept gambles which have even chances of gains and losses unless the gains are twice as big as the losses. On the other hand, in decisions under certainty, loss aversion leads to a systematic discrepancy in the assessment of advantages and disadvantages. Thus, Kahneman and Tversky (2000:481) argue that when an option is compared to the reference point, the comparison is presented in terms of advantages and disadvantages of that option. When the reference point is the status quo, the preservation of that status quo becomes an option leading to loss aversion. Thus, Tversky and Kahneman (2000:147) assert that loss aversion implies status quo bias.

As discussed earlier, according to PT, losses are weighted substantially more than objectively commensurate gains in the evaluations of various prospects and trade. The implication of this asymmetry is that if goods are evaluated as a loss when they are given up and as a gain when they are acquired, loss aversion will, on average, induce a higher dollar value for owners than for potential buyers, leading to a fall in the set of mutually beneficial trades (Kahneman et al. 2004:57). In other words, when preferences are not independent of endowments, the existence of loss aversion produces inertia in the economy because potential traders are more reluctant to trade than is normally assumed. However, this does not mean that in such cases there is no Pareto optimal trade possible, only that fewer mutually advantageous exchanges are possible (Kahneman et al. 2004:71).

According to PT, loss aversion can also explain other phenomena, such as the equity premium puzzle where investors are not averse to variability of returns (risks) but are averse to loss (the chance that the returns are negative). The result is that they require very high returns for stocks compared to bonds, and other behaviours such as asymmetries in consumer reactions to price increases and decreases, and the insensitivity of consumption to bad income news are exhibited (Camerer 2004:150-158).

In contrast to PT's contention, in much of the economics literature dealing with risk and uncertainty, including the Von Neumann-Morgenstern expected utility theory and some nonexpected utility theories, the carriers of value consist of final, net asset positions (Battalio et al.1990:29). However, in PT, the primary carriers of value are considered to be changes in wealth rather than final wealth. Tversky and Kahneman's (1986) argument of risk seeking over losses and risk aversion over gains is based on their different assumptions regarding carriers of value. When carriers of value are assumed to be changes in wealth, choices over losses will generally be risk seeking and choices over gains will be risk averse. In contrast to the prospect theory, most economic studies postulate a predominant pattern of risk aversion over both gains and losses with no changes in preferences whether questions are posed in terms of gains or losses, as long as the net balance position remains the same (Battalio et al. 1990:29).

6.4.3.1.2 The Reflection effect

Using empirical analysis Kahneman and Tversky (2000:22-23) compared preferences between positive and negative prospects. They argue that the preference between negative prospects is the mirror image of the preference between positive prospects and this reflection of the prospects around 0 reverses the preference order. They call this phenomenon the *reflection effect.*

In terms of prospect theory, the value function $v(x\text{-}r)$ exhibits diminishing marginal sensitivity to deviations from the reference point r creating a *reflection effect* because v(x-r) is convex for losses and concave for gains (i.e. $v''(x\text{-}r) > 0$ for $x < r$ and $v''(x\text{-}r) < 0$ for $x > r$).

The first implication of the *reflection effect* is that investors are risk averse when considering prospects with positive outcomes but risk seeking when considering prospects with only negative outcomes. Thus, investor behaviour is characterized by the maximization of the expected value of an S-shaped value function, $v(x\text{-}r)$, which is convex for negative x but concave for positive x (Levy at al. 2000:199-200).

The second implication of the reflection effect is the difference in preference patterns between positive prospects and negative prospects. For positive prospects, the certainty effect leads to a risk averse preference for a sure gain over a large gain which is probable while for negative prospects the same effect leads to a risk seeking preference for a loss which is probable over a smaller loss that is certain (Kahneman and Tversky 2000:23).

Finally, the reflection effect eliminates aversion to uncertainty as an explanation of the certainty effect. The PT also suggests that agents do not prefer sure loss to probable loss. This implies that certainty is not generally desirable. Instead, certainty increases aversiveness of losses as well as desirability of gains (Kahneman and Tversky 2000:23).

6.4.3.1.3 Nonlinear Weighting of Probability

In prospect theory, decisions that yield a risky outcome x_i with probability p_i are valued by $\sum \pi(p_i)v(x_i - r)$, where $\pi(p)$ is a function that weighs probability nonlinearly, overweighting probability below 0.3 or so and underweighting larger probabilities.

Since PT uses the nonlinear weighting of probability in describing preferences, these preferences are represented by nonlinear indifference curves shown in figure 6.3. Because the probability weighting function, $\pi(p)$, is discontinuous near 0 and 1 the indifference curves are discontinuous along the edges, as shown in figure 6.3. Moreover, the theory assumes that people will overweight the small probabilities of winning X_L near the hypotenuse or X_H near the lower edge, and they will much prefer points just inside those edges to near points that are exactly on the edge as a result of which indifference curves on the hypotenuse and lower edge will appear very steep and flat respectively (Camerer 1989:75).

The overweighting of outcomes that can be obtained with certainty, relative to outcomes that are merely probable, gives rise to violations of invariance and dominance. This is because of the certainty effect which shows that the reduction in the probability of winning from certainty to a probable level has a greater effect than the corresponding reduction in probabilities of already risky alternatives.

The implication of the nonlinear weighting of probability is that investors systematically distort probabilities and base their decisions on subjective rather than on objective probabilities, in such a way that low probabilities are subjectively overestimated (Levy et al. 2000:199-200). However, the nonlinear weighting of probability follows the same

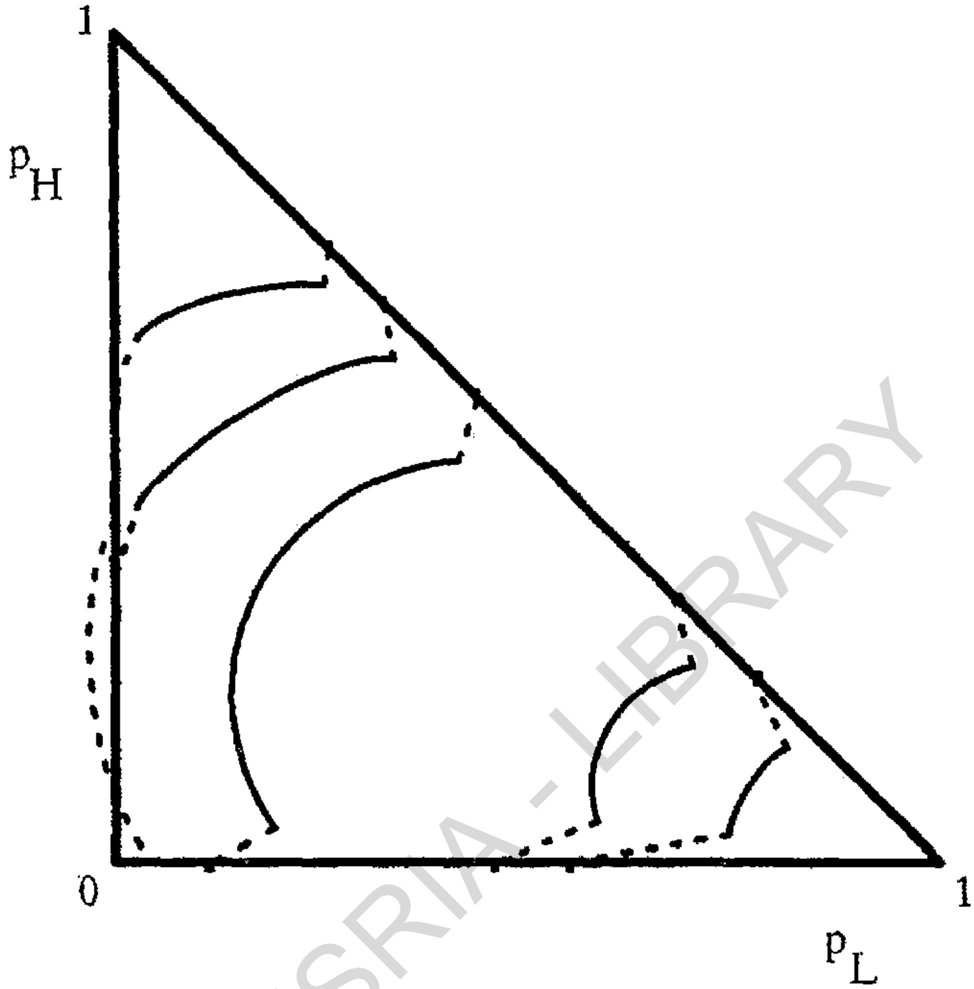

Figure 6.3 Indifference curves suggested by the prospect theory in a unit probability triangle

argument forwarded by Allais (1953, 1979b) and cannot be considered as unique to the prospect theory. Moreover, it has been shown that probability distortion does not have a clear, one way effect on asset pricing and hence on investment (Levy et al. 2000:225).

6.4.3.1.4 Empirical tests of prospect theory

Battalio et al.'s (1990:33) empirical test of the loss aversion hypothesis of PT yielded results which support the general economic postulate of risk aversion. More specifically, they observed risk loving for questions with gains involving all positive outcomes, and a willingness on the part of subjects questioned to accept fair gambles involving both gains and losses. That is, there is strong risk seeking in some questions with strictly positive payoffs where prospect theory predicts risk aversion. These findings contradict the Tversky-Kahneman hypothesis of loss aversion and consequently their theory's predictions regarding the conditions under which the gains function is concave and the loss function relatively steeper than the gains function. This is also related to the prospect theory's reflection effect.

Battalio et al. (1990) also tested PT's reflection effect empirically. The reflection effect is the central argument of the prospect theory in terms of its adequacy as an alternative descriptive theory of choice. However, the empirical evidence revealed that the reflection effect is not a universal phenomenon. Battalio et al. (1990:45) show that when one of the alternatives has a positive probability of a zero payoff, prospect theory's reflection effects (risk seeking for high probability losses and risk aversion over high probability gains and vice versa for low probability gains and losses) are present on the average but are not universal. This is because there is strong risk seeking over some questions with strictly positive payoffs where prospect theory predicts risk aversion. Thus, the reflection effect and the fourfold characterisation of risk attitude do not seem to be a reflection of the universal behaviour of economic agents. This undermines the prospect theory's claim that it is an adequate descriptive theory of choice under uncertainty.

6.4.3.2 Cumulative prospect theory as analysis of decisions under uncertainty

Cumulative prospect theory (CPT) represents an advance in prospect theory in that it employs cumulative instead of separable decision weights in the analysis of decisions involving both risky as well as uncertain prospects with any number of outcomes. This theory uses different weighting functions for gains and for losses, while the curvature of the value function and the weighting function are explained by diminishing sensitivity and loss aversion. These distinctions are not made in the standard cumulative model (Tversky and Kahneman 2000:44-45).

Cumulative prospect theory was not formulated by Tversky and Kahneman alone, however. Other authors formulated the theory before this, based on the Choquet expected utility. Sarin and Wakker (1998:223) argue that the cumulative prospect theory generalizes the Choquet expected utility by permitting decision weights for gains to be different from decision weights for losses. The theory was also formulated by Starmer and Sugden (1989) and Luce and Fishburn (1991) before Tverksy and Kahneman's (1992) formulation. Birnbaum (2005:266) lists Gonzalez and Wu (1999), Tversky and Wakker (1995), Wakker and Tversky (1993) and Wu and Gonzalez (1996) as other authors who also formulated cumulative prospect theory. Among the various types of cumulative prospect theories, that of Kahneman and Tversky (1992) implies that stochastic dominance must be satisfied. From this point on, we will make a distinction between the cumulative prospect theory of Tversky and Kahneman (1992) (CPT) and other cumulative prospect theories, known henceforth as "other CPTs".

Tversky and Kahneman's CPT identifies five phenomena of choice which it claims violate the standard EU theory but which can be fully explained as normal by an alternative descriptive theory of choice. These are: framing effects, nonlinear preference, source dependence, risk seeking and loss aversion, all of which the theory claims are confirmed by empirical experiments (Tversky and Kahneman 1992:298).

Source dependence is a phenomenon that explains the violation of the assumptions of the standard model of choice. People's choice to gamble on an uncertain event depends not only on the degree of uncertainty but also on its source (Tversky and Kahneman 2000:45). Heath and Tversky (1991:7) propose an alternative account of uncertainty preference which they call competence hypothesis, suggesting that individuals often prefer to bet on an event in their area of competence with a vague probability rather than a bet on a matched chance event with clear probability.

Another argument of Tversky and Kahneman's CPT is the phenomenon of risk seeking. However, risk seeking is assumed both in cumulative and non cumulative prospect theories. The expected utility theory and most related theories of individual behaviour toward risk assume that agents are usually risk averse. However, according to cumulative prospect theory there are consistent risk seeking choices in two classes of decision problems. These are: (a) individuals always prefer the small probability of winning a large prize to the expected value of that prospect, and (b) when presented with the choice of sure loss and a substantial probability of larger loss, individuals exhibit risk seeking behaviour (Tversky and Kahneman 1992:298). This is not always the case, however, because empirical evidence provided earlier showed that individuals may display risk seeking behaviour for substantial amounts of gains in contradiction to the prediction of the prospect theory.

Tversky and Kahneman's CPT suggests two major modifications of the standard theory, which maintains that the utility of an uncertain prospect is the sum of the utilities of the outcomes, each weighted by its probability. These modifications are: (a) the carriers of value are gains and losses, not final assets; and (b) the value of each outcome is multiplied by a decision weight, but not an additive probability (Tversky and Kahneman 1992:299). Empirical evidence has shown that the prospect theory's hypothesis of loss aversion is the direct result of its assumption regarding the carriers of value. If carriers of value are assumed to be gains and losses, instead of the final or net asset positions as assumed by the expected utility and many other expected utility and nonexpected utility theories, agents show a tendency to seek risk for losses and avoid risk for gains. This seriously limits the universal applicability of the cumulative prospect theory as an alternative theory of choice under uncertainty.

While the original prospect theory uses monotonic transformation of outcome probabilities as in other models, this presents a problem in that it does not satisfy the requirement of stochastic dominance and cannot be extended to outcomes with a large number of prospects. CPT solves this problem by transforming the entire cumulative distribution function or by applying the cumulative function separately to gains and losses, thereby extending prospect theory to uncertain as well as to risky prospects with any number of outcomes (Tversky and Kahneman 2000:47).

The Tversky and Kahneman CPT considers a finite set of states of nature S where only one state obtains and this is unknown to the decision maker. The outcomes are represented by the set of consequences called X, which includes a neutral outcome denoted, 0. An uncertain prospect f is a function from S into X that assigns to each state $s \in S$ an outcome $f(s) = x$ in X where a prospect f can be positive, negative and mixed and is represented as a sequence of pairs (x_i, A_i) which yields x_i when A_i occurs. The theory represents the value function for a risky and uncertain prospect f using the concept of capacity in the Choquet (1955) utility function for both negative and positive prospects. Thus, for a strictly increasing value function v: X $\rightarrow$ Re, satisfying $v(x_0) = v(0) = 0$, such that for $f = (x_i, A_i)$, $-m \leq i \leq n$, and assuming that $\pi_i = \pi_i^+$ if $i \geq 0$ and $\pi_i = \pi_i^-$ if $i < 0$, and where π_i s are functions of capacities Ws, then the Tversky and Kahneman's CPT becomes:

$$V(f) = \sum_{i=-m}^{n} \pi_i v(x_i) \qquad (6.12)$$

For both negative and positive outcomes the decision weights π_i s add up to 1, while for mixed prospects the sum of the decision weights can be either greater or smaller than one, because the decision weights for gains and for losses are defined by separate capacities (Tversky and Kahneman 2000:48). The above formulation of Tversky and Kahneman's CPT is similar to the Choquet expected utility we reviewed in chapter four of the present study. According to Sarin and Wakker (1998:227), the only difference between the two is that the capacity for gains may be different from the capacity for losses in prospect theory formulation.

The formulation of the CPT model based on an inverse-S shaped value function has been criticized by several authors. Birnbaum (2004a:104) argues that the CPT inverse-S relationship between certainty equivalents and probability in binary gambles is taken as a cumulative probability weighting function, which can be tested by exploring its implied violations of restricted branch independence and distribution independence. The CPT model with its inverse-S weighting function implies a violation of both these properties.

Moreover, as in PT, the analysis of Tversky and Kahneman's CPT assumes that carriers of value are changes in wealth instead of net asset positions, in contrast to the assumptions of most expected and nonexpected utility theories. Moreover, both PT and Tversky and Kahneman's CPT rely on the fundamental assumption that economic agents are not maximizers of expected utility of profit but of value. This assumption excludes consideration of costs of decision-making and hence the associated cost uncertainty. This limits the universal applicability of the prospect theories as alternative theories of choice under uncertainty.

One feature of Tversky and Kahneman's CPT that distinguishes it from the earlier PT is that it clearly differentiates between risk and uncertainty. In the valuation of uncertainty there are two natural boundaries. These are certainty and impossibility, corresponding to the end points of the certainty scale. For uncertain prospects, the principle of diminishing sensitivity yields subadditivity for very unlikely events and superadditivity near certainty, corresponding to a weighting function that is concave near 0 and convex near 1 (Tversky and Kahneman 1992:303).

6.4.3.2.1 Empirical tests of cumulative prospect theory

Recent and more detailed tests of selected alternative nonexpected utility models, including Tversky and Kahneman's CPT, were carried out by Birnbaum and Navarrete (1998) and Birnbaum (2004a, 2005). These authors tested Tversky and Kahneman's CPT for the violation of stochastic dominance and cumulative independence by using the four recipes for violations, i.e. transitivity, consequence monotonicity, coalescing and probability branch independence.

According to Birnbaum (2005:265), stochastic dominance is the relation between non-identical gambles such as $P(x > t/G) \geq p(x > t/F)$ for all t. This relation is also known as first order stochastic dominance, distinct from other relations described as types of stochastic dominance. He further argues that the assertion that choices satisfy first order stochastic dominance means that if G dominates F, then G *is preferred to* F and hence we need to decide whether the observed violations of first order stochastic dominance are due to "chance errors" or are "real" (Birnbaum 2005:265). The violation of stochastic dominance occurs when G dominates F but F is preferred to G.

Tversky and Kahneman's CPT claims that it satisfies stochastic dominance. Thus, unlike PT, CPT attempts to explain the violation of stochastic dominance for nontransparent prospects. However, empirical evidence has shown that CPT itself violates stochastic dominance and negates its own assumption (Birnbaum and Navarrete 1998:62; Birnbaum 2004a:101).

Birnbaum (2005:283) argues that the CPT of Tversky and Kahneman cannot account for systematic violation of stochastic dominance. Of the 32 choices between three branch gambles presented by Birnbaum (2005), the CPT model of Tversky and Kahneman (1992) made 23 erroneous predictions of the model choice and was wrong in all tests with mixed gambles. Birnbaum (2005:265-266) argues that if choices satisfy transitivity, coalescing and consequence monotonicity, then they will not violate first order stochastic dominance and that since CPTs assume these three principles, they cannot explain systematic violations of stochastic dominance.

Birnbaum (2005:283) goes on to argue that, considering the evidence presented and the growing mass of evidence contradicting cumulative prospect theory (CPT), it is time to set it aside and move on to evaluate models that can describe the empirical phenomena of risky decision-making.

As can be seen from the preceding analysis, a growing body of literature has proved that prospect theories, both cumulative and noncumulative, are not capable of describing the empirical choice problems under risk and uncertainty, as their authors initially claimed. Furthermore, any theory that satisfies comonotonic independence, monotonicity, transitivity and coalescing, must also satisfy cumulative independence conditions (Birnbaum and Navarrete 1998:53). However, the empirical evidence suggests that CPT violates both stochastic dominance and cumulative independence. Thus, these findings represent similar paradoxes to the CPT as the Allais paradox does to the expected utility model.

The cumulative empirical evidence that shows the violation of cumulative prospect theories has now reached a critical threshold where these theories must be questioned as descriptive of human decision-making. The weight of evidence against CPT now exceeds the case against EU theory reviewed by Kahneman and Tversky (1979) (Birnbaum 2004a:100). CPT has performed so poorly in several empirical experiments that suggestions have been made for it to be replaced by more accurate models. These models do not use more parameters but account for seven different results that refute this class of theories. According to Birnbaum (2004a:100-1001), it has been empirically proved that CPT (with any choice of functions and parameters) cannot account for violation of coalescing (event splitting effect), violation of stochastic dominance, violation of lower cumulative independence, violation of upper cumulative independence, or violation of 3-branch tail independence, nor violations of branch independence and distribution independence in the opposite direction from that observed when CPT tries to accommodate the Allais paradox. Each of these seven phenomena has been well established by systematic experiment and most have been replicated in various studies. Thus, the CPT's claim to be an alternative descriptive theory of choice under uncertainty has been empirically refuted.

6.4.4 Nontransitive preference theories: the regret theory

Regret theory is one of the alternative nonexpected utility formulations to EU theory that focuses on non-transitive pairwise choice between alternatives under uncertainty. The most prominent formulation is the regret theory model of Loomes and Sugden (1982, 1987). The initial Loomes and Sugden (1982) model was confined to choice problems involving a set of only two actions. However, the later Loomes and Sugden (1987) model developed a version of regret theory that applies to any finite set of acts.

The starting point of regret theory is psychological intuition related to regret and rejoice where preferences are defined for actions rather than prospects. In regret theory, consequences are assumed to take the form of monetary payments and preferences over outcomes reflect the usual preference for more money over less. Consequences are a result of the interactions between individual choices and the occurrence of state of the world denoted by s, where s is the set of S possible states of the world occurring with probability π_s. Consider the situation where there are two choices, A_i and A_j, for an individual. If the individual chooses an action A_i in preference to A_j and that the state of the world s occurs then, he receives x_{is}. On the other hand, if he had chosen A_j he would have received x_{js}. The fundamental intuition behind the regret theory is that "having x_{is} and missing out on x_{js}" is a composite experience and the utility the individual derives from this experience depends on x_{is} as well as on x_{js} (Loomes and Sugden 1987:272).

According to the regret theory the individual chooses so as to maximize utility denoted by $m(x_{is}, x_{js})$. The regret theory model can be expressed as

$$m^*(x_{is}, x_{js}) = m(x_{is}, x_{js}) - m(x_{js}, x_{is}) \qquad (6.13)$$

which refers to the net gain, in utility terms, arrived at, taking into account the regret and rejoicing in choosing A_i instead of A_j in the event that state s occurs. The function m and m^* are increasing in their first argument and nonincreasing in their second. This is related to the two conditions of the regret theory called ordering of pure consequences (OPC) and increasingness (I). According to OPC if $g \succ h$, the experience of "having x_g and missing out on x_h" is strictly more pleasurable the more preferred x_g is, and weakly less pleasurable, the more preferred x_h is. Moreover, assuming I, the function $m(x_g, x_h)$ can be written as a function of $f(x_g)$ and $f(x_h)$, increasing in $f(x_g)$ (Loomes and Sugden 1987:273).

In this theory, A_i is chosen if the expectation of m^* given by

$$E[m^*(x_{is}, x_{js})] = \sum_{s=1}^{S} \pi_s m^*(x_{is}, x_{js}) \tag{6.14}$$

is positive. Quiggin (1994:155) contends that the two functional forms above guarantee that regret theory will yield statewise stochastic dominance for independent prospects, which means that if one prospect yields a better outcome in every state it will be chosen. Using independent three-consequence prospects x_1, x_2 and x_3 with the set of all probability mix of p, q, r, Loomes and Sugden (1987:277-278) showed that if OPC holds and an additional condition of convexity is imposed, the regret theory generates a family of upward sloping indifference lines, which intersect at a single point and which have Machina's "fanning-out" property. They further argue that regret theory generates preferences over three-consequence actions that are similar to those generated by Chew and MacCrimmon's theory, as well as that of Machina, both of which generate transitive preference ordering over all sets of prospects. However, for prospects with four or more consequences, prospects regret theory can generate nontransitive preferences over statistically independent prospects (Loomes and Sugden 1987:278). Loomes and Sugden (1986:12) argue that, because regret theory makes comparisons across actions but within states of the world, it can predict the violations of the transitivity axiom but not the violation of the sure-thing principle.

If the assumption of independent prospects is relaxed, regret theory can yield situations of preference cycles and preference reversal. Moreover, according to Loomes and

Sugden (1987:280-81), since regret theory can generate strict preference between stochastically equivalent actions, it cannot be formulated in terms of preferences over prospects that stochastically dominate one another. Thus the theory does not generally satisfy the condition of stochastic dominant preference. Instead, regret theory has the property of statewise stochastic dominance preference which implies that $A_i \succ A_j$ is true whenever A_i statewise dominates A_j.

The fact that pairwise choices generated by regret theory are intransitive means that an individual with regret theoretic preferences is open to manipulation by a "money pump" (endless chain of trade) using successive pairwise choices over three or more prospects. However, Loomes and Sugden (1987:285) argue that in such situations individuals with regret theoretic preferences will take the entire choice set into account and will not therefore be vulnerable to a money pump. Moreover, it is assumed that choices should not be influenced by the availability of alternatives which are statewise dominated (Quiggin 1994:154). In regret theory, to remedy the problems of cyclical preferences and money pump arguments, the condition of irrelevance of statewise dominated alternative (ISDA), which is a weak rationality criterion for the model, must hold. It is often argued that an individual with cyclical preferences can be trapped in an endless chain of trade. However, because of the transitivity of the relation indicating strict preference (if he chooses one, he has to reject others and the feasible set for this would only be limited), it is not logically possible for an individual to be induced to go round the cycles of trading more than once. Thus, the problem of money pumping cannot prevail. Loomes and Sugden (1987:285-286) do argue, though, that this result depends on the retrospective nature of regret and rejoicing and may not apply to other theories of choice in which preferences are set specific. They add that regret and rejoice are *ex post* experiences in that they do not occur at the moment of choice but only after the relevant uncertainty has been resolved. At the moment of choice these experiences are only anticipated in terms of a psychological response to having chosen one action and rejected others in each state of the world in which, *ex post*, the set of rejected actions contains all actions during the period before the uncertainty was resolved and the order in which the actions were rejected seems to have no particular significance.

