The Economics of Risk and Time

Christian Gollier
GREMAQ and IDEI, University of Toulouse

May 27, 1999

Contents

Introduction

Uncertainty is everywhere. There is no field in economics in which risk is not an important dimension of the decision-making environment. The theory of finance provides the most obvious example of this. Similarly, most recent developments in macroeconomics have been made possible by recognizing the importance of risk in explaining individual decisions. Consumption patterns, investments and labor decisions can only be understood completely if uncertainty is taken into account into the decision-making process. Environmental economics provides another illustration. Public opinion is now very sensitive to the presence of potentially catastrophic risks related, for example, to the greenhouse effect and genetic manipulations. Environmental economists introduced probabilistic scenarios in their models to exhibit socially efficient levels of prevention efforts. Finally, the extraordinary contributions of asymmetric information to game theory have heightened interest in uncertainty among economists.

We are lucky enough to have a well-accepted and unified framework to introduce uncertainty in economic modelling. Namely, we are indebted to John von Neumann and Oskar Morgenstern who in the mid-forties developed the expected utility theory building on Daniel Bernoulli's idea that agents facing risk maximize the expected value of the utility of their wealth. Expected utility theory (EU) is now fifty years old. It is a ubiquitous instrument in economic modelling. Most economists recognize that the theory has been very useful for explaining the functioning of our economies. The aim of this book is to provide a detailed analysis of the implications of the expected utility model to the economic theory.

It is important to note that expected utility theory does not provide any hint as to which specific utility function should be used by decision makers. We must confess that the empirical estimation of utility functions of real world decision makers is still in the early stage of development.

This precludes most economists from using particular utility functions to solve their specific problems. For example, because expected utility the study of optimal behavior in the presence of several sources of risk, very complex financial theory relies heavily on the mean-variance approximation — in which interactions between different risks are easily analyzed — of expected utility. This approximation is exact only for very specific utility functions, like quadratic or exponential ones. The same observation holds for macroeconomics, in which it has been possible to treat dynamic uncertainties only by restricting utility functions to very specific subsets of increasing and concave functions. Whereas the general framework of expected utility relies on a set of simple axioms that are intuitively

appealing, the selection of a subset of utility functions to apply this theory has been done on the basis of a purely technical convenience.

One of the innovation of this book is its refusal to use convenient but less sophisticated utility functions to solve in complex problems of decision and market equilibrium under uncertainty. The benefit of this method is twofold. First, it allows us to better understand the relevant basic concepts that determine optimal behavior under uncertainty. Second, it reestablishes the richness and diversity of the EU framework that tends to be forgotten in the strongly reductionist limitations imposed on the widely-used set of utility functions. In particular, this book can help in understanding some puzzles that have been exhibited in relation to observed financial prices by combining the EU theory with commonly used utility functions in finance.

This book provides an overview of the most recent developments in expected utility theory. It is aimed at any audience that is interested in problems related to efficient/optimal public/private strategies for dealing with risk. The book heavily relies on concepts that are standard in the theory of finance. But this fact should not hide the point that most findings presented in this book are useful for our understanding of various economic problems ranging from macroeconomic fluctuations to public policies towards global warming.

Part I of this book is devoted to the presentation of the EU model and of its basic related concepts. **Chapter 1** presents the axioms that underlie the EU approach. It proves the von Neumann-Morgenstern EU Theorem and discusses its main limitations. **Chapter 2** shows that the linearity of the EU objective function with respect to probabilities implies that simple mathematical tools may be used in this framework. The existence of such simple technical tools in the case of EU is central to our understanding of the success of the EU model in economics. Yet it raises doubt about the possibility that more general non-EU models will be adaptable for solving similar problems. These basic technical tools will be used in the next four parts of the book. The two important techniques originating from the diffidence theorem – a tool richer than Jensen's inequality – and from the properties of log-supermodular functions are examined. The use of these tools will allow us to unify a large range of different results that have appeared in the literature during the last thirty years. In **Chapter 3** we explain how to introduce risk aversion, and "more risk aversion", in the EU model. The notion of risk premium is introduced here, together with the familiar utility functions that are used in macroeconomics and finance. The discussion also helps the reader to assess his/her own degree of risk aversion. **Chapter 4** is devoted to the dual question of how to define risk, and "more risk", in this model. It also presents

some recent developments in the literature on stochastic dominance.

In Part II, we present probably the simplest decision problem of decision-making under uncertainty. More specifically, **Chapter 5** discusses the standard portfolio problem in which an investor has to allocate his wealth between a risk-free asset and a risky asset. We review the effect of a change in wealth, in the degree of risk aversion or in the degree of risk on the optimal demand for the risky asset. This model is also helpful for understanding the behavior of a policyholder insuring a given risk at an unfair premium, or the behavior of a risk-averse entrepreneur who must determine her productions capacity under price uncertainty. **Chapter 6** is devoted to an equilibrium version of this problem. We present the simplest version of the equity premium puzzle.

We examine several extensions to the standard portfolio problem in Part III. The standard portfolio model has been developed up to the early eighties by assuming that the portfolio risk was the unique source of risk borne by the consumer. In the real world, people have to manage and control several sources of risk simultaneously. To make our models more realistic, Part III deals with how the presence of one risk affects behavior towards other independent risks. **Chapter 7** starts with an analysis of how the presence of an exogenous risk affecting background wealth may force us to refine the notion of comparative risk aversion that has been developed in Chapter 3. **Chapter 8** discusses how the presence of an exogenous risk in background wealth affects the demand for other independent risks. New notions like properness, standardness and risk vulnerability are introduced and discussed. **Chapter 9** addresses the difficult problem of whether independent risks can be substitute. To illustrate this, we characterize the conditions under which the opportunity to invest in one risky asset reduces the optimal demand for another independent risky asset.

The remaining of this Part deals with how to use dynamic programming methods to show how the time horizon length affects the optimal in behavior towards repeated risks. **Chapter 10** does this for the simplest case which financial markets are frictionless, whereas in **Chapter 11**, we explore the effect of various limitations in dynamic trading strategies on optimal dynamic portfolio management.

The canonical decision problem in the theory of finance, i.e., the Arrow-Debreu portfolio problem, is developed in Part IV. In **Chapter 12**, we assume that the investor can exchange any contingent claim contract in completely competitive markets. We determine the impact of risk aversion on the investors optimal portfolio. Several properties of the value function for the investor's wealth are derived in **Chapter 13**.

We introduce consumption and savings in Part V. We start with the consump-

tion problem under certainty in **Chapter 14**. In this chapter, we also discuss how people evaluate their welfare when they consume their wealth over several different periods. We show that the structure of this problem is equivalent to the structure of the static Arrow-Debreu portfolio. We also explore the properties of the marginal propensity to consume and the concept of ”time diversification”. The assumption of a time-separable utility function makes the treatment of time very similar to the treatment of uncertainty in classical economics. In **Chapter 15,** the important concept of prudence is introduced. Saving is examined as a way to build a precautionary reserve to forearm oneself against future exogenous risks. In these two chapters, it is assumed that financial markets are perfect, i.e., that there is a single risk free interest rate at which consumers can borrow and lend. This is obviously not a realistic assumption. A market imperfection is introduced in **Chapter 16** by the way of a liquidity constraint. The existence of a liquidity constraint provides another incentive to save. Finally, we discuss the Merton-Samuelson problem of the joint decisions on consumption and risk-taking. This is done in **Chapter 17**.

In Part VI, we gather several of the previous results to determine the equilibrium price of risk and time in an Arrow-Debreu economy. More generally, it provides an analysis of how risks are traded in our economies. **Chapter 18** starts with the characterization of socially efficient risk sharing arrangements. **Chapter 19** shows how competition in financial markets can generate an efficient allocation of risks. We also determine the equilibrium price of risk and time as a function of the characteristics of the particular economy. We examine the standard asset pricing models like the CAPM. This chapter also provides an introduction to decision-making problems for corporate firms, and to the Modigliani-Miller Theorem. **Chapter 20** shows how the complete markets model has been calibrated to yield the equity premium puzzle and the risk free rate puzzle. **Chapter 21** is devoted to the analysis of the term structure of interest rates.

The last part of the book focuses on dynamic models of decisions-making under uncertainty when a flow of information on future risks is expected over time. The basic tools and concepts (value of information and Blackwell’s theorem) are provided in **Chapter 22**. **Chapter 23** addresses the important question of the optimal timing of an investment decision when there is a quasi-option value to waiting for future information about the investments profitability. The degree of irreversibility of the investment is taken into account. The last chapter, **Chapter 24**, offers a smorgasbord of analyses related to the effect of the expectation of future information on current decisions and an equilibrium prices of financial assets.

Audience

I direct my work to researchers dealing with economic uncertainty. The main fields of interest are in finance, macroeconomics and environmental economics. The only required mathematics background is standard calculus. No specific knowledge on the economics of uncertainty is necessary (I start from the beginning). This should be a good book for a graduate course on the economics of uncertainty.

Acknowlegment

This book would not have been possible without the exceptional environment in my life. I am particularly grateful for this to my wife, Dominique, and to my kids, Quentin, Simon, Vincent and Louise. I also want to thank my colleagues in GREMAQ and IDEI at Toulouse for the quality of the work atmosphere that I enjoyed during the last 6 years.

I am definitely indebted to Louis Eeckhoudt for all these years of joint work. Moreover, he took a great responsability in reading the first versions of the book in a very professional way. I would also like to thank Rose-Anne Dana, Jean-Charles Rochet and Nicolas Treich for helpful comments.

Part I

General theory

Chapter 1

The expected utility model

Before addressing any decision problem under uncertainty, it is necessary to build a preference functional that evaluates the level of satisfaction of the decision maker who bears risk. If such a functional is obtained, decisions problems can be solved by looking for the decision that maximizes the level of satisfaction. This first chapter provides a way to evaluate the level of satisfaction under uncertainty. The rest of this book deals with applications of this model to specific decision problems.

1.1 Simple and compound lotteries

The description of an uncertain environment contains two different types of elements. First, one must enumerate all possible outcomes. An outcome is a list of parameters that affect the well-being of the decision maker. It can contain a health status, some meteorological parameters, a level of pollution, and the quantity of different goods that are consumed. Most of the time, we will assume that the outcome can be measured in a one-dimensional unit, namely money. But, at this stage, there is no cost to allow for a multidimensional outcome. Let X be the set of possible outcomes. To avoid technicalities, we assume that the number of possible outcomes is finite, i.e. $X \equiv \{x_s\}_{s=1,...,S}$.

The second characteristic of an uncertain environment is the vector of probabilities of the possible outcomes. Let $p_s \geq 0$ be the probability of occurrence of x_s,with $\sum_{s=1}^{S} p_s = 1$. We hereafter assume that these probabilities are objectively known.

A lottery L is described by a vector $(x_1, p_1; x_2, p_2; ...; x_S, p_S)$. Since the set

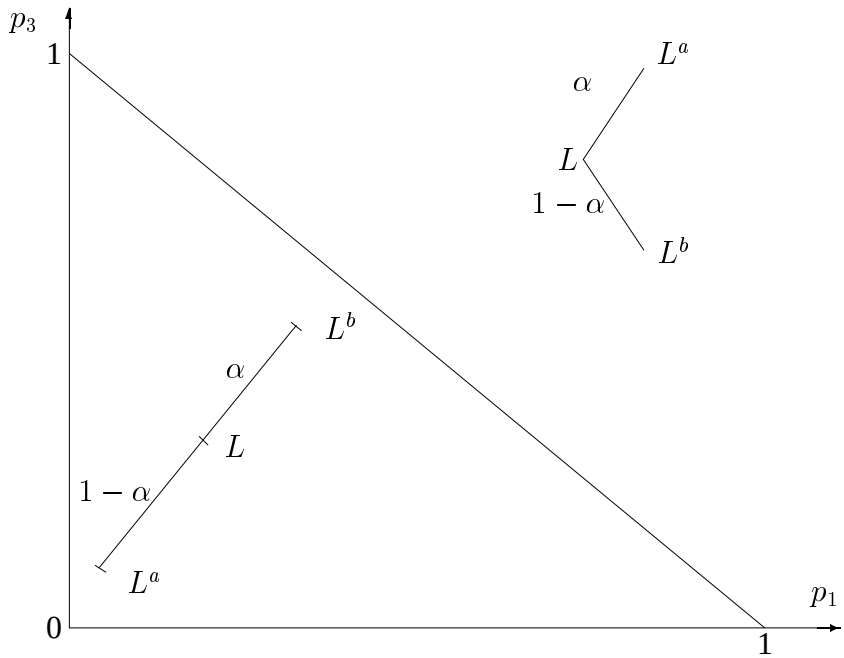

Figure 1.1: A compound lottery $L = \alpha L^a + (1-\alpha)L^b$ in the Machina triangle.

of potential outcomes is invariant, we will define a lottery L by its vector of probabilities $(p_1, ..., p_S)$. The set of all lotteries on outcomes X is denoted $\mathcal{L} \equiv \left\{(p_1, ..., p_S) \in R^S_+ \mid p_1 + ... + p_S = 1\right\}$.

When $S = 3$, we can represent a lottery by a point (p_1, p_3) in R^2, since $p_2 = 1 - p_1 - p_3$. More precisely, in order to be a lottery, this point must be in the so-called Machina triangle $\left\{(p_1, p_3) \in R^2_+ \mid 1 - p_1 - p_3 \leq 1\right\}$. If the lottery is on a edge of the triangle, one of the probabilities is zero. If it is at a corner, the lottery is degenerated, i.e., it takes one of the values x_1, x_2, x_3 with probability 1. This triangle is represented in Figure 1.1.

A compound lottery is a lottery whose outcomes are lotteries. Consider a compound lottery L which yields lottery $L^a = (p^a_1, ..., p^a_S)$ with probability α and lottery $L^b = (p^b_1, ..., p^b_S)$ with probability $1 - \alpha$. The probability that the outcome of L be x_1 is $p_1 = \alpha p^a_1 + (1-\alpha)p^b_1$. More generally, we obtain that

$$L \text{ has the same vector of probabilities as } \alpha L^a + (1-\alpha)L^b. \tag{1.1}$$

A compound lottery is a convex combination of simple lotteries. Such a compound lottery is represented in Figure 1.1. From condition (1.1), it is natural to confound L with $\alpha L^a + (1-\alpha)L^b$. Whether a specific uncertainty comes from a simple lottery or from a complex compound lottery has no significance. Only the probabilities of potential outcomes matter. This idea is called consequentialism, which states that L is equivalent to $\alpha L^a + (1-\alpha)L^b$.

1.2 Axioms on preferences under uncertainty

We assume that the decision maker has a rationale preference relation $\succeq$ over the set of lotteries $\mathcal{L}$. This means that order $\succeq$ is complete and transitive. For any pair (L^a, L^b) of lotteries in $\mathcal{L}$, either L^a is preferred to L^b ($L^a \succeq L^b$), or L^b is preferred to L^a ($L^b \succeq L^a$) (or both). Moreover, if L^a is preferred to L^b which is itself preferred to L^c, then L^a is preferred to L^c. To this preference order $\succeq$, we associate the indifference relation $\sim$, with $L^a \sim L^b$ if and only if $L^a \succeq L^b$ and $L^b \succeq L^a$.

A standard hypothesis that is made on preferences is that they are continuous. This means that small changes in probabilities do not change the nature of the ordering between two lotteries. Technically, this axiom is written as follows:

Axiom 1 (Continuity) *The preference relation $\succeq$ on the space of simple lotteries $\mathcal{L}$ is such that for all $L^a, L^b, L^c \in \mathcal{L}$ such that $L^a \succeq L^b \succeq L^c$, there exists a scalar $\alpha \in [0,1]$ such that*

$$L^b \sim \alpha L^a + (1-\alpha)L^c.$$

As is well-known in the theory of consumer choice under certainty, the continuity axiom implies the existence of a functional $U : \mathcal{L} \to R$ such that

$$U(L^a) \geq U(L^b) \;\Leftrightarrow\; L^a \succeq L^b. \tag{1.2}$$

The preference functional U is an index that represents the degree of satisfaction of the decision maker. It assigns a numerical value to each lottery, ranking them

in accordance to the individual's preference $\succeq$. Notice that U is not unique: take any increasing function $g : R \to R$. Then, the functional V that is defined as $V(.) = g(U(.))$ also represents the same preferences $\succeq$. The preference functional is ordinal in the sense that it is invariant to any increasing transformation. It is only the ranking of lotteries that matters. A preference functional can be represented in the Machina triangle by a family of continuous indifference curves.

The above assumptions on preferences under uncertainty are minimal. If no other assumption is made on these preferences, the theory of choice under uncertainty would not differ from the standard theory of consumer choice under certainty. The only difference would be on how to define the consumption goods. Most developments in the economics of uncertainty and its applications have been made possible by imposing more structure on preferences under uncertainty. This additional structure originates from the independence axiom.

Axiom 2 (Independence) *The preference relation $\succeq$ on the space of simple lotteries $\mathcal{L}$ is such that for all $L^a, L^b, L^c \in \mathcal{L}$ and for all $\alpha \in [0,1]$:*

$$L^a \succeq L^b \;\Leftrightarrow\; \alpha L^a + (1-\alpha)L^c \succeq \alpha L^b + (1-\alpha)L^c.$$

This means that, if we mix each of two lotteries L^a and L^b with a third one L^c, then the preference ordering of the two resulting mixtures is *independent* of the particular third lottery L^c used. The independence axiom is at the heart of the classical theory of uncertainty. Contrary to the other assumptions made above, the independence axiom has no parallel in the consumer theory under certainty. This is because there is no reason to believe that if I prefer a bundle A containing 1 cake and 1 bottle of wine to a bundle B containing 3 cakes and no wine, I also prefer a bundle A' containing 2 cakes and 2 bottles of wine to a bundle B' containing 3 cakes and 1.5 bottles of wine, just because

$$(2,2) = 0.5(1,1) + 0.5(3,3) \text{ and } (3,1.5) = 0.5(3,0) + 0.5(3,3).$$

1.3 The expected utility theorem

The independence axiom implies that the preference functional U must be linear in the probabilities of the possible outcomes. This is the essence of the expected utility theorem, which is due to von Neumann and Morgenstern (1944).

Theorem 1 (Expected Utility) *Suppose that the rationale preference relation $\succeq$ on the space of simple lotteries $\mathcal{L}$ satisfies the continuity and independence axioms. Then, $\succeq$ can be represented by a preference functional that is linear. That is, there exists a scalar u_s associated to each outcome x_s, $s = 1, ..., S$, in such a manner that for any two lotteries $L^a = (p_1^a, ..., p_S^a)$ and $L^b = (p_1^b, ..., p_S^b)$, we have*

$$L^a \succeq L^b \;\Leftrightarrow\; \sum_{s=1}^{S} p_s^a u_s \geq \sum_{s=1}^{S} p_s^b u_s.$$

Proof: We would be done if we prove that for any compound lottery $L = \beta L^a + (1-\beta)L^b$, we have

$$U(\beta L^a + (1-\beta)L^b) = \beta U(L^a) + (1-\beta)U(L^b). \tag{1.3}$$

Applying this property recursively would yield the result. To do this, let us consider the worst and best lotteries in $\mathcal{L}$, $\underline{L}$ and $\overline{L}$. They are obtained by solving the problem of minimizing and maximizing $U(L)$ on the compact set $\mathcal{L}$. By definition, for any $L \in \mathcal{L}$, we have $\overline{L} \succeq L \succeq \underline{L}$. By the continuity axiom, we know that there exist two scalars, α^a and α^b, in $[0, 1]$ such that

$$L^a \sim \alpha^a \overline{L} + (1-\alpha^a)\underline{L}$$

and

$$L^b \sim \alpha^b \overline{L} + (1-\alpha^b)\underline{L}.$$

Observe that $L^a \succeq L^b$ if and only if $\alpha^a \geq \alpha^b$. Indeed, suppose that $\alpha^a \geq \alpha^b$ and take $\gamma = \frac{\alpha^a - \alpha^b}{1-\alpha^b} \in [0, 1]$. Then, we have

$$\begin{aligned}
\alpha^b \overline{L} + (1-\alpha^b)\underline{L} &\sim \gamma \overline{L} + (1-\gamma)\left[\alpha^a \overline{L} + (1-\alpha^a)\underline{L}\right] \\
&\succeq \gamma\left[\alpha^a \overline{L} + (1-\alpha^a)\underline{L}\right] + (1-\gamma)\left[\alpha^a \overline{L} + (1-\alpha^a)\underline{L}\right] \\
&\sim \alpha^a \overline{L} + (1-\alpha^a)\underline{L}.
\end{aligned}$$

The two equivalences are direct applications of consequentialism. The second line of this sequence of relations is due to the independence axiom together with the fact that $\overline{L} \succeq \alpha^a \overline{L} + (1-\alpha^a)\underline{L}$, by definition of $\overline{L}$.

We conclude that $U(L) = \alpha$, where α is such that $L \sim \alpha\overline{L} + (1-\alpha)\underline{L}$, perfectly fits the definition (1.2) of a preference functional associated to $\succeq$. Thus, $U(L^a) = \alpha^a$ and $U(L^b) = \alpha^b$. It remains to prove that

$$U(\beta L^a + (1-\beta)L^b) = \beta\alpha^a + (1-\beta)\alpha^b,$$

or, equivalently, that

$$\beta L^a + (1-\beta)L^b \sim (\beta\alpha^a + (1-\beta)\alpha^b)\overline{L} + (\beta(1-\alpha^a) + (1-\beta)(1-\alpha^b))\underline{L}.$$

This is true since, using the independence axiom twice, we get

$$\begin{aligned}
\beta L^a + (1-\beta)L^b &\sim \beta\left[\alpha^a\overline{L} + (1-\alpha^a)\underline{L}\right] + (1-\beta)L^b \\
&\sim \beta\left[\alpha^a\overline{L} + (1-\alpha^a)\underline{L}\right] + (1-\beta)\left[\alpha^b\overline{L} + (1-\alpha^b)\underline{L}\right] \\
&\sim (\beta\alpha^a + (1-\beta)\alpha^b)\overline{L} + (\beta(1-\alpha^a) + (1-\beta)(1-\alpha^b))\underline{L}.
\end{aligned}$$

This concludes the proof.■

The consequence of the independence axiom is that the family of indifference curves in the Machina triangle is a set of parallel straight lines. Their slope equals $\frac{u_1-u_2}{u_{3-u_2}}$. This has the following consequences on the attitude towards risk. Consider four lotteries, L^a, L^b, M^a, M^b, that are depicted in the Machina triangle of Figure 1.2. Suppose that the segment L^aL^b is parallel to segment M^aM^b. It implies that L^a is preferred to L^b if and only if M^a is preferred to M^b.

In the remaining of this book, an outcome is a monetary wealth. The ex-post environment of the decision maker is fully described by the amount of money that he can consume. A lottery on wealth can be expressed by means of a random variable $\widetilde{w}$ whose realization is the outcome, i.e. an amount of money w that can be consumed. This random variable can be expressed by a cumulative distribution function F where $F(w)$ is the probability that $\widetilde{w}$ be less or equal than w. This covers the case of continuous, discrete or mixed random variables. By the expected utility theorem, we know that to each wealth level w, there exists a scalar $u(w)$ such that

$$\widetilde{w}_1 \succeq \widetilde{w}_2 \;\Leftrightarrow\; Eu(\widetilde{w}_1) \geq Eu(\widetilde{w}_2) \;\Leftrightarrow\; \int u(w)dF_1(w) \geq \int u(w)dF_2(w),$$

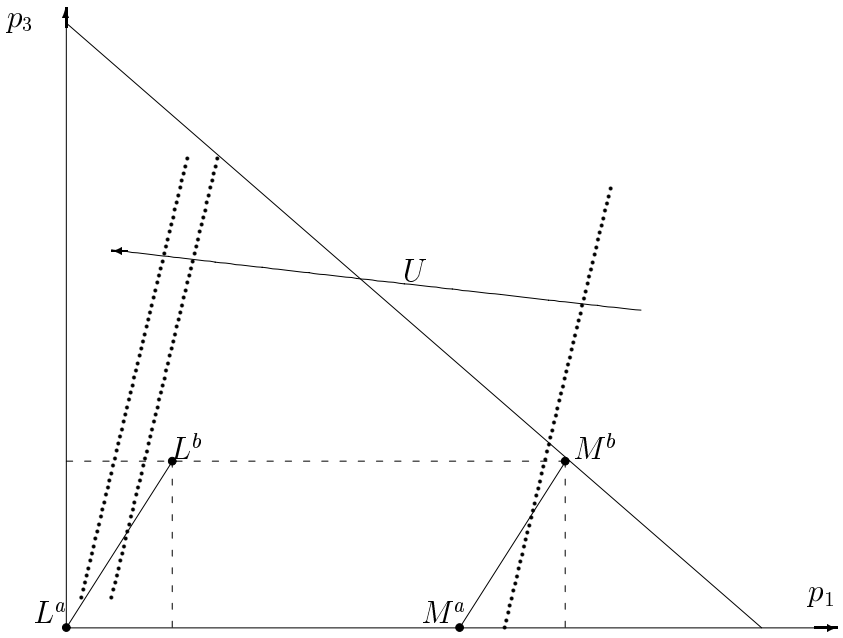

Figure 1.2: The Allais paradox in the Machina triangle.

where F_i is the cumulative distribution function of $\widetilde{w}_i$. It is intuitive in this context to assume that the utility function is nondecreasing.

Notice that the utility function is *cardinal*: an increasing *linear* transformation of u, $v(.) = au(.) + b$, $a > 0$, will not change the ranking of lotteries: if $\widetilde{w}_1 \succeq \widetilde{w}_2$, we have

$$Ev(\widetilde{w}_1) = E\left[au(\widetilde{w}_1) + b\right] = aEu(\widetilde{w}_1) + b \geq aEu(\widetilde{w}_2) + b = Ev(\widetilde{w}_2).$$

To sum up, expected utility is ordinal, whereas the utility function is cardinal. Differences in utility have meaning, whereas differences in expected utility have no significance.

1.4 Discussion of the expected utility model

The independence axiom is not without difficulties. The oldest and most famous challenge to it has been proposed by Allais (1953). Allais proposes the following experiment. An urn contains 100 balls that are numbered from 0 to 99. They are four lotteries whose monetary outcomes depend in different ways on the number of the ball that is taken out of the urn. The outcome, expressed say in thousands of dollars, are described in Table 1.1.

Lottery	0	1-10	11-99
L^a	50	50	50
L^b	0	250	50
M^a	50	50	0
M^b	0	250	0

Table 1.1: Outcome as a function of the number of the ball

Decision makers are subjected to two choice tests. In the first test, they are asked to choose between L^a and L^b, whereas in the second test, they must choose between M^a and M^b. Many people report that they prefer L^a to L^b, but they prefer M^b to M^a. Notice that since L^a and L^b have the same outcome when the number of the ball is larger than 10, the independence axiom tells us that these people prefer $L^{a'}$ that takes value 50 with certainty to $L^{b'}$ which takes value 0 with probability 1/11 and value 250 with probability 10/11. The (Allais) paradox is that the same argument can be used with the opposite result when considering

the preference of M^b over M^a! Thus, these people do not satisfy the independence axiom.

This problem can be analyzed with the Machina triangle in Figure 1.2. Let us define

$$\begin{aligned} x_1 &= 0 \\ x_2 &= 50 \\ x_3 &= 250. \end{aligned}$$

The four lotteries in the Allais experiment are respectively $L^a = (0, 1, 0)$, $L^b = (0.01, .89, .10)$, $M^a = (.89, .11, 0)$ and $M^b = (.90, 0, .10)$. Segment $L^a L^b$ is parallel to segment $M^a M^b$. Because indifference curves are parallel straight lines under expected utility, it must be that $M^a \succeq M^b$ if $L^a \succeq L^b$. Recognizing that the independence axiom makes sense, a lot of those who ranked $L^a \succeq L^b$ and $M^b \succeq M^a$ reverse their second ranking after having been explained the inconsistency of their choice with respect to the axiom.

An argument in favor of the independence axiom is that individuals who violate it would be subject to the acceptance of so-called "Dutch books". Suppose that a gambler is offered three lotteries L^a, L^b and L^c such that the gambler has ranking $L^a \succ L^b$ and $L^a \succ L^c$, but, contrary to the independence axiom, $L^d = 0.5L^b + 0.5L^c \succ L^a$. L^d is a compound lottery in which a fair coin is used to determine which lottery, L^b or L^c, will be played. In front of the choice between L^a, L^b and L^c, the gambler rationally selects L^a. Since L^d is preferred to L^a, we know that he is willing to pay some positive fee to replace L^a by the compound lottery L^d. The deal is thus rationally accepted by the gambler. But as soon as the coin is tossed, giving the gambler either L^b or L^c, you can get him to pay another fee to trade this lottery for L^a. Hence, at that point, the gambler has paid two positive fees but would otherwise be back to his original position! This is dynamically inconsistent.

1.5 Concluding remark

We will see in the next chapters that the linearity of the preference functional with respect to probabilities plays a central role to find simple solutions to several decision problems under uncertainty. If we relax the independence axiom, most problems presented in this book cannot be solved anymore. Our approach is thus pragmatic: we recognize that the independence axiom may fail to be verified in

some specific risky environments, in particular when there are low probability events. But the combination of the facts that the independence axiom makes much sense to us and that it makes most problems solvable is enough to justify the exploration of its implications. This is the aim of this book.

The search for an alternative model of decision under uncertainty that does not rely on the independence axiom has been a lively field of research in economics for more than 15 years. There are several competing models, each of them with its advantages and defaults. Many of them entail the expected utility model as a particular case. Nevertheless, the use of the expected utility model is pervasive in economics.

Chapter 2

Some basic tools

In spite of the diversity of applications to the economics of uncertainty, a handful technical tools are necessary to analyze them. This chapter is devoted to their presentation.

This chapter is the hard part of this book. But the reward for the reader who swallows this pill is high. Specialists in this field used to repetitively use similar techniques to solve different problems. By presenting and proving the results in a unified way, we will be able to save much time and effort in the remaining of this book. Moreover, we will be able to better understand the links between apparently very different results.

Path breaking economists who have made this unified approach possible are Susan Athey, Ian Jewitt, Miles Kimball and John Pratt.

2.1 The diffidence theorem

Since the well-being of an individual is measured by the expected utility of his final wealth, some problems in expected utility theory simplify to determining the impact of a change in a parameter in the model on the expectation of a function of a random variable. Let $\widetilde{x}$ be this random variable, and let f_1 and f_2 be this real-valued function respectively before and after the change in the parameter of the model. The question becomes whether we have

$$Ef_2(\widetilde{x}) \leq Ef_1(\widetilde{x}) \ \ \forall \widetilde{x} \tag{2.1}$$

In general, no restriction is imposed on the distribution of the random variable, except that its support is bounded in $[a, b]$. Obviously, a necessary condition is that

$$f_2(x) \leq f_1(x) \qquad \forall x \in [a, b]\,. \tag{2.2}$$

Indeed, condition (2.2) is nothing else than condition(2.1) for the degenerated random variable which takes value x with probability 1. Condition (2.2) is also sufficient for condition (2.1), since it implies that

$$\int_a^b f_2(x)dG(x) \leq \int_a^b f_1(x)dG(x) \tag{2.3}$$

for any cumulative distribution function G.

To conclude, in order to verify that condition (2.1) holds for any random variable, it is enough to verify that it holds for any *degenerated* random variable. This greatly reduces the dimensionality of the problem. Most of the time, the technical problem is a little more complex than the one just presented. More specifically, it happens that one has already some information about the random variable under consideration. For example, one knows that

$$Ef_1(\widetilde{x}) = 0,$$

and one inquires about whether condition (2.1) holds for this subset of random variables. Obviously, since one wants a property to hold for a smaller set of objects, the necessary and sufficient condition will be weaker than condition (2.2). We now determine this weaker condition.

A simpler way to write the problem is as follows:

$$\forall \widetilde{x}: \qquad Ef_1(\widetilde{x}) = 0 \qquad \Longrightarrow \qquad Ef_2(\widetilde{x}) \leq 0. \tag{2.4}$$

If one does not know the results presented below, the best way to approach this problem would probably be to try to find a counter-example. A counter-example would be found by characterizing a random variable $\widetilde{x}$ such that $Ef_1(\widetilde{x}) = 0$, with $Ef_2(\widetilde{x})$ positive. Our best chance is obtained by finding the cumulative distribution function F of $\widetilde{x}$ that satisfies the constraint and which maximizes the expectation of f_2. Thus, we should first solve the following problem:

$$\begin{array}{ll} \max_{dG\geq 0} & \int_a^b f_2(x)dG(x) \\ s.t. & \int_a^b f_1(x)dG(x) = 0 \\ & \int_a^b dG(x) = 1. \end{array} \tag{2.5}$$

The second constraint simply states that G is indeed a cumulative distribution function. Now, observe that the objective and the two (equality) constraints are linear in dG,which is required to be nonnegative. Thus, this is a standard linear programming problem. Linearity entails corner solutions. More specifically, the basic lesson of linear programming is that the number of choice variables that are positive at the optimal solution does not exceed the number of constraints. Thus, we obtain the following Lemma.

Lemma 1 *Condition (2.4) is satisfied if and only if it is satisfied for any binary random variable, i.e. for any random variable whose support contains only two points.*

For readers who are not trained in operations research, we hereafter provide an informal proof of this result. Suppose by contradiction that the solution of program (2.5) has three atoms, namely at x_1, x_2 and x_3, respectively with probability $p_1 > 0,\ p_2 > 0$ and $p_3 > 0$. We can represent this random variable by a point A in the Machina triangle where we use the last constraint in (2.5) to replace p_2 by $1 - p_1 - p_3$. This is drawn in Figure 2.1. Because this random variable satisfies condition $Ef_1(\widetilde{x}) = 0$, we have that point A is an interior point on segment CD that is defined by

$$(f_1(x_1) - f_1(x_2))p_1 + (f_1(x_3) - f_1(x_2))p_3 + f_1(x_2) = 0. \tag{2.6}$$

The problem is to find (p_1, p_3) that maximizes the following function

$$(f_2(x_1) - f_2(x_2))p_1 + (f_2(x_3) - f_2(x_2))p_3 \tag{2.7}$$

As usual, this problem is graphically represented by iso-objective curves in Figure 2.1. The important point is that these curves are straight lines. As we know, this is specific to expected utility, where the well-being of an agent is linear in probabilities. It clearly appears that an interior point on segment CD is never optimal.

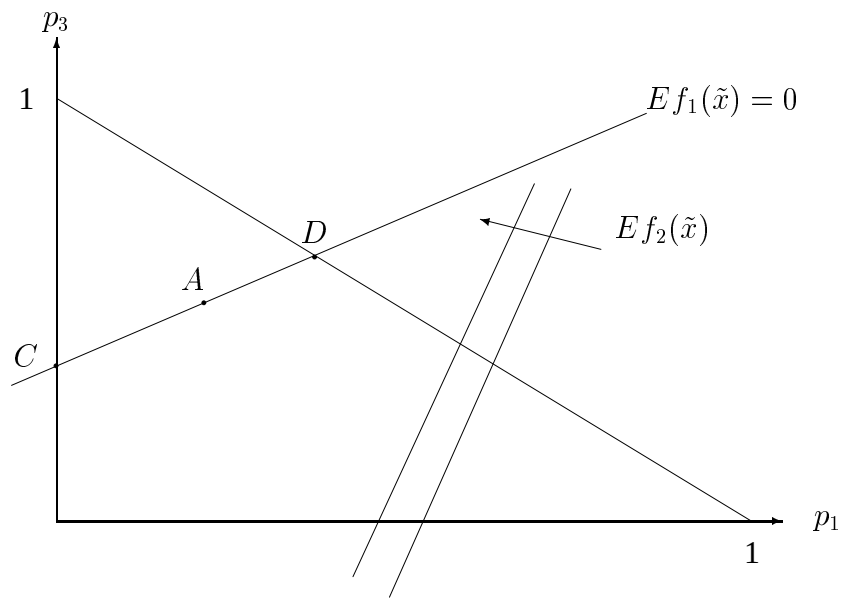

Figure 2.1: Interior points on segment CD are not optimal.

At the limit, the iso-objective lines are parallel to segment CD, and it is enough to check the condition at C or D, as stated in the Lemma.

Thus, to check that condition (2.4) holds for any random variable, it is enough to check that it holds for any binary random variable $\widetilde{x} \sim (x_1, p; x_2, 1-p)$. This problem is rewritten:

$$pf_1(x_1) + (1-p)f_1(x_2) = 0 \qquad \Longrightarrow \qquad pf_2(x_1) + (1-p)f_2(x_2) \leq 0 \quad (2.8)$$

for all $p \in [0,1]$, and $(x_1, x_2) \in [a,b] \times [a,b]$. In order for the left condition to be satisfied, we need $f_1(x_1)$ and $f_1(x_2)$ to alternate in sign. Without loss of generality, let $f_1(x_1) < 0 < f_1(x_2)$. By eliminating p from this problem, it can be rewritten as

$$\frac{f_2(x_1)}{f_1(x_1)} \geq \frac{f_2(x_2)}{f_1(x_2)}.$$

Since this must be true for all x_1 such that $f_1(x_1) < 0$, the above inequality is equivalent to

$$\min_{x|f_1(x)\leq 0} \frac{f_2(x)}{f_1(x)} \geq \frac{f_2(x_2)}{f_1(x_2)}.$$

This must hold for any x_2 such that $f_1(x_2) > 0$. In a similar way, the condition is rewritten as:

$$\min_{x|f_1(x)\leq 0} \frac{f_2(x)}{f_1(x)} \geq \max_{x|f_1(x)\geq 0} \frac{f_2(x)}{f_1(x)}$$

This can be true only if there exists a scalar m, positive or negative, such that

$$\min_{x|f_1(x)\leq 0} \frac{f_2(x)}{f_1(x)} \geq m \geq \max_{x|f_1(x)\geq 0} \frac{f_2(x)}{f_1(x)}$$

The two above inequalities hold if and only if

$$\exists m : \qquad f_2(x) \leq m f_1(x) \qquad \forall x \in [a,b]. \qquad (2.9)$$

In short, condition (2.4) holds for any binary random variable if and only if condition (2.9) holds. Combining this result with the Lemma yields the following central result, which is due to Gollier and Kimball (1997). We hereafter refer to this Proposition as the Diffidence Theorem. The origin of this term will be explained in the next chapter.[1]

Proposition 1 *The following two conditions are equivalent:*

1. $\forall \widetilde{x}$ *with support in* $[a, b]$: $\quad Ef_1(\widetilde{x}) = 0 \Longrightarrow Ef_2(\widetilde{x}) \leq 0.$
2. $\exists m$: $\quad f_2(x) \leq mf_1(x) \quad \forall x \in [a, b]$.

Observe that, as expected, the necessary and sufficient condition for (2.4) is weaker that the necessary and sufficient condition for (2.1). Indeed, m is not forced to be equal to 1 in condition (2.9), contrary to what happens in condition (2.2).

There still remains a difficulty to verify whether condition (2.4) is true: one has to look for a scalar m such that

$$H(x, m) = f_2(x) - mf_1(x)$$

be nonpositive for all x. This search may be difficult. However, the problem is much simpler if f_1 and f_2 satisfy the following conditions:

a. there is a scalar x_0 such that $f_1(x_0) = f_2(x_0) = 0$;

b. f_1 and f_2 are twice differentiable at x_0;

c. $f_1'(x_0) \neq 0$.

¿From the first condition, we have that $H(x_0, m) = 0$. Since, from the second condition, H is differentiable at $x = x_0$, it must be true that

$$\frac{\partial H}{\partial x}(x_0, m) = 0, \text{ and } \frac{\partial^2 H}{\partial x^2}(x_0, m) \leq 0$$

[1]There are various ways to prove the Diffidence Theorem. It is for example a standard application of the Separating Hyperplane Theorem. In a similar spirit, we can see m as the Lagrangian multiplier associated to the first constraint in program (2.5). We preferred proving the result by using Lemma 1 for pedagogical reasons.

in order to have H nonpositive in the neighborhood of x_0. The first necessary condition yields

$$m = \frac{f_2'(x_0)}{f_1'(x_0)},$$

whereas the second necessary condition is rewritten as

$$f_2''(x_0) \leq \frac{f_2'(x_0)}{f_1'(x_0)} f_1''(x_0). \tag{2.10}$$

We obtain the following Corollary, which is due to Gollier and Kimball (1997).

Corollary 1 *Suppose that conditions (a)-(b)-(c) are satisfied. Then, a necessary and sufficient condition for (2.4) is*

$$\forall x \in [a, b] : \quad f_2(x) \leq \frac{f_2'(x_0)}{f_1'(x_0)} f_1(x). \tag{2.11}$$

A necessary condition for (2.4) is condition (2.10).

Again, the dimensionality of the problem is much reduced by this result. One had initially to check whether any random variable that verifies condition $Ef_1(\widetilde{x}) = 0$ also verifies condition $Ef_2(\widetilde{x}) \leq 0$. We end up with checking whether a function of a single variable, $H(x, \frac{f_2'(x_0)}{f_1'(x_0)})$, is nonpositive.

This problem is even further simplified by necessary condition (2.10) that is usually referred to as the *local* diffidence condition. This terminology comes from the fact that (2.10) is the necessary and sufficient condition for any *small* risk around x_0 to verify property (2.4). Indeed, suppose that $\widetilde{x} = x_0 + k\widetilde{\epsilon}$. If k is small, condition $Ef_1(\widetilde{x}) = 0$ can be rewritten as

$$f_1(x_0) + kf_1'(x_0)E\widetilde{\epsilon} + k^2 f_1''(x_0)\frac{E\widetilde{\epsilon}^2}{2} \simeq 0.$$

Using the assumption that $f_1(x_0) = 0$ implies that

$$E\widetilde{\epsilon} \simeq k\frac{E\widetilde{\epsilon}^2}{2}\frac{f_1''(x_0)}{f_1'(x_0)}. \tag{2.12}$$

In the same vein, condition $Ef_2(\widetilde{x}) \leq 0$ is equivalent to

$$f_2(x_0) + kf_2'(x_0)E\widetilde{\epsilon} + k^2 f_2''(x_0)\frac{E\widetilde{\epsilon}^2}{2} \leq 0.$$

Replacing $E\widetilde{\epsilon}$ by its expression from (2.12) makes the above inequality equivalent to condition (2.10). Thus, this condition is necessary and sufficient for condition (2.4) to hold for any small $\widetilde{x}$. Inequality (2.10) is only necessary for it to be true for any risk.

In some applications, the equality condition $Ef_1(\widetilde{x}) = 0$ in problem (2.4) is replaced by an inequality:

$$\forall \widetilde{x}: \qquad Ef_1(\widetilde{x}) \leq 0 \qquad \Longrightarrow \qquad Ef_2(\widetilde{x}) \leq 0. \tag{2.13}$$

The necessary and sufficient condition for this problem is that there exists a *non-negative* scalar m such that $f_2(x) \leq mf_1(x)$ for all x. The sufficiency of this condition is immediate: if this condition holds for any x, taking the expectation yields

$$Ef_2(\widetilde{x}) \leq mEf_1(\widetilde{x}).$$

The right-hand side of the above inequality is the product of a nonnegative scalar and of a nonpositive one. It implies that $Ef_2(\widetilde{x})$ is nonpositive. Lemma (1) proves the necessity of the existence of a m . The fact that it must be nonnegative is most easily shown when conditions (a)-(b)-(c) are satisfied, where $m = \frac{f_2'(x_0)}{f_1'(x_0)}$. By contradiction, suppose that $\frac{f_2'(x_0)}{f_1'(x_0)}$ is negative. It means that $f_1'(x_0)$ and $f_2'(x_0)$ have opposite signs. Suppose that $f_1'(x_0)$ is positive (resp. negative) and $f_2'(x_0)$ is negative (resp. positive). It implies that condition (2.13) is violated for some $\widetilde{x}$ that is degenerated at a value slightly to the left (resp. right) of x_0.

Corollary 2 *Suppose that conditions (a)-(b)-(c) are satisfied. Then, a necessary and sufficient condition for (2.13) is*

$$\forall x \in [a,b]: \quad f_2(x) \leq \frac{f_2'(x_0)}{f_1'(x_0)} f_1(x) \quad \textit{and} \quad \frac{f_2'(x_0)}{f_1'(x_0)} \geq 0.$$

A necessary condition for (2.13) is condition (2.10).

2.2 Applications of the Diffidence Theorem

2.2.1 Jensen's inequality

Another kind of problem is to determine the effect of risk on the expectation of a function of a random variable. More specifically, we want to determine the condition under which the expectation of this function is smaller than the function evaluated at the expected value of the random variable:

$$Ef(\widetilde{x}) \leq f(E\widetilde{x}). \tag{2.14}$$

In fact, this problem can be solved by using the diffidence theorem. Let us rewrite it as

$$E\widetilde{x} = x_0 \quad \Longrightarrow \quad Ef(\widetilde{x}) \leq f(x_0)$$

for any scalar x_0. Using Proposition 1 with $f_1(x) = x - x_0$ and $f_2(x) = f(x) - f(x_0)$ yields the following condition:

$$f_2(x) \leq f_2(x_0) + m(x - x_0).$$

The left-hand side of the above inequality is the equation of a straight line with slope m that intersects the curve associated to function f at point x_0. The inequality means that there must exists such a straight line that is always above curve f as is the case in Figure 2.2. Since this must be true for any point x_0 on curve f, it is clear that condition (2.14) holds if and only if function f is concave. This is Jensen's inequality, which is a direct application of the Diffidence Theorem.

Proposition 2 (Jensen's Inequality) *The two following conditions are equivalent:*

1. $Ef(\widetilde{x}) \leq (\geq) f(E\widetilde{x})$ *for all* $\widetilde{x}$;
2. f *is concave (resp. convex).*

If f is neither concave nor convex, the result is ambiguous in the sense that it is always possible to find a pair $(\widetilde{x}_1, \widetilde{x}_2)$ such that inequality (2.14) holds with $\widetilde{x} = \widetilde{x}_1$ and the opposite inequality holds with $\widetilde{x} = \widetilde{x}_2$.

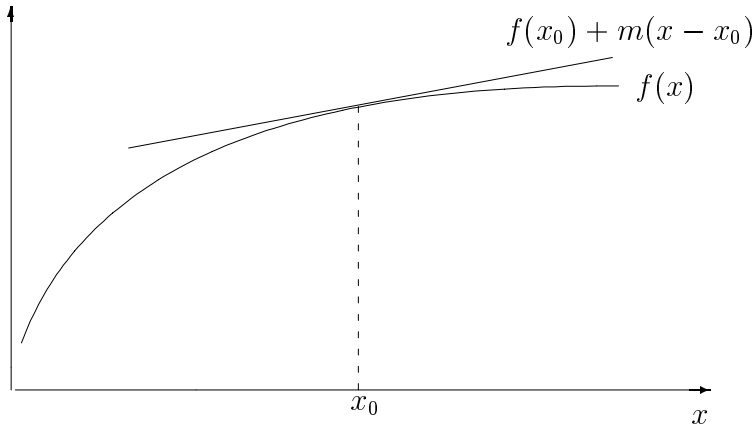

Figure 2.2: A concave function is below any of its tangent lines.

2.2.2 The covariance rule

One of the most used formula in the economics of uncertainty is

$$cov(\widetilde{z}, \widetilde{y}) = E\widetilde{z}\widetilde{y} - E\widetilde{z}E\widetilde{y}.$$

Thus, a sufficient condition for

$$E\widetilde{z}\widetilde{y} \leq E\widetilde{z}E\widetilde{y}$$

is that $\widetilde{z}$ and $\widetilde{y}$ covary negatively. In the particular case where z is a deterministic transformation of y, this means that z increases when y decreases. Let us define $\widetilde{z} = f(\widetilde{x})$ and $\widetilde{y} = g(\widetilde{x})$. We have that

$$Ef(\widetilde{x})g(\widetilde{x}) \leq Ef(\widetilde{x})Eg(\widetilde{x}) \tag{2.15}$$

for all $\widetilde{x}$ if f and g are monotonic and have opposite slopes: $f'(x)g'(x) \leq 0$ for all x. It means that z and y go in opposite direction when x is increased. We now show that this condition is necessary and sufficient if no restriction is put on $\widetilde{x}$. This can be seen by rewriting condition (2.15) as follows:

$$Ef(\widetilde{x}) = f(x_0) \quad \Longrightarrow \quad Ef(\widetilde{x})g(\widetilde{x}) \leq f(x_0)Eg(\widetilde{x}) \tag{2.16}$$

for any random variable $\widetilde{x}$ and any scalar x_0. A direct application of the Diffidence Theorem is that this property holds if and only if there exists a scalar $m(x_0)$ such that

$$g(x)\left(f(x) - f(x_0)\right) \leq m(x_0)\left(f(x) - f(x_0)\right)$$

for all x and x_0. Without loss of generality, let us assume that f is nondecreasing. Then, the above inequality is equivalent to

$$g(x) \begin{cases} \geq m(x_0) & \text{if } x < x_0 \\ \leq m(x_0) & \text{otherwise.} \end{cases} \tag{2.17}$$

The above condition is called a single-crossing condition: g can cross the horizontal axis at level $m(x_0)$ only once, from above.[2] Since this must be true for any x_0, this is possible only if g is decreasing.

Suppose alternatively that condition (2.15) must not hold for any random variable, but only for those that satisfy property $Ef(\widetilde{x}) = f(x_0)$ for a *specific* scalar x_0. Then, single-crossing condition (2.17) for *that* x_0 is the necessary and sufficient condition for (2.15) to hold for this set of random variables. We depict in Figure 2.3 a pair (f, g) satisfying the property.

We summarize these results in the following Proposition.

Proposition 3 *Suppose that $f : R \to R$ is nondecreasing. Condition*

$$Ef(\widetilde{x})g(\widetilde{x}) \leq Ef(\widetilde{x})Eg(\widetilde{x})$$

1. *holds for any $\widetilde{x}$ if and only if g is nonincreasing.*

2. *holds for any $\widetilde{x}$ such that $Ef(\widetilde{x}) = f(x_0)$ if and only if there exists a scalar $m(x_0)$ such that the single crossing condition (2.17) holds.*

[2] If g is continuous, the only candidate for $m(x_0)$ is $g(x_0)$.

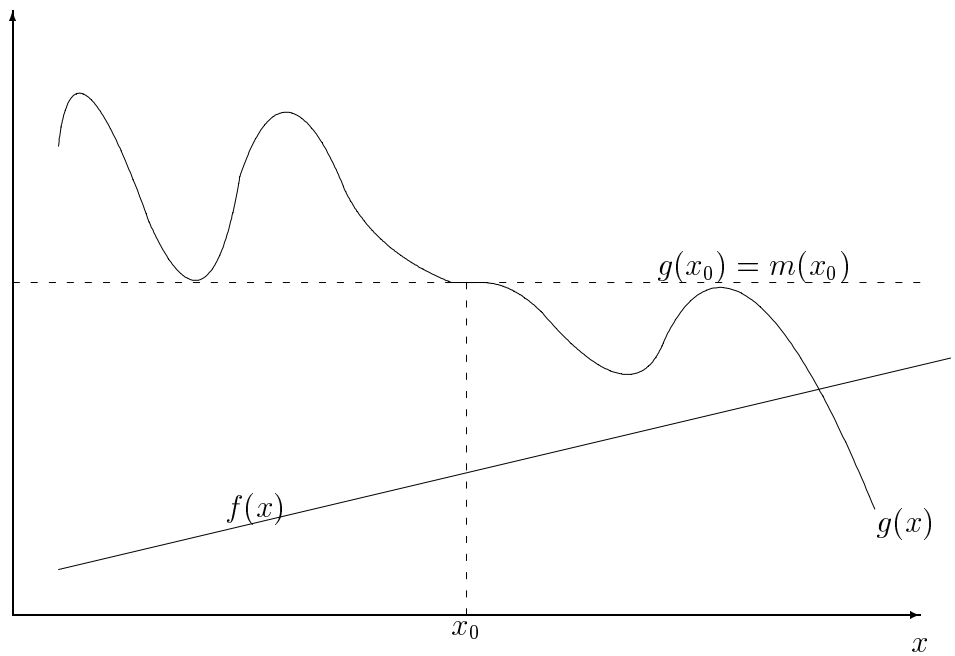

Figure 2.3: (f, g) satisfies condition $Ef(\tilde{x})g(\tilde{x}) \leq Ef(\tilde{x})Eg(\tilde{x})$ for all $\tilde{x}$ such that $Ef(\tilde{x}) = f(x_0)$.

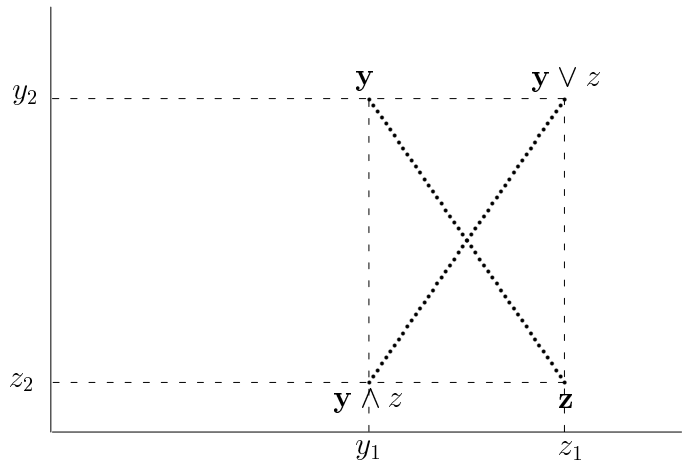

Figure 2.4: Representation of the ”meet” and ”join” operators.

2.2.3 Log-supermodularity and single-crossing

In order to introduce the useful concept of log-supermodularity, let us consider two vectors in R^n, $\mathbf{y}$ and $\mathbf{z}$. The operation ”meet” $(\vee)$ and ”join” $(\wedge)$ are defined as follows:

$$\mathbf{y} \vee \mathbf{z} = \inf \{\mathbf{t} \in R^n \mid \mathbf{t} \geq \mathbf{y}; \mathbf{t} \geq \mathbf{z}\}$$

$$\mathbf{y} \wedge \mathbf{z} = \sup \{\mathbf{t} \in R^n \mid \mathbf{t} \leq \mathbf{y}; \mathbf{t} \leq \mathbf{z}\} .$$

These two operators from R^n to R^n are represented in Figure 2.4 for the case $n = 2$.

Consider now a real-valued function h: $R^n \longrightarrow R$, $n \geq 2$, that is nonnegative.

Definition 1 *Function h is log-supermodular (LSPM) if for all $\mathbf{z}, \mathbf{y} \in R^n$,*

$$h(\mathbf{y} \vee \mathbf{z})h(\mathbf{y} \wedge \mathbf{z}) \geq h(\mathbf{y})h(\mathbf{z}).$$

Using Figure 2.4 for the case $n = 2$, it means that, whatever the rectangle under consideration in the plan, the product of h evaluated at the southwest-northeast corners of the rectangle is larger than the product of h at the northwest-southeast corners. Observe that it is a direct consequence of the above definition that the product of two log-supermodular functions is log-supermodular. An important example of LSPM functions in R^2 is the indicator function $h(x, \theta) = I(x \leq \theta)$ which takes value 1 if $x \leq \theta$ is true and 0 otherwise.

In the two-dimension case ($n = 2$), we can rewrite the definition of LSPM as:

$$x_0 < x \text{ and } \theta_L < \theta_H \implies h(x_0, \theta_L)h(x, \theta_H) \geq h(x_0, \theta_H)h(x, \theta_L).$$

If we now assume in addition that h is strictly positive, h is LSPM if and only if

$$\forall x, x_0 \in R,\ \forall \theta_H > \theta_L :\ (x - x_0)\frac{h(x, \theta_H)}{h(x_0, \theta_H)} \geq (x - x_0)\frac{h(x, \theta_L)}{h(x_0, \theta_L)},$$

or, equivalently,

$$(x - x_0)\left[\log(h(x, \theta_H)) - \log(h(x_0, \theta_H))\right] \geq (x - x_0)\left[\log(h(x, \theta_L)) - \log(h(x_0, \theta_L))\right].$$

If h is differentiable with respect to its first argument, this is equivalent to the condition that

$$(x - x_0)\int_{x_0}^{x} \frac{\partial}{\partial s}\log(h(s, \theta_H))ds \geq (x - x_0)\int_{x_0}^{x} \frac{\partial}{\partial s}\log(h(s, \theta_L))ds.$$

Since this must be true for any x and x_0 and for any $\theta_H > \theta_L$, this inequality holds if and only if

$$\frac{\partial}{\partial s}\log(h(s, \theta)) = \frac{\frac{\partial h}{\partial s}(s, \theta)}{h(s, \theta)}$$

is nondecreasing in θ. This proves the following Lemma, which is a particular case of a result obtained by Topkis (1978).

Lemma 2 *Suppose that $h : R^2 \to R^+$ is differentiable with respect to its first argument. Then, h is LSPM if and only if one of the following two equivalent conditions holds:*

1. $\forall x, x_0 \in R,\ \forall \theta_H > \theta_L :\ (x - x_0)\frac{h(x,\theta_H)}{h(x_0,\theta_H)} \geq (x - x_0)\frac{h(x,\theta_L)}{h(x_0,\theta_L)}$.

2. $\frac{\frac{\partial h}{\partial x}(x,\theta)}{h(x,\theta)}$ *is nondecreasing in* θ.

We see that, LSPM is basically that the cross-derivative of the logarithm of the function is nonnegative. As every economists knows, the sign of cross-derivatives are essential for comparative static analyses. Without surprise thus, log-supermodularity will play an important role for solving comparative statics problems of decision under uncertainty. This is also due to two properties of LSPM functions.

Consider a function g that single-crosses from below the horizontal axis at point x_0, i.e., $(x - x_0)g(x) \geq 0$ for all x. Consider also a positive function $h(x, \theta)$. We want to determine the conditions under which

$$\forall \widetilde{x} \in [a, b] :\ Eg(\widetilde{x})h(\widetilde{x}, \theta_L) = 0 \implies \forall \theta_H > \theta_L\ :\ Eg(\widetilde{x})h(\widetilde{x}, \theta_H) \geq 0. \quad (2.18)$$

In words, we want to know whether function $G(\theta) = Eg(\widetilde{x})h(\widetilde{x}, \theta)$ also single-crosses the horizontal axis from below. To solve this problem, one can use the Diffidence Theorem with $f_1(x) = -g(x)h(x, \theta_L)$ and $f_2(x) = -g(x)h(x, \theta_H)$. Using Corollary 1, it yields

$$\forall x :\ g(x)h(x, \theta_H) \geq \frac{h(x_0, \theta_H)}{h(x_0, \theta_L)} g(x)h(x, \theta_L).$$

Since, by assumption, $g(x)$ and $(x - x_0)$ have the same sign, the above inequality holds if and only if

$$\forall x :\ (x - x_0)\frac{h(x, \theta_H)}{h(x_0, \theta_H)} \geq (x - x_0)\frac{h(x, \theta_L)}{h(x_0, \theta_L)}.$$

If x_0 is arbitrary, this condition is equivalent to the LSPM of h. LSPM is necessary in the sense that if h is not LSPM, it is possible to find a single-crossing g and a random variable $\widetilde{x}$ such that condition (2.18) fails. The way the Diffidence Theorem has been proved provides the technique to build such a counter-example. In a similar fashion, if g does not single-crosses, there must exist a LSPM function h that violates condition (2.18).

This is the essence of a result obtained by Karlin (1968). It has been used in economics initially by Jewitt (1987) and Athey (1997). We summarize this result as follows:

Proposition 4 *Take any real-valued function g:$R \to R$ that satisfies the single-crossing condition: $\exists x_0 : \forall x : (x - x_0)g(x) \geq 0$. Then, condition (2.18) holds if and only if h is log-supermodular. Moreover, if g does not satisfy the single crossing condition, there exists a LSPM function h that violates condition (2.18).*

2.2.4 Expectation of a log-supermodular function

Consider now a function $h : R^n \to R^+$ with $n \geq 3$. Consider also a vector $\mathbf{z}$ in R^{n-1}. Suppose finally that h is log-supermodular. We are interested in determining whether

$$H(\mathbf{z}) = Eh(\mathbf{z}, \widetilde{\theta})$$

is also log-supermodular. In words, does the expectation operator preserve log-supermodularity? This question has been examined by Karlin and Rinott (1980), Jewitt (1987) and Athey (1997). They obtained the following positive result, the proof of which relies directly on the Diffidence Theorem.

Proposition 5 $H(\mathbf{z}) = Eh(\mathbf{z}, \widetilde{\theta})$ *is log-supermodular if h is log-supermodular.*

Proof: See the Appendix.

Notice that the log-supermodularity of h is sufficient, but not necessary, for the log-supermodularity of H.

2.3 Concluding remark

All the results presented in this chapter will be extensively used in the remaining of this book. It appears that they all rely on the Diffidence Theorem. We showed that this Theorem heavily depends on Lemma 1 where the linearity of expected utility with respect to probabilities plays a central role. It implies that we can limit in many instances the analysis of risky situations by just looking at binary risks. This in turn generates a lot of simplifications whose the different Propositions of this chapter are the expression. A side product of this presentation is that none of the tools presented here can be used for examining the implications of non-expected utility models. This is probably why the NEU literature is currently poor in terms of comparative statics analysis of specific problems of decision under uncertainty.

Appendix

Proof of Proposition 5

For the sake of simplicity, we prove this result in the particular case of $n = 3$. Notice that all applications of this Proposition in this book are for case $n = 3$.

Take an arbitrary vector (x_L, x_H, y_L, y_H). We have to prove that

$$x_L < x_H \text{ and } y_L < y_H \implies H(x_L, y_L)H(x_H, y_H) \geq H(x_L, y_H)H(x_H, y_L).$$

The Proposition is a direct application of the following Lemma:

Lemma 3 *Suppose that $h(.,.,\theta)$, $h(x_L,.,.)$ and $h(.,y_H,.)$ are all LSPM in R^2. Then, the following inequality holds for all $\widetilde{\theta}$ and all (x_H, y_L) such that $x_L < x_H$ and $y_L < y_H$:*

$$Eh(x_L, y_L, \widetilde{\theta})Eh(x_H, y_H, \widetilde{\theta}) \geq Eh(x_L, y_H, \widetilde{\theta})Eh(x_H, y_L, \widetilde{\theta}).$$

Proof: Using the Mean Value Theorem, we rewrite this condition as

$$\frac{Eh(x_L, y_L, \widetilde{\theta})}{Eh(x_L, y_H, \widetilde{\theta})} = \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)} \implies \frac{Eh(x_H, y_L, \widetilde{\theta})}{Eh(x_H, y_H, \widetilde{\theta})} \leq \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}.$$

Using the Diffidence Theorem, a necessary and sufficient condition for this to be true is that there exists a function $m(\theta_0)$ such that

$$h(x_H, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)} h(x_H, y_H, \theta)$$

$$\leq m(\theta_0; x_L, x_H, y_L, y_H) \left[h(x_L, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)} h(x_L, y_H, \theta) \right] \tag{2.19}$$

for all θ. This is what we have to prove. Without loss of generality, let us assume that $\theta > \theta_0$. Because we know that $h(.,.,\theta)$ is LSPM, we have that

$$h(x_H, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)} h(x_H, y_H, \theta)$$

$$
\begin{aligned}
&= h(x_H, y_H, \theta)\left[\frac{h(x_H, y_L, \theta)}{h(x_H, y_H, \theta)} - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}\right] \\
&\leq h(x_H, y_H, \theta)\left[\frac{h(x_L, y_L, \theta)}{h(x_L, y_H, \theta)} - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}\right] \\
&\leq \frac{h(x_H, y_H, \theta)}{h(x_L, y_H, \theta)}\left[h(x_L, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}h(x_L, y_H, \theta)\right]
\end{aligned}
$$

But, because $h(x_L, ., .)$ is LSPM, we know that

$$
\left[h(x_L, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}h(x_L, y_H, \theta)\right] \leq 0
$$

and, because $h(., y_H, .)$ is also LSPM,

$$
\frac{h(x_H, y_H, \theta)}{h(x_L, y_H, \theta)} \geq \frac{h(x_H, y_H, \theta_0)}{h(x_L, y_H, \theta_0)}.
$$

Combining the last three inequalities yields

$$
\begin{aligned}
&h(x_H, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}h(x_H, y_H, \theta) \\
&\leq \frac{h(x_H, y_H, \theta_0)}{h(x_L, y_H, \theta_0)}\left[h(x_L, y_L, \theta) - \frac{h(x_L, y_L, \theta_0)}{h(x_L, y_H, \theta_0)}h(x_L, y_H, \theta)\right].
\end{aligned}
$$

Thus, condition (2.19) holds with

$$
m(\theta_0; x_L, x_H, y_L, y_H) = \frac{h(x_H, y_H, \theta_0)}{h(x_L, y_H, \theta_0)}.
$$

This concludes the proof. ■

Chapter 3

Risk aversion

In Chapter 1, we have seen that the behavior of an expected-utility maximizer satisfies a set of interesting properties. Still, those properties are not sufficient to determine the effect of a change in his risky environment on his welfare or on his decisions. This chapter is devoted to the analysis of the effect of risk on the level of satisfaction of the agent. The basic reference for this chapter is the very dense paper by John Pratt (1964).

3.1 Diffidence and risk aversion

The impact of risk on welfare depends upon the combination of the characteristics of the risk, the wealth level of the agent and his utility function. We focus here on the effect of a *pure* risk, i.e. a *zero-mean* risk. It is widely believed that agents dislike pure risks, i.e. that agents are averse to risk. We now introduce two new concepts.

Definition 2 *An agent is diffident at wealth w_0 if he dislikes any zero-mean risk, given his initial wealth level w_0. An agent is risk-averse if he is diffident at all wealth level.*

Technically, an agent with sure wealth w_0 and utility function u dislikes any zero-mean risk if the following condition is satisfied:

$$E\widetilde{x} = 0 \qquad \Longrightarrow \qquad Eu(w_0 + \widetilde{x}) \leq u(w_0). \tag{3.1}$$

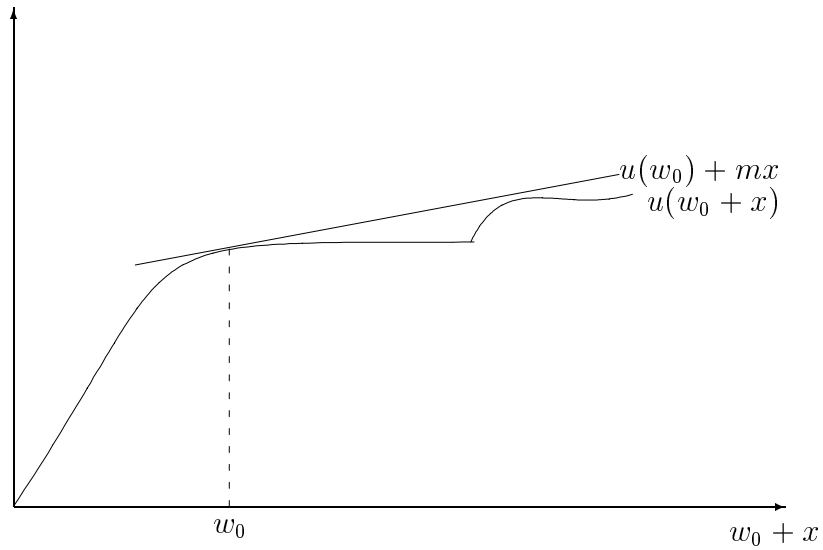

Figure 3.1: An agent who is diffident at w_0.

¿From the definition, the agent is *diffident at* w_0 if this property holds for any pure risk. Using the Diffidence Theorem with $f_1(x) = x$ and $f_2(x) = u(w_0 + x) - u(w_0)$, we obtain the following necessary and sufficient condition:

$$\exists m \in R : \forall x : \qquad u(w_0 + x) \leq u(w_0) + mx. \tag{3.2}$$

Graphically, it means that there exists a straight line containing $(w_0, u(w_0))$ that is entirely above the curve representing u. This is the case for example for the utility function depicted in Figure 3.1 at wealth w_0. If u is differentiable at w_0, $m = u'(w_0)$ is the only candidate to satisfy condition (3.2). A necessary condition for the diffidence of u at w_0 is that u be locally concave at that point. If u is twice differentiable, it means that $u''(w_0)$ is nonpositive.

The problem is different if one requires that the agent dislikes any pure risk, *whatever his initial wealth*:

$$\forall w_0,\ \exists m(w_0) \in R : \forall x : \qquad u(w_0 + x) \leq u(w_0) + m(w_0)x. \tag{3.3}$$

Obviously, this is not true for the agent with utility u in Figure 3.1 since the diffidence condition (3.2) is not satisfied for all initial wealth levels. If condition (3.1) is satisfied for any w_0, the agent is called "risk-averse". In other words, an agent is risk-averse if he is diffident at any wealth level. The difference between risk aversion and diffidence at some wealth level looks subtle at first sight. It will play an important role in more complex risky environments.

Proposition 6 *An agent is diffident at w_0 if and only if condition (3.2) is satisfied. An agent is risk-averse if and only if his utility function is concave.*

Proof: The first part of the Proposition is a direct application of the Diffidence Theorem, as shown above. We prove the second part of the Proposition. Sufficiency is a direct consequence of Jensen's inequality. If u is concave, $Eu(w_0 + \widetilde{x})$ is less than $u(w_0 + E\widetilde{x}) = u(w_0)$. Necessity is obtained by contradiction. Suppose that u is not concave, i.e. there exists a w_0 and a $\epsilon > 0$ such that u is strictly convex in $[w_0 - \epsilon, w_0 + \epsilon]$. Consider a zero-mean random variable $\widetilde{x}$ whose support is in $[-\epsilon, \epsilon]$. Then, using Jensen's inequality for convex functions yields $Eu(w_0+\widetilde{x}) > u(w_0)$. This is a contradiction. Another way of proving necessity is to observe that risk aversion implies local diffidence at any wealth level z, which implies that u be locally concave at any z. As is well-known, pointwise (local) concavity is equivalent to global concavity. ■

Similarly, an agent is risk-lover if his utility function is convex. He is risk-neutral if u is linear. The observation of human behaviours are strongly in favor of the assumption that human beings are risk-averse.[1] For example, most households would want to insure their physical assets if an actuarially fair insurance premium would be offered to cover them. Or, most investors would not purchase risky assets if they would not yield a larger expected return than the risk free asset. These strategies are compatible with the assumption of risk aversion. However, horse betting and other unfair gambling seem to contradict that assumption. Several reasons could explain it, as the possibility of optimism in the assesment of the probability of wining, or the willingness to participate to the financing of public goods often provided by the organizers of lotteries. Anyway, the size of markets for unfair lotteries is marginal with respect to the importance of insurance and financial markets, where risk aversion is a necessity to explain observed strategies and equilibrium prices.

[1] Observing animal behaviours leads to the same conclusion. See Battalio, Kagel and MacDonald (1985) for experiments with rats.

3.2 Comparative diffidence and risk aversion

Human beings are heterogenous in their genes and preferences. Some will value security at a high price, whereas others will not. The first ones will be reluctant to purchase stocks, and they will be willing to insure their endowed risk, even at a high insurance premium. In a word, they are highly risk-averse with respect to the others. It is crucial to make this notion of more risk aversion more precise. Calling agents by their utility function, we have the following two definitions.

Definition 3 *Agent u_1 is more diffident than agent u_2 at wealth level w_0 if the former dislikes all risks for which the latter is indifferent. Agent u_1 is more risk-averse than agent u_2 if the first is more diffident than the other at any wealth level.*

Technically, agent u_1 is more diffident than agent u_2 at w_0 if, for any $\widetilde{x}$,

$$Eu_2(w_0 + \widetilde{x}) = u_2(w_0) \qquad \Longrightarrow \qquad Eu_1(w_0 + \widetilde{x}) \leq u_1(w_0). \tag{3.4}$$

Using the Diffidence Theorem with $f_1(x) = u_2(w_0 + x) - u_2(w_0)$ and $f_2(x) = u_1(w_0 + x) - u_1(w_0)$, the necessary and sufficient condition for v to be more diffident than u at w_0 can be written as:

$$\exists m \in R : \forall x : \;\; u_1(w_0 + x) - u_1(w_0) \leq m \left[u_2(w_0 + x) - u_2(w_0)\right].$$

Observe that $f_1(x_0) = f_2(x_0) = 0$ for $x_0 = 0$. If, in addition, we assume that u_1 and u_2 are differentiable at w_0, we know that $m = u_1'(w_0)/u_2'(w_0)$, so that this condition is rewritten as:

$$\forall x : \;\; \frac{u_1(w_0 + x) - u_1(w_0)}{u_1'(w_0)} \leq \frac{u_2(w_0 + x) - u_2(w_0)}{u_2'(w_0)} \tag{3.5}$$

A couple of functions satisfying this condition is presented in Figure 3.2, where we use the normalization $u_1(w_0) = u_2(w_0)$ and $u_1'(w_0) = u_2'(w_0)$. If the two functions are twice differentiable, the necessary condition (2.11) becomes

$$\frac{u_1''(w_0)}{u_1'(w_0)} \leq \frac{u_2''(w_0)}{u_2'(w_0)},$$

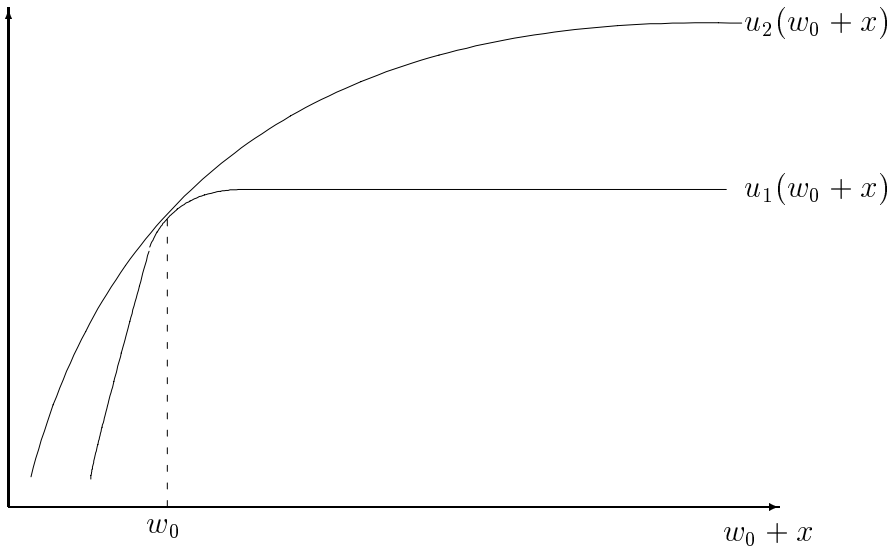

Figure 3.2: u_1 is more diffident than u_2 at initial wealth level w_0.

which is usually rewritten as

$$A_1(w_0) \geq A_2(w_0) \tag{3.6}$$

where $A_i(z) = -u_i''(z)/u_i'(z)$ is the coefficient of absolute risk aversion of agent i. Condition (3.6) is the necessary and sufficient condition for agent 2 to reject any *small* risk for which agent 1 is indifferent, assuming that they have the same wealth w_0. If the risk is not "small" the necessary and sufficient condition is condition (3.5).

As we defined the concept of risk aversion as more diffidence at any wealth level, we defined the notion of "more risk aversion" as being "more diffident at any w_0". Assuming differentiability, u_1 is more risk-averse than u_2 if and only if inequality (3.5) holds not only for any x but also for any w_0. Of course, a necessary condition is that inequality (3.6) holds for any w_0. This set of conditions is made much simpler by observing that these two conditions are in fact equivalent. This is seen by proving that the latter implies the former. To do this, define function ϕ such that $u_1(z) = \phi(u_2(z))$ for all z. Function ϕ transforms u_2 into u_1. We check that

$$\phi'(u_2(z)) = \frac{u_1'(z)}{u_2'(z)} > 0,$$

and

$$\phi''(u_2(z)) = \frac{1}{u_2'(z)}\left[A_2(z) - A_1(z)\right]$$

which is negative if we assume that condition (3.6) holds at any wealth level. Thus, condition (3.6) means that u_1 is a concave transformation of u_2. But then, using Jensen's inequality directly implies that

$$Eu_2(w_0 + \widetilde{x}) = E\phi(u_1(w_0 + \widetilde{x})) \leq \phi(Eu_1(w_0 + \widetilde{x})) = \phi(u_1(w_0)) = u_2(w_0).$$

In conclusion, condition (3.6) for any w_0 is not only necessary, but is also sufficient, for u_1 to be more risk-averse than u_2. A final remark is that this necessary and sufficient condition is nothing but the log-supermodularity of function $h(z, i) = u_i'(z)$. This is a direct application of condition 2 in Lemma 2.

We summarize these results in the next Proposition.

Proposition 7 *Suppose that u_1 and u_2 are twice differentiable. The following conditions are equivalent:*

1. *The agent with utility function u_1 is more risk-averse than the agent with utility function u_2, i.e., the former rejects all lotteries for which the latter is indifferent;*
2. *$h(z, i) = u_i'(z)$ is log-supermodular, i.e., $-u_1''(z)/u_1'(z) \geq -u_2''(z)/u_2'(z)$ for all z;*
3. *u_1 is a concave transformation of u_2;*
4. *Condition (3.5) holds for any w_0.*

This defines the notion of an increase in risk aversion. It is a notion stronger than an increase in diffidence at some specific level w_0, as it appears in Figure 3.2, where u_1 is more diffident than u_2 for an initial wealth level w_0, but obviously A_1 is not uniformly larger than A_2. In particular, $A_1(w_0 + x)$ is zero for large absolute values of x, whereas A_2 is always positive.

3.3 Certainty equivalent and risk premium

In the remaining of this book, we assume that agents are risk-averse: pure risks reduce the level of satisfaction of risk-bearers. In this section, we want to quantify this effect. This can be done by evaluating the maximum amount of money π that one is ready to pay to escape a pure risk. For a risk-averse agent, this amount is positive. π is called the risk premium. It can be calculated by solving the following equation:

$$Eu(w_0 + \widetilde{x}) = u(w_0 - \pi) \tag{3.7}$$

One is indifferent between retaining the risk and paying the risk premium to eliminate the risk. The risk premium is a function of the parameters appearing in the above equation: $\pi = \pi(w_0, u, \widetilde{x})$. A large part of the remaining of this chapter and of the next chapter are devoted to the analysis of this function.

For the risk premium to be a valid measure of risk aversion, it must be that an increase in risk aversion increases the risk premium. Intuitively, a more risk-averse agent should be willing to pay more to escape risk. Suppose that u_1 is more risk-averse than u_2. We show that it implies that $\pi(z_0, u_1, \widetilde{x}) = \pi_1$ is larger than $\pi(z_0, u_2, \widetilde{x}) = \pi_2$, or

$$Eu_2(z_0 + \widetilde{x}) = u_2(z_0 - \pi_2) \implies Eu_1(z_0 + \widetilde{x}) \leq u_1(z_0 - \pi_2) \tag{3.8}$$

for all $\widetilde{x}$, π_2 and w_0. The left-hand side equality means that π_2 is the risk premium that agent u_2 is ready to pay to get rid of the risk, whereas the right-hand side inequality means that π_1 is larger than π_2. Define $w_0 = z_0 - \pi_2$ and $\widetilde{y} = \widetilde{x} + \pi_2$. Then, the above condition is formally equivalent to condition (3.4), which we know is equivalent to u_1 being more risk-averse than u_2.

Proposition 8 *Agent u_1 is always ready to pay more than agent u_2 to escape risk, i.e., $\pi(w_0, u_1, \widetilde{x}) \geq \pi(w_0, u_2, \widetilde{x})$ for any w_0 and $\widetilde{x}$, if and only if u_1 is more risk-averse than u_2.*

We now define the concept of certainty equivalent. Consider a lottery whose net gain is described by random variable $\widetilde{y} = \mu + \widetilde{x}$, with $E\widetilde{x} = 0$. Suppose that you are offered a deal in which you can either play the lottery $\widetilde{y}$, or receive a sure gain C. What is the sure amount C that makes you just indifferent between the

two options? The answer to this question is called the certainty equivalent of the lottery. It verifies condition (3.9).

$$Eu(w_0 + \widetilde{y}) = u(w_0 + C) \tag{3.9}$$

The certainty equivalent is the value of the lottery. It is a function $C = C(w_0, u, \widetilde{y})$. Comparing conditions (3.7) and (3.9) yields

$$C(w_0, u, \mu + \widetilde{x}) = \mu - \pi(w_0 + \mu, u, \widetilde{x}). \tag{3.10}$$

The certainty equivalent of a lottery equals its expected gain minus the risk premium evaluated at the expected wealth of the agent. It implies in particular that an increase in risk aversion reduces certainty equivalents.

3.4 The Arrow-Pratt approximation

We examine in this section the characteristics of the risk premium for small risks. To do this, let us consider a pure risk $k\widetilde{x}$. Let $g(k)$ denote its associated risk premium $\pi(w_0, u, k\widetilde{x})$, i.e.,

$$Eu(w_0 + k\widetilde{x}) = u(w_0 - g(k)).$$

We want to characterize the properties of g around $k = 0$. Observe that $g(0) = 0$. If we assume that u is twice differentiable, fully differentiating the above equality with respect to k allows us to write

$$E\widetilde{x}u'(w_0 + k\widetilde{x}) = -g'(k)u'(w_0 - g(k)),$$

so that $g'(0) = 0$, since $E\widetilde{x} = 0$. Differentiating once again with respect to k yields

$$E\widetilde{x}^2 u''(w_0 + k\widetilde{x}) = [g'(k)]^2 u''(w_0 - g(k)) - g''(k)u'(w_0 - g'(k)).$$

It implies that

$$g''(0) = \frac{-u''(w_0)}{u'(w_0)} E\widetilde{x}^2.$$

Finally, using a Taylor expansion of g around $k = 0$, we obtain that

$$\pi(w_0, u, k\widetilde{x}) \quad = \quad g(k) \backsimeq g(0) + kg'(0) + 0.5k^2 g''(0), \tag{3.11}$$

or, equivalently,

$$\pi(w_0, u, \widetilde{y}) \simeq \quad 0.5\ E\left[\widetilde{y}\right]^2 A(w_0). \tag{3.12}$$

This is the so-called Arrow-Pratt approximation, which allows us to disentangle the characteristics of risk and preferences to evaluate the impact of the former on welfare. In fact, it should have been called the "de Finetti-Arrow-Pratt" approximation, since de Finetti (1952) found it first. This approximation states that the risk premium for a small pure risk is approximately proportional to its variance. As the size k of this risk tends to zero, its risk premium tends to zero as k^2. This property is called "second order risk aversion" by Segal and Spivak (1990). Using condition (3.10), it implies that

$$C(w_0, u, k(\mu + \widetilde{x})) \backsimeq k\mu - k^2 A(w_0) E\widetilde{x}^2.$$

For small k, the certainty equivalent of risk $k(\mu + \widetilde{x})$ equals $k\mu$, and the riskiness of the situation does not matter. Second order risk aversion means that risk yields a second-order effect on welfare, compared to the effect of the mean of the corresponding lottery. This is an important property of expected utility theory.

We now turn to the analysis of the relationship between the risk premium and preferences. The Arrow-Pratt approximation (16.8) tells us that the risk premium is approximately proportional to the coefficient of absolute risk aversion measured at the initial wealth. In crude terms, $A(w_0)$ measures the maximum amount that an agent with wealth w_0 and utility u is ready to pay to get rid of a small risk with a variance of 2. Because the variance has the dimensionality of the square of the monetary unit, A has the dimensionality of the inverse of the monetary unit.

In Figure 16.16, we draw the risk premium as a function of the size of the risk. We took the logarithmic utility function as an illustration. We assumed that

Figure 3.3: The risk premium (solid line) and its approximation (dashed line) as a function of the size k of the risk.

w_0=10 and that $\widetilde{x}$ takes value +1 and -1 with equal probabilities. The Arrow-Pratt approximation gives us $\pi \simeq k^2/20$, which is represented by the dashed parabol in the Figure. We see that this approximation is excellent as long as k is not larger than, say, three. It underestimates the actual risk premium for larger sizes of the risk. Notice that the exact risk premium tends to 10 when the k tends to 10, but the approximation gives only $\pi \simeq 5$!

One can also examine the risk premium associated to a multiplicative risk, that is when final wealth $\widetilde{w}_f$ is the product of initial wealth and a random factor: $\widetilde{w}_f = w_0(1 + k\widetilde{y})$. We define the relative risk premium $\widehat{\pi}(w_0, u, k\widetilde{y})$ as the maximum share of one's wealth that one is ready to pay to escape a risk of losing a random share $k\widetilde{y}$ of one's wealth:

$$\widehat{\pi}(w_0, u, k\widetilde{y}) = \frac{\pi(w_0, u, w_0 k\widetilde{y})}{w_0}$$

If $E\widetilde{y} = 0$ and k is small, then the Arrow-Pratt approximation can be rewritten as

$$\widehat{\pi}(w_0, u, k\widetilde{y}) \backsimeq 0.5\, E\left[k\widetilde{y}\right]^2 \frac{-w_0 u''(w_0)}{u'(w_0)}. \tag{3.13}$$

The relative risk premium is approximately proportional to

$$R(w_0) = \frac{-w_0 u''(w_0)}{u'(w_0)} = w_0 A(w_0)$$

which is called the coefficient of relative risk aversion. It has the advantage over the coefficient of absolute risk aversion to be independent of the monetary unit for wealth.

3.5 Decreasing absolute risk aversion

We now turn to the relationship between the risk premium and initial wealth. It is widely believed that the more wealthy we are, the smaller is the maximum amount we are ready to pay to escape a given additive risk. This corresponds to the notion of decreasing absolute risk aversion (DARA), which is fromally defined as follows.

Definition 4 *Preferences exhibit decreasing absolute risk aversion if the risk premium associated to any risk is a decreasing function of wealth:* $\frac{\partial \pi(w_0,u,\widetilde{x})}{\partial w_0} \leq 0$, *for any* $w_0, \widetilde{x}$.

¿From the Arrow-Pratt approximation, we know that this is true for small risks if and only if $A(w_0)$ be decreasing in w_0. We prove now that it is in fact the necessary and sufficient condition for π to be decreasing in w_0, independent of the size of $\widetilde{x}$. Fully differentiating condition (3.7) with respect to w_0 yields

$$\frac{\partial \pi(w_0,u,\widetilde{x})}{\partial w_0} = \frac{u'(w_0 - \pi(w_0,u,\widetilde{x})) - Eu'(w_0+\widetilde{x})}{Eu'(w_0+\widetilde{x})}$$

This is negative if $Eu'(w_0+\widetilde{x}) \geq u'(w_0-\pi)$, where $\pi = \pi(w_0,u,\widetilde{x})$. Thus, the risk premium is decreasing in wealth if the following property holds:

$$\forall w_0, \pi, \widetilde{x}: \; Eu(w_0+\widetilde{x}) = u(w_0-\pi) \implies Eu'(w_0+\widetilde{x}) \geq u'(w_0-\pi). \tag{3.14}$$

This condition is formally equivalent to condition (3.8) with $u_2 \equiv u$ and $u_1 \equiv -u'$. But we know that the necessary and sufficient condition for (3.8) is that u_1 be more risk-averse than u_2. In consequence, the necessary and sufficient condition for the risk premium to be decreasing in wealth is that $-u'$ be more concave than u. Let $P(w_0) = -u'''(w_0)/u''(w_0)$ denote the degree of concavity of $-u'$. P is called the degree of absolute prudence and is analyzed in more details in Chapter 14. To sum up, we have a risk premium that is decreasing in wealth if and only if

$$P(w_0) \geq A(w_0). \tag{3.15}$$

This is equivalent to the condition that A be decreasing in wealth, since it is easily verified that

$$A'(w_0) = A(w_0)\left[A(w_0) - P(w_0)\right]. \tag{3.16}$$

Finally, observe that condition (3.15) is equivalent to $h(w_0, x) = u'(w_0 + x)$ being log-supermodular.

Proposition 9 *Suppose that u is three times differentiable. The following conditions are equivalent:*

1. *Preferences exhibit decreasing absolute risk aversion;*

2. *$u'(w_0 + x)$ is log-supermodular with respect to (w_0, x), i.e., $A'(z) \leq 0$ or $-u'''(z)/u''(z) \geq -u''(z)/u'(z)$ for all z;*

3. *$-u'$ is a concave transformation of u.*

Notice that DARA requires that u' be (sufficiently) convex, or $u''' \geq 0$.

A similar exercise could be performed on the notion of relative risk aversion. We would get that the relative risk premium for a given multiplicative risk is decreasing in wealth if and only if relative risk aversion is decreasing in wealth. However, there is no strong evidence in favor of such hypothesis, contrary to what we have for DARA. Notice that decreasing relative risk aversion is stronger than DARA, since

$$R'(w_0) = w_0 A'(w_0) + A(w_0).$$

The intuition is that, as wealth increases, the risk itself increases, which tends to raise the risk premium.

3.6 Some classical utility functions

It is often the case that problems in the economics of uncertainty are intractable if no further assumption is made on the form of the utility function. A specific subset of utility functions has emerged as being particularly useful to derive analytical results. This is the set of utility functions with an harmonic absolute risk aversion (HARA), i.e., with an inverse of absolute risk aversion that is linear in wealth. The inverse of absolute risk aversion is called absolute risk tolerance and is denoted T with:

$$T(z) = \frac{1}{A(z)} = -\frac{u'(z)}{u''(z)}.$$

Utility functions exhibiting an absolute risk tolerance that is linear in wealth take the following form:

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma} \tag{3.17}$$

These functions are defined on the domain of z such that $\eta + \frac{z}{\gamma} > 0$. We have that

$$u'(z) = \zeta\frac{1-\gamma}{\gamma}(\eta + \frac{z}{\gamma})^{-\gamma} \tag{3.18}$$

$$\begin{aligned} u''(z) &= -\zeta\frac{1-\gamma}{\gamma}(\eta + \frac{z}{\gamma})^{-\gamma-1} \\ u'''(z) &= \zeta\frac{(1-\gamma)(\gamma+1)}{\gamma^2}(\eta + \frac{z}{\gamma})^{-\gamma} \end{aligned}$$

In order to guarantee that $u' > 0$ and $u'' < 0$, we need to have $\zeta(1-\gamma)\gamma^{-1} > 0$. The different coefficients related to the attitude towards risk are thus equal to:

$$A(z) = (\eta + \frac{z}{\gamma})^{-1} \tag{3.19}$$

$$P(z) = \frac{\gamma+1}{\gamma}(\eta + \frac{z}{\gamma})^{-1} \tag{3.20}$$

$$R(z) \quad = \quad z(\eta + \frac{z}{\gamma})^{-1} \tag{3.21}$$

Observe that the inverse of absolute risk aversion is indeed a linear function of z. By changing parameters η and γ, we are able to generate all utility functions exhibiting an harmonic absolute risk aversion. Three special cases of HARA utility functions have been analyzed in particular:

- **Constant relative risk aversion (CRRA)**: If $\eta = 0$, then $R(z)$ equals γ, which must be assumed to be nonnegative. Thus, the case $\eta = 0$ corresponds to the situation where relative risk aversion is independent of wealth: the relative risk premium that one is ready to pay to escape a multiplicative risk does not depend upon wealth. Now, use condition (3.18) by selecting ζ in such a way to normalize $u'(1) = 1$. It implies that his set of utility functions is defined as

$$u'(z) = z^{-\gamma},$$

which means that

$$u(z) = \begin{cases} \frac{z^{1-\gamma}}{1-\gamma} & \text{if } \gamma \neq 1 \\ \ln(z) & \text{if } \gamma = 1. \end{cases} \tag{3.22}$$

Some of these functions are represented in Figure 3.4. The thick curve is the logarithmic function. The curves above (resp. below) it represent utility functions with a relative risk aversion larger (resp. smaller) than 1. It is noteworthy that the asymptotic properties of these functions are quite different depending upon relative risk aversion is larger or smaller than unity. When γ is less than 1, utility comes from $-\infty$ to zero as wealth goes from zero to infinity. But when γ is larger than unity, utility goes from 0 to infinity. However, all these functions exhibit DARA, since $A'(z) = -\gamma/z^2$.

Figure 3.4: CRRA utility functions

- **Constant absolute risk aversion (CARA)**: We see from (3.19) that absolute risk aversion $A(z)$ is independent of z if $\gamma \to +\infty$. We obtain $A(z) = A = 1/\eta$ in this case. It is easy to check that the set of functions u that satisfy the differential equation $-u''(z) = Au'(z)$ corresponds to:

$$u(z) = -\frac{\exp(-Az)}{A}. \tag{3.23}$$

All these functions exhibit increasing relative risk aversion.

Figure 3.5: CARA utility functions

- **Quadratic utility functions**: We obtain quadratic utility functions by selecting $\gamma = -1$. There are two problems with this set of functions. First,

they are defined only on the interval of wealth $z < \eta$, since they are decreasing above η. Second, they exhibit increasing absolute risk aversion, which is not compatible with the observation that risk premia for additive risks are decreasing with wealth.

We want to stress here that these functions have been considered in the literature because they are easy to manipulate. There is no reason to believe that they represent the attitude towards risk of agents in the real world.

Another particularly useful class of utility function is represented by piecewise linear functions with a single kink. They take the following form:

$$u(z) = \begin{cases} z & \text{if } z \leq z_0 \\ z_0 + a(z - z_0) & \text{if } z > z_0. \end{cases} \tag{3.24}$$

This function is concave if $a \leq 1$. Absolute risk aversion is zero everywhere except at kink z_0 where it is not defined, since u is not differentiable at z_0. The only utility functions that are more risk-averse than u are those similar to u, but with a constant a being replaced by $b < a$. One can verify that the latter function is a concave transformation of the former. Notice that a risk-neutral agent who has a gross income z that is taxed at a marginal rate $1 - a$ above z_0 will have the same behavior toward risk as an agent with utility u who is not taxed. More generally, this analogy tells us that an increasing marginal tax rate increases the aversion toward risk of the taxpayer.

3.7 Test for your own degree of risk aversion

Several authors tried to build the utility function of economic agents through questionnaires and tests in laboratories. The kind of questions was as follows:

> ”Suppose that your wealth is currently equal to 100. This wealth is composed of a sure asset and a house whose value is 40. There is a p=5% probability that a fire totally destroys your house. How much would you be ready to pay to fully insure this risk?”

If we normalize the utility function in such a way that $u(100) = 1$ and $u(60) = 0$, an answer k to this question would allow us to compute

$$u(100 - k) = 0.05u(60) + 0.95u(100) = 0.95.$$

Changing the value of p in the questionnaire would then allow to draw curve u point by point. However, the sensitivity of the answer to the environment and the difficulty to make the test as much real as possible make the results relatively weak. This direction of research has long been abandoned. A less ambitious goal has been pursued, which is simply to estimate the local degree of risk aversion of economic agents, rather than to estimate their full utility function. This has been done via market data where the actual decisions made by agents have been observed. Insurance markets and financial markets are particularly useful, since risk and risk aversion play the driving role for the exchanges in those markets. Because we did not yet link decisions of agents in these markets to their degree of risk aversion, we postpone the presentation of these analyses to Chapter .

Still, you will not be in a situation to apply any theory presented here to your own life if you have no idea of your own degree of risk aversion. Our aim is thus to offer a very simple exercise that allows you to make a crude estimation of your degree of risk aversion. Also, for the remaining of this book, it is important that you have a notion of what is a reasonable degree of risk aversion. We stick to an estimation of relative risk aversion, since it is without any dimension.

Since we don't try to estimate the functional form of u, let us just make an assumption on it. More specifically, let us assume that your utility function takes the form (3.22). As said earlier, this is a strong assumption without any empirical justification. This assumption is done just for simplicity. Then, answer the following question: What is the share of your wealth that you are ready to pay to escape the risk of gaining or losing a share α of it with equal probability? Think about it for a few seconds before pursuing this reading.

Table 3.1 allows you to translate your answer, $\widehat{\pi}$, to your degree of relative risk aversion (RRA) γ, if $\alpha = 10\%$ or $\alpha = 30\%$. You can do the computation by yourself by solving the following equation:

$$0.5\frac{(1-\alpha)^{1-\gamma}}{1-\gamma} + 0.5\frac{(1+\alpha)^{1-\gamma}}{1-\gamma} = \frac{(1-\widehat{\pi})^{1-\gamma}}{1-\gamma}$$

RRA	$\alpha = 10\%$	$\alpha = 30\%$
$\gamma = 0.5$	0.3%	2.3%
$\gamma = 1$	0.5%	4.6%
$\gamma = 4$	2.0%	16.0%
$\gamma = 10$	4.4%	24.4%
$\gamma = 40$	8.4%	28.7%

Table 3.1: Relative risk premium $\widehat{\pi}$ associated to the risk of gaining or losing a share α of wealth, when the utility function is as in (3.22).

If we focus on the risk of gaining or losing 10% of one's wealth, I would consider an answer $\widehat{\pi} = 0.5, ..., 2\%$ as something reasonable. It implies that it is reasonable to believe that relative risk aversion is somewhere between 1 and 4. Saying it differently, a relative risk aversion superior to 10 seems totally foolish, as it implies very high relative risk premium. Notice in particular that a relative risk aversion of 40 implies that one would be ready to pay as much as 8.4% to escape the risk of gaining or losing 10% of one's wealth!

Notice that a side product of this exercise is a test on the validity of the Arrow-Pratt approximation (3.13) which states that the relative risk premium must be approximately equal to half the product of the variance of the multiplicative risk and γ. For $\alpha = 0.1$, the variance of the multiplicative risk equals 0.01. For $\gamma = 1$, it would predict a $\widehat{\pi}$ equaling to 0.005 which is the true value, at least for the first three decimals. For $\gamma = 10$, it would predict $\widehat{\pi} = 0.05$, which is 10% larger than the true value. The reader can verify that the Arrow-Pratt approximation entails a larger error for the larger risk $\alpha = 30\%$.

3.8 An application: The cost of macroeconomic risks

Let us consider a very simplistic representation of our Society, in which every agent would be promised the same share of the aggregate production in the future. This means that the risk borne by every agent is the risk affecting the growth of GDP per head. All other risks are washed away by diversification. How this can be made possible in more complex, decentralized economies is the subject of other chapters in this book. In order to assess the degree of risk associated to every one's income in the future, let us examine how volatile the growth of GDP per capita as been in the past. We reproduce in Figure 3.6 the time serie of US GDP per capita in 1985 dollars for the period from 1963 to 1992. By long-term historical standards, per capita GDP growth is high: around 1.86% per year. However, this growth rate has been variable over time, as shown in Figure 3.7. This serie appear stationary and ergodic. The standard deviation of the growth rate has been around 2.41%.

Figure 3.6: U.S. Real GDP per capita in 1985 dollars. (Source: Penn World)

Figure 3.7: Growth rate of the US real GDP per capita, in %.

Figure 3.8: Frequency table of the growth of real GDP per capita, USA, 1963-1992.

Suppose that we attach an equal probability that any of the annual growth rate observed during the period 1963-92 occurs next year. That is, transform frenquencies into probabilities. Let $\widetilde{g}$ be the random variable with such probability distribution. If we assume that agents have constant relative risk aversion γ, one can estimate the certainty equivalent growth rate, g_c, by using the following equality:

$$E\frac{(100+\widetilde{g})^{1-\gamma}}{1-\gamma} = \frac{(100+g_c)^{1-\gamma}}{1-\gamma} \tag{3.25}$$

The social cost of risk can be measured by the risk premium associated to $\widetilde{g}$. It is the reduction in the growth rate one would be ready to pay to get rid of the macroeconomic risk. We computed the certainty equivalent growth rate and the social cost of risk for different values of γ. The results are reported in Table 3.2. We see that it does not differ much from the expected growth rate for acceptable values of γ. If we take $\gamma = 4$ for a reasonable upper bound, *the macroeconomic risk does not cost us more than one-tenth of a percent of growth*. The implication of this is that not too much should be done to reduce this uncertainty. Policy makers should rather concentrate their effort in maximizing the expected growth. The existing macroeconomic risk is just not large enough to justify spending much of our energy to reduce it.

RRA	certainty equivalent growth rate	social cost of macroeconomic risk
$\gamma = 0$	1.86%	–
$\gamma = 0.5$	1.85%	0.01%
$\gamma = 1$	1.83%	0.03%
$\gamma = 4$	1.74%	0.12%
$\gamma = 10$	1.56%	0.30%
$\gamma = 40$	0.50%	1.36%

Table 3.2: Social cost of macroeconomic risk

3.9 Conclusion

Two important notions have been introduced in this chapter. The first one is risk aversion, which means that one rejects any zero-mean risk, whatever one's initial wealth. Risk aversion is equivalent to the concavity of the utility function. We also

determined whether an agent is more risk-averse than another, i.e. whether the first rejects any lottery for which the former is indifferent, whatever their identical initial wealth level. This is equivalent to requiring that the utility function of the first is a concave transformation of the utility function of the second. We also explained that weaker conditions hold if one wants these properties to hold only for a specific initial wealth level.

Chapter 4

Change in risk

In the previous chapter, we examined the notion of risk premium. The risk premium π depends basically upon the characteristics of the utility function and the characteristics of the risk. We also characterized the changes in the utility function that have an unambiguous effect on the risk premium, independently of the risk under consideration. This corresponded to the concept of ”more risk aversion”. In this chapter, we perform the symmetric exercise, which is to look for changes in risk that have an unambiguous effect on the risk premium, independently of the utility function under consideration.

Let us normalize w_0 to zero. Observe that $\pi(0, u, \widetilde{x}_1)$ is larger than $\pi(0, u, \widetilde{x}_2)$ if and only if

$$Eu(\widetilde{x}_1) \leq Eu(\widetilde{x}_2). \tag{4.1}$$

The theory of stochastic dominance is devoted to determining under which condition on the distribution functions of $\widetilde{x}_1$ and $\widetilde{x}_2$ the above inequality holds for any utility function u in a function set Υ. Of course the answer to this question depends upon which set Υ is considered. This problem is complex because the dimensionality of Υ can be very large. In the next section, we provide a general method to solve this kind of problem. As we will see, the technique consists again in reducing the dimensionality of the problem.

At this stage, it is important to observe that as soon as Υ contains more than one utility function, the associated stochastic dominance ordering is incomplete in the sense that there exist pairs $(\widetilde{x}_1, \widetilde{x}_2)$ such that $\widetilde{x}_1$ does not dominate $\widetilde{x}_2$ and $\widetilde{x}_2$ does not dominate $\widetilde{x}_1$. This is the case when two agents in Υ diverge on their ordering of these two risks, so that no unanimity emerges.

To escape potential problems of convergence of the integrals, we limit the analysis to random variables whose supports are in interval $[a, b]$.

4.1 The basis approach

The general technique used in stochastic dominance is based on the fact that the expectation operator is additive:

$$\left.\begin{array}{l} Eu_0(\widetilde{x}_1) \leq Eu_0(\widetilde{x}_2) \\ Eu_1(\widetilde{x}_1) \leq Eu_1(\widetilde{x}_2) \end{array}\right\} \Longrightarrow \forall \lambda \in [0,1] : Ev_\lambda(\widetilde{x}_1) \leq Ev_\lambda(\widetilde{x}_2)$$

where $v_\lambda(z) = (1-\lambda)u_0(z) + \lambda u_1(z)$ for all z. This means that if condition (4.1) is satisfied for some utility functions $(u_0, u_1, ..., u_n)$, it is also satisfied for any utility function which is a convex combination of them. The basis approach consists in searching for a small subset B of Υ such that any element in Υ can be written as a convex combination of elements in B. In our terminology, B is called a basis for Υ. Requiring that condition (4.1) holds for elements in B will not only be necessary, but also sufficient, for condition (4.1) to hold for any element in Υ. If the dimensionality of B is smaller than the dimensionality of Υ, our task is made simpler.

Let us illustrate the basis approach on the following simple example. Suppose that Υ is the set of increasing, concave piecewise linear functions with two kinks, one at z_0 and the other at z_1, and such that utility is constant for wealth levels larger than z_1. That is, Υ is the set of functions u such that

$$\exists a \in [0,1] : \; u(z) = \begin{cases} z & \text{if } z < z_0 \\ z_0 + a(z - z_0) & \text{if } z_0 \leq z < z_1 \\ z_0 + a(z_1 - z_0) & \text{if } z \geq z_1. \end{cases} \tag{4.2}$$

This function is parametrized by a scalar a. Υ has an infinite number of functions, but its dimensionality is 1. Notice that any function in Υ is a convex combination of $u_0(.) = \min(., z_0)$ and $u_1(.) = \min(., z_1)$. These two basic functions together with a function u of type (4.2) are depicted in Figure (4.1). These two functions represent a basis for Υ since function (4.2) equals $(1-a)u_0 + au_1$. We conclude that it is enough to verify that

$$E \min(\widetilde{x}_1, z_0) \leq E \min(\widetilde{x}_2, z_0)$$

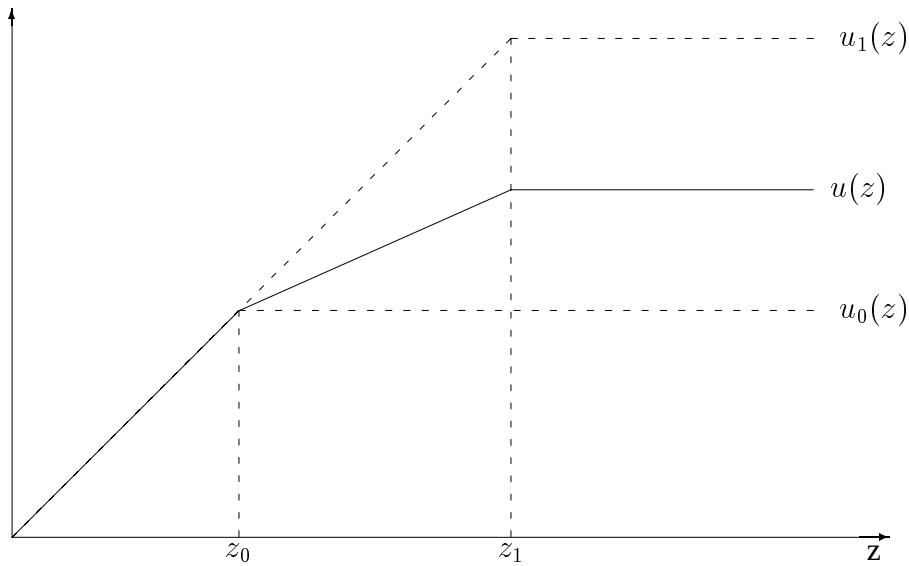

Figure 4.1: Function u is a convex combination of functions u_0 and u_1.

and

$$E\min(\widetilde{x}_1, z_1) \leq E\min(\widetilde{x}_2, z_1)$$

to guarantee that all functions of type (4.2) satisfy condition (4.1).

In general, the function set Υ is not as simple as this one. In the remaining of this chapter, $b(., \theta)$ will denote a function in the basis, where θ is an index in some index set $\Theta \subset R$. It means that for any utility function in set Υ, there exists a nondecreasing function $H : \Theta \to [0, 1]$ such that

$$u(z) = \int_\Theta b(z, \theta) dH(\theta)$$

for all $z \in [a, b]$, where H satisfies the condition that

$$\int_\Theta dH(\theta) = 1.$$

H is the transform of u with respect to basis B. $dH(\theta)$ can be interpreted as the weight associated to function $b(.,\theta)$. It must be positive and it must sum up to 1. In fact, H can be interpreted as a cumulative distribution function on a random variable $\widetilde{\theta}$, and u can be written as

$$u(z) = Eb(z, \widetilde{\theta}).$$

To sum up, we replace problem (4.1) for all $u \in \Upsilon$ by the following one:

$$Eb(\widetilde{x}_1, \widetilde{\theta}) \leq Eb(\widetilde{x}_2, \widetilde{\theta})$$

for all random variable $\widetilde{\theta}$ whose support is in Θ. The analogy with problem (2.1) now appears clearly if we define $f_i(\theta) = Eb(\widetilde{x}_i, \theta)$. The necessary and sufficient condition is written as

$$Eb(\widetilde{x}_1, \theta) \leq Eb(\widetilde{x}_2, \theta) \qquad \forall \theta \in \Theta. \tag{4.3}$$

This is another way to say that $\widetilde{x}_2$ is unanimously preferred to $\widetilde{x}_1$ by population Υ if it is unanimously preferred by the subset of agents in B.