6.4.4.1 Empirical tests of regret theory

Regret theory is the psychological assessment of the behaviour of the decision maker in terms of the pleasure and pain involved in making one desirable choice instead of another. Battalio et al. (1990:38-39) empirically examined regret-rejoicing effects involving the first analysis of its kind over losses. They use different experimental procedures from Looms and Sugden's graphical representation of prospects, employing instead numerical formations. They conclude that they were unable to find any regret effect in the data (Battalio et al. 1990:46). This suggests serious limitations to the applicability of the theory as an alternative descriptive model of choice under uncertainty.

6.5 PREFERENCE REVERSAL

The phenomenon of preference reversal first observed by Lichtenstein and Slovic (1971) and Lindman (1971) explains the failure of procedure invariance that refers to the situation where preferences over prospects are independent of the method used to elicit them. In the empirical experiments of preference reversal, individuals are required to carry out two distinct tasks. The first task involves the choice between two prospects. The first prospect, often called the \$-bet prospect, offers a small chance of winning a greater prize while the second prospect, called the "P-bet", provides a good chance of winning a smaller prize. Several studies have shown that individuals tend to choose the P-bet while placing a higher value on the \$-bet, i.e. M(\$) > M(P). This is called preference reversal and presents a puzzle for economic analysis. The ordering revealed appears to depend on the elicitation procedure (Starmer 2004:111).

Various explanations have been offered for preference reversal. According to Slovic (1995:366), preference reversal implies either the intransitivity of preference relations or the failure of procedure invariance, or both. Consider the \$-bet (the low probability gamble) and the P bet (the high probability gamble). Preference reversal implies the existence of the following relations:

$$\text{P-bet} \succ \text{\$-bet and } C_{\text{\$-bet}} \succ C_{\text{P-bet}}$$

where $C_{\text{\$-bet}}$ and $C_{\text{P-bet}}$ refer to the cash equivalent or the minimum selling price of \$-bet and P-bet respectively. Suppose the cash equivalents for \$-bet and P-bet are X and Y respectively. According to Slovic (1995:366-267), if procedure invariance holds, an individual will be indifferent between his stated price (cash equivalent) X and the bet \$; i.e.

$$\$ \succ X \text{ iff } C_{\text{\$-bet}} > X \text{ and } C_{\text{\$-bet}} = X \text{ iff } C_{\text{\$-bet}} \approx X,$$

And, therefore, if invariance holds, preference reversal implies the following intransitive cycle:

$$C_{\text{P-bet}} \approx \text{P-bet} \succ \text{\$-bet} \approx C_{\text{\$-bet}} > C_{\text{P-bet}}$$

where the two inequalities follow from the preference reversal, and the two equivalences follow from procedure invariance.

When preference reversal is involved, the rankings observed in choice and valuation tasks cannot be explained with reference to a single preference ordering. However, Tversky et al. (1990:204) argue that observed preference reversal cannot be adequately explained by violations of the independence axiom or transitivity. They believe that the primary cause of preference reversal is the failure of procedure invariance, associated with the nonlinear weighting of probability involving overpricing of low probability high payoff bets.

It must be kept in mind, however, that the elicitation of choice problems considered in the experiments of preference reversal and other nonexpected utility choice problems are limited only to benefit measures. They do not incorporate the costs of making such choices. This limits the application of psychological experiments to wider economic choice problems. In real life situations, every economic decision involves both benefits and costs and the associated benefit and cost uncertainties. When all sources of these uncertainties are included in the analysis of the economic decisions of agents, conclusions of psychological experiments can be substantially altered. On the other hand, the expected utility theory assumes that the basic objective of economic agents is the maximization of the expected utility, or the expected utility of profit in the case of investment decisions, which automatically brings both benefits and costs and the associated uncertainties into the elicitation of choice problems. It is for this reason that the Von Neumann-Morgenstern expected utility theory remains the most dominant theory of economic decisions under uncertainty to this day.

6.6 INEFFICIENT CHOICES AND INVESTOR IRRATIONALITY IN DECISION-MAKING

As stated earlier, economic theory is based on the assumption that investors and other economic agents are rational and efficient in that they make the best choices from the available alternatives. But some recent studies have suggested that investor behaviour is much more complicated than most economic theories would predict in the sense that investors may display inefficient and irrational behaviour.

Explicit definition of investor irrationality and inefficiency is not widely adopted in economic literature, but Levy et al. (2000: 67-68) have provided the first definition of *weak* and *strong* investor irrationality. They define *weak* investor irrationality as the deviation from expected utility maximization while *strong* investor irrationality as the more severe violation of the monotonicity axiom of EU theory, or the principle of preferring more to less. They use the CAPM model to illustrate weak investor irrationality. Over and above the key assumptions used to derive the CAPM model, e.g. that all investors are risk averse, that distribution of rates of return is normal and that there are no taxes and no transaction costs, it is also implicitly assumed that investors are rational and hence act to maximize their expected utility (they are not weakly irrational) and that they do not make mistakes and always choose their portfolios from the efficient set, i.e. they choose portfolios located on the capital market line (CML) (Levy et al. 2000:68). Since the CAPM model has been based on the expected utility frame work, the Levy et al. (2000) model uses the expected utility theory as the benchmark for the analysis of investor inefficiency and irrationality.

In a similar manner, strong investor irrationality is explained within the framework of first degree stochastic dominance (FSD). If there are two prospects A and B and if A dominates B by FSD, under the monotonicity axiom the expected utility of each investor, regardless of his preferences, is higher with A relative to B. However, some investors may choose B, either because they do not know how to choose efficiently or because their goal is not to maximize expected utility. In this case investors are strongly irrational in the sense that their behaviour contradicts expected utility theory and, in particular, the monotonicity axiom which states that investors prefer more wealth to less (Levy et al. 2000:68).

Because investors are not perfectly rational in their investment decisions, they often extrapolate bad news from stocks that have been big losers and avoid investing in them for long periods. This means that investors use past rates of return to evaluate the future rate of return and believe that the past will repeat itself. Thus, investors undervalue stocks with bad *ex post* performances. This implies that investors make systematic errors in their investment decisions and a sophisticated entrepreneur who understands this phenomenon can benefit from these errors by establishing a mutual fund that invests in these undervalued stocks and reaps abnormal profit from such market inefficiency (Levy et al. 2000:15). However, if several other investors discover this information, they will make similar rational decisions and hence the Levy et al. (2000) argument of strong investor irrationality does not hold.

Using 34 MBA students, Gordon et al. (1972) conducted laboratory tests on investment allocation between risky assets and riskless assets for various levels of wealth, as well as on the choice among various assets. Each student was allowed to make 11 investment decisions, the total number of decisions being 374. They were allowed to lend or borrow at zero interest rate with restrictions on overborrowing to avoid bankruptcy. Diversification was permitted for a single game on each trial and the chosen risky asset was mixed with a riskless asset (Gordon et al. 1972:110). Gordon et al. (1972) used the mean-variance (M-V) rule to analyze investment inefficiency. Gordon et al.(1972:117) argue that the participants in their experiment exhibited decreasing absolute risk aversion (DARA) and increasing relative risk aversion (IRRA), in that the amount invested rose while the proportion invested fell as wealth increased. They however, qualified their argument by stating that such a summary statement does not fully explain the participants' behaviour. On the other hand, if the proportion of wealth invested in risky assets is constant, there is CRRA. The following table shows the Gordon et al. outcome for maximum investment on gambles:

Table 6.1 The Gordon et al. Outcomes for Maximum Investment on Five Gamblesψ

		Wealth outcome if	
Gamble number	Amounts of money invested	win	lose
1	\$500,000	\$250,000	0
2	\$133,333	\$266,667	0
3	\$166,667	\$250,000	0
4	\$100,000	\$250,000	0
5	\$100,000	10^7	0

ψ the investor's initial wealth is \$100,000, and he borrows the difference between his investment and \$100,000 at a zero interest rate.
Note: Table taken form Levy et al. (2000: 71)

The outcome of the gambles for gambles number 1-4 is based on 50%-50% probability of winning \$1.30 and losing \$0.80 for gamble 1; \$1.50 and \$.70 for gamble 2; \$1.90 and \$ .40 for gamble 3; \$2.50 and \$0.0 for gamble 4; and 0.005 and .955 chance of winning

100 and losing 0 for gamble 5. The cash flow is then calculated as follows: for investment of $100,000 in gamble 1, the investor could end up either with $130,000 or $80,000. If the investor borrows and invests the $500,000 available, he will end up with either $500,000X1.3 = $250,000 or $500,000X0.8- $400,000 = 0. The same procedure is used for the other gambles. Gordon et al. (1972) analyzed the efficiency of choices according to Markowitz's (1952) mean-variance rule. Although, the M-V criterion is not optimal when the outcomes of the bets are discrete and the returns are not normally distributed, the above bets are very simple and their results can be interpreted in the expected utility framework with no need to assume normality of returns (Levy et al. 2000:72).

When the cumulative distribution of the five gambles is considered, distribution 2 dominates by FSD all other distributions with the exception of investment 5, because $F_2(X) \leq F_i(X)$ for all X and for i = 2,3,4. In the first four gambles, the probability of winning and losing are equal, but in gamble five, the probability of winning is 0.005 and the probability of losing is 0.995. Thus, F_2 dominates gamble 5 by second degree stochastic dominance (SSD). The frequency distribution of gambles selected was 10 for gamble 1, 20 for gamble 3 and 78 for gamble 4 which implies that 108 investment decisions out of 374 decisions were wrong decisions, i.e. about 29% of investment decisions were inferior, not only by M-V criteria, but also in terms of FSD for all utility functions (Levy et al. 2000:72-73).

It is often argued that actual investor behaviour and the investor behaviour predicted by most economic and financial models differ substantially. Investors differ in their preferences, in their investment horizons, the information at their disposal and their interpretation of this information. However, such investor heterogeneity is difficult to incorporate into any analytical framework. Levy et al. (2000:141) argues that empirical and experimental evidence suggest that most investors are characterized by constant relative risk aversion (CRRA) which implies a power (myopic) utility function. This is in contradiction to the hypothesis of prospect theory which claims that agents exhibit the behaviour of risk loving for losses and risk aversion for gains, based on nonlinear weighting of probability and the reflection effect.

Some authors still maintain that investors exhibit some kind of irrationality against the assumptions of homogenous and rational-representative-agent models, which include assumptions such as (a) there are no trading volumes, (b) there is zero autocorrelation of returns; and (c) there is no price volatility. However, some authors report empirical findings against these assumptions. These include those who indicate the presence of heavy trading volumes (Admati and Pfleiderer 1988) and those who indicate short run momentum (positive autocorrelation) and long run mean reversion (negative autocorrelation) in stock returns (Fama and French 1988; Jagdeesh and Titman 1993; Levy and Lim1998; and Shiller 1981). These puzzles are considered to be a sign of the weakening of the position of the expected utility model as the descriptive theory of choice under risk and uncertainty. These isolated anomalies do not fully explain deviations in the rational agent models, however, as various empirical evidence presented in the previous section has indicated.

Levy at al. (2000) investigated empirically the effects of various behavioural elements on the investment behaviour of agents. They used the so called Levy-Levy-Solomon (LLS) microscopic simulation model that allowed them to incorporate the experimental findings regarding the behaviour of investors and to evaluate the effects of various behavioural elements on market dynamics and asset pricing. They used two approaches: in the first approach their investors are characterized by the Von Neumann-Morgenstern expected utility with power utility function: $U(W) = W^{1-a} / 1-a$, where a is the risk aversion parameter. This form of utility function uniquely satisfies the condition of constant relative risk aversion (CRRA). This implies that investors limit their horizon to a single period when making investment decisions, although their actual horizon may be different. This property is also known as myopia or "short vision", hence the term myopic utility function for the power utility function above. For a utility function other than the power function, the investment horizon of the investor does influence the portfolio choice and in such cases, dynamic programming issues must be considered (Levy et al. 2000:147-48).

In this approach two types of investors are modelled: these are rational, informed identical (RII) investors and efficient market believers (EMB). Levy et al. analyzed their investment behaviour in the stock market with two investment alternatives: investment in risky stock and in riskless bonds. The main nonstandard assumption of the LLS microscopic simulation model is that there is a small minority of investors in the market who are uninformed about the dividend process and who believe in market efficiency. The investment decision of these investors is reduced to the optimal diversification between stocks and bonds. According to Levy et al. (2000:152), the main characteristics of such efficient market believers (EMBs) are that: they employ the *ex post* return distribution in order to estimate the *ex ante* return distribution; they may be heterogeneous in the way they perform their expectations, and; their investment decisions are aimed at expected utility maximization. However, they may deviate to some extent from optimality.

The LLS model generated some of the empirically documented market phenomena that are considered as puzzles in the expected utility model of rational agents. These are short term momentum, long term mean reversion, excess volatility, heavy trading volume, positive correlation between volume and contemporaneous absolute returns, positive correlation between volumes and lagged absolute returns and endogenous market crash (Levy et al. 2000:178-79). They further argue that this small group of investors could have a dramatic impact on the market and are not wiped out by the majority of rational investors, neither are they dominated nor do they dominate other groups by FSD and SSD; these groups coexist in equilibrium and none of them vanishes.

Therefore, in spite of the above puzzles, the findings of the LLS model indicate that the majority of investors do not make irrational decisions and those who do are a small minority. However, their suggestion that these minorities could have a dramatic effect on the market is not substantiated by any empirical evidence. Thus, the conclusions of the nonexpected utility theories regarding agent behaviour that contradict the expected utility theories are based on the behaviour of this small group of minority investors. The standard expected utility model still explains the behaviour of the majority of identical and informed investors and efficient market believers who make rational decisions.

The second approach Levy at al. used included investors who are characterized by prospect theory's S-shaped value function (rather than the utility function) and who base their decisions on subjective probability weights rather than on objective probabilities. In prospect theory, individual preferences are homogenous and the unique value function which is consistent with homogenous preference is a power function. Since prospect theory makes a distinction between positive prospects and negative prospects it suggests a two part value function of the form:

$$V(x) = \begin{cases} x^{\alpha} & \text{if } x \geq 0 \\ -\lambda(-x)^{\beta} & \text{if } x < 0 \end{cases} \qquad (6.15)$$

where x is the change in wealth, and α, β, and $-\lambda$ are constants that satisfy $0 < \alpha < 1$, $0 < \beta < 1$ and $0 < \lambda$. This value function satisfies $V'(x) > 0$ for all $x \neq 0$, $V''(x) > 0$ for $x < 0$ and $V''(x) < 0$ for $x > 0$ and hence it is S-shaped. This is because the value function is convex for negative x and concave for positive x, implying that agents are risk seeking for negative x ($x < 0$) and show risk aversion for positive prospects ($x > 0$). From their experiment, Tversky and Kahneman estimated the typical parameters of this value function as $\alpha = 0.88$, $\beta = 0.88$ and $\lambda = 2.25$ (Tversky and Kahneman 1992:311-312).

Using the LLS model with two investment alternatives of risky stocks and riskless bonds, Tversky and Kahneman's power value function implies extreme investor diversification policy regarding optimal portfolio allocation. For a given arbitrary rate of return distribution, the investor will either invest fully in bonds or in stocks and there is a sharp crossover between the state of full investment in stocks and full investment in bonds. In other words, a small change in one of the return distribution parameters can lead to a shift from full investment in one asset to full investment in the other (Levy et al. 2000:206).

Levy et al. calculate the optimal investment proportion for expected utility maximizer with a log utility function and for prospect theory expected value maximizer with $\alpha = \beta = 0.88$ and $\lambda = 2.25$. The EU optimal investment proportion is calculated by numerically maximizing the EU with $U(\bullet) = \log(\bullet)$, the

$$EU = \int U[1-p)r + pR]f(R)dR \qquad (6.16)$$

where p is the investment proportion, R is rate of return, and r is rate of interest. The PT optimal investment proportion is calculated by numerically maximizing the expected value

$$EV = W^{\alpha}{}_0 \left[\int_{R0(p)}^{\infty} \{((1-p)r + pR)^a f(R)dR\} - \lambda \int_{-\infty}^{R0(p)} (-(1-p)r + pR)^a f(R)dR\right] \qquad (6.17)$$

The theoretical analysis of the effects of PT on asset allocation, asset pricing and market dynamics yielded the following results (Levy et al. 2000: 224-225):

a) Given a risky stock with a given rate of return distribution and a riskless asset, the asset allocation decision of a PT investor characterized by Tversky and Kahneman's value function where $\alpha = \beta$ is independent of the investor's wealth. In other words, optimal investment proportion in the stock is not the function of the investor's wealth.

b) The diversification policy implied by PT is characterized by a sharp crossover between full investment in the bond to full investment in the stock.

c) A risky security with some distribution of end of period value V_1 will be priced higher by EV maximizers (relative to the pricing by EU maximizers). In contrast, if V_1 is low, EV maximizers will price the security lower than EU maximizers.

Based on the above theoretical results, the LLS microscopic simulation model led to the following findings (Levy et al. 2000: 225):

a) Although PT implies higher sensitivity to losses this does not necessarily mean that risky securities are priced lower by PT investors (relative to the pricing by EU maximizers). The relative pricing depends not only on the value function and utility function parameters but also on the nature of the risky asset. Generally, PT investors price assets with low expected returns lower than EU maximizers, while they price assets with high expected returns higher than EU maximizers.

b) Probability distortion does not have a clear, one way effect on asset pricing. Low levels of probability distortion ($0.75 < p < 1$) cause the stocks to be underpriced relative to the pricing based on the objective probabilities. A higher level of probability distortion ($p > 0.75$) leads to overpricing. Levy et al.'s experiential value of $p = 0.6$ led to overpricing by about 2%.

c) When some investors in the market employ the *ex post* return distribution in order to estimate the *ex ante* distribution (as the EMBs do), EV maximization leads to more frequent price deviations from the fundamental value. This enhances the phenomenon of heavy trading volume, excess volatility, short term momentum, long term mean reversion and the correlation of volume with contemporaneous and lagged absolute values.

The formulation of the investment decision problems in terms of the changes in wealth and the maximization of value, as in the prospect theory, instead of the maximization of the expected utility of profit does not yield any superior investment decisions. In some cases, decisions taken by the expected utility maximizers (in line with EU theory) are more efficient than those of expected value maximizers (in line with PT). The key assumptions of loss aversion and nonlinear weighting of probability in prospect theory do not seem to have clear impacts on firms' asset price and investment decisions.

6.7 SUMMARY AND CONCLUSION

Expected utility theories analyze choice problems based on the assumption of rationality. The premises of rationality are based on the belief that economic agents always make well behaved decisions when confronted with various choices. The dominant economic theory, neoclassical economics, is usually based on the postulate of instrumental (global) rationality which represents the utility maximizing rational economic agent. However, more recent theories have highlighted the limitations of instrumental rationality and have suggested alternative postulates of rationality. These are: procedural rationality, behavioural rationality based on rules and norms developed by society, and expressive (bounded) rationality involving decisions based on limited knowledge owing to uncertainty about the future. The postulate of bounded rationality laid the foundation for the analysis of rational choices under uncertainty.

The emergence of the expected utility theory goes back to 1738 when, motivated by the *St Petersburg paradox,* Daniel Bernoulli challenged the principle of expected value maximization and suggested an alternative evaluation of the decision using the expected utility maximization. Two centuries after the emergence of the expected utility theory, von Neumann and Morgenstern developed an axiomatic expected utility theory of choice under risk and uncertainty with the main axiom of independence as the postulate of rationality of decisions.

Several nonexpected utility theories have challenged the independence axiom of the EU theory as the postulate of rational decisions. These theories argue that economic agents systematically violate the independence axiom of EU theory. The Allais nonlinear intensity theory and the prospect theory are the best known examples in this respect. The Allais nonlinear intensity theory has shown that owing to *certainty effect*, agents value probability nonlinearly and hence violate the independence axiom of EU theory, challenging the rationality assumption. In addition to the nonlinear weighting of probability in prospect theory, there are two other conditions that explain the violation of the independence axiom of the EU theory. These conditions are loss aversion and the reflection effects related to the fourfold characterisation of risk attitude by cumulative prospect theory. However, more recent empirical tests of these alternative theories have shown that, while the Allais paradoxes exist in certain cases, they can be reversed or completely eliminated when coalescing or event splitting is used in the analysis. Moreover, empirical tests have shown that cumulative prospect theory not only violates its own assumption of stochastic dominance, but also several other assumptions required by any adequate descriptive theory of choice under uncertainty. Further empirical testing has shown that the prospect theory's reflection effect is present on average, but that it is not universal, thus leading to the questioning of the subcertainty of the probability weighting function. Loss aversion, on the other hand, is found to be the result of prospect theory's assumption that carriers of value are changes in wealth rather than net asset positions, as assumed by expected utility and some nonexpected utility theories. Moreover, the conclusions of the nonexpected utility models depend on their basic assumption of nonexpected utility maximization by economic agents. At present, the cumulative evidence against prospect theories is so great that some authors suggest completely abandoning these theories.

Regret theory is another alternative to EU theory that focuses on the nontransitive choices between alternatives under uncertainty. This theory is based on psychological intuition of regret and rejoicing where preferences are defined over actions rather than prospects. The objective of the agent in regret theory is to maximize the utility of the actions of choosing the prospects and the theory yields statewise stochastic dominance for independent prospects. However, the theory generates nontransitive preference over statistically independent prospects for only four or more consequence prospects and empirical testing of this theory could not observe the regret effects in the data.

Other authors have suggested that instead of trying to discard EU theory as a nondescriptive theory of choice, it is possible to generalize it by ignoring the independence axiom. One of these suggestions is Machina's *fanning-out* hypothesis that suggests that the fanning out indifference curves can generate predictions consistent with cases that show the violations of the independence axiom, as shown by the Allais paradox as well as the St Petersburg paradox. Likewise, Quiggin's rank-dependent expected utility model modifies the expected utility model by dropping the independence axiom, while preserving the standard properties of continuity, transitivity and first order stochastic dominance. Empirical tests of Machina's *fanning-out* hypothesis and rank-dependent expected utility model showed conditions leading to large systematic violations of hypothesis II in the areas of the unit probability triangle and found common ratio violation against the argument of the rank dependent expected utility model. In particular, rank dependent models have performed as poorly as prospect theories in various other empirical tests.

In general, empirical tests have indicated that none of the nonexpected utility theories reviewed in the present study organized choices consistently and further research is needed to develop alternatives to EU theory as a descriptive theory of choice under uncertainty. The expected utility theory continues to be dominant in economic analysis.

The most recent economic literature considers the possibility that some investors may make irrational and inefficient choices. The deviation from expected utility maximization is considered to be a reflection of weak irrationality while the more serious violation of the monotonicty axiom is considered to be a reflection of strong investor irrationality explained by using the framework of first degree stochastic dominance. However, in spite of certain puzzles in the findings, the Levy et al. empirical experiments show that the majority of investors make rational decisions, and only a small group of investors makes irrational ones. The impact of this small group of investors on the market process as a whole has not been empirically tested.

Based on these and other previous arguments, the I-L model developed in chapter 8 of this thesis will rely on the assumption of expected utility maximization rather than expected value maximization. However, before we present the I-L model we review in the following chapter the theories of banking firm investment decisions under uncertainty.

CHAPTER SEVEN

INVESTMENT DECISIONS OF BANKING FIRMS UNDER UNCERTAINTY

7.1 INTRODUCTION

Banking firms' investment decisions are interwoven with various other decisions. These involve the need to ensure liquidity by holding a certain amount of cash, compliance with the reserve and other requirements of the regulatory authorities, decisions regarding what proportion of the asset should be invested directly and what proportion should be provided as loans to borrowers, and coping with customer deposit uncertainty. Thus, among other things, banking firms are required to make simultaneous decisions in two segments of the market: the investment segment and the lending segment.

Traditional theories suggest that there are three main elements in banks' decision problems. These are return maximization, risk minimisation, and liquidity needed to take care of any reserve losses or loan demands that may occur. According to these models, in attempting to maximize return on portfolio, the only decision variable banks face is return on portfolio, while risk and liquidity will be taken care of by meeting the satisfaction of the regulatory authorities regarding the constraints imposed by law and customer demands (Beazer 1975:3). Recent econometric and noneconometric models of bank investment decisions have focused on the impact of greater uncertainty on the investment and lending decisions of banking firms. This approach is important as it assists in analysis of banking firms' investment and lending choice problems under various sources of uncertainty.

Theories of the firm have been dominated by a neoclassical profit maximizing economic agent that operates in an environment that is beyond its control. The neoclassical, as well as most other alternative theories of the firm, neglect the importance of risk and uncertainty in economic decisions taken by firms. Theories of banking firm investment decisions follow similar patterns. However, in real life decisions by banking firms are affected both by measurable risks and unmeasurable uncertainty. The present chapter focuses on analyses of various theories of commercial banking firm investment decisions under risk and uncertainty, following a brief overview of the theories of the firm.

The chapter is organized as follows: section 7.2 presents an overview of the theory of the firm to serve as a basis for the analysis of various theories of the banking firm. Section 7.3 presents an overview of banking firm models. Section 7.4 presents detailed assessment of the literature on characteristics of commercial bank investment and loan markets. Section 7.5 deals with bank portfolio choice under certainty, followed by an investigation of the risk return portfolio choice models in section 7.6. Section 7.7 critically assesses the literature on bank portfolio choice models under uncertainty. Section 7.8 investigates the Chavas model of portfolio selection under uncertainty. Section 7.9 deals with Yaari's comparative statics model of portfolio choice. Section 7.10 analyzes bank lending and the management of risk. Section 7.11 examines empirical econometric models of banking firm investment decisions under uncertainty. The problem of lending monitoring and costly state verification, with particular emphasis on the Gale-Hellwig model, is assessed in section 7.12, while section 7.13 presents a summary and conclusion.