4.2 Second-order stochastic dominance

In this section, we determine the conditions under which inequality (4.1) holds for any increasing and concave utility function. In that case, we say that $\widetilde{x}_2$ dominates $\widetilde{x}_1$ in the sense of second-order stochastic dominance (SSD). Let Υ_2 be the set of increasing and concave functions from $[a, b]$ to R. It is a convex set. As is well-known, a simple basis for Υ_2 is the set B_2 of min-functions:

$$B_2 \equiv \{b(.,\theta) = \min(.,\theta) \mid \theta \in \Theta \equiv [a, b]\}.$$

To show that any increasing and concave function can be expressed as a convex combination of $\min$ functions, let us take an arbitrary $u \in \Upsilon_2$. The weighting function H must satisfy:

$$u(z) = \int_a^b \min(z, \theta) dH(\theta)$$

for any $z \in [a, b]$. Integrating by parts yields

$$u(z) = zH(b) - aH(a) - \int_a^b I(\theta \leq z)H(\theta)d\theta = zH(b) - aH(a) - \int_a^z H(\theta)d\theta,$$

where $I(x)$ is the indicator function which takes value 1 if x is true, and 0 otherwise. Differentiating the above equality with respect to z yields $u'(z) = H(b) - H(z)$. If u is twice differentiable, taking $H'(\theta) = u''(\theta)$ will give us the weight function. At wealth levels z where u' is not differentiable, H discontinuously jumps by $u'(z_-) - u'(z_+)$, which corresponds to an atom in the distribution of $\widetilde{\theta}$.

This shows that the necessary and sufficient condition (4.3) is written here as:

$$E\min(\widetilde{x}_1, \theta) \leq E\min(\widetilde{x}_2, \theta) \qquad \forall \theta \in [a, b]\,.$$

Let F_i denote the cumulative distribution function of $\widetilde{x}_i$. Then the above inequality is equivalent to

$$\int_a^\theta x dF_1(x) + \theta(1 - F_1(\theta)) \leq \int_a^\theta x dF_2(x) + \theta(1 - F_2(\theta)),$$

or, integrating by parts, to

$$\theta - \int_a^\theta F_1(x)dx \geq \theta - \int_a^\theta F_2(x)dx.$$

This is in turn equivalent to

$$\int_a^\theta F_1(x)dx \leq \int_a^\theta F_2(x)dx \qquad \forall \theta \in [a, b]\,. \tag{4.4}$$

This set of inequalities is called the Rothschild-Stiglitz's (1970) integral condition for SSD, although this condition already appeared in Hardy, Littlewood and Polya (1929) and Hadar and Russell (1969). This condition means that all min-utility individuals prefer $\widetilde{x}_2$ over $\widetilde{x}_1$. This is sufficient to guarantee that all risk-averse agents prefer $\widetilde{x}_2$ over $\widetilde{x}_1$.

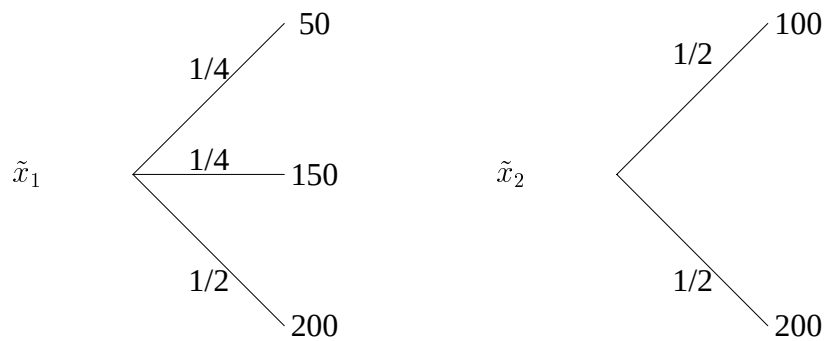

Figure 4.2: $\tilde{x}_1$ is an increase in risk of $\tilde{x}_2$.

Notice that a necessary condition for a SSD deterioration of a risk is that the mean is not increased. This is a consequence of definition (4.1) where we take u as the identity function, which is in Υ_2. This can also be seen by recalling the following statistical relationship that can be obtained by integrating by parts:

$$E\widetilde{x}_i = b - \int_a^b F_i(x)dx.$$

Combining this with condition (4.4) for $\theta = b$ yields the result. A limit case is when the expectation of the risk is unchanged by the shift in distribution. A SSD deterioration in risk that preserves the mean is called an increase in risk (IR). We present in Figure 4.2 an example of a pair $(\widetilde{x}_1, \widetilde{x}_2)$ such that $\widetilde{x}_1$ is an increase in risk with respect to $\widetilde{x}_2$. Observe that we split the probability mass at 100 into two equal probability masses respectively at 50 and 150. This is an example of a mean-preserving spread (MPS) of probability mass. More generally, define an interval $X \subset [a, b]$. A MPS is obtained by taking some probability mass from X to be distributed outside X, in such a way to preserve the mean. In the case of a continuous random variable, a spread around X is defined by the condition that $(F_1'(x) - F_2'(x))$ is positive (resp. negative) outside (resp. inside) X. The reader can easily verify that this is a sufficient condition for (4.4). Rothschild and Stiglitz (1970) showed that any increase in risk can be obtained by a sequence of such mean-preserving spreads.

Another way to look at the change in risk presented in Figure 4.2 is to say that

$\widetilde{x}_1$ is obtained from $\widetilde{x}_2$ by adding a white noise $\widetilde{\epsilon}$ to the low realization of $\widetilde{x}_2$, with $\widetilde{\epsilon}$ being distributed as (-50,1/2;50,1/2). More generally, if

$$\widetilde{x}_1 =_d \widetilde{x}_2 + \widetilde{\epsilon} \text{ with } E\left[\widetilde{\epsilon} \mid \widetilde{x}_2 = x\right] \leq 0 \; \forall x, \tag{4.5}$$

a direct use of Jensen's inequality yields

$$Eu(\widetilde{x}_1) = E_{\widetilde{x}_2}\left[E_{\widetilde{\epsilon}} u(\widetilde{x}_2 + \widetilde{\epsilon}) \mid \widetilde{x}_2\right] \leq E_{\widetilde{x}_2}\left[u(\widetilde{x}_2 + E\left[\widetilde{\epsilon} \mid \widetilde{x}_2\right])\right] \leq Eu(\widetilde{x}_2).$$

Thus, condition (4.5) is sufficient for SSD. Rothschild and Stiglitz (1970) showed that it is also necessary.

To sum up, an increase in risk from $\widetilde{x}_2$ to $\widetilde{x}_1$ can be defined by any of four equivalent statements:

- the means are the same and risk-averse agents unanimously prefer $\widetilde{x}_2$ over $\widetilde{x}_1$: $\forall$ concave: $Eu(\widetilde{x}_1) \leq Eu(\widetilde{x}_2)$.
- the Rothschild-Stiglitz's condition (4.4) is satisfied; it is satisfied as an equality for $\theta = b$.
- $\widetilde{x}_1$ is obtained from $\widetilde{x}_2$ by a sequence of mean-preserving spreads.
- $\widetilde{x}_1$ is obtained from $\widetilde{x}_2$ by adding a white noise $\widetilde{\epsilon}$ to it, with $E\left[\widetilde{\epsilon} \mid \widetilde{x}_2 = x\right] = 0$ for all x.

Several other preference orderings than expected utility with $u'' < 0$ are such that agents unanimously dislike a SSD change in distribution. Such preference orderings are said to satisfy the SSD property.

4.3 Diversification

There are many ways to reduce risk. The most well-known one is diversification. Consider a set of n lotteries whose net gains are characterized by $\widetilde{x}_1, \widetilde{x}_2, ...\widetilde{x}_n$ that are assumed to be independent and identically distributed. You are offered to participate to these gambles. If you are allowed to split your stake in different gambles, how should you do it? Common wisdom suggests that one should not "put all one's eggs in the same basket". This is because diversification is an efficient way to reduce risk, as we now show.

A feasible strategy in this game is characterized by a vector $A = (\alpha_1, ..., \alpha_n)$ where α_i is one's share in gamble i, with $\sum_{i=1}^n \alpha_i = 1$. It yields a net payoff equaling $\sum_{i=1}^n \alpha_i \widetilde{x}_i$. We have the following Proposition, which is a generalization of Rothschild and Stiglitz (1971).

Proposition 10 *The distribution of final wealth generated by the perfect diversification strategy* $(\frac{1}{n}, ..., \frac{1}{n})$ *SSD-dominates the distribution of final wealth generated by any other feasible strategy.*

Proof: Let $\widetilde{y}$ denote final wealth under the perfect diversification strategy. Consider now any alternative feasible strategy $A = (\alpha_1, ..., \alpha_n)$. We have that

$$\sum_{i=1}^n \alpha_i \widetilde{x}_i = \widetilde{y} + \sum_{i=1}^n (\alpha_i - \frac{1}{n}) \widetilde{x}_i.$$

We would be done if noise $\sum_{i=1}^n (\alpha_i - 1) \widetilde{x}_i$ has a zero mean conditional to $\widetilde{y}$. We have that

$$E\left[\sum_{i=1}^n (\alpha_i - \frac{1}{n}) \widetilde{x}_i \mid \widetilde{y}\right] = \sum_{i=1}^n (\alpha_i - \frac{1}{n}) E\left[\sum_{i=1}^n \widetilde{x}_i \mid \widetilde{y}\right].$$

By symmetry, $E\left[\sum_{i=1}^n \widetilde{x}_i \mid \widetilde{y}\right]$ is independent of i. Let us denote it k. It implies that

$$E\left[\sum_{i=1}^n (\alpha_i - \frac{1}{n}) \widetilde{x}_i \mid \widetilde{y}\right] = k \sum_{i=1}^n (\alpha_i - \frac{1}{n}) = 0.$$

This proves that the alternative strategy is second-order stochastically dominated by the perfect diversification strategy.■

It implies that all risk-averse agents will select the perfect diversification strategy. At the limit, when n tends to infinity, this strategy allows to replace the unit risky gamble $\widetilde{x}_1$ by its expectation. This is the Law of Large Numbers.

4.4 First-order stochastic dominance

A more restrictive stochastic order is obtained by enlarging the set of utility functions that must satisfy condition (4.1). This is the case when we allow risk-loving behaviors.. Let Υ_1 denote the set of all increasing functions from $[a, b]$ to R. If condition (4.1) holds for all utility functions in Υ_1, we say that $\widetilde{x}_2$ dominates $\widetilde{x}_1$ in the sense of the first-order stochastic dominance (FSD) order. Using the same technique as above, it is straightforward to verify that the set B_1 of step-functions

$$B_1 \equiv \{b(.,\theta) \mid b(x,\theta) = I(x \leq \theta),\ \theta \in \Theta \equiv [a,b]\}\ .$$

is a basis for Υ_1. For a differentiable utility function, this can be seen by observing that

$$u(z) = u(a) + \int I(t \leq z)u'(t)dt.$$

It implies that $\widetilde{x}_2$ is preferred to $\widetilde{x}_1$ for all expected-utility maximizers (risk-averse or risk-lover) if and only if $\widetilde{x}_2$ is preferred to $\widetilde{x}_1$ for all expected-utility maximizers with a step utility function, i.e., if

$$F_1(\theta) \geq F_2(\theta) \qquad \forall \theta \in \Theta \equiv [a,b]\ . \tag{4.6}$$

In words, the change in risk must increase the probability that the realized value of the risk be less than a threshold, whatever the threshold. Such a change in risk can be obtained by shifting probability masses to the left, or by adding a noise $\widetilde{\epsilon}$ to the original risk $\widetilde{x}_1$ such that $\Pr\left[\widetilde{\epsilon} \leq 0 \mid \widetilde{x}_2 = x\right] = 1$ for all x.

A particular case of FSD shifts in distribution is when $k(\theta) = F_1(\theta)/F_2(\theta)$ is nonincreasing in θ. This is equivalent to say that $F_i(\theta)$ is log-supermodular in (θ, i). This property defines the monotone probability ratio (MPR) stochastic order. Since by definition $k(b)$ equals 1, the fact that k is nonincreasing implies that $k(\theta) \geq 1$ for all θ less than b. This is equivalent to the FSD condition (4.6).

A more restrictive condition is the so-called monotone likelihood ratio (MLR) property. Limiting our analysis to random variables $\widetilde{x}_i$ having a density F_i' which is positive on $[a, b]$,[1] this stochastic dominance order is defined by the condition

[1] See Athey (1997) and Gollier and Schlee (1997) for more general definitions of the MLR order.

that $l(\theta) = F_1'(\theta)/F_2'(\theta)$ be nonincreasing, or that $F_i'(\theta)$ be log-supermodular in (θ, i). In the next Proposition, we show that MLR is a particular case of MPR.

Proposition 11 *MLR $\Longrightarrow$ MPR $\Longrightarrow$ FSD $\Longrightarrow$ SSD.*

Proof: We have just to prove that MLR implies MPR. This simple proof is taken from Athey (1997). Let $g(x, i)$ denote $g_i(x)$ with $g = F$ or F'. Then, observe that

$$F(x, i) = \int I(\theta \leq x) F'(\theta, i) d\theta. \tag{4.7}$$

We know that the indicator function $I(\theta \leq x)$ is LSPM in (x, i, θ) and, by assumption, that $F'(\theta, i)$ is LSPM in (x, i, θ). Remember that the product of two LSPM functions is LSPM. Thus, the integrand in equation (4.7) is LSPM. Using Proposition 5 with $\mathbf{z} = (x, i)$ yields the result that $F(x, i)$ is LSPM.■

When risks have support in $\{x_1, x_2, x_3\}$, we can use the Machina triangle to represent all lotteries that are dominated by lottery A with respect to the dominance orders defined in this Chapter. This is done in Figure 4.3 by assuming without loss of generality that $x_1 < x_2 < x_3$.

4.5 Jewitt's preference orders

Consider an individual u that is indifferent between $\widetilde{x}_1$ and $\widetilde{x}_2$. It is intuitive that $\widetilde{x}_1$ is more risky than $\widetilde{x}_2$ if all individuals who are more risk-averse than u prefer $\widetilde{x}_2$ over $\widetilde{x}_1$, or, equivalently, if all those who are more risk-lover than u prefer $\widetilde{x}_1$ over $\widetilde{x}_2$. If this is true for any $u \in \Upsilon$, we say that $\widetilde{x}_2$ dominates $\widetilde{x}_1$ in the sense of Jewitt's stochastic dominance associated to Υ. We address this question for the case $\Upsilon = \Upsilon_1$. Let the utility function $u(x, \theta)$ be indexed by a risk-loving parameter θ, i.e., an increase in θ represents a reduction in the degree of concavity à la Arrow-Pratt. This means that u' is LSPM. Our problem can then be written as: $\forall u(., \theta_L) \in \Upsilon_1$:

$$Eu(\widetilde{x}_1, \theta_L) = Eu(\widetilde{x}_2, \theta_L) \implies \forall \theta_H > \theta_L : \; Eu(\widetilde{x}_1, \theta_H) \geq Eu(\widetilde{x}_2, \theta_H) \tag{4.8}$$

Let $G \equiv F_2 - F_1$. The above inequality can be rewritten as follows: $\forall u(., \theta_L) \in \Upsilon_1$:

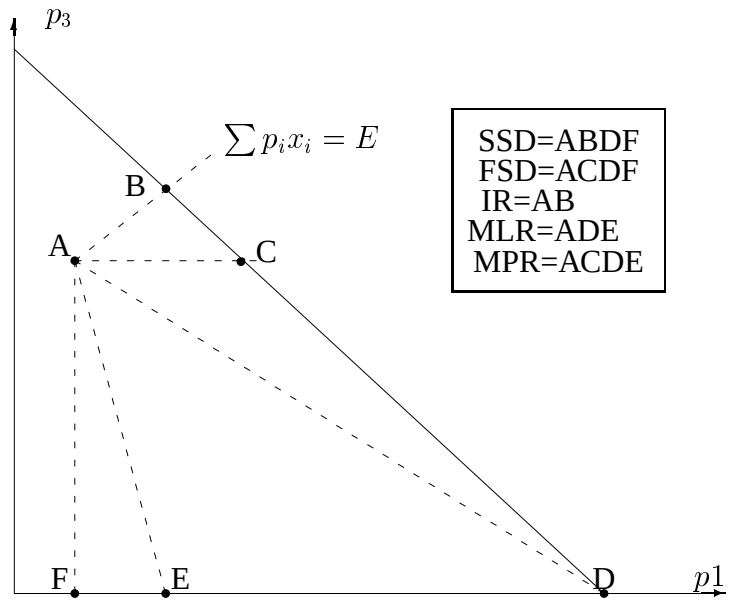

Figure 4.3: The random variables that are dominated by A in the Machina triangle.

$$\int u(x,\theta_L)dG(x) = 0 \Longrightarrow \forall \theta_H > \theta_L : \int u(x,\theta_H)dG(x) \leq 0, \tag{4.9}$$

or, integrating by parts,

$$\int u'(x,\theta_L)G(x)dx = 0 \Longrightarrow \forall \theta_H > \theta_L : \int u'(x,\theta_H)G(x)dx \geq 0. \tag{4.10}$$

By assumption, we know that u' is log-supermodular. Using Proposition 4 directly yields the result.

Proposition 12 *Condition (4.8) holds if and only if F_1 crosses F_2 only once, from above:*

$$\exists x_0 \forall x : (x - x_0)(F_2(x) - F_1(x)) \geq 0.$$

Examples of changes in risk that satisfy the Jewitt's stochastic dominance order associated to Υ_1 are MPS shifts. This is in general not the case for SSD shifts, as F_1 and F_2 may cross more than once.

Jewitt (1989) and Athey (1997) derive a similar condition for the Jewitt's order associated to Υ_2.

4.6 Concluding remark

In this chapter, we showed that a utility function is basically the same object as a random variable. Indeed, using a basis $\{b(.,\theta)\}_{\theta\in\Theta}$ of the set of functions under consideration, to each utility function in the function set generated by the basis, there exists an associated "random variable" $\widetilde{\theta}$ such that

$$u(x) = Eb(x,\widetilde{\theta})\ \forall x.$$

It implies that expected utility can be written as

$$Eu(\widetilde{x}) = Eb(\widetilde{x},\widetilde{\theta}) = \int\int b(x,\theta)dH(\theta)dF(x) = EU(F,H).$$

Our objective here is to show that the notions of change in risk developed in this chapter are linked to the notions of change in preference developed in the previous chapter. For example, consider the set Υ_1 of increasing utility functions, where $b(x,\theta) = I(x \leq \theta)$ and $H(\theta) = u(\theta)$, with the normalization $u(a) = 0$, $u(b) = 1$. In this case, the monotone likelihood ratio order which states that $-F_1''(\theta)/F_1'(\theta) \geq -F_2''(\theta)/F_2'(\theta)$ for all θ is dual to the notion of more risk aversion, which states that $-H_1''(\theta)/H_1'(\theta) \geq -H_2''(\theta)/H_2'(\theta)$. The symmetry is not perfect, however. Whereas it is economically meaningful to address the problem of the effect of a change of F on EU,

$$\forall H : \; EU(F_1, H) \leq EU(F_2, H),$$

there is no sense to determine the effect of a change of H on EU :

$$\forall F : \; EU(F, H_1) \leq EU(F, H_2),$$

or, $Eu_1(\widetilde{x}) \leq Eu_2(\widetilde{x})$. This is because expected utility is ordinal. It does not allow to make inter-personal comparisons of welfare. Otherwise, all results that have been presented in this chapter could have been used to make such comparisons. Just replace F_i by u_i in the case of Υ_1, or by $-u_i'$ in the case of Υ_2! Because expected utility is ordinal, the previous chapter extensively relied on the Diffidence Theorem where no inter-personal comparison of expected utility is made.

Part II

The standard portfolio problem

Chapter 5

The standard portfolio problem

One of the classical problems in the economics of uncertainty is the standard portfolio problem. An investor has to determine the optimal composition of his portfolio containing a risk-free and a risky asset. This is a simplified version of the problem of determining whether to invest in bonds or equities. In fact, the structure of this problem may be applied to several problems in which a risk-averse agent has to choose the optimal exposure to a risk. This is the case for example when an insured person has to determine the optimal coinsurance rate to cover a risk of loss, or when an entrepreneur has to fix his capacity of production without knowing the price at which he will be able to sell the output. The standard portfolio problem has applications in the day-to-day life.

In this chapter, we examine the static version of this model. This model has been analyzed first by Arrow (1964) and Pratt (1964). Several extensions to this basic model will be considered later on.

5.1 The model and its basic properties

Consider an agent with an increasing and concave utility function $u \in \Upsilon_2$. He has a sure wealth W that he can invest in a risk-free asset and in a risky asset. The return of the risk-free asset is r. The return of the risky asset over the period is a random variable $\widetilde{x}_0$. The problem of the agent is to determine the optimal composition $(w_0 - \alpha, \alpha)$ of his portfolio, where $w_0 - \alpha$ is invested in the risk-free asset and α is invested in the risky asset. The value of the portfolio at the end of the period is thus

$$(W-\alpha)(1+r)+\alpha(1+\widetilde{x}_0)=W(1+R)+\alpha(\widetilde{x}_0-r)=w_0+\alpha\widetilde{x}, \tag{5.1}$$

where $w_0 = W(1+R)$ is future wealth and $\widetilde{x} = \widetilde{x}_0 - r$ is the excess return. Let F be the cumulative distribution of random variable $\widetilde{x}$. We do not consider here the existence of short-sale constraints, i.e., α may be larger than W or less than zero. The problem of the investor is to choose α in order to maximize the expected utility $V(\alpha)$:

$$\max_{\alpha}\ V(\alpha) = Eu(w_0+\alpha\widetilde{x}). \tag{5.2}$$

Let us assume that u is differentiable. The first-order condition for this problem is written as

$$V'(\alpha^*) = E\widetilde{x}u'(w_0+\alpha^*\widetilde{x}) = 0, \tag{5.3}$$

where α^* is the optimal demand for the risky asset. Notice that this is a well-behaved maximization problem since the objective function V is a concave function of the decision variable α:

$$V''(\alpha) = E\widetilde{x}^2u''(w_0+\alpha\widetilde{x}) \leq 0.$$

Since V is concave, the sign of $V'(0)$ determines the sign of α^*. But we have

$$V'(0) = u'(w_0)E\widetilde{x}.$$

In consequence, $E\widetilde{x}$ and α^* have the same sign. In addition, it is optimal to invest everything in the risk-free asset if $E\widetilde{x} = 0$. This later result is obvious: buying some of the risky asset generates an increase in risk of the value of the portfolio. When the expected excess return of the risky asset is positive, the optimal portfolio is a best compromise between the expected return and the risk.

Proposition 13 *In the standard portfolio problem with a differentiable utility function, risk-averse agents invest a positive (resp. negative) amount in the risky asset if and only if the expected excess return is positive (resp. negative).*

This is an important and somewhat surprising result. One could have conjectured that if the excess return is positive in expectation but very uncertain, it could be better not to invest in the risky asset. This is not possible: as soon as the expected xcess return is positive, even if very small, it is optimal to purchase some of the risky asset. However, in the real world, a large proportion of the population does not hold any stock, although they have a large expected return.[1]

It is noteworthy that there is an important prerequisite in Proposition 13: the utility function must be differentiable. If it is not, it may be possible that $\alpha^* = 0$ even if $E\widetilde{x} > 0$. There is an intuition for this result. As we have seen in Chapter 3, if the utility function is differentiable, the risk premium tends to zero as the *square* of the size of the risk. This has been refered to as risk aversion being of the second order. In such a case, when considering purchasing a small amount of the risky asset, the expected benefit (return) is of the first order whereas the risk is of the second order. This is not the case anymore if the utility function is not differentiable.

¿From now on, we assume that $E\widetilde{x}$ is positive. If it is not, just replace $\widetilde{x}$ by $-\widetilde{x}$ and α by $-\alpha$.

Observe that the first-order condition (5.3) has no solution if $\widetilde{x}$ does not alternate in sign. In particular, if potential realizations of $\widetilde{x}$ are all larger than zero, V is increasing in α and the optimal solution is unbounded. This is because the risky asset has a return larger than the risk-free asset with probability 1. The risky asset is a money machine in this case. This cannot be an equilibrium. Therefore, we hereafter assume that $\widetilde{x}$ alternates in sign. Notice that this assumption does not guarantee that α^* is bounded. Suppose that the domain of u is R. Then, α^* is bounded only if $\lim_{\alpha\to\infty} V'(\alpha)$ is negative. This condition is rewritten as follows:

$$\left[\lim_{z\to-\infty} u'(z)\right]\int^0 xdF(x) \;+\; \left[\lim_{z\to+\infty} u'(z)\right]\int_0 xdF(x) < 0,$$

or, equivalently,

$$\frac{\lim_{z\to-\infty} u'(z)}{\lim_{z\to+\infty} u'(z)} > \frac{\int_0 xdF(x)}{-\int^0 xdF(x)}. \tag{5.4}$$

If u' tends to zero when wealth tends to plus infinity, or when u' tends to $+\infty$ when wealth tends to minus infinity, the above inequality will automatocally be satisfied.

[1]This can be due to the existence of transaction and learning costs that are inherent to financial markets.

A similar argument can be used when the domain of u is bounded upward or downward.

5.2 The case of a small risk

It can be useful to determine the solution to problem (5.2) when the portfolio risk is small. The problem with this approach is that the size of the risk is endogenous in this problem. The simplest method to escape the difficulty is to define

$$\widetilde{x} = k\mu + \widetilde{y}$$

where $E\widetilde{y} = 0$ and $\mu > 0$. The optimal investment in the risky asset is $\alpha^*(k)$, which is a function of k, with $\alpha^*(0) = 0$. When k is positive, we obtain $\alpha^*(k)$ as the solution of the following equation:

$$E(k\mu + \widetilde{y})u'(w_0 + \alpha^*(k)(k\mu + \widetilde{y})) = 0.$$

Fully differentiating this equation yields

$$[\mu E u'(\widetilde{w}) + \alpha^*(k)\mu E(k\mu + \widetilde{y})u''(\widetilde{w})] + \alpha^{*'}(k)E(k\mu + \widetilde{y})^2 u''(\widetilde{w}) = 0,$$

where $\widetilde{w} = w_0 + \alpha^*(k)(k\mu + \widetilde{y})$. Evaluating this expression at $k = 0$, we obtain

$$\alpha^{*\prime}(0) = \frac{\mu}{E\widetilde{y}^2}\frac{1}{A(w_0)}.$$

A first-order Taylor expansion of $\alpha^*(k)$ around $k = 0$ yields $\alpha^*(k) \simeq \alpha^*(0) + k\alpha^{*\prime}(0)$, or

$$\alpha^* \simeq \frac{E\widetilde{x}}{E(\widetilde{x} - E\widetilde{x})^2}\frac{1}{A(w_0)} \tag{5.5}$$

where α^* is the solution to program (5.2) with $\widetilde{x} = k\mu + \widetilde{y}$. The *relative* share of wealth to invest in the risky asset is thus approximatly equal to

$$\frac{\alpha^*}{w_0} \simeq \frac{E\widetilde{x}}{E(\widetilde{x}-E\widetilde{x})^2}\frac{1}{R(w_0)}$$

The optimal share of wealth to invest in the risky asset is approximately proportional to the ratio of the expectation and variance of the excess return. The coefficient of proportionality is the inverse of relative risk risk aversion. Let us consider a logarithmic investor ($R = 1$). If the excess return has an expectation of 5% and a standard deviation of 30%, the optimal portfolio contains around $5/9$ in the risky asset. This seems to be a rather large proportion. We will come back to this problem when we will discuss the equilibrium price of assets.

It is noteworthy that approximation (5.5) is exact if u is CARA and $\widetilde{x}$ is normally distributed. If

$$u(z) = -\exp(-Az) \text{ and } \widetilde{x} \sim N(\mu, \sigma^2),$$

we obtain

$$\begin{aligned} V(\alpha) &= \tfrac{-1}{\sigma\sqrt{2\pi}} \int \exp(-A(w_0+\alpha x)) \exp(-\tfrac{(x-\mu)^2}{2\sigma^2}) dx \\ &= -\exp(-A(w_0+\alpha\mu - \tfrac{A\alpha^2\sigma^2}{2})) \tfrac{1}{\sigma\sqrt{2\pi}} \int \exp(-\tfrac{(x-(\mu-0.5A\alpha\sigma^2))^2}{2\sigma^2}) dx \\ &= -\exp(-A(w_0+\alpha\mu - \tfrac{A\alpha^2\sigma^2}{2})). \end{aligned}$$

This shows that the Arrow-Pratt approximation is exact in this case. It implies that the optimal α is the one which maximizes $\alpha\mu - \frac{A\alpha^2\sigma^2}{2}$. This yields

$$\alpha^* = \frac{\mu}{\sigma^2}\frac{1}{A}.$$

5.3 The case of HARA functions

We have just seen that the portfolio problem is much simplified if we assume that the utility function of the investor is CARA and if $\widetilde{x}$ is normally distributed. Another type of simplification can be made if we assume that the utility function belongs to the class of HARA functions, i.e., if

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma}.$$

The value of parameters ζ, η and γ are selected in order to have u increasing and concave (see Chapter 2). The first-order condition (5.3) is rewritten in this case as

$$E\widetilde{x}\left(\eta + \frac{w_0 + \alpha^*\widetilde{x}}{\gamma}\right)^{-\gamma} = 0. \tag{5.6}$$

Let us first solve the following problem:

$$E\widetilde{x}\left(1 + \frac{a\widetilde{x}}{\gamma}\right)^{-\gamma} = 0.$$

Notice that a is the optimal investment if $\eta + w_0/\gamma = 1$. It is a function of the distribution of $\widetilde{x}$ and of the coefficient of concavity γ. Now, observe that

$$\alpha^* = a(\eta + w_0/\gamma)$$

is the general solution to equation (5.6). Indeed, we have

$$E\widetilde{x}\left(\eta + \frac{w_0 + a(\eta + w_0/\gamma)\widetilde{x}}{\gamma}\right)^{-\gamma} = (\eta + w_0/\gamma)E\widetilde{x}\left(1 + \frac{a\widetilde{x}}{\gamma}\right)^{-\gamma} = 0.$$

Thus, by solving the first-order condition for a single value of w_0, we obtain the optimal solution for any w_0. When the utility function is HARA, there is a linear relation between the optimal exposure to risk and the wealth level. Two special cases are usually considered. When the utility is CARA ($\gamma \to \infty$), α^* is independent of wealth. When the utility function is CRRA ($\eta = 0$), α^* is proportional to wealth, i.e., the share of wealth invested in the risky asset is a constant independent of wealth.

Judging by empirical evidence, constant relative or absolute risk aversion does not seem to be a reasonable assumption. On the first hand, given CRRA, the optimal share of wealth invested in various assets would be independent of wealth. However, as Kessler and Wolff [1991] show, the portfolios of households with low wealth contain a disproportionately large share of risk free assets. Measuring by wealth, over 80 % of the lowest quintile's portfolio was in liquid assets, whereas the highest quintile held less than 15 % in such assets. This suggests that relative risk aversion is decreasing with wealth. On the other hand, CARA functions are disqualified by the fact that would induce the relative share of wealth invested in risky assets to be *decreasing* in wealth! More generally, this is the case for all HARA utility functions with $\eta > 0$.

5.4 The impact of risk aversion

Under which condition does a change in preferences reduce the optimal exposure to risk, whatever its distribution? To answer this question, let us consider a change of the utility function from u_2 to u_1. By the concavity of the objective function V with respect to α, this change reduces the optimal investment in the risky asset if and only if

$$E\widetilde{x}u_2'(w_0+\widetilde{x})=0 \implies E\widetilde{x}u_1'(w_0+\widetilde{x})\leq 0, \tag{5.7}$$

where we normalized the optimal exposure to risk of u_1 to unity.

To begin with, consider the case where w_0 is arbitrary. It means that we want condition (5.7) to hold for any $\widetilde{x}$ and any w_0. This problem is a direct application of Proposition 4 with $g(x)=x$ and $h(x,i)=u_i'(w_0+x)$. We know that $u_i'(w_0+x)$ is log-supermodular if and only if u_1 is more risk-averse than u_2 in the sense of Arrow-Pratt. Thus, condition (5.7) holds for any $\widetilde{x}$ and w_0 if and only if u_1 is more risk-averse than u_2. This is a very intuitive result: whatever their identical wealth, the more risk-averse of two agents has a smaller demand for the risky asset.

This condition is however too strong if the initial wealth of the two agents is not arbitrary. In this case, the Diffidence Theorem (Corollary 1) must be used to solve problem (5.7). It yields

$$\forall x:\ x\frac{u_1'(w_0+x)}{u_1'(w_0)}\leq x\frac{u_2'(w_0+x)}{u_2'(w_0)} \tag{5.8}$$

as a necessary and sufficient condition. If condition (5.8) holds, we say that u_1 is "centrally more risk averse" than u_2 with respect to w_0. If we normalize $u_2'(w_0)=u_1'(w_0)$, this means that the slope of u_1 is smaller (resp. larger) than the slope of u_2 to the right (resp. to the left) of w_0. This condition is weaker than u_1 being more concave than u_2. However, if it holds for any w_0, u_1 is more risk-averse than u_2.

Proposition 14 *Consider two risk-averse investors with the same initial wealth w_0 and with respectively utility function u_1 and u_2. For any distribution of returns, investor u_1 invests less in the risky asset than investor u_2*

1. *independent of w_0 if and only if u_1 is more risk-averse than u_2, i.e., $\forall z$: $-u_1''(z)/u_1'(z)\geq -u_2''(z)/u_2'(z)$.*

2. *for a given w_0 if and only if u_1 is centrally more risk-averse than u_2 with respect to w_0:* $\forall x: \; xu_1'(w_0+x)/u_1'(w_0) \leq xu_2'(w_0+x)/u_2'(w_0)$.

At this stage, it is important to notice that central risk aversion is a concept that is stronger than diffidence. If an agent u_1 is centrally more risk-averse than agent u_2 with respect to w_0, then agent u_1 is more diffident than agent u_2 with respect to w_0, but the opposite is not true. In other words, condition (5.8) implies condition (3.5). This is proven by integrating condition (5.8) (divided by x) with respect to x. Since the direction of integration cancels out the direction of the inequality when x is positive or negative, we obtain that condition (5.8) implies

$$\int_0^x \frac{u_1'(w_0+\xi)}{u_1'(w_0)} d\xi \leq \int_0^x \frac{u_2'(w_0+\xi)}{u_2'(w_0)} d\xi$$

for all x. But this condition is precisely equivalent to condition (3.5).

Proposition 15 *If agent u_1 is centrally more risk-averse than agent u_2 with respect to w_0, then he is also more diffident than u_2 with respect to w_0.*

The condition that an agent asks always less of the risky asset than another is a more demanding condition than the one that the first rejects more lotteries than the second, when these conditions are for a given wealth level. When this must be true for any wealth level, these two conditions become equivalent. It is the notion of more risk aversion.

A similar exercise can be performed on the effect of a change in w_0 on the optimal structure of the portfolio. This exercise can be seen as an application of Proposition 14 by defining $u_2(z) = u_1(z-k)$. Using result 1 above directly yields the Corollary.

Corollary 3 *A reduction in wealth always reduces the investment in the risky asset if and only if absolute risk aversion is decreasing.*

Similarly, a reduction in wealth always reduces the share of wealth invested in the risky asset if and only if relative risk aversion is decreasing.

5.5 The impact of a change in risk

The dual question to the one examined in the previous section is the following one. Under which condition does a change in the distribution of return increase the demand for the risky asset? To address this question, let us consider two random returns $\widetilde{x}_1$ and $\widetilde{x}_2$ respectively with cumulative distribution function F_1 and F_2. In general terms, the problem is to determine the conditions under which we have:

$$E\widetilde{x}_1 u'(w_0 + \widetilde{x}_1) = 0 \implies E\widetilde{x}_2 u'(w_0 + \widetilde{x}_2) \geq 0. \tag{5.9}$$

To begin with, suppose that $\widetilde{x}_1$ and $\widetilde{x}_2$ are continuous random variables. Then the above condition can be rewritten as:

$$\int xu'(w_0 + x)F_1'(x)dx = 0 \implies \int xu'(w_0 + x)F_2'(x)dx \geq 0.$$

A first result can be obtained by applying Proposition 4 with $g(x) = xu'(w_0 + x)$ and $h(x, i) = F_i'(x)$. Remember that the log-supermodularity of $F_i'(x)$ corresponds to the notion of a monotone likelihood ratio (MLR). The following result is in Milgrom (1981), Landsberger and Meilijson (1990) and Ormiston and Schlee (1993).

Proposition 16 *A MLR-dominant change in the distribution of the risky asset always increases the demand for it.*

Many other sufficient conditions for the comparative statics of a change in the distribution of the return of a risky asset have been proposed in the literature.[2] The necessary and sufficient condition of the change in risk that guarantees that all risk-averse investors increase their demand for the risky asset is obtained as follows. We want condition (5.9) to hold for any $u \in \Upsilon_2$. We know that to any such a utility function u, there is a random variable $\widetilde{\theta}$ such that $u(z) = E\min(z, \widetilde{\theta})$ for all z. It implies that $u'(z) = EI(z \leq \widetilde{\theta})$. Let us define

$$f_i(\theta) = E\widetilde{x}_i I(w_0 + \widetilde{x}_i \leq \theta) = \int^{\theta - w_0} x dF_i(x),\ i = 1, 2.$$

[2] See for example Meyer and Ormiston (1985), Black and Bulkley (1989), Dionne, Eeckhoudt and Gollier (1993) and Gollier and Schlesinger (1996).

It implies that one may rewrite condition (5.9) as

$$Ef_1(\widetilde{\theta}) = 0 \implies Ef_2(\widetilde{\theta}) \geq 0.$$

¿From the Diffidence Theorem, we know that this is true for any $\widetilde{\theta}$, i.e. for any $u \in \Upsilon_2$, if and only if there exists a scalar m such that $f_2(\theta) \geq mf_1(\theta)$ for all θ. This yields the following Proposition.

Proposition 17 *All risk-averse investors increase their demand for the risky asset due to a change in distribution of returns from F_1 to F_2 if and only there exists a scalar m such that for all θ,*

$$\int^{\theta} xdF_2(x) \geq m \int^{\theta} xdF_1(x). \tag{5.10}$$

If this condition is satisfied, one says that $\widetilde{x}_2$ dominates $\widetilde{x}_1$ in the sense of central dominance (CD). Contrary to the intuition, second-order stochastic dominance is not sufficient for central dominance: a change in risk can improve the expected utility of all risk-averse investors although some of them reduce their demand for this risk! This was shown by Rothschild and Stiglitz (1971). As pointed out by Gollier (1995), second-order stochastic dominance is not necessary either: all risk-averse investors can reduce their demand for the risky asset due to a change in distribution of returns although some of them like this change! SSD and CD cannot be compared.

Meyer and Ormiston (1985) found a subset of changes in risk which is in the intersection of SSD and CD. They called it a strong increase in risk (SIR). A SIR is a mean-preserving spread in which all the probability mass that is moved is transfered *outside* the initial support of the distribution. The reader can verify that a SIR is a special case of CD, with condition (5.10) satisfied for $m = 1$. This is also true for MLR changes in risk, with $m = 1$. Eeckhoudt and Gollier (1995) and Athey (1997) have shown that the MPR order is also a special case of CD.

As observed by Rothschild and Stiglitz (1971) and Hadar and Seo (1990), the origin of the problem of SSD not being sufficient is that $\phi(x) = xu'(w_0+x)$ is not in Υ_2, i.e. it is not increasing and concave in x. Otherwise, $\widetilde{x}_1$ being dominated by $\widetilde{x}_2$ in the sense of SSD would have been sufficient to guarantee that $E\widetilde{x}_2u'(w_0+\widetilde{x}_2)$ be larger than $E\widetilde{x}_1u'(w_0 + \widetilde{x}_1)$, which is in turn sufficient for condition (5.9). Observe that

$$\phi'(x) = u'(w_0 + x)\left[1 - R(w_0 + x) + w_0 A(w_0 + x)\right].$$

Thus, assuming that the domain of u is R^+, ϕ is increasing if relative risk aversion R is less than unity. This condition will be sufficient to guarantee that any FSD-dominant change in the distribution of returns will increase the demand for the risky asset. We can also verify that

$$\begin{aligned}\phi''(x) &= u''(w_0 + x)\left[1 - R(w_0 + x) + w_0 A(w_0 + x)\right] \\ &\quad + u'(w_0 + x)\left[-R'(w_0 + x) + w_0 A'(w_0 + x)\right].\end{aligned}$$

The first term is the right-hand side is negative if ϕ' is positive. The second term is negative if relative risk aversion is increasing and absolute risk aversion is decreasing. Notice that ϕ'' may also be written as:

$$\phi''(x) = -u''(w_0 + x)\left[(P^r(w_0 + x) - 2) - w_0 P(w_0 + x)\right]$$

where $P^r(z) = zP(z)$ is relative prudence. This condition implies that ϕ is concave if relative prudence is less than 2, and prudence is positive. We thus conclude this section by the following Proposition.

Proposition 18 *Suppose that the domain of u is R^+. Then, a shift of distribution of returns increases the demand for the risky asset if:*

1. *this shift is FSD-dominant, and if relative risk aversion is less than unity;*
2. *this shift is SSD-dominant shift, relative risk aversion is less than unity and increasing and absolute risk aversion is decreasing;*
3. *this shift is SSD-dominant shift, relative prudence is positive and less than 2.*

The problem with this Proposition is that it is rather unlikely that relative risk aversion is less than unity. Moreover, for HARA utility functions relative prudence is less than 2 if and only if relative risk aversion is less than 1.

5.6 Concluding remark

The results in this chapter are important. The central one is that an increase in risk aversion reduces the demand for the risky asset. It is noteworthy that this condition is necessary: if the change in preference is not an increase in concavity, it is possible to find an initial wealth and a distribution of excess returns such that the investor reduces his demand. This result will frequently be used in the rest of this book. The effect of a change in distribution is much more problematic.

These results are important also because the standard portfolio problem (5.2) has more than one economic interpretation. In addition to the portfolio investement problem, it can be interpreted as the problem of an entrepreneur with a linear production function who has to determine his production before knowing the competitive price at which he will be able to sell it. In that case, $\widetilde{x}$ is the unit profit margin (output price minus marginal cost) and α is the production. This application has first been examined by Sandmo (1971).

Alternatively, consider a risk-averse agent who possesses an asset which is subject to a random loss $\widetilde{y}$ with an expected value equaling μ. Suppose that insurance companies face transaction costs that are proportional to the the actuarial value of the contracts that they sell. Suppose that the agent can select the share α of the risk that he will retain. Then, the actuarial value of the contract is $(1-\alpha)\mu$ and the premium equals $\lambda(1-\alpha)\mu$. The final wealth is $w_0 - \alpha\widetilde{y} - \lambda(1-\alpha)\mu = w_0 - \lambda\mu + \alpha(\lambda\mu - \widetilde{y})$. In consequence, problem (5.2) can also be interpreted as a (proportional) insurance problem where α is the retention rate on the insurable risk. For more details on this application of the standard portfolio problem, see Mossin (1968).

All results in this chapter can be reinterpreted for these two applications. In the next two chapters, we examine decisions problems under uncertainty that cannot be expressed as a standard portfolio problem.

Chapter 6

The equilibrium price of risk

We have shown in the previous chapter how people should manage their portfolio given the current price of the assets. The equilibrium price of the risky asset relative to the price of the risk free asset is determined by the degree of risk aversion of the population of investors and by the degree of risk in the economy. In this chapter, we examine this relationship in the most simplistic model similar to the one proposed by Lucas (1978). Following Mehra and Prescott (1985) and many others afterward, we also calibrate this model. Without much surprise, the model does not fit the data. The surprise, if there is any, comes from how big is the discrepancy between the theoretical equilibrium prices and the observed prices. A large part of this book is devoted to explore different ways of solving this puzzle.

6.1 A simple equilibrium model for financial markets

We assume that the economy is composed of several identical agents who live for one period. Consumption takes place at the end of the period. Agents are identical not only on their preferences but also on their endowment. They have the same utility function u which is increasing and concave. Each agent is endowed with one unit of a production good. This could be called "a firm". Each identical firm produces a certain quantity of a consumption good at the end of the period. Those firms are facing the same exogenous risk that affects their production in the same way. Let $\widetilde{y}$ be the random variable representing the production of each firm. Everyone agrees on the distribution of $\widetilde{y}$. Since productions are perfectly correlated across firms, there is no way to diversify this risk. We see that $\widetilde{y}$ is the

GDP per capita.

There is a simple financial market in this economy. Before the realization of $\widetilde{y}$, agents can sell their property rights on their firm against a certain quantity of the consumption good, which is the numeraire. Since all firms are identical, there is a single price P for 100% of the property rights of a firm. We see that there are two assets on this market. the first one is a stock, or equity, which is a property right on all firm's production. The other is a bond, which guarantee the delivery of one unit of the consumption good. In fact, P can be seen as the relative price of the stock with respect to the bond.

The problem of each agent is to determine the proportion α of property rights on his firm that he will retain, given price P of those rights. If α is larger than 1, this means that the agent is purchasing shares from other firms on the market. The maximization problem is written as

$$\max_{\alpha} \qquad Eu(\alpha\widetilde{y} + (1-\alpha)P). \tag{6.1}$$

This problem is formally equivalent to the one examined in the previous chapter , where $\widetilde{x} = \widetilde{y} - P$ is the net payoff of holding the firm, and $w_0 = P$ is the net wealth of each agent. We know that the solution to this problem is uniquely determined by the its first-order condition:

$$E(\widetilde{y} - P)u'(\alpha\widetilde{y} + (1-\alpha)P) = 0. \tag{6.2}$$

In order to determine the equilibrium price of firms, we need to impose a market-clearing condition on financial markets. The net demand for property rights must be zero. Notice that the optimal demand for them is the same for all agents in this model with identical preferences and identical endowments. It implies that the equilibrium condition is $\alpha = 1$: no one wants to sell his firm at equilibrium. This condition together with condition (6.2) form a system of equations that must satisfied for α and P to characterize a competitive equilibrium. Combining them yields

$$P = \frac{E\widetilde{y}u'(\widetilde{y})}{Eu'(\widetilde{y})}. \tag{6.3}$$

The equity premium ϕ is the expected outperformance of a stockholder over a bondholder. It is a measurement of the bonus that must be given to those agents

that accept to take risk in our society. It is an equilibrium price of risk. The return of the portfolio of the stockholder is given by $(y - P)/P$. The performance of the portfolio of the bondholder is just zero with probability 1. The equity premium is thus equal to

$$\phi = \frac{E\widetilde{y}Eu'(\widetilde{y})}{E\widetilde{y}u'(\widetilde{y})} - 1. \tag{6.4}$$

Using the diffidence method, it can be checked that the equity premium is positive and that it is an increasing function of risk aversion.

6.2 The equity premium puzzle

We see that computing the equity premium is a very simple task in this framework. Remember that $\widetilde{y}$ in equation (6.4) is nothing else than the real Gross Domestic Product per capita. We can believe that agents form their expectation about the future growth of GDP per capita by looking at the past experience. We presented the data for the growth of real GDP per capita in section 3.8. Let us assume that agents believe that each realized growth rate in the past will occur with equal probability. Let us also assume that agents have a constant relative risk aversion γ. Using equation (6.4) allow us to derive the equity premium in that economy. These computations are reported in the following Table.

RRA	Equity premium
$\gamma = 0.5$	0.03%
$\gamma = 1$	0.06%
$\gamma = 2$	0.11%
$\gamma = 4$	0.23%
$\gamma = 10$	0.61%
$\gamma = 40$	2.86%

Table: Equity premium with CRRA and $\widetilde{y}$ based on the actual growth of GDP per capita, USA, 1963-1992.

Notice that we can use a simpler method to calculate the equity premium in this economy, since the risk associated to the annual change of real GDP is so

small. Using Taylor approximations, the equity premium is approximately equal to

$$\phi \sim \frac{E\widetilde{y}^2}{[E\widetilde{y}]^2} R, \tag{6.5}$$

where $R = -[E\widetilde{y}]\,[u''(E\widetilde{y})]\,/u'(E\widetilde{y})$ is the index of relative risk aversion evaluated at the mean. We see that the equity premium is approximately proportional to relative risk aversion and to the variance of the risk. In our data, we have $\frac{E\widetilde{y}^2}{[E\widetilde{y}]^2} = 0.056\%$. Using this rule of thumb yields equity premia with only a very small margin of error with respect to the exact theoretical values.

We see that in the range of acceptable values of relative risk aversion ($\gamma \in [1, 4]$), the equity premium does not exceed one-fourth of a percent. With $\gamma = 2$, the equity premium equals one-tenth of a percent. This means that those who purchase stocks rather than bonds will end up at the end of the year with a wealth that is 0.1% larger on average. This is the compensation at equilibrium for accepting a variable income stream.

Is this prediction about how risks should be priced in our economy compatible with the data? Historical data on asset returns are available several sources. Shiller (1989) and Kocherlakota (1996) provides statistics on asset returns for the U.S. over the period from 1889 to 1978. The average real return to Standard and Poor 500 is 7% per year, whereas the average short-term real risk free rate is 1%. The observed equity premium has thus been equal to 6% over the century. We reproduce the data for the period corresponding to the period examined above in Figures 6.1, 6.2 and 6.3. Over the period 1963-1995, the real return of the S&P500 has been 7.72% per year, which is similar to its historical average. Second, by historical standards, the real interest rate has been high, averaging 3.14%. This is particularly true for the eighties. The bottom line is an equity premium around 4.5%.

Clearly, this simple version of the model proposed by Lucas (1978) does not fit the data. The observed equity premium is just 20 to 50 time larger than its theoretical value. This puzzle has first been stated by Mehra and Prescott (1985). It can be restated by the fact that the existing equity premium is compatible with the model only if relative risk aversion is assumed to be larger than 40. Remember that with such a large risk aversion, agents are ready to pay as much as 29% of their wealth to get rid of the risk of losing or gaining 30% of it with equal probability.

Figure 6.1: The annual real return of Standard & Poor 500.

Figure 6.2: Annual real return to nominally risk free bonds

Figure 6.3: Excess returns of Standard & Poor 500.

There are two ways to interpret this puzzle. To put it simply, either markets do not work as they should, or the model is wrong. If you believe that the first answer is correct, you should invest in stock like crazy, to take advantage of the incredibly large equity premium, given the associated low risk. The other explanation seems more plausible, because a lot of simplifying assumption have been made. Let us list just a few of them:

- Agents bear no other risk than their portfolio risk. There is no risk on human capital.
- Agents cannot internationally diversify their portfolio. There is no other risky investment opportunity than US stocks.
- Agents consume all their wealth at the end of the period. There is no intend to invest for the long term.
- There is no cost for trading stocks and bonds.
- There is just one risky asset. Agents cannot tailor their final consumption plan in a nonlinear way by using options.
- All agents have identical preferences and there is no wealth inequality.
- Agents have constant relative risk aversion.

The question is whether relaxing these assumptions could explain the puzzle. The next chapters can be seen as different attempts of answering this question. A first step in that direction is made in the next section.

6.3 The equity premium with limited participation

It is a well-known fact that only a small proportion of households actually are stockholders. Could this fact explain the equity premium puzzle?

Following Vissing-Jorgensen (1998), let us suppose that for some exogenous reason, some agents are not allowed to hold stocks. In other words, contrary to the assumption made in the basic model, these agents have to sell the property right that they initially owned on their individual firm. This could be due to their lack of knowledge about how to manage a portfolio of stocks, or to their inability to manage their business. The consequence is that they are fully insured against the macroeconomic risk. How the exclusion of this category of agents from financial markets would affect the equilibrium price of property rights on firms? The intuition is that this exclusion will increase the supply of stocks, thereby reducing their equilibrium price. It would yield an increase in their returns. The equity premium would be larger as a consequence.

The decision problem for those who can invest in stock is not different from program (6.1). The only difference in the modelling comes from the market clearing condition. The exogenously given proportion of agents allowed to participate is denoted by k. This is the participation rate on financial markets. Since all participants are alike, they will all hold the same number α of property rights. The equilibrium condition is that $k\alpha = 1$, or $\alpha = 1/k$. Since this α is larger than 1, all investors will go short on the risk free asset to purchase stocks. Using the first-order condition (6.2) with this α yields the following condition on the equilibrium price $P(k)$:

$$E(\widetilde{y} - P)u'(\frac{\widetilde{y}}{k} + (\frac{k-1}{k})P) = 0. \tag{6.6}$$

Notice that the equilibrium price $P(k)$ cannot in general be analytically extracted from this condition, contrary to the previous case where α equaled 1. Numerical methods must be used to get $P(k)$ from which the equity premium $\phi(k) = (E\widetilde{y}/P(k)) - 1$ can be obtained. We have done this exercise by using a power utility function with $\gamma = 2$, together with the data of the growth of real

Figure 6.4: The equity premium with limited participation (with $\gamma = 2$).

GDP per capita for the period 1963-92. The results are drawn in Figure 6.4. We see that the participation rate has no sizeable effect on the equity premium as long as it exceed, say, 30%. If it is less than that, the effect becomes big enough to explain the puzzle. The macroeconomic risk is borne by such a small fringe of the population that a large bonus must be given to them as a compensation.

Part III

Multiple risks

Chapter 7

Risk aversion with background risk

Up to now, we assumed that the decision maker faces a single source of risk. In the real world, people face many risks simultaneously. This raises the question of whether the optimal decision to one problem is independent of the existence of other risks in the agent's portfolio. This is generally not true, even if risks are independent. As an illustration, we may question the idea that all individual risks are traded on financial markets. Suppose in contrast that houeholds bear an uninsurable risk on their wealth. It cannot be sold on markets. The best examples are risks associated to human capital. Those risks are reputed not to be insurable for obvious problems related to asymmetric information and incentives. In such a situation, the portfolio problem is more complex, since the decision makers must now determine their attitude toward one risk — the portfolio risk — in the presence of another risk, i.e. the risk on human capital. In this chapter, we examine whether this another risk does affect the qualitative results obtained earlier in this book. In the next chapter, we quantify its effect.

The exogenous risk, $\widetilde{x}$, is called the "background risk". It cannot be insured. The portfolio problem is now rewritten as

$$\max_{\alpha} \; Eu(w_0 + \alpha\widetilde{y} + \widetilde{x}) \tag{7.1}$$

where $\widetilde{y}$ is the return of the risky asset, which is assumed to be independent of background risk $\widetilde{x}$. Observe that the expectation operator is on $\widetilde{x}$ and $\widetilde{y}$: the decision on α must be taken prior to the realization of the two sources of risk.