7.2 THEORIES OF THE FIRM

Since the birth of modern economics in 1776, theories of the firm have been dominated by neoclassical theory that considers the firm as a profit maximizing entity that operates in an exogenously given environment that is beyond its control. This neglect has been attributed to the excessive preoccupation of economists with the study of the workings of the price system. However, different schools of thought have attempted to provide limited alternative theories of the firm. There are currently five such fully developed theories. These are: *neoclassical, principal-agent, transaction cost, evolutionary, and managerial theory* (Jacobson and Andréosso 1996:24).

The neoclassical theory of the firm neglects the complexities of the decision-making process, the problem of incomplete information and the organizational complexity of the firm. Its premise of the objective of profit maximization by the firm achieved by equating marginal cost with marginal revenue may face difficulties because a firm's objective may not necessarily be profit maximization. Firms may not make decisions based on the principles of equating marginal cost with marginal revenue.

Authors such as Elyasiani (1983) have applied the neoclassical theory of the firm to analyze the investment decisions of banking firms under uncertainty. This application of the theory implies that banking firms are considered to be productive firms with neoclassical production functions and the associated resources costs. According to this approach, the investment decisions of banking firms are modelled in a microeconomic firm theoretic context. The nonintermediary portion of banking activity, namely, the clearance output production, as well as resources costs and production function constraints, are incorporated (Elyasiani 1983:1002). The present study will model the investment decision problems of the investor-lender firms in a portfolio choice context instead of a microeconomic firm theoretic context, but includes the cost-of-funds constraint.

Managerial theory, the main proponent of which is Marris (1963, 1964), emphasizes the complex nature of the modern corporate firm and states that, for the managerially controlled firm, growth of the firm becomes a more important objective than profit maximization. Marris (1963:186) argues "…the various sources of positive managerial utility would appear to be strongly correlated with a single observable attribute of the firm, that is, its size". He argues that managers aim to increase both supply growth and demand growth where demand growth determines profit and profit determines supply growth. This theory emphasizes the separation between ownership and control that may lead to a divergence of interest between the owners and controlling managers in the same manner as in the principal-agent problem. Marris (1964:41-42) argues that managers of modern corporate firms focus on securing both sustainable and safe firm growth by ensuring both the growth of demand for the firm's products and its supply of capital. He further argues that such pressures as the need to maximize profit and conditions of market imperfections lead managers to maximize the rate of growth of the firm in which they are employed subject to a constraint imposed by the security motive (Marris 1964:42).

The principal-agent theory focuses on the relationship between ownership and control. As applied to the firm, the theory identifies the owner as the principal and the manager as the agent. In this model, there is information asymmetry between the principals and agents, where the objectives of the agent are different from that of the principal and the latter is unable to tell the agent's degree of commitment. This leads to the problem of moral hazard (Jacobson and Andréosso 1996:35-36).

The transaction cost theory originated with the seminal contribution of Coase's (1937) article "*The nature of the firm*", which argues that firms exist because of the existence of transaction costs. This theory emphasizes the central question of the need for organization in the market system. If prices determine allocation of resources for production of specific types of goods and services through the market mechanism, why do we need organizations? The answer provided by this theory is that transactions between individuals will be too difficult or expensive and inefficient, and thus a need arises for an organization to coordinate these activities and reduce transaction costs by internalising them. Demsetz (1991:159) argues that since the emergence of modern economics two centuries ago, only two theories have been written that have altered the perspective of the profession: Knight's (1921) "Risk, uncertainty and profit" and Coase's (1937) "The nature of the firm". He further argues, however, that even so the theory of the firm remains incomplete, for it fails to give greater weight to the cost of information. The transaction cost theory, however, assumes bounded rationality and in that it at least recognises the existence of imperfect knowledge or an uncertain future faced by firms.

Knight's theory focuses on risk and uncertainty and their impact on the decisions of the firm. Using the framework of perfect competition in economics, he showed how the presence of risk and uncertainty can create a wedge between actual and theoretical competition. According to Knight, firms face true or unmeasurable uncertainty instead of measurable risk in their decisions and it is this uncertainty that forms the basis of a valid theory of profit and accounts for divergence between actual and theoretical competition (Knight 1921:20). Many other theories of the firm ignore the importance of risk and uncertainty on the firm's economic decisions.

Evolutionary theory focuses on two key aspects of the firm. These are organizational routines and organizational capabilities, where the former are considered to be not only the building blocks of the latter, but serve as the genetic codes of the firm that carry the adaptive information required for competition and survival (Jacobson and Andréosso 1996:41-42). The earliest version of this theory was regarded as social Darwinist theory, as it emphasized competition and survival in a firm's growth process, to the neglect of the internal workings of the firm. Later writers of evolutionary theory such as Chandler (1992:93) argue that growth of firms is based on the ability to utilise competitive advantages created by the coordinated learnt routines in production, distribution, marketing and improving existing products and processes, a concept related to Best's (1990) notion of Schumpeterian competition of creative destruction.

With the exception of transaction cost theory and Knight's risk, uncertainty and profit, all theories of the firm briefly reviewed here have failed to recognise the importance of risk and uncertainty in firms' decision-making processes that are essential for their existence and growth.

7.3 BANKING FIRM MODELS

Over the past three decades, various models have been developed to assess the optimal behaviour of banking firms. These models focus on the production, as well as the intermediation, aspect of banking firm activities. According to Santomero (1984:584-599), the most popular banking firm models include: (a) asset allocation models which in turn are divided into: (i) reserve management models and (ii) portfolio composition models; (b) liability choice models which involve deposit modelling and capital decisions; (c) the two sided (asset-liability) models; and (d) credit rationing models.

Most of the above models do not, however, incorporate the problem of risk and uncertainty in the economic decisions of the banking firms. The present study will focus on the asset allocation problems of the investor-lender firm, with particular emphasis on the optimal choice between investment in riskfree assets and lending decisions under risk and uncertainty. Before analyzing alternative bank portfolio choice models, we will review the literature on the characteristics of investment and loan markets.

7.4 CHARACTERISTICS OF INVESTMENT AND LOAN MARKETS

7.4.1 Investment markets

Investments by commercial banks involve the purchase of various types of investment securities. Investment securities are defined as those securities with maturities exceeding one year. In the United States, two categories of securities dominate over 90% of the commercial bank investment portfolios. These are government and agency securities and municipal securities (Gup and Kolari 2005:181). Unlike discount money market instruments such as treasury bills, commercial papers and bankers acceptances, Treasury securities such as Treasury notes and bonds purchased by the banks with a maturity of one to five years, are coupon or interest bearing instruments. They serve as an alternative income generation scheme for banks. Agency, municipal and corporate securities such as bonds also form an important component of banks' investment in securities.

However, investments in securities involve various risks. These include security specific risk, portfolio risk and risk of potential inflation on investment values. Security specific risk for bonds includes default risk, price risk, and marketability risk. Default risk refers to the probability that promised payment of the principal and interest will not be made on time. In general, municipal, corporate and some agency bonds have greater credit risk than central government bonds. Price risk refers to the inverse relationship between changes in the level of interest rates and the prices of securities, in that an increase in interest rates leads to a fall in the prices of securities. Marketability risk refers to the ability of commercial banks to sell their bonds without loss of principal. If the bank has to sell investment securities to meet liquidity demands, it will find more readily available markets for central government securities than municipal or corporate securities, because the secondary market for the former types of securities is deeper and broader than for the latter group (Gup and Kolari 2005:187-196).

7.4.2 Loan markets

The loan market is much more complex than the investment market. While there are various types of loans, such as consumer credit and home mortgage loans, our main concern will be business (or commercial) loans provided to those entities that invest the borrowed funds. The lending process begins with the evaluation of loan requests.

7.4.2.1 Determinants of lending

Although the use of credit scoring models is growing in importance, particularly for consumer and mortgage loans, most business loans are evaluated using the traditional method involving the six Cs of credit. According to Gup and Kolari (2005:263) these are:

a. Character: personal characteristics of the borrower such as honesty, and attitudes of willingness and commitment to pay debts.
b. Capacity: the borrower's success in running the business, measured by smooth cash flows.
c. Capital: the financial condition of the borrower measured by its net worth.
d. Collateral: assets pledged for security in a credit transaction intended to reduce default risk by borrowers.
e. Conditions: economic conditions that are beyond the control of the firm. These include recessions, interest rate volatilities, asset price deflation and so on. These factors affect the borrower's ability to repay loans.
f. Compliance: compliance with laws and regulations. While the other conditions apply to the borrower, compliance applies to the lender. The lender must lend in accordance with the laws and regulations of the concerned authorities.

These six Cs are used by banking firms to determine the riskiness of borrower firms. Depending on the performance of the borrowers regarding each of the Cs, banks can classify the firms as very low, low, average or high risk borrowing firms. Of the six Cs, all except character and economic conditions are measurable and attainable either by the borrower or the lender. The character can be estimated based on past client information. However, economic conditions are outside the control of both the borrower and the lender.

Similarly, an investor-lender firm is uncertain about future economic conditions at the time it makes investment or lending decisions. The firm cannot form an objective probability distribution of economic conditions. It may assign subjective probability. Thus, this factor forms the core of the decision problem for an investor-lender firm under uncertainty. The lending decisions depend not only on the expected return, which is uncertain: they depend also on the uncertain costs associated with the future level of deposit liability, which is unknown at the time of the lending decision. This will be analyzed in detail under the investor-lender (I-L) firm model in the next chapter.

7.4.2.2 Types of loans

The principal lending activities of commercial banks consist of loans and leases. There are various types of loans. These include lines of credit, revolving loans, term loans and bridge loans (Gup and Kolari 2005:156). A line of credit is a loan provided for a period of one year or less and used to finance seasonal variations in inventory and accounts receivable. This loan is payable by the borrower on demand by the bank or within 90 days. Revolving loans are also provided to finance the borrower's temporary and seasonal working capital needs. However, unlike lines of credit, they are provided for medium to long term periods with a maturity of two or more years.

Term loans refer to loans used for activities of a more permanent nature such as acquiring machinery, renovating a building, refinancing debt and the like. Term loans usually have maturities of five or more years. Another important aspect of term loans is that the loan provided should not exceed the value of the asset being financed, nor should the maturity of the loan exceed the economic life of the asset, especially if the asset is being used as collateral for the loan (Gup and Kolari 2005:258). The bridge loan is short term financing that is made in anticipation of obtaining longer term financing in some or other form. It is temporary financing provided by commercial banks for firms that already have an agreed arrangement for longer term financing of their ventures. For the purpose of the present study, the types of loans do not matter, as long as these are used for business purposes. The study is not concerned with nonbusiness loans. The principal issue to be investigated is why investor-lender firms or commercial banks decide to lend their funds to other firms that invest directly in risky projects, instead of investing them in riskfree assets.

7.4.2.3 Pricing of business loans

An important issue in the loan market is the pricing of commercial loans. This refers to the determination of what interest rate to charge the borrower and how to calculate this. The interest rate may be set by using a loan pricing model to determine the interest rate that a bank should charge on a commercial loan. The interest rate that is stated in the loan agreement is usually called the nominal interest rate and this differs from the effective yield which is calculated by taking into account the payment accrual basis and the payment frequency into account (Gup and Kolari 2005:265). Effective yield is the product of the total interest paid divided by principal amount on the one hand and the number of days in a year divided by the term of the loan in days on the other. That is,

$$EY = \frac{TR}{PR} \times \frac{365}{TL} \qquad (7.1)$$

where, EY = effective yield, TR = total interest paid, PR = principal amount and TL = the term of loan in days. The payment frequency also affects the interest income earned by the bank. The bank earns more when interest is collected frequently than when it is collected annually, because of the discounting involved.

When pricing their loans, banks should take care not to overprice or underprice them. For a given level of risk, underpricing of loans leads to lower earnings, while overpricing leads to loss of customers. Many banks price commercial loans by using an index rate, i.e. using prime rate plus a markup of one or more percentage points. Other banks use the cost of borrowed funds plus a markup. Markups are meant to compensate the bank for the risk it takes in making a loan and to provide a return on its investment on loans. However, such markups may not properly account for the risk, the cost of funds or the operating expenses. An alternative is to use loan pricing models that properly account for risk, costs and return (Gup and Kolari 2005:268).

One variant of loan pricing models uses return on net funds employed. This model focuses on the required rate of return and the net income the loan should generate to allow the bank to earn that return. In this model:

Marginal cost of capital + profit goal = (loan income - loan expense)/net bank capital employed.

The left hand side of the equation shows the required rate of return. The marginal cost of capital is the rate of return required by the debt and equity investors on the newly issued funds they provide to the bank. Some authors assume that the marginal cost of capital is equal to the *weighted average cost of capital* (WACC). According to Gup and Kolari (2005:268), the weighted average cost of capital of new funds (*Kw*) is modelled as:

$$Kw = Kd\,(1-T)L + Ke\,(1-L) \qquad (7.2)$$

Where

Kd = cost of interest bearing liabilities

T = corporate tax rate

L= ratio of liabilities to assets

Ke = cost of equity

The above equation states that the cost of capital for all new funds raised is equal to the proportionate after tax cost of liabilities plus the proportionate cost of equity. The cost of equity may be calculated using the capital asset pricing model (CAPM), which states that the cost of equity capital is equal to the riskfree rate of interest plus beta times the market premium. The market premium is the difference between the expected return on the stock market and the riskfree rate of interest (Gup and Kolari 2005:279). The riskfree rate of interest is that paid on default free Treasury securities. Beta is a measure of systematic risk that is common to all stocks discussed in chapter two of the present study.

In the above model, high risk loans require larger markups than low risk loans. Thus, given information on the marginal cost of capital, profit goal, loan expense and net bank funds employed (which is the average amount of loan over its life funds provided by the borrower, net of reserve requirement for the bank), the amount of loan income required to earn the stated return can be obtained.

7.5 BANK PORTFOLIO CHOICE UNDER CERTAINTY

Wood (1975) studied commercial bank loan and investment behaviour with particular emphasis on aggregate loan and securities holdings of American commercial banks, using the assumption of certainty. He observed that for the period between the end of WWII and 1973, both the loan rate, r, and the average yield on securities, y, rose during expansions and fell during recessions. However, the cyclical variation of y exceeds that of r so the difference r-y tends to decrease in expansions and increase in recessions. This movement in rate differential during expansions would appear to cause securities, G, to become more attractive relative to loans, L, as bank investments (Wood 1975:1). However, bank holdings of securities relative to loans decline during expansions. Similarly, bank holdings of loans relative to securities fall at a time when rates of return on loans are rising relative to return on securities (Wood 1975:1). Is this a paradox? Wood argues that this is not, for the following reasons: first, banks show public spiritedness in meeting local credit demands. A sense of community responsibility compels them to accommodate the community's requirements during a period of rising credit demand. Second, it is argued that banks play a relatively passive role in the sense that "the initiative in the bank loan market lies with the borrower, not with the banker" (Galbraith 1963:20). However, Wood and Galbraith's analyses ignore the role of risk and uncertainty associated, for example, with interest rates and security yields and hence are not useful to the present study.

7.6 THE RISK-RETURN PORTFOLIO CHOICE MODELS

In a state of certainty banks' decisions will be straightforward. In such cases the banks will hold cash or maturing securities sufficient to satisfy all transaction requirements and invest the remainder of their portfolio in the highest yielding assets. However, in the real world this is not the case. Banks' operating environment is characterized by risk and uncertainty in the sense defined in chapter one of the present study. Repetitive decision problems involving risk are theoretically easier to handle than those involving uncertainty. In the former case the decision maker can obtain the expected value of the variable under consideration by using objective probability distribution of repetitive events such as, for example, seasonal variations of loan demand or deposits. In the case of the latter, however, a decision maker can only assign subjective probability at best (Beazer 1975:3-5). Thus, although there are those who postulate a major operational distinction between risk and uncertainty with respect to investment decisions, the distinction is not clear cut (Beazer 1975:7). Ellsberg's (1961) experiment has shown that decision makers prefer known (objective) probability to unknown (subjective) probability, even though those acts violate rational choice. However, the experiment does not show how agents behave if they are presented only with one or the other probability distributions or when there is the need to take both into account. There seems, therefore, to be no strong argument to make an operational distinction between objective and subjective probabilities (Beazer 1975:7).

The risk-return portfolio choice models that study the investment behaviour of firms argue that investors follow an efficient portfolio approach in making their investment decisions. The most popular of these is the Markowitz (1952, 1959) portfolio selection model, which states that investors choose a set of efficient portfolios when there is no other set which gives lower variance of return and the same expected return.

Moreover, portfolio theory states that it is possible for a securities portfolio to decrease the portfolio risk of the bank's assets, especially if the returns on securities over time are not perfectly correlated with the returns on loans. For instance, if loan rates of return are falling due to declining interest rates, but the rates of return on securities are rising due to increasing capital gains, bank earnings will become smoother over time and it will be preferable to hold both securities and loans instead of holding loans alone. This type of risk reduction is called the diversification effect (Gup and Kolari 2005:196). Gup and Kolari (2005:196) further argue, "Few banks make a conscious effort to set up a securities portfolio to reduce the total risk of bank assets due to greater emphasis on credit risk, interest rate risk and liquidity risk objectives. However, most banks consciously purchase securities to help protect themselves from potential downturns that could increase loan losses. In these circumstances securities offset to some extent falling earnings in the loan portfolio. This income stream smoothing is an important benefit of the diversification of bank asset portfolios." However, banks may not always behave in this manner.

Markowitz's (1959) individual investment choice under uncertainty established modern portfolio theory as a foundation of financial market investment behaviour. This theory suggests that investors evaluate both the expected rate of return on investment and its associated risk before making investment decisions. Assume a bank holds two assets: securities and loans. Modern portfolio theory investigates the effects of the securities portfolio j on the total risk of a bank's assets after taking into account the loan portfolio k. The expected rate of return is calculated as follows:

$$E(R_j) = \sum_{i=1}^{n} P_i R_{ji} \tag{7.3}$$

where P_i is the probability that a particular random state of nature i will occur out of n states, (n is usually considered to represent pessimistic, average and optimistic business scenarios). R_{ji} is the rate of return on the security portfolio in the ith state of nature and R_j is the expected rate of return on the security portfolio. The risk of the security portfolio is measured by σR_j. Assuming the rates of return are normally distributed the mean $E(R_j)$ and σR_j are used as measures of risk of and the return to the portfolio.

The portfolio effect measures the effect of purchasing a security portfolio while the bank already holds a loan portfolio. Gup and Kolari (2005:213) argue that purchasing a securities portfolio not only increases a bank's total expected rate of return above that earned by the loan portfolio but also reduces the bank's risk by reducing the σ of rates of return to less than that of the loan portfolio. Bank risk declines, both because the standard deviation of the securities portfolio was less than the loan portfolio's standard deviation and because the covariance term was negative. The covariance of rates of return on loans and securities affects the bank's total risk. When returns are perfectly positively correlated, risk is not reduced; when correlation is less than 1, risk is reduced where maximum variance or risk reduction occurs with perfect negative correlation (i.e. $\rho = -1$). Diversification is the reduction of variance caused by less than perfect positive correlation.

This approach has not been widely applied in the analysis of commercial bank investment and lending decisions. Most studies in this area focus only on a portion of a bank's portfolio problem, i.e. the selection of investment assets, particularly government securities. One of the reasons for the nonapplication of the efficient portfolio approach to commercial banks is the fact that the loan segment of the portfolio trades in a completely different kind of market than the investment segment (Beazer 1975:12). The concern of the present study is to suggest an alternative model that explains how banking firms make simultaneous decisions in the loan and investment market under risk and uncertainty in a manner that maximizes the benefit of the firm and how this decision can be interpreted in terms of rational choice models. This is the concern of the next chapter.

7.7 BANK PORTFOLIO CHOICE MODELS UNDER UNCERTAINTY

There are two categories of models that deal with bank portfolio decisions. These are: (a) those who are concerned with the improvement of the rate of return (portfolio optimisation models), which use a linear programming approach (e.g. Beazer 1975; Chambers and Charnes 1961) and (b) those concerned with the portfolio choice process (e.g. Porter 1961; Pyle 1971). Kamales (1983) developed a banking firm model, but focused exclusively on commercial bank lending and the impact of interaction of real and financial resources on the banking activities. The present study is exclusively concerned with the actual investment and lending decision or portfolio choice process of commercial banks under uncertainty. There are limited previous studies that attempted to develop alternative models of portfolio choice, involving simultaneous decisions to invest and lend by commercial banks. Among such limited studies Porter (1961) and Pyle (1971) are the most commonly cited examples. The following sections review these models.

7.7.1 Porter bank portfolio model

Porter (1961) made an early attempt to develop a bank portfolio choice model. His model was purely theoretical and no comparison was made with actual bank portfolios. His portfolio consists of three elements: cash, bonds and loans. He stated that cash is fixed as a proportion of deposits and hence there are only two decision variables, securities and loans, and the determination of one automatically determines the other. He considers the bank to be a profit maximizing firm with profits interpreted in terms of the following: the money return, liquidity and capital certainty the portfolio offers (Porter 1961:323).

According to Porter, there are two important areas of uncertainty for the bank. These are the level of the future deposit liabilities and the market value of the nonmatured securities in the bank's portfolio at any future point in time, for both of which a linear probability distribution is presumed known. The bank has only two ways of obtaining funds to meet deposit withdrawals. It can either sell its securities or discount its loans at the central bank. Thus, for Porter, the amount of the securities the bank holds is the function of the discount rate and the "deposit low" which he defines as the peak demand for liquidity. Porter's analysis is based on the assumption that it is always cheaper to sell securities than to discount loans.

Porter's analysis attributes risk, in terms of an objective probability distribution, to bonds and deposit liabilities and focuses only on the determination of the amount of bonds held by the bank. Although this indirectly determines the number of loans the bank can hold, in line with the assumption of the model it does not provide any possibility for assessing the efficiency and rationality of the bank's decisions to hold bonds or lend its assets. Moreover, it does not indicate whether loans are risky or riskfree assets and hence no comparisons can be made between the two decisions of investing in bonds and/or in loans. Thus, it is not clear whether the decision to invest in bonds or in loans reflects rational or irrational investor behaviour. Therefore, Porter's model of portfolio choice is not useful for our purpose which is to focus primarily on the rationality of the investment and lending decisions of the commercial banking firms.

7.7.2 Pyle's bank portfolio model

Pyle (1971) considered an intermediary that makes choices from three securities: a riskless security and two securities with uncertain yield over the decision period. The two risky securities are considered to be loans and deposits. Pyle's analysis focuses on the conditions under which a firm would be willing to sell deposits in order to buy loans, i.e. he analyzes the trade off between deposits and loans. That is, denoting the amount of the three securities as x_0 (the intermediary's position in riskfree assets), x_1 (the intermediary's position in loans), and x_2 (the intermediary's position in deposits), he analyzes simultaneous decisions of sale of deposits ($x_2 < 0$) and purchase of loans ($x_1 > 0$) (Pyle 1971:738). and c, r and i the corresponding yields per decision period, where c is certain, r and i are random variables with given means and joint distribution.

The investment decision is subject to the balance sheet constraint: $x_0 + x_1 + x_2 = 0$

The profit of the firm for the decision period is

$$\pi = c \bullet x_0 + r \bullet x_1 + i \bullet x_2 = x_1(r-c) + x_2(i-c) \quad (7.4)$$

If the firm's objective function is defined by $F(x_1, x_2)$ the assumed expected utility maximization can be given by:

$$F(x_1, x_2) = E[U(\pi)] \quad (7.5)$$

The first derivatives of the objective function with respect to x_1 and x_2 are:

$$F_{X1}\ (x_1, x_2) = E[U'(\pi)(r-c)] \quad (7.6)$$

and

$$F_{x2}(x_1, x_2) = E[U'(\pi)(i-c) \quad (7.7)$$

With $U(\pi)$ and $F(x_1, x_2)$ strictly concave functions, in line with the risk aversion principle of the expected utility model, if the expected marginal utility of loans evaluated at $x_1 = 0$ is positive for all non positive x_2 and the expected marginal utility of deposits evaluated at $x_2 = 0$ is negative for non negative x_1, the optimum for $U(\pi)$ will imply intermediation (i.e. positive loan position and negative deposit position) (Pyle 1971:739).

By focusing on the trade off between deposits and loans, both of which are characterized by risky returns, Pyle showed that banks prefer to hold more loans than deposits, depending on their marginal utility. However, the model does not indicate whether there is a possibility of increasing bank profit by investing in the alternative riskfree asset instead of loans. Pyle's model focuses on the existence of financial intermediation and provides strong proof of this. The model does not deal with the trade off between investment and lending decisions and the rationality or otherwise of such behaviour. Moreover, neither model considers the partial irreversibility of loans in terms of default risk as an important element in portfolio choice processes. This will be the concern of the next chapter which develops an alternative investor-lender firm model.

Portfolio choice models, such as Pyle's (1971), consider profit and its variability in the analysis of banking firms' investment decisions, but do not analyze the effect of deposit cost uncertainty. In these models the investor or manager is viewed as maximizing a concave profit function with a quadratic or exponential function often used to represent the firm's preference ordering. In these models, the asset choice is restricted to the efficient frontier, where additional return is only achievable at the expense of added variance (Santomero 1984:590).

In the above risk-return models the return characteristics of bank assets are assumed to be exogenously given and independent of bank decision-making. However, in practice, banks manipulate lending terms and conditions and can somehow influence the return characteristics of the assets Even though they have limited power to minimise the associated risk and uncertainty about the future returns. Santomereo (1984:590) argues that such approaches are theoretically correct if the entire set of assets from which the bank constructs its portfolio includes multiple pricing options for each loan category.

Harte and Jaffe (1974:141) state that the efficient risk-return combination for portfolio choice by a financial intermediary is determined by its attitude towards risk. The result of the separation theorem depends essentially on the ability of the intermediary to expand assets and liabilities in proportion and not on specific properties of risk in borrowing and lending. Santomero (1984:590), states that the results of Harte and Jaffe's (1974) bank portfolio selection models showed that under some restrictive assumptions, one can segment the scale of bank operations from the risk-return choice and a separation theorem can be developed for the bank in an environment in which no riskfree asset exists and a fully liability funded portfolio structure is assumed. Since Harte and Jaffe assume that the intermediary's borrowing and lending decisions are independent of risk attitudes, their model is not appropriate for the analysis of the investor-lender firm behaviour where firms are assumed to face both price and cost of funds uncertainty in the process of their decision-making.

7.8 THE CHAVAS MODEL OF PORTFOLIO SELECTION UNDER UNCERTAINTY

Chavas (2004) studied the behaviour of a firm confronted with investment decisions with two investment alternatives. These are a riskless asset and a risky asset. The investor has a one period planning horizon. The investment decision is made at the beginning of the period, yielding monetary returns at the end of the period. For each dollar invested, the riskless asset yields a sure return at the end of the period. The riskless asset can be considered to be a government bond which exhibits no risk of default. On the other hand, the risky asset yields an uncertain return at the end of the period.

What should the investor decide? At the beginning of the period, let I denote initial wealth of the investor. Let y denote the amount of money invested in the risky asset y and let z denote the amount of money invested in the riskless asset. According to Chavas (2004:124), the investor faces a budget constraint:

$$I = y + z, \tag{7.8}$$

Denote by p the monetary return per unit of risky asset y and r the monetary return per unit of riskless asset z. While r is known ahead of time, p is uncertain at the time of the investment decision. The uncertain rate of return on y is given by (p-1), while the sure rate of return on z is given by (r-1). The uncertain variable p is treated as a random variable and the investor has a subjective probability distribution on p. At the end of the period, let W denote the terminal wealth of the firm which satisfies:

$$W = py + rz \tag{7.9}$$

Let $p = \mu + \sigma e$, where $\mu = E(p)$ and e is a random variable satisfying $E(e) = 0$. The parameters μ and σ are the mean and standard deviation (or mean preserving spread) of r respectively.

Using the expected utility model, let the preference function of the decision maker be U(W). For a risk averse investor, we assume that $U' > 0$ and $U''<0$. Then the investment decisions are given by:

$$\text{Max}_{y,z} \{EU(W): I = y + z, W = py + rz\} \quad (7.10)$$

or $$\text{Max}_y \{EU[py + r.(I-y)]\} \quad (7.10')$$

or $$\text{Max}_y\{EU[rI + py-ry)]\} \quad (7.10'')$$

Chavas's model is similar to Sandmo's (1971) model of the competitive firm under price uncertainty. If the optimal choice of a risky asset in the above optimisation problem is given by $y^*(I, \mu, \sigma, r)$, (and if $c = rI$, the two models become equivalent) implying that the result obtained for output price uncertainty by Sandmo (1971) applies to this model too. Thus, Chavas (2004:124-125) showed that

a) $\partial y^* / \partial I >, =, < 0$ under DARA, CARA or IARA.
b) $\partial y^* / \partial \mu = \partial y^c / \partial \mu + (\partial y^* / \partial c)\ y^* > 0$ under DARA. This is the "Slutsky equation" where $\partial y^c / \partial \mu$ is the compensated price effect and $[(\partial y^* / \partial c)\, y^*]$ is the income or wealth effect.
c) $\partial y^* / \partial \sigma < 0$ under DARA.
d) If we denote the proportion of wealth invested in the risky asset as $Y = y / I$, the maximization problem can be alternatively written as

$$Max_Y \{EU[I \bullet (r + p - rY)]\} \quad (7.10''')$$

This model is similar to Sandmo's firm under price uncertainty when I = 1-t, t being the tax rate. Thus, the following result applies (Chavas 2004:125): $\partial Y^* / \partial I = \partial(y^* / I) / \partial I$ >, =, < 0 under DRRA, CRRA, IRRA respectively.

According to Chavas (2004:125), result (a) shows that under DARA preferences, a higher income tends to increase investment in risky assets and reduce investment in riskfree assets because higher income reduces the implicit cost of risk. Result (b) shows that, under DARA, increasing the expected rate of return on risky asset tends to increase its demand. Result (c) shows that, under DARA and risk aversion, increasing the riskiness of y (as measured by the standard deviationσ) tends to reduce its demand. As the implicit cost of risk rises, the risk averse investor has the incentive to decrease his investment in the risky asset and increase it in the riskless asset. We will compare this result with the I-L firm model to be developed in the next chapter. Result (d) shows how risk preference affects the proportion of investor wealth held in a risky asset in the $y^* /$ I. It also shows that this proportion does not depend on income, I, under CRRA and it declines with income under IRRA. These results provide useful linkages between risk, risk aversion and investment behaviour.

7.9 YAARI'S COMPARATIVE STATICS MODEL OF PORTFOLIO CHOICE

Yaari (1987) analyzes the portfolio selection decision using his dual theory approach. The model begins with Tobin's (1958) basic liquidity preference problem. The model considers two assets: a riskfree asset (cash) and a risky asset (security). The rate of return on cash is 0 and the rate of return on the risky security is θ where θ is a random variable distributed on the interval [-1, a), for some a > 0. One must assume that Eθ > 0. A decision maker wishes to invest a fixed amount, K, which satisfies the following condition $[0 \leq K \leq 1] / (1+a)$, and faces the problem of dividing this amount between cash and the risky security. Let x be the amount invested in the risky security, $0 \leq x \leq K$. Then the decision maker's gross return from his portfolio is given by the random variable K + θx. This belongs to Yaari's class of *V* random variables.

Yaari (1987:96) defines V as a set of all random variables defined on some given probability space with values in unit interval. For each $v \in V$, he defines a decumulative distribution function (DDF) of v to be denoted Gv as follows:

$$Gv(t) = \Pr\{v, t\}, \quad 0 \leq t \leq 1 \qquad (7.11)$$

Where Gv is always nonincreasing, right continuous, and satisfies $Gv(1) = 0$. For all $v \in V$, the following relationship holds:

$$Ev = \int_0^1 G_v(t)dt, \qquad (7.12)$$

where Ev is the expected value of v.

Yaari interprets the value of the random variables in V as payments denominated in some monetary units, which make each $v \in V$ interpretable as a gamble or lottery that a decision maker might consider holding. Restricting the value of the random variables in V to the unit interval can be interpreted to mean that (a) no gamble which involves a possible loss exceeding the decision maker's total wealth can be considered, (b) no gamble exists which offers prizes exceeding some predetermined large number.

Assuming that $=\succ$ is the decision maker's preference order on V, and assuming that $=\succ$ satisfies Yaari's axioms A1- A5, then by Yaari's theorem 1 there exists a continuous and nondecreasing real function, f, satisfying the following preference equation (Yaari 1987:99):

$$[1; p] \sim [f(p); 1].$$

such that picking the best portfolio is equivalent to selecting an x in the interval [0, K] so as to maximize the quantity

$$\Phi(x) = \int_0^1 f\,(G_{K+\theta x}(t))\,dt, \qquad (7.13)$$

where $G_{K+\theta x}$ is the DDF of GK+ θx.

The function $\Phi(\bullet)$, defined by Yaari, is of the form

$$\Phi(x) = K + cx\,,\ 0 \leq x \leq K \tag{7.14}$$

Equation (7.14) refers to Yaari's proposition 4, where the constant c is given by

$$c = \int_{-1}^{a} f\,(G_{\theta}(t)\ dt\text{-}1 \tag{7.14'}$$

where G_θ is the DDF of θ.

Since in dual theory Φ(x) is linear in x, this theory predicts plunging rather than diversification. Thus, letting x^* be the maximizer of Φ(x) under $0 \leq x \leq K$, Yaari (1987:109) concludes from his proposition 4 that

$$x^* \quad \begin{cases} = 0 & \text{if } \int_{-1}^{\infty} f(G_{\theta}(t))\, dt < 1 \\ = \text{any value in } [0, K] & \text{if } \int_{-1}^{\infty} f(G_{\theta}(t))\, dt = 1 \\ = K & \text{if } \int_{-1}^{\infty} f(G_{\theta}(t))\, dt > 1 \end{cases} \tag{7.15}$$

The above means no investment in risky assets if the return is less than 1, any amount of investment if the return is 1 and full investment of the entire wealth if the return is greater than 1. This is how Yaari's dual theory deduces plunging instead of diversification. Yaari's (1987) portfolio choice theory distinguishes between portfolio diversification of expected utility model and "plunging", which must not be confused with risk seeking. According to the dual theory, this concept refers to the behaviour of an agent who waits until the rate of return is high enough and then invests the whole resource. Conversely, under expected utility theory, investors always diversify in the sense that the amount invested in the risky asset is always positive, sometimes covering the entire wealth available for investment. However, both positions are extreme and real investment behaviour probably lies somewhere in between. As such, the dual theory produces corner solutions in optimisation problems under comparative statics (Yaari 1987:109-110).

Although Yaari's dual theory produced some results that are not possible under the expected utility theory, as seen in chapter 6 of the present study, his application of the theory to portfolio choice problems has produced results that are useful for further analysis of the investment decisions of firms under uncertainty.

7.10 ECONOMETRIC MODELS OF BANKING FIRM INVESTMENT DECISIONS UNDER UNCERTAINTY

Among the econometric models of banking firm investment decisions under uncertainty, Baum et al. (2005), based on the Lucas (1973) model, analyzed bank lending decisions under uncertainty using the econometric portfolio model. They begin from the point where the bank manager allocates x percent of the total asset as loans to the private sector and (100-x) to securities (investment) to maximize bank profit. Securities are assumed to provide a riskfree return of ($r_{f,t}$). The risky loans yield a stochastic return based on a time varying risk premium generated by the generalized conditional autoregressive hetroscedasticity (GARCH) model and denoted by

$$R_{i,t} = r_{f,t} + \text{premium}_{i,t} \qquad (7.16)$$

the expected risk premium E(premium$_{i,t}$) = ρ and its variance Var(premium$_{i,t}$) = $\sigma^2_{\varepsilon,t}$. Hence, the true return on risky loans will take the form:

$$\sigma R_{i,t} = r_{f,t} + \rho + \varepsilon_{i,t}, \qquad (7.17)$$

where the random component $\varepsilon_{i,t}$ is distributed as $\varepsilon_{i,t} \sim N(0, \sigma^2_{\varepsilon,t})$.

The decision maker, a bank manager, observes a noisy signal on $\varepsilon_{i,t}$ in the form of $S_{i,t} = \varepsilon_{i,,t} + v_t$ before allocating bank assets to risky and riskfree alternatives. The random variable v_t denotes the noise which is normally distributed as $v_t \sim N(0, \sigma^2_{v,t})$ and independent of $\varepsilon_{i,t}$. This noise signal is the proxy of macroeconomic uncertainty. Thus conditional on the signal $S_{i,t}$, the decision maker can form an optimal forecast of the return from risky loans as $E_t(\varepsilon_{i,,t}|S_{i,t}) = \lambda_t S_{i,t}$ where $\lambda_t = \sigma^2_{\varepsilon,t}/(\sigma^2_{\varepsilon,t} + \sigma^2_{v,t})$. Thus, at any point in time total expected return conditional on the signal takes the form (Baum et al. 2005:24-25):

$$E(\hat{Y}_{i,t}|\ S_{i,t}) = x_{i,t}(r_{f,t} + \rho + \lambda_t S_{i,t}) + (1 - x_{i,t})\ r_{f,t} \qquad (7.18)$$

where $\hat{Y}_{i,t}$ is total returns.

The conditional variance of returns will be:

$$Var(\hat{Y}_{i,t}|\ S_{i,t}) = \lambda_t \sigma^2_{v,t}\, x^2_{i,t} \qquad (7.19)$$

The decision maker's objective function using a simple expected utility framework $E(\tilde{U}_{i,t}|\ S_{i,t})$ which increases in expected return and decreases in the variance of return conditional on the signal $S_{i,t}$ is given by

$$E(\tilde{U}_{i,t}|\ S_{i,t}) = E(\hat{Y}_{i,t}|\ S_{i,t}) - \alpha/2\ Var(\hat{Y}_{i,t}|\ S_{i,t}) \qquad (7.18)$$

where α is the coefficient of risk aversion. From this we can easily derive the *ith* bank's optimal loan to asset (LTA) ratio as:

$$x_{i,t} = (\rho + \lambda_t S_{i,t}) / (\alpha\, \lambda_t \sigma^2_{v,t}) \qquad (7.19)$$

The bank's lending behaviour is affected by σ^2_v which is the measure of macroeconomic uncertainty. An increase in macroeconomic uncertainty or σ^2_v leads to a decrease in the LTA ratio. i.e. $[\partial Var(x_{i,t})]/\partial\sigma^2_{v,t} < 0$ (Baum et al. 2005: 25).

The above indicates that, as macroeconomic uncertainty increases, the cross section dispersion of the share of risky loans to total assets decreases, as uncertainty hinders the bank's ability to foresee investment opportunities. In other words, higher uncertainty renders noisier the signal that banks receive on expected returns, thereby pushing the banks to rebalance the composition of their assets according to new and worse signals provided by credit markets adversely affecting the allocation of financial resources. This fosters herding behaviour and leads banks to behave more homogenously than in quiet periods (Quagliariello 2006:3).

Quagliariello (2006) conducted similar econometric modelling of the impact of macroeconomic uncertainty on lending behaviour by Italian banks. He investigated the impact of both idiosyncratic (firm specific) uncertainty and aggregate macroeconomic uncertainty on the lending behaviour of banking firms. He found that during a period of increasing turmoil the allocation of bank credit became less efficient and concluded that macroeconomic uncertainty is an important determinant of banks' lending decisions and a cause of potential disturbance in financial resources allocation (Quagliariello 2006:17-18).

7.11 BANK LENDING AND THE MANAGEMENT OF RISK

Banking operations involve catering for contradictory and opposing objectives, such as maintaining liquidity, maximizing profitability and improving solvency. In addition to these, financial markets are in general characterized by the phenomena of moral hazard and adverse selection. Because of these frictions, banks are forced to hold a capital buffer of sufficient size, to hold enough liquid assets and engage in risk management (Cebenoyan and Strahan 2001:1-2). On the other hand, Froot and Stein (1998:56) argue that, although such frictions in financial markets affect banks' lending and risk taking decisions, active risk management can allow banks to hold less capital and to invest more aggressively in risky and illiquid assets such as loans. In this regard it is argued that banks can manage credit risk better if they engage in trading credit risks in the loans selling markets.

Thus, the existence of loan purchase and selling markets is considered to be crucial to banks' credit risk management. Banks that purchase and sell their loans hold lower levels of capital in relation to their assets than banks not engaged in loan buying or selling or loan selling or buying, but not in both. Similarly, banks that are engaged in loan sales and purchases hold lower levels of liquid assets as a percentage of their total assets (Cebenoyan and Strahan 2001:2-3).

The most striking result of these recent studies in bank lending and risk management behaviour is that banks which are involved in credit risk management through active loan purchase and sales hold more risky loans as a percentage of their balance sheet than other banks. This is partly because banks use the risk reducing benefit of the advanced risk management (via loan buying and selling) to take on more profits, engage in higher risk activities and operate with greater financial leverage in that they are ready to lend more of their assets to risky borrowers. Thus, the benefit of the advances in risk management in banking will probably be greater credit availability rather than reduced risk in the banking system (Cebenoyan and Strahan 2001:2-3). The present study will investigate whether the investment decisions of investor- lender firms correspond with the actual preference patterns followed by the commercial banks.

7.12 LENDING, MONITORING AND COSTLY STATE VERIFICATION -THE GALE-HELLWIG MODEL

The Gale and Hellwig (1985) model of an optimal credit contract under a competitive capital market was based on Townsend's (1979) model of *Optimal Contracts and Competitive Markets under Costly State Verification* and considers two types of economic agents: investors and entrepreneurs. Entrepreneurs are defined as the agents who wish to undertake risky ventures, but lack resources and hence turn to investors for external finance. Investors are defined as banks, and deposit taking financial institutions who are assumed to hold sufficiently large and diversified portfolios of investment to achieve perfect risk pooling (Gale and Hellwig 1985:650). In this case investors behave as if they are risk neutral due to risk pooling. Entrepreneurs are also considered to be risk neutral. However, in reality most of them are risk averse. Furthermore, the model assumes that investors (banks) can obtain deposits by paying the fixed rate of interest on riskless securities. The investor also assumes that he can obtain whatever funds he needs at that fixed interest rate and that this rate will serve as an opportunity cost of funds. The model considers a single representative investor-entrepreneur pair from a large number of individually insignificant investors and entrepreneurs in the competitive capital market. Competitive pressure ensures that each investor-entrepreneur pair writes a contract which maximizes the expected utility or expected profit of the entrepreneur subject to the constraint that the expected return to the investor must cover the opportunity cost of funds (Gale and Hellwig 1985:650-651).

The implementation of the contract requires the entrepreneur to reveal true information about the state to the investor. He will do so if he has no incentive to lie. That is why the contract is said to be incentive compatible. The optimal contract challenge is to choose an incentive compatible contract to maximize the entrepreneur's expected utility, subject to the investor's zero profit (i.e. zero economic profit) condition expected to be satisfied under competitive equilibrium.

Monitoring of loan funded projects is important because asymmetric information and adverse selection play an important role in lending or credit risk. Asymmetric information, usually captured by labour market *implicit contract models* (ICMs), refers to the fact that the borrowers know more about their business prospects than banks and that banks tend to attract high risk borrowers. This and increased competition for lending from nonbank lenders has resulted in banks shifting their portfolio to higher risk loans in the hope of increasing their profitability (Gup and Kolari 2005:171-272).

The problem of state verification under lending and investment decisions arises from the existence of such asymmetric information. Gale and Hellwig (1985) analyzed the problems of incentive compatible debt contracts using the model of borrowing and lending with asymmetric information. They argue that the revenue of a firm depends both on its investment in inputs and on the state of nature. While the firm (the borrower) can observe the state of nature directly at no cost, other agents (e.g. lenders) cannot.

Thus, there is a positive state observation cost involved for the lender to verify the state of the borrower. Investment takes place before the state is observed and hence the impact of asymmetric information falls directly on the distribution of revenue between the borrower and lender and only indirectly on the level of investment. This implies that the Gale-Hellwig model is the description of risk sharing with asymmetric information and positive observation costs (Gale and Hellwig 1985:648).

In the Gale-Hellwig (1985) optimal credit contract model, the act of the observation of state can be considered as "bankruptcy" which means that the state is observed if and only if the firm cannot repay the loan in full or when the firm is insolvent. The lender cannot observe the state when the firm is not bankrupt. Since bankruptcy is a costly business (due to loss of valuable employees, goodwill and reputation until the new management takes over the firm) state verification of the lender is costly. This argument implies that the optimal credit contract between lenders and borrowers is a standard debt contract under bankruptcy (Gale and Hellwig 1985:648).

There are four possible scenarios regarding outstanding loan or investor entrepreneur debt: these are: (a) it can be repaid on schedule, (b) it can be renewed or extended, (c) the bank can sell the loan to another investor, and (d) the loan can go into default and the bank can sustain losses. While the first three scenarios are desirable, it is the last one that causes trouble to the bank. The concern of the present study is this last scenario and its impact on the investment and lending behaviour of the I-L firm. This is where the irreversibility of investment in loans, discussed in chapter five of the present study, enters the frame. State verification in the form of bankruptcy implies two possibilities; first, the lender makes as much recovery of his loans as he can by selling the assets of the firm, or, second, he fails to make any recovery of loans and incurs a complete loss.

7.13 SUMMARY AND CONCLUSION

Theories of the firm generally neglect the importance of risk and uncertainty in investment decisions. Other than the transaction cost theory, which assumes bounded rationality, and Knight's work on risk and uncertainty, most theories of the firm ignore these important factors in firms' investment decisions.

Theories of the banking firm follow a similar pattern. Many theories assume certainty in banking firm investment decisions. Under certainty banks' decisions would be simple. In such cases the banks would hold cash or maturing securities sufficient to satisfy all transaction requirements and invest the remainder of their portfolio in the highest yielding asset. However, in the real world, banking firms face risk and uncertainty in their decision process although the operational distinction between these two concepts with respect to investment decisions is not clear cut.

The risk-return portfolio choice models, particularly the Markowitz model, stipulate that investors choose a set of efficient portfolios that give lower variance of return and the same expected return for any available alternative. This model also suggests that firms prefer to hold diversified portfolios to minimise risk. This model shows that when there is less than perfect positive correlation diversification leads to reduction in return variances or risks. However, the risk-return portfolio choice models have not been widely applied in the analysis of the commercial bank investment and lending decisions.

There are two categories of models that deal with bank portfolio decisions. These are those concerned with the improvement of the rate of return (portfolio optimisation models) that use a linear programming approach and those concerned with the portfolio choice process. The latter analyze bank portfolio choice problems under uncertainty.

The Porter bank portfolio model identifies two important areas of uncertainty for the bank: the future deposit liabilities and the market value of the future nonmatured securities in the bank's portfolio, for both of which linear probability distribution is known. However, Porter focuses only on the determination of the level of security investment. The model assumes that the lending decision is automatically determined after the investment decision on security is made and hence does not analyze the lending decision on its own. Thus, the objective of Porter's model and that of the investor-lender firm model developed in the next chapter of the present study differ significantly.

Pyle's bank portfolio model analyzes the problem of trade off between loans and deposits instead of lending and investment decisions and concludes that the maximization of the expected utility of the banking firm with a positive loan position and negative deposit position will imply intermediation. The objective of this model is to provide proof of the existence of financial intermediation but not the rationale for the lending decision that is the objective of the investor-lender firm model of the present study.

The Chavas model of bank portfolio selection identifies riskfree and risky investment options. The model investigates the effects of attitudes toward risk on the level of investment in riskfree and risky assets. However, the model is concerned with the maximization of the expected utility of terminal wealth instead of the utility of profit.

The existence of loan purchase and selling markets is considered to be crucial in banks' credit risk management. It is argued that banks that purchase and sell their loans hold lower levels of capital in relation to their assets than banks not engaged in loan buying or selling. However, the presence of loan buying and selling markets is more likely to lead to greater credit availability than to reduce risk in the banking system. This has an important implication for the study of the behaviour of the investor-lender firm. In practice, the main reason for the collapse of several banks is lending risk. However, banks obtain more profit from their lending activities and continuously target high risk borrowers as risk management methods improve.

The econometric models of bank economic decisions under uncertainty use generalized autoregressive conditional hetroscedasticity (GARCH) models to measure macroeconomic uncertainty. These models conclude that the latter is an important determinant of banks' lending decisions and a cause of potential disturbance in financial resources allocation.

Banking firms operate in two different segments of markets: investment markets and lending markets. In portfolio choice problems, investment markets usually involve securities that are characterized by various risks, the least risky security investment being investment in government bonds. The lending market is more complex. In these markets, firms face return uncertainty because of the risk of loan default. Banking firms try to incorporate risks and the costs of their fund when they price their loans. Lending decisions also involve the problem of state verification owing to the existence of asymmetric information usually captured by labour market implicit contract models (ICMs). The lender cannot observe the state when the firm is not bankrupt. Since bankruptcy is an expensive business state verification is costly to the lender.

None of the bank portfolio choice models investigated in this chapter has incorporated the problem of cost uncertainty into the analysis of the investment decisions made by banking firms. The investor-lender firm model to be developed in the following chapter will incorporate the effects of cost uncertainty in the investment decisions of investor-lender firms.

CHAPTER EIGHT

INVESTOR-LENDER FIRM MODEL

8.1 INTRODUCTION

Existing investment theories have given little attention to the study of the behaviour of firms that choose between lending their funds to other firms and investing it themselves, i.e. investor-lender firms. In addition, most models that analyze investment behaviour of such firms ignore the problems of risk and uncertainty. Those that incorporate the latter focus only on price and revenue uncertainty. Little attention has been given to cost-of-funds uncertainty. The objective of this chapter is to develop a model of an investor-lender (I-L) firm that captures decisions of investment and lending that incorporates both return uncertainty and cost-of-funds uncertainty into a single decision process.

As indicated in chapter one, an investor-lender firm can be thought of as a commercial bank involved with mobilisation of funds from customer deposits and the use of these funds either to lend to other firms or invest itself. Cantillo and Wright (2000:158) argue that if banks decide not to lend, they can invest their capital in a riskfree asset and also avoid intermediation costs. We also distinguish between commercial banks and investment banks. The present study focuses on the activities of the commercial banking firm as an investor-lender (I-L) firm and analyzes its behaviour based on the assumption that the I-L firm is a rational portfolio investor under risk and uncertainty with cost-of-funds constraint.

As indicated in the preceding chapter, economic theory suggests that economic agents prefer certainty to risk and risk to uncertainty or ignorance. This argument is based on the most commonly accepted economic assumption of rationality propagated mainly by the expected utility model. According to this model, economic agents are expected always to exhibit rational choice and firms strive to maximize the expected utility of profit. However, various alternative nonexpected utility models have argued that, in reality, economic agents do not behave in the manner postulated by the expected utility model. These theories argue that economic agents systematically violate the basic assumption of the EU theory, i.e. linearity in probability or the independence axiom.