How does the presence of $\widetilde{x}$ affect the properties of the optimal demand for the risky asset? For example, is decreasing absolute risk aversion still sufficient to

guarantee that an increase in w_0 reduces the optimal demand for the risky asset? Does a more risk-averse agent always demand less of the risky asset? The standard method to solve these questions is to refer to the notion of the indirect utility function, as introduced by Kihlstrom, Romer and Williams (1981). In fact, the analysis of the effect of an independent background risk on an optimal decision under risk can be transformed into a problem related to a change in preference by defining the following indirect utility function:

$$v(z) = Eu(z + \widetilde{x}). \tag{7.2}$$

The decision problem (7.1) is then equivalent to

$$\max_{\alpha} \; Ev(w_0 + \alpha\widetilde{y}). \tag{7.3}$$

The optimal decision of agent u with background risk $\widetilde{x}$ is the same as for agent v who faces no background risk. Thus the question of the effect of $\widetilde{x}$ is the same as the effect of a change in preference from u to v. The advantage of this method is that we know very well how a change in the shape of the utility function affects the optimal decision under uncertainty.

7.1 Some basic properties

We start with a simple first remark on the properties that the indirect utility function defined by (7.2) inherits from the utility function. It comes from the observation that

$$v^{(n)}(z) = Eu^{(n)}(z + \widetilde{x})$$

where $u^{(n)}$ is the nth derivative of u. In consequence, if $u^{(n)}$ does not change sign, it transfers its sign to $v^{(n)}$. Thus, risk aversion and prudence for example are preserved by background risk.

7.1.1 Preservation of DARA

A less obvious question is whether the indirect utility function inherits decreasing absolute risk aversion from u. Let A_v and P_v denote respectively the coefficient of absolute risk aversion and absolute prudence of function v. We have that

$$A_v(z) = \frac{-v''(z)}{v'(z)} = \frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} \quad \text{and} \quad P_v(z) = \frac{-v'''(z)}{v''(z)} = \frac{-Eu'''(z+\widetilde{x})}{Eu''(z+\widetilde{x})}.$$

It is important to notice that A_v and P_v are not just the expected value of $A(z+\widetilde{x})$ and $P(z+\widetilde{x})$. Thus, it is not a priori obvious that P_v is uniformly larger than A_v if P is uniformly larger than A. This is however true, as stated in the following more general Proposition.

Proposition 19 *The indirect utility function v defined by condition (7.2) inherits the property that P is uniformly larger than kA:*

$$\frac{-u'''(z)}{u''(z)} \geq k\frac{-u''(z)}{u'(z)} \ \ \forall z \Longrightarrow \ \frac{-v'''(z)}{v''(z)} \geq k\frac{-v''(z)}{v'(z)} \ \ \forall z.$$

Proof: This is another application of the Diffidence Theorem. Condition $\frac{-v'''(z)}{v''(z)} \geq k\frac{-v''(z)}{v'(z)}$ may be rewritten as follows:

$$E\left[u''(z+\widetilde{x}) + \lambda u'(z+\widetilde{x})\right] = 0 \quad \Longrightarrow \quad E\left[u'''(z+\widetilde{x}) + k\lambda u''(z+\widetilde{x})\right] \geq 0$$

Using the Diffidence Theorem, we know that this condition holds for any z and tx if and only if there exists $m = m(\lambda, k, z)$ such that

$$u'''(z+x) + k\lambda u''(z+x) \geq m\left[u''(z+x) + \lambda u'(z+x)\right] \tag{7.4}$$

for all x.

By risk aversion, condition $P \geq kA$ implies that

$$\begin{aligned} u'''(z+x) + k\lambda u''(z+x) &= -A(z+x)u'(z+x)\left[k\lambda - P(z+x)\right] \\ &\geq -A(z+x)u'(z+x)k\left[\lambda - A(z+x)\right]. \end{aligned} \tag{7.5}$$

We are looking for a m such that

$$-A(z+x)u'(z+x)k\left[\lambda - A(z+x)\right] \geq m\left[u''(z+x) + \lambda u'(z+x)\right] \tag{7.6}$$

Combining conditions (7.5) and (7.6) would yield the necessary and sufficient condition (7.4). Taking $m = -k\lambda$ is a good candidate, since condition (7.6) is then equivalent to

$$u'(z+x)k\left[\lambda - A(z+x)\right]^2 \geq 0,$$

which is true for all $k \geq 0$. ■

It is noteworthy that the opposite property is not true: the value function does not in general inherit property $P \leq kA$ from u. A first illustration of the above result is the preservation of decreasing absolute risk aversion. As a consequence, by using Proposition 14, if u exhibits DARA, the optimal solution of program (7.1) is increasing in w_0. But it is possible to build an example of a utility function with increasing absolute risk aversion with an optimal exposure to risk that is locally increasing in w_0.

The preservation of DARA by the expectation operator is a special case of a property found by Pratt (1964) that a sum of DARA utility functions is DARA. Again, the opposite is not true: the sum of two functions that exhibit increasing absolute risk aversion may exhibit decreasing absolute risk aversion locally. Notice also that the sum of two CARA functions is DARA.

A generalization of Proposition 19 is as follows: suppose that $u^{(n-1)}$ does not alternate in sign and that $-u^{(n)}/u^{(n-1)}$ is uniformly positive. Then, using the same method to prove the Proposition for $n = 2$, we obtain that

$$\frac{-u^{(n+1)}(z)}{u^{(n)}(z)} \geq k\frac{-u^{(n)}(z)}{u^{(n-1)}(z)} \ \forall z \Longrightarrow \frac{-v^{(n+1)}(z)}{v^{(n)}(z)} \geq k\frac{-v^{(n)}(z)}{v^{(n-1)}(z)} \ \forall z.$$

An application of this result is that decreasing absolute prudence is inherited by the indirect utility function. This result has first been stated by Kimball (1993).

7.2 The comparative risk aversion is not preserved

The positive results obtained in Proposition 19 could induce us to believe that all the characteristics of u are transferred to v. This is not true. For example, we don't know whether the convexity of absolute risk tolerance is preserved. Another example is the following one, which is due to Kihlstrom, Romer and Williams (1983). Consider two agents respectively with utility u_1 and u_2. Suppose that u_1

is more risk-averse than u_2 in the sense of Arrow-Pratt, i.e. $A_1(z)$ is larger than $A_2(z)$ for all z. Does it imply that agent u_1 will behave in a more risk-averse way than agent u_2 if they both face the same background risk $\widetilde{x}$? Technically, the question is equivalent to whether v_1 is more concave than v_2 in the sense of Arrow-Pratt, where v_i is defined as

$$v_i(z) = Eu_i(z + \widetilde{x}).$$

Remember that v_1 is more risk-averse than v_2 if and only if $H(z, i) = v_i'(z)$ is log-supermodular in (z, i). Observe also that

$$H(z, i) = Eh(z, \widetilde{x}, i)$$

where $h(z, x, i) = u_i'(z + x)$. Finally, remember that, by Proposition 5, log-supermodularity is preserved by the expectation operator. The problem is that the use of this Proposition for our purpose here requires h to be LSPM. Function h is LSPM if indeed u_1 is more risk-averse than u_2 *and* u_i is DARA. An application of Proposition 5 is thus that u_1 will behave in a more risk-averse way than u_2 in the presence of background risk $\widetilde{x}$ if A_1 is uniformly larger than A_2 and A_i is decreasing. Using Lemma 3 rather than the Proposition allows for a slight improvement of this result.

Proposition 20 *Comparative risk aversion is preserved if one of the two utility functions exhibits nonincreasing absolute risk aversion.*

Technically, we mean the following: suppose that $A_1(z) = -u_1''(z)/u_1'(z)$ is larger than $A_2(z) = -u_2''(z)/u_2'(z)$ for all z, and that A_1 or A_2 is nonincreasing in z. Then, agent u_1 will behave in a more risk-averse way than u_2 in the presence of background risk, i.e. $-v_1''(z)/v_1'(z)$ is larger than $-v_2''(z)/v_2'(z)$ for all z. Kihlstrom, Romer and Williams (1983) provide a counter-example when absolute risk aversion is increasing for both utility functions. Although decreasing absolute risk aversion is a simple and natural assumption, it is stronger than necessary. Pratt (1988) obtained the necessary and sufficient condition for the preservation of the comparative risk aversion. The idea is to replace condition

$$\forall \widetilde{x} : \ \frac{-Eu_1''(z + \widetilde{x})}{Eu_1'(z + \widetilde{x})} \geq \frac{-Eu_2''(z + \widetilde{x})}{Eu_2'(z + \widetilde{x})} \tag{7.7}$$

by the following equivalent one:

$$\forall \widetilde{x}, k: \; Eu_1'(z+\widetilde{x}) - kEu_2'(z+\widetilde{x}) = 0 \implies -Eu_1''(z+\widetilde{x}) + kEu_2''(z+\widetilde{x}) \geq 0. \tag{7.8}$$

Then, the Diffidence Theorem would directly yield the result. We prefer using Lemma 1 that is at the origin of this Theorem. It happens that it is easier in this case not to make the additional algebraic manipulations that lead to the Theorem. What Lemma 1 says is that the two equivalent conditions (7.7) and (7.8) hold for any random variable if they hold for any binary random variable $\widetilde{x} \sim (x_1, p; x_2, 1-p)$. The necessary and sufficient condition is thus written as: $\forall x_1, x_2, \forall p \in [0,1]$:

$$-\frac{pu_1''(z+x_1) + (1-p)u_1''(z+x_2)}{pu_1'(z+x_1) + (1-p)u_1'(z+x_2)} \geq -\frac{pu_2''(z+x_1) + (1-p)u_2''(z+x_2)}{pu_2'(z+x_1) + (1-p)u_2'(z+x_2)}. \tag{7.9}$$

This condition is not easy to manipulate. It is satisfied by a transformation of preference that is not covered by Proposition 20.Suppose that $u_1 \equiv \lambda u_2 + g$, with λ positive and g being decreasing and concave. In this case, we say that u_1 is more risk-averse than u_2 in the sense of Ross (1981). Then, inequality (7.7) is rewritten as

$$-\frac{\lambda Eu_2''(z+\widetilde{x}) + Eg''(z+\widetilde{x})}{\lambda Eu_2'(z+\widetilde{x}) + Eg'(z+\widetilde{x})} \geq \frac{-Eu_2''(z+\widetilde{x})}{Eu_2'(z+\widetilde{x})}$$

which is equivalent to

$$-\frac{Eg''(z+\widetilde{x})}{Eg'(z+\widetilde{x})} \leq \frac{-Eu_2''(z+\widetilde{x})}{Eu_2'(z+\widetilde{x})}.$$

This condition is automatically satisfied, since the left-hand side of this inequality is negative whereas the right-hand side is positive.

Proposition 21 *Suppose that u_1 is more risk-averse than u_2 in the sense of Ross, i.e. that there exist a positive scalar λ and a decreasing and concave function g such that $u_1(z) = \lambda u_2(z) + g(z)$ for all z. Then, agent u_1 behaves in a more risk-averse way than agent u_2 in the presence of background risk.*

The notion of more risk aversion in the sense of Ross is very strong. It is in particular stronger than the notion of Arrow-Pratt. But this new concept will play an important role in the next section.

7.3 Dependent background risk

The important advantage of the independence assumption is that it allows us to use the technique of the indirect value function. The derivation of the characteristics of the indirect value function generates in one single shot a full description of the effect of background risk for a wide set of decision problems under uncertainty. This is not possible if we allow the background risk to be correlated with the endogenous risk. In the analysis that we provide now for a dependent background risk, we consider the simplest decision problem under uncertainty. Namely, we examine a take-it-or-leave-it offer to gamble, the solution of which depends directly upon the risk premium of the corresponding lottery.

Let us define the premium $\pi_i(\widetilde{w}_1 \to \widetilde{w}_2)$ as the price that agent u_i is ready to pay to replace lottery $\widetilde{w}_1$ by lottery $\widetilde{w}_2$. It is obtained by:

$$Eu_i(\widetilde{w}_1) = Eu_i(\widetilde{w}_2 - \pi_i). \tag{7.10}$$

If

$$\widetilde{w}_1 \equiv w_0 + \widetilde{y} + \widetilde{x} \text{ and } \widetilde{w}_2 \equiv w_0 + \widetilde{x},$$

price π_i is the risk premium that u_i is ready to pay for the elimination of noise $\widetilde{y}$ to the initial risk $\widetilde{x} + \widetilde{y}$. If $E\widetilde{y} = 0$ and $\widetilde{x}$ is degenerated at zero, π_i is the risk premium analyzed in section 3.3. In the previous section, we obtained conditions for π_1 to be larger than π_2 when the two sources of risk are independent.

Ross (1981) and Jewitt (1986) examined the case where $\widetilde{w}_1$ and $\widetilde{w}_2$ are stochastically dependent. The natural extension is to assume that $\widetilde{w}_1$ is an increase in risk with respect to $\widetilde{w}_2$, or equivalently that $E\left[\widetilde{y} \mid \widetilde{x}\right] = 0$. In this case, π_i is the premium that agent u_i is ready to pay for the reduction of risk from $\widetilde{w}_1$ to $\widetilde{w}_2$, or equivalently, for the elimination for the correlated noise $\widetilde{y}$. It is intuitive that a more risk-averse agent will be ready to pay more for a given reduction of risk. Because this condition is stronger than when we are in the particular case of an uncorrelated noise, this condition will require restrictions on preferences that are stronger than condition (7.9).

We want to guarantee that $\pi_1 = \pi_1(\widetilde{w}_1 \to \widetilde{w}_2)$ be larger than $\pi_2 = \pi_2(\widetilde{w}_1 \to \widetilde{w}_2)$ for all pair $(\widetilde{w}_1, \widetilde{w}_2)$ such that $\widetilde{w}_1$ is an increase in risk of $\widetilde{w}_2$. This question is in fact dual to the concept of the Jewitt's stochastic dominance order that we described in section 4.5. This would be true if $Eu_1(\widetilde{w}_1) \leq Eu_1(\widetilde{w}_2 - \pi_2)$ where π_2 is defined by equation (7.10) for $i = 2$. Notice that a sufficient condition is that u_1 be more risk-averse than u_2 in the sense of Ross. Indeed, if $u_1 \equiv \lambda u_2 + g$, with $\lambda > 0$ and $g', g'' \leq 0$, we have that

$$\begin{aligned} Eu_1(\widetilde{w}_1) &= \lambda Eu_2(\widetilde{w}_1) + Eg(\widetilde{w}_1) \\ &= \lambda Eu_2(\widetilde{w}_2 - \pi_2) + Eg(\widetilde{w}_1) \\ &\leq \lambda Eu_2(\widetilde{w}_2 - \pi_2) + Eg(\widetilde{w}_2) \\ &\leq \lambda Eu_2(\widetilde{w}_2 - \pi_2) + Eg(\widetilde{w}_2 - \pi_2) \\ &= Eu_1(\widetilde{w}_2 - \pi_2). \end{aligned}$$

The first equality is by assumption. The second equality is due to the definition of π_2. The first inequality is due to the concavity of g together with the fact that $\widetilde{w}_1$ is an increase of risk with respect to $\widetilde{w}_2$. The second inequality is obtained by observing that π_2 is positive and g is decreasing.

The more difficult step is to prove that the Ross' notion of more risk aversion is necessary. To do this, let us assume that $\widetilde{w}_2$ is distributed as $(\widetilde{z}, 1-p; z_0, p)$ and that the added noise to produce the riskier $\widetilde{w}_1$ is degenerated to zero for any realization of $\widetilde{z}$, and is equal to $k\widetilde{\epsilon}$ if $\widetilde{w}_2$ takes value z_0. Define $\Pi_i(k)$= $\pi_i(\widetilde{w}_1 \to \widetilde{w}_2)$ as a function of the size k of the noise. Obviously, $\Pi_i(0) = 0$ and, by first-order risk aversion, $\Pi_i'(0) = 0$. Thus, a necessary condition is that $\Pi_1''(0)$ be larger than $\Pi_2''(0)$. This condition is written as

$$\frac{-pu_1''(z_0)}{Eu_1'(\widetilde{w})} \geq \frac{-pu_2''(z_0)}{Eu_2'(\widetilde{w})},$$

or

$$E\frac{u_2'(\widetilde{w})}{-u_2''(z_0)} \geq E\frac{u_1'(\widetilde{w})}{-u_1''(z_0)}.$$

This condition holds for any distribution of $\widetilde{w}$ if and only if it holds for any degenerate distribution, that is, if and only if

$$\frac{u_2'(w)}{-u_2''(z_0)} \geq \frac{u_1'(w)}{-u_1''(z_0)}.$$

This condition must hold for any pair (w, z). Thus, there must exist a scalar λ such that

$$\forall z_0, w: \quad \frac{u_1''(z_0)}{u_2''(z_0)} \geq \lambda \geq \frac{u_1'(w)}{u_2'(w)}.$$

It is easily verified that this is equivalent to the condition that $u_1 \equiv \lambda u_2 + g$, with g being decreasing and concave. This proves the following Proposition, which is due to Ross (1981).

Proposition 22 *The following two conditions are equivalent:*

1. *Agent u_1 is ready to pay more than agent u_2 for any Rothschild-Stiglitz reduction is risk.*

2. *Agent u_1 is more risk-averse than agent u_2 in the sense of Ross.*

This result is a first example of the difficulty to extend existing results on the effect of an introduction/elimination of a pure risk to the effect of a marginal increase/reduction in risk. In this case, the fact that an agent is always willing to pay more than another for the elimination of a risk does not necessarily imply that he is also willing to pay more for any marginal reduction of risk. In other words, even under DARA, a more risk-averse agent (in the sense of Arrow-Pratt) may be more reluctant to pay for a risk-reducing investment.

7.4 Conclusion

The presence of an uninsurable background risk in the wealth of a decision maker affects his behaviour towards other independent sorces of risk. But under the standard assumptions that his utility function is concave and exhibits DARA, it is still true that he will reject all unfair lotteries, that he will increase his saving if future risks are increased, and that he will take more risk if his sure wealth is increased. Under DARA, an increase in risk aversion, i.e., concavifying the utility function, always increases the risk premium that the agent is ready to pay to escape an independent risk. However, this result does not hold if we consider the risk premium that the agent is ready to pay for a marginal reduction of risk, i.e., for the elimination of a correlated noise. The Ross's concept of an increase in risk aversion is the relevant concept to solve this last problem.

Chapter 8

The tempering effect of background risk

In the previous chapter, we were interested in whether the basic properties of the utility function were transferred to the indirect utility function in the presence of an exogenous background risk. More recent developments in this field changed the focus to the impact of background risk on the degree of risk aversion. Does the presence of an exogenous background risk reduce or increase the demand for other independent risks? In other words, are independent risks substitutes or complementary?

This is an old question that was first raised first by Samuelson (1963):

> I offered some lunch colleagues to bet each $200 to $100 that the side of a coin they specified would not appear at the first toss. One distinguished scholar (...) gave the following answer: ”I won’t bet because I would feel the $100 loss more than the $200 gain. But I’ll take you on if you promise to let me make 100 such bets”.

This story suggests that independent risks are complementary. However, Samuelson went ahead and asked why it would be optimal to accept 100 separately undesirable bets. The scholar answered:

> ”One toss is not enough to make it reasonably sure that the law of averages will turn out in my favor. But in a hundred tosses of a coin, the law of large numbers will make it a darn good bet.”

Obviously, this scholar misinterprets the Law of Large Numbers! It is not by accepting a second independent lottery that one reduces the risk associated to a first one. If $\widetilde{x}_1, \widetilde{x}_2, ..., \widetilde{x}_n$ are independent and identically distributed, $\widetilde{x}_1 + \widetilde{x}_2 + ... + \widetilde{x}_n$ has a variance n time as large as each of these risks. What is stated by the Law of Large Numbers is that $\frac{1}{n}\sum_{i=1}^{n} \widetilde{x}_i$ — not $\sum_{i=1}^{n} \widetilde{x}_i$ — tends to $E\widetilde{x}_1$ almost surely as n tends to infinity. It is by subdividing — not by adding — risks that they are washed away by diversification, as shown in section 4.3. An insurance company does not reduce its aggregate risk by accepting more independent risks in its portfolio.

Once this fallacious interpretation of the Law of Large Numbers is explained and understood, the intuition suggests that independent risks should rather be substitute. The presence of one risk should have an adverse effect on the demand for another independent risk. In this Chapter, we examine the conditions under which the presence of an *exogenous* risk increases the aversion to other independent risks. This branch of the literature is referred to as the "background risk" problem. In the next Chapter, we will turn to the effect of the presence of an *endogenous* risk on the aversion to other independent risks.

8.1 Risk vulnerability

In this section, we examine the impact of a background risk $\widetilde{x}$ that has a nonpositive mean. The intuition is strong to suggest that such a background risk must increase the aversion to other independent risks. It should induce the agent to reject more lotteries, and to reduce his demand for any risky asset that has an independent return, whatever his initial wealth.

Definition 5 *Preferences exhibit "risk vulnerability" if the presence of an exogenous background risk with a nonpositive mean raises the aversion to any other independent risk.*

In spite of its intuitive content, all expected utility preferences are not necessarily risk vulnerable, as shown in the following counter-example. Consider a risk-averse individual with utility $u(w)$ which equals w if w is less than 100, otherwise it equals $50 + 0.5w$. If his initial wealth is 101, he dislikes a lottery $\tilde{x}$ offering -11 and $+14$ with equal probabilities, i.e. $\widetilde{y} = (-11, 1/2; 14, 1/2)$. But if one adds the mean-zero independent "background" risk $\widetilde{x} = (-20, +20; 1/2, 1/2)$ to his initial wealth, the individual would find $\widetilde{y}$ to be desirable. It is a paradox that an

undesirable lottery can be made desirable by the presence of another (mean-zero) risk. Another illustration of this point is obtained by considering the standard portfolio problem. The risk-free rate is zero, whereas the return of the risky asset equals -100% or $+190\%$, with equal probabilities. With a sure background wealth of 101, the agent would optimally purchase one unit of the risky asset. The same individual would purchase 10.5263 units of the risky asset in the presence of the undesirable risk background risk $\widetilde{x}$ in addition to his 101 units of wealth!

¿From Propositions 7 and 14, intuition is true if and only if the indirect utility function defined by condition (7.2) is more concave than the original utility function u, or if

$$E\widetilde{x} \leq 0 \quad \Longrightarrow \quad \forall z : \frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} \geq \frac{-u''(z)}{u'(z)}, \tag{8.1}$$

or, equivalently,

$$E\widetilde{x} \leq 0 \quad \Longrightarrow \quad \forall z : EA(z+\widetilde{x})u'(z+\widetilde{x}) \geq A(z)Eu'(z+\widetilde{x}) \tag{8.2}$$

This condition is the technical condition for risk vulnerability, a concept introduced by Gollier and Pratt (1996): risk aversion is vulnerable to universally undesirable risks.

The problem is decomposed in two parts: we need first that any sure reduction in wealth raises risk aversion. Second, we also need that any zero-mean background risk raises risk aversion. The combination of these two conditions is necessary and sufficient for risk vulnerability, since any risk with a non-positive mean can be decomposed into the sum of a degenerated random variable that takes a non-positive value with probability 1, and a pure risk. The easy step is the first condition, which is trivially equivalent to decreasing absolute risk aversion. Thus, DARA is necessary for risk vulnerability.

We now present two sufficient conditions for risk vulnerability. The first one is that absolute risk aversion $A(z) = -u''(z)/u'(z)$ be decreasing and convex. Decreasing absolute risk aversion, combined with Proposition 3, implies that

$$EA(z+\widetilde{x})u'(z+\widetilde{x}) \geq EA(z+\widetilde{x})Eu'(z+\widetilde{x}). \tag{8.3}$$

If A is convex, Jensen's inequality implies that

$$EA(z+\widetilde{x}) \geq A(z+E\widetilde{x}) \geq A(z), \tag{8.4}$$

the second inequality being due to DARA combined with $E\widetilde{x} \leq 0$. The combination of the inequalities (8.3) and (8.4) implies condition (8.2), which defines risk vulnerability. The assumption of a convex absolute risk aversion is a natural assumption. It means in particular that the reduction of risk premium due to an increase in wealth is a decreasing function of wealth, at least for small risks (use the Arrow-Pratt approximation). Moreover, absolute risk aversion cannot be positive, decreasing and concave everywhere.

Another sufficient condition for risk vulnerability is standardness. The concept of standardness has been introduced by Kimball (1993) to refer to the condition that absolute risk aversion and absolute prudence are decreasing with wealth. To prove the sufficiency of standardness for risk vulnerability, let us define the precautionary premium ψ associated to risk $\widetilde{x}$ at wealth z as follows:

$$Eu'(z+\widetilde{x}) = u'(z+E\widetilde{x}-\psi(z,u,\widetilde{x})).$$

The analogy with the risk premium is obvious: ψ is the risk premium of $\widetilde{x}$ for the agent with utility function $-u'$. In the terminology of Chapter 3, ψ is equal to $\pi(z,-u',\widetilde{x})$. It implies that

$$Eu''(z+\widetilde{x}) = (1-\frac{\partial\psi}{\partial z}(z,u,\widetilde{x}))u''(z+E\widetilde{x}-\psi(z,u,\widetilde{x}))$$

and

$$\frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} = (1-\frac{\partial\psi}{\partial z}(z,u,\widetilde{x}))A(z+E\widetilde{x}-\psi(z,u,\widetilde{x})).$$

By Proposition 9, we know that ψ is decreasing in z because the decreasing absolute risk aversion of $-u'$ is equivalent to the decreasing absolute prudence of u. It implying that

$$\frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} \geq A(z+E\widetilde{x}-\psi(z,u,\widetilde{x})).$$

Now, remember also that prudence is necessary for DARA. Thus the assumption that $E\widetilde{x} \leq 0$ implies that $E\widetilde{x} + (-\psi(z, u, \widetilde{x}))$ is negative, because it is the sum of two negative terms. DARA concludes the proof that standardness is sufficient for risk vulnerability.

These two sufficient conditions are not equivalent. Indeed, decreasing absolute prudence is itself equivalent to

$$A''(z) \geq -A'(z)\left[P(z) - 2A(z)\right]. \tag{8.5}$$

Under DARA, this condition shows that, when $P \geq 2A$, decreasing absolute prudence implies the convexity of absolute risk aversion. But nothing can be said when P is smaller than $2A$.

Proposition 23 *The presence of a non-positive background risk raises the aversion to other independent risks, i.e. risk aversion is risk vulnerable, if at least one of the following two conditions is satisfied:*

1. *Absolute risk aversion is decreasing and convex;*

2. *Absolute risk aversion and absolute prudence are decreasing.*

Notice that the subset of HARA utility functions defined by (22.5) that are DARA ($\gamma \geq 0$) satisfies these two conditions.

The necessary and sufficient condition is obtained by using the Diffidence Theorem on condition (8.2). It yields

$$\forall z, x: \quad \left[A(z+x) - A(z)\right] u'(z+x) \geq xA'(z)u'(z). \tag{8.6}$$

This condition is not easy to manipulate. In addition, contrary to the other applications of the diffidence Theorem presented earlier in this book, it does not simplify to the unidimensional condition (similar to $A > 0$, $P > 2A$ for example). Here, we must verify that a function of *two* variables is positive everywhere. A necessary condition is derived from the local necessary condition (2.10). It is written here as

$$\forall z: \quad A''(z)u'(z) + 2A'(z)u''(z) \geq 0,$$

or,

$$\forall z: \quad A''(z) \geq 2A'(z)A(z). \tag{8.7}$$

This means that absolute risk aversion may not be too concave. Careful differentiation shows that this condition can be rewritten as

$$\forall z: \quad \frac{-u''''(z)}{u'''(z)} \geq \frac{-u''(z)}{u'(z)}$$

To sum up, the necessary and sufficient condition for risk vulnerability is rather complex. We have two necessary conditions, and two sufficient conditions that are much easier to check. The two necessary conditions are DARA and condition (8.7). The two sufficient conditions are stated in Proposition 23.

8.2 Risk vulnerability and increase in risk

8.2.1 Increase in background risk

Risk vulnerability guarantees that the introduction of a pure risk increases the aversion to other independent risks. Does it also guarantee the same result for any marginal second order stochastically dominated change in background risk? In other terms, suppose that $\widetilde{y}_2$ dominates $\widetilde{y}_1$ in the sense of SSD. Does it imply that $v_1(z) = Eu(z + \widetilde{y}_1)$ is more concave than $v_2(z) = Eu(z + \widetilde{y}_2)$, or

$$\frac{-Eu''(z + \widetilde{y}_1)}{Eu'(z + \widetilde{y}_1)} \geq \frac{-Eu''(z + \widetilde{y}_2)}{Eu'(z + \widetilde{y}_2)}$$

for all z? Because $\widetilde{y}_2$ dominates $\widetilde{y}_1$ in the sense of SSD, we know that $\widetilde{y}_1$ is distributed as $\widetilde{y}_2 + \widetilde{x}$, with $E\left[\widetilde{x} \mid \widetilde{y}_2\right] = 0$. Let us first suppose that noise $\widetilde{x}$ is independent of $\widetilde{y}_2$. In this case, the problem simplifies to whether $\widehat{v}(z) = Ev(z+\widetilde{y}_2)$ is more concave than $\widehat{u}(z) = Eu(z+\widetilde{y}_2)$. Observe that risk vulnerability implies that v is more concave than u. Does it imply that $\widehat{v}$ is more concave than $\widehat{u}$? Or, does an exogenous background risk preserve the comparative risk aversion order? This is precisely the question that we raised in section 7.2. We showed in Proposition 20 that DARA of one of the two utility functions is sufficient for the preservation

of the comparative risk aversion order. Since DARA is necessary for risk vulnerability anyway, this proves that adding an independent noise to background risk raises the aversion to other independent risks under risk vulnerability. The problem is less obvious in the case of a correlated noise. A full characterization of the solution to this problem is in Eeckhoudt, Gollier and Schlesinger (1996).

8.2.2 Increase in the endogenous risk

Risk vulnerability brings another view on the effect of an increase in risk of the return of a risky asset on its demand, something that we already examined in section 5.5. In the standard portfolio problem without background risk, suppose that the return of the risky asset undergoes an increase in risk from $\widetilde{x}_2$ to $\widetilde{x}_1 = \widetilde{x}_2 + \widetilde{x}$, with $E\left[\widetilde{x} \mid \widetilde{x}_2\right] = 0$. We have shown in that section that this does not necessarily imply that risk-averse investors will reduce their demand for the risky asset. We provide now another intuition for this phenomenon. Normalizing the initial demand to unity with $E\widetilde{x}_2 u'(w_0 + \widetilde{x}_2) = 0$, they reduce their demand if

$$0 \geq E\widetilde{x}_1 u'(w_0 + \widetilde{x}_1) = E\widetilde{x}_2 u'(w_0 + \widetilde{x}_2 + \widetilde{x}) + E\widetilde{x} u'(w_0 + \widetilde{x}_2 + \widetilde{x}). \tag{8.8}$$

We see that we can decompose the effect of an increase in risk of return into a wealth/background risk effect on one hand, and a substitution effect on the other hand. Let us start with the substitution effect, which is represented by the second term in the right-hand side of condition. This term is always negative, under risk aversion. Indeed, using Proposition 3, we have that

$$\begin{aligned} E\widetilde{x} u'(w_0 + \widetilde{x}_2 + \widetilde{x}) &= E_{\widetilde{x}_2}\left[E\widetilde{x} u'(w_0 + \widetilde{x}_2 + \widetilde{x}) \mid \widetilde{x}_2\right] \\ &\leq E_{\widetilde{x}_2}\left[E\left[\widetilde{x} \mid \widetilde{x}_2\right] E\left[u'(w_0 + \widetilde{x}_2 + \widetilde{x}) \mid \widetilde{x}_2\right]\right] = 0. \end{aligned}$$

The first term in the right-hand side of condition (8.8) is typically a background risk effect: how does the introduction of risk $\widetilde{x}$ to wealth affect the demand for $\widetilde{x}_2$? Because the effect of a background risk is ambiguous, the global effect of an increase in risk of return is also ambiguous. Moreover, this background risk $\widetilde{x}$ is potentially correlated to the initial risk $\widetilde{x}_2$. However, in the special case of a Rothschild-Stiglitz increase in risk that takes the form of adding an independent noise, this second term is unambiguously negative under risk vulnerability. We obtain the following Proposition, which is due to Gollier and Schlesinger (1996).

Proposition 24 *Suppose that u is risk vulnerable. Then, in the standard portfolio problem, adding an independent zero-mean noise to the distribution of returns of the risky asset reduces the demand for it.*

8.3 Risk vulnerability and the equity premium puzzle

Does the existence of an uninsurable risk on human capital could explain the equity premium puzzle? This question was raised by Mehra and Prescott (1985) and it has been examined by Weil (1992). What the above analysis shows is that it can help solving the puzzle if one assumes that preferences are risk vulnerable. The question now is to determine whether it yields a sizeable effect.

The model is as follows. In addition to the one unit of physical capital, each agent in the economy is endowed with some human capital. As in the basic model, the revenue generated by each unit of physical capital is denoted $\widetilde{y}$. It is perfectly correlated across firms. The revenue generated by the human capital is denoted $\widetilde{x}$. It is assumed that it is independently distributed across agents, and is also independent of $\widetilde{y}$. The equity premium in this economy will be equal to

$$\phi = \frac{E\widetilde{y}Eu'(\widetilde{y}+\widetilde{x})}{E\widetilde{y}u'(\widetilde{y}+\widetilde{x})} - 1. \tag{8.9}$$

To keep things simple, let us assume that background risk $\widetilde{x}$ is distributed as $(-k, 1/2; +k, 1/2)$. Notice that k is the standard deviation of the growth of labour incomes. Again, using the historical frequency of the growth of GDP per capita for $\widetilde{y}$, and using a CRRA utility function with $\gamma = 2$, we obtain the equity premium as a function of the size k of the uninsurable risk. We report these results in Figure 8.1. We see that a very large background risk is required to explain the puzzle, with a standard deviation of the annual growth of individual labour income exceeding 80%! We conclude that the existence of idiosyncratic background risks cannot explain in itself the equity premium puzzle.

8.4 Generalized risk vulnerability

A consequence of risk vulnerability is that any background risk with a non-positive mean raises the risk premium that one is ready to pay to escape another inde-

Figure 8.1: The equity premium with background risk

pendent risk. It implies in particular that the presence of a non-positive-mean background risk will never transform another undesirable independent risk into a desirable one. As Pratt and Zeckhauser (1987) and Kimball (1993), one could ask the same question for other types of background risk. Let $\Sigma_i(z, u)$ be a set of background risks that satisfy some properties that can depend upon the preference u of the agent and his initial wealth z. We say that the von Neumann-Morgenstern utility function u is vulnerable at z with respect to $\Sigma_i(z, u)$ if an undesirable risk given background wealth z can never be made desirable by the introduction of any independent background risk $\widetilde{x}$ in $\Sigma_i(z, u)$:

$$Eu(z + \widetilde{y}) \leq u(z) \text{ and } \widetilde{x} \in \Sigma_i(z, u) \Longrightarrow Eu(z + \tilde{x} + \tilde{y}) \leq Eu(z + \widetilde{x}). \quad (8.10)$$

If property (8.10) holds for every z, then u is said to be vulnerable to Σ_i ("$u \in V_i$").

There are four sets Σ_i, $i = 1, 2, 3, 4$, that lead to known restrictions on the utility function.

1. Decreasing absolute risk aversion: V_1.

   DARA is necessary and sufficient for a reduction in wealth to enlarge the set of undesirable risks. DARA is thus equivalent to vulnerability to $\Sigma_1 \equiv \{\widetilde{x} \mid \exists x_0 \leq 0 : \widetilde{x} = x_0 \text{ with probability } 1\}$. It is equivalent to $A' \leq 0$, or $P \geq A$.

2. Risk vulnerability: V_2.

   By definition, risk vulnerability is equivalent to vulnerability to $\Sigma_2 \equiv \{\widetilde{x} \mid E\widetilde{x} \leq 0\}$. Obviously, $\Sigma_1 \subset \Sigma_2$: risk vulnerability is more demanding than DARA, since the same property is required to hold for a larger set of background risks under risk vulnerability than under DARA. It is noteworthy that the concepts V_1 and V_2 rely on background risk sets Σ_1 and Σ_2 whose characterizations are independent of z and u, in contrast to the following two concepts.

3. Properness: V_3.

   The concept of proper risk aversion proposed by Pratt and Zeckhauser [1987] corresponds in our terminology to vulnerability to $\Sigma_3(z,u) \equiv \{\widetilde{x} \mid Eu(z+\widetilde{x}) \leq u(z)\}$, the set of undesirable risks at w. It answers to the question raised by Samuelson (1963) that two independent risks that are separately undesirable are never jointly desirable. Observe that, by risk aversion, $\Sigma_3(z,u)$ is a subset Σ_2. Therefore, Pratt and Zeckhauser's concept of properness is a special case of risk vulnerability V_2. It is known that all HARA utility functions are proper. Pratt and Zeckhauser showed that condition $A'' \geq A'A$ is necessary for properness.

4. Standardness: V_4.

   The concept of standard risk aversion introduced by Kimball [1993] corresponds to vulnerability to $\Sigma_4(z,u) \equiv \{\widetilde{x} \mid Eu'(z+\widetilde{x}) \geq u'(z)\}$, i.e. the set of expected-marginal-utility-increasing risks. Given risk aversion, $\Sigma_4(z,u)$ includes Σ_1 as a subset. Therefore, decreasing absolute risk aversion is necessary for V_4. Given this fact, $\Sigma_3(z,u)$ is a subset of $\Sigma_4(z,u)$, since decreasing absolute risk aversion means that $-u'$ is more concave than u. Therefore, as shown by Kimball, standardness implies properness (which itself implies risk vulnerability). A specific presentation of standardness is presented in the next section.

To sum up, we have that

$$\text{Standardness} \Longrightarrow \text{Properness} \Longrightarrow \text{Risk vulnerability} \Longrightarrow \text{DARA}.$$

We will not write down the necessary and sufficient condition for Pratt and Zeckhauser's properness. The reader can do it directly by using the Diffidence

Theorem after having observed that vulnerability at z with respect to $\Sigma_i(z,u)$ is equivalent to the condition that the indirect utility function $v(.) = Eu(. + \widetilde{x})$ is more diffident at z than the original utility function, i.e., that

$$\widetilde{x} \in \Sigma_i(z,u) \Longrightarrow \frac{Eu(z+y+\widetilde{x}) - Eu(z+\widetilde{x})}{Eu'(z+\widetilde{x})} \leq \frac{u(z+y)-u(z)}{u'(z)} \tag{8.11}$$

for all y. It is a straightforward exercise to apply the Diffidence Theorem to these problems, for $i = 3$ or 4.

Thus, vulnerability at z with respect to $\Sigma_i(z,u)$ just says that the indirect utility function v is more diffident at z than the original utility function u for any background risk in $\Sigma_i(z,u)$. As we know, this does not necessarily imply that v is more concave than u. We can just infer of it that, locally at z, $-v''(z)/v'(z)$ is larger than $-u''(z)/u'(z)$. In particular, we cannot infer anything about whether the agent with wealth z will invest less in the risky asset due to the presence of a background risk in $\Sigma_i(z,u)$. That would require that the indirect utility function be centrally more risk-averse than the original utility function at z. As we know from Proposition 15, this is a condition stronger than more diffidence at z. One may thus define a notion of "strong vulnerability" that is stronger than the notion defined above. We say that the utility function u is strongly vulnerable at z with respect to $\Sigma_i(z,u)$ if any background risk $\widetilde{x}$ in $\Sigma_i(z,u)$ makes the indirect utility function centrally more risk-averse at z, i.e., reduces the demand for any independent risky asset.

$$\widetilde{x} \in \Sigma_i(z,u) \Longrightarrow y\frac{Eu'(z+y+\widetilde{x})}{Eu'(z+\widetilde{x})} \leq y\frac{u'(z+y)}{u'(z)} \tag{8.12}$$

for all y.

We know that strong vulnerability and vulnerability do not differ for the case Σ_1 and Σ_2. We show in the next section that it is also true for Σ_4. No general result has been established for case $\Sigma_{3.}$

8.5 Standardness

In this section, we take a closer look at the notion of standard risk aversion. We show that decreasing absolute risk aversion and decreasing absolute prudence are

necessary and sufficient for both conditions (8.11) and (8.12) for $\Sigma_4(z,u) \equiv \{\widetilde{x} \mid Eu'(z+\widetilde{x}) \geq u'(z)\}$. These results are in Kimball (1993).

Any risk in $\Sigma_4(z,u)$ can be decomposed into an expected-marginal-utility p-reserving risk and a sure reduction in wealth. DARA takes care of the second effect. Thus, we now assume DARA, and we show that decreasing absolute prudence is necessary and sufficient for a expected-marginal-utility preserving risk to yield the result. Consider any background risk $\widetilde{x}$ such that $Eu'(z+\widetilde{x}) = u'(z)$, or $E\widetilde{x} - \psi(z,u,\widetilde{x}) = 0$ where ψ is the precautionary premium.

The necessity of decreasing prudence is immediate. Remember that a necessary condition for both more diffidence at z and central more risk aversion at z is an increase in risk aversion locally at z. As we already observed, the absolute risk aversion of the indirect utility function satisfies the following property:

$$\frac{-v''(z)}{v'(z)} = \frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} = (1 - \frac{\partial\psi}{\partial z}(z,u,\widetilde{x}))A(z).$$

Because we need this to be larger than $-u''(z)/u'(z) = A(z)$, a necessary condition is that $\frac{\partial\psi}{\partial z}$ be negative. By Proposition 53, this condition is equivalent to decreasing absolute prudence.

We now prove the sufficiency of decreasing absolute prudence for condition (8.12) to hold for any expected-marginal-utility preserving risk. Remember that decreasing absolute prudence means that $-u''(z+y+\xi)$ is log-supermodular with respect to (y,ξ). It means that

$$yu''(z+y+\xi)u''(z) \geq yu''(z+y)u''(z+\xi)$$

for all $\xi > 0$, with the inequality reversed for $\xi \leq 0$. Since the direction of integration is opposite the direction of the inequality, it can be integrated from $\xi = 0$ to $\xi = x$ to yield

$$y\left[u'(z+y+x) - u'(z+y)\right]u''(z) \geq yu''(z+y)\left[u'(z+x) - u'(z)\right].$$

Since this condition holds for all x, taking the expectation yields

$$y\left[Eu'(z+y+\widetilde{x}) - u'(z+y)\right]u''(z) \geq yu''(z+y)\left[Eu'(z+\widetilde{x}) - u'(z)\right] = 0.$$

In consequence, decreasing absolute prudence is sufficient for $yEu'(z+y+\widetilde{x}) \leq yu'(z+y)$. Dividing by $Eu'(z+\widetilde{x}) = u'(z)$ implies that

$$y\frac{Eu'(z+y+\widetilde{x})}{Eu'(z+\widetilde{x})} \leq y\frac{u'(z+y)}{u'(z)}$$

for all y, which means that the indirect utility function is centrally more risk-averse at z than the original utility function. We conclude that standardness is necessary and sufficient for any expected-marginal-utility-increasing background risk to reduce the demand for the risky asset.

Because more diffidence is a weaker concept than central more risk aversion, this also proves that standardness is sufficient for the indirect utility function to be more diffident at z than the original utility function, for any $\widetilde{x} \in \Sigma_4(z,u)$. Since we already proved that standardness is necessary, we obtain the following Proposition.

Proposition 25 *Standard risk aversion is necessary and sufficient for any of the two following properties to hold:*

1. *any expected-marginal-utility-increasing background risk reduces the demand for an independent risky asset;*
2. *any expected-marginal-utility-increasing background risk makes the agent to reject more independent lotteries.*

Kimball (1993) provides a variant to this Proposition. A problem with the result above is that it deals with a background risk that increases the expected marginal utility given sure wealth z, whereas we assume that the agent has the opportunity to invest this wealth in risky activities. Kimball shows that the same result applies for background risks that raises Eu' evaluated with the optimal exposure to the endogenous risk. More precisely, the problem is written as

$$\left.\begin{array}{l} E\widetilde{y}u'(z+\widetilde{y}) = 0 \\ Eu'(z+\widetilde{y}+\widetilde{x}) \geq Eu'(z+\widetilde{y}) \end{array}\right\} \Longrightarrow E\widetilde{y}u'(z+\widetilde{x}+\widetilde{y}) \leq 0. \tag{8.13}$$

The first condition states that the investor invests one dollar in the risky asset whose return is $\widetilde{y}$. The second condition states that background risk $\widetilde{x}$ increases the expected marginal utility in the presence of the optimal exposure to risk $\widetilde{y}$. The implication is that this background risk must reduce the demand for $\widetilde{y}$.

Proposition 26 *Standard risk aversion is necessary and sufficient for property (8.13) which states that any background risk that raises Eu' in the presence of the initial optimal investment in the risky asset reduces the demand for this asset.*

Proof: We focus on the proof of sufficiency, which is based again on the fact that decreasing absolute prudence is equivalent to the log-supermodularity of $-u''(z+y+\xi)$ with respect to (y,ξ). It yields

$$u''(z+y_2+\xi)u''(z+y_1)-u''(z+y_1+\xi)u''(z+y_2)$$

has the same sign as $\xi(y_2-y_1)$. Integrating with respect to ξ,

$$\begin{aligned} h(y_1,y_2,x) &= [u'(z+y_2+x)-u'(z+y_2)]\,u''(z+y_1) \\ &\quad -[u'(z+y_1+x)-u'(z+y_1)]\,u''(z+y_2) \end{aligned}$$

has the same sign as (y_2-y_1). It implies that $E(\widetilde{y}_2-\widetilde{y}_1)h(\widetilde{y}_1,\widetilde{y}_2,\widetilde{x})$ is nonnegative for any distribution of $\widetilde{y}_1,\widetilde{y}_2$ and $\widetilde{x}$. Taking $\widetilde{y}_1$ and $\widetilde{y}_2$ i.i.d. and distributed as $\widetilde{y}$, this condition becomes:

$$\begin{aligned} & [E\widetilde{y}u'(z+\widetilde{y}+\widetilde{x})-E\widetilde{y}u'(z+\widetilde{y})]\,Eu''(z+\widetilde{y}) \\ \geq\ & [Eu'(z+\widetilde{y}+\widetilde{x})-Eu'(z+\widetilde{y})]\,E\widetilde{y}u''(z+\widetilde{y}). \end{aligned} \tag{8.14}$$

Remember now that DARA means that $-u'$ is more concave than u. It yields

$$E\widetilde{y}u'(z+\widetilde{y})=0 \Longrightarrow E\widetilde{y}u''(z+\widetilde{y})\geq 0.$$

In consequence, the right-hand side of inequality (8.14) is nonnegative. A necessary condition for 8.14) is thus $E\widetilde{y}u'(z+\widetilde{y}+\widetilde{x})\leq E\widetilde{y}u'(z+\widetilde{y})=0$.■

To sum up, standardness generates the same comparative statics properties as risk vulnerability, at least for the standard portfolio problem and for the take-it-or-leave-it lottery problem. Because standardness does not imply that the indirect utility function be more concave, there is no guarantee that standardness can generate other comparative statics results. Still, standardness has two important advantages. First, it is easy to characterize: contrary to risk vulnerability that is characterized by a two-dimension inequality (8.6), standardness just require that a one-variable function, P', to be nonnegative. Second, the comparative statics property holds for a larger set of background risks ($\Sigma_2\subset\Sigma_4(z,u)$). This is at the cost of a stricter condition on preferences (standardness implies risk vulnerability).

8.6 Conclusion

It is widely believed that the presence of a risk in background wealth has an adverse effect on the demand for other independent risks. This is not true in general, and the conditions that must be imposed on preferences to guarantee this result depends upon the kind of background risk we have in mind. Without surprise, because it raises the question of the effect of risk on the degree of risk aversion $-u''/u'$, all these conditions rely in one way or another on the fourth derivative of u. In spite of this fact, some of these conditions are quite intuitive. For example, a sufficient condition for a pure background risk to raise the aversion to other independent risks is that absolute risk aversion be decreasing and convex. This is a quite natural assumption, as it means that the risk premium is decreasing with wealth at a decreasing rate.

Chapter 9

Taking multiple risks

In the previous chapter, we examined the effect of an exogenous risk borne by an agent on his behavior towards other independent risks. We now turn to the analysis of the behavior towards independent risks that are endogenous. Typically, we address here the question of how does investing in a portfolio influence the household's behavior toward insuring their car. Or, how does the opportunity to invest in security A affect the demand for another independent security B? A last illustration is about the effect of the opportunity to invest in risky assets on the decision to invest in human capital.

9.1 The interaction between asset demand and gambling

We start with a model in which the agent can invest in a risky asset and, at the same time, gamble a fixed amount in a risky activity. We examine the interaction between these two decisions. The problem is written as

$$\max_{\delta \in \{0,1\}, \alpha} \quad Eu(w_0 + \delta \widetilde{y} + \alpha \widetilde{x}).$$

Random variable $\widetilde{x}$ is the return of the risky asset whereas $\widetilde{y}$ is the net gain in gambling, which is assumed to be independent of $\widetilde{x}$. Decision variable α is the amount invested in the risky asset, whereas δ is the amount invested in gambling, which is constrained to be 1 ("take-it") or 0 ("leave-it"). The intuition suggests that the opportunity to invest in financial markets should reduce the incentive to

take other risks. This means that the investor would reject more gambles $\widetilde{y}$ than if he would not be allowed to invest in financial markets.

In this section, we limit the analysis to the case where the gamble is small with respect to portfolio risk. Thus, the problem simplifies to determining whether the optimal portfolio risk $\widetilde{x}$ raises *local* risk aversion:

$$\forall z, \widetilde{x} : E\widetilde{x}u'(z+\widetilde{x}) = 0 \quad \Longrightarrow \quad \frac{-Eu''(z+\widetilde{x})}{Eu'(z+\widetilde{x})} \geq \frac{-u''(z)}{u'(z)}. \tag{9.1}$$

Applying the Diffidence Theorem generates the following necessary and sufficient condition:

$$\forall z, x : -u''(z+x)u'(z) + u''(z)u'(z+x) \geq \frac{-u'''(z)u'(z) + (u''(z))^2}{u'(z)} xu'(z+x).$$

Rearranging, this is equivalent to:

$$\forall z, x : A(z+x) \geq A(z) + xA'(z)$$

which is true if and only if absolute risk aversion is convex.

Proposition 27 *Investing in a risky asset raises local risk aversion if and only if absolute risk aversion is convex.*

Remember that the convexity of absolute risk aversion, combined with DARA, is sufficient for risk vulnerability, which itself is a natural assumption.

There is an alternative way to prove Proposition 27. Let F be the cumulative distribution function of $\widetilde{x}$. Let the cumulative distribution function of $\widetilde{y}$ be defined by

$$dG(x) = \frac{u'(z+kx)dF(x)}{\int u'(z+ks)dF(s)}.$$

Condition (9.1) may then be rewritten as:

$$\forall z, \widetilde{y} : \quad E\widetilde{y} = 0 \quad \Longrightarrow \quad EA(z+\widetilde{y}) \geq A(z). \tag{9.2}$$

The parallelism with condition (3.1) is perfect when one replaces u by $-A$. The convexity of A is clearly necessary and sufficient for the result. We can further define the substitution premium ι of $\widetilde{y}$, defined as

$$EA(z+\widetilde{y}) = A(z-\iota(z)).$$

In words, the substitution premium associated to the opportunity to invest in a risky asset is the sure reduction in wealth that has the same effect on the demand for another small risk. It is positive if absolute risk aversion is decreasing and convex. Using the Arrow-Pratt approximation, it is approximately equal to

$$\iota(z) \simeq \frac{E\widetilde{y}^2}{2}\frac{A''(z)}{-A'(z)}.$$

9.2 Are independent assets substitutes?

We now turn to the demand interaction between independent assets. It is intuitive to think that the demand for a risky asset should be reduced by the presence of other independent risky assets, i.e., that independent risky assets are substitutes. We start with an analysis of the i.i.d. case, which is easier to treat.

9.2.1 The i.i.d. case

Suppose that there are two risky assets available in the economy, in addition to the risk free asset. Their returns $\widetilde{x}$ and $\widetilde{y}$ are independent and identically distributed. Echoing the Samuelson's story presented in the introduction of chapter 8, suppose that you are offered to invest in $\widetilde{x}$ alone. Contrary to Samuelson (1963), we assume here that you can gamble how much you want on $\widetilde{x}$. Suppose now that you are also offered the opportunity to invest how much you want in another identical and independent gamble. Would you invest less in $\widetilde{x}$ as a consequence?

Because we assume here that $\widetilde{x}$ and $\widetilde{y}$ are i.i.d., we can use Proposition 10 to simplify the analysis. Indeed, this Proposition implies that in the presence of two i.i.d. assets, it is optimal to invest the same amount in each of them. Normalizing the demand to unity when there is only 1 asset, the presence of a second i.i.d. asset reduces the demand for the first one if

$$\left.\begin{array}{l} E\widetilde{x}u'(z+\widetilde{x})=0 \\ \widetilde{x},\widetilde{y} \text{ i.i.d.} \end{array}\right\} \quad \Longrightarrow \quad E(\widetilde{x}+\widetilde{y})u'(z+\widetilde{x}+\widetilde{y})\leq 0 \tag{9.3}$$

where the condition to the right states that the derivative of $h(\alpha) = Eu(z + \alpha(\widetilde{x}+\widetilde{y}))$ is negative when evaluated at the initial optimal exposure $\alpha = 1$. By symmetry, this right condition is equivalent to $E\widetilde{y}u'(z+\widetilde{x}+\widetilde{y}) \leq 0$. In fact, we will look for a weaker condition:

$$\left.\begin{array}{l} E\widetilde{x}u'(z+\widetilde{x})=0 \\ E\widetilde{y}u'(z+\widetilde{y})=0 \end{array}\right\} \quad \Longrightarrow \quad E\widetilde{y}u'(z+\widetilde{x}+\widetilde{y})\leq 0, \tag{9.4}$$

relaxing the condition that $\widetilde{x}$ and $\widetilde{y}$ must be identically distributed.

This condition states that the presence of background risk $\widetilde{x}$ *that is optimal when taken in isolation* reduces the demand for any other independent asset. This is true only if the indirect utility function $Eu(.+\widetilde{x})$ is centrally more risk-averse than u at z, for any background risk satisfying condition $E\widetilde{x}u'(z+\widetilde{x})=0$:

$$E\widetilde{x}u'(z+\widetilde{x})=0 \quad \Longrightarrow \quad y\frac{Eu'(z+y+\widetilde{x})}{Eu'(z+\widetilde{x})}\leq y\frac{u'(z+y)}{u'(z)}. \tag{9.5}$$

Let us define a new set of random variables: $\widehat{\Sigma}(z,u) = \{\widetilde{x} \mid E\widetilde{x}u'(z+\widetilde{x})=0\}$. Condition (9.5) defines the notion of "strong vulnerability" with respect to $\widehat{\Sigma}(z,u)$, as defined in general terms by condition (8.12). Using the Diffidence Theorem, this condition is equivalent to

$$H(z,x,y) = y\left\{u'(z+y+x)u'(z) - u'(z+x)u'(z+y)\right\} \tag{9.6}$$

$$+xyu'(z+x)u'(z+y)\left[A(z+y)-A(z)\right].$$

being nonpositive for all (z,x,y) in the feasible domain.

It is easy to prove that the convexity of absolute risk aversion is a necessary condition for (9.6). This is done by verifying that

$$H(z,x,0)=0,$$

$$\left.\frac{\partial H}{\partial y}\right|_{y=0} = 0,$$

and

$$\left.\frac{\partial^2 H}{\partial y^2}\right|_{y=0} = 2u'(z)u'(z+x)\left[A(z) + xA'(z) - A(z+x)\right].$$

The first two properties imply that a necessary condition is $\left.\frac{\partial^2 H}{\partial y^2}\right|_{y=0}$ be non-positive, which in turn necessitates that $A(z+x) \geq A(z) + xA'(z)$. A similar exercise may be performed with respect to x, with the same result.

Condition (9.6) is not easy to manipulate. It can thus be useful to try to provide a simpler sufficient condition. We use the same method as in the previous chapter where we showed for example that (strong) vulnerability with respect to $\Sigma_4(z,u)$ (standardness) is sufficient for (strong) vulnerability with respect to $\Sigma_2(z,u)$ (risk vulnerability) because $\Sigma_2(z,u) \subset \Sigma_4(z,u)$. If we were able to show that $\widehat{\Sigma}(z,u)$ is in $\Sigma_i(z,u)$, we would have in hand the sufficiency of V_i. The problem is that any risk satisfying condition $E\widetilde{x}u'(z+\widetilde{x})$ must have a positive mean, together with being desirable (otherwise $\alpha = 0$ would dominate strategy $\alpha = 1$). Thus, the only potential candidate is $\Sigma_4(z,u)$. We would have $\widehat{\Sigma}(z,u)$ in $\Sigma_4(z,u)$ if

$$E\widetilde{x}u'(z+\widetilde{x}) = 0 \quad \Longrightarrow \quad Eu'(z+\widetilde{x}) \geq u'(z). \tag{9.7}$$

In words, we are looking for the condition under which the option to purchase stocks increases the marginal value of wealth. As stated in the following Proposition, this is the case if prudence is larger than twice the risk aversion ($P \geq 2A$).

Proposition 28 *Condition (9.7) holds for any z and $\widetilde{x}$ if and only if absolute prudence is larger than twice absolute risk aversion: $P(z) \geq 2A(z)$ for all z.*

Proof: By the Diffidence Theorem (Corollary 1), condition (9.7) holds if and only if

$$u'(z+x) - u'(z) \geq \frac{u''(z)}{u'(z)} xu'(z+x)$$

for all (z, x). If we divide the above condition by $u'(z+x)u'(z)$, this inequality is rewritten as

$$f(z+x) \leq f(z) + xf'(z),$$

where $f(w) = 1/u'(w)$. Obviously, this condition holds for any (z, x) if and only if f is concave. But it is easily checked that $f = 1/u'$ is concave if and only if

$$\frac{-u'''(z)}{u''(z)} \geq 2\frac{-u''(z)}{u'(z)} \ \forall z, \tag{9.8}$$

or $P(z) \geq 2A(z)$.[1] This concludes the proof of the following Proposition, which is due to Gollier and Kimball (1997).[2]■

We conclude that this condition joint with standardness is sufficient for conditions (9.3), (9.4), (9.5) and (9.6). Because decreasing absolute prudence means that $A'' \geq -A'(P-2A)$, we verify that this sufficient condition satisfies the necessary condition $A'' \geq 0$.

Proposition 29 *Two risky assets with i.i.d. returns are substitutes if absolute prudence is decreasing and larger than twice absolute risk aversion. A necessary condition for the substitutability of independent risky assets is that absolute risk aversion be convex.*

To examine how far away this sufficient condition is from the necessary and sufficient condition (9.6), we can look at a specific set of utility function. As usual, let us take the set of HARA functions, with

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma} \tag{9.9}$$

We know that these functions have a convex absolute risk aversion and are standard if γ is positive. They satisfy condition $P \geq 2A$ if γ is less than 1. Thus a

[1] This condition is equivalent to $T'(z) \geq 1$, where $T = 1/A$ is absolute risk tolerance.

[2] This result can also be proven by using Proposition 4 with $g(x) = x$, $h(x, 1) = u'(z+x)$ and $h(x, 2) = x^{-1}\left[u_1'(z) - u_1'(z+x)\right]$.

Figure 9.1: Evaluation of $H(1,1,y)$ for the CRRA utility function, with $\gamma = 5$.

sufficient condition is $0 < \gamma \leq 1$. Can we expect condition (9.6) to be satisfied for other values of γ? In the case $\eta = 0$, this sufficient condition is rewritten as

$$H(z,x,y) = y\left\{(z+y+x)^{-\gamma}z^{-\gamma} - (z+x)^{-\gamma}(z+y)^{-\gamma} - \frac{\gamma y x(z+x)^{-\gamma}(z+y)^{-\gamma}}{z(z+y)}\right\} \tag{9.10}$$

being nonpositive for all (z,x,y) in the feasible domain

$$\{(z,x,y) \mid z+x>0; z+y>0 \text{ and } z+x+y>0\}.$$

Fix x and y positive, and let z approach 0. Then, the first and the third terms in the right-hand side of (9.10) tends to infinity. If γ is larger than 1, this is the first which dominates, and H tends to $+\infty$. Function H is depicted in Figure 9.1 for $\gamma = 5$, $z = 1$, $x = 1$ and $\eta = 0$. We conclude that, in the set of HARA utility functions, condition $P \geq 2A$ is necessary and sufficient. Observe that we have that $H(1,1,0) = 0$, $\frac{\partial H}{\partial y}(1,1,0) = 0$, $\frac{\partial^2 H}{\partial y^2}(1,1,0) < 0$, but H is positive for large values of y.

A counter-example is thus easy to obtain. Take the CRRA utility function with $\gamma = 5$, $z = 1$ and $\widetilde{x}$ and $\widetilde{y}$ distributed as $(-0.1, 4/10; 10, 6/10)$. After some computations, we get that the optimal investment when only one risky asset is offered is 0.1678. When the second risky asset is also offered, the investor increases his demand for the first asset up to 0.1707.

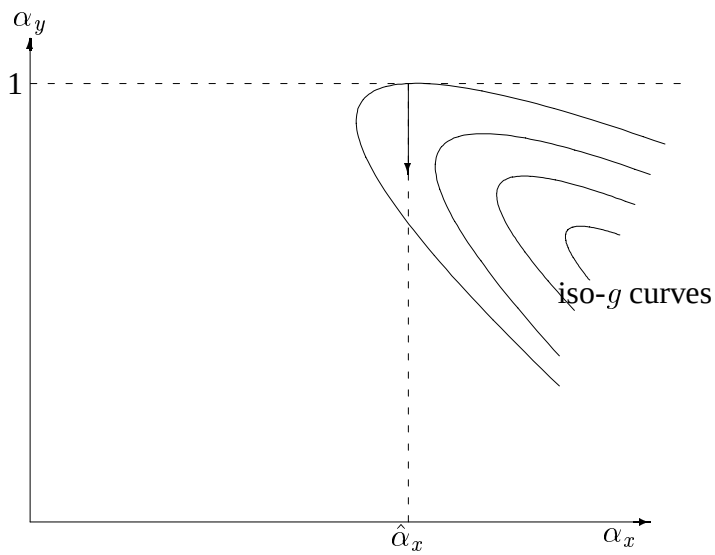

Figure 9.2: α_y^* is less than 1 if $\frac{\partial g}{\partial \alpha_y}(\widehat{\alpha}_x, 1)$ is negative.

9.2.2 The general case

In this paragraph, we remove the assumption that $\widetilde{x}$ and $\widetilde{y}$ are identically distributed. This makes the problem less easy to solve because we cannot use the diversification argument anymore. The absence of symmetry forces us to perform a comparative statics analysis with two decision variables.

Consider a function $g(\alpha_x, \alpha_y)$ that is concave with respect to (α_x, α_y). Let (α_x^*, α_y^*) denote the optimal solution of the maximization of $g(\alpha_x, \alpha_y)$. Under which condition is α_y^* less than 1? The standard method consists in first obtaining the α_x that maximizes $g(\alpha_x, 1)$. Let us denote it $\widehat{\alpha}_x$. Then, α_y^* is less than 1 if $\frac{\partial g}{\partial \alpha_y}$ is negative when evaluated at $(\widehat{\alpha}_x, 1)$. This is confirmed by Figure 9.2.

We apply this technique for $g(\alpha_x, \alpha_y) = Eu(z + \alpha_x \widetilde{x} + \alpha_y \widetilde{y})$, which is concave under risk aversion. As before, normalize to 1 the demand for $\widetilde{y}$ when it is the only asset on the market. Let us also normalize $\widehat{\alpha}_x$ to unity. This means that the optimal demand for $\widetilde{x}$ in the presence of one unit of $\widetilde{y}$ equals 1. Thus, the demand for $\widetilde{y}$ is reduced by the opportunity to also invest in $\widetilde{x}$ if and only if the following property holds:

$$\left.\begin{array}{ll}(i) & E\widetilde{y}u'(z+\widetilde{y})=0\\ (ii) & E\widetilde{x}u'(z+\widetilde{x}+\widetilde{y})=0\end{array}\right\} \quad \Longrightarrow \quad (iii) \quad E\widetilde{y}u'(z+\widetilde{x}+\widetilde{y})\leq 0. \tag{9.11}$$

The condition (i) states that the optimal demand for asset $\widetilde{y}$ equals 1, when it is the only asset offered on the market. Condition (ii) states that the optimal demand for asset $\widetilde{x}$ when the investor already has one unit of $\widetilde{y}$ is also 1. Condition (iii) requires that the investor then demands less than one unit of $\widetilde{y}$.

Except for condition (ii), this problem is the same as problem (9.4). Applying the Diffidence Theorem twice, once by taking $\widetilde{x}$ as given, and then again over $\widetilde{y}$, would yield a rather complex necessary and sufficient condition. We will limit our analysis to showing that the combination of standardness with condition $P > 2A$, is also sufficient for (9.11). We would be done if we could show that background risk $\widetilde{x}$ in condition (iii) raises the expected marginal utility of wealth. The trick here is to use Proposition 26 rather than Proposition 25, by showing that

$$E\widetilde{x}u'(z+\widetilde{x}+\widetilde{y})=0 \quad \Longrightarrow \quad Eu'(z+\widetilde{x}+\widetilde{y})\geq Eu'(z+\widetilde{y}) \tag{9.12}$$

when P is larger than $2A$. Knowing that condition (8.13) holds under standardness would yield the result.

To prove that condition (9.12) holds under $P > 2A$, define the indirect utility function $v(w) = Eu(w+\widetilde{y})$. The problem can then be rewritten as

$$E\widetilde{x}v'(z+\widetilde{x})=0 \quad \Longrightarrow \quad Ev'(z+\widetilde{x})\geq v'(z). \tag{9.13}$$

Using Proposition 28, this is true if and only if the indirect utility function satisfies condition $P_v > 2A_v$. But we also know by Proposition 19 that the indirect utility function inherits this property from u. We conclude that, under condition $P > 2A$, property (9.11) holds if property (8.13) holds. The use of Proposition 26 concludes the proof of the following Proposition, which generalizes Proposition 29 to assets whose returns are not identically distributed.

Proposition 30 *Independent risky assets are substitutes if absolute prudence is decreasing and larger than twice absolute risk aversion. A necessary condition for the substitutability of independent risky assets is the convexity of absolute risk aversion.*

9.3 The equity premium with internationally diversified portfolios

Can we link this analysis to the equity premium puzzle? In our simple version of the Lucas' model, we assumed that US households could just purchase US stocks. In the real world, US citizens can also invest their money in Japan or in Europe. Let us consider an economy with two countries of equal size with agents having the same attitude towards risk expressed by utility function u. Suppose that firms in country 1 generate revenues $\widetilde{y}_1$ whereas firms in country 2 generate revenues $\widetilde{y}_2$. For the sake of simplicity, let us assume that $\widetilde{y}_1$ and $\widetilde{y}_2$ are i.i.d.. In the Lucas tree economy, it means that the two countries have the same climate, but are exposed to independent meteorological shocks. The question is whether assuming that citizens can invest only in their own country leads to an underestimation of the equity premium. Or alternatively, does liberalizing capital flows between the two countries raise the equity premium?

We see that the question is closely related to the one that we examined in this chapter. Here, the question is not about how the opportunity to invest in another independent risky asset affect the demand for the first asset, but rather about how it affect the aggregate demand for stocks. Our analysis above does not provide any information on this question. Indeed, if independent risky assets are substitute, it may be that the reduction of the demand for the first asset due to the present of the other asset be so large that the global demand for risky assets be reduced. That would raise the equity premium in the two countries. This scenario, however, is not likely to occur. There is indeed a strong diversification effect that will take place with the liberalization of capital flows.

Observe first that, because of the symmetry, asset prices will be the same at equilibrium in the two countries. Then, as we have seen from Proposition 10, it will optimal for all agents to perfectly diversify their portfolio. The problem is thus to determine the demand for an asset with payoff $0.5(\widetilde{y}_1 + \widetilde{y}_2)$ compared to the demand for the asset yielding payoff $\widetilde{y}_1$. But we know that $0.5(\widetilde{y}_1 + \widetilde{y}_2)$ is a SSD-dominant shift in the distribution of risk with respect to $\widetilde{y}_1$. We can then use Proposition 18 which provides a sufficient condition for any such shift to increase the demand for the risky asset. This yields the following result:

Proposition 31 *Suppose that countries faces idiosyncratic shocks on their revenues. Suppose also that relative risk aversion is less than unity and nondecreasing, and absolute risk aversion is nonincreasing. Then, liberalizing capital flows*

will raise the aggregate demand for stocks. At equilibrium, it will reduce the equity premium.

We see that under these conditions, the equity premium puzzle is even stronger than as been stated in Chapter 6. Taking into account of international diversification of portfolio would reduce the theoretical value of the equity premium. But the assumptions that are proposed in the above Proposition are not satisfactory. In particular, it requires that relative risk aversion be less than unity, a very doubtful assumption. Eeckhoudt, Gollier and Levasseur (1994) provides other sufficient conditions. Also, it should be noted that individual portfolios do not seem to be as much internationally diversified than they should. This makes this discussion less important, but it raises another puzzle of financial markets.

We conducted some numerical computations on the effect of liberalizing capital flows in the simple two-country model that we presented in this section. With liberalized financial markets, the equity premium is:

$$\phi = \frac{E\left[\frac{\widetilde{y}_1+\widetilde{y}_2}{2}\right] E\left[u'\left(\frac{\widetilde{y}_1+\widetilde{y}_2}{2}\right)\right]}{E\left[\frac{\widetilde{y}_1+\widetilde{y}_2}{2} u'\left(\frac{\widetilde{y}_1+\widetilde{y}_2}{2}\right)\right]} - 1. \tag{9.14}$$

Once again, using historical data for the growth of real GDP in the U.S., we were able to compare the equity premia in the two situations for different values of the constant relative risk aversion. It yields the following Table:

RRA	Equity premium in autharky	Equity premium with international diversification
$\gamma = 0.5$	0.03%	0.015%
$\gamma = 1$	0.06%	0.03%
$\gamma = 2$	0.11%	0.06%
$\gamma = 4$	0.23%	0.12%
$\gamma = 10$	0.61%	0.32%
$\gamma = 40$	2.86%	1.42%

Table: The comparison of equity premia before and after the liberalization of capital flows between two countries.

Of course, this is an academic exercise, since real GDP of different countries are affected by common random factors. Still it provides an upper bound on the

reduction of the equity premium that can be obtained by organizing better ways to allow people to diversify their portfolio. In the case of i.i.d. shocks on national productions, allowing two equally sized countries to exchange capital ventures reduces the equity premium by two. This is almost uniquely due to diversification, since the variance of individual portfolio returns has been divided by two. Indeed, it should be recalled that the equity premium is approximately proportional to the variance of the portfolio revenues.

9.4 Conclusion

We can feel in this chapter the difficulty to determine the optimal composition of a portfolio when departing from mainstream finance in which the mean-variance paradigm is the dogma. We looked here at the simplest case that we can imagine after the standard one-risk-free-one-risky-asset model. Namely, we assumed the existence of one risk free and two risky assets with independent returns. Except when absolute risk aversion is constant, the presence of the second risky asset affects the demand for the first. The intuition suggests that it should reduce it. We showed that this is true if absolute risk aversion is standard and smaller than half the absolute prudence.

Chapter 10

Horizon length and portfolio risk

How should the length of one's investment horizon affect the riskiness of his portfolio? This question confronts start-up companies choosing which ventures to pursue before they go public, ordinary investors building a nest egg, investment managers concerned with contract renewals and executives seeking strong performance before their stock options come due, among others.

Portfolio decisions are the focus of this analysis. However, the horizon-risk relationship extends beyond finance. Thus, students may vary their strategy on grades – say how venturesome a paper to write – over the course of the grading period; presidents may adjust the risk of their political strategies from early through mid-term and then as they approach an election.

Popular treatments suggest that short horizons often lead to excessively conservative strategies. Thus, the decisions of corporate managers, judged on their quarterly earnings, are said to focus too much on safe, short-term strategies, with underinvestment say in risky R & D projects. Privately-held firms, it is widely believed, secure substantial benefit from their ability to focus on longer-term projects. Mutual fund managers, who get graded regularly, are also alleged to focus on strategies that will assure a satisfactory short-term return, with long-term expectations sacrified.[1]

Economists and decision theorists, speculators and bettors, have long been fascinated by the problem of repeated investment. Thus, long ago Bernoulli provided the first motivation of utility theory when he confronted the St. Petersburg

[1]The experience of the U.S. mutual fund Twentieth Century Giftrust is instructive. It requires monies to be left with it for 10 years at least. Managers of the fund suggest that thanks to this 10-year no-withdrawal rule the fund was able to return 24 percent annually since 1985, nearly 10 points better than the S & P500.(See Newsweek 6/19/95, page 60)

Paradox, whose components can be reformulated as payoffs from an infinite series of actuarially fair, double-till-you-lose bets. Successful speculators must manage their money effectively even when making a series of bets that are actuarially favorable. They must determine how much to allocate to each gamble given its odds, future prospects, and the time horizon.

In recent years, this class of problems has been pursued in two different literatures, one in utility theory the other in dynamic investment strategies. A recurring theme in both is that the opportunity to make further investments affects how one should invest today. These two literatures have not merged, in part because a key question has gone unresolved. How should the length of the investment horizon affect the riskiness of one's investments? Some special cases have yielded results, as in the case of HARA utility functions. And with any particular utility function and set of investment opportunities actual calculations, perhaps using a simulation, could answer that question within a dynamic programming framework. But the central theoretical question of the link between the structure of the utility function and the horizon-riskiness relationship remained unresolved. Guiso, Jappelli, and Terlizzese (1996) test the relation between age and risk-taking in a cross-section of Italian households. Their empirical results show that young people, presumably facing greater income risk than old, actually hold the smallest proportion of risky assets in their portfolios. The share of risky assets increases by 20% to reach its maximum at age 61 (Guiso et al, p. 165).

In the formal literature, the horizon-riskiness issue has received the greatest attention addressing portfolios appropriate to age. Samuelson (1989) and several others have asked: "As you grow older and your investment horizon shortens, should you cut down your exposure to lucrative but risky equities?" Conventional wisdom answers affirmatively, stating that long-horizon investors can tolerate more risk because they have more time to recoup transient losses. This dictum has not received the imprimatur of science, however. As Samuelson (1963, 1989a) in particular points out, this "time-diversification" argument relies on a fallacious interpretation of the Law of Large Numbers: repeating an investment pattern over many periods does not cause risk to wash out in the long run.

Early models of dynamic risk-taking, such as those developed by Samuelson [1969] and Merton (1969) find no relationship between age and risk-taking. This is hardly surprising, since the models, like most of them in continuous time finance, employ HARA utility functions. This choice is hardly innocuous: If the risk-free rate is zero, myopic investment strategies are rational iff the utility function is HARA (Mossin, 1968). Given positive risk-free rates, empirically the normal case, only constant relative risk aversion (CRRA) utility functions allow for

myopia. A myopic investor bases each period's decision on that period's initial wealth and investment opportunities, and maximizes the expected utility of wealth at the end of the period. The value function on wealth exhibits the same risk aversion as the utility function on consumption; the future is disregarded. This chapter is devoted to the characterization of optimal dynamic strategies when the utility function is not HARA. Most of the results presented in this chapter are from Gollier and Zeckhauser (1998).

10.1 Static versus dynamic optimization

How does the possibility to take risk tomorrow affect the optimal exposure to today risk? This problem is close to the one that we raised in chapter 9, i.e., how does the possibility to take risk $\widetilde{y}$ today affect the optimal exposure to independent risk $\widetilde{x}$ that is also available today? The difference lies on the timing of the decisions. The problem that we examined in chapter 9 was static in the sense that the decisions on $\widetilde{x}$ and $\widetilde{y}$ had to be taken simultaneously. We now consider the case where the decision on $\widetilde{x}$ can be taken after having observed the realization of $\widetilde{y}$. Because we assume all along this chapter that $\widetilde{x}$ and $\widetilde{y}$ are independent, the benefit of this relaxed assumption is not on a better knowledge of the distribution of $\widetilde{x}$. Rather, the benefit comes from the added of available strategies for the investor. Indeed, the investor can here adapt α_x to losses or benefits that he incurred on $\widetilde{y}$. For example, under DARA, the investor will reduce his exposure to risk $\widetilde{x}$ if he incurred a loss on $\widetilde{y}$, and he will increase it otherwise. There is thus no doubt that this increased flexibility raises the lifetime expected utility of the investor. What is more difficult to examine is the comparative statics effect.

We consider a simple model with no intermediary consumption: the investor is willing to maximize the expected utility of his final wealth, which is the initial wealth plus the accumulated benefits and losses on the subsequent risks undertaken. Investors choose their portfolio to maximize the expected utility of the wealth that they accumulated at retirement. There are three dates $t = 0, 1, 2$. Young investors invest at dates $t = 0$ and 1, and they retire at date $t = 2$. Old investors invest at date $t = 0$, and they retire at $t = 1$. Each investor is endowed with wealth w_0 at date 0, and has utility function u, which is assumed to be twice differentiable, increasing and concave. There is no serial correlation on the return of risky assets.

The problem of young investors is solved by backward induction. Namely, one first looks at the investment problem at date $t = 1$, which is a static one. One

does it for any potential wealth level z that is attained at that date. One can thus define the value function $v(z)$ as the maximal expected utility that can be obtained on financial markets with this wealth level at date 1. The problem of the young at date $t = 0$ is then to find the investment strategy at that date which maximizes the expectation of the value v of wealth z generated over the first period. This is again a static investment problem, but with the utility function u that is replaced by the value function v. Thus, the impact of time horizon on the optimal investment today can be measured by the change in the degree of concavity between u and v. If v is less concave than u, we will say that duration enhances risk.

10.2 The standard portfolio problem

10.2.1 The model

In this model, the active investors at date $t = 0$ or 1 have the opportunity to invest in a risk free asset with a zero return and in a risky asset whose return is distributed as $\widetilde{x}_t$. Again, we assume that $\widetilde{x}_0$ and $\widetilde{x}_1$ are independent. The problem of the investors that are still active at date $t = 1$ with an accumulated wealth z is written as:

$$v(z) = \max_{\alpha} \quad Eu(z + \alpha\widetilde{x}_1). \tag{10.1}$$

The problem of young investors at date $t = 0$ is thus written as

$$\max_{\alpha} \quad Ev(w_0 + \alpha\widetilde{x}_0)$$

whereas the problem of older investors who retire at date $t = 1$ is to maximize $Eu(w_0 + \alpha\widetilde{x}_0)$. Using Proposition 14, we know that young investors will invest more than older ones, whatever w_0 and $\widetilde{x}_0$, if and only if v is less concave than u.

The existence of the second-period investment opportunity exerts two effects on first-period risk taking, which we label the flexibility and background risk effects. The background risk effect is similar to the one that we examined in the previous chapter. But an important difference appears in this dynamic framework. That is, the investor is flexible on how much of risk $\widetilde{x}_1$ he will accept. More specifically, the investor will adjust optimal risk exposure in the second period to the outcome in the first. With decreasing absolute risk aversion, for example,

the better the first-period outcome the more of the risky asset is purchased in the second period. The opportunity to adjust one's portfolio is an advantage; this flexibility effect always reduces aversion to current risks. The relationship between the optimal exposure at date 1 and the accumulated wealth z can be evaluated by solving the first-order condition to problem (10.1) for every z:

$$E\widetilde{x}_1 u'(z + \alpha(z)\widetilde{x}_1) = 0. \tag{10.2}$$

Let us assume that u is twice differentiable. Fully differentiating this condition yields

$$\frac{d\alpha}{dz} = -\frac{E\widetilde{x}_1 u''(z + \alpha\widetilde{x}_1)}{E\widetilde{x}_1^2 u''(z + \alpha\widetilde{x}_1)}. \tag{10.3}$$

As stated in Proposition 3, $\alpha(z)$ is increasing, constant or decreasing depending upon whether absolute risk aversion is decreasing, constant or increasing. The envelope theorem yields

$$v'(z) = Eu'(z + \alpha(z)\widetilde{x}_1). \tag{10.4}$$

Fully differentiating again this equality yields

$$v''(z) = Eu''(z + \alpha\widetilde{x}_1) + \frac{d\alpha}{dz} E\widetilde{x}_1 u''(z + \alpha\widetilde{x}_1) \tag{10.5}$$

where $\alpha = \alpha(z)$. Combining conditions (10.3), (10.4) and (10.5) allows us to write

$$\begin{aligned} -\frac{v''(z)}{v'(z)} &= -\frac{Eu''(z + \alpha\widetilde{x}_1)}{Eu'(z + \alpha\widetilde{x}_1)} + \frac{d\alpha}{dz}\frac{-E\widetilde{x}_1 u''(z + \alpha\widetilde{x}_1)}{Eu'(z + \alpha\widetilde{x}_1)}. \\ &= -\frac{Eu''(z + \alpha\widetilde{x}_1)}{Eu'(z + \alpha\widetilde{x}_1)} + \frac{[E\widetilde{x}_1 u''(z + \alpha\widetilde{x}_1)]^2}{E\widetilde{x}_1^2 u''(z + \alpha\widetilde{x}_1) Eu'(z + \alpha\widetilde{x}_1)}. \end{aligned} \tag{10.6}$$

Our aim now is to compare $-v''(z)/v'(z)$ to $-u''(z)/u'(z)$. Our two effects emerge from condition (10.6). The flexibility effect is expressed by the second term in the right-hand side of (10.6), which is negative. The background risk effect

corresponds to the first term in the right-hand side of (10.6). Future risk $\alpha\widetilde{x}_1$ can be interpreted as a background risk with respect to the independent current risk. If α would be fixed and independent of the realization of the first-period risk (i.e. $d\alpha/dz = 0$), the degree of risk aversion of the young investor would equal $-Eu''(z+\tilde{y})/Eu'(z+\tilde{y})$. Notice that we know by Proposition 27 that this is larger (resp. smaller) than $-u''(z)/u'(z)$ if and only if absolute risk aversion is convex (resp. concave). In consequence, the two effects are in accordance with the hypothesis that duration enhances risk if absolute risk aversion is concave. The concavity of A is sufficient, but not necessary for this result.

Proposition 32 *Consider the two-period investment problem with a zero yield risk-free asset and another risky asset. Young investors are less risk-averse than old investors if absolute risk aversion is concave.*

10.2.2 The HARA case

The above model is not tractable to extract the solution to the problem of the young investor, expect in the HARA case with

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma}.$$

We know that absolute risk aversion is convex for HARA utility functions. The Proposition 32 is not helpful to solve the problem. But using results derived in section (5.3), we know that the optimal second period strategy $\alpha(z)$ is such that

$$\alpha(z) = a(\eta + z/\gamma)$$

with

$$E\widetilde{x}_1 \left(1 + \frac{a\widetilde{x}_1}{\gamma}\right)^{-\gamma} = 0.$$

It implies that

$$v(z) = \zeta E(\eta + \frac{z}{\gamma} + \frac{\alpha(z)\widetilde{x}_1}{\gamma})^{1-\gamma} = \zeta K(\eta + \frac{z}{\gamma})^{1-\gamma} = Ku(z)$$

where

$$K = E(1 + \frac{a\widetilde{x}_1}{\gamma})^{1-\gamma}.$$

We conclude that, in the HARA case, the value function v represents the same attitude towards risk than the original utility function u, since the first is a linear transformation of the second. Duration has no effect on risk-taking in this case. An interpretation of this result is that the flexibility effect just compensate the background risk effect which goes the opposite direction.

10.2.3 The necessary and sufficient condition

We are now looking for the necessary and sufficient condition for duration to enhance risk in the standard portfolio problem. Consider any specific z. Without loss of generality, assume $\alpha(z) = 1$, i.e., $E\widetilde{x}_1 u'(z+\widetilde{x}_1) = 0$. We have to determine under what conditions

$$\frac{-v''(z)}{v'(z)} = -\frac{Eu''(z+\widetilde{x}_1)}{Eu'(z+\widetilde{x}_1)} + \frac{[E\widetilde{x}_1 u''(z+\widetilde{x}_1)]^2}{E\widetilde{x}_1^2 u''(z+\widetilde{x}_1)Eu'(z+\widetilde{x}_1)} \leq \frac{-u''(z)}{u'(z)}$$

for all z and $\widetilde{x}_1$ satisfying $E\widetilde{x}_1 u'(z+\widetilde{x}_1) = 0$. Let F denote the cumulative distribution function of $\widetilde{x}_1$. Let also define $\tilde{x}$ as the random variable with cumulative distribution function G, with

$$dG(x) = \frac{u''(z+x)dF(x)}{\int u''(z+s)dF(s)}.$$

We verify that dG is positive under risk aversion, and that $\int dG(x) = 1$. Using this change of variable, the problem is rewritten as

$$E\tilde{x}T(z+\tilde{x}) = 0 \implies \frac{1}{T_v(z)} = \frac{1}{ET(z+\tilde{x})} - \frac{(E\tilde{x})^2}{E\tilde{x}^2 ET(z+\tilde{x})} \leq \frac{1}{T(z)}. \tag{10.7}$$

Notice that we derive from the characterization of T_v above that it is nonnegative: duration does never transform an agent with a concave utility function into a risk-lover. This is seen by observing that $(E\widetilde{x})^2 \leq E\widetilde{x}^2$, yielding

$$\frac{1}{T_v(z)} = \frac{1}{ET(z+\tilde{x})} - \frac{(E\tilde{x})^2}{E\tilde{x}^2 ET(z+\tilde{x})} \geq 0.$$

Let us rewrite property (10.7) as

$$E\tilde{x}T(z+\tilde{x}) = 0 \implies T(z)\left[1 - \frac{(E\tilde{x})^2}{E\tilde{x}^2}\right] \leq ET(z+\tilde{x}). \tag{10.8}$$

Our main result is the consequence of the two following Lemmas.

Lemma 4 *Condition (10.8) holds for any $\widetilde{x}$ with a two-point support if and only if the absolute risk tolerance of u is convex.*

Proof: Let us assume that $\widetilde{x}$ is distributed as $(x_-, p; x_+, 1-p)$, with $x_- < 0 < x_+$. We use the following notation:

$$T_- = T(z+x_-) \qquad T_0 = T(z) \qquad T_+ = T(z+x_+) \tag{10.9}$$

According to condition (10.8), we have to verify that

$$\Gamma = T_0\left[1 - \frac{[px_- + (1-p)x_+]^2}{px_-^2 + (1-p)x_+^2}\right] - [pT_- + (1-p)T_+] \tag{10.10}$$

is nonpositive whenever

$$px_-T_- + (1-p)x_+T_+ = 0. \tag{10.11}$$

Eliminating p from (10.10) by using equation (10.11) and reorganizing the expression yields

$$\Gamma = \frac{-T_-T_+(x_+ - x_-)}{(x_+T_+ - x_-T_-)(x_+T_- - x_-T_+)}\left[-T_0(x_+ - x_-) - x_-T_+ + x_+T_-\right] \tag{10.12}$$

Since the fraction in the right-hand side of this expression is negative, Γ is non-positive if and only if

$$-T_0(x_+ - x_-) - x_-T_+ + x_+T_- = (x_+ - x_-)\left[-T_0 + \frac{x_+}{x_+ - x_-}T_- + \frac{-x_-}{x_+ - x_-}T_+\right] \tag{10.13}$$

is positive. This is equivalent to require that

$$T(\lambda y_+ + (1-\lambda)y_-) \leq \lambda T(y_+) + (1-\lambda)T(y_-) \tag{10.14}$$

where $\lambda = \dfrac{-x_-}{x_+ - x_-}$, $y_+ = z + x_+$ and $y_- = z + x_-$. Since this must be true for any λ, y_+ and y_-, the necessary and sufficient condition is that T be a convex function.■

When returns follow a binary process with just one ”up” state and one ”down” state, the convexity of absolute risk tolerance is enough to guarantee that younger people should invest more in stocks. However, the assumption that $\widetilde{x}$ is binary is not satisfactory, and we want to obtain a condition that does not rely on any limitation on the distribution of returns. In the next Lemma, we show that just looking at binary distributions is not enough to guarantee that the result presented above holds for any distribution. We must go one step further.

Lemma 5 *Condition (10.8) holds for any $\tilde{x}$ if it holds for any $\tilde{x}$ with a three-point support.*

Proof: The structure of this Lemma is the same as for Lemma 1. We look for the characteristics of the $\tilde{x}$ which is the most likely to violate condition (10.8). To do so, we solve the following problem for any scalar λ:

$$\begin{array}{ll} \max_{dG\geq 0} & T(z)(1-\lambda\int x dG(x)) - \int T(z+x)dG(x) \\ \text{subject to} & \int xT(z+x)dG(x) = 0; \\ & \int x(1-\lambda x)dG(x) = 0; \\ & \int dG(x) = 1. \end{array} \tag{10.15}$$

If the solution of this problem for any λ is negative, condition (10.8) would be proven. Observe that the above problem is a standard linear programming problem. Therefore the solution contains no more than three x that are such that

$dG(x) > 0$, since they are three equality constraints in the program. Thus, condition (10.8) would hold for any $\tilde{x}$ if it holds for any $\tilde{x}$ with three atoms. ■

Going from two-point supports to three-point supports raises an interesting difficulty. Take any three-point $\tilde{x}$ that satisfies the first-order condition $E\widetilde{x}T(z + \widetilde{x}) = 0$. It is easy to check that we can find a pair of two-point random variables $(\widetilde{y}_1, \widetilde{y}_2)$ and a probability $\lambda \in [0, 1]$ such that $\widetilde{y}_i$ satisfies the first-order condition $E\widetilde{y}_i T(z + \widetilde{y}_i) = 0, \ \ i = 1, 2,$ and $\widetilde{x}$ is the compound lottery $(\lambda, \widetilde{y}_1; 1 - \lambda, \widetilde{y}_2)$. Suppose that absolute risk tolerance be convex. By Proposition 1, it implies that

$$T(z)\left[1 - \frac{(E\tilde{y}_i)^2}{E\tilde{y}_i^2}\right] \leq ET(z + \tilde{y}_i), \tag{10.16}$$

for $i = 1, 2$. Observe that

$$\lambda ET(z + \tilde{y}_1) + (1 - \lambda)T(z + \tilde{y}_2) = ET(z + \tilde{x}) \tag{10.17}$$

The expectation operator is linear in probabilities. Thus, the weighted sum of the right-hand side of inequalities (10.16) for $i = 1$ and $i = 2$ equals the right-hand side of inequality (10.8). We would be done if the same operation could be applied with the left-hand sides, i.e., if the left-hand sides would also be the expected value of a function of the random variable. This is not the case, and that explains why we cannot extend the result from two-point supports to three-point supports and, therefore, to all random variables. It happens that $(E\widetilde{x})^2/E\widetilde{x}^2$ is a convex function of the vector of probabilities of $\widetilde{x}$. It implies that [2]

[2] The proof of this claim is obtained by easy manipulations of (10.18). Denoting $a_i = E\tilde{y}_i$ and $b_i = E\tilde{y}_i^2 > 0$, condition (10.18) is equivalent to

$$2a_1a_2 \leq a_1^2\frac{b_2}{b_1} + a_2^2\frac{b_1}{b_2}.$$

This is in turn equivalent to

$$\left[a_1\sqrt{\frac{b_2}{b_1}} + a_2\sqrt{\frac{b_1}{b_2}}\right]^2 \geq 0,$$

which is always true.

$$\lambda\frac{(E\tilde{y}_1)^2}{E\tilde{y}_1^2}+(1-\lambda)\frac{(E\tilde{y}_2)^2}{E\tilde{y}_2^2}\geq\frac{(E(\lambda\tilde{y}_1+(1-\lambda)\tilde{y}_2))^2}{E(\lambda\tilde{y}_1^2+(1-\lambda)\tilde{y}_2^2)}=\frac{(E\tilde{x})^2}{E\tilde{x}^2}. \tag{10.18}$$

¿From conditions (10.16) and (10.18), we conclude that the convexity of absolute risk tolerance implies that

$$T(z)\left[1-\left\{\lambda\frac{(E\tilde{y}_1)^2}{E\tilde{y}_1^2}+(1-\lambda)\frac{(E\tilde{y}_2)^2}{E\tilde{y}_2^2}\right\}\right]\leq ET(z+\widetilde{x}) \tag{10.19}$$

for all three-point $\tilde{x}$ that is distributed as $(\tilde{y}_1,\lambda;\tilde{y}_2,1-\lambda)$, with $E\widetilde{y}_iT(z+\widetilde{y}_i)=0$. But it does not imply condition (10.8) that is more demanding. This is confirmed by the following counterexample. Take $T(w)=w^{-2}$, which is convex, and $z=1$. Let $\widetilde{x}$ be distributed as $(-0.5,0.10659;1,0.83266;10,0.06075)$.[3] It is easily verified that $E\widetilde{x}T(z+\widetilde{x})=0$, but

$$T(z)\left[1-\frac{(E\tilde{x})^2}{E\tilde{x}^2}\right]=0.723>0.635=ET(z+\tilde{x}).$$

Condition (10.8) is not satisfied for this three-point distribution, although absolute risk tolerance is convex.

This approach also provides a positive result. We have seen that the convexity of $(E\widetilde{x})^2/E\widetilde{x}^2$ with respect to the vector of probabilities acts in the opposite direction when considering the condition that younger people be less risk-averse. It goes in the good direction when considering the condition that younger people be *more* risk-averse, which holds if

$$E\tilde{x}T(z+\tilde{x})=0\implies T(z)\left[1-\frac{(E\tilde{x})^2}{E\tilde{x}^2}\right]\geq ET(z+\tilde{x}). \tag{10.20}$$

Decompose again the three-point $\widetilde{x}$ into two binary $\widetilde{y}_i$ that satisfies the first-order condition. Assuming that T is concave, Proposition 1 implies condition (10.16) with the inequality reversed. It implies in turn condition (10.19) with the inequality reversed. Combining this with condition (10.18) and Lemma 5 yields the following Proposition.

[3]These probabilities are the solution to program (10.15) when G has its support in (-0.5,1,10).

Proposition 33 *Consider the two-period investment problem with a zero yield risk-free asset and another risky asset. Young investors are more risk-averse than old investors if the absolute risk tolerance of final wealth is concave.*

To sum up, the concavity of absolute risk tolerance is sufficient for younger people to purchase less of the risky asset. The age of the investor does not influence the optimal portfolio composition when absolute risk tolerance is linear. But the convexity of absolute risk tolerance does not implies that younger people purchase more of the risky asset, as confirmed by the counter-example!

The above propositions provide qualitative results. We would like to know the quantitative magnitude of the duration or age effect on risk taking, and could determine that if we knew the first four derivatives of the utility function. For now, consider an illustration for the case of $u(z) = z + \ln z$, with $\tilde{s} = (-1, 2; 1/2, 1/2)$. After some tedious computations, we get $T(5) = 30.0$, whereas $T_v(5) = 64.1$: the young investor is more than twice as risk-tolerant than the old investor. This implies that if the expected excess return of the risky asset is small, the young will invest twice as much in it as will the old!

10.3 Discussion of the results

10.3.1 Convex risk tolerance

Convex absolute risk tolerance is equivalent to

$$A''(z) \leq 2\frac{[A'(z)]^2}{A(z)}. \tag{10.21}$$

If absolute risk aversion is concave, absolute risk tolerance is automatically convex. But under the familiar condition of decreasing A, the concavity of A is not plausible, since a function cannot be positive, decreasing and concave everywhere. Condition (10.21) means in fact that absolute risk aversion may not be too convex. There is no obvious relationship between the necessary and sufficient condition (10.21) and the necessary condition for risk vulnerability, which is $A'' \geq 2A'A$. In order to relate (10.21) to standardness, one can easily verify that convex risk tolerance is equivalent to the condition that function

$$\phi(z) \equiv \frac{P(z)}{A(z)}, \tag{10.22}$$

be decreasing in z, where P denotes absolute prudence. This condition should be related to standardness which is characterized by the condition that both A and P are decreasing.

10.3.2 Positive risk free rate and intermediary consumption

Thus far we have assumed that the investor's utility function applies solely to terminal wealth; he has a pure investment problem. In real world contexts, investors consume a portion of their lifetime wealth each period. Moreover, the risk free rate is in general positive. As observed in section **??**, allowing for intermediate consumption makes young people potentially more willing to take risks than in the pure investment problem because current risks can be attenuated by spreading consumption over time. It is easily checked by combining Propositions 51 and **??** that, with complete markets, duration enhances risk when the risk free rate is nonnegative and intermediary consumption is allowed if and only if T is convex and subhomogeneous.

10.3.3 Non-differentiable marginal utility

Throughout this analysis we have assumed that utility was twice differentiable. Absent this property, absolute risk tolerance may not be defined correctly and the above theory is no longer relevant. We now show an example of a utility function that is not twice differentiable and that yields the property that young people are less risk-averse than older ones. Take

$$u(z) = \min(z, (1-t)(z-D) + D), \tag{10.23}$$

with $t \in [0, 1]$ and any scalar D. This function is continuous, piecewise linear and concave. It is drawn in Figure 10.1. Consider the standard portfolio problem with no intermediate consumption, $\rho = 1$ together with $\widetilde{x}_1 = (x_-, x_+; 1/2, 1/2)$, $x_- < 0 < x_+$ and $0.5(x_- + x_+) > 0$. It can be shown that a bounded solution exists for this problem with utility function (10.23) if and only if

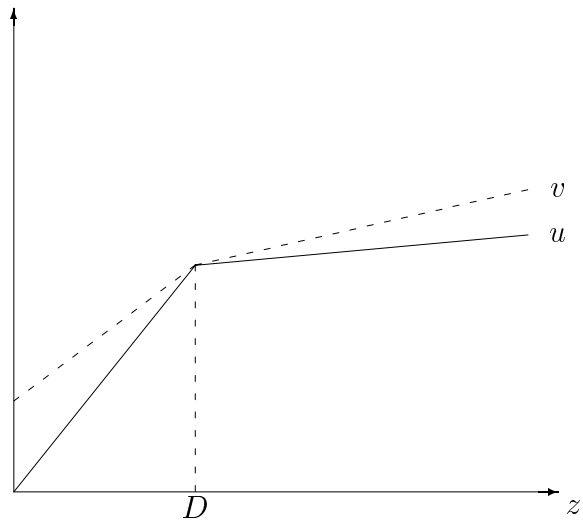

Figure 10.1: The value function with a piecewise utility function

$$t > \frac{x_- + x_+}{x_+}.$$

Under this condition, the optimal demand for the risky asset equals

$$\alpha(z) = \begin{cases} \dfrac{D-z}{x_+} & \text{if } z < D, \\ \dfrac{z-D}{\mid x_- \mid} & \text{if } z \geq D. \end{cases}$$

The value function is then written as

$$v(z) = \min\left[\frac{x_+ - x_-}{2x_+}z + D\frac{x_+ + x_-}{2x_+}, (1-t)\frac{x_+ - x_-}{2 \mid x_- \mid}(z-D) + D\right],$$

whose graph is depicted in Figure 10.1 .

The value function is obviously less concave than the utility function in the usual sense: v is a convex transformation of u. Young investors purchase more

of the risky asset than old investors. However, the absolute risk tolerance is not convex in the proper sense. Notice also that v is piecewise linear as is u, so that the argument can be reproduced recursively over more than two periods.

10.4 Background risk and time horizon

Up to now, we assumed that investors are exposed only to their portfolio risk. It is interesting to examine how the time horizon affects the demand for stocks when investors also face a risk on their human capital. In the first part of this section, we assume that this independent risk occurs only once, at the time of retirement. We also look at the case when the flow of labor incomes is a random walk.

10.4.1 Investors bear a background risk at retirement

The two-period model that we consider here is exactly as before except for the fact that the investor bears a background risk $\widetilde{y}$ when he is old. No such risk is borne when young. We know that this risk will have an adverse effect on the demand for stock from old investors under risk vulnerability.[4] We want to determine how this risk influences the relationship between time horizon and portfolio risk. We know that the answer to this question depends upon the concavity or convexity of the absolute risk tolerance of the following interim indirect utility function:

$$h(z) = Eu(z + \widetilde{y})$$

The agent having a utility function u and facing risk $\widetilde{y}$ at retirement will have the same optimal dynamic portfolio strategy than agents having a utility function h and no such risk. The absolute risk tolerance of h equals

$$T_h(z) = -\frac{Eu'(z + \widetilde{y})}{Eu''(z + \widetilde{y})}$$

when evaluated at z. Differentiating with respect to z yields

[4] It is not necessarily the case that it also reduces the demand from young investors. To obtain such a result, we need to check that v_2 is more risk-averse than v_1 whenever u_2 is more risk-averse than u_1, where v_i is defined as in (10.1). See Roy and Wagenvort (1996) for a counter-example.

$$T_h'(z) = \frac{Eu'(z+\widetilde{y})Eu'''(z+\widetilde{y})}{\left[Eu''(z+\widetilde{y})\right]^2} - 1.$$

It implies that T_h is concave if

$$\begin{aligned} K(z) &= \left[Eu''(z+\widetilde{y})\right]^2 Eu'''(z+\widetilde{y}) + Eu'(z+\widetilde{y})Eu''(z+\widetilde{y})Eu''''(z+\widetilde{y}) \\ &\quad -2Eu'(z+\widetilde{y})\left[Eu'''(z+\widetilde{y})\right]^2 \end{aligned}$$

is nonnegative. We are not aware of any sufficient condition for these properties to hold other than the ones presented in the next Proposition.

Proposition 34 *Suppose that the utility function u is HARA and that the background risk $\tilde{y}$ is small. Then, the interim indirect utility function $h(.) = Eu(.+\widetilde{y})$ has a concave absolute risk tolerance. Longer time horizons reduce the demand for stocks.*

Proof: We have to prove that K is nonnegative. Without loss of generality, we suppose that the expectation of $\widetilde{y}$ is zero. Let $\widetilde{y}$ be distributed as $k\widetilde{s}$, with $E\widetilde{s} = 0$ and $Var(\widetilde{s}) = 2$. Then, a second order Taylor expansion around z yields

$$Eu^{(n)}(z+k\widetilde{s}) = u^{(n)}(z) + k^2u^{(n+2)}(z) + o(k^3).$$

It implies that

$$\begin{aligned} K(z) &= (u^{(2)}(z)+k^2u^{(4)})^2(u^{(3)}(z)+k^2u^{(5)}) \\ &\quad +(u^{(1)}(z)+k^2u^{(3)})(u^{(2)}(z)+k^2u^{(4)})(u^{(4)}(z)+k^2u^{(6)}) \\ &\quad -2(u^{(1)}(z)+k^2u^{(3)})(u^{(3)}(z)+k^2u^{(5)}) + o(k^3), \end{aligned}$$

or, equivalently,

$$K(z) = a_0 + a_1k^2 + o(k^3).$$

It is easy to verify that $a_0 = 0$ (this is due to the linearity of T) and

$$a_1 = (u^{(2)})^2 u^{(5)} + 3u^{(2)}u^{(3)}u^{(4)} + u^{(1)}(u^{(4)})^2 + u^{(1)}u^{(2)}u^{(6)} - 2(u^{(3)})^3 - 4u^{(1)}u^{(3)}u^{(5)}$$

where $u^{(n)} = u^{(n)}(z)$. Following the assumption that preferences are HARA, suppose that u is defined by condition (22.5). Let x denote $a + \frac{z}{\gamma}$. It implies that

$$u^{(1)} = x^{-\gamma};\ \ u^{(2)} = -x^{-\gamma-1};\ \ u^{(3)} = \frac{\gamma+1}{\gamma}x^{-\gamma-2};\ \ u^{(4)} = \frac{(\gamma+1)(\gamma+2)}{\gamma^2}x^{-\gamma-3}; \ldots$$

After some tedious manipulations, we obtain that

$$a_1 = \frac{4(\gamma+1)x^{-3\gamma-6}}{\gamma^4} > 0.$$

We conclude that $K(z)$ is positive if k is small enough. ■

Thus, the existence of a small idiosyncratic risk tends to bias the effect of time horizon in favor of more conservative portfolios. This is the good news of this exercise. However, there are two bad news. The first one is that this result is not true for background risk that are not small. This is shown by the following example: take a CRRA utility function with $\gamma = 4$. In Figures 10.2 and 10.3 , we have drawn T_h' when background risk is distributed as $(-k, 1/2; k, 1/2)$. When $k = 1$, T_h' is uniformly decreasing, implying that the indirect utility function has a concave absolute risk tolerance. But for the larger background risk with $k = 5$, it happens that the absolute risk tolerance is first convex and then concave. This second graph thus provides a counterexample to the above Proposition.

The other bad news is that the effect of background risk on the relation between time horizon and portfolio risk is at best a second order effect under HARA. The following numerical simulation is an illustration of this fact. As above, let us consider a CRRA utility function with a relative risk aversion $\gamma = 4$. Suppose that the background risk at retirement is $\widetilde{y} \sim (-1, 1/2; +1, 1/2)$ and that the return on stocks at each period is $\widetilde{x} \sim (-1, 1/2; 2, 1/2)$. Take a wealth level $z = 5$. If there is only one period to go, the optimal investment in the stock is $\alpha = 0.251693$. It goes down when more time remains before retirement. However, the effect is hardly observable since the optimal demand is $\alpha = 0.251477$ with six periods to go.

Figure 10.2: T_h' when $\widetilde{y} \sim (-1, 1/2; 1, 1/2)$

Figure 10.3: T_h' when $\widetilde{y} \sim (-5, 1/2; 5, 1/2)$

Figure 10.4: Optimal demand for stocks as a function of wealth. The highest curve is for the last period before retirement. The lowest curve is for six period before retirement.

10.4.2 Stationary income process

In the real world, the risk on human capital takes the form of recurrent shocks on labor incomes. The simplest way to model this is to assume that, at each period t, investors hold a lottery ticket $\widetilde{y}_t$ that is actuarially fair. Assume also that $\widetilde{y}_0$,$\widetilde{y}_1, \ldots$ are i.i.d. and are independent of the returns of capital investments. We see that an additional element comes into the picture now: younger people bear a larger risk on their human capital than older agents who already observed some of the shocks on their lifetime labor incomes. The general idea is that if utility is vulnerable to risk , the larger risk on human capital increases aversion to other independent risks, this will inhibit risk taking while young.

It is difficult to extract the qualitative effect of time horizon on portfolio risk when a recurrent background risk is present. We hereafter present a numerical estimation of this model. We took a CRRA utility function with $\gamma = 2$. At each period, the investor faces a random shock on his labor income which takes the form $\widetilde{y}_t \sim (-1, 1/2; +1, 1/2)$. The per period rate of return on stocks in $\widetilde{x}_t \sim (-1, 1/2; 1.2, 1/2)$. In Figure 10.4, we draw the optimal demand $\alpha_t(z)$ for stocks as a function of accumulated wealth for different time horizon. The main feature of this figure is that the demand for stocks goes down sharply as retirement is more distant, at least for low wealth levels. Risk vulnerability is the main reason for this effect.

10.5 Final Remark

This chapter has been devoted to the characterization of efficient dynamic portfolio strategies. To do that, we answered the following question: how does the presence of an option to reinvest in a portfolio in the future affect the optimal portfolio today? By solving this problem, we have been able to adapt the static portfolio problem to its intrinsically intertemporal nature. This dynamic nature does not dramatically transform the portfolio strategy. This is particularly true if the utility function is HARA, in which case the sequence of optimal static strategies, i.e., myopia, are dynamically optimal.

There is a direct application of this result for the equity premium puzzle. In Chapter 6, we calculated the equity premium based on the assumption that investors live just for one year. Under HARA, this assumption has no effect on the resulting equity premium. But if absolute risk tolerance is concave, taking into account of the longevity of investors will make them more risk-averse than assumed in the static one-year model, thereby it will increase the equity premium.

These results do not take into account of the fact that younger people usually bear a larger risk on their human capital. Under risk vulnerability, this is a strong motive for younger investors to purchase a more conservative portfolio.

Chapter 11

Special topics in dynamic finance

The standard portfolio model and the complete markets model represent an idealization of the real world. They do not take into account the presence of transaction costs and of constraints on financial strategies available to investors. The effects of these imperfections are in general trivial when considering a static model. They are extremely complex in dynamic models. Our aim in this chapter is to explore the effect of these imperfections on the optimal dynamic portfolio strategies.