The most common examples of these violations are given by the Allais nonlinear intensity theory and prospect theory, which also add the failure of the invariance principle, i.e. descriptive invariance (as a result of framing effects) and procedure invariance (due to preference reversal). The violation of independence indicates that individuals prefer certainty to higher probabilities but prefer taking risks for smaller probabilities. The prospect theory analyzes this problem using the fourfold characterization of attitudes toward risk based on nonlinear weighting of probabilities in economic decisions. Other alternative theories try to generalize the EU theory by weakening or dropping the independence axiom. Such theories include Machina's type II or *fanning-out* hypothesis where preferences are represented by indifference curves which fan out in a unit probability triangle (see ch 6). Based on the above argument, recent experimental studies have therefore argued for the presence of investor irrationality in terms of weak irrationality when the decisions of investors deviate from expected utility maximization, and strong irrationality when their decisions show more severe violations of the monotonicity (preferring more to less) axiom of the EU theory (see ch 6).

Empirical testing of the four nonexpected utility theories reviewed in the present study has shown that none of these theories is capable of fully explaining the violations of some assumptions of the EU theory. In many cases the assumptions of the alternative theories are themselves violated. Therefore, none of these theories can serve as an adequate alternative theory of choice under uncertainty. Moreover, the above conclusions of nonexpected utility theories, particularly the Allais nonlinear intensity theory and prospect theory, are based on two crucial assumptions. First, the objective of economic agents is considered to be not the maximization of the expected utility of profit, but rather the maximization of other benefits or values. Second, while these theories incorporate revenue uncertainty into their analysis of the economic decisions of agents, they ignore the cost uncertainty associated with these decisions. This is a direct result of the first assumption of these models. In addition to these, unlike the expected utility theories and most nonexpected utility theories, prospect theory makes a further assumption. It assumes that carriers of value in economic decision processes are not final, net positions in wealth but, rather, changes in wealth. Its conclusions regarding different risk attitudes for losses and gains are the direct consequence of this assumption. Since the evidence suggests that none of the alternative theories provide a comprehensive alternative tool to analyze economic decisions taken by agents under risk and uncertainty, the I-L model developed in the next section of the present study will be based on an assumption of the maximization of expected utility of profit in line with the EU theory. However, results obtained will be compared with those of the nonexpected utility theories, particularly the Allais nonlinear intensity theory and prospect theory for similar decision problems.

The rest of the chapter is organized as follows: section 8.2 develops the I-L firm model. Section 8.3 presents the analysis of the result of the I-L model. Section 8.4 provides the nonexpected utility interpretation of the result of the I-L model with particular emphasis on the Allais nonlinear intensity theory and prospect theory, while section 8.5 concludes.

8.2 THE MODEL

The I-L firm faces an investment and lending choice problem. It can lend either all or part of its capital as loans (L) to other firms, or invest all or part of its capital in riskfree securities. Assume that the I-L firm is a rational investor or portfolio holder and not a neoclassical productive firm, where the firm faces cost constraint but not production function constraint, i.e. it incurs cost-of-funds or deposit cost.

For investment decisions regarding loans we assume that (a) there are no loan brokers in the market and borrowers (which we shall define henceforth as entrepreneurs) borrow directly from the I-L firms, (b) there is no subsidy to entrepreneurs, (c) entrepreneurs are internal finance constrained; that is, they have limited retained earnings and equity financing options, (d) entrepreneurs face minimum transaction cost during loan processing, (e) the I-L firms face uncertainty about the future returns from L. That is, there is no state verification by I-L firms of projects undertaken by the borrower firms; the optimal contract between entrepreneurs and I-L firms is debt, (f) there is no credit rationing: firms do not refuse loans to entrepreneurs based on their perceived default rate. Thus, entrepreneurs are not credit constrained and hence aim to maximize the expected profit. Finally, (g) we assume that there is no friction between the I-L firms and their depositors and the opportunity cost of holding capital by the firms does not increase.

Another assumption of this model is that individuals and firms prefer certainty to both risk and uncertainty. They also prefer risk to ambiguity. We assume further that they prefer one type of uncertainty to other types of uncertainty, i.e. they prefer familiar uncertainty to nonfamiliar uncertainty. The question this model attempts to answer is: why do the I-L firms decide to invest in loans which bring higher uncertain returns rather than in riskfree financial assets with lower certain returns?

Assume that there are n investor-lender firms operating in a perfectly competitive market. Assume also that the I-L firm invests only in homogenous government securities (B) with riskfree returns when it decides not to lend its capital. The monetary return per unit of risky asset L is r and the monetary return per unit of riskless asset B is b. The I-L firms are price takers for loans r and riskfree bonds b. While b is known ahead of time and is a fixed rate with the probability P(b) = 1, r is uncertain at the time of the investment decision. The variable r is treated as a random variable indicating uncertainty of returns from loans and the investor has a subjective probability distribution $f(r)$ on r and the $E(r) = \mu_r$. The random variable r is defined as:

$$r = \mu_r + \sigma_r e \qquad (8.1)$$

where e is a random variable with mean zero.

The capital of the I-L firm is generated by collecting deposits from customers. However, the level of the future deposit liability faced by the I-L firm is uncertain, because the transaction pattern of the deposit holders and deposit turnover rate is not known in advance. Deposit uncertainty is captured by the randomness of customer deposits. The associated deposit cost per unit of deposit is δ and is random. The I-L firms are price takers for deposit δ. Sealey (1980:1143) argues that if deposits are assumed to be *ex ante* random, then deposit rates must always be set prior to observing the actual quantity of deposits forthcoming and in this case, both interest and resource costs should be random. The random deposit cost per unit of D has a probability distribution $f(\delta)$ with the $E(\delta) = \mu_\delta$. The random deposit rate is defined as

$$\delta = \mu_\delta + \sigma_\delta e \qquad (8.2)$$

where e is a random variable with mean zero, i.e. $E(e) = 0$. The deposit cost δD is random and hence uncertain. Therefore, the I-L firm faces both return uncertainty from loans as well as cost-of-funds uncertainty.

Under the assumption of certainty, individuals or firms will borrow whenever their time preference is greater than the market rate, and will lend whenever it is less than the market rate (Haberler 1931:502). However, under uncertainty, this preference pattern no longer holds.

The investment decisions of the I-L firm is subject to the following budget or balance sheet constraint

$$B + L + D = 0 \qquad (8.3)$$

Assume that the objective of the firm is to maximize the expected utility of profits in line with the expected utility model where the utility function of the I-L firm is a concave, continuous and differentiable function of profits implying

$$U'(\Pi) > 0, \quad \text{and} \quad U''(\Pi) < 0$$

which refers to the concave Von Neumann-Morgenstern utility function for the risk averse individual or firm.

The I-L firm is confronted with choices regarding the level of investment in the riskfree security B and risky loans L. The revenue of the I-L firm at the end of the decision period is given by

$$R = bB + rL \tag{8.4}$$

The revenue, R, of the I-L firm has both a riskfree component bB and a risky and uncertain component rL. This implies that if the objective of the I-L firm is the maximization of revenue or some value, as nonexpected utility theories argue, the I-L firm can avoid revenue uncertainty by fully investing in the riskfree security B. However, when the objective of the I-L firm is the maximization of profit under uncertainty, the issue of cost-of-funds uncertainty becomes crucial.

The cost of the I-L firm when it invests in both riskfree bonds and risky loans is given by

$$C = \delta D \tag{8.5}$$

where C is an uncertain deposit cost or cost-of-funds. Since the I-L firm is assumed to be a rational portfolio investor, it faces no other costs except the uncertain deposit cost. The deposit cost C is positive for positive deposit level.

The profit of the I-L firm at the end of the decision period, when it makes this investment decision, is given by revenue minus cost i.e., π = R-C and is stated as

$$\pi = (bB + rL) - \delta D \tag{8.6}$$

Subsequently, solving for B in the balance sheet constraint (8.3), and substituting the result in equation (8.6) we obtain the following profit function for I-L firm

$$\pi = L(r - b) + D(\delta - b) \tag{8.6'}$$

where π is a concave, continuous and differentiable profit function for the I-L firm.

The profit function (8.6′) is similar to the famous Pyle (1971:739) terminal wealth function for a financial firm involved in the investment of two risky assets and one riskfree asset with the same balance sheet constraint. This is exactly the same problem faced by the investor-lender firm.

The objective of the I-L firm is to maximize the expected utility of profit in (8.6′). The maximization of the expected utility of profit as the objective of the banking firms is assumed by several authors (Kane and Malkiel 1965; Parkin, 1970; Pyle, 1971; Hyman 1972; Stigum, 1976 and Sealey, 1980).

The expected utility of profit for I-L firm becomes:

$$E[U(L(r - b) + D(\delta - b)] \tag{8.7}$$

where E is the expectation operator. Then, the profit maximization problem of the I-L firm is associated with both the random variable of the return on loans r and the random variable δ, the deposit cost.

The investment decision of the I-L firm can be represented by the following problem of the maximization of the expected utility of profit:

$$\underset{L,D}{Max} F(L,D) = \underset{L,D}{Max} E\{U(L(r-b) + D(\delta -b))\} \tag{8.8}$$

That is, the I-L firm maximizes the expected utility of profit when it invests in given loans (risky securities) given the risky deposit liability (D).

The first order conditions for the existence of the maximum are

$$\frac{\partial F}{\partial L} = E[U'(r-b)] = 0 \qquad (8.9)$$

$$\frac{\partial F}{\partial D} = E[U'(\delta-b)] = 0 \qquad (8.10)$$

Expanded forms of equations (8.9) and (8.10) become:

$$E[U'(r\text{-}b)] = E(U')\,\mu_r + Cov(U',r) - [E(U')b + \mathrm{Cov}(U',b)] = 0 \qquad (8.9')$$

$$E[U'(r\text{-}b)] = \mu_r - b + \frac{Cov(U',r) - Cov(U',b)}{E(U')} = 0 \qquad (8.9'')$$

And

$$E[U'(\delta\text{-}b)] = E(U')\,\mu_\delta + Cov(U',\delta) - [E(U')b + \mathrm{Cov}(U',b)] = 0 \qquad (8.10')$$

$$E[U'(\delta\text{-}b)] = \mu_\delta - b + \frac{Cov(U',\delta) - Cov(U',b)}{E(U')} = 0 \qquad (8.10'')$$

U' is the marginal utility of profit and hence can also be written as $U'(\pi)$. Under risk neutrality, marginal utility of profit is constant (Sealey 1980:1144) and hence all four of the covariance terms become zero. Thus, as in the case of other models based on the assumptions of profit maximization by the banking firms, the utility function in the I-L model plays a part in the analysis.

Thus, the first condition (8.9") means that the I-L firm will invest in loans until the expected return from risky and riskfree assets are equalized while the second condition (8.10") means that the I-L firm invests in deposits until the expected marginal deposit cost and the return from riskfree assets are equalized.

Putting the two conditions together, i.e. putting (8.9") and (8.10") together and rearranging, we get the key result that the optimal investment for the I-L firm occurs when the expected return from loans equals the expected marginal deposit cost. That is:

$$\frac{\mu_r}{\mu_\delta} = \frac{b}{b} = 1 \qquad (8.11)$$

And therefore,

$$\mu_r = \mu_\delta \qquad (8.12)$$

We know that μ_r is the expected return from lending. From equation (8.5) δ is the marginal deposit cost. Since this variable is random, its expected value, i.e. $E(\delta) = \mu_\delta$

Therefore, the I-L firm continues to invest in loans until the expected return from loans and the expected marginal deposit cost are equalised. This an equilibrium solution. As indicated in equations (1) and (2), r and δ are determined as follows:

$r = \mu_r + \sigma_r e$, and

$\delta = \mu_\delta + \sigma_\delta e$

Therefore, these two variables can be stated as functions of uncertainty, σ (which can be measured as time varying variance using the GARCH model) and the random variable e. That is

$$r = f(\sigma_r, e) \qquad (13)$$

And

$$\delta = f(\sigma_\delta, e) \tag{14}$$

Thus, $\frac{\partial r}{\partial \sigma_r} > 0$, $\frac{\partial r}{\partial e} > 0$, $\frac{\partial \delta}{\partial \sigma_\delta} > 0$ and $\frac{\partial \delta}{\partial e} > 0$.

That is both the return from loans and deposit costs increase as their respective uncertainties increase. The increase in the random variable e also increases both the required return from loans and costs of deposits. The positive relationship between uncertainty and the required rate of return or the threshold rate is explained in chapter 5 of the present study.

However, the relationship between the parameters of the equilibrium solution of the I-L model, i.e. between μ_r and μ_δ, and uncertainty is negative. That is

$$\mu_r = r - \sigma_r e \tag{15}$$

And

$$\mu_\delta = \delta - \sigma_\delta e \tag{16}$$

Equations (15) and (16) indicate that given a positive random variable e, increase in uncertainty reduces the expected return from loans as well as the expected marginal deposit cost.

The I-L firm model is the extension of the famous Pyle (1971) financial intermediary model. Moreover, the result of the Investor-Lender firm model in (8.12) is the direct counterpart of another financial intermediary model under risk and uncertainty by Sealey (1980). See Sealey (1980:1142-1144) for comparison with the result of the I-L model. Sealey's (1980) result differs from the result of the I-L firm model in two important respects. First, Sealey (1980) was concerned only with investment decisions involving lending and deposits. He does not deal with any riskfree assets. Second, he incorporates resource and liquidity costs into the analysis of the investment decisions while for the I-L model the only cost involved is the uncertain deposit cost.

Results similar to those of the I-L model are also obtained by Klein (1971:213) whose equation (30) indicates that the optimal loan policy is chosen at the point at which the marginal expected return on loans is equal to the average (and marginal) expected return on government securities. This is very close to the solution of the I-L model.

The second order sufficient condition for the optimality of the I-L firm investment decision is assumed to be negative. In the following section, we interpret the above result of the I-L model.

8.3 THE RESULT OF THE I-L MODEL

The result of the I-L model indicates that when the objective of the I-L firm is the maximization of profit, the optimal investment decision occurs when the expected return from lending is equal to the expected marginal deposit cost. This finding is completely different from the decision criteria of the classical certainty models, which equate price with marginal cost in production decisions of competitive firms. For decisions under uncertainty, the "price equals marginal cost" principle does not hold. Therefore, the I-L firm's investment decisions are based on the comparisons of random price and random costs which are both uncertain at the time of decision-making.

Thus, for the I-L firm, the optimal investment decision occurs when the expected return from loans equals the expected marginal deposit cost. This is the direct result of the incorporation of deposit cost or cost-of-funds uncertainty into the problem of the maximization of the expected utility of profit by I-L firms. This optimal investment decision involves the investment of the entire capital on loans i.e. the decision involves plunging. This means that when the objective of the firm is the maximization of the expected utility of profit, investment in risky loans emerges as a dominant decision. Thus, investment in a riskfree asset does not maximize profit.

In other words, the presence of deposit cost uncertainty reduces the superior riskfree investment opportunity to an inferior option. Riskfree investment is associated with risky cost which leads to profit uncertainty. Revenue certainty implied by investment in riskfree securities does not dominate the investment decision because the presence of deposit cost uncertainty leads to profit uncertainty rendering the riskfree investment decision inferior. On the other hand, the I-L firm views investment decisions made when the expected return from loans equals the expected marginal deposit cost as superior and that which maximizes its expected utility of profit. Therefore, the I-L firm does not see any reason for investing its funds in riskfree financial assets which will not maximize its expected utility of profit. That is why I-L firms lend their capital to other firms instead of investing it themselves.

Moreover, as long as the objective of the I-L firm is the maximization of profit, its investment decisions always entails risk even when the actual return on the investment is riskfree. Therefore, the firm faces this major uncertainty whether it decides to invest in riskfree assets or in risky assets. In essence, there is no riskfree economic decision when the objective of that decision entails the maximization of profit. The results of the I-L model imply that economic agents are confronted with various sources of risk and uncertainty in their day to day economic decisions and hence all sources of uncertainty must be included in the analyses of their economic decisions.

Based on the objective of the maximization of the expected utility of profit, the I-L model predicts that the lending decision of the firm represents a rational choice under uncertainty. This decision does not imply the violation of the independence axiom or the linearity in probability assumption of the EU theory.

The result also explains why, in actual practice, the I-L firms (commercial banks) continue to lend more and more of their capital to high risk entrepreneurs in pursuit of higher profits. With banking's increased sophistication in risk management in loans purchasing and selling markets, the I-L firms tend to invest their capital in increasingly risky ventures because they equate their expected return from these with the expected marginal deposit cost. However, this result does not necessarily imply that these firms are risk and uncertainty loving; rather, it implies that, given the real life situation of risk and uncertainty, firms can make optimal choices by internalising them into their decision processes.

Another implication of the results of the I-L firm model under cost uncertainty is that the investment decision of the I-L firm involves plunging instead of diversification. This is similar to the result of Yaari's model which is based on a completely different approach. However, the conclusion of plunging, both in the I-L model and Yaari's model, is based on two asset portfolio investment problems. The result for investment decisions involving more than two assets needs further anaysis.

Previous studies on banking firm investment decisions ignored the importance of cost uncertainty in these investment decisions, mainly because of the importance attached to price and revenue uncertainty in the study of the investment decisions. By introducing cost-of-funds uncertainty to the analysis of investment decisions of I-L firms, the present study makes an original contribution to the literature in this area. Thus, the model developed in this chapter can be termed the "Theory of the Investor-Lender Firm under Cost-of-Funds Uncertainty".

The problem of the firm level financial investment decisions under uncertainty analysed in this chapter is extended to the problem of aggregate fixed investment decisions under uncertainty in the next chapter. This problem is investigated using the GARCH generated measures of uncertainty and econometric modelling of the investment equation.

8.4 THE NONEXPECTED UTILITY INTERPRETATION OF THE RESULT OF THE I-L MODEL

Consider again the investment decision of the I-L firm:

$$\underset{L,D}{Max}F(L,D) = \underset{L,D}{Max}EU\{(L(r\text{-}b) + D(\delta\text{-}b))\}$$

with the optimal investment decision accruing when the expected return from loans is equal to the expected marginal deposit cost, i.e.

$$\mu_r = \mu_\delta$$

with the following budget or balance sheet constraint:

$$L + B + D = 0$$

The above conclusion holds when the objective of the firm is the maximization of the expected utility of profit. However, nonexpected utility theories argue that the objective of economic agents is not always maximization of the utility of profit. Based on this assumption, these theories provide alternative explanations of the decision process of economic agents.

Let us now present the results of the I-L firm choice problem under uncertainty in the form of individual preferences for two pairs of gambles in line with the two famous nonexpected utility theories: the Allais nonlinear intensity theory and the prospect theory. In this case, we have to make two crucial assumptions. First, the objective of the I-L firm is not the maximization of the expected utility of profit. The objective is simply the maximization of the expected value, or the maximization of the expected revenue in the case of the I-L firm. Second, investment decisions now involve choices between one riskfree investment or investment in two risky assets. Assume the riskfree asset to be government security (*B*), and the risky assets to be loans (*L*) and common stocks (*S*). Since the return on investment on the riskfree government security is certain, we set its probability of return from B at $P_B = 1$. We assume that the benefit (revenue) from this investment is \$600,000. Since loans and common stocks are risky investment options, we set the probability of the return from L at $0 < P_L < 1$ and the probability of return from the common stocks S at $0 < P_S < 1$. The revenue from investment in loans is \$3,000,000 while the revenue from investment in common stocks is \$600, 000, the same as that from investment in government security. In this case the investment decision is constrained by the following balance sheet constraint.

$$L + B + S + D = 0$$

The choice problem can now be presented as follows:

h_1 :{\$600,000 benefit (B) with P = 1} versus

h_2: {\$600,000 benefit (S) with P = 0.89, \$0 benefit with P = 0.01, \$3000,000 benefit (L) with P = 0.1}

and

h_3: {\$3,000,000 benefit (L) with P = 0.10, \$0 benefit with P = 0.90} versus

h_4: {\$600,000 benefit (S) with P = 0.11, \$0 benefit with P = 0.89}

In the above choice problem there are three prospects of benefits. These are: {\$0, \$600,000, \$3,000,000}. See Machina (1987:128) for a similar formulation of the above game problem. Nonexpected utility theories argue that under the expected utility theory, any risk averse investor would prefer h_1 or to invest in the riskfree government security (B) in the first case and h_4 or investing in common stocks (S) in the second pair of gambles. This is the result expected under the assumption of parallel indifference curves or linearity in probability under the expected utility model. Thus, it does not represent a paradox. For Allais (1953, 1979b), Kahneman and Tversky (1979) and Tversky and Kahneman (1986), the paradox occurs because the majority of individuals choose h_1 in the first pair and h_3 in the second pair, indicating the nonlinearity of the probabilities or the violation of the independence axiom of the expected utility theory. The application of the above game theory to the I-L firm investment decisions with choice problems among riskfree asset (B), risky asset common stocks (S) and lending (L) indicates that the firm chooses to invest in riskfree bonds in the first pair but decides to lend its capital to other firms in the second pair of gambles thereby indicating the violation of the independence axiom. For the independence axiom to hold if $h_1 \succ h_2$ then $h_4 \succ h_3$. According to the nonexpected utility theories, individuals violate this in their decisions because the difference between the payoff probabilities between h_3 and h_4 of 0.01 is outweighed by the larger payoff for h_3. The independence axiom is considered to be the most important postulate of rationality. Therefore, according to the nonexpected utility theories, its violation challenges this basic assumption of rational choice in economic analysis. The violation of independence as described above indicates that individuals prefer certainty to higher probabilities but prefer taking risks for smaller probabilities. However, this result depends crucially on the first assumption of the nonexpected utility approach that the objective of the I-L firm is not the maximization of profit.

This result implies that the I-L firm's decision to lend its capital to other firms represents risk seeking behaviour and hence nonrational choice that contradicts the basic axiom of independence of the expected utility theory. For prospect theory, this result represents the nonlinear weighting of probability where economic agents over weight lower probabilities and under weight higher probabilities, demonstrating risk seeking behaviour for lower probabilities and risk averse behaviour for higher probabilities. However, as we have shown above, the conclusions of the nonexpected utility theories (Allais's nonlinear intensity and prospect theories in this case) in the above decision problems involving two pairs of gambles is the direct result of the fundamental assumption of these models that the basic objective of economic agents is not the maximization of the expected utility of profit but the maximization of benefits or some values. In the above investment choice problem presented in the form of two pairs of gambles, there cannot be a riskfree investment option that corresponds to $P = 1$ if the objective of the investor is the maximization of the expected utility of profit. Thus, h_1 cannot serve as an alternative in the first pair of the gamble because h_1 will not maximize the expected utility of profit. This has been shown by the results of the I-L model. All alternatives in the two pairs would reflect risky options and the firm would choose from among these an option which maximizes its expected utility of profit. In so doing, it would equate the expected return from this risky option with the expected marginal cost-of-funds. The idea that no alternatives in economic decisions may be riskfree has been echoed by Ross (1981:622) who states that all choices compared in economic decisions involve risk and that there is no riskfree situation *per se* in the sense that all situations involve some form of uncertainty. Thus, when the objective of the economic agents is the maximization of the expected utility of profit, the I-L firm's decision to lend its assets to other firms represents a rational decision under uncertainty. The findings of the I-L model support Battalio et al. as well as Birnbaum and Birnbaum-Navarrete's findings and the findings of other authors who have performed empirical tests of the main nonexpected utility models. None of these theories serve as a comprehensive alternative to EU theory as descriptive theories of choice under uncertainty.

Therefore, by extending the expected utility theory and introducing cost-of-funds uncertainty into the analysis of the investment decisions of firms with the objective of the maximization of the expected utility of profit, the results of the I-L model show that the I-L firm's decision to invest in risky ventures as opposed to seemingly riskfree ventures represents a rational choice under uncertainty. The result of the model also implies that when the objective of the firm is maximization of the expected utility of profit, no investment decision can be riskfree. All investment decisions would involve risk because in the absence of revenue uncertainty, the firm is confronted with an alternative cost uncertainty that should be included in the analysis of the investment decisions of the firm.

8.5 CONCLUSION

The present chapter developed an investor-lender firm model. The I-L firm is considered to be a rational portfolio investor with cost-of-funds uncertainty. The firm is faced with two choices: investment in riskfree assets (homogenous government bonds) with certain returns or investment in risky loans with uncertain returns. The capital of the firm is raised through deposit collection from customers. The future deposit liability is uncertain. Therefore, the associated deposit cost or cost-of-funds faced by the firm is also uncertain. The objective of the firm is the maximization of the expected utility of profit.

Given the objective of the I-L firm, its optimal investment decision is obtained by maximizing the expected utility of profit. The optimal investment decision of the I-L firm occurs when the expected return from loans is equal to the expected marginal deposit cost. This means that when the objective of the firm is the maximization of the expected utility of profit, investment in risky loans emerges as a dominant decision. Thus, investment in riskfree assets does not maximize profit. Thus, the decision to lend reflects a rational choice under uncertainty. As long as the objective of the I-L firm is the maximization of the expected utility of profit, its investment decisions always entail risk even when the actual return on the investment is riskfree. The firm faces this major uncertainty whether it decides to invest in riskfree assets or in risky assets.

This result also reflects the actual preference pattern these kinds of firms reveal when making investment decisions. Commercial banks continue to lend an increasing amount of their capital to riskier entrepreneurs because their objective is the maximization of the expected utility of profit. However, this does not necessarily mean that they are risk and uncertainty seeking. It reflects the fact that in a real life situation, firms are forced to take the presence of risk and uncertainty into consideration when they make investment decisions.