11.1 The length of periods between trade

Most dynamic strategies entail a rebalacement of the portfolio composition at each date, almost surely. Even with a CRRA utility function where the investor follows the simple strategy to maintain constant the relative share of the portfolio value that is invested in the risky asset, the investor must trade whenever the excess return of the risky asset differs from zero. He must sell some of it if the excess return is positive, whereas he must buy some of it if the excess return has been negative.[1] Thus, we should observe that households are active on financial markets at every instant where relative prices of different securities diverge. That is, in fact, at every instant. This is obviously not the kind of investment behaviour that we observe in the real world. One reason is that, contrary to what we assumed, transaction costs prevail on financial markets. Investors must pay transaction costs every time that they want to rebalance their portfolio. The most obvious illustration of such costs is the bid-ask spread, i.e. the fact that the price at which one can sell an asset

[1] We suppose here that there is no dividend. The return of an asset comes from an increase in its value.

is smaller than the price at which one can purchase it. One should also take into account of the time/cost spent to get financial information, to learn prices and to submit an order to the market. Sometimes, distorting taxes prevail that limit the possibility to trade frequently. Most of these costs have a fixed component and a variable component. The presence of transaction costs will induce trading only if the benefit to trade exceeds its cost.

The effect of transaction costs on the *static* optimal portfolio strategy is in general trivial. If there is a fixed cost to invest in risky assets, it may be optimal to be 100% in the risk free asset, although $E\widetilde{x} > 0$. The variable component of transaction costs may be analyzed as a leftward shift in the distribution of the return of the risky asset. In an intertemporal framework, the effect of a transaction cost is less obvious to tackle. In short, knowing that one will have to pay cost to rebalance one's portfolio in the future may have an impact on the optimal structure of the portfolio today. In this section, we consider extreme case where there is an infinite cost to rebalance portfolios in the future. A good motivation for examining the infinite case is that it answer to the following question: what is the impact of the length of periods between trade on the optimal portfolio risk? To be more precise, we consider the model with three dates and two assets that has been examined in the previous chapter.

In this analysis, we examine more generally the relationship between flexibility – we label its opposite rigidity – and the risk of portfolios optimized to maximize expected utility. Common wisdom suggests that flexibility enhances risk. The strong form of rigidity arises when the investor is not able to modify his exposure to risk from period to period. We hereafter examine a somewhat weaker form of rigidity, i.e., the investor can modify his risk exposure from one period to another, but he is forced to fix it a priori, i.e., before observing realizations of the early returns on his investment. The individual invests at date 0, receives investment returns, and then must invest at date 1. In the rigid economy, date 1 investment decisions must be made before date 0 results are learned. The problem is written as follows:

$$\alpha^r \in \arg\max_{\alpha} \quad \max_{\beta} \quad Eu(w_0 + \alpha\widetilde{x}_0 + \beta\widetilde{x}_1)$$

where $\widetilde{x}_t$ is the return of the risky asset between t and $t+1$. As before, we assume that the returns of the risky asset follows a random walk, i.e., that $\widetilde{x}_0$ and $\widetilde{x}_1$ are independent. Observe that the decision β on the exposure on $\widetilde{x}_1$ must be taken simultaneously with α. The optimal exposure to risk $\widetilde{x}_0$ is denoted α^r, where

the upper index refers to the rigid economy. This problem has been examined in section 9.2.

We want to compare α^r to the optimal investment in $\widetilde{x}_0$ when the investors live in the flexible economy, i.e., when he may determine β after the observation of $\widetilde{x}_0$. The investment problem in the flexible economy is written as

$$\alpha^f \in \arg\max_{\alpha} \quad Ev(w_0 + \alpha\widetilde{x}_0) \quad \text{with} \quad v(z) = \max_{\beta} \quad Eu(z + \beta\widetilde{x}_1).$$

This problem is the one that we already examined in section 10.2. The intuition suggests that the opportunity to rebalance one's portfolio in the future contingent to the realization of the first period portfolio should induce more risk-taking: $\alpha^f \geq \alpha^r$. Let us compare α^r and α^f to the solution of the myopic problem:

$$\alpha^m \in \arg\max_{a} \quad Eu(w_0 + \alpha\widetilde{x}_0).$$

¿From Proposition 30, we know that α^r is less than α^m if absolute prudence is decreasing and larger than twice the absolute risk aversion. Under these conditions, $\widetilde{x}_1$ is substitute to $\widetilde{x}_0$. If the utility function is HARA, this is equivalent to the condition that γ be less than 1. Moreover, we know that for the class of HARA functions, α^m equals α^f: myopia is optimal in that case. It yields the following result, which is due to Gollier, Lindsey and Zeckhauser (1997).

Proposition 35 *Suppose that the utility function is HARA, with $\gamma \leq 1$. Then, α^f is larger than α^r, i.e., flexibility enhances risk.*

We can extend this Proposition by using sufficient conditions for α^m being smaller than α^f. These sufficient conditions are in Proposition 32.

Proposition 36 *Suppose that absolute prudence is decreasing and larger than twice the absolute risk aversion. Suppose also that absolute risk aversion is concave. Then, α^f is larger than α^r, i.e., flexibility enhances risk.*

The sufficient conditions presented in this Proposition are in fact sufficient for the stronger property that $\alpha^r \leq \alpha^m \leq \alpha^f$. We see that there are two effects of transforming a rigid economy into a more flexible one. These effects are obtained by decomposing the change via the single-period, myopic economy. Going from the rigid economy to the myopic one introduce a substitution effect. It implicitly

eliminates the opportunity to invest in asset $\widetilde{x}_1$, which increases the demand for asset $\widetilde{x}_0$ if they are substitutes. The second effect is obtained by moving now from the single-period economy to the flexible one, which raises the duration of the investment. This has a positive effect on risk-taking if duration enhances risk.

We now go back to the initial motivation of this analysis, which was the analysis of the effect of the length of periods between trade on the optimal portfolio composition. If we assume that $\widetilde{x}_0$ and $\widetilde{x}_1$are identically distributed, then, by using Proposition 10, we get that

$$\arg\max_{\alpha} \quad \max_{\beta} \quad Eu(w_0 + \alpha\widetilde{x}_0 + \beta\widetilde{x}_1) \quad = \quad \arg\max_{\alpha} \quad Eu(w_0 + \alpha(\widetilde{x}_0 + \widetilde{x}_1)).$$

This is due to the fact that the optimal values of α and β coincide when the two risk are identically distributed. The problem to the right is an investment problem where the investor may not rebalance his portfolio at date 1. Its solution α^r is thus interpreted now as the optimal investment in the risky asset when the period between trade can take place is between $t = 0$ and $t = 2$. Financial markets are closed at date $t = 1$. Comparing α^r to α^f is the relevant question to determine the effect of the length of periods between trade. The following result is a straightforward reinterpretation of the above result.

Corollary 4 *Suppose that the return of the risky asset follows a stationary random walk. Under the conditions of Proposition 36, reducing the length of periods between trade raises the demand for the risky asset.*

11.2 Dynamic discrete choice

In this section, we examine the discrete choice version of the model presented in section 10.2. We analyze the effect of an option to gamble on a lottery with payoff $\tilde{x}_1$ in the future on the decision to gamble today on a lottery with payoff $\tilde{x}_0$. The important restriction that is imposed here is that the investor can either accept or reject the lottery. This is a take-it-or leave-it problem. Contrary to the model presented in section 10.2 where the agent could determine the size of his exposure to risk, the size of the risk is exogenous in this section. An illustration of this problem is when an entrepreneur faces a sequence of real investment projects. It is often the case that the entrepreneur has no choice about the size of the investment; he can just do it or not, due to the presence of large fixed costs to implement the

project. The attitude towards the current lottery is characterized by the degree of concavity of the value function v defined as:

$$v(z) \;=\; \max_{\alpha \in K} \; Eu(z + \alpha \tilde{x}_1), \tag{11.1}$$

where $K \equiv \{0, 1\}$ is the set of acceptable values for the decision variable.

The value function is the upper envelope of two concave functions of z, respectively $u(z)$ and $Eu(z + \widetilde{x}_1)$. Bell (1988a), considering the same specific model, showed that v is not concave, except in the trivial case in which $\tilde{x}_1$ is never (or always) desirable. We reproduce here his figure (Figure 11.1) in the specific case of decreasing absolute risk aversion.[2] Function v is locally convex at z_0 where $u(.)$ crosses $Eu(. + \tilde{x}_1)$. The intuition of this result comes from the option-like payoff that is generated when the gambler gets the opportunity to bet on $\tilde{x}_1$. This is shown by rewriting v as

$$v(z) = u(z + \max(0, CE(z))), \tag{11.2}$$

where CE is the certainty equivalent of $\tilde{x}_1$, i.e. $u(z + CE(z)) = Eu(z + \tilde{x}_1)$. The non-concavity of v implies that the gambler could rationally accept unfair lotteries at date 0. As it is intuitively sensible, "an entrepreneur with an idea she believes will work, but without financial backers who agree, would be justified trying to raise the necessary capital in Atlantic City"(Bell 1988).

We conclude from this discussion that v may never be more concave than u in all cases where the choice on $\widetilde{x}_1$ depends upon the wealth z of the agent. Is it possible that v be globally less risk-averse than u? The answer is no, at least for proper utility functions. The existence of an option to gamble on $\tilde{x}_1$ may induce gamblers to reject some lotteries $\tilde{x}_0$ that would have been accepted in the absence of an option.

Proposition 37 *Suppose that $\tilde{x}_1$ is desirable at some levels of wealth and undesirable at other levels. If risk aversion is proper, indirect utility function $v(.)$ defined by equation (11.1) is not a convex transformation of u. That is, there exist some lotteries $\tilde{x}_0$ that are accepted in the absence of an option to gamble on $\tilde{x}_1$, but that are rejected if such an option is offered.*

[2] Use condition (3.14) with $\pi = 0$ to prove that, under DARA, curve $u(z)$ crosses $Eu(z + \widetilde{x}_1)$ from above.

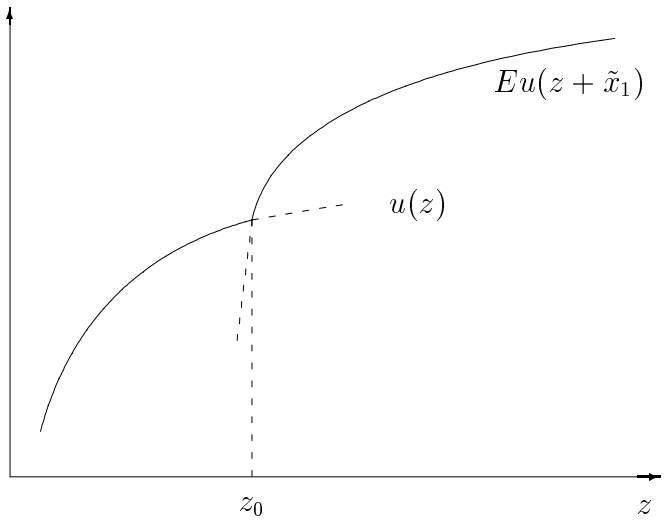

Figure 11.1: The value function is the upper envelope of $u(.)$ and $Eu(. + \tilde{x}_1$.

Proof: Define $m(w) = Eu(w + \tilde{x}_1)$. Because DARA is necessary for properness, there exists a unique z_1 such that $m(z_0) = u(z_0)$. We first show that $-m''(z_0)/m'(z_0)$ is larger than $-u''(z_0)/u'(z_0)$. Because u is proper and $Eu(z_0 + \tilde{x}_1) = u(z_0)$, we get that, for any $\tilde{y}$,

$$Eu(z_0 + \tilde{y}) \leq u(z_0) \implies Eu(z_0 + \tilde{x}_1 + \tilde{y}) \leq Eu(z_0 + \tilde{x}_1),$$

or, equivalently,

$$Eu(z_0 + \tilde{y}) \leq u(z_1) \implies Em(z_0 + \tilde{y}) \leq m(z_1).$$

At z_0, m rejects all lotteries that u rejects: m is more diffident than u around z_0. A necessary condition is that $-m''(z_0)/m'(z_0)$ is larger than $-u''(z_0)/u'(z_0)$. By definition of z_0, m and u must cross at z_0. From definition (3.14) of DARA, m crosses u upwards. Thus, v is locally more concave than u to the right of z_0. ■

In order to illustrate Proposition 37, suppose that the gambler with initial wealth $w_0 = 2.1$ and utility function $u(w) = ln(w)$ has the opportunity to bet on $\tilde{x}_0 = (-0.1, 0.107; 1/2, 1/2)$. One can easily verify that it is optimal for him

to bet. Suppose now that in addition to the opportunity to bet on $\tilde{x}_0$ today, the gambler knows that he could bet tomorrow on $\tilde{x}_1 = (-1, 2; 1/2, 1/2)$. In that case, it is not optimal anymore to bet on $\tilde{x}_2$. The option to bet in the future makes some current desirable risks undesirable.

When u is proper, the absolute risk aversion index for v is decreasing everywhere except at z_1 where an upwards jump occurs. Thus, the value function defined by (11.1) does not inherit the property of DARA from the utility function u. This jump in risk aversion at the wealth level where the option value becomes positive may yield counter-intuitive comparative statics effects. To illustrate, let us consider again the logarithmic gambler facing the current lottery $\tilde{x}_0 = (-0.1, 0.107; 1/2, 1/2)$ and the option to bet on $\tilde{x}_1 = (-1, 2; 1/2, 1/2)$. It is easy to verify that $z_0 = 2$. It is noteworthy that there is more than one critical wealth level in the first period at which the gambler switches from rejecting to accepting lottery $\tilde{x}_0$.[3] One can verify that the gambler will not bet on $\tilde{x}_0$ if his wealth at the start of the game is less 1.529 or if it belongs to interval $[2.094, 2.282]$. As wealth increases, the gambler switches thrice by first rejecting, then accepting, then rejecting again and finally accepting the initial bet, despite DARA!

These paradoxes disappear if one assumes that $\widetilde{x}_0$ and $\widetilde{x}_1$ are identically distributed as $\widetilde{x}$. In that case, the opportunity to gamble again on the same risk in the future always induce more risk-taking today. This is seen by proving that

$$Eu(w + \widetilde{x}) \geq u(w) \quad \Longrightarrow Ev(w + \widetilde{x}) \geq v(w). \tag{11.3}$$

Consider an agent with wealth w at $t = 0$. The condition to the left means that if he does not accept to gamble today, it is optimal to gamble tomorrow. The condition to the right means that it is optimal to gamble at $t = 0$. To prove this property, observe that

$$Ev(w + \widetilde{x}_0) = E_0 \max\left[u(w + \widetilde{x}_0), E_1 u(z + \widetilde{x}_0 + \widetilde{x}_1)\right],$$

where E_t is the expectation with respect to $\widetilde{x}_t$. Obviously, it implies that

[3] Bell [1988b] considered the problem of the number of switches to the sign of $h(w) = Eu(w + \tilde{y}_a) - Eu(w + \tilde{y}_b)$, for any couple of random variables $(\tilde{y}_a, \tilde{y}_b)$. When the problem simplifies to accepting or rejecting a lottery with a sure initial wealth, i.e. when $\tilde{y}_b \equiv 0$, DARA is sufficient to yield at most one switch. Thus, the fact that the logarithmic utility function is not "one-switch" in the sense of Bell is not relevant to explain the paradox.

$$Ev(w+\widetilde{x}_0) \geq E_0 u(w+\widetilde{x}_0) = Eu(w+\widetilde{x}).$$

The left condition in (11.3) implies that $v(w) = \max\left[u(w), Eu(w+\widetilde{x})\right] = Eu(w+\widetilde{x})$. Thus, the above inequality implies that $Ev(w+\widetilde{x}) \geq v(w)$. This result is easily extended for a larger sequence of i.i.d. risks. This proves the following Proposition.

Proposition 38 *Consider a sequence of take-it-or-leave-it offers to gamble on i.i.d. lotteries. It is never optimal to postpone the acceptance of a desirable lottery.*

That corresponds to the common sense that one should not "put off till tomorrow what could be done today". The intuition of this result comes from the fact that postponing the acceptance of a desirable risk does nothing else than to reduce the opportunity set of the gambler. In other words, the optimal strategy with t periods remaining can always be duplicated when there are $t+1$ periods remaining. This is done by acting as if the horizon is a period shorter, and by rejecting the last lottery. This strategy is clearly weakly dominated by the strategy of doing as if the horizon be a period shorter, and by doing what is optimal to do at the last period depending upon the state of the world at that time.

The consequence of this result is that once the gambler decides to reject a lottery, he will afterwards never gamble again. Two scenarios are then possible:

(i) the agent never gambles;

(ii) the agent gambles from the first period till some period s, and then never gambles again.

In short, the problem simplifies to a stopping time. Let $\mathcal{S}_t$ be the set of wealth levels for which it is optimal to gamble when there are t lotteries remaining. Proposition 38 states that $\mathcal{S}_1 \subset \mathcal{S}_2 \subset ... \subset \mathcal{S}_T$: the acceptance set is smaller for shorter horizon problems. As the horizon recedes, the gambler is more and more willing to gamble. This gives a potential rationale to the Samuelson [1963]'s lunch colleagues who refused to gamble on a single risk, but who would have accepted to gamble if many such i.i.d. risks were offered. Our interpretation of the game is different from the one proposed by Samuelson in the sense that we do not assume that the player is forced to determine ex-ante the number of i.i.d. gambles. Notice that the intuition of our result is in no way related to the Law

of Large numbers because, as in Samuelson [1963], i.i.d. risks are added through time rather than subdivided.

Gollier (1996) shows that acceptance sets $\mathcal{S}_t$ are of the form $\mathcal{S}_t = \{z \geq z_t\}$ if u is DARA. As suggested by the intuition, wealth must be large enough to gamble. From Proposition 38, it must be that $z_0 \geq z_1 \geq z_2 \geq ...$: the larger the number of option remaining, the smaller is the minimum wealth level at which one starts to accept to gamble. As in the standard static model, a DARA gambler accepts to gamble only if his wealth at the time of the decision is large enough. There is a critical wealth level at every period separating gambling states from no-gambling states. But the existence of one or more options to bet in the future implies that the critical wealth level is reduced. The positive effect of the existence of options implies the counter-intuitive property that a DARA gambler could well decide to stop gambling after the occurrence of a net gain on the previous lottery. The point is that the reduction in risk aversion due to the increase in wealth is not sufficient to compensate for the negative effect of the reduction of the opportunity set.

To illustrate this result, let us consider the case of the logarithmic utility function with $\widetilde{x} \sim (-1; 0.5; 2, 0.5)$. The critical/minimal wealth level when there is no option to gamble again is $z_0 = 2.000$. When there is a single option to gamble remaining, this critical wealth level goes down $z_1 = 1.812$.[4] We also obtain $z_2 = 1.702$, $z_3 = 1.634$,..., $z_{17} = 1.316$.

11.3 Constraints on feasible strategies

The model presented above differs significantly from the one in section 10.2 by the fact that α must be either 0 or 1 in the former case, whereas it could be anything in the latter case. Suppose more generally that the demand for the risky asset at date 1 must be in some subset $K(z)$ of R, where z is the wealth level of the investor at that date. If $K(z)$ is the set of integers, this constraint means that the investor may not purchase fractions of the risky asset. If $K(z)$ is the set of real numbers less than z, this constraint is a no-short-sale constraint stating that the investor cannot borrow at the risk free rate to invest in the risky asset.

For the time being, we are just willing to determine whether the value function

$$v_K(z) = \max_{\alpha \in K(z)} \quad Eu(z + \alpha\widetilde{x}) \tag{11.4}$$

[4] It is obtained by solving $Ev(z_1 + \widetilde{x}) = v(z_1)$.

is concave. We have shown in the previous chapter that it was the case when α is not constrained, i.e., when $K \equiv R$. On the contrary, when $K = \{0, 1\}$ as in the previous section, it was not the case. We provide now a sufficient condition for the value function defined by (11.4) to be concave.

Suppose that $K(z)$ is convex:

$$\begin{aligned} \alpha_1 \quad &\in \quad K(z_1) \text{ and } \alpha_2 \in K(z_2) \\ &\Longrightarrow \quad \forall \lambda \in [0,1] : \lambda\alpha_1 + (1-\lambda)\alpha_2 \in K(\lambda z_1 + (1-\lambda)z_2). \end{aligned} \tag{11.5}$$

Let us show that this condition together with the concavity of u imply that the value function is concave. Let us take any (z_1, z_2) with $z_1 < z_2$, and any $\lambda \in [0, 1]$. By definition, we have that

$$\lambda v_K(z_1) + (1-\lambda)v_K(z_2) = \lambda Eu(z_1 + \alpha_1\widetilde{x}_1) + (1-\lambda)Eu(z_2 + \alpha_2\widetilde{x}_1), \tag{11.6}$$

with

$$\alpha_i \in \arg\max_{\alpha \in K(z_i)} \quad Eu(z_i + \alpha\widetilde{x}).$$

By the concavity of u, applying Jensen's inequality to the right-hand side of condition (11.6) for every realization of $\widetilde{x}_1$ yields

$$\lambda v_K(z_1) + (1-\lambda)v_K(z_2) \leq Eu(\lambda z_1 + (1-\lambda)z_2 + (\lambda\alpha_1 + (1-\lambda)\alpha_2)\widetilde{x}_1). \tag{11.7}$$

Since $\alpha_1 \in K(z_1)$ and $\alpha_2 \in K(z_2)$, the convexity of K implies that $\lambda\alpha_1 + (1-\lambda)\alpha_2$ is in $K(\lambda z_1 + (1-\lambda)z_2)$. It implies that

$$Eu(\lambda z_1 + (1-\lambda)z_2 + (\lambda\alpha_1 + (1-\lambda)\alpha_2)\widetilde{x}_1) \leq \max_{\alpha \in K(\lambda z_1 + (1-\lambda)z_2)} \quad Eu(\lambda z_1 + (1-\lambda)z_2 + \alpha\widetilde{x}_1). \tag{11.8}$$

The right-hand side of the above inequality is nothing but $v_K(\lambda z_1 + (1-\lambda)z_2)$. Combining conditions (11.6), 11.7) and (11.8) yields the result:

$$\lambda v_K(z_1) + (1-\lambda)v_K(z_2) \leq v_K(\lambda z_1 + (1-\lambda)z_2).$$

This means that v_K is concave.

Proposition 39 *The value function* $v_K(z) = \max_{\alpha \in K(z)} \quad Eu(z+\alpha\widetilde{x})$ *is concave if* u *is concave and* K *is convex (property (11.5)).*

This result is a generalization of a result by Pratt (1987). Observe that in the take-it-or-leave-it model with $K(z) = \{0, 1\}$, the condition that K be convex is not satisfied, and the Proposition does not apply.

11.4 The effect of a leverage constraint

In the previous section, we examined whether the value function $v(z) = \max_{\alpha \in K(z)} \quad Eu(z+\alpha\widetilde{x})$ is concave. We now turn to the comparison of the degrees of concavity of u and v. We consider two types of leverage constraints: $\alpha \geq k(z)$ and $\alpha \leq k(z)$.

11.4.1 The case of a lower bound on the investment in the risky asset

Consider the case with

$$K(z) = \{\alpha \mid \alpha \geq k(z)\}$$

for some real-valued continuous function $k(.)$. There is a minimum requirement on the size of the investment in the risky asset. This is the case for example in a planned french law where pension plans will be required to be invested in stocks for at least 50% of their market value. This is also the case in the coinsurance model (see section 5.6) when the insured person must retain a positive deductible on his insurance contract.

Let $\alpha(z)$ be the unconstrained solution of the static portfolio problem with wealth z:

$$E\widetilde{x}_1 u'(z + \alpha(z)\widetilde{x}_1) = 0.$$

Consider a wealth level $\widehat{z}$ where the constraint becomes binding, yielding $k(\widehat{z}) = \alpha(\widehat{z}) = \widehat{\alpha}$. We examine the case where the constraint induces less flexibility, i.e., $k'(\widehat{z})$ is smaller than $\alpha'(\widehat{z})$: the constrained exposure to risk is less flexible to changes in wealth than the optimal exposure. It implies that the constraint is binding for wealth levels slightly smaller than $\widehat{z}$. We obtain that

$$v_K(z) = \begin{cases} Eu(z + k(z)\widetilde{x}_1) & \text{if} \quad z \leq \widehat{z} \\ v(z) = Eu(z + \alpha(z)\widetilde{x}_1) & \text{if} \quad z > \widehat{z} \end{cases}$$

in a neighborhood of $\widehat{z}$. We want to determine the impact of the minimum requirement constraint on the optimal portfolio at date $t = 0$. It is useful to start with the analysis of the degree of concavity of v_K around $\widehat{z}$. It is the same as the degree of concavity of v — the value function of the unconstrained problem — to the right of $\widehat{z}$. For values of z slightly to the left of $\widehat{z}$, we obtain that

$$\left.\frac{-v_K''(z)}{v_K'(z)}\right|_{z=\widehat{z}_-} = -\frac{Eu''(z + \widehat{\alpha}\widetilde{x}_1)}{Eu'(z + \widehat{\alpha}\widetilde{x}_1)} + k'(\widehat{z})\frac{-E\widetilde{x}_1 u''(z + \widehat{\alpha}\widetilde{x}_1)}{Eu'(z + \widehat{\alpha}\widetilde{x}_1)} \tag{11.9}$$

whereas the degree of concavity of the unconstrained value function equals

$$\left.\frac{-v''(z)}{v'(z)}\right|_{z=\widehat{z}_-} = -\frac{Eu''(z + \widehat{\alpha}\widetilde{x}_1)}{Eu'(z + \widehat{\alpha}\widetilde{x}_1)} + \alpha'(\widehat{z})\frac{-E\widetilde{x}_1 u''(z + \widehat{\alpha}\widetilde{x}_1)}{Eu'(z + \widehat{\alpha}\widetilde{x}_1)} \tag{11.10}$$

Remember now that DARA means that $-E\widetilde{x}_1 u''(z + \widehat{\alpha}\widetilde{x}_1)$ is negative. The comparison of (11.9) and (11.10) together with $k'(\widehat{z}) \leq \alpha'(\widehat{z})$ directly implies that

$$\left.\frac{-v_K''(z)}{v_K'(z)}\right|_{z=\widehat{z}_-} \geq \frac{-v''(\widehat{z})}{v'(\widehat{z})}$$

under DARA. We conclude that DARA is necessary and sufficient for the introduction of a minimum requirement on the investment in the risky asset to raise the aversion to risk of the value function in the neighborhood of a critical wealth level where the constraint becomes binding. The fact that such a minimum requirement raises aversion to date 0 risks is a very natural property. It is compatible with the idea that independent risks are substitutes. The fact that an agent is potentially forced to take more risk in the future induces the agent to undertake less risk today.

The above analysis shows that DARA is necessary for the introduction of a minimum requirement to induce more risk aversion. This condition is not sufficient for this result. The comparison of $\frac{-v_K''(z)}{v_K'(z)}$ and $\frac{-v''(z)}{v'(z)}$ is not trivial when they are not evaluated at a critical wealth level. The search for a sufficient condition will be limited now to the case where $k(z)$ is a constant k that is independent of z.

The easiest way to treat this problem is again to decompose the problem into first comparing v_K to u and, second, comparing u to v. We know that v_K is more concave than u if

$$\forall z, \widetilde{x}_1 : \quad E\widetilde{x}_1 u'(z + k\widetilde{x}_1) \leq 0 \quad \Longrightarrow \quad \frac{-Eu''(z + k\widetilde{x}_1)}{Eu'(z + k\widetilde{x}_1)} \geq \frac{-u''(z)}{u'(z)}.$$

The condition to the left means that we are in the binding region. This condition is close to condition (9.1), except that the left condition is an inequality rather than an equality. Using Lemma 2 with $f_1(x) = xu'(z + kx)$ and $f_2(x) = u''(z + kx)u'(z) - u'(z + kx)u''(z)$ yields the following set of necessary and sufficient conditions:

$$\begin{cases} \forall z, x : \quad A(z + x) \geq A(z) + xA'(z) \\ \forall z : \quad -A'(z) \geq 0 \end{cases}$$

Thus, v_K is more concave than u if absolute risk aversion is decreasing and convex. We would be done if u be (weakly) more concave than v. This is the case under HARA.

Proposition 40 *Suppose that u be HARA, with decreasing absolute risk aversion. Then, a minimum requirement on the size of the investment in the risky asset ($\alpha \geq k$) in the future reduces the demand for the risky asset today.*

The restrictive condition HARA can be relaxed by using weaker sufficient conditions for u to be more concave than v. Using Proposition 32, we obtain for example the following result:

Proposition 41 *Suppose that the risk on returns be binary. A minimum requirement on the size of the investment in the risky asset ($\alpha \geq k$) in the future reduces the demand for the risky asset today if absolute risk tolerance is convex and absolute risk aversion is decreasing and convex.*

¿From these three conditions, only DARA is necessary. The two other conditions are probably too strong. The convexity of absolute risk aversion has already been found in Proposition 27 as a condition for independent risks to be substitutes when taken simultaneously.

11.4.2 The case of an upper bound on the investment in the risky asset

We now turn to the case where the investor faces an upper limit constraint $\alpha \leq k(z)$ on his risky investment at date 1.

As before, we first look at the effect of the leverage constraint on the degree of concavity of the value function around a critical wealth level $\widehat{z}$. Again, we examine the case where the constraint induces less flexibility, i.e., $k'(\widehat{z})$ is smaller than $\alpha'(\widehat{z})$. An exercise parallel to the one performed above would show that DARA is necessary and sufficient to guarantee that v_K be more concave than the unconstrained value function around $\widehat{z}$.

The intuition that a constraint limiting the size of the risky investment in the future should induce a reduction in the demand for the risky asset today is not obvious. Indeed, given the idea that independent risks are rather substitutes, the investor could prefer to increase his risky investment today to compensate for the anticipated limitation in the size of the risky investment tomorrow. This effect goes the opposite direction than the flexibility effect mentioned above, under DARA. We may thus expect an ambiguous global effect under assumptions compatible with the substituability of independent risks. To show this, let us again consider the special case with $k(z) = k$. We decompose the global effect of the leverage constraint into a change from v_K to u first, and, then, a change from u to v. The first effect is an increase in risk aversion if

$$\forall z, \widetilde{x}_1 : \quad E\widetilde{x}_1 u'(z + k\widetilde{x}_1) \geq 0 \quad \Longrightarrow \quad \frac{-Eu''(z + k\widetilde{x}_1)}{Eu'(z + k\widetilde{x}_1)} \geq \frac{-u''(z)}{u'(z)}. \tag{11.11}$$

The condition to the right means that the investor with wealth z would like to invest more than k in the risky asset. One could use Corollary 2 to get the necessary and sufficient condition that A be *increasing* and convex. It is more intuitive to consider the trick presented after Proposition 27: let the cumulative distribution function of $\widetilde{y}$ be defined by

$$dG(x) = \frac{u'(z+kx)dF(x)}{\int u'(z+ks)dF(s)}.$$

Condition (11.11) may then be rewritten as:

$$\forall z, \widetilde{y}: \quad E\widetilde{y} \geq 0 \quad \Longrightarrow \quad EA(z+\widetilde{y}) \geq A(z). \tag{11.12}$$

The ambiguity appears then clearly. The Jensen's inequality combined with the assumption that A is convex implies that $EA(z+\widetilde{y})$ is larger than $A(z+E\widetilde{y})$. This goes into the intuitive direction. But DARA plus $E\widetilde{y}$ positive goes the opposite direction, with $A(z+E\widetilde{y}) \leq A(z)$. Suppose in particular that $A''(w)/A'(w)$ tends to zero when w tends to infinity. Then, using standard results on the symmetric concept of risk premium, the substitution equivalent ι of $\widetilde{y}$, defined as

$$EA(z+\widetilde{y}) = A(z+E\widetilde{y}-\iota(z)),$$

tends to zero with z. It implies that condition 11.12 is violated for large values of z.

Notice that conditions $A' < 0$ and $A''/A' \to 0$ are satisfied for HARA utility functions that exhibit DARA ($\gamma > 0$). Thus, for HARA functions, the absolute risk aversion of v_K is smaller than the absolute risk aversion of u when z is large. Since u and v are confounded for HARA functions, we obtain the following result:

Proposition 42 *Suppose that u is HARA. Then , the introduction of a leverage constraint $\alpha \leq k$ in the future has an ambiguous effect on the attitude towards current risks in the following sense: it raises local risk aversion around the critical wealth level where the leverage constraint becomes binding, and it reduces local risk aversion for large wealth levels.*

We represented in Figure 11.2 the typical form of the absolute risk aversion of the value function with a leverage constraint.

11.5 Concluding remarks

The previous chapter characterized the efficient dynamic portfolio strategies when investors face no constraint on the way that they can rebalance their portfolio. This

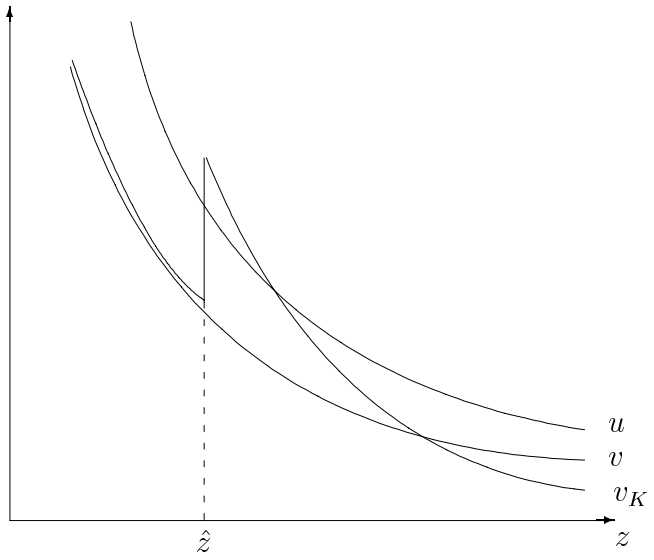

Figure 11.2: The absolute risk aversion of the value function with a leverage constraint.

was a world of full flexibility. This chapter has been devoted to various limitations to flexibility. The general intuition is that a reduction of flexibility should make investors more averse to invest in risky assets. We first examined the effect on the optimal portfolio of the impossibility to rebalance one's portfolio through time. In a second part of the chapter, we looked at the effect of some limitations on the amount that can be invested in the risky asset in the future. We first showed that it may lead to a risk-loving behavior if the choice set is not convex. We then examined the effect of a leverage constraint. Our findings have been quite mixed. Restrictions on preferences are strong, and sometimes clearly unrealistic, to guarantee that less flexibility leads to less risk taking.

Part IV

The complete market model

Chapter 12

The demand for contingent claims

In the standard portfolio problem, the final wealth of the investor is linear in the market risk $\widetilde{x}$. This is because we assumed that there is a single risky asset which condenses all risks in the economy. In the real world, people can make more sophisticated plans by purchasing specific assets whose returns are not perfectly correlated to the market. This is done for example by purchasing options, or by negotiating contingent (insurance) contracts for specific risks with other agents on the market. It implies that people can select a risky position in a much wider class than the class of linear positions. A specific risky position is described by a function $c(x)$ which determines the consumption (or final wealth) if the outcome of the market is x. In this chapter, we assume that agents can select *any* risky position $c(.)$ that is compatible with their budget constraint. Investors can also organize gambles on possible future events.

We examine the portfolio problem by assuming that markets are complete. This is a rather extreme view which is at the other end-point of the spectrum describing a financial market structure. We do this in accordance with the theory of finance in which this assumption played a central role in the progresses made in this field for the last thirty years. In fact, financial markets are incomplete, but not as much as by limiting the investors' choice to linear risky positions. Introducing an idiosyncratic, uninsurable background risk as we did in Chapter 8 is one way to relax the assumption that markets are complete.

12.1 The model

The uncertainty in the economy is represented by a set of possible states of the world that could prevail at the end of the period. This uncertainty is quantified by a random variable $\widetilde{x}$ whose cumulative distribution function is objectively known. The definition of a state of the world must describe all the elements of the environment that affect the economy. If the only source of uncertainty is the aggregate wealth available in the economy, $\widetilde{x}$ could be the GDP per capita, or any deterministic one-to-one function of it.

By definition, the Arrow-Debreu asset associated to state x has a unit value in state x and has no value otherwise. The assumption that markets are complete means that to each state of the world, there is an associated Arrow-Debreu asset that is traded on financial markets. As is well-known, this unrealistic assumption is equivalent to the condition that there are enough real assets on the market to build portfolios that duplicate the Arrow-Debreu assets.

At the time of the portfolio decision, Arrow-Debreu securities are traded, and equilibrium prices emerge. Let $\Pi(x)$ denote the equilibrium price of the Arrow-Debreu asset associated to state x. It is also called the "state price". We take them as given in this chapter. At the end of the period, a state x is revealed, and agents consume their income generated by the liquidation of their portfolio. The problem of a price-taking investor with wealth z is to select a portfolio of Arrow-Debreu securities that maximizes his expected utility under a budget constraint. A portfolio is represented by function $c(.)$, where $c(x)$ is the number of units purchased of the Arrow-Debreu security x. By definition, it is also the level of consumption in state x, since we assume that investors have no other contingent incomes than the ones generated by their portfolio. In this sense, the portfolio is also a consumption plan.

In the discrete version of the model where $\widetilde{x}$ has its support in $\{x_1, x_2, ..., x_n\}$, the problem is written as follows:

$$\max \sum_{i=1}^{n} p_i u(c_i) \tag{12.1}$$

$$s.t. \sum_{i=1}^{n} p_i \pi_i c_i = z$$

where p_i is the probability of state x_i, $c_i = c(x_i)$, $\pi_i = \Pi(x_i)/p_i$ and z is the initial wealth of the agent. Notice that we replaced the state price Π_i by $\pi_i = \Pi_i/p_i$

which is a state price density. It is the price of asset x_i per unit of probability. Observe that the expected gross return of asset x_i is the inverse of its price density:

$$\frac{p_i.1 + (1 - p_i).0}{\Pi_i} = [\pi_i]^{-1}.$$

To put a bridge between the model developed here and the standard portfolio problem, consider the case with $n = 2$. Let us introduce the following change in variable:

$$w_0 = \frac{z}{\pi_1 + \pi_2},$$

$$\alpha = w_0 - c_1$$

and

$$\widetilde{y} \sim (-1, p_1; \frac{\pi_1}{\pi_2}, p_2).$$

The elimination of c_2 by using the budget constraint makes problem (12.1) equivalent to

$$\max\ Eu(w_0 + \alpha\widetilde{y}).$$

This is the standard portfolio problem, with the risky asset return having two possible values, -1 and $\frac{\pi_1}{\pi_2}$. This equivalence does not exist anymore if they are more than 2 independent states of the world. This is because a nonlinear consumption plan cannot be duplicated by a portfolio of the risk free asset and the market risk.

12.2 Characterization of the optimal portfolio

For the general model where $\widetilde{x}$ can be discrete, continuous or mixed, the portfolio problem with complete markets is written as:

$$\max\ Eu(c(\widetilde{x})) \tag{12.2}$$

$$s.t.\ E\pi(\widetilde{x})c(\widetilde{x}) = z.$$

The Lagrangian takes the form

$$L = E\left[u(c(\widetilde{x})) - \xi\pi(\widetilde{x})c(\widetilde{x})\right] + \xi z,$$

where ξ is the Lagrangian multiplier associated with the budget constraint. The first-order condition is written as follows:

$$u'(c(x)) = \xi\pi(x)\ \forall x. \tag{12.3}$$

Since the problem is concave under risk aversion, this is the necessary and sufficient condition.

It is a tautology in this model that the consumption in state x depends only upon the price density $\pi = \pi(x)$ in that state. Without loss of generality, let us assume that there is a monotone relationship between π and x, so that function π admits an inverse. We can thus define a function $C(\pi)$ such that

$$u'(C(\pi)) = \xi\pi \qquad \forall\pi, \tag{12.4}$$

so that the optimal consumption plan is $c(x) = C(\pi^{-1}(x))$. If $\widetilde{\pi}$ denotes the random variable with the same distribution as $\pi(\widetilde{x})$, the budget constraint can be rewritten as

$$E\widetilde{\pi}C(\widetilde{\pi}) = z.$$

By the monotonicity and concavity of u and condition (12.4), C is decreasing in π: the larger the state price density, the smaller the demand for the corresponding Arrow-Debreu security and consumption. Or, the smaller the expected return of an Arrow-Debreu security, the smaller the demand for it. Under risk aversion, Arrow-Debreu securities are not Giffen goods.

One can determine the optimal exposure to risk by differentiating condition (12.4). It yields

$$u''(C)C'(\pi) = \xi,$$

or, eliminating ξ,

$$C'(\pi) = -\frac{T(C(\pi))}{\pi}, \tag{12.5a}$$

where $T(C) = -u'(C)/u''(C)$ is the index of absolute risk tolerance. The sensitivity of consumption to the state-price density equals the absolute risk tolerance of the agent evaluated at that level of consumption. Observe that $-C'$ is a local measure of the risk exposure that the agent is willing to take. A constant C means that the agent is not willing to take any risk. It yields the same level of consumption in all states of the world, although consuming in some states is more costly than in others. As shown by equation (12.5a), this is optimal only if the agent has a zero risk tolerance, i.e. only if he is infinitely risk-averse.[1] In all other cases, T is not zero and so is C': the agent bears risk. The finitely risk-averse agent is willing to take advantage of cheaper consumption in more favorable states despite it generates a risky consumption plan ex ante. These results parallel those obtained in section 5.1 for the standard portfolio problem. In both cases, risk-averse agents always take some risk.

Proposition 43 *In an economy with complete markets and risk-averse investors, the demand for Arrow-Debreu securities satisfies the following properties:*

1. *the demand for security x is decreasing in its price;*

2. *the local exposure to risk measured by $-C'(\pi)$ equals the ratio of the absolute risk tolerance measured in the corresponding state over the state price (equation (12.5a));*

3. *the local exposure to risk is decreasing (resp. increasing) in the state price if u' is convex (resp. concave).*

[1] This is the case only if u' is not differentiable at C. That is, if risk aversion is of order 1.

Proof: We just have to prove that $-C''(\pi)$ is negative when u' is convex. From equation (12.5a), this is true if $T'(C(\pi))C'(\pi)\pi - T(C(\pi))$ is negative. Since, from (3.16),

$$T'(C) = -\frac{A'(C)}{(A(C))^2} = T(C)(P(C) - A(C)), \tag{12.6}$$

this condition becomes

$$T(C(\pi))\left[(P(C(\pi)) - A(C(\pi)))\frac{-T(C(\pi))}{\pi}\pi - 1\right] \leq 0. \tag{12.7}$$

Under risk aversion, this is true if and only if $P(C(\pi))$ is nonnegative, i.e., if u' is convex.■

In other words, under risk aversion and prudence, the optimal portfolio strategy characterized by function $C(\pi)$ is decreasing and convex. We draw two typical optimal portfolio strategies in Figure 12.1.

As Mitchell (1994) for example, one may also inquire about how the total budget of the investor is allocated for the purchase of the different Arrow-Debreu securities. Since the budget associated to all contingent assets whose state price density is π equals $B(\pi) = \pi C(\pi)$, an increase in π increases B if $C(\pi) + \pi C'(\pi)$ is positive. Using condition (12.5a), this is equivalent to $C - T(C) > 0$, or $C/T(C) > 1$, i.e., to relative risk aversion being larger than unity. Reciprocally, the amount invested in a Arrow-Debreu security is decreasing in its price if relative risk aversion is less than unity. This result should be put in relation with property 1 in Proposition 18.

12.3 The impact of risk aversion

In the standard portfolio problem, an increase in risk aversion always reduces the optimal risk exposure, where the risk exposure is measured by the amount invested in the risky asset. In the complete markets model, a global measurement of the exposure to risk is less obvious, since $C'(\pi)$ just provides a measurement of the exposure to a local change in π. What we can show is in the following Proposition. It is a direct consequence of property (12.5a).

Proposition 44 *Consider two agents with utility function u_1 and u_2 that are twice differentiable. If u_1 is more risk-averse than u_2, then the consumption plan of u_1, $C_1(\pi)$, single-crosses from below the consumption plan of u_2, C_2.*

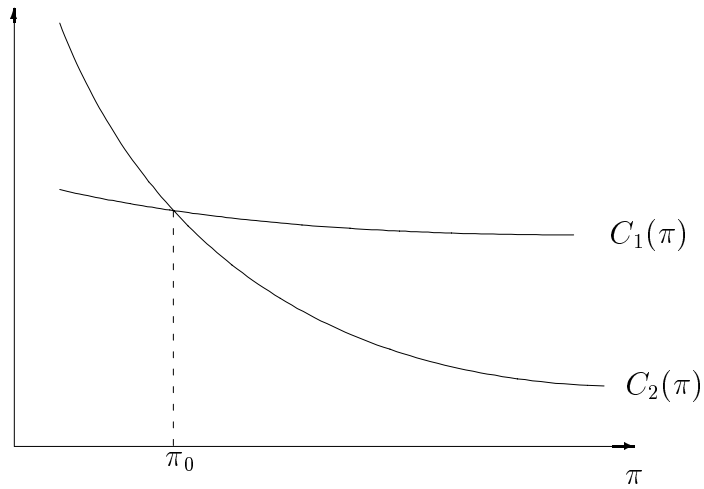

Figure 12.1: Optimal consumption plans for u and v, v being more risk-averse than u.

Proof: Suppose that there exists a π_0 such that $C_1(\pi_0) = C_2(\pi_0)$. Then, using condition (12.5a), we have

$$C_1'(\pi_0) = -\frac{T_1(C_1(\pi_0))}{\pi_0} \geq -\frac{T_2(C_2(\pi_0))}{\pi_0} = C_v'(\pi_0)$$

where T_1 and T_2 are the absolute risk tolerance of respectively u_1 and u_2. The inequality is a direct consequence of the assumption that u_1 is more risk-averse than u_2. We conclude that whenever C_1 crosses C_2, it must cross it from below. By a standard continuity argument, it implies that the two curves can cross only once. ■

We draw two representative consumption plans in Figure 12.1. It clearly appears that u_1 has a safer position than u_2. It is true in particular if the distribution of $\widetilde{\pi}$ is in a small neighborhood of π_0.

A direct application of Proposition 44 is that the presence of a non-tradable unfair background risk that is independent of $\widetilde{\pi}$ will induce risk-vulnerable investors to select a safer consumption plan C. More generally, this Proposition allows us to use all results that we already obtained about the conditions under

which a well-defined indirect utility function is more concave than the original utility function.

12.4 Conclusion

We assumed in this chapter that there are as many independent securities than the number of possible states of the world. It implies that investors can decide in an independent way how much to consume in each state. Investors can select their exposure to risk in a much broader set than the set we considered in the alternative standard portfolio model. In this paradigm, there is no much difference between the portfolio problem and the problem of determining how many bananas and apples to consume under certainty. The decision problems are technically the same: maximizing the consumer's objective under a linear budget constraint. The only difference comes from expected utility, which makes the objective function separable. This is different for the consumption problem under certainty, where the marginal utility of bananas is a function of the number of apples already in the bundle. The property of separability due to the independence axiom allowed us to derive simple properties of the optimal structure of the portfolio by using standard algebra.

The problem is that the complete markets world is unrealistic. Many risks are not tradable due to the existence of asymmetric information and transaction costs. This makes the complete markets model an interesting benchmark for its tractability. Not for its realism.

Chapter 13

Properties of the optimal behavior with complete markets

In the previous Chapter, we examined how a change in price and a change in risk aversion do affect the demand for contingent claims. We now turn to the analysis of the effect of a change in initial wealth z. This will allow us to determine the impact of z on the optimal expected utility. We will then extract some properties of the optimal portfolio strategy with complete markets similar to those that we obtained in the case of the standard portfolio model.

Let v denote the optimal expected utility that can be attained by an agent with wealth z who invests his wealth in the stock market. We called this the value function. In a slightly generalized version of the complete markets model that we examined in the previous Chapter, we allow for the utility function on consumption to depend upon the state of the world. This means that the utility depends upon C and π. To each state of the world π, there is an associated function $u(., \pi)$ that is increasing and concave. With such an extension, the program of the investor is now written as:

$$v(z) = \max \; Eu(C(\widetilde{\pi}, z), \widetilde{\pi}) \tag{13.1}$$

$$s.t. \quad E\widetilde{\pi}C(\widetilde{\pi}, z) = z \tag{13.2}$$

Notice that $u' = \partial u/\partial C$ and $-u'/u''$ measure respectively the marginal utility of consumption and the absolute tolerance to risk associated to consumption. By

defining v, we are now able to characterize the marginal value of *wealth* by v' and the absolute tolerance to risk on *wealth* by $-v'/v''$. One of the main aims of this Chapter is to examine the relationships that exist between the characteristics of the utility function on consumption and those of the value function of wealth.

Whereas z was previously considered as a constant, we are now interested in a sensitivity analysis of v with respect to a change in z. Obviously, both the Lagrangian multiplier and the optimal solution are a function of z. They are denoted respectively $\xi(z)$ and $C(\pi, z)$. The solution to this program satisfies the following condition:

$$u'(C(\pi, z), \pi) = \xi(z)\pi \tag{13.3}$$

where u' denotes the marginal utility of consumption. Solving system of equations (13.2) and (13.3) allow us to obtain $C(z, \pi)$ and $\xi(z)$. Once $C(\pi, z)$ is obtained, one can calculate the value function by:

$$v(z) = Eu(C(\widetilde{\pi}, z), \widetilde{\pi}).$$

13.1 The marginal propensity to consume in state π

In order to examine the impact of a change in wealth on the maximal expected utility, it is first useful to see how it affects the optimal consumption plan $C(\pi, z)$. To do this, we fully differentiate the above system. It yields

$$\frac{\partial C}{\partial z}(\pi, z) = \frac{\pi \xi'(z)}{u''(C(\pi, z), \pi)}$$

and

$$E\widetilde{\pi}\frac{\partial C}{\partial z}(\widetilde{\pi}, z) = 1. \tag{13.4}$$

Combining these two conditions implies that

$$\xi'(z) = \frac{1}{E\left[\frac{\widetilde{\pi}^2}{u''(C(\widetilde{\pi},z),\widetilde{\pi})}\right]} \leq 0. \tag{13.5}$$

In consequence, C is increasing in z, for all π. According to the intuition, an increase in initial wealth raises the consumption level in all states of the world. We now have an explicit formula to calculate the marginal propensity to consume in state π:

$$\frac{\partial C}{\partial z}(\pi, z) = \frac{T(C(\pi, z), \pi)}{E\widetilde{\pi}T(C(\widetilde{\pi}, z), \widetilde{\pi})}. \tag{13.6}$$

It can be useful to determine whether the marginal propensity to consume is increasing or decreasing, i.e. whether the consumption function $C(\pi, z)$ is convex or concave in z. Notice first that, by condition (13.6), the weighted average of $\frac{\partial C}{\partial z}$ is independent of z. Thus, if there exists some π where $\frac{\partial C}{\partial z}$ is increasing in z, there must exist other values of π where it is decreasing in z. Consider now a specific π. By differentiating condition (13.6), we obtain that the marginal propensity to consume $C(\pi, z)$ is convex in z if the following inequality is satisfied:

$$T'(C(\pi, z), \pi) \geq \overline{T}'(z) \equiv \frac{E\widetilde{\pi}T(C(\widetilde{\pi}, z), \widetilde{\pi})T'(C(\widetilde{\pi}, z), \widetilde{\pi})}{E\widetilde{\pi}T(C(\widetilde{\pi}, z), \widetilde{\pi})}. \tag{13.7}$$

Observe that the right-hand side of this inequality that we denoted $\overline{T}'(z)$ is a weighted average of $T'(z)$. In consequence, $\overline{T}'(z)$ is somewhere between the minimum and maximum realizations of $T'(\widetilde{C}, \widetilde{\pi})$.

Thus, $C(\pi, z)$ is locally convex (resp. concave) in z wherever $T'(C(\pi, z))$ is larger (resp. smaller) than $\overline{T}'(z)$. The limit case is when the utility function is HARA and state independent, i.e., when T' is a constant independent of C and π. It implies that condition (13.7) is always an equality. It implies in turn that the marginal propensity to consume is constant.

Proposition 45 *The marginal propensity to consume $\frac{\partial C}{\partial z}(\pi, z)$ is independent of z for state independent HARA utility functions. Otherwise, there exists a function $\overline{T}'(z)$ defined by condition (13.7) such that*

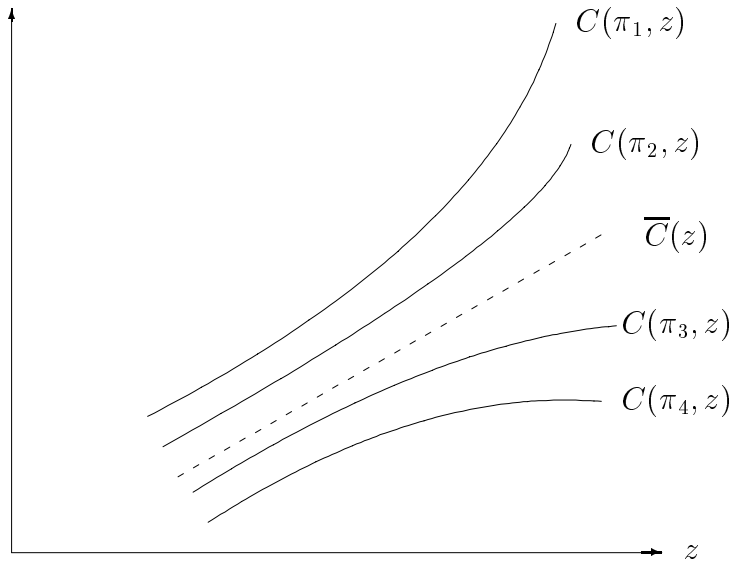

Figure 13.1: The optimal consumption plan as a function of wealth when absolute risk tolerance is convex.

$$\frac{\partial^2 C}{\partial z^2}(\pi, z) \begin{cases} \leq 0 & \text{if} \quad T'(C(\pi, z), \pi) \leq \overline{T}'(z) \\ \geq 0 & \text{if} \quad T'(C(\pi, z), \pi) \geq \overline{T}'(z). \end{cases}$$

Consider the case of a state independent utility function that is not HARA and suppose that T is convex. Then, the above Proposition states that the marginal propensity to consume out of wealth is decreasing if $T'(C(\pi, z))$ is smaller than $\overline{T}'(z)$. Let us define function $\overline{C}(z)$ such that $T'(\overline{C}(z)) = \overline{T}'(z)$. A restatement of the result for a state-independent convex T function is that the marginal propensity to consume is decreasing if $C(\pi, z)$ is smaller than $\overline{C}(z)$. A representative illustration of this case is drawn in Figure 13.1. Notice that it may be possible that a curve $C(\pi, .)$ crosses curve $\overline{C}(.)$, implying that it is alternatively locally concave and convex.

13.2 The preservation of DARA and IARA

It implies that

$$v'(z) = Eu'(C(\widetilde{\pi}, z), \widetilde{\pi})\frac{\partial C}{\partial z}(\widetilde{\pi}, z) = \xi(z)E\widetilde{\pi}\frac{\partial C}{\partial z}(\widetilde{\pi}, z) = \xi(z) \tag{13.8}$$

and

$$v''(z) = \xi'(z) = \frac{u''(C(\pi, z), \pi)\frac{\partial C}{\partial z}(\pi, z)}{\pi}.$$

Combining these equations, we obtain that

$$\frac{\partial C}{\partial z}(\pi, z)T_v(z) = T(C(\pi, z), \pi) \tag{13.9}$$

for all π, where $T(., \pi)$ and $T_v(.)$ are the absolute risk tolerance of respectively $u(., \pi)$ and $v(.)$. Multiplying both side of the equality by π, taking the expectation and using the property that $E\widetilde{\pi}\frac{\partial C}{\partial z}(\widetilde{\pi}, z) = 1$ implies that

$$T_v(z) = E\widetilde{\pi}T(C(\widetilde{\pi}, z), \widetilde{\pi}). \tag{13.10a}$$

This means that the absolute risk tolerance is a martingale.

We can go one step further by differentiating condition (13.10a), which implies that

$$T_v'(z) = E\widetilde{\pi}\frac{\partial C}{\partial z}(\widetilde{\pi}, z)T'(C(\widetilde{\pi}, z), \pi).$$

We conclude that T_v' is a weighted average of the ex post T', i.e., that the derivative of the absolute risk tolerance is also a martingale. Indeed, by (13.4), $E\widetilde{\pi}\frac{\partial C}{\partial z}(\widetilde{\pi}, z)$ equals 1. We thus have that T_v' is between the lower bound and the upper bound of T'. Moreover, because $T' = -1 + (P/A)$, we have that

$$T'(c, \pi) \geq k - 1 \qquad \Leftrightarrow \qquad P(c, \pi) \geq kA(c, \pi) \tag{13.11}$$

and

$$T'(c, \pi) \leq k - 1 \qquad \Leftrightarrow \qquad P(c, \pi) \leq kA(c, \pi).$$

This proves the following Proposition.

Proposition 46 *Consider any pair (k_1, k_2) in R^2. If the utility function on consumption satisfies condition $k_1 P(z, \pi) \geq k_2 A(z, \pi)$ for all (z, π), so does the value function v defined by condition (13.1).*

This means that gambling on complete markets preserves the property that P is uniformly larger or uniformly smaller than kA. Several special cases of Proposition 46 are noteworthy:

- $(k_1, k_2) = (0, -1)$: as we already know, risk aversion is preserved.
- $(k_1, k_2) = (1, 0)$: prudence is preserved.
- $(k_1, k_2) = (1, 1)$: decreasing absolute risk aversion is preserved.
- $(k_1, k_2) = (-1, -1)$: increasing absolute risk aversion is preserved.
- $(k_1, k_2) = (1, 2)$: condition $P \geq 2A$ is preserved.

These results should be put in contrast with the fact that even DARA is not preserved in the standard portfolio problem. Roy and Wagenvort (1996) obtained a counter-example where $v(z) = \max_\alpha Eu(z + \alpha\widetilde{x})$ is not DARA despite u is DARA.

¿From now on, we assume that u is not state-dependent, i.e., that $u(., \pi) = u(.)$ for all π.

13.3 The marginal value of wealth

In this section, we are interested in determining how the opportunity to gamble on complete markets affect the marginal value of wealth. In order to eliminate a wealth effect, we assume hereafter that $E\widetilde{\pi} = 1$: if the agent takes a risk free position, he would end up consuming his initial wealth z. Remember that we already addressed the question of how the opportunity to take risk affect the marginal value of wealth. More specifically we addressed this question in the framework of the standard portfolio problem. In Proposition 28, we showed that the option to invest in stocks raises the marginal value of wealth if and only if prudence is larger than twice the absolute risk aversion. We obtain the same result when the option is on investing in complete markets, as stated in the following Proposition.

Proposition 47 *Suppose that $E\widetilde{\pi} = 1$. Then, the option to invest in a complete set of Arrow-Debreu securities raises (resp. reduces) the marginal value of wealth if and only if absolute prudence is larger (resp. smaller) than twice the absolute risk aversion: $-u'''(C)/u''(C) \geq (\leq) 2\left[-u''(C)/u'(C)\right]$ for all C.*

Proof: Define function $\phi(y) = u'^{-1}(1/y)$. The combination of conditions (13.2) and (13.3) yields

$$E\widetilde{\pi}\phi(\frac{1}{\xi\widetilde{\pi}}) = z.$$

If ϕ is convex, Jensen's inequality implies that

$$z \geq \phi(1/\xi) = u'^{-1}(\xi).$$

Since, by condition (13.8), $\xi = v'(z)$, we obtain $u'(z) \leq v'(z)$. Finally, it is easy to verify that the convexity of ϕ is equivalent to $P(z) \geq 2A(z)$, for all z. The proof of necessity is by contradiction, by selecting $\widetilde{\pi}$ such that the support of $1/\xi\widetilde{\pi}$ be in the zone where ϕ is concave.■

13.4 Aversion to risk on wealth

We earlier obtained that the absolute risk tolerance on wealth is characterized by

$$T_v(z) = E\widetilde{\pi}T(C(\widetilde{\pi}, z)). \tag{13.12}$$

If $E\widetilde{\pi} = 1$, i.e. if the risk free rate is zero, then the degree of tolerance to risk of the value function is a weighted average of the ex post absolute risk tolerance of the investor. The use of Jensen's inequality together with the fact that $E\widetilde{\pi}C(\widetilde{\pi}, z) = z$ directly yields the following result.

Proposition 48 *Suppose that $E\widetilde{\pi} = 1$. Then, a complete set of Arrow-Debreu securities raises (resp. reduces) the absolute risk tolerance of the agent if and only if the absolute risk tolerance $T(C) = -u'(C)/u''(C)$ is convex (resp. concave) in C.*

The reader certainly made a connection between this result and the one obtained in Proposition 33 in the case of the standard portfolio problem. To make this link more precise, consider the following dynamic investment problem. At the beginning of each period, the investor can bet on any gamble giving him a dollar if and only if a prespecified event x occurs during the period. For each possible exclusive event x, there is a price $\Pi(x)$ for betting on it. These prices can be a deterministic function of time, but it is assumed that $\sum p(x)\Pi(x) = 1$ in all periods. How does the option to gamble another time in the future affect the optimal gambling strategy? Combining Propositions 44 and 48 yields a complete answer to this question: if T is convex, the investor will take a riskier position, whereas he will take a safer position if T is concave.

Two remarks must be made when comparing the effect of time horizon on the optimal dynamic investment strategy in the two frameworks. First, we obtain a much more clear-cut result in the case of the complete market model. Indeed, in the standard portfolio problem, the effect of time horizon is ambiguous when absolute risk tolerance is convex. Second, it is noteworthy that the method of proof just use standard differential calculus in the complete markets framework. This is to be compared to the heavy artillery that we had to use to solve the standard portfolio problem.

13.5 Concluding remark

In a sense, the optimal portfolio decision problem is easier to solve in the case of complete security markets than in the one-risk-free one-risky assets case. We just have to solve a system of $n+1$ equations in the first case, where n is the number of states of the world. At the limit, it is not necessary to know what is a probability or an expectation operator to do it. Thus, the Diffidence Theorem and its corollaries are just useless in this world. The assumption that markets are complete allows for the most efficient use of the separability of the decision-maker's objective function across states.[1]

[1]Chateauneuf, Dana and Tallon (1998) explores the consequences of non-additive expected utility on optimal portfolio decisions and equilibrium prices.

Part V

Consumption and saving

Chapter 14

Consumption under certainty

14.1 Time separability

In this part of this book, we examine how an expected utility maximizer allocates his consumption over time when future incomes are uncertain. We have to examine the objective function of an agent who lives from date $t = 0$ to date $t = n$. Let c_t denote consumption at date t. Suppose that the agent faces different feasible uncertain consumption streams $\{\widetilde{c}^i = (\widetilde{c}_0^i, \widetilde{c}_1^i, ..., \widetilde{c}_n^i)\}_{i=1,2,...}$ where $\widetilde{c}_t^i$ is the random variable characterizing the uncertain consumption at date t if stream i is selected. The problem of the agent is to rank them. If the agent satisfies the continuity axiom and the independence axiom for intertemporal lotteries, that is, for lotteries whose outcomes are in $X \subset R^{n+1}$, then we know from Chapter 1 that there exists a function $U : R^{n+1} \to R$ such that consumption stream $\widetilde{c}^2$ is preferred to consumption stream $\widetilde{c}^1$ if and only if

$$EU(\widetilde{c}_0^2, \widetilde{c}_1^2, ..., \widetilde{c}_n^2) \geq EU(\widetilde{c}_0^1, \widetilde{c}_1^1, ..., \widetilde{c}_n^1).$$

This general approach raises several difficulties due to the fact that the marginal utility of consumption at date t is a function of past and future consumptions. Backward induction is not easy to use when such kind of interrelations are allowed. In consequence, it is customary in intertemporal economics to assume that U is time-separable, i.e. that there exist n+1 functions $u_t : R \to R, t = 0, 1, ..., n$, such that

$$U(c_0, c_1, ..., c_n) = \sum_{t=0}^{n} u_t(c_t).$$

Time separability implies that the ordering of consumption streams from $t = 0$ to t_0 is independent of whatever consumption levels from $t_0 + 1$ onward. This looks very much like the independence axiom. In fact, this condition plays exactly the same role with respect to time than the role played by the independence axiom with respect to the states of the world. In particular, this independence condition *implies* time separability (see Blackorby, Primont and Russell (1978)). This condition is restrictive since it does not allow for habit formations, a term for the idea that the marginal utility of future consumption is increasing with the level of past consumption. Despite this deficiency, we will maintain this assumption throughout.

14.2 Exponential discounting

A further restriction on intertemporal preferences is often considered in the literature. Namely, it is assumed that

$$u_t(c) = \beta^t u_0(c)$$

where u_0 is called the felicity/utility function on consumption, and β is the discount factor. This assumption is often referred to as "exponential discounting", since its equivalent formula with continuous time is $u_t(c) = e^{-rt}u(c)$.

Several arguments are in favor of discounting. First, without going into details, this assumption forces the time consistency of preferences.[1] Second, if we

[1]Time consistency in a world of certainty means that what is optimal to do tomorrow, seen from today, is still optimal when tomorrow becomes today. Consider a problem in which you can save from $t = 1$ to $t = 2$ at a gross return ρ. Seen from today ($t = 0$), the optimal saving solves the following problem:

$$\max_s \; u_0(w_0) + u_1(w_1 - s) + u_2(w_2 + \rho s) + ...$$

where w_t is the income at t. One period later ($t = 1$), when the time to go to the bank arises, the optimal saving solves the somewhat different problem:

normalize the utility if death at zero and if we interpret β as the survival probability from one date to another, $u_1(c) = \beta u_0(c)$ has the interpretation of the expected utility of the consumption flow in period 1, seen from period 0. Finally, discounting felicity is a simple way to guarantee that the objective function of the agent is well-defined in the case of an infinite horizon model ($n \to \infty$).

It is noteworthy that β is used to discount felicity, not money. It is commonly assumed that β is less than unity. This means that one unit of felicity tomorrow is valued less than one unit of felicity today. This is a notion of time impatience for happiness. Because β^t tends to zero exponentially, it implies that the distant future does not matter for current decisions.

In the next sections, we reinterpret the concavity of the felicity function in relation with the substitution of sure consumptions over time. We will reintroduce uncertainty on future incomes in the next Chapter.

14.3 Consumption smoothing under certainty

There are two motives to save in a risk free environment: on can save to smooth consumption over time or one can save to take advantage of a high return on savings. We first examine the income smoothing argument. Consider a two-date model with $u_0 \equiv u_1 \equiv u$. Thus there is no time impatience. Furthermore, we assume that the return on saving is zero. This is to isolate the consumption smoothing effect. A justification of these assumptions is that we are considering a short-term model where the second date is "tomorrow". Suppose finally that the income flow is constant (smooth) over time: the agent earns w at each date, and has no reserve of liquidity at date 0. The problem is to determine the optimal saving s at that time:

$$\max_s \; V(s) = u(w - s) + u(w + s).$$

$$\max_s \; u_0(y_1 - s) + u_1(y_2 + \rho s) + ...$$

Obviously, what is optimal to implement at $t = 1$ is what it was planned one period earlier only if $u_t = \beta^t u_0$. This is the only case where the marginal rate of substitution between consumptions at $t = 1$ and $t = 2$ is the same today and tomorrow. See also Laibson (1996) an analysis of sophisticated consumption behaviours with hyperbolic discounting.

Observe that $V'(s) = u'(w+s) - u'(w-s)$ is zero for $s = 0$ and is decreasing in s if u is concave. Thus, under the concavity of the utility function, the optimal saving is zero. This means that the agent has a preference for a constant consumption over time. This would be the worst solution if the utility function were convex since V would be minimized at $s = 0$. Thus, the concavity of the utility function represents the willingness to smooth consumption over time.

One can measure the intensity of the willingness to smooth consumption over time by considering a situation where the income w_0 at date 0 is strictly less than the income w_1 at date 1. Since the marginal utility of consumption is larger at date 0 than at date 1 ($u'(w_0) \geq u'(w_1)$), we know that the agent will not accept a deal to exchange one unit of consumption today for one unit of consumption tomorrow. He will require more than one unit tomorrow for the sacrifice of one unit today. Accepting the deal would increase the discrepancy between the consumption levels at the two dates. So, the degree of *resistance* to intertemporal substitution can be measured by the additional reward k that must be given to the agent at date 1 to compensate the loss in consumption at date 0. If the monetary unit is small with respect to the consumption level, k is defined by the following condition:

$$u'(w_0) = (1+k)u'(w_1).$$

This condition states that the marginal loss in utility at date 0 must equal the marginal increase in utility at date 1. If w_1 is close to w_0, we can use a first order Taylor expansion of $u'(w_0)$ around w_1. It yields

$$k \simeq (w_1 - w_0)\frac{-u''(w_1)}{u'(w_1)}.$$

Thus, the resistance to intertemporal substitution is proportional $-u''/u'$. This ratio is hereafter called the degree of resistance to intertemporal substitution. The reader did certainly not miss the point that this index is equivalent to the index of absolute risk aversion. We give more flesh to the relationship between the consumption problem under certainty and the risk-taking problem in the next section.

14.4 Analogy with the portfolio problem

The willingness to smooth consumption is only one element of the picture. If the interest rate is not zero, the desire to smooth consumption is in conflict with the

possibility to increase total wealth by investing money over time. To simplify, let us assume that the interest rate ρ in the economy is the same for all time horizons (the yield curve is flat). There are zero-coupon bonds available for each date. A zero-coupon bond for date t guarantees to its holder 1 unit of the consumption good at date t and nothing for the other dates. The price of this zero-coupon is $1/\rho^t$. At date $t = 1$, the saving problem under certainty is thus written as follows:

$$\max \sum_{t=1}^{n} \beta^{t-1} u(c_t) \tag{14.1}$$

$$s.t. \sum_{t=1}^{n} \frac{c_t}{\rho^{t-1}} = z \equiv z_0 + \sum_{t=1}^{n} \frac{w_t}{\rho^{t-1}}$$

where w_t is the sure income at date t and z_0 is the current gross wealth. We defined z as the current net wealth, taking into account of the discounted value of future incomes. The budget constraint means that the discounted value of the consumption stream equals the current net wealth. Alternatively, writing the budget constraint as

$$\sum_{t=1}^{n} \frac{c_t - w_t}{\rho^{t-1}} = 0,$$

means that the portfolio of zero-bonds that the agent purchases has a zero value.

The above problem can be rewritten as

$$\max \sum_{t=1}^{n} \frac{\beta^{t-1}}{\sum_{\tau=1}^{n} \beta^{\tau-1}} u(c_t) \tag{14.2}$$

$$s.t. \sum_{t} \frac{\beta^{t-1}}{\sum_{\tau=1}^{n} \beta^{\tau-1}} \left[(\beta\rho)^{-t+1} \sum_{\tau=1}^{n} \beta^{\tau-1} \right] c_t = z.$$

Now, observe that this problem is formally equivalent to the portfolio problem (12.1) when there is a complete set of Arrow-Debreu securities. The table of correspondence is

$$p_t = \frac{\beta^{t-1}}{\sum_{\tau=1}^{n} \beta^{\tau-1}} \tag{14.3}$$

$$\pi_t = (\beta\rho)^{-t+1} \sum_{\tau=1}^{n} \beta^{\tau-1}. \tag{14.4}$$

To pursue the table of correspondence, the zero-coupon bond is the exact twin of the Arrow-Debreu security. The assumption of a state-independent utility function is technically equivalent to the assumption of a time-independent felicity function. This is the additive nature of the objective function with respect to time and uncertainty that generates these equivalences. All results that have been obtained in the static Arrow-Debreu economy can thus be used in this context.

Here is a first illustration of the correspondence with the portfolio problem: we know that full insurance (constant consumption) is optimal with a differentiable utility function only if the state price density π is the same in all states. Since π_t is proportional to $(\beta\rho)^{-t+1}$ in this model, the state price density is constant only if $\beta\rho = 1$. In the terminology of the saving problem, it is optimal to smooth consumption completely if and only if the discount rate equals the inverse of the gross return on saving. This is a generalization of the result presented in the previous section where we assumed that $\beta = \rho = 1$.

More generally, by Proposition 43, we know that the level of consumption in different states is a decreasing function of the state price density π in the corresponding state, under the concavity and differentiability of u. Again, using the table of correspondence with $\pi_t = \xi(\beta\rho)^{-t+1}$, we obtain that consumption increases with time if $\beta\rho$ is larger than unity. This is when the willingness to increase wealth by saving is stronger than time impatience. We accept to consume less at the early stages of life to benefit from the positive return on savings. This is true even if $\beta\rho$ is just above unity. This means that consumption smoothing is only a second order effect if u is differentiable, exactly as risk aversion is a second order effect for the demand of a risky asset. By condition (12.5a), we see that the rate at which consumption will increase over time is proportional to $-u'(c_t)/u''(c_t)$. This means that saving under certainty is proportional to the absolute risk tolerance of the agent.

Arrow portfolio problem	**consumption problem**
separability across states	separability across time
risk aversion	resistance to intertemporal substitution
Arrow-Debreu securities	zero-coupon bonds
full insurance is optimal only if $\pi_s = \pi$	income smoothing is optimal only if $\beta\rho = 1$
the demand is decreasing in π	consumption is increasing in time if $\beta\rho > 1$

Table: Correspondences between the portfolio problem and the consumption problem under certainty

We also obtain the following Proposition, which is a direct application of Proposition 44.

Proposition 49 *Consider the consumption-saving problem under certainty. An increase in the concavity of the felicity function in the sense of Arrow-Pratt increases (resp. reduces) early consumption if $\beta\rho$ is larger (resp. smaller) than unity.*

This is a confirmation that the degree of concavity of the felicity function is a measure of the resistance to intertemporal substitution. In the normal case where $\beta\rho$ is larger than 1, the agent accept to substitute current consumption by future consumption in order to take advantage of the relatively large return on investments. An increase in the concavity of u tends to reduce this substitution of consumption over time. It increases early consumption.

14.5 The social cost of volatility

The understanding of consumption behavior is probably one of the most important challenges in modern macroeconomics. There has been a lot of developments in this area of research since the seminal papers of Modigliani and Brumberg (1954) and Friedman (1957). Life cycle models of consumption rely on the assumption that households solve forward looking intertemporal consumption problems. Program (14.2) belongs to this family.

Lucas (1987) raised an interesting puzzle of life cycle models of consumptions. This puzzle is the twin of the puzzle that we raised in section 3.8 about the low social cost of the macroeconomic risk. In order to present Lucas' puzzle, let us solve the consumption problem (14.2) under certainty in the case of a CRRA utility function. Using the first-order condition (12.4) and the definition (14.4) for

π in the context of intertemporal consumption, it is easily verified that the optimal consumption path is

$$c_t = c_0 a^t, \tag{14.5}$$

where a is equal to $(\beta\rho)^{1/\gamma}$. Constant c_0 can be obtained by using the lifetime budget constraint of the agent. The important point here is that the growth rate g of consumption is constant over time. It equals

$$g = (\beta\rho)^{1/\gamma} - 1 \simeq \frac{\beta\rho - 1}{\gamma}. \tag{14.6}$$

Under CRRA, it is optimal to let consumption growth at a constant rate. If there is no impatience ($\beta = 1$), the optimal growth rate of consumption is approximately equal to the ratio of the interest rate $\rho - 1$ and the relative resistance to intertemporal substitution γ. With a risk free rate around 1%, as observed in the period 1963-1992, the optimal growth rate should be somewhere one-fourth of a percent ($\gamma = 4$) and one percent ($\gamma = 1$).

For the sake of illustration, consider an agent borne on January 1, 1963 who would be perfectly aware of the sequence of his disposable income for the next 30 years. Because he is a representative agent, the growth rate of his income is exactly equal to the growth rate of real U.S. GDP per capita that has actually been observed in period 1963-1992, as reported in section 3.8. The geometric mean of the growth has been $g = 1.849\%$ per year. Without any impatience and an observed risk free rate around $\rho - 1 = 1\%$, this is compatible with a coefficient $\gamma = (\rho - 1)/g \simeq 0.5$. We draw on Figure 14.1 the optimal consumption path of this agent respectively for $\gamma = 0.5$ and $\gamma = 2$. We see that the actual consumption path is close to the optimal one for $\gamma = 0.5$. We also check that a more concave felicity function leads to more consumption smoothing.

Of course, the actual consumption path is not optimal because the growth rate of consumption has not been constant over time. Business cycles occurred which forced household to adapt their consumption accordingly. Those cycles makes people worse off because of the volatility that it generates on their consumption. Given the concavity of the felicity function, they would prefer to smooth their consumption around a constant trend. What is the social cost of business cycles, assuming that they are deterministic? Let us focus on the agent with $\gamma = 0.5$, since this agent would have selected a growth rate of consumption similar to that

Figure 14.1: Actual and optimal consumption paths with $\gamma = 0.5$ (dotted line) and $\gamma = 2$ (solid line)

observed on average over the thirty years. One can measure the cost of business cycle by the reduction in the growth rate of consumption that this agent would accept in exchange for the complete elimination of business cycles. His optimal consumption path would have started at $c_0 = 10708$ dollars with a growth rate of 2.01% per year. The actual path provides the same discounted utility than a path starting at the same c_0, but with a growth rate equaling 2.006%. In consequence, business cycles "costs" reduction of 0.004% in the annual growth rate of the economy. This is totally insignificant! Lucas concluded that economists should deal with the determinants of long term growth rather than with the reduction of volatility.

14.6 The marginal propensity to consume

An important macroeconomic question since Keynes is to determine how agents react to an increase in their wealth in terms of consumption and saving. The problem is thus to derive the vector of consumption functions $(c_1(z), ..., c_n(z))$ that solves program 14.2 for all z. The marginal propensity to consume is the share of the increase in wealth that is immediately consumed, i.e., $c_1'(z)$. To examine the effect of a change in wealth on the optimal consumption plan, we use the analogy presented above and the results that we derived in section 13.1. We have that

$$c_t(z) = C(\pi_t, z)$$

where π_t is defined by equation (14.4) and function C is the function that has been studied in section 13. By analogy to condition 13.6, we obtain than the marginal propensity to consume equals

$$c_1'(z) = \frac{T(c_1(z))}{E\widetilde{\pi}T(C(\widetilde{\pi}, z))}$$

where $T(.)$ is absolute risk tolerance of $u(.)$ and $\widetilde{\pi}$ is the random variable which takes value π_t defined by (14.4) with probability p_t defined by (14.3), $t = 1, ..., n$. Using these definitions, we obtain that

$$c_1'(z) = \frac{T(c_1(z))}{\sum_{\tau=1}^{n} \rho^{1-\tau} T(c_\tau(z))} \tag{14.7}$$

An immediate consequence of this condition is that c_1' is less than unity. This is because the denominator is a sum of nonnegative elements whose the first is just $T(c_1(z))$. Moreover, a longer longevity reduces the marginal propensity to consume. The intuition is that the same discounted wealth must be consumed with more parsimony to finance consumption at the old age. The limit case is when $\beta = \rho = 1$ in which case the marginal propensity to consume is equal to $1/n$. Indeed, we know that perfect income smoothing is optimal in this case, so that any increase in wealth is equally distributed over time. When $\beta\rho = 1$ but β and ρ are each different from one, the same rule apply, but one must take into account of returns on the additional saving leading to the fair distribution of the additional wealth. It yields $c_1'(z) = 1/\sum_{\tau=1}^{n} \rho^{1-\tau}$. We conclude that the length of the time horizon has a strong effect on the marginal propensity to consume.

14.6.1 The concavity of the consumption function

Another question is to determine whether the marginal propensity to consume is decreasing with wealth. The idea that the consumption function is concave in wealth is an old one, with Keynes (1935) writing that "...the marginal propensity to consume [is] weaker in a wealthy community..." . Lusardi (1992) and Souleles (1995) found a substantially smaller marginal propensity to consume for

consumers with high wealth. This provides some support to the idea that absolute risk tolerance should be convex in wealth.

¿From Proposition 45, the marginal propensity to consume is a constant if the utility function exhibits the HARA property. If absolute risk tolerance is convex rather than linear, function $c_1'(z)$ will be decreasing if $c_1(z)$ is less than a threshold $\overline{C}(z)$ defined by condition $\overline{T}'(z) = T'(\overline{C}(z))$. Notice that this threshold is in between the smallest element and the largest element of vector $(c_0(z), ..., c_n(z))$. In the normal case where $\beta\rho$ is larger than 1, we know that this vector is ordered in an increasing way. That implies that $c_1(z)$ is necessarily less than the threshold. The other cases follow in the same way.

Proposition 50 *We consider the consumption problem under certainty in the normal case where $\beta\rho$ is larger than 1. The marginal propensity to consume is decreasing if and only if absolute risk tolerance is convex.*

14.7 Time diversification

In Chapter 10, we documented the argument that younger people should take more portfolio risk. We used a model in which the agent could repetitively take risk over time. Intermediary consumption was not allowed, and all these risks are added to each others at the end of the ”game”. No time diversification is possible in that framework. The idea of time diversification is better described by a model in which an agent can take risk today, and can allocate the induced shock on his wealth by adapting his consumption accordingly over several periods. We provide a theoretical justification to this idea. Young people are potentially more willing to take risks because current risks can be attenuated by spreading consumption over time. For example, in the case of the absence of future risk opportunity and two dates, if complete consumption-smoothing is optimal, a $\$1$ loss on current investment will be split into a fifty cent reduction in current consumption and a fifty cent reduction in future consumption. Given the concavity of the utility of consumption, that has a smaller effect on total utility than a straight $\$1$ reduction in final consumption in the investment problem.

The timing of the problem is as follows:

- at date 0, the agent makes a decision under uncertainty. It generates an uncertain payoff $\widetilde{x}$ at date 1;

- at date 1, the agent observes the realization x of $\widetilde{x}$. He allocates his discounted wealth under certainty over the n remaining dates by selecting a consumption stream $(c_1, ..., c_n)$ that is feasible.

We use backward induction to solve this problem. We start with the solution to the consumption problem at date 1. The problem at date 1 is written as follows:

$$h(y) = \max_{(c_1,...,c_n)} \sum_{t=1}^{n} \beta^{t-1} u(c_t) \tag{14.8}$$

$$s.t. \sum_{t=1}^{n} \frac{c_t}{\rho^{t-1}} = y + \sum_{t=2}^{n} \frac{w_t}{\rho^{t-1}}, \tag{14.9}$$

where $y = w_1 + x$ is the available wealth at date 1. Variable y is often called the cash-on-hand at date $t = 1$. It is important at this stage to stress the fact that y is not total wealth, which would take into account the discounted value of future incomes. Total wealth is the right-hand side of budget constraint (14.9).

In order to solve the first-period problem, it is useful to inquire about the properties of function h. Again, we can use the equivalence of this problem with problem (13.1), using the table of equivalence used in section 14.4, together with $h(y) = v(y + \sum_{t=2}^{n} \frac{w_t}{\rho^{t-1}})$. Notice that whereas h is the value function of cash-on-hand, v is the value function of aggregate wealth.

The equivalence implies that we can directly use results that have been developed in section 13 for the value function of the complete market model. For example, we know that h inherits the properties of monotonicity and concavity from u: the possibility to smooth shocks on incomes do not transform a risk-averter into a risk-lover. Thus, if the decision problem at date 0 is a take-it-or-leave-it offer to gamble, the agent should reject all unfair gambles.

The value function h also inherits prudence from u. This means that if the decision problem at date 0 is a decision to save in face of an exogenous risk that will occur at date 1, the existence of this risk will raise his (precautionary) saving. It also implies that the model developed in section 15 is automatically extended to models with more than two dates.[2]

[2]This is also due that function $\mu(.) = Eu(. + \widetilde{x})$ has a positive third derivative if u''' is positive.

Since DARA is also preserved, an increase in wealth at date 0 induces more risk taking. In short, the *qualitative* effects of risk on welfare and optimal decisions under uncertainty are unaffected by the possibility to smooth the effect of the shocks on wealth over time.

We now come back to the important question of time diversification. Consider an agent who has a stable flow of income over time: $w_1 = w_2 = ... = w_n = w$. Using equation (13.12), his tolerance to the risk offered at $t = 1$ is given by

$$T_h(y) = T_v(z) = \sum_{t=1}^{n} \rho^{1-t} T(c_t(z)),$$

with $z = y + \sum_{t=2}^{n} \frac{w}{\rho^{t-1}}$. Suppose first that $\beta = \rho = 1$, which implies that $c_\tau(z) = \frac{z}{n} = \frac{y + (n-1)w}{n}$ for all τ. Then, the above equation can be rewritten as

$$T_h(y) = nT(\frac{y + (n-1)w}{n}). \tag{14.10}$$

If the risk that is offered at $t = 0$ is small, it is enough to examine the value of T_h at $y = w$. We obtain that

$$T_h(w) = nT(w). \tag{14.11}$$

The tolerance to an instantaneous small risk is proportional to the time horizon of the agent! This is a strong argument in favor of younger people to be less risk-averse.

Remark 1 *Another approach is based on the assumption that $w = 0$. This means that h is the value function of aggregate wealth, and $T_h(y) = nT(y/n)$. It is not clear here whether a longer time horizon should induce more risk-taking. In addition to the time diversification effect, there is a dissemination effect: an agent with a longer time horizon and no flow of incomes must share his wealth among a larger number of periods. It makes him implicitly less wealthy. Under DARA, that raises risk aversion. There is a mathemitical concept for that problem, which is named subhomogeneity. A function g is subhomogeneous if $kg(z/k) > g(z)$ for all z and all $k > 1$. Thus, in the case of perfect consumption smoothing, the subhomogeneity of absolute risk tolerance guarantees the property that, given the*

same aggregate wealth, the younger investor will take more risk. It guarantees that the dissemination effect dominates the wealth effect. In the case of CRRA, absolute risk tolerance is homogeneous and the two effects neutralize each other.

When β and ρ are not necessarily equal to unity, another effect comes into the picture. Suppose that T is convex. Then, using this assumption together with Jensen's inequality yields

$$\begin{aligned} T_h(y) = T_v(z) &= \left[\sum_{\tau=1}^n \rho^{1-\tau}\right]\left[\sum_{t=1}^n \frac{\rho^{1-t}}{\sum_{\tau=1}^n \rho^{1-\tau}} T(c_t(z))\right] \\ &\geq \left[\sum_{\tau=1}^n \rho^{1-\tau}\right]\left[T\left(\sum_{t=1}^n \frac{\rho^{1-t} c_t(z)}{\sum_{\tau=1}^n \rho^{1-\tau}}\right)\right]. \end{aligned}$$

Using the budget constraint, we obtain that

$$T_h(y) \geq \left[\sum_{\tau=1}^n \rho^{1-\tau}\right]\left[T\left(\frac{y + \sum_{t=2}^n \frac{w}{\rho^{t-1}}}{\sum_{\tau=1}^n \rho^{1-\tau}}\right)\right]. \tag{14.12}$$

Evaluated at $y = w$, we obtain that

$$T_h(w) \geq \left[\sum_{\tau=1}^n \rho^{1-\tau}\right] T(w). \tag{14.13}$$

This concludes the proof of the following Proposition:

Proposition 51 *Suppose that a small risk is borne by an agent only once. Under HARA, his tolerance to this risk is proportional to $\sum_{\tau=1}^n \rho^{1-\tau}$, where n is his remaining lifetime. This is a lower (resp. upper) bound if the absolute risk tolerance is convex (resp. concave).*

The introduction of the convexity of T has been necessary to control for the optimal variability of consumption. Selecting an optimal consumption path under certainty is similar to choosing a portfolio of Arrow-Debreu securities. It reduces risk aversion if and only if T is convex.

14.8 Concluding remark

A special case of the consumption problem under certainty is the so-called ”cake-eating” problem, which can be stated in a somewhat unorthodox way as follows:

> An agent lives for n periods. There is only one state of the world. He is endowed with a cake in each period. The size of the cake can vary over time. He can transfer cakes across time at some rate of decay that can be negative (if the older cakes are better) or positive (if it is a perishable good). What is the consumption plan that maximizes the discounted value of the felicity generated by it?

Now, change some words to obtain the following question:

> An agent faces n possible states of the world. There is only one period. He is endowed with a cake in each possible state. The size of the cake can vary over different states. He can transfer cakes across states of the world by purchasing insurance contracts with an actuarially favorable or unfavorable premium rate. What is the consumption plan that maximizes the expected utility generated by it?

We see that the change in the wording does not affect the concept that are behind them. In fact, the maximization problems are strictly equivalent. Since we already characterized the solution to the second problem, which is nothing else than the Arrow-Debreu portfolio problem, we had not much to do in this chapter to characterize the solution to the cake-eating consumption-saving problem under certainty.

Chapter 15

Precautionary saving and prudence

Under certainty, the optimality of a consumption path depends upon two behavioral factors: the degree of impatience of the consumer measured by his discount factor and his resistance to intertemporal substitution measured by the degree of concavity of his felicity function. When some uncertainty affects future incomes, the optimal consumption path will also be influenced by the degree of prudence of the agent. The principal innovation in life cycle models in the past decade has been to introduce uncertainty, thereby allowing for the precautionary motive of saving. One of the main benefit is that we can accommodate a much wider range of behavior in the precautionary model. This has been at the cost of a loss of tractability, as we already observed when we examined the dynamic portfolio problem.

It is widely believed that the uncertainty affecting future incomes raises savings. Agents who behave in this way are said to be *prudent*. This an old idea originating from Keynes and Hicks, but it got the imprimatur of science only in the late sixties with the seminal works by Leland (1968) and Sandmo (1970). We now introduce uncertainty in the basic model presented in the previous sections. This allows us to examine the precautionary saving motive.

15.1 Prudence

We have the following definition:

Definition 6 *An agent is prudent if adding an uninsurable zero-mean risk to his future wealth raises his optimal saving.*

In order to determine the impact of a future income risk on optimal saving, we compare the optimal solution of the saving problem with or without this risk. Because we consider a two-date model, there is no problem of time inconsistency. Thus, the analysis does not require that $u_1(z) = \beta u_0(z)$. Under certainty, the agent has an optimal saving s^* which is the solution to problem (15.1):

$$s^* \in \arg\max_s \; u_0(w_0 - s) + u_1(w_1 + \rho s). \tag{15.1}$$

This problem is a rewriting of the one where one maximizes $u_0(c_0) + u_1(c_1)$ under the budget constraint that $c_0 + (c_1/\rho)$ equals $w_0 + (w_1/\rho)$. Using this constraint with $c_0 = w_0 - s$ yields $c_1 = w_1 + \rho s$, which can be injected in the objective function. Assuming the concavity of u_0 and u_1, the necessary and sufficient condition for s^* is written as

$$u_0'(w_0 - s^*) = \rho u_1'(w_1 + \rho s^*). \tag{15.2}$$

It is easily checked by differentiating this condition with respect to s^* and w_1 that s^* is decreasing in w_1 and increasing in w_0. This expresses the preference for consumption smoothing.

Suppose now that, at the time of the decision on s, there is uncertainty about the income that will be earned in period 1. Namely, the income at date 1 is $w_1 + \widetilde{x}$, not w_1. In order to isolate the effect of risk from the effect of smoothing expected consumption that has been analyzed in the previous sections, we assume that $E\widetilde{x} = 0$. We also assume that this risk cannot be transferred to the market. Namely, it is uninsurable. The new problem is thus to maximize

$$H(s) = u_0(w_0 - s) + Eu_1(w_1 + \rho s + \widetilde{x}).$$

We inquire whether the optimal solution to this problem is larger than s^*, the difference being the level of precautionary saving. It is noteworthy that H is concave in s if we assume that the agent has preferences for consumption smoothing, i.e., if u_0 and u_1 are concave. In consequence, there is precautionary saving if $H'(s^*) \geq 0$. Using condition (15.2), this is true if

$$Eu_1'(w_1 + \rho s^* + \widetilde{x}) \geq u_1'(w_1 + \rho s^*).$$

Intuitively, the willingness to save is increased if the expected marginal utility of future wealth is increased. To summarize, we want to guarantee that

$$E\widetilde{x} = 0 \implies Eu_1'(z + \widetilde{x}) \geq u_1'(z), \tag{15.3}$$

where $z = w_1 + \rho s^*$. Because w_0, w_1 and ρ are arbitrary, so is z. Thus we want this condition to hold for any acceptable z. This condition is formally equivalent to condition (3.1) which defines the notion of risk aversion, *except that* u_1 is replaced by $-u_1'$. Using the Diffidence Theorem or using this analogy together with Proposition (6) yields that condition (15.3) holds for any z and $\widetilde{x}$ if and only if $-u_1'$ is concave, i.e., if and only if u_1' is convex. This result is due to Leland (1968).

Proposition 52 *An agent is prudent if and only if the marginal utility of future consumption is convex.*

Prudence and risk aversion are independent concepts in the sense that one can be locally risk-averse and locally prudent, locally risk-averse and locally imprudent, locally risk-lover and locally prudent, and, finally, locally risk-lover and locally imprudent.

Kimball (1990) proposes to measure prudence — or the intensity of the precautionary saving motive — by the *precautionary equivalent premium* ψ. This has already been introduced in section (8.1) to prove that standardness implies risk vulnerability. The precautionary equivalent premium is the certain reduction in w_1 that has the same effect on optimal saving as the addition of the random noise to w_1. It depends upon z and the distribution of $\widetilde{x}$, together with the degree of convexity of u_1'. It is obtained by solving the following equation:

$$Eu_1'(z + \widetilde{x}) = u_1'(z - \psi(z, u_1, \widetilde{x})). \tag{15.4}$$

Observe again the equivalence with the symmetric notion in risk aversion, which is the risk premium π defined by condition (3.7). We have that

$$\psi(z, u_1, \widetilde{x}) = \pi(z, -u_1', \widetilde{x}).$$

This equivalence has several important consequences. For example, one can approximate ψ by a formula that parallels the Arrow-Pratt approximation for the risk premium:

$$\psi(z, u_1, \widetilde{x}) \simeq \frac{E\widetilde{x}^2}{2} P(z)$$

where $P(z)$ is the index of absolute prudence that is defined as:

$$P(z) = \frac{-u_1'''(z)}{u_1''(z)}. \tag{15.5}$$

One can also define the notion of "more prudence". An agent with utility function on future consumption v_1 is more prudent than the one with utility function u_1 if and only if $\psi(z, v_1, \widetilde{x})$ is larger than $\psi(z, u_1, \widetilde{x})$ for all $(z, \widetilde{x})$. Using Proposition 8, this is true if and only if $-v_1'''(z)/v_1''(z)$ is larger than $-u_1'''(z)/u_1''(z)$ for all z.

Finally, from Proposition 9, we obtain that ψ is decreasing in z if and only if P is decreasing in z. This notion is called decreasing absolute prudence, which states that the sensitivity of saving to future risks is smaller for wealthier people.

Proposition 53 *Define the precautionary equivalent premium $\psi(z, u_1, \widetilde{x})$ by condition (15.4) and absolute prudence P by condition (15.5). They satisfy the following properties:*

1. *$\psi(z, u_1, \widetilde{x})$ is nonnegative for all $(z, \widetilde{x})$ if and only if $P(z)$ is nonnegative for all z.*

2. *$\psi(z, v_1, \widetilde{x})$ is larger than $\psi(z, u_1, \widetilde{x})$ for all $(z, \widetilde{x})$ if and only if $-v_1'''(z)/v_1''(z)$ is larger than $-u_1'''(z)/u_1''(z)$ for all z.*

3. *$\psi(z, u_1, \widetilde{x})$ is decreasing in z for all $\widetilde{x}$ if and only if P is decreasing in z.*

Is prudence an assumption as realistic as risk aversion or decreasing absolute risk aversion? There are two main arguments in favor of prudence. First, many empirical studies have been conducted to verify whether people that are more exposed to future income risks save more. The results are in favor of this assumption. See for example Guiso, Jappelli and Terlizesse (1996) and Browning and Lusardi (1996). The second argument is based on the property that

$$A'(z) = A(z)\left[A(z) - P(z)\right],$$

which implies that prudence is *necessary* for the widely accepted view that absolute risk aversion is decreasing.

We finish with a short analysis of decreasing absolute prudence. Notice first that, exactly as $u''' > 0$ is necessary for DARA, $u'''' < 0$ is necessary for decreasing absolute prudence. Another necessary condition for decreasing absolute prudence is decreasing absolute risk aversion, as shown by Kimball (1993). This is shown easily by observing that decreasing absolute prudence means that $-u''(z+x+\theta)$ is log-supermodular with respect to (z, x, θ). By Proposition 5, it implies that $E - u''(z + x + \widetilde{\theta})$ is LSPM with respect to (z, x). Using this result with $\widetilde{\theta}$ having a uniform distribution on $[0, \alpha - z - x]$ for some $\alpha > z + x$, we get that

$$u'(z+x) - u'(\alpha)$$

is log-supermodular in (z, x)., i.e. u is DARA.

Proposition 54 *Decreasing absolute prudence is stronger than decreasing absolute risk aversion.*

15.2 The marginal propensity to consume under uncertainty

Up to now, we examined how a future risk affects the level of current saving at a given level of wealth. Another question is to determine how a future risk affects the sensitivity of saving with respect to change in wealth, i.e., the marginal propensity to consume (MPC). Consider again the two-period saving problem under uncertainty:

$$s^* \in \arg\max_s \; u(w_0 - s) + \beta E u(\rho s + \widetilde{x}) = u(w_0 - s) + \beta v(\rho s),$$

where $v(.) = Eu(. + \widetilde{x})$ is the usual indirect utility function with an exogenous background risk. Observe that we restrict our analysis here to the case where the felicity function is time-independent. Let $c^*(w_0) = w_0 - s^*(w_0)$ be the optimal consumption under uncertainty. Having excluded the background risk by using this technic, this problem becomes a particular case of the portfolio problem with complete markets, with a state dependent utility function. This will simplify the analysis of the problem.

15.2.1 Does uncertainty increase the MPC?

Using conditions (13.9) and (13.10a), the marginal propensity to consume out of wealth under uncertainty equals

$$\frac{\partial c^*}{\partial w_0} = \frac{T(c^*)}{T(c^*) + \rho^{-1} T_v(\rho(w - c^*))}. \tag{15.6}$$

We compare it to the marginal propensity to consume under certainty. The consumption problem under certainty is written as

$$\widehat{s} \in \arg\max_s \; u(w_0 - s) + \beta u(\rho s),$$

which yields an optimal consumption $\widehat{c}(w_0) = w_0 - \widehat{s}(w_0)$. We know that $\widehat{c}(w_0)$ is smaller than $c^*(w_0)$ under prudence. The marginal propensity to consume under certainty equals

$$\frac{\partial \widehat{c}}{\partial w_0} = \frac{T(\widehat{c})}{T(\widehat{c}) + \rho^{-1} T(\rho(w - \widehat{c}))} \tag{15.7}$$

The comparison of equation (15.6) and (15.7) indicates that there is little hope to obtain an unambiguous answer to the question of whether uncertainty increases the MPC. Indeed, there are two contradictory effects of uncertainty. To understand that, observe that the MPC is increasing with the current risk tolerance and decreasing with the future risk tolerance. The first effect is due the direct impact of the risk on the future risk tolerance. Under risk vulnerability, it reduces it, i.e. $T_v(z)$ is less than $T(z)$ for all z. This *raises* the MPC. The second effect is indirect: under prudence, the risk reduces the current consumption, and it increases the expected future consumption. Under DARA, that reduces the current risk tolerance, and it increases the future risk tolerance. This *reduces* the MPC. It is unclear which of the two effects dominates the other. Kimball (1998) provides more insight on this problem.

15.2.2 Does uncertainty make the MPC decreasing in wealth?

We showed in Proposition 50 that the marginal propensity to consume is decreasing with wealth under certainty if and only if absolute risk tolerance is convex. If

absolute risk tolerance is linear, then the marginal propensity to consume under certainty is not sensitive to a change in wealth. Carroll and Kimball (1996) showed that introducing uncertainty into the basic model will concavify the consumption function of HARA agents. The explanation of this phenomenon is obtained by combining results contained in Propositions 45 and 19.

By Proposition 45, we know that the consumption in the first period will be concave in wealth if $T'(w_0 - s^*)$ is smaller than $T'_v(\rho s^*)$. This condition is proven in two steps. First, we know that $T'_v(\rho s^*) \geq T'(\rho s^*)$. Indeed, a HARA utility function is such that $T'(z) = 1/\gamma$, or $P(z) = \frac{\gamma+1}{\gamma}A(z)$. By Proposition 19, it implies that $P_v(z) \geq \frac{\gamma+1}{\gamma}A_v(z)$, or equivalently, $T'_v(z) \geq 1/\gamma = T'(z)$ for all z. In particular, we have that $T'_v(\rho s^*) \geq T'(\rho s^*)$. Second, because T' is a constant in the HARA case, we obtain that $T'_v(\rho s^*) \geq T'(\rho s^*) = T'(w_0 - s^*)$. This concludes the proof of the Carroll-Kimball's result.

Proposition 55 *Under HARA, the optimal consumption rule is concave in wealth, or* $s^{*''}(w_0) \geq 0$.

The reader can easily extend this result to the case where absolute risk tolerance is convex and expected consumption is increasing through time.

15.3 More than two periods

15.3.1 The Euler equation

A dynamic version of the model is obtained by assuming that the labor income flow of the agent follows a random walk. Let $\widetilde{x}_t$ be the income at date $t = 0, 1, 2, ..., n$, with $\widetilde{x}_0, \widetilde{x}_1, ..., \widetilde{x}_n$ being independently distributed. Assuming a time-independent felicity function u, the optimal consumption strategy is obtained by backward induction. At the last date $t = n$, the agent consumes its remaining wealth. At date $t = n - 1$, after observing his income, his decision problem is written as:

$$v_{n-1}(z) = \max_{c_{n-1}} \qquad u(c_{n-1}) + \beta E u(\rho(z - c_{n-1}) + \widetilde{x}_n), \tag{15.8}$$

where z is his "cash-on-hand" at $t = n - 1$, which is the sum of his accumulated financial wealth at that time and its current labor income. Solving this problem

for each z yields the optimal state-dependent consumption $c_{n-1}(z)$ together with the value function of wealth. With this function, one can go one step backward by solving the decision problem at date $t = n - 2$. The decision problem at that date is written as

$$v_{n-2}(z) = \max_{c_{n-2}} \qquad u(c_{n-2}) + \beta E v_{t-1}(\rho(z - c_{n-1}) + \widetilde{x}_{n-1}), \tag{15.9}$$

The first-order condition to these problems problem are respectively:

$$u'(c_{n-1}) = \beta \rho E u'(\rho(z - c_{n-1}) + \widetilde{x}_n), \tag{15.10}$$

and

$$u'(c_{n-2}) = \beta \rho E v'_{n-1}(\rho(z - c_{n-2}) + \widetilde{x}_{n-1}). \tag{15.11}$$

Observe also that the envelope theorem applied to program (15.8) yields

$$v'_{n-1}(z) = \beta \rho E u'(\rho(z - c_{n-1}(z)) + \widetilde{x}_n) = u'(c_{n-1}(z)). \tag{15.12}$$

Combining all this implies that the first-order condition at $t = n - 2$ is rewritten as

$$u'(c_{n-2}) = \beta \rho E u'(\widetilde{c}_{n-1}), \tag{15.13}$$

where c_{n-2} and $\widetilde{c}_{n-1}$ are the optimal consumptions at the corresponding dates conditional to a given cash-on-hand at $t = n - 2$. This is often called the Euler equation, which states that the conditional (discounted) expectations of marginal utility must be equalized over time. There is a simple intuition behind the Euler equation: any small increase in saving at $n - 2$ that is used for increasing the consumption at $n - 1$ has no effect on the lifetime expected utility.

More generally, at date $t = 0, 1, ..., n - 1$, the decision problem is as follows:

$$v_t(z) = \max_{c_t} \qquad u(c_t) + \beta E v_{t+1}(\rho(z - c_t) + \widetilde{x}_{t+1}), \tag{15.14}$$

with $v_n \equiv u$. The Euler equation is obtained using the same method as above. It yields

$$u'(c_t) = \beta\rho E v'_{t+1}(\rho(z - c_t) + \widetilde{x}_{t+1}) = \beta\rho E u'(\widetilde{c}_{t+1}), \tag{15.15}$$

which yields the optimal consumption $c_t(z)$ at t as a function of the cash-on-hand z at that date.

15.3.2 Multiperiod precautionary saving

The extension to more than two periods raises several questions. Several of them are still unexplored. This extension of the basic model implies that one must determine which of the properties of u are inherited by the value function v_{n-1}. Because the v_{n-1} function combines an optimization operator with an expectation operator, we have Propositions 19 and 46 to assist us in solving the various problems related to this extension.

The first obvious problem is whether the uncertainty on the income at $t + 1$ increases saving at t . The question has been answered for $t = n - 1$. We just need u to be prudent. Let us go one step backward now: $t = n - 2$. Building on the equivalence between programs (15.1) and (15.9), the existence of risk $\widetilde{x}_{n-1}$ increases savings at $t = n - 2$ if and only if v_{n-1} is prudent. This is true if u itself is prudent. Indeed, this is a consequence of the fact that the maximization operator preserves prudence (Proposition 46), together with the property that $Eu'''(z + \widetilde{x})$ is positive if u''' is positive. By recurrence, v_t is prudent for all t. In the same way, v_t is concave and DARA if u is concave and DARA.

Our numerical example is based on the assumption that the $\beta = \rho = 1$. If labor income would be $w = 10$ with probability 1, perfect consumption smoothing would be optimal, with a consumption path given by $\widehat{c}_t(z) = w + (1 + n - t)^{-1}(z - w)$: consumption is set equal to the mean income plus a fraction of cash-on-hand above the mean income. This strategy in the range $z \in [5, 25]$ and $t = n - 1, n - 2, n - 3, n - 4$ is represented by the dotted lines in Figure **??**. Suppose alternatively that $\{\widetilde{x}_t\}_{t=1,\ldots,n}$ are random and i.i.d.. More precisely, suppose that the labor income at each period is either 5 or 15 with equal probability. The optimal consumption strategy is represented in the same graphics by using plain curves. We see that the uncertainty on the flow of incomes reduces the consumption at all levels of wealth and at all dates.

Figure : The optimal consumption as a function of cash-on-and.

A few other observations can be made from Figure **??**. First, it appears that the uncertainty raises the marginal propensity to consume at all dates. Second, the consumption functions are slightly concave. Whereas we know that this is true for $t = n - 1$, this is not necessarily true for more periods to go. Third, we can inquire about the intensity of precautionary savings as a function of the number of periods remaining. The level of precautionary saving is measured in this Figure by the vertical distance between the dotted curve (optimal consumption under certainty) and the plain curve (optimal consumption under uncertainty). We reported the precautionary savings in Figure **??**. It is interesting to observe that, at least for low wealth levels, the level of precautionary saving is first increasing and then decreasing in the time horizon. The reason for the ambiguous effect of time horizon on the level of precautionary saving is as follows. An increase in the number of periods to go raises the uncertainty on the aggregated net present value of future incomes. Under prudence, that should increase the level of precautionary saving. But there is also a "time-diversification" effect that goes the opposite direction: with more periods to go, the consumer is able to share each individual risk over more periods.[1]

[1]See the next chapter for a formal presentation of the effect of time horizon on precautionary

Precautionary saving as a function of cash-on-hand for different time horizons.

15.3.3 Another look at the concept of time diversification

We use this opportunity to provide an additional insight to the notion of time diversification that we explored in the previous Chapter. When the risk is borne only once in time, the tolerance to it is basically proportional to the number of periods remaining. If we consider the more realistic model that we examined in this section, when the risk repeats itself at each period, the intuition suggests that the concept of time diversification is less powerful. With two periods remaining, we can write

$$T_{v_{n-1}}(z) = T(c_{n-1}) + \rho^{-1} T_{\widehat{u}}(\rho(z - c_{n-1}) + E\widetilde{x}), \tag{15.16}$$

where $\widehat{u}(.) = Eu(. + \widetilde{x}_n - E\widetilde{x}_n)$. Suppose that u is risk vulnerable, yielding that $\widehat{u}$ is more concave than u. Then, the above inequality implies that

$$T_{v_{n-1}}(z) \leq T(c_{n-1}) + \rho^{-1} T(\rho(z - c_{n-1}) + E\widetilde{x}), \tag{15.17}$$

saving when the risk on income occurs only once in time.