Nonexpected utility theories interpret the above type of result as a violation of the independence axiom or the linearity in probability assumption of the expected utility theory. By presenting choice problems in the form of pairs of gambles with the objective of benefit (revenue) maximization, the Allais nonlinear intensity theory and the prospect theory conclude that agents violate the independence axiom of the EU theory and hence challenge the basic assumption of rational choice in economic analysis. This violation of independence indicates that individuals choose certainty over higher probabilities but prefer taking risks for smaller probabilities, indicating nonlinearity in probability. However, empirical tests of the nonexpected utility theories have shown that they do not explain most of the violations of EU theory and hence cannot serve as adequate alternative theories of choice under uncertainty.

CHAPTER NINE

THE EFFECTS OF UNCERTAINTY ON AGGREGATE FIXED INVESTMENT OF FIRMS: AN EMPIRICAL ECONOMETRIC ANALYSIS*

9.1 INTRODUCTION

The current chapter is an empirical counterpart of chapter five that deals with the effects of investment irreversibility, uncertainty and financial constraints on investment decisions and an extension of chapter eight that deals with the investment decisions by an investor-lender financial firm faced with cost-of-funds uncertainty. Since investment is divided into investment in financial assets such as stocks and other securities (see ch 7 and 8) and investment in fixed assets such as machinery, equipment and buildings (see ch 1), the extension of the previous chapters to include fixed investment of private business firms makes the problem of investment decisions under uncertainty more complete. The aggregation of fixed investment, i.e. investment in machinery, equipment and non-residential buildings by investor-lender firms and other private businesses firms gives the aggregate fixed private investment in an economy. The concern of this chapter is the empirical econometric analysis of the effects of uncertainty on aggregate fixed investment of private business firms using data from South Africa.

Contrary to the assumptions of certainty by classical economics, economic agents face various forms of uncertainty when they make important economic decisions. There are various sources of uncertainty. These are: political uncertainty, policy uncertainty, market or demand uncertainty, cost uncertainty and technical uncertainty. Measures of political uncertainty include, among others, lack of government stability, lack of

* This chapter is an updated version of Kumo (2006).

property rights, fear of conflict and so on while policy uncertainty refers to constant shifts in macroeconomic policy such as interest rate and exchange rates. Cost uncertainty refers to fluctuations of input prices and hence cost of investment. It is also called input cost uncertainty. On the other hand, technical uncertainty refers to non-factor cost elements of investment such as time, effort and material required to complete a project (Dixit and Pindyck 1994:47). Some of these uncertainties are firm specific while others are macro level uncertainties.

Although the importance of uncertainty in economic decisions has been emphasized since the early twentieth century by Frank Knight (1921) and Keynes (1921,1936, 1937), the quantitative analysis of its effect on investment has been limited owing to the problem of perceived unmeasurability of uncertainty. However, the importance of uncertainty in investment decision has been re-emphasized by the new theory of investment under uncertainty which linked uncertainty to the irreversibility of most investment decisions and the real option of waiting or deferring investment to obtain more information about the future (see ch 5)

However, different theoretical models of the theory of irreversible investment under uncertainty have come up with different predictions regarding the impact of uncertainty on investment making the uncertainty-investment relationship an empirical problem. Most empirical studies on this issue have focused on the firm level effects of uncertainty. Few studies that have captured the impact of uncertainty on aggregate investment have used unconditional sample variance of a single proxy variable to measure uncertainty.

The objective of the present chapter is to investigate the link between macro level uncertainty and aggregate fixed investment of private business firms in South Africa using five different measures of time varying uncertainty.

South Africa is one of the emerging market economies that is characterized by a high level of investment and economic growth. However, it was also a country characterized by a unique history of prolonged political instability and a high degree of government economic control until 1994 and a democratic transition during the post 1994 period. Thus, South Africa is a transitional economy that is characterized by a high level of risk and macroeconomic uncertainty. The current political stability means not only that increased government expenditure may increase volatility of aggregate demand but also various policy shifts corresponding to the transitional nature of the economy may lead to more uncertainty to the private investors.

Previous empirical studies on the impact of uncertainty on investment in South Africa have focused primarily on the impact of political uncertainty on investment. Bleaney (1994) studied the impact of political uncertainty on private investment in South Africa during the pre-1994 period and concluded that political uncertainty measured by the financial rand discount and real net capital flows had a negative impact on investment. On the other hand, Fielding (1997) investigated the impact of uncertainty on investment on traded and non-traded capital goods by using indexes of political and economic uncertainty. Fielding (1999, 2002) investigated the impact of uncertainty on aggregate investment by using three measures of political uncertainty, i.e. political instability, political rights and property rights in addition to economic indices. Fedderke (2000) used the same political indexes in his study of the manufacturing sector investment. These studies have shown that the impact of political uncertainty on investment was negative.

The present chapter is an updated version of Kumo (2006) which differs from the previous studies in two respects. First, while all previous empirical studies for South Africa have focused on political uncertainty, Kumo (2006) focused entirely on macroeconomic uncertainty. Second, the latter study employed different measures of uncertainty. Instead of using a sample variance of a single proxy variable, the study employed GARCH generated volatility measures of five different macroeconomic variables as measures of uncertainty. These variables are output growth, real interest rate, real effective exchange rate, producer price inflation and relative prices of export to import or terms of trade. The current chapter updates Kumo (2006) in two ways: (a) in using general-to-specific model selection instead of using four lags based on quarterly data, and (b) by extending the period of time series data used in the analysis by 10 quarters from 2003 (Q3) to 2006(Q1).

The chapter is organized as follows: Section 9.2 presents a review of macroeconomic uncertainty and aggregate investment. Section 9.3 deals with the sources and measures of macroeconomic uncertainty where empirical estimation of these measures is carried out using the GARCH volatility model. Section 9.4 is concerned with the econometric methodology and the data while section 9.5 presents the results of econometric estimation of the error correction model. Finally, section 9.6 presents a conclusion.

9.2 MACRO-LEVEL UNCERTAINTY AND AGGREGATE FIXED INVESTMENT OF FIRMS

While a growing literature has focused on uncertainty and irreversibility and their impact on investment decisions at the level of individuals firms, few studies investigated the impacts of irreversibility and uncertainty at macro level. However, some authors have emphasized that since capital goods cannot retain their full values on the second hand market, investment irreversibility is more realistic at the aggregate level than at the micro level (Bertola and Cabbalero 1991:1).

However, there are two main concerns in dealing with investment irreversibility at macro level. First, while incorporating irreversibility into econometric models of aggregate investment is not a simple task, more problems arise if we want to explain the long run equilibrium relationship between investment and uncertainty (Dixit and Pindyck 1994:421). Another concern is the problem of aggregation. The theory of irreversible investment applies most directly to firm level investments because different firms have different technologies as well as managerial capacities and subject to different uncertainties and have different action thresholds (Dixit and Pindyck, 1994:421). Under such difficulties, the available literature suggests two alternatives. The first is to use the threshold that triggers investment and assess its relation with uncertainty while the second option is to use the process of stochastic aggregation and assess the direct relationship between uncertainty and investment at micro or macro level.

Yet another concern is the importance of macro-level uncertainty. For instance, Carruth, et al (1998:7) state that if firm level uncertainty occurs at different times for different firms, then it is possible that these fluctuations may cancel each other out at the aggregate level. However, Bernanke (1983:103-104) argues that the effects of uncertainty may not disappear at the aggregate level because, first, macroeconomic factors such as uncertainty about future price levels, interest rates and exchange rates are important in determining firm-level investment decisions and second, if a firm is uncertain about the nature of the future uncertainty, it may delay investment and wait to get more information about the permanency of the uncertainty and this act of delay in the investment decision is propagated by the presence of the irreversibility constraint. Thus, in the words of Ben Bernanke (1983:103), as long as macro-uncertainty is important, then aggregate investment instability follows from our micro-analysis. This provides one of the justifications for the extension of firm level analysis of investor-lender behaviour to the analysis of the role of uncertainty on the aggregate investment of firms. The argument that micro-uncertainty does not disappear at the aggregate level is supported by previous empirical studies and the empirical results of the present study.

9.3 SOURCES AND MEASURES OF MACROECONOMIC UNCERTAINTY: THE GARCH ESTIMATION OF VOLATILITY MODELS

As stated in the introduction to this chapter, the present study identifies five main sources of uncertainty. These are: (a) uncertainty about changes in real effective exchange rate (REER), (b) uncertainty about changes in output growth (GDP growth) (c) uncertainty about changes in producer price inflation (Producer inflation), (d) uncertainty about changes in real interest rate (Real interest rate) and (e) terms of trade uncertainty (TOT). These sources of uncertainty are measured using the generalised autoregressive conditional hetroscedasticity (GARCH) volatility models.

The most popular model of volatility measure is the GARCH (1,1) process due to Bollerslev (1986). Using Bollerslev's specification, the GARCH (1,1) model can be stated as:

$$Y_t = a_{0,t} + a_{1,t}Y_{t-1} + \varepsilon_t \qquad (9.1)$$

$$\varepsilon_t = H^{1/2}{}_t u_t, \quad u_t \sim N(0,1), \qquad (9.2)$$

and

$$H_t = a_0 + a_1\varepsilon^2{}_{t-1} + b_1 H_{t-1} \qquad (9.3)$$

where, $\varepsilon_t \sim N(0, H_t)$

In this model conditional variance is measured as a sum of lagged squared residual and the lagged conditional variance itself. Some empirical specifications of the GARCH(1,1) models add a trend variable in equation (9.1). However, preliminary estimation for the present study has shown that there is no major difference in the values of the variances obtained by adding or omitting the trend variable. Equations (9.1)–(9.3) are used to estimate conditional variance measures of volatility which are used as measures of uncertainty for each of the five macroeconomic variables for the period 1975(4)-2006(1).

The PcGive 10.0 econometric package was used to estimate these equations to measure the volatility of the five macroeconomic variables. The estimation is carried out using the maximum likelihood estimation (MLE) technique with two restrictions. These are (a) the imposition of stationarity, i.e. $\alpha + \beta < 1$ and (b) the imposition of nonnegativity, i.e. $\alpha i + \beta i \geq 0$. The unit root test was carried out using the augmented Dickey-Fuller (ADF) test for variables under consideration before the GARCH estimation. Stationarity for variables with unit roots was achieved by differencing once. The result of the GARCH (1,1) estimation for South African macroeconomic time series is reported in Table 9.1 below.

Table 9.1 The estimation results of GARCH (1,1) volatility models for South Africa, 1975(4)-2006(1)

Sources of uncertainty	Coefficients and t-values of GARCH (1,1) model[a]				
	Intercept	Variable_1	alpha_0	alpha_1	beta_1
REER	7.46244 (3.47)***	0.941196 (55.3)***	40.8770 (2.94)***	0.349218 (2.50)**	-0.172535 (-2.34)**
GDP growth	0.0125098 (0.202)	-0.298963 (-3.14)***	9.21462e-016 (00)	0.0551409 (1.17)	0.926923 (17.0)***
Producer inflation	-0.0206031 (-0.176)	-0.217235 (-2.35)**	0.573976 (0.679)	0.115630 (1.29)	0.535414 (0.980)
Real interest rate	0.106921 (0.327)	0.996087 (44.2)***	0.756975 (1.39)	0.249608 (1.19)	0.255109 (0.833)
Terms of trade	0.220778 (3.89)***	0.794948 (15.1)***	0.0002319 (1.37)	0.368681 (1.96)*	0.618987 (3.94)***

a values in brackets are t-values

***, ** and * refer to significance at 1%, 5% and 10% respectively

The PcGive 10.0 output of the volatility estimation for the five macroeconomic variables reported in Table 9.1 above indicates that most variables have strong GARCH effects. The strongest GARCH (1,1) effect is observed for the REER time series. The coefficients of both alpha_1 and beta_1 are significant at 5% level while the coefficient of the lagged REER term is significant at 1% level. Although the coefficient of beta_1 is negative for REER, the nonnegativity imposition is still maintained because alpha(1) + beta(1) is greater than zero.

Three tests are carried out for GARCH (1,1) model adequacy. These are: normality test, ARCH 1-2 test and portmanteau (12) test. For REER time series, although the portmanteau (12) and ARCH 1-2 tests show that residuals are not serially correlated and there are no problems of autoregressive hetroscedasticity respectively, the residuals are nonnormally distributed. However, the nonnormality of residuals is not expected to cause a major problem for the current sample size.

Another strong GARCH effect is observed in the terms of trade time series. For this series, both the coefficients of the lagged terms of trade and beta_1 are significant at 1% level while the alpha_1 coefficient is significant at 10% level. Moreover, for this time series the sum of alpha_1 and beta_1 is close to unity indicating that the persistence of the conditional variance on terms of trade time series is very high. The model adequacy test for this series did not show any problem except the nonnormality of residuals.

The third macroeconomic time series with significant GARCH effects is the GDP growth time series. For this series, both the lagged coefficient for GDP growth and the beta_1 coefficients are significant at 1% level. Although, the alpha_1 coefficient is not significant, the sum of the alpha and beta parameters are found to be quite close to unity indicating that the persistence of the conditional variance on GDP growth is also high. The model adequacy test indicated no deficiencies except the nonnormality of residuals.

On the other hand, the producer inflation and the real interest rate time series have GARCH effects but the coefficients of both alpha_1 and beta_1 are not significant. However, the coefficients of the lagged producer inflation and lagged real interest rate are significant at 5% and 1% respectively. Although the alpha and beta coefficients indicate that there are no significant GARCH effects, it is possible to obtain time varying conditional variance for both series which are included in our investment equations for the purpose of comparison. Some econometricians ignore the significance of alpha and beta coefficients and use the resulting conational variance as an input into the regression equations. But the result of this must be interpreted cautiously.

The conditional hetroscedasticity (Ht's) obtained by the above five volatility measures are used to estimate the neoclassical investment equation (9.4) that includes both conventional and uncertainty augmented determinants of fixed investment. The identification of GARCH effects in the South African macroeconomic times series is one of the original contributions of the present study.

9.4 ECONOMETRIC METHODOLOGY AND THE DATA

Empirical studies of uncertainty-investment relationship have used various econometric methods based on the assumption that this relationship is linear. For instance, Asteriou and Price (2000) and Fedderke (2000) used dynamic panel models while Fielding (1999, 2002) used time series models. Most empirical studies used cross section investment models with the linearity assumption. The only exception in the empirical literature is Bo and Lensink (2001) where they modeled investment-uncertainty relationship as nonlinear.

The present study models the relationship between investment and conventional determinants in an error correction model (ECM) framework controlling for the effects of uncertainty. The advantage of the ECM approach is that it combines both the short-term dynamics and the long run equilibrium relationship between investment and its determinants.

9.4.1 The model

Empirical studies investigating the relationship between uncertainty and investment used various measures of investment. Almost all studies used either the ratio of investment to GDP, or both the level and the change in fixed capital stock as measures of investment. For instance, Asteriou and Price (2000) used change in capital stock (ΔK) as a measure of investment in their study of the impact of uncertainty on investment and growth while Fedderke (2000) equates net investment with changes in capital stock (ΔK) in his investment equation. On the other hand, Fielding (1999, 2002) used equation of capital stock in his study of the impact of political uncertainty on aggregate fixed investment in South Africa. Using the equations of capital stock instead of equations of investment is more appropriate in time series studies because in many cases quarterly (or even annual) changes in the levels of fixed investment are so insignificant that they mask the true relationship between investment and its determinants. Moreover, since fixed investment is a net addition to capital stock and is a short-term phenomenon, its time series behaviour is better captured by using models of capital stock which provides a long-term measure of the act of accumulating fixed capital over time. The present study uses a model of capital stock in line with Fielding (1999, 2002) to investigate the impact of macroeconomic uncertainty on aggregate fixed investment in South Africa.

The simple neoclassical type empirical specification of the equation of capital stock relating aggregate private fixed capital stock to conventional and uncertainty-augmented determinants can be specified as follows:

$$K_t = f(K_{t-1}, Z_t, H_t) + \varepsilon_t \qquad (9.4)$$

Where K is the log of aggregate private fixed capital stock at constant prices, Z is the set of conventional determinants of accumulation of private fixed capital stock related to the accelerator and neoclassical theories of investment, H is a set of measures of uncertainty related to the new theory of investment and ε is a random disturbance. The relationship between accumulation of capital stock and uncertainty is assumed to be linear. The expanded form of equation (9.4) becomes:

$$\ln K_t = \beta_0 + \beta_1 \ln K_{t-1} + \beta_2 \ln GDP_t + \beta_3 \ln REALr_t + \beta_4 \ln REALdcp_t +$$

$$\sum_{i=5}^{9}\sum_{j=1}^{5}\beta_i H_{jt} + u_t \qquad (9.5)$$

Where lnK is the log of private fixed capital stock, lnGDP is the logarithm of gross domestic product at constant prices, lnREALr is the logarithm of real interest rate representing part of the user cost of capital in line with the neoclassical theory of investment, lnREALdcp is the logarithm of real domestic credit to the private sector; this variable is used in the investment equation to measure the impact of financial constraints on fixed investment under uncertainty, $H_{t's}$ are the five measures of macroeconomic uncertainty obtained using the GARCH (1,1) process. The uncertainty measures refer to volatility of real effective exchange rate (H_{tREER}), volatility of GDP growth ($H_{tGDPgrth}$), terms of trade volatility (H_{tTOT}), volatility of real interest rate (H_{tREALr}), and volatility of producer inflation ($H_{tProdINF}$) for the period 1975(4)-2006(1).

Equation (9.5) is estimated using the Ordinary Least Squares (OLS) technique of estimation. Before the estimation of the equation unit root tests were carried for the time series using Augmented Dickey-Fuller (ADF) test. Tests were also carried out to determine the presence of long run equilibrium co-integrating relationship among the conventional variables of the long run model. Pesaran et al (2001) bounds test was used to analyze the long run cointegrating relationship among the conventional macroeconomic determinants of investment.

9.4.2 Testing for the stationarity of the time series

To determine the stationarity of the conventional (nonuncertainty) macroeconomic variables used in the regression equations, unit root test was carried out using Augmented Dickey-Fuller (ADF) test. The variables are the logarithm of private business fixed capital stock, lnK, the logarithm of real domestic credit to the private sector, lnREALdcp, the logarithm of real interest rate, lnREALr, and the logarithm of the GDP, lnGDP. Table 9.2 presents the results of the Augmented Dickey-Fuller unit root tests for the four macroeconomic variables. The measures of uncertainty $H_{t's}$ are all stationary because the restricted GARCH (1,1) process imposes both the stationarity as well as nonnegativity of the conditional variances obtained by this process.

Table 9.2 Augmented Dickey-Fuller unit root tests for the conventional macroeconomic variables in levels including constant and trend

Variable	ADF statistic	Point (%)	ADF critical value	Comment
lnK	-1.78	5	-3.449	unit root
lnGDP	-1.102	5	-3.449	unit root
lnREALr	-2.558	5	-3.449	unit root
lnREALdcp	-3.117	5	-3.449	unit root

As indicated in Table 9.2 all the macroeconomic time series have unit roots. However, from this information alone it is not possible to determine the order of integration of the variables. For this reason we differenced the variables once and carried out unit root test using the augmented Dickey-Fuller unit root tests with constant but no trend. The result of the test is presented in Table 9.3 below.

Table 9.3 Augmented Dickey-Fuller unit root tests for the conventional macroeconomic variables in first differences including constant but no trend

Variable	ADF statistic	Point (%)	ADF critical value	Comment
ΔlnK	-3.129*	5	-2.887	no unit root
ΔlnGDP	-4.316**	1	-3.488	no unit root
ΔlnREAL	-4.731**	1	-3.488	no unit root
ΔlnREALdcp	-5.045**	1	-3.488	no unit root

** and * refer to significance at 1% and 5% level respectively.

The result showed that the variables do not have unit roots in their first differences. That is all the variables are integrated of order one, i.e. they are all I(1) processes. The lagged length used in the test of unit root for both level values and their first differences is 4.

However, it is not possible to carry out regression in levels for variables that follow an I(1) process. Thus, we need to check whether they are co-integrated or not. If they are co-integrated, it is possible to carry out regression analysis with levels without worrying about spurious regression (see Banerjee et al. 1993). The following section presents tests for cointegration using bounds testing based on the unrestricted error correction model of Pesaran et al (2001).

9.4.3 Testing for Co-integration

The results of the unit root tests in the previous section indicate that all of the variables have unit roots in levels but all of them are stationary in their first differences. That is all the variables are I(1) processes. Thus, when I (1) variables or the combinations of I(0) and I(1) variables are used in the regression equations, bounds test is the most appropriate technique to analyze the long run relationships between these variables.

Therefore, the present study uses Pesaran et al (2001) bounds test based on the unrestricted error correction model (UECM) for the test of cointegration of the variables in question. The empirical UECM model for the capital stock equation and its conventional macroeconomic determinants for South Africa is specified as follows:

$$\Delta \ln(K)_t = a_0 + \sum_{1}^{K} a_i \Delta \ln(K)_{t-i} + \sum_{0}^{K} b_i \Delta \ln(GDP)_{t-i} + \sum_{0}^{K} d_i \Delta (REALr)_{t-i} +$$

$$\sum_{0}^{K} \gamma_i \Delta \ln(REALdcp)_{t-i} + a^* \ln(K)_{t-1} + b^* \ln(GDP)_{t-1} +$$

$$d^*(\text{REALr})_{t-1} + \gamma^*\ln(\text{REALdcp})_{t-1} + u_t \qquad (9.6)$$

Where the variables are as defined earlier. K is the lag length. The lag length (K) for equation (9.6) is determined using the general-to-specific modelling approach (see the next section for the details). Accordingly, K is determined to be 9. All the letters preceding the variables are parameters while u_t is the regression residual.

Pesaran et al (2001:300-301) carried out bounds test (calculated upper and lower bounds) for the critical values of an F–test for the joint significance of the stated parameters. The bounds test is carried out as follows: the null hypothesis is tested by considering the UECM for the capital stock (9.6) excluding the lagged variables, i.e. $\ln(K)_{t-1}$, $\ln(GDP)_{t-1}$, $Ln(REALr)_{t-1}$, and $\ln(REALdcp)_{t-1}$. In other words, we perform a joint significance test for the coefficients of the lagged levels of the variables where the null and alternative hypotheses are:

H_o: $a^* = b^* = d^* = \gamma^* = 0$ against the alternative

H_A: $a^* \neq b^* \neq d^* \neq \gamma^* \neq 0$

The null hypothesis of the test is that there is no long run relationship between the I(1) variables. Pesaran et al (2001:300-301) provide two asymptotic critical values. These are lower bound values or I(0) which assumes that the regressors are I(0) and the upper bound values I(1) which assumes that the regressors are I(1). Thus, for some α level of significance, if the F statistic falls out side the critical bound, a conclusive inference can be made without considering the order of integration of the explanatory variables. Thus, if the F-statistic is higher than the critical bound, the null hypothesis of no cointegration is rejected. However, if the calculated F-statistic falls in-between the upper and the lower bounds, the result is inconclusive. Since the upper bound corresponds to the case that all variables are I(1) while the lower bound to the case that all the variables are I(0), in the latter case the order of integration of the explanatory variables must be known before any inferences can be made. On the other hand, if the calculated value of F falls below the lower critical bound we do not reject the null hypothesis. If it falls above the upper critical bound, then the variables are cointegrated.

Table 9.4 Bounds Testing for Cointegration Analysis, 1979(2)-2006(1)*

Dependent Variable	R^2 unrestricted	R^2 restricted	F-stat	Critical bounds value**		Comment
				I(0)	I(1)	
LnK	0.85016	0.817283	4.28	2.14	3.30	Cointegrated
LnREALr	0.659303	0.612242	2.66	2.14	3.30	Inconclusive
LnREALdcp	0.517437	0.441963	3.01	2.14	3.30	Inconclusive
LnGDP	0.671682	0.605893	3.86	2.14	3.30	Cointegrated

* No of observations = 107; Number of parameters = 44
**Critical bounds values are at 5% level of significance and lag length K = 9 for all UECM equations

Critical values of bounds tests are taken from Pesaran et al (2001) P.300 Table CI(iii) Case III.

The result of the bounds test indicates that the capital stock (lnK) and its macroeconomic determinants are cointegrated. We also find another cointegrating vector when output (LnGDP) is the dependent variable. However, the evidence is inconclusive when real interest rate and real domestic credit to the private sector are dependent variables. Since the capital stock (LnK) and its macroeconomic determinants are cointegrated, we can formulate the investment equation in an error correction framework. This is done in the next section.

9.4.4 The empirical error correction equation

The long run form of capital stock equation (9.5) relating the logarithm of aggregate private fixed capital stock to the neoclassical determinants as well as the uncertainty-augmented determinants takes the form of:

$$\ln K_t = \beta_0 + \beta_1 \ln GDP_t + \beta_2 \ln Realr_t + \beta_3 \ln REALdpc_t + \beta_4 H_{tGDPgrth} + \beta_5 H_{tREER}$$

$$+ \beta_6 H_{tREALr} + \beta_7 H_{tTOT} + \beta_8 H_{tProdINF} + u_t \qquad (9.7)$$

The implied error correction model for conventional macroeconomic determinants is presented in equation (9.8). Since information criteria do not often provide conclusive results for lag length selection when the sample size is small we employed general-to-specific (*gets*) modeling in selecting the lag length for the empirical ECM. The *gets* approach was developed during the last decades predominately in a single equation framework and has been applied mainly for single equation time series modelling (Lütkepohl 2007:319). Thus, since the present study uses single equation time series modelling, the *gets* approach is appropriate for the model selection for the ECM equation. Moreover, Hendry and Krolzig (2001) programmed the *gets* modelling in an automated format in PcGets software. The *gets* modelling is reliable in lag length selection because it is a data-dependent method involving modelling starting with a very general model having a large number of lagged values of the variables in question and using a testing down procedure based on the significance of the coefficient of the variable with the longest lag. Accordingly, the lag length for the empirical ECM equation (9.7) is determined to be 12, i.e. K= 12. For this model, most coefficients are significant at the 12th lag and the model passed most of the diagnostic tests. Using 12 quarterly lags (the maximum of 3 years lag) for each variable is reasonable because the period is not too long for the effects of the variables to disappear. On the other hand, in the previous article, Kumo (2006), we selected the lag length based on the number of quarters used in the analysis.

$$\Delta \ln K_t = \sum_{i=0}^{K} \delta_{1i} \Delta \ln GDP_{t-i} + \sum_{i=0}^{K} \delta_{2i} \Delta \ln REALr_{t-i} + \sum_{i=0}^{K} \delta_{3i} \Delta \ln REALdcp_{t-i}$$

$$+ \sum_{i=0}^{K} \beta_{1i} H_{tREERt-i'} + \sum_{i=0}^{K} \beta_{2i} H_{tREALr_{t-i}} + \sum_{i=0}^{K} \beta_{3i} H_{tTOT_{t-i}} + \sum_{i=o}^{K} \beta_{4i} H_{tGDPgrth_{t-i}}$$

$$+\sum_{i=0}^{K}\beta_{5i}H_{t\,\mathrm{ProdINF}_{t-i}} + \gamma Z_t + u_t \qquad (9.8)$$

Where Z_t is the error correction term for conventional determinants and γ is the speed of adjustment, i.e. how the variable K_t changes in response to disequilibrium.

The ECM equation (9.8) contains both dynamic short run changes and the long run adjustment processes for conventional macroeconomic determinants of the accumulation of fixed capital stock. It also contains the level values of five macroeconomic uncertainty measures obtained using the GARCH (1,1) process. The uncertainty measures are entered in their levels instead of changes because they are all stationary variables. We used empirical ECM model (9.8) to estimate the impact of these variables and the variables of macroeconomic uncertainty on the behaviour of private fixed capital accumulation in South Africa during the stated period.

However, the error correction term, Z_t is generated with out including the 5 uncertainty variables included in equation (9.8). Thus, to test for the significance of the error correction term, equation (9.8) is estimated with out the five uncertainty variables. The result is reported in Table 9.5 below.

The result of the ECM estimation for the capital stock and its macroeconomic determinants supports the results of our bounds test that the capital stock and its macroeconomic determinants are cointegrated. This is confirmed by the significance of the error correction term, Z_t. This variable is significant at the 5% level.

The diagnostic tests for the error correction model estimated excluding the 5 measures of uncertainty does not reject the presence of ARCH 1-4 effects and the nonnormality of residuals. However, there are no problems of hetroscedasticity and the functional form is correct.

Table 9.5 The estimated result of the ECM model for capital stock and macroeconomic determinants, 1978(3)-2006(1)

Variable	Coefficient	Std.Error	t-value [prob]
Intercept	0.00682372	0.0004481	15.2 [0.000]
ΔlnREALdcp	0.00934655	0.008599	1.09[0.280]
ΔlnREALdcp_1	0.00669169	0.008612	0.777[0.439]
ΔlnREALr	0.0115778	0.005695	2.03[0.045]
ΔlnREALr_1	0.00767718	0.005704	1.35[0.181]
ΔlnGDP	0.0305032	0.04371	0.698[0.487]
ΔlnGDP_1	-0.0523786	0.04459	-1.17[0.243]
z_t	-0.0249631	0.01073	-2.33[0.022]

Dependent variable = ΔlnK Mean dependent variable = 0.00681598
sigma = 0.00341325 Residual Sum of Squares = 0.00119998082
R Squared = 0.173574 F(7,103) = 3.09 [0.005]**
No of observations included =111 Durbin-Watson Statistic = 0.538

Diagnostic tests

Test Statistic		F –version[Prob}
ARCH 1-4 test:	F(4,95) =	33.470 [0.0000]**
Normality test:	Chi^2(2) =	77.981 [0.0000]**
hetero test:	F(14,88) =	0.92990 [0.5306]
hetero-X test:	F(35,67) =	0.43729 [0.9955]
RESET test:	F(1,102) =	1.0872 [0.2996]

9.4.5 The data

The present study made use of adjusted quarterly time series data for South Africa for the period 1975:Q1-2006:Q1, i.e. a total of 126 observations. The data used in the present study are obtained from various issues of IFS of the IMF and Quarterly Bulletin of the South African Reserve Bank.

Aggregate private business fixed capital stock (K) includes investment in machinery, equipment and non-residential structures. These figures are available only annually. These figures are converted from lower frequency to higher frequency using the cubic match method of Eviews 5.1 econometric package. The figures are at 2000 constant prices.

GDP is expressed in seasonally adjusted annualized figures at constant 2000 prices. Real interest rate (REALr) is calculated using nominal interest rate minus the inflation rate divided by 1 plus the inflation rate*. Real domestic credit to the private sector (REALdcp) is obtained by deflating the nominal domestic credit to the private sector by the consumer price index, i.e. by dividing the nominal quarterly values by the consumer price index divided by 100.

GDP growth (GDPgrth) refers to the annual average growth rate of gross domestic product at constant prices. Producer inflation rate (ProdINF) is measured by quarterly change in the producer price index provided in the IFS database of the IMF. The exchange rate is measured by quarterly index of real effective exchange rates (REER) provided in the IFS database. The terms of trade (TOT) is calculated as the ratio of the quarterly index of the price of exports over that of imports.

9.5 THE ESTIMATION RESULTS

The result of the empirical estimation of the investment model showed that macroeconomic uncertainty has significant negative impact on aggregate private fixed investment in South Africa. The result is reported in Table 9.6. In particular, the uncertainty about changes in real effective exchange rate (H_{tREER}) is significant and persistent. Both current REER uncertainty and lagged REER uncertainty is significant. REER uncertainty during the past 3 years seem to be as important as the current

---

* More specifically real interest rate, $REALr_t$, is calculated as follows:
$REALr_t = i_t - [(p_t - p_{t-1})/p_{t-1}] / [1 + (p_t - p_{t-1})/p_{t-1}]$,
Where, i_t is nominal lending rate.

uncertainty since the coefficients for the 12^{th} lag and the current uncertainty are significant at 10% and 5% respectively. This is in line with the previous empirical finding in Kumo (2006).

Another source of uncertainty that has significant effects on the private fixed investment is the terms of trade uncertainty (H_{tTOT}). The coefficients of this measure of uncertainty are significant at 10% for the first, seventh and twelfth lags. On the other hand, output growth uncertainty has the wrong sign but no significant effect. The difference between this result and the result in Kumo (2006) could be attributed to difference in model selection criteria as well as the changes in the period of the time series data used for the analysis.

In addition to the three measures of uncertainty, the uncertainty about the producer price inflation ($H_{tProdINF}$) and the uncertainty about changes in real interest rate (H_{tREALr}) have significant negative effects on the aggregate private fixed investment in South Africa for the period under consideration. For the uncertainty about producer price inflation, both current and the lagged measures of uncertainty have significant negative effects on investment. The coefficients of current and the fifth lag for uncertainty about producer price inflation are significant at 5% level. For uncertainty about changes in real interest rate, the coefficients of the first as well as the twelfth lag are significant at 10% level.

The results of the estimated ECM model also show that conventional determinants such as output, real interest rate and availability of domestic credit to the private sector have significant effect on the private fixed investment. In the short run, both output and domestic credit to the private sector have significant positive effects on the level of fixed capital investment in South Africa.

On the other hand, the rise in the real interest rate depresses private fixed investment in the short run. However, the short term elasticities of fixed investment with respect to

Table 9.6 Private Fixed Capital Stock ECM equation for South Africa, 1979(2)-2006(1)

Variable	Coefficient	Std.Error	t-value[prob]
Intercept	0.0308126	0.009312	3.31 [0.080]
ΔlnREALdcp	0.0571853	0.01225	4.67 [0.043]
ΔlnREALdcp_1	0.0434956	0.01267	3.43 [0.075]
ΔlnREALdcp_2	0.0614679	0.01468	4.19 [0.053]
ΔlnREALdcp_4	0.0402296	0.01961	2.05 [0.177]
ΔlnREALr_12	-0.0498283	0.008419	-5.92 [0.027]
ΔlnGDP_4	0.326472	0.08641	3.78 [0.063]
ΔlnGDP_11	0.156202	0.07799	2.00 [0.183]
H_{tREER}	-9.56152e-005	1.377e-005	-6.94 [0.020]
H_{tREER_10}	-5.26746e-005	2.464e-005	-2.14 [0.166]
H_{tREER_12}	-6.02466e-005	1.841e-005	-3.27 [0.082]
H_{tREALr_1}	-0.00250856	0.0007975	-3.15 [0.088]
H_{tREALr_12}	-0.00118367	0.0002900	-4.08 [0.055]
$H_{tProdINF}$	-0.00732060	0.001449	-5.05 [0.037]
$H_{tProdINF\_5}$	-0.00928890	0.001367	-6.80 [0.021]
H_{tTOT_1}	-1.15863	0.3433	-3.37 [0.078]
H_{tTOT_7}	-0.876468	0.2672	-3.28 [0.082]
H_{tTOT_8}	-0.609294	0.2136	-2.85 [0.104]
H_{tTOT_12}	-0.779156	0.2502	-3.11 [0.089]
$H_{tGDPgrth}$	0.00425400	0.009010	0.472 [0.683]
Z_t	-0.118203	0.07468	-1.58 [0.254]

Dependent variable = ΔlnK | Mean dependent variable = 0.00684526
sigma = 0.000875 | Residual Sum of Squares = 1.53124973e-006
R Squared = 0.998943 | F-statistic F(105,2) = 18 [0.054]
No of observations included =108 | Durbin-Watson Statistic = 2.8
Log-likelihood = 822.619 | Akaike Information Criteria = -16.1086
Hannan-Quinn Criteria = -15.0412 | Schwartz criteria = -13.4761

Diagnostic Tests

Test statistic		F and Chi^2[prob.]
AR 1-1 test:	F(1,1) =	1.1885 [0.4725]
ARCH 1-1 test:	Chi^2(1) =	0.00000 [1.0000]
Normality test:	Chi^2(2) =	14.316 [0.0008]**
RESET test:	F(1,20) =	1.7457 [0.2013]

both domestic credit and real interest rate are very low, less than 0.1 while the short term elasticity with respect to output is only 0.3. Thus, the size of the short term effects of the conventional determinants on fixed investment is limited. However, the estimated investment model indicated that, there is short run empirical evidence in support of the accelerator and neoclassical models of investment while the long term determinants are still the measures of macroeconomic uncertainty. However, it is difficult to state the same about the long run impacts of conventional determinants.

The error correction term Z_t is not significant when the uncertainty variables are included in the error correction equation. This is because the latter variables are not used to generate the error correction term. They are included in the empirical model to control for the effects of macroeconomic uncertainty on the investment.

The diagnostic tests of the model estimated indicate that the model does not suffer from the AR 1-1 effects and ARCH 1-1 effects. Moreover, the estimated model has a correct functional form as indicated by the result of Ramsay Reset test. However, the residuals are not normally distributed.

Moreover, simple correlation analysis showed that various measures of macroeconomic uncertainty and aggregate private fixed investment are negatively correlated for South Africa during the stated period. The correlation matrix is inline with the results of the ECM model. Except the real interest rate uncertainty (H_{tREALr}), the remaining GARCH (1,1) measures of macroeconomic uncertainty are negatively correlated with the aggregate private fixed capital stock (see Table 9.7).

Table 9.7 Correlation Matrix for aggregate private fixed capital stock and macroeconomic uncertainty, i.e. GARCH (1, 1) conditional variances

	H_{tREER}	H_{tRealr}	H_{tTOT}	$H_{tGDPgrth}$	$H_{tProdINF}$	lnK
H_{tREER}	1.0000	-------	--------	-----	------	--------
H_{tREALr}	0.30369	1.0000	--------	-------	------	---------
H_{tTOT}	0.068695	-0.016981	1.0000	-------	------	---------
$H_{tGDPgrth}$	0.24227	0.0069589	0.40358	1.0000	------	---------
$H_{tProdINF}$	-0.043379	-0.011217	0.24617	0.15695	1.0000	--------
lnK	-0.15123	0.040551	-0.41502	-0.90052	-0.18935	1.0000

On the other hand, most measures of macroeconomic uncertainty are found to be positively correlated among themselves (see Table 9.8). Real interest rate uncertainty is negatively associated with both the terms of trade and producer inflation uncertainty while the later is negatively associated with the real exchange rate uncertainty. However, most measures of macroeconomic uncertainty are positively associated. This implies that the uncertainty variables reinforce each other and move in a mutually positive direction.

The data used in the empirical estimation was tested for the presence of structural break using a dummy variable with values 1 for the post-1994 period and 0 for the pre-1994 period. The coefficient of the dummy variable included in the empirical ECM equation (9.8) was found to be insignificant and hence the dummy variable was dropped from the model. This means that the same parameters are appropriate for both periods.

Table 9.8 Correlation Matrix among measures of macroeconomic uncertainty, i.e. GARCH (1, 1) conditional variances

	H_{tREER}	H_{tRealr}	H_{tTOT}	$H_{tGDPgrth}$	$H_{tProdinf}$
H_{tREER}	1.0000	-------	------	-------------	-------
H_{tREALr}	0.30369	1.0000	-----	------------	------
H_{tTOT}	0.068695	-0.016981	1.0000	--------	------
$H_{tGDPgrth}$	0.24227	0.0069589	0.40358	1.0000	-------
$H_{tProdINF}$	-0.043379	-0.011217	0.24617	0.15695	1.0000

9.6 CONCLUSION

The present chapter investigated the relationship between uncertainty and aggregate fixed investment of private business firms in South Africa. This chapter is an extension of the uncertainty and irreversible investment problem of chapter 5 and the investor-lender firm model of chapter 8. The results of the empirical estimation of the investment equation indicated that macroeconomic uncertainty had significant negative effects on aggregate fixed investment of private business firms in South Africa during the stated period. Uncertainty about changes in real effective exchange rates has highly significant negative effects on aggregate fixed investment of firms. The current as well as lagged measures of real effective exchange rate uncertainty have significant negative effects on investment. This implies that managing the volatility of the real effective rand exchange rate is more important in boosting private sector investment than controlling its actual rates in relation to foreign currencies.

Terms of trade uncertainty has significant negative effects on investment at both first and twelfth lag. Thus, uncertainty during the previous quarter and the past three years seem to be important in influencing current investment decisions.

Output growth uncertainty has the wrong sign but has no significant effects on investment. The change in the result obtained for output growth uncertainty for the previous study and for the current one could be because of changes in the model selection criteria and changes in the period of the time series data used in the analysis.

Uncertainty about producer price inflation and real interest rate uncertainty also had significant negative effects on investment during the period under consideration. However, this result must be interpreted cautiously because these two time series do not have significant time varying variance in the GARCH(1,1) model estimated. Further research is required to determine why GARCH models with insignificant beta coefficient have time varying variance.

The simple correlation analysis indicated that most measures of macroeconomic uncertainty are negatively associated with investment consistent with most empirical findings. Similarly, most of the uncertainty variables are found to be positively associated with themselves with the implication that they reinforce each other creating a possibility to use a single principal component to estimate the impact of over all uncertainty. Thus, results obtained by using real effective exchange rate uncertainty or terms of trade uncertainty as a principal component in the empirical model may be equally relevant.

The short term results of the ECM also provide some empirical support to the conventional views of accelerator and neoclassical theories on the determinants of the private investment behaviour. The result shows that both output and domestic credit to the private sector have significant positive effects on the level of fixed investment of private business firms. Real interest rate, a partial measure of user cost of capital, also has a significant depressing effect on aggregate private fixed investment in South Africa. However, it is difficult to make the same conclusions about the long run impact of conventional determinants on investment. In the long run, the main determinants of aggregate fixed investment of private business firms are still measures of macroeconomic uncertainty. This is in line with previous empirical findings.

Although the above results are encouraging, there are still some gaps that need to be addressed. The present study assumes linear relationship between uncertainty and investment. The assumption of nonlinear relationship between uncertainty and investment requires different modeling approach than employed here. Moreover, the uncertainty measures are approximated by the volatility of financial time series. Finding more accurate measures of uncertainty (if any) would need further research.

CHAPTER TEN

FINAL CONCLUSION

Contrary to the assumptions of certainty of old and modern classical economics, in real life economic agents face uncertainty in the form of both risk and ambiguity in their day to day decision-making processes. The use of the concepts of benefit and distance functions in choices under certainty links certainty and risk through equivalent and compensating benefit measures, where the latter is equated with the risk premium. For choices under risk, when preferences exhibit constant absolute risk aversion, economic decisions are independent of the level of initial wealth; for decreasing and increasing absolute risk aversion, on the other hand, economic decisions depend on the level of initial wealth. But risk aversion cannot be equated with smooth concavity of the utility curve. The utility function for an individual may exhibit a different curvature, i.e. it may be concave for lower levels of income, convex for medium levels and concave again for higher levels of income indicating risk aversion, risk seeking and risk aversion at different levels of wealth.

Individual preferences between acts cannot always be represented by additive probabilities. In the context of expected utility maximization, nonadditive probabilities reflect uncertainty (ambiguity) aversion. In this case, investors value their assets based on nonadditive probability measures. Moreover, when investments are irreversible and firms face financial constraints, uncertainty (ambiguity) can act as a major deterrent to investment because firms have an option to wait until the uncertainty is fully resolved before they commit their resources.

The present study investigated the behaviour of an investor-lender firm confronted with two types of decision problems. The problems involved choosing either to invest in riskfree financial assets, obtaining riskfree benefits, or investing in risky financial assets with uncertain future returns and facing the consequences of making such decisions. The study investigated the optimal investment decisions of the investor-lender (I-L) firm based on the assumption that the objective of the firm is the maximization of the expected utility of profit in line with the expected utility theory.

The results of the I-L firm model have shown that when the objective of the firm is the maximization of the expected utility of profit, the firm faces not only price and revenue uncertainty but also cost-of-funds uncertainty associated with the firm's uncertain deposit liability used to finance its lending activities. In the presence of cost-of-funds or deposit cost uncertainty, the optimal decision of the I-L firm involves lending its capital to other firms instead of investing it in riskfree financial assets. This decision takes place when the expected return from lending is equal to the expected marginal deposit cost.

The main weaknesses of previous studies on the investment decisions of financial firms is that while they accept that these firms face return risks, they assume that the associated costs are all perfectly known. The I-L firm model of the present study has shown that this is not the case and that optimal investment decisions of a profit maximizing financial firm under risk and uncertainty cannot be determined without incorporating cost-of-funds uncertainty in its decision processes.

The nonexpected utility models are not suitable for the analysis of investment behaviour of I-L firms. This is the case not only because of the more serious empirical deficiencies of these models, as alternative theories of choice under uncertainty, but also because, for these models, the objective of economic agents does not involve maximization of the expected utility of profit. This implies that, under these models, agents do not face cost uncertainty in their economic decisions. The result of the I-L model implies not only that the objective of profit maximization is an important real life decision problem, but also that when the firm strives to achieve this objective in the face of risk and uncertainty, it does not have optimal riskfree alternatives to choose from. In the words of Ross (1981:622), in real life, there are no riskfree situations in the sense that all situations involve some uncertainty in the form of "the background noise of life".

Although the nonexpected utility theories question the validity of the predictions of economic behaviour based on the assumptions of expected utility, there is no other alternative theory that serves as a comprehensive descriptive theory of choice under uncertainty at present. Further research is needed to develop a comprehensive alternative to expected utility theory.

The present study also investigated the effects of uncertainty on aggregate fixed investment of private business firms using the GARCH generated measures of uncertainty on five major macroeconomic time series in South Africa. These are: uncertainty about changes in real effective exchange rate, GDP growth uncertainty, terms of trade uncertainty, uncertainty about changes in real interest rate and producer price uncertainty. The results of the estimation of empirical error correction model for conventional determinants controlling for the effects of uncertainty showed that macroeconomic uncertainty had significant negative effects on aggregate fixed investment in South Africa during the period under consideration. This result is in line with previous empirical studies.

However, the results of the study are based on one crucial assumption. That is that the relationship between investment and uncertainty is assumed to be linear. An alternative assumption of nonlinear relationship between investment and uncertainty requires a different modelling approach. This is an avenue for further research.

REFERENCES

Abel, A. 1983. Optimal investment under uncertainty. *American Economic Review*, 73 (1):228-233.

Abel, A. and Eberly, JC. 1999. The effects of irreversibility and uncertainty on capital accumulation. *Journal of Monetary Economics*, 44(3):339-377.

Admati, A. and Pfleiderer, P. 1988. A theory of intraday patterns: volume and price variability. *Review of Financial Studies*, 1(1):3-40.

Allais, M. 1953. Le Comportement de l'homme rationnel devant le risque: critique des axioms et postulats de l'école Americaine. *Econometrica*, 21 (4): 503-546. French version with English summary.

Allais, M. 1979b. The so called Allais paradox and rational decisions under uncertainty. In *Expected utility hypothesis and the Allais paradox* edited by M. Allais and O. Hagen, Dordrecht, Holland: Reidel, 437-681.

Amariglio, J. and Ruccio, DF. 1995. Keynes, postmodernism, uncertainty, In *Keynes, knowledge and uncertainty,* edited by S. Dow & J. Hillard, Aldershot, England: Edward Elgar, 334-356.

Arrow, KJ. 1964. The role of securities in the optimal allocation of risk-bearing. *Review of Economic Studies*, 31(2):91-96.

Arrow, KJ. 1965. *Aspects of the Theory of Risk-Bearing.* Helsinki: Yrjö Hahnsson Foundation.

Arrow, KJ. and Debreu, G. 1954. Existence of an equilibrium for a competitive economy. *Econometrica*, 22 (3):265-290.

Arrow, KJ. and Hahn, FH. 1971. *General competitive analysis.* Amsterdam: North Holland.

Asteriou, D. and Price, S. 2000. Uncertainty, investment and economic growth: Evidence from a dynamic panel. (Online). Available from: http://www.cass.city.ac.uk/conferences/esrc2001/cp1.pdf

Banerjee, A., Dolado, J., Galbraith, J. and Hendry, DF. 1993. *Co-integration, error correction, and the econometric analysis of the non-stationary data.* Oxford: Oxford University Press.

Bar-Ilan, A. and Strange, WC. 1996. Investment lags. *American Economic Review*, 86 (3):610-622.

Battalio, RC., Kagel, JH. and Jiranyakul, K. 1990. Testing between alternative models of choice under uncertainty: some initial results. *Journal of Risk and Uncertainty*, 3:25-50.

Baum, C., Caglayan, M. and Ozkan, N. 2005. The second moment matters: the response of bank lending behaviour to macroeconomic uncertainty. *Boston College-Working Paper Number 521.*

Beazer, WF. 1975. *Optimization of bank portfolios.* London: Lexington Books.

Bernanke, BS. 1983. Irreversibility, uncertainty and cyclical investment. *Quarterly Journal of Economics*, 98:85-106.

Bertola, G. and Caballero, RJ. 1991. Irreversibility and aggregate investment. *NBER Working Paper* #3865.

Best, MH. 1990. *The new competition.* Cambridge: Polity Press.

Birnbaum, MH. 2005. A comparison of five models that predict violations of first-order stochastic dominance in risky decision-making. *Journal of Risk and Uncertainty*, 31(3):263-287.

Birnbaum, MH. 2004a. Causes of Allais common consequence paradoxes: an experimental dissection. *Journal of mathematical psychology*, 48:87-106.

Birnbaum, MH. and Navarrete, JB. 1998. Testing descriptive utility theories: violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty*, 17: 49-78.

Bleaney, M. 1994. Political uncertainty and private investment in South Africa. *South African Journal of Economics*, 62(3): 188-197.

Bo, H. and Lensink, R. 2001. Is investment-uncertainty relationship non-linear? An empirical analysis for the Netherlands. (Online). Available from: www.eco.rug.nl/medewerk/lensink/download/hong.pdf

Brealey, RA. and Myers, SC. 2003. *Capital investment and valuation.* New York: McGraw-Hill.

Brown, WS. 1988. *Macroeconomics.* New Jersey: Princeton Hall.

Butos, WN. 1997. Hayek and rational expectations, In *Austrian economics in debate*, edited by WB. Keizer, Tieben, B. and van Zijp, R. London: Routledge, 220-242.

Caballero, R. 1997. Aggregate investment. *NBER Working Paper* # 6264.

Caballero, R. 1999. Aggregate investment, In *Handbook of Macroeconomics*, edited by JB. Taylor and M.Woodford, Elsevier Science B.V., 1: 813-892.

Cantillo, C. and Wright, J. 2000. How do firms choose their lenders? An empirical investigation. *The Review of Financial Studies*, 13(1):155-189.

Camerer, C.1989. An experimental test of various generalized expected utility theories, *Journal of Risk and Uncertainty*, 2: 61-104.

Camerer, C. 2004. Prospect theory in the wild: evidence from the field, In *Advances in behavioural economics*, edited by CF. Camerer, Loewenstein, G. and Rabin, M. Princeton: Princeton University Press, 149-161.

Carruth, A., Dickerson, A. and Henley, A. 1998. Econometric modelling of UK aggregate investment: the role of profits and uncertainty. (Online). Available from: http://www.kent.ac.uk/economics/papers/papers-pdf/1998/9812.pdf

Cebenoyan, AS. and Strahan, PE. 2001. Risk management, capital structure and lending at banks. (Online). Available from: http://fic.wharton.upenn.edu/fic/papers/02/0209.pdf

Chambers, D. and Charnes, A. 1961. Intertemporal analysis and optimisation of bank portfolios. *Management Science*, 7:393-410.

Chandler, AD. 1992. Organizational capabilities and the economic history of the industrial enterprise. *The Journal of Economic Perspectives*, 6 (3):79-100.

Chavas, JP. 2004. *Risk analysis in theory and practice*. Amsterdam: Elsevier Academic Press.

Choquet, G. 1953. Theory of capacities. *Annales de l'Institut Fourier*, 5:131-292.

Clark, JM. 1917. Business acceleration and the law of demand: a technical factor in business cycle. *Journal of Political Economy*, 25:217-235.

Clark, JM. 1944. Additional note on business accelerator and the law of demand. *Readings in Business Cycle Theory*, Philadelphia: The Blakiston Company, 254-260.

Coase, RH. 1937. The nature of the firm. *Economica*, 4:386-405.

Coddington, A. 1982. Deficient foresight: A troublesome theme in Keynesian economics. *American Economic Review*, 72(3):480-487.

Davidson, P. 1991. Is probability theory relevant for uncertainty? A post Keynesian perspective. *Journal of Economic Perspectives*, 5(1):129-143.

Davidson, P. 1995. Uncertainty in economics, In *Keynes, knowledge and uncertainty*, edited by S. Dow and J. Hillard, Aldershot, England: Edward Elgar, 107-116.

Davidson, P. 1996. Some misunderstanding on uncertainty in modern classical economics, In *Uncertainty in economic thought*, edited by C. Schmidt, Cheltenham UK: Edward Elgar, 21-37.

Debreu, G. 1959. *Theory of value: an axiomatic analysis of economic equilibrium*. New York: Wiley.

Demsetz, H. 1991. The theory of the firm revisited, In *The theory of the firm* edited by OE. Williamson and SG. Winter, New York and Oxford: Oxford University Press, 159-178.

Diamond, PA. and Stiglitz, JE. 1974. Increases in risk and risk aversion. *Journal of Economic Theory*, 8(3):337-360.

Dixit, AK. and Pindyck, RS. 1994. *Investment under uncertainty*. Princeton: Princeton University Press.

Dow, J. and Werlang, SRC. 1992. Uncertainty aversion, risk aversion and, the optimal choice of portfolio. *Econometrica*, 60(1):197-204.

Dow, S. 1995. Uncertainty about uncertainty, In *Keynes, knowledge and uncertainty*, edited by S. Dow and J. Hillard, Aldershot, England: Edward Elgar, 117-136.

Drèze, JH. 1987. *Essays on economic decisions under uncertainty*. Cambridge: Cambridge University Press.

Driver, C. and Moreton, D. 1992. *Investment, expectations and uncertainty*. Oxford: Blackwell.

Eisner, R. 1963. Investment: facts and fancy. *American Economic Review*, 53(2): 237-246.

Ellsberg, D.1961. Risk, ambiguity and the Savage axioms. *Quarterly Journal of Economics*, 75:643-669.

Elyasiani, E. 1983. The two product banking firm under uncertainty. *Southern Economic Journal*, 49(4): 1002-1017.

Erb, CB., Harvey, CR. and Viskanta, TE. 1996. Political risk, economic risk and financial risk. *Financial Analysts Journal*, 6:29-46.

Erb, CB., Harvey, CR. and Viskanta, TE. 2000. The risk and expected returns of African equity investment, In *Investment and risk in Africa,* edited by P. Collier and C. Pattillo, London: Macmillan, 122-145.

Epstein, L. 1999. A definition of uncertainty aversion. *The Review of Economic Studies*, 66(3):579-608.

Fama, E. and French, K. 1988. Permanent and temporary components of stock prices. *Journal of Political Economy*, 96(2):246-273.

Farmer, M. 1995. Knowledgeability, actors and uncertain worlds, In *Keynes, knowledge and uncertainty*, edited by S. Dow and J.Hillard. Aldershot, England: Edward Elgar, 67-76.

Fedderke, J. 2000. Investment in fixed capital stock: the South African manufacturing industry 1970-1997. ERSA *Working Paper* No.16.

Ferderer, JP. 1993. The impact of uncertainty on aggregate investment spending: an empirical analysis. *Journal of Money, Credit and Banking*, 25(1): 30-48.

Ferson, WE. and Harvey, CR. 1998. Fundamental determinants of national equity market returns: a perspective on country risk and asset pricing. *Journal of Banking and Finance,* 21:1625-1665.

Fielding, D. 1997. Aggregate investment in South Africa: a model with implications for political reform. *Oxford Bulletin of Economics and Statistics,* 59(3): 349-369.

Fielding, D. 1999. Manufacturing investment in South Africa: a time series model. *Journal of Development Economics*,58: 405-27.

Fielding, D. 2002. Human rights, political instability and investment in South Africa: a note, *Journal of Development Economics*, 67: 173-180.

Fishburn, PC. 1988. *Nonlinear preference and utility theory*. London: The John Hopkins University Press.

Fisher, I. 1907. *The rate of interest,* New York: Macmillan, reprinted in 1982 by Garland, New York.

Fisher, I. 1930. *The theory of interest.* New York: Macmillan.

Fitzgibbons, A. 1995. Keynes's Policy Model, In *Keynes, knowledge and uncertainty*, edited by S. Dow and J. Hillard, Aldershot, England: Edward Elgar, 211-220.

Friedman, M. 1968. The role of monetary policy. *American Economic Review*, 58:1-17.

Friedman, M. and Savage, LJ. 1948. The utility analysis of choices involving risk. *The Journal of Political Economy*, 56(4):279-304.

Froot, KA. and Stein, JC. 1998. Risk management, capital budgeting and capital structure policy for financial institutions: an integrated approach. *Journal of Financial Economics*, 47:55-82.

Galbraith, JA. 1963. *The economics of banking operations: a Canadian study*. Montreal: McGill University Press.

Gale, D. and Hellwig, M. 1985. Incentive-compatible debt contracts: the one period problem. *Review of Economic Studies*, 52:647-663.

Geankoplos, JD. and Polemarchakis, HM. 1986. Existence, regularity, and constrained suboptimality of competitive allocations when the asset market is incomplete, In *Uncertainty, information and communication: essays in honour of Kenneth J. Arrow, Vol. III* edited by WP.Heller, Starr, RM. and Starret, DA., Cambridge: Cambridge University Press, 65-95.

Gerrard, B. 1995. Probability, uncertainty and behaviour: a Keynesian perspective, In *Keynes, knowledge and uncertainty*, edited by S. Dow and J. Hillard. Aldershot, England: Edward Elgar, 177-196.

Ghosal, V. and Loungani, P. 2000. The differential impact of uncertainty on investment in small and large business. *Review of Economics and Statistics*, 82:338-349.

Gilboa, I. 1987. Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, 16(1):65-88.

Gonzalez, R. and Wu, G. 1999. On the shape of probability weighting function. *Cognitive Psychology*, 38:129-166, Cited in MH.Birnbaum, 2005. A comparison of five models that predict violations of first-order stochastic dominance in risky decision-making. *Journal of Risk and Uncertainty*, 31(3):263-287.

Gordon, MJ., Paradis, GE. and Rorke, CH. 1972. Experimental evidence on alternative portfolio decision rules. *American Economic Review*, 62 (½): 107-118.

Guiso, L. and Parigi, G. 1999. Investment and demand uncertainty. *Quarterly Journal of Economics*, 114(1):185-227.

Gul. F. 1991. The theory of disappointment aversion. *Econometrica*, 59(3):667-686.

Gup, BE. and Kolari, JW. 2005. *Commercial banking: the management of risk.* New Jersey: John Wiley and sons Inc.

Haberler, G. 1931. Irving Fisher's "Theory of Interest". *The Quarterly Journal of Economics*, 45(3):499-516.

Harris, R. and Sollis, R. 2003. *Applied time series modeling and forecasting.* Willey and Sons.

Harsanyi, JC. 1977. *Rational behaviour and bargaining equilibrium in games and social situations.* Cambridge: Cambridge University Press.

Hart, OD. and Jaffe, DM.1973. On the application of portfolio theory to depository financial intermediaries. *Review of Economic Studies*, 41(1):129-147.

Hayashi, F.1982. Tobin's marginal Q and average Q: a neoclassical interpretation. *Econometrica*, 50(1):213-224.

Hayek, FA. 1945. The use of knowledge in society. *American Economic Review*, 35(4):519-530.

Heap, HS. 1989. *Rationality in economics.* Oxford: Basil Blackwell.

Heath, C. and Tversky, A . 1991. Preference and belief: ambiguity and competence in choice under uncertainty. *Journal of risk and uncertainty*, 4:5-28.

Heiner, RA. 1983. The origin of predictable behaviour. *American Economic Review*, 75:576-578.

Heiner, RA. 1985(a). The origin of predictable behaviour: further modelling and applications. *American Economic Review Proceedings*, 75:391-396.

Heiner, RA. 1985(b). Predictable behaviour: reply. *American Economic Review*, 75(3):579-585.

Heller, WP., Starr, RM. and Starret, DA. 1986. (eds). *Uncertainty, information and communication: essays in honour of Kenneth J. Arrow, Vol. III*, Cambridge: Cambridge University Press.

Hendry, DF. and Krolzig, HM. 2001. *Automatic Econometric Model Selection with PcGets.* Timberlake Consultants Press, London

Hirshleifer, J. 1966. Investment decisions under uncertainty: applications of the state-preference approach. *The Quarterly Journal of Economics*, 80(2):252-277.

Holt, RW. 2000. *Investment, irreversibility and financial constraint*. Ph.D Thesis, Chapter Five. University of Edinburgh.

Hyman, DN. 1972. A behavioural model for financial intermediation. *Economics and Business Bulletin*, 24:9 -17.

IMF, International Financial Statistics, Various Issues.

Jacobsson, D. and Andréosso, B. 1996. *Industrial economics and organization: a European perspective*. London: McGraw Hill.

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

Jorgenson, D. 1963. Capital theory and investment behaviour. *American Economic Review*, 53(2):247-259.

Kahneman, D. and Tversky, A. 1979. Prospect theory: An analysis of decision under risk. *Econometrica*, 47:263-291.

Kahneman, D. and Tversky, A. 2000. Advances in prospect theory: cumulative representation of uncertainty, In *Choices, values and frames*, edited by D. Kahneman and A. Tversky, Cambridge: Cambridge University Press, 44-66.

Kahneman, D., Knetsch, JL. and Thaler, RH. 2004. Experimental tests of the endowment effects and the Coase theorem, In *Advances in behavioural economics*, edited by CF. Camerer, Loewenstein, G. and Rabin, M., Princeton: Princeton University Press, 55-74.

Kamales, S. 1983. *Banking firm model: an integrated view of real resource and financial aspects of banking firms*, Ph.D. thesis. Oklahoma State University, University Microfilms International.

Kane, EJ. and Malkiel, EG. 1965. Bank portfolio allocation, deposit variability and the availability doctrine. *Quarterly Journal of Economics*, 79:113-134.

Karni, E. 1985. *Decision-making under uncertainty: the case of state-dependent preference*. Cambridge: Harvard University Press.

Keynes, JM. 1921. *A treatise on probability*. Reprinted as Collected Writings, 1973, Vol. VIII. London: Macmillan for the Royal Economic Society.

Keynes, JM. 1936. *The general theory of employment, interest and money*. London: Macmillan.

Keynes, JM. 1937. The general theory of employment. *Quarterly Journal of Economics*, 51:209-23.

Kirzner, IM. 1973. *Competition and entrepreneurship,* Chicago: University of Chicago Press.

Kirzner, IM. 1979. *Perception, opportunity and profit: studies in the theory of entrepreneurship*, Chicago: University of Chicago Press.

Kirzner, IM. 1985. *Discovery and the capitalist process*. Chicago: University of Chicago Press.

Klein, MA. 1971. A theory of the banking firm. *Journal of Money, Credit and Banking,* 3(1):205-218.

Knight, F. 1921. *Risk, uncertainty and profit.* Boston: Houghton Mifflin.

Knox, AD. 1952. The acceleration principle and the theory of investment: a survey. *Economica*, 19(75):269-297.

Kreps, DM. 1990. *Game theory and economic modeling*. Oxford: Clarendon Press.

Kumo, WL. 2006. Macroeconomic uncertainty and aggregate private investment in South Africa, *South African Journal of Economics,* 74(2): 190-204. Economic Society of South Africa and Blackwell publishers.

Lawson, T. 1985. Uncertainty and economic analysis. *The Economic Journal*, 95:909-927.

Levy, M., Levy, H. and Solomon, S. 2000. *Microscopic simulation of financial markets: from investor behaviour to market phenomenon.* London: Academic Press.

Levy, H. and Lim, KC. 1998. The economic significance of cross-sectional autoregressive model: further analysis. *Review of Quantitative Finance and Accounting*, 11:37-51.

Lichtenstein, S. and Slovic, P. 1971. Reversals of preference between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89:46-55.

Lindman, HR. 1971. Inconsistent preference among gambles. *Journal of Experimental Psychology,* 89: 390-397.

Littlechild, SC. 1986. Three types of market process, In *Economics as a process*, edited by RN. Langlois, New York: Cambridge University Press, 27-40.

Loomes, G. and Sugden, R. 1982. Regret theory: an alternative theory of rational choice under uncertainty. *Economics Journal*, 92:805-824.

Loomes, G. and Sugden, R.1986. Disappointment and dynamic consistency in choice under uncertainty. *Review of Economic Studies*, 53:271-282.

Loomes, G. and Sugden, R.1987. Some implications of a more general form of regret theory. *Journal of Economic Theory*, 41:270-287.

Lucas, RE. 1973. Some international evidence on output inflation tradeoffs. *American Economic Review*, 63(3): 326-34.

Lucas, RE. 1980. Methods and problems in business cycle theory. *Journal of money, credit and banking*, 12 (4), Part 2, Rational Expectations: 696-715.

Luce, RD. and Fishburn, PC. 1991. Rank-and-sign dependent linear utility models for finite first-order gambles. *Journal of Risk and Uncertainty*, 4:29-59.

Luenberger, DG. 1992. Benefit functions and duality. *Journal of Mathematical Economics*, 21:461-481.

Luenberger, DG. 1994. Dual Pareto efficiency. *Journal of Economic Theory*, 62:70-85.

Luenberger, DG. 1995. *Microeconomic theory*. New York: McGraw-Hill.

Lütkepohl, H. 2007. General-to-specific or specific-to-general modelling? An opinion on current econometric terminology. *Journal of Econometrics*, 136(1):319-324

Machina, JM. 1982. "Expected Utility" analysis without the independence axiom. *Econometrica*, 50(2):277-324.

Machina, JM. 1987. Choice under uncertainty: problems solved and unsolved. *Economic Perspectives*, 1(1):121-154.

Machina, JM. and Neilson, WS. 1987. The Rose characterization of risk aversion: strengthening and extension. *Econometrica*, 55(5):1139-1149.

Malmquist, S. 1953. Index numbers and indifference surfaces. *Trabajos de Estatistica*, 4: 209-242.

Markowitz, HM. 1952. Portfolio Selection. *Journal of Finance*, 7:77-91.

Markowitz, HM. 1959. *Portfolio selection: Efficient diversification of investments.* New York: Wiley and Sons.

Marris, R. 1963. A model of "managerial" enterprise. *Quarterly Journal of Economics*, 77:185-209.

Marris, R.1964. *The economic theory of managerial capitalism*, London: Macmillan.

Marsh, P. 1982. The choice between equity and debt: an empirical study. *The Journal of Finance*, 37(1):121-144.

McDonald, R. and Siegel, DR. 1986. The value of waiting to invest. *Quarterly Journal of Economics*, 101:707-728.

Mises, L. von. 1949. *Human action: A treatise on economics*, London: William Hodge and Co. Ltd, 3rd ed. 1966.

Modigliani, F. and Miller, MH. 1958. The cost of capital, corporation finance and theory of investment. *American Economic Review*, 48:261-297.

Muller, A.1998. Another tale of two tails: on characterization of comparative risk, *Journal of Risk and Uncertainty*, 16(2):187-197.

Neumann, J. von. and Morgenstern, O.1944. *Theory of games and economic behaviour.* Princeton, NJ: Princeton University Press; 2nd ed. 1947.

Nickell, SJ. 1978. *The investment decisions of firms.* Cambridge: Cambridge University Press.

O'Driscoll, JP. and Rizzo, MJ. 1985. *The economics of time and ignorance.* Oxford: Basil Blackwell.

Online Dictionary of the Social Sciences. 2002. (Online). Available from: http://bibucket.icaap.org/dict.pl

Palacios-Huerta, I. 1999. The aversion to the sequential resolution of uncertainty. *Journal of Risk and Uncertainty*, 18:249-269.

Parkin, M. 1970. Discount house portfolio and debt selection. *Review of Economic Studies*, 37:469-497.

Pattillo, C. 2000. Risk, financial constraint and equipment investment in Ghana: a firm level analysis, In *Investment and risk in Africa,* edited by P. Collier and C. Pattillo, London: Macmillan, 96-121.

Perlman, M. and McCain, CR. 1996. Varieties of uncertainty, In *Uncertainty in economic thought,* edited by C. Schmidt, Cheltenham, UK: Edward Elgar, 9-20.

Pesaran, MH., Shin, Y. and Smith, RJ. 2001. Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3):289-326.

Phelps, ES.1967. Phillips curves, expectations of inflation and optimal unemployment over time. *Econometrica*, 34:254-281.

Pindyck, RS. and Solimano, A.1993. Economic instability and aggregate investment. *NBER Working Paper* # 4380.

Porter, RC. 1961. A model of bank portfolio selection. *Cowles Foundation Paper* 168, Reprinted from *Yale Economic essays*, 1(2):323-359.

Pratt, JW. 1964. Risk aversion in the small and in the large. *Econometrica*, 32(½):122-136.

Price, S., 1995. Aggregate uncertainty, capacity utilization and manufacturing investment. *Applied Economics*, 27:147-154.

Pyle, DH. 1971. On the theory of financial intermediation. *Journal of Finance*, 26:737-747.

Quagliariello, M. 2006. Macroeconomic uncertainty and banks' lending decisions: the case of Italy. (Online). Available from:

http://www.york.ac.uk/depts/econ/documents/dp/0602.pdf

Quiggin, J. 1982. A theory of anticipated utility. *Journal of Economic Behaviour and Organization*, 3(4):323-343.

Quiggin, J. 1993. *Generalized expected utility theory: the rank-dependent model.* Boston: Kluwer Academic Publishers.

Quiggin, J. 1994. Regret theory with general choice sets. *Journal of Risk and Uncertainty*, 8:153-165.

Quiggin, J., and Chambers, RG. 1998. Risk premium and benefit measures for generalized-expected-utility theories. *Journal of Risk and Uncertainty*, 17:121-137.

Raiffa, H. 1968. *Decision analysis: introductory lectures on choices under uncertainty.* Reading, Massachusetts: Addison-Wesley.

Ross, S. 1981. Some stronger measures of risk aversion in the small and the large with applications. *Econometrica*, 49(3):621-638.

Rothschild, M. and Stiglitz, J. 1970. Increasing risk: I. a definition. *Journal of Economic Theory*, 2(3):225-243.

Rothschild, M. and Stiglitz, J. 1971. Increasing risk: II. Its economic consequences. *Journal of Economic Theory*, 3:66-84.

Runde, J. 1995. Risk, uncertainty and Bayesian decision theory: a Keynesian view, In *Keynes, knowledge and uncertainty,* edited by S. Dow and J. Hillard, Cheltenham: Edward Elgar, 197-210.

Sandmo, A. 1971. On the theory of the competitive firm under price uncertainty. *The American Economic Review*, 61(1):65-73.

Santomero, AM. 1984. Modeling the banking firm: a survey. *Journal of Money, Credit and Banking*, 16(4):576-602. Part 2: Bank Market Studies.

Sarin, R. and Walker, P .1998. Revealed likelihood and Knightian uncertainty. *Journal of Risk and Uncertainty*, 16:223-250.

Savage, LJ. 1954. *The foundations of statistics.* New York: Dover Publications.

Savage, LJ. 1972. *The foundations of statistics*, 2nd ed. New York: Dover Publications.

Schmeidler, D. 1989. Subjective probability and expected utility without additivity, *Econometrica*, 57:571-87.

Schmidt, C. 1996: Introduction: what is certainty in economic uncertainties? In *Uncertainty in economic thought*, edited by C. Schmidt. Cheltenham, UK: Edward Elgar, 1-6.

Sealey, CW. 1980. Deposit rate-setting, risk aversion and the theory of depository financial intermediaries. *Journal of Finance*, 35(5):1139-1154.

Segal, U. 1987. Some remarks on Quiggin's Anticipated Utility. *Journal of Economic Behaviour and Organization*, 8:145-154. North-Holland.

Sen, AK. 1985. Rationality and uncertainty. *Theory and Decision*, 18:109-127.

Serven, L. 1997. Uncertainty, instability and irreversible investment: theory, evidence and lessons for Africa. *Policy Research Working Paper 1722.* The World Bank.

Serven, L. 1998. Macroeconomic uncertainty and private investment in LDCs. (Online). Available from:

http://www.worldbank.org/html/dec/publications/workpapers/wp2000series/wp2035/wp2035.pdf

Shackle, GLS. 1961. *Decision order and time in human affairs.* Cambridge: Cambridge University Press.

Shepard, R. 1953. *Cost and Production Functions.* Princeton, NJ: Princeton University Press. Reprint of first edition 1981, Berlin: Spirnger-Verlag.

Shiller, RJ. 1981. Do stock returns move too much to be justified by subsequent changes in dividends? *American Economic Review,* 71(3):421- 436.

Simon, HA. 1978. Rationality as a process and as product of thought. *American Economic Review*, 68:1-16.

Simon, HA. 1982. *Models of bounded rationality: behavioural economics and business organization,* vol.2, Cambridge, MA: MIT Press.

Slovic, P. 1995. The construction of preferences. *American Psychologist,* 50:364-371.

Smith, A. 1776. An enquiry into the nature and causes of the wealth of nations. The Adam Smith Institute. (Online). Available from: http://www.adamsmith.org/smith/won-intro.htm

Starmer, C. 2004. Developments in nonexpected-utility theory: the hunt for a descriptive theory of choice under risk. In *Advances in behavioural economics*, edited by Camerer, et al., Princeton NJ: Princeton University Press, 104-147.

Starmer, C. and Sugden, R 1989. Violations of the independence axiom in common ratio problems: an experimental test of some competing hypotheses. *Annals of Operations Research*, 19:79-101.

Stigum, ML. 1976. Some further implications of profit maximization by a savings and loan association, *Journal of Finance*, 31:1405-1426.

Tobin, J. 1958. Liquidity preference as behaviour toward risk. *Review of Economic Studies*, 25:65-86.

Tobin, J. 1969. A general equilibrium approach to monetary theory. *Journal of Money Credit and Banking*, 1:15-29.

Townsend, R. 1979. Optimal contracts and competitive markets with costly state verification. *Journal of Economic Theory*, 21:265-293.

Trigeorgis, L. 1999. Real options: a primer. In *The new investment theory of real options and its implications for telecommunications economics*, edited by J. Alleman and E. Noam, Boston: Kluwer Academic Publishers, 3-33.

Tversky, A., Slovic, P. and Kahneman, D.1990. The causes of preference reversal. *American Economic Review*, 80(1):204-217.

Tversky, A. and R. Thaler. 1990. Preference reversals. *Journal of Economic Perspectives*, 4:201- 211.

Tversky, A. and Kahneman, D. 1986. Rational choices and the framing of decisions. *Journal of Business*, 59(4):S251-S278.

Tversky, A. and Kahneman, D. 1992. Advances in Prospect Theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5:297-323.

Tversky, A. and Wakker, P. 1995. Risk attitude and decision weights. *Econometrica*, 63, 1255-1280, Cited in MH. Birnbaum, 2005. A comparison of five models that predict violations of first-order stochastic dominance in risky decision-making. *Journal of Risk and Uncertainty*, 31(3):263-287.

Tversky, A. and Kahneman, D. 2000. Prospect theory: an analysis of the decision under risk. In *Choices, values and frames,* edited by D. Kahneman, and A. Tversky, Cambridge: Cambridge University Press, 17-43.

Tversky, A. and Fox, CR. 2000. Weighting risk and uncertainty. In *Choices, values, and frames* edited by D. Kahneman, and A. Tversky, Cambridge: Cambridge University Press, 93-117.

Vercammen, J. 2000. Irreversible investment under uncertainty and the threat of bankruptcy. *Economics Letters*, 66:319-325.

Wakker, P. and Tversky, A. 1993. An axiomatization of cumulative prospect theory. *Journal of Risk and Uncertainty*, 7: 147-176, Cited in MH. Birnbaum, 2005. A comparison of five models that predict violations of first-order stochastic dominance in risky decision-making. *Journal of Risk and Uncertainty*, 31(3):263-287.

Wood, JH. 1975. *Commercial bank loan and investment behaviour*. London: John Willey and Sons.

Worthington, PR. 1995. Investment, cash flow and sunk costs. *The Journal of Industrial Economics*, 43(1):49-62.

Wu, G. and Gonzalez, R. 1996. Curvature of the probability weighting function. *Management Science*, 42: 1676-1690, Cited in MH. Birnbaum, 2005. A comparison of five models that predict violations of first-order stochastic dominance in risky decision-making. *Journal of Risk and Uncertainty*, 31(3):263-287.

Wubben, E. 1995. Austrian economics and uncertainty: on non-deterministic but non-haphazard future. In *New perspectives on Austrian economics*, edited by G. Meijer, London: Routledge, 106-145.

Yaari, M.1987. The dual theory of choice under risk. *Econometrica*, 55(1): 95-115.

Zijp, van R. and Visser, H.1995. On the non-neutrality of mathematical formulation: Austrian vs. New Classical business cycle theories. In *New perspectives on Austrian economics*, edited by G. Meijer, London: Routledge, 178-194.