Under HARA or concave risk tolerance, the right-hand side of this inequality is not larger than

$$(1+\rho^{-1})T\left(\frac{z+\rho^{-1}E\widetilde{x}}{1+\rho^{-1}}\right). \tag{15.18}$$

Assuming that the recurrent risk is small, we can limit or evaluation at $z = E\widetilde{x}$. We conclude that

$$T_{v_{n-1}}(E\widetilde{x}) \leq (1+\rho^{-1})T(E\widetilde{x}). \tag{15.19}$$

The time diversification effect is still there, which multiply the absolute risk tolerance by a factor $1+\rho^{-1}$. But in addition to this effect together with the relation with the convexity of T that we already examined in the previous chapter, there is an adverse effect generated by the risk vulnerability of the consumer's preferences. This takes the accumulation of risks into account. We expect however that the time diversification effect dominates. This is confirmed in the numerical illustration that we considered in this section. In Figure **??**, we draw the risk premium associated to risk $\widetilde{x}_t \sim (5, 1/2; 15, 1/2)$ as a function of cash-on-hand for different dates. We observe that the risk premium is not only positive and decreasing in cash-on-hand (DARA), but it is also a decreasing function of the number of periods remaining. The size of the time-diversification effect is quite amazing.

The risk premium associated to $\widetilde{x}_t$ as a function of cash-on-hand.

15.4 Long-term saving under uncertainty

Up to now, we supposed that the decision on saving can be made after observing the income earned at date each date. This is a natural assumption in a flexible economy where agents can cash on their savings whenever an adverse shock occurs on their incomes. However, a large part of households' savings is put in long-term funds in which withdrawal are costly. For various reasons, tax incentives are attached to long-term saving instruments, which make withdrawals very difficult or impossible. Good examples are mortgages and the money saved for retirement. Suppose that we are living in a extremely rigid world in which the long-term saving plan must be chosen at an early age. It implies that households consume at every period the difference between their actual income and the prespecified contribution to their retirement plan. In such a world, saving looses most of its ability to forearm households against shocks to their incomes.

The problem is to determine the optimal saving in such a rigid economy. More precisely, we want to determine whether the presence of risk on the flow of incomes affect the level of long-term saving. The benchmark is again the two-period saving problem (15.1) under certainty. In this section, we need to assume an exponential discounting $u_1 = \beta u_0$.

Let us now assume that the agent faces serially independent shocks on his income. Namely, his income at date t is $w_t + \widetilde{x}_t$, where $\widetilde{x}_0$ and $\widetilde{x}_1$ are independent and identically distributed and $E\widetilde{x}_t = 0$. The optimal long-term saving in this uncertain rigid economy is the solution to the following program:

$$\max_s \qquad H(s) = Eu_0(w_0 + \widetilde{x}_0 - s) + \beta Eu_0(w_1 + \widetilde{x}_1 + \rho s). \tag{15.20}$$

As usual, one can use the indirect utility function $v(.) = Eu_0(. + \widetilde{x}_0) = Eu_0(. + \widetilde{x}_1)$ in order to rewrite this problem as follows:

$$\max_s \qquad H(s) = v(w_0 - s) + \beta v(w_1 + \rho s). \tag{15.21}$$

This problem is formally equivalent to the initial problem (15.1), except for the replacement of the original felicity function u_0 by the indirect felicity function v. We know from Chapter 8 that v is more concave than u_0 if the original felicity function exhibits risk vulnerability. In this context, more concavity means a larger resistance to intertemporal substitution of consumption. If we assume the normal case $\beta\rho > 1$, a larger resistance to intertemporal substitution implies a smaller saving, according to Proposition 49. The symmetric case is proven in the same way.

Proposition 56 *Suppose that the felicity function exhibits risk vulnerability. The presence of risk on the flow of income reduces (resp. increases) long-term saving in the rigid economy if $\beta\rho$ is larger (resp. smaller) than unity.*

In simpler words, the presence of risk on incomes makes the agent more resistant to intertemporal substitution. It is noteworthy that in the normal case we obtain exactly the opposite result than in the previous section: here, risk reduces saving! Whether the rigid economy that we examined here is a better representation of the real world than the flexible economy presented in the previous section is an empirical question that remains to be addressed. There is no doubt that a better model would be in between these two extreme models.

15.5 Conclusion

In this Chapter, we addressed the question of how does a risk on incomes affect the optimal level of saving. We gave more flesh to the assumption that human

beings are prudent, i.e. that the third derivative of their utility function is positive. Otherwise, the presence of a risk on their future income would have the counter-intuitive effect to reduce the optimal level of saving. We also discussed the idea that human beings have standard risk aversion. Otherwise, wealthier agents would react more than poor ones to the presence of a risk on future income. This also seems counter-intuitive.

We also examine a world in which saving cannot be used to smooth consumption over time. This corresponds to the idea that a large part of savings are motivated by the willingness to accumulate wealth for retirement (the life cycle motive). If agents are forced to freeze their long-term saving, the presence of risk on the flow of future incomes is likely to reduce saving. This is true in the normal case with a rate of return on investments which is larger than the rate of impatience if the agent is risk vulnerable.

Chapter 16

The equilibrium price of time

Accepting to delay consumption is in general rewarded by a positive real return on the risk free asset. As is well-known, the level of the risk free rate is one of the most important instruments to control the level of activity of the economy. It affects the consumers' willingness to save and the entrepreneurs' willingness to invest. It is *the* variable which organizes a balance between current and future pleasures. In this sense, it determines the growth of the economy. Through its use as a discount factor to compute the net present value of potential investments, it guides our efforts for improving the future.

Let us take a specific example to illustrate this. At the current rate of emission of greenhouse gases on earth, there will be enough of such gases accumulated in the atmosphere in, say, between 50 to 200 years from now, to generate a potentially important global warming on earth. That will generate a potentially large adverse effect on GDP. Suppose that we perfectly know the size of these damages. At the Conferences of Rio and Kyoto, we discussed about how much efforts should be accomplished today to reduce these damages in the future. Notice however that there are two ways to improve the welfare of future generations: we can either make efforts to limit our emissions in order to reduce future damages, or we can increase our physical investments to increase the stock of capital available for these generations. It implies that we should make the financial effort to reduce our emission only if this strategy dominates the strategy to invest more in physical capital. Therefore, we should do that only if the internal rate of return of reducing our emissions of greenhouse gases exceeds the rate of return of risk free physical investment. This is the argument to justify the use of the Net Present Value in Cost-Benefit analyses. The discount rate must be equalized to the equilibrium return of risk free investments, i.e. to the risk free rate.

The present value of a damage D in t years equals $PV = D(1+r)^{-t}$ where r is the risk free rate. If a technology is available today to eliminate this future damage, it should be used only if its present cost is less than PV. In Table 1, we computed PV for different values of r and t, with a real damage D of one million. The exponential effect of discounting appears clearly. For example, one should spend no more than twenty cents today to eliminate a one-million damage happening 200 years from now if one uses a discount rate of 8%. Notice that 8% is considered as an acceptable discount rate. Several countries like the USA, France and Germany recommend the use of a discount rate somewhere between 5% and 8% for their public investment policy.

	50 years	100 years	200 years	500 years
1%	$608,039$	$369,711$	$136,686$	$6,907$
2%	$371,528$	$138,032$	$19,053$	50
5%	$87,203$	$7,605$	57	$2.5\ 10^{-5}$
8%	$21,321$	454	0.2	$1.9\ 10^{-11}$

Table 1: Present value of one million in t years with discount rate r.

Why is the value of a far distant benefit so low? Which discount rate should we use? To answer these questions, we need to examine the determinants of the interest rate. To do this, we will examine an economy à la Lucas (1978).

16.1 Description of the economy

There is a fixed set of identical consumers who live from date 0 to T. Because we assume that T is very large, we have in mind economic agents being dynasties of paternalistic consumers, rather than the consumers themselves, who are expected to die before T. At the origin, each economic agent is endowed with one tree of size z_0. All trees have the same size z_t at each date t. At each date, a tree produces a quantity of fruits that is proportional to its current size. There is no possibility to plant additional trees in the economy, i.e. the stock of capital is exogenous. Fruits are perishable, i.e. they cannot be transferred from one period to the other: they must be consumed instantaneously. Thus, z can also be interpreted as the GDP per capita in this economy.

Agents extract utility by consuming fruits, which are the numeraire. The grow rate of the size of the tree from date $t-1$ to t equals $\widetilde{g}_t$. We assume that there is a perfect correlation in the growth rate of the different trees in the economy during

a specific period, but there is no correlation in the growth rate of a tree in different periods. Consumers are allowed to borrow and lend fruits over time. The equilibrium exchange rate for fruits tomorrow against fruits today is the equilibrium price of time. In this simplest version of the model, we assume that there is no markets for trees, but the existence of such a market would not affect the results.[1]

With a time separable utility function, a consumer with a tree of size z and b fruits to reimburse at date t would choose his consumption of fruits c_t as follows:

$$v_t(z,b) = \max_{c_t} \qquad u(c_t) + \beta E v_{t+1}(z(1+\widetilde{g}_{t+1}), \rho_t(c_t + b - z)), \tag{16.1}$$

with $v_T(z,b) = u(z-b)$. At date t, the agent gets z fruits from his tree; he consumes c_t fruits; he reimburses his short term debt b; and he borrows the difference $c_t + b - z$ at the gross risk free rate ρ_t. Next period, the size of his tree will be $z(1+\widetilde{g}_{t+1})$, and he will have to reimburse $\rho_t(c_t + b - z)$. The optimal consumption strategy is characterized by function $c_t(z,b)$. The first-order condition to this problem is written as

$$u'(c_t) = -\beta \rho_t E v'_{t+1}(z(1+\widetilde{g}_{t+1}), \rho_t(c_t + b - z)), \tag{16.2}$$

where v'_{t+1} is the derivative of v_{t+1} with respect to its second argument.

In this model with undiversifiable risks and with agents that are identical in their preferences and in their endowment, the autharky is a competitive equilibrium: $c_t(z,b) = z - b$ for all z, b and t is the market-clearing condition on the credit market at date t. The agent consumes the crop of his tree at each date, and he never borrows or lends fruits on the credit market. Because the agent consumes his crop at each date, the value function takes the following form:

$$v_t(z,b) = u(z-b) + \sum_{\tau=t+1}^{T} \beta^{\tau-t} E\left[u(\widetilde{z}_\tau) \mid \widetilde{z}_t = z\right], \tag{16.3}$$

where $\widetilde{z}_\tau$ is the size of the tree at date τ. The risk free rate that sustains this equilibrium is obtained by using the first-order condition (16.2), together with the characterization (16.3) of the value function. It yields

[1]The non-existence of a market for trees is easy to justify when the capital called ”tree” is a human capital. Slavery is prohibited by law.

$$\rho_t(z) = \frac{u'(z)}{\beta E u'(z(1+\widetilde{g}_{t+1}))}. \tag{16.4}$$

Notice that the risk free rate is a function of the state of the economy at the corresponding date. In this simple economy, the state is given by the size z of the trees.

16.2 The determinants of the interest rate

The short term interest rate $\rho_t(z)$ given by condition (16.4) is an equilibrium because it makes individual consumption plans compatible with each other in the economy. Three components are at play in the determination of the price of time that generates this outcome: impatience, the willingness to smooth consumption over time, and precautionary saving. These three components can be introduced separately by looking at specific distributions for $\widetilde{g}_t$, as we do now.

16.2.1 The interest rate in the absence of growth

We can isolate the effect of impatience on the interest rate by assuming that there will be no growth. When $\widetilde{g}_{t+1} = 0$ almost surely, the equilibrium risk free rate ρ_t is equal to $1/\beta$. It is generally assumed that agents are impatient, i.e. that β is less than unity. This means that agents value future utils less than current ones. In the absence of growth, the discount rate on utils is the same than the discount rate on the numeraire.

With a zero interest rate on credit markets, agents will want to borrow to increase current consumption at the cost of a smaller consumption later on, just because they are impatient to consume. Because all agents would do that, this cannot be an equilibrium. Impatience must be counterbalanced by a positive interest rate to reduce the willingness to borrow. Indeed, we have seen in section **??** that the optimal saving is increasing in the interest rate if the optimal saving is zero.

Arrow (1996) suggests that this argument of pure preference for the present yields an interest rate of one to two percents.

16.2.2 The effect of a sure growth

When $\widetilde{g}_{t+1} = g_{t+1} > 0$ almost surely, the discount rate on the numeraire is not equal anymore to the discount rate on utility: an argument of income smoothing enters into the picture. Expecting larger revenues in the future, agents want to borrow today in order to smooth consumption over time. An additional increase in the interest rate is necessary to induce agents not to borrow. The equilibrium interest rate equals

$$\rho_t(z) = \frac{u'(z)}{\beta u'(z(1+g_{t+1}))}. \tag{16.5}$$

An increase in the concavity of the utility function raises the willingness to smooth consumption over time. When the growth is positive, this means that agents want to save less. The market will react by raising the interest rate. When the growth is negative, this means that agents want to save more. It implies a reduction of the equilibrium risk free rate.

Proposition 57 *Suppose that there is no uncertainty on the growth of the economy. Then, an increase in the concavity of the utility function of the representative agent*

- *increases the equilibrium risk free rate if the growth is positive ($\beta\rho > 1$);*
- *has no effect on the equilibrium risk free rate if there no growth ($\beta\rho = 1$);*
- *reduces the equilibrium risk free rate if the growth is negative ($\beta\rho < 1$).*

If g_{t+1} is small, first-order Taylor approximations of the denominator around z yield

$$\rho_t(z) \cong \frac{1}{\beta} + g_{t+1}R(z), \tag{16.6}$$

where $R(z) = -zu''(z)/u'(z)$ is the relative fluctuation aversion. We see that the impact of a sure growth g_{t+1} on the equilibrium interest rate is approximately equal to the product of g_{t+1} by $R(z)$. With an historical growth rate of real GDP per capita around 2% per year, and with R between 1 and 4, we obtain an effect increasing the interest rate by 2% to 8%.

It is noteworthy that with a growth rate of 2% per year forever, the real GDP per capita 200 years from now will be 52 times larger than today. This puts a perspective to Table 1: we should not make too much effort to improve the GDP of future generations because they will be so much wealthier than us anyway! This is a wealth effect. Because a constant growth rate of the economy means that the GDP per capita growth exponentially, we understand why the NPV of a future cost or benefit should decrease exponentially with its maturity. However, this effect heavily relies on the assumption that the growth rate is known with certainty, even for a distant future.

16.2.3 The effect of uncertainty

We now turn to the general case where $\widetilde{g}_{t+1}$ is a nondegenerated random variable. We see immediately from equation (16.4) that the presence of uncertainty on $\widetilde{g}_{t+1}$ reduces the risk free rate ρ_t if and only if u' is convex. The intuition is simple: when future incomes are uncertain, agents want to save for the precautionary motive. This must be compensated by a reduction of the risk free rate.

One way to quantify the effect of the uncertain growth on the interest rate is to define the "precautionary equivalent"[2] growth rate, the certain growth rate that yields the same interest rate. The precautionary equivalent growth rate $\widehat{g}_{t+1}(z)$ is implicitly defined as

$$Eu'(z(1+\tilde{g}_{t+1})) = u'(z(1+\widehat{g}_{t+1})). \tag{16.7}$$

As suggested by Kimball (1990), the precautionary equivalent growth rate can be approximated by

$$\widehat{g}_{t+1}(z) \cong E\tilde{g}_{t+1} - \frac{1}{2}\sigma^2_{\tilde{g}_{t+1}} \frac{-zu'''(z)}{u''(z)}. \tag{16.8}$$

Condition (16.8) indicates that the effect of the uncertainty on growth on the interest rate is the same as a sure reduction of the growth rate by the product of half its variance by relative prudence. The more prudent people, the smaller the precautionary equivalent growth rate, and the smaller the equilibrium interest rate. If

[2]This should not be confused with the certainty equivalent growth rate, the certain growth rate that generates the same expected utility of the representative agent in the future.

the precautionary effect is larger than the wealth effect combined with the effect of impatience, the equilibrium interest rate will be negative.

To sum up, two different characteristics of the utility function affect the level of the equilibrium risk free rate. A first question is to determine the effect of uncertainty. Which certain growth rate should we consider as equivalent to the uncertain growth rate we face? We showed that the degree of relative prudence determines the impact of the riskiness of growth on the precautionary equivalent growth rate. The more prudent we are, the smaller should the equivalent certain growth rate be. Prudence is measured by the degree of convexity of u'. The second question is to determine by how much we should substitute consumption today by consumption tomorrow in the face of this precautionary equivalent growth rate. This depends upon the degree of fluctuation aversion. This is measured by the degree of concavity of u. The more resistant to intertemporal substitution we are, the larger should the discount rate be, for a given certainty equivalent growth rate.

From this, we understand that one can isolate the precautionary effect on the risk free rate of a change in the utility function by assuming that the initial precautionary equivalent growth is zero, which implies that $\beta\rho = 1$. From Proposition 57, we know that there is no consumption smoothing effect of a change in the utility function in such a situation. We obtain the following result.

Proposition 58 *Suppose that the precautionary equivalent growth $\widehat{g}$ defined by equation (16.7) vanishes ($\beta\rho = 1$). Then an increase in the degree of prudence reduces the risk-free rate.*

We can easily use Proposition 57 to extend this result to the case where the precautionary equivalent growth is not zero. For exemple, if the precautionary equivalent growth is positive a change in u reduces the risk free rate if u becomes less concave and u' becomes more convex. This can be seen more explicitly by combining conditions (16.6) and (16.8). This yields

$$\rho_t(z) \cong \frac{1}{\beta} + \left[E\widetilde{g}_{t+1} - 0.5Var(\widetilde{g}_{t+1})\frac{-zu'''(z)}{u''(z)}\right]\frac{-zu''(z)}{u'(z)}. \tag{16.9}$$

Hansen and Singleton (1983) obtained a similar formula with power utility functions and Lognormal growth under continuous time. An advantage of this formula is to exhibit the three determinants of the equilibrium risk free rate, with the three

terms in the right-hand side of the above equality being respectively the pure preference for the present, the consumption smoothing effect and the precautionary effect.

16.3 The risk free rate puzzle

We can calibrate this model by using CRRA utility functions for which $\frac{-zu''(z)}{u'(z)} = \gamma$ and $\frac{-zu'''(z)}{u''(z)} = \gamma + 1$. Notice that CRRA implies that the risk free rate is independent of the wealth level: it is scale invariant.

We use data for the U.S. growth of GDP per capita over the period 1963-1992, for which we have $E\widetilde{g} = 1.86\%$ and $Var(\widetilde{g}) = (2.41\%)^2$. With $\beta = 1$, condition (16.9) is rewritten as

$$\rho - 1 = 0.0183\gamma - 0.00029\gamma^2. \tag{16.10}$$

Thus, the equilibrium risk free rate is increasing in the degree of risk aversion as long as γ is less than 31. Then, the risk free rate is decreasing in γ to eventually become negative when relative risk aversion is larger than 63. For $\gamma = 1$, the Hansen-Singleton formula yields an equilibrium risk free rate approximately equal to 1.80%, whereas, for $\gamma = 4$, it yields a risk free rate around 6.86%.[3]

If we accept the idea that reasonable values of relative risk aversion are between 1 and 4, and if we take into account of a pure preference for the present around 1%, we obtain an equilibrium risk free rate in an interval between 2.80% and 7.86%. This should be compared to the historical risk free rate that fluctuated around an average value of 1% over the last century (see Figure 6.2). Thus, following Weil (1989), we conclude that the theoretical model overpredicts the equilibrium risk free rate. The growth of consumption has been so large and its volatility has been so low than the model is unable to explain why households saved so much during this period. A much larger risk free rate than the actual one would have been necessary to explain the actual consumption growth. This is the risk free rate puzzle.

[3]The exact values are 1.80% and 7.04% respectively for $\gamma = 1$ and $\gamma = 4$. More generally, we found that condition (16.9) provides an excellent approximation of the true equilibrium risk free rate.

Another interesting point is to compare the risk free rate puzzle to the equity premium puzzle. We showed in chapter 6 that solving the equity premium requires a large γ around 40. But this choice of γ would yield an equilibrium annual risk free rate around 25%! In short, we need a low γ to solve the risk free rate puzzle, whereas we need a large γ to solve the equity premium puzzle. Solving the two puzzles simultaneously would require a large γ together with a negative rate of pure preference ($\beta >> 1$).

This illustrates the lack of flexibility either of the expected utility model, or of the utility functions with constant relative risk aversion. In Chapter 19, we explore a generalized expected utility model that is more flexible. Alternatively, we can relax the assumption that relative risk aversion is constant with wealth. Considering other utility functions will make it possible to change the relation between risk aversion and prudence, which is linked by condition $-zu'''/u'' = (-zu''/u') + 1$. In order to explore this possible solution, let us consider the set of HARA utility functions with

$$u'(z) = (z - k)^{-\gamma}, \tag{16.11}$$

for $\gamma > 0$ and k is some fixed minimum level of subsistence. Notice that relative risk aversion and relative prudence are increasing in k. Moreover, the difference between risk aversion and prudence is decreasing with k. Normalizing z_0 to unity, the Hansen-Singleton formula (16.9) becomes

$$\rho \cong \frac{1}{\beta} + 0.0186\frac{\gamma}{1-k} - 0.00029\frac{\gamma(\gamma+1)}{(1-k)^2}. \tag{16.12}$$

Remember that the second term of the right-hand side of the above equality represents the consumption smoothing effect, whereas the third term is the precautionary effect. When γ is small, only the consumption smoothing effect matters. Since the relative fluctuation aversion is increasing in k, we obtain that the equilibrium risk free rate is increasing in the relative level of minimum subsistence. This is only when γ is large that the precautionary effect enters into the picture. Because relative prudence is increasing in k, a positive k reduces the certainty equivalent growth rate that can eventually become negative. We depict the relationship between the risk free rate and γ in Figure 16.1, for different values of the relative level of minimum subsistence. From this figure, we see that allowing k to be different from 0 is not really helpful to solve the risk free rate puzzle. Moreover, a *negative* minimum level of subsistence would be necessary.

Figure 16.1: The equilibrium risk free rate when $u'(z) = (z-k)^{-\gamma}$ and $z_0 = 1$.

16.4 The yield curve

16.4.1 The pricing formula

Up to now, we limited the analysis to the short-term interest rate at date t. This is the return at t of a zero-coupon bond that makes a single payment at time $t+1$. More generally, we can define the interest rate at time t of a zero coupon bond maturing at time $t+n$. The gross return per period of such an asset is denoted y_{tn}. At equilibrium, it must satisfy the following condition:

$$(y_{tn}(z))^n \beta^n E u'\left(z \prod_{\tau=t+1}^{t+n} (1+\widetilde{g}_\tau)\right) = u'(z), \tag{16.13}$$

where z is the current consumption per capita. The right-hand side of this expression is the marginal utility cost of consuming one real dollar less at time t. Suppose that this additional dollar is invested in the zero-coupon bond with maturity at $t+n$. It will yield an additional income at that date which is equal to $(y_{tn}(z))^n$. The left-hand side is thus the discounted expected marginal utility benefit generated by consuming this additional income at $t+n$. At equilibrium, the discounted marginal benefit must equal the marginal cost, i.e., condition (16.13) must be satisfied. This condition is rewritten as

$$(y_{tn}(z))^n = \frac{u'(z)}{\beta^n E u'\left(z \prod_{\tau=t+1}^{t+n}(1+\widetilde{g}_\tau)\right)}. \tag{16.14}$$

When $n = 1$, we obtain $y_{t1}(z) = \rho_t(z)$. In general, the interest rate of a zero-coupon bond will depend upon its maturity, that is $y_{tn}(z) \neq \rho_t(z)$ for $n > 1$. The term structure of interest rates is the set of $\{y_{tn}\}_{n=1,2,...}$ at a given time t. The yield curve is a plot of the term structure, that is a plot of y_{tn} against n. The observed yield curve is most commonly upward-sloping, but it may happen that it is inverted, sloping down over some or all maturities.

There is a wide literature on the equilibrium form of the yield curve. The most cited references on this topic are Vasicek (1977) and Cox, Ingersoll, and Ross (1985a,b). The form of the yield curve is a complex function of the attitude towards risk and time, and of the statistical relationships that may exist in the temporal growth rates of the economy. The properties of the yield curve are not yet completely understood, even in the HARA case. In this section, we assume that there is no serial correlation in the growth of the economy. This is an unrealistic assumption because the instantaneous growth rate at t and $t+n$ are in general influenced by the same underlying factors. It implies that $\widetilde{g}_t$ and $\widetilde{g}_{t+n}$ are correlated in the real world. We also assume that $\widetilde{g}_1, \widetilde{g}_2, ...$ are identically distributed. It implies that the interest rate $y_{tn}(z) = y_n(z)$ is independent of t.

The structure of interest rate is governed by rules that link future short term rates to the current long term rate. They can be derived either directly from equation (16.14), or by using an arbitrage argument. For example, we have the following condition:

$$(y_2(z))^{-2} = (y_1(z))^{-1} E\left[\frac{u'(z(1+\widetilde{g}_1))}{Eu'(z(1+\widetilde{g}_1))}(y_1(z(1+\widetilde{g}_1)))^{-1}\right]. \tag{16.15}$$

This equation will be useful to determine whether $y_2(z)$ is smaller than $y_1(z)$, i.e., whether the yield curve is decreasing. The yield curve will be decreasing if if

$$u'(z)Eu'(z(1+\tilde{g}_1)(1+\tilde{g}_2)) \geq Eu'(z(1+\tilde{g}_1))Eu'(z(1+\tilde{g}_2)). \tag{16.16}$$

16.4.2 The yield curve for some specific classes of utility functions

The simplest case is obtained with CRRA functions. When $u'(z) = z^{-\gamma}$, we obtain that

$$\begin{aligned}(y_n(z))^n &= \beta^n E\left[\prod_{\tau=t+1}^{t+n}(1+\tilde{g}_\tau)^{-\gamma}\right]^{-1} \\ &= \left[\prod_{\tau=t+1}^{t+n}\beta E(1+\tilde{g}_\tau)^{-\gamma}\right]^{-1} \\ &= \left[\beta E(1+\tilde{g})^{-\gamma}\right]^n = (y_1(z))^n.\end{aligned} \tag{16.17}$$

For CRRA function, the short term interest rate is independent of GDP. It implies that the future short term interest rates are perfectly known even if there is some uncertainty about the growth rate of the economy. Moreover, they are all equal to the current short term interest rate. By a simple arbitrage argument, or using equation (16.15), the long term interest rate must be equal to it.

The case of One-Switch utility functions is interesting also because there is a simple proof for the property that the yield curve is not increasing. Following Bell (1988), function u is One-Switch if $u'(z) = a + z^{-\gamma}$ with $a > 0$ and $\gamma > 0$. It yields relative risk aversion $R(z) = -zu''(z)/u'(z) = \gamma\left[az^{-\gamma}+1\right]^{-1}$, which is decreasing in z. In this case, condition (16.16) is rewritten as

$$(a+z)(a+zE(1+\widetilde{g}_1)^{-\gamma}E(1+\widetilde{g}_2)^{-\gamma}) \geq (a+zE(1+\widetilde{g}_1)^{-\gamma})(a+zE(1+\widetilde{g}_1)^{-\gamma}), \tag{16.18}$$

This can be rewritten as

$$1+\left[E(1+\widetilde{g})^{-\gamma}\right]^2 \geq 2E(1+\widetilde{g})^{-\gamma}, \tag{16.19}$$

or

$$\left[1-E(1+\widetilde{g})^{-\gamma}\right]^2 \geq 0. \tag{16.20}$$

This is always true, which implies that the yield curve is nonincreasing for the set of One-Switch utility functions. Notice that the yield curve can be flat if $E(1+\widetilde{g})^{-\gamma} = 1$, which implies that the three above conditions are actually equalities. The yield curve would be decreasing for any random variable with $E(1+\widetilde{g})^{-\gamma} \neq 1$.

Finally, let us consider a calibration using the family of HARA functions (16.11) with $\gamma = 1$. Let us also consider the scenario in which the growth rate of the GDP per capita is either -0.55% or 4.27% with equal probabilities. This random variable has an expectation of 1.86% and a standard deviation of 2.41%,

Figure 16.2: The yield curve for $u'(z) = (z-k)^{-1}$, $\beta = 1$ and $\widetilde{g} \sim (-0.55\%, 1/2; 4.27\%, 1/2)$.

as the mean and standard deviation of the yearly U.S. growth rate of GDP per capita over the period 1963-1992. Current consumption is normalized to unity.

We computed the interest rate y_n for horizons up to n =45 years, using various values of the parameter k. The results are drawn in Figure 16.2. As a benchmark, consider the case $z = 1$ and $k = 0.5$ for which relative risk aversion equals 2 at current consumption and goes to 1 as consumption goes to infinity. The interest rate for a zero-coupon bond with maturity of 1 year is 3.50%. A zero-coupon bond with a maturity in 45 years has an equilibrium rate of return of only 2.80% per year.

16.4.3 A result when there is no risk of recession

For the sake of a simple notation, let x_t denote the gross rate of growth $1 + g_t$. Let function h from R^2_+ to R be defined as $h(x_1, x_2) = u'(zx_1x_2)$. Suppose that this function be log supermodular. This means that

$$h(\min(x_1, x_1'), \min(x_2, x_2'))\, h(\max(x_1, x_1'), \max(x_2, x_2')) \geq h(x_1, x_2)\, h(x_1', x_2') \tag{16.21}$$

for all (x_1, x_2) and (x_1', x_2') in R^2_+. Taking $x_1 = x_2' = 1$, the log supermodularity of h implies that

$$h(1,1)\ h(x_1', x_2) \geq h(1, x_2)\ h(x_1', 1) \tag{16.22}$$

for all x_1' and x_2 that are both larger than 1. This inequality is equivalent to

$$u'(z)u'(z(1+g_1)(1+g_2)) \geq u'(z(1+g_1))u'(z(1+g_2)) \tag{16.23}$$

for all $g_1 = x_1' - 1$ and $g_2 = x_2 - 1$ that are both positive. Suppose now that the growth rate $\widetilde{g}_t$ per period is positive almost surely. Then, the log supermodularity of h is sufficient to guarantee that condition (16.23) holds almost everywhere. Taking the expectation of this inequality directly yields inequality (16.16), which implies in turn that y_2 is less than y_1.

If the utility function is three time differentiable, the log supermodularity of h also means that the cross derivative of $\log h$ is positive. It is easily seen that this is equivalent to require that relative risk aversion is decreasing (DRRA). Notice that all inequalities from (16.16) to (16.23) are reversed if relative risk aversion is increasing.

Proposition 59 *Suppose that the growth rate of consumption is nonnegative almost surely. The long term discount rate is smaller (resp. larger) than the short term one if relative risk aversion is decreasing (resp. increasing).*

There is a simple reason why should the yeild curve be decreasing when relative risk aversion is decreasing, in the absence of recession. The combination of these two conditions implies that the future short term rate is decreasing in GDP per capita. Indeed, we have that

$$\begin{aligned} zy_1'(z) &= y_1(z)\left[\frac{-Ez\widetilde{x}_2 u''(z\widetilde{x}_2)}{Eu'(z\widetilde{x}_2)} - \frac{-zu''(z)}{u'(z)}\right] \\ &= y_1(z)\left[E\left[\frac{u'(z\widetilde{x}_2)}{Eu'(z\widetilde{x}_2)}R(z\widetilde{x}_2)\right] - R(z)\right], \end{aligned} \tag{16.24}$$

It is then clear that y_1' is negative if x_2 is larger than unity almost surely and $R' \leq 0$. Now, this implies that the future short term interest rate will be smaller than today almost surely , since we assume that the GDP per capita in the future will be at least as large as today. By the standard arbitrage argument, it must be the case that the current long term rate is smaller than the current short term one. Notice that we require that both $\widetilde{g}_1$ and $\widetilde{g}_2$ be nonnegative to get the result.

Decreasing relative risk aversion is compatible with the well-documented observation that the relative share of wealth invested in risky assets is an increasing function of wealth. Kessler and Wolf (1991) for example show that the portfolios of U.S. households with low wealth contains a disproportionately large share of risk free assets. Measuring by wealth, over 80% of the lowest quintile's portfolio was in liquid assets, whereas the highest quintile held less than 15% in such assets. Guiso, Jappelli and Terlizzese (1996), using cross-section of Italian households, observed portfolio compositions which are also compatible with decreasing relative risk aversion. This suggests that the yield curve should be decreasing in an economy without serial correlation in the growth rate of GDP per capita, and without any friction on the credit market. Thus one should select a smaller rate to discount far distant benefits than the rate to discount benefits realized in the near future. Growth risks are mutually aggravating.

16.4.4 Exploring the slope of the yield curve when there is a risk of recession

DRRA is not sufficient for a decreasing yield curve when there is a risk of recession. To illustrate this, let us assume that $\beta y_1(z_0) = 1$: the precautionary effect just compensate the wealth effect, which implies that the short term interest rate is equal to the rate of pure preference for the present. This also means that $Eu'(z_0(1+\widetilde{g}_1)) = u'(z_0)$. This may be true only if there is a positive risk of recession. The yield curve would be decreasing if $\beta y_2(z_0)$ is smaller than unity. This property is summarized as follows:

$$\left.\begin{array}{rcl} Eu'(z_0\widetilde{x}_1) & = & u'(z_0) \\ Eu'(z_0\widetilde{x}_2) & = & u'(z_0) \end{array}\right\} \quad \Longrightarrow \quad Eu'(z_0\widetilde{x}_1\widetilde{x}_2) \geq u'(z_0). \tag{16.25}$$

The two conditions to the left correspond to our assumption that $\beta y_1(z_0) = 1$. The condition to the right means that $\beta y_2(z_0)$ is less than unity. The interpretation of condition (16.25) in terms of saving behaviour is simple. Suppose that the agent does not want to save when the uncertain growth per period of her income is either $\widetilde{x}_1$ or $\widetilde{x}_2$. Does it imply that she would save in the presence of growth risk $\widetilde{x}_1\widetilde{x}_2$? Condition (16.25) states that she wants to save more. That would reduce the equilibrium interest rate.

Define function v in such a way that $v(z) = -u'(z)$ for all z. Next, we define function V as $V(Z) = v(\exp Z)$, and $Z = \ln z_0$ and $X_t = \ln x_t$. We can then rewrite property (16.25) as follows:

$$\left.\begin{array}{rcl} EV(Z+\widetilde{X}_1) & = & V(z_0) \\ EV(Z+\widetilde{X}_2) & = & V(z_0) \end{array}\right\} \quad \Longrightarrow \quad EV(Z+\widetilde{X}_1+\widetilde{X}_2) \leq V(Z). \quad (16.26)$$

In words, condition (16.26) means that two lotteries on which the agent with utility function V is indifferent when taken in isolation are jointly undesirable. This means that V is weak proper, as defined by Pratt and Zeckhauser (1987). They showed that this condition is satisfied for all function V that are $HARA$. Now, remember that $V(Z) = -u'(\exp Z)$. It implies for example that One-switch utility functions satisfy condition (16.25). Indeed, with $u'(z) = a + z^{-\gamma}$, we have $V(Z) = -a + b \exp Z$, which belongs to the class of HARA functions.

More interestingly, we know that a necessary condition for V to be weak proper is that

$$\left[\frac{-V''(Z)}{V'(Z)}\right]'' \geq \left[\frac{-V''(Z)}{V'(Z)}\right]' \left[\frac{-V''(Z)}{V'(Z)}\right]. \quad (16.27)$$

This condition is necessary in the sense that its violation would imply the existence of a pair (X_1, X_2) that would violate condition (16.26). Pratt and Zeckhauser also showed that this condition is necessary and sufficient for condition (16.26) to hold when $\widetilde{X}_1$ and $\widetilde{X}_2$ are small risks. With $V(Z) = -u'(\exp Z)$, we have that

$$\frac{-V''(Z)}{V'(Z)} = P(\exp Z) - 1 \quad (16.28)$$

$$\left[\frac{-V''(Z)}{V'(Z)}\right]' = (\exp Z)P'(\exp Z) \quad (16.29)$$

$$\left[\frac{-V''(Z)}{V'(Z)}\right]'' = (\exp Z)P'(\exp Z) + (\exp Z)^2 P''(\exp Z). \quad (16.30)$$

This yields the following result.

Corollary 5 *A necessary condition for the yield curve to be nonincreasing under the stationary condition (**??**) is that*

$$zP''(z) \geq P'(z)(P(z) - 2) \quad (16.31)$$

for all z, where $P(z) = -zu'''(z)/u''(z)$ is relative prudence. This condition is necessary and sufficient when the risk on growth is small and the short term interest rate equals the rate of pure preference for the present.

Necessary condition (16.31) is sophisticated, as it requires conditions on the fifth derivative of the utility function. This means that introducing the risk of recession in the long term and in the short term makes it really a hard task to guarantee that long term discount rates are smaller than short term ones.

Up to now, we have not been able to fully characterize the set of utility functions generating a nonincreasing yield curve, whatever the distribution of $\widetilde{g}_t$. In the next Proposition, we provide the necessary and sufficient condition when $\widetilde{g}_t$ is binary.

Proposition 60 *Define function H as $H(x,y) = u'(z)u'(zxy) - u'(zx)u'(zy)$. Suppose that u exhibits DRRA. A necessary condition for the yield curve to be nonincreasing independent of the distribution of $\widetilde{g}_t$ is written as*

$$[H(x,y)]^2 \leq H(x,x)H(y,y) \tag{16.32}$$

for all (x,y) such that $x < 1 < y$. Limiting the analysis to binary distributions for $\widetilde{g}_t$ makes condition (16.32) necessary and sufficient.

Proof: We normalize z to unity. Let $1+\widetilde{g}$ be distributed as $(x,p;y,1-p)$. Condition (16.16) is rewritten as

$$u'(1)\left[p^2u'(x^2) + 2p(1-p)u'(xy) + (1-p)^2u'(y^2)\right] \geq \left[pu'(x) + (1-p)u'(y)\right]^2, \tag{16.33}$$

or, equivalently,

$$K(x,y,p) \equiv p^2H(x,x) + 2p(1-p)H(x,y) + (1-p)^2H(y,y) \geq 0. \tag{16.34}$$

Remember that DRRA means that $(x-1)(y-1)H(x,y) \geq 0$. In particular, it means that $H(x,x)$ and $H(y,y)$ are nonnegative. When x and y are both larger than unity, we also have that $H(x,y)$ is nonnegative, yielding $K(x,y,p) \geq 0$. This is another proof of Proposition 59 for binary distributions of $\widetilde{g} \geq 0$. Again, the difficulty is when $x < 1 < y$, implying $H(x,y) < 0$. This is the case that we examine now.

Observe that K is quadratic in p with $\partial^2K/\partial p^2 = H(x,x) - 2H(x,y) + H(y,y) \geq 0$. It is minimum at

$$p^* = \frac{H(y,y) - H(x,y)}{H(x,x) - 2H(x,y) + H(y,y)} \tag{16.35}$$

which is between 0 and 1. The minimum value of $K(x,y,p)$ for $p \in [0,1]$ is

$$K(x,y,p^*) = \frac{H(x,x)H(y,y) - [H(x,y)]^2}{H(x,x) - 2H(x,y) + H(y,y)}. \tag{16.36}$$

It will be nonnegative for all $x < 1 < y$ if and only if condition (16.32) is satisfied.■

The reader can check that, symmetrically, increasing relative risk aversion together with condition $[H(x,y)]^2 \geq H(x,x)H(y,y)$ for all $x < 1 < y$ are necessary for the yield curve to be nondecreasing. It can also easily be checked that condition (16.32) is satisfied as an equality for One-Switch utility functions. Finally, observe that a sufficient condition for (16.32) is that H itself be log supermodular. This condition should not be confounded with DRRA, which means that $h(x,y) = u'(zxy)$ be log supermodular.

16.5 Concluding remark

We showed that the equilibrium price of time is influenced by three distinct characteristics of the preferences of the representative agent:

- his pure preference for the present which is characterized by β;
- his willingness to smooth consumption, if the economy is growing. This is characterized by the degree of concavity of his utility function;
- his willingness to accumulate wealth, if the rate of growth of the economy is uncertain. This is characterized by the degree of convexity of his marginal utility.

A larger expected growth increases the equilibrium interest rate, whereas more uncertainty on the growth rate reduces it. When we calibrate the model to the actual mean and standard deviation of the historical economic growth, we obtain a short-term interest rate around 4 to 5 percents per year, which is much larger than

its historical level (around 1%). The large historical growth rate of the economy and its low variability around the mean are such that economic agents should have accumulated much less saving than they actually did at the historical rate of interest. This is the risk free rate puzzle.

Another intruiging phenomenon is that under the reasonable assumption that relative risk aversion is decreasing, the yield curve should be decreasing, i.e. the long-term interest rate should be less than the short-term one. This means that the so-called ”inverted” yield curve should be the rule rather than the exception. A possible explanation of this puzzle is that ordinary households face a liquidity constraint. This would generate an illiquidity premium for long term investments.

Chapter 17

The liquidity constraint

Up to now, we assumed that agents can borrow and save how much they want at the interest rate ρ. This is not a realistic assumption. the evidence is that the rate at which we can borrow money is larger than the rate at which one can save (the opposite would generate a money machine). This is due to the cost of intermediation together with the asymmetric information that prevails with respect to the solvency issue. In this Chapter, we take the extreme position that agent may not be a net borrower: net saving must be positive at any time. This is the strong version of the liquidity constraint.

The existence of a liquidity constraint reduces the ability to smooth consumption over time. In particular, if agents expect a large increase in their earnings in the future, their inability to borrow forces them to have a rate of increase in consumption over time that is larger than the optimal one. Deaton (1991), Deaton and Laroque (1992) and Hubbard, Skinner and Zeldes (1994) among others, used numerical simulations to show that the risk to face a liquidity constraint introduces another motive to save that may have a large impact on the optimal level of saving, and on the impossibility to smooth consumption over time. In this case, saving is a ”buffer stock” which reduces the probability that the liquidity constraint will be binding in the future. Everyday life provides a lot of examples that illustrates the fact that a liquidity constraint may have a devastating effect on the lifetime welfare of a household. It just prohibits consumption smoothing.

We hereafter examine some of the consequences of the liquidity constraint on the saving behavior and on the attitude towards risk.

17.1 Saving as a buffer stock

It is trivial to say that a binding liquidity constraint reduces saving. But even when current earnings plus current cash are large enough to cover current expenses, the risk that the liquidity constraint will be binding in the future because of a sequence of adverse shocks on incomes can induce a larger saving rate. The existence of the liquidity constraint is costly for agents with a concave felicity function because of the misallocation of consumption that it generates. These agents are willing to pay for the reduction of the risk that this constraint will be binding in the future. The simplest way to reduce this risk is to accumulate more wealth. This is the intuition for why the liquidity constraint increases saving even when the constraint is currently not binding. Another way to explain the same idea is that the possibility of borrowing provides some insurance. Removing this opportunity is like introducing a downside risk which increases the willingness to save. This excess of saving with respect to its level when agents can freely save and borrow under the lifetime budget is called the "buffer stock" saving. The term has been introduced by Deaton (1992) because agents accumulate assets only to insulate themselves from income fluctuations.

In this section, we consider the simplest possible model to examine the effect of the liquidity constraint on the saving behavior. Our model is such that the consumer faces risk only once in his life. The specific timing of the problem is as follows:

- At date 0, the consumer has a cash-on-hand w_0. He decides how much to save (s), and he consumes the difference $w_0 - s$.
- At date 1, the agent observe his random income $w_1 + \widetilde{\varepsilon}$. He determines his consumption plan $(c_1, ..., c_n)$, given his cash-on-hand $s + w_1 + \widetilde{\varepsilon}$, and given his sure flow of income $(w_2, ..., w_n)$ for the remaining periods.

Observe that we have two consumption decisions here. One is done ex ante, before observing $\widetilde{\varepsilon}$. The other is made ex post. In order to understand the first, we must first examine the second. For simplicity, let us assume that $w_2 = ... = w_n = w$. Let us assume first that and $\beta = \rho = 1$.

Without any liquidity constraint, its optimal level of saving s at date 0 solves the following problem:

$$\max_s \; u(w_0 - s) + v(s + w + \widetilde{\epsilon}), \tag{17.1}$$

where v is defined by condition (14.8). Let z denote the cash-on-hand at $t = 1$, i.e., $z = s+w+\varepsilon$. Under our assumptions, the value function v takes the following form:

$$v(z) = nu(\frac{z + (n-1)w}{n}) \tag{17.2}$$

because perfect smoothing is optimal here. If z is larger than w, the consumer will consume w plus only $1/n$ of the difference between z and w at date 1. The remaining is equally allocated over the $n-1$ other dates. On the contrary, if z is smaller than w, this negative shock is split equally over the n dates by borrowing $(n-1)/n$ of the $|z-w|$ from an intermediary. Observe that in any case, the marginal propensity to save is only $1/n$.

The willingness to save at date $t = 0$ is measured by the expectation of function v', with $v'(z) = u'(\frac{z+(n-1)w}{n})$ for all z. Thus, if z is random, we see that the consumer should take into account only $1/n$ of it. This is the time diversification effect that we already illustrated in the previous chapter. It reduces the precautionary saving accordingly.

Suppose now that the agent is prohibited to be a net borrower. The problem can here be written as (17.1), but with a value function v that is replace by function v_c that is defined as

$$v_c(z) = \begin{cases} u(z) + (n-1)u(w) & \text{if} \quad z \leq w \\ nu(\frac{z+(n-1)w}{n}) & \text{if} \quad z > w \end{cases}$$

If z is smaller than w, the liquidity constraint is binding and the agent is forced to absorb all the shock immediately, by reducing his consumption to $c_1 = z$. Observe that an immediate effect of the liquidity constraint is to raise the MPC from $1/n$ to 1 in the range where it is binding. Another remark is that the ex post cost of the liquidity constraint can be measured by the difference between $v(z)$ and $v_c(z)$, which is nonnegative under risk aversion.

To make the problem interesting, suppose that the liquidity constraint is not binding at date 0 (s is positive). This can be done by assuming that w_0 is large enough. However, the risk that it may be binding in the future if $\widetilde{\epsilon}$ is unfavorable may have a positive impact on the optimal saving at date 0. This is the notion of a buffer stock. The buffer stock is positive if Ev_c' is larger than Ev'. Since v_c' and v' coincide when $s + \varepsilon$ is positive, we have to check whether v_c' is larger than v' for all negative $s + \varepsilon$, that is if

$$z < w \implies u'(z) \geq u'(\frac{z + (n-1)w}{n}).$$

This is always true under the assumption of the concavity of the utility function. We conclude that the presence of a liquidity constraint always induces risk-averse households to increase their saving when $\beta = \rho = 1$. The inability to smooth negative shocks over time raises the expected marginal utility at the date of the shock. No other assumption than risk aversion is required in this case.

The analysis is less straightforward when we do not assume that $\beta = \rho = 1$. We know that selecting a consumption path under certainty when $\beta\rho \neq 1$ is like selecting a portfolio of Arrow-Debreu securities under uncertainty. Introducing a liquidity constraint is thus like prohibiting the purchase of such portfolio. We know from Proposition 47 that it raises the expected marginal utility if prudence is smaller than twice the absolute risk aversion. In fact, the following Proposition basically extends Proposition 47 to the case where $E\widetilde{\pi}$ is larger than 1.

Proposition 61 *Suppose that $\beta\rho$ is larger than unity. The existence of a liquidity constraint increases savings if absolute prudence is less than twice the absolute risk aversion.*

Proof: See the Appendix to this Chapter.

Another important effect of the liquidity constraint on the optimal consumption is that it concavify the consumption function. Indeed, the marginal propensity to consume out of wealth at date $t = 1$ equals

$$\frac{\partial c_1}{\partial z} = \begin{cases} 1 & \text{if} \quad z \leq w \\ [\sum_{t=1}^{n} \rho^{1-t}]^{-1} & \text{if} \quad z > w. \end{cases} \tag{17.3}$$

Because $[\sum_{t=1}^{n} \rho^{1-t}]^{-1}$ is (much) smaller than unity, the MPC is decreasing, and the consumption is concave.

17.2 Liquidity constraint and risk aversion

We now turn to the effect of a liquidity constraint on the degree of risk tolerance. It is intuitive that an agent who is able to smooth a shock over several periods should be more willing to take risk than an agent who is forced to absorb the shock

immediately through a reduction of its current consumption. This is because the liquidity constraint eliminates the time diversification effect for downside risks.

We consider the same problem as above, except that the decision problem at date 0 is on an optimal exposure to risk, not a saving problem. It can be for example a standard (or an Arrow-Debreu) portfolio that will be liquidated at date 1. Using Proposition 14 (or Proposition 44), we know that the effect of a change in the environment has a negative impact on the optimal risk exposure if it increases the degree of concavity of the value function. The problem is thus to compare T_v to T_{v_c}. Since v and v_c coincide when z is larger than w, so do T_v and T_{v_c}. Thus, we just have to look at what happens in region $z < w$. Differentiating condition (17.2) twice yields

$$T_v(z) = nT(\frac{z + (n-1)w}{n}).$$

For v_c, we directly obtain that

$$T_{v_c}(z) = T(z).$$

Thus, a liquidity constraint reduces the willingness to take risk if T_v is larger than T_{v_c} in the relevant domain, i.e., if

$$z < w \implies nT(\frac{z + (n-1)w}{n}) \geq T(z).$$

When z is close to w, we have that the unconstrained agent has an absolute risk tolerance that is n times larger than the absolute risk tolerance of the agent facing a binding liquidity constraint. When z is smaller than w, the liquidity constraint induces a smaller consumption at date 1. This generates a ”wealth effect” that can go both directions depending upon whether absolute risk tolerance is increasing or decreasing. If it is increasing, we verify immediately that

$$nT(\frac{z + (n-1)w}{n}) \geq nT(z) \geq T(z).$$

We conclude that DARA is sufficient to guarantee that the a liquidity constraint induces more risk aversion, if $\beta = \rho = 1$. This effect is quite large, since a binding liquidity constraint divide risk tolerance by more than a factor n under DARA.

If assumption $\beta = \rho = 1$ is relaxed, the convexity of absolute risk tolerance must be added to DARA to guarantee the result. This result parallels the one obtained in the previous section: introducing a liquidity constraint plays the same role than prohibiting the purchase of a portfolio of Arrow-Debreu securities. From Proposition 48, it reduces the tolerance to risk if T is convex. The proof of the following Proposition follows the same line than the proof of Proposition 61. It is let to the reader.

Proposition 62 *The liquidity constraint reduces the willingness to take risk if the absolute risk tolerance is increasing and convex.*

17.3 Numerical simulations

The problem is more complex to analyze if all future incomes are uncertain. By backward induction, the decision problem at t is written as

$$v_t(z) = \max_{c_t < z} \qquad u(c_t) + \beta E v_{t+1}(\rho(z - c_t) + \widetilde{x}_{t+1}), \tag{17.4}$$

with $v_n \equiv u$. The Euler equation is written as

$$u'(c_t) \geq \beta \rho E v'_{t+1}(\rho(z - c_t) + \widetilde{x}_{t+1}) \tag{17.5}$$

with an equality if the liquidity constraint is not binding.

As in previous chapters, the effect of adding more risks must be treated by looking at the preservation of the relevant properties of the felicity function. This research remains to be done and will not be covered here. Rather, we provide numerical simulations based on the example presented in Section 15.3.2. As a reminder, we assume that $\beta = \rho = 1$ and that the income at every period is either 5 or 15 with equal probability. Finally, we consider an agent with a constant relative risk aversion equaling 2. The benchmark case is the optimal consumption strategy when there is no liquidity constraint, as described in Section 15.3.2. In Figure 17.1, we draw the optimal strategy with a liquidity constraint as a function of cash-on-hand for different time horizons.

We see that the liquidity constraint becomes binding at a cash-on-hand around 6, for all time horizons that are considered here. The concavity of the consumption

Figure 17.1: The optimal consumption as a function of cash-on-hand with a liquidity constraint.

function comes from the kink generated by the constraint, but also from Proposition 55. If zero assets are carried forward, as is the case to the left of the graph, consumption changes are set equal to income changes over the period.

One can measure the size of the buffer stock behavior by the difference between the optimal consumption without the liquidity constraint and the optimal consumption with the liquidity constraint. These two functions of cash-on-hand are depicted in Figure 17.2 for different time horizons. When the quantity of assets carried forward is large, the risk of the liquidity constraint being binding in the future is small, in particular if the time horizon is short. Then, the buffer stock need not be large. This confirmed by this Figure, since the optimal consumption quickly converge to the optimal consumption without the constraint. For example, with three periods to go and a cash-on-hand of 10, it is optimal to consume around 8 at that date as it appears in Figure 17.2b. The agent will carry 2 forward. If the bad state of the world occurs, it makes a cash-on-hand of 7. We see on Figure17.2a that the liquidity constraint does not bind with that wealth level at that time. This explains why the buffer stock saving is zero with $z = 10$ with three periods to go. A striking feature of these Figures is the small size of the buffer stock. This point

Figure 17.2: Comparing optimal consumptions with and without the liquidity constraint.

has first been made by Deaton (1992).

We also measured the attitude toward risk as a function of the length of time horizon. We did it by computing the willingness to pay for the elimination of the of the current risk on income. We report these results on Figure 17.3, which should be compared to what we obtained when there is no liquidity constraint. We see a huge increase in risk aversion generated by the partial removal of the opportunity to time diversify due to the liquidity constraint.

17.4 Conclusion

The existence of a liquidity constraint is extremely useful to explain several aspects of saving and risk-taking behaviors. First, it provides an additional motive to save. The ”buffer stock” motive to save relies on the idea that a larger wealth accumulation is useful to get rid of the risk of the liquidity constraint to be binding in the future. It generates a better smoothing of consumption over time. Second,

Figure 17.3: The risk premium for the current income risk as a function of cash-on-hand.

it explains why the marginal propensity to consume is so large in the real world. A life cycle model with a perfect financial market would predict a very low MPC for young households, in particular if there is no serial correlation in the flow of labor incomes. The existence of the liquidity constraint can explain the fact that actual consumption flow is much parallel to the actual flow of incomes.

The liquidity constraint also explain why young consumers may be more risk-averse than would have been expected with a perfect financial market. Remember that in such an environment, the young consumer's tolerance to risk is basically proportional to her remaining lifetime. This is due to her better position to time diversify intertemporally independent risks. But in the real world, we don't see many young agents insuring older ones. This suggests that time diversification cannot be used in a large extend by most consumers. The liquidity constraint is one explanation for that phenomenon to prevail.

Proof of Proposition 61

We cannot directly use Proposition 47 because the implicit price density does not satisfy condition $E\widetilde{\pi} = 1$. We escape this difficulty by defining the following program:

$$\overline{v}(y) = \max_{(c_1,\ldots,c_n)} \sum_{t=1}^{n} \frac{\beta^{t-1}}{\sum_{\tau=1}^{n} \beta^{\tau-1}} u(c_t) \tag{17.6}$$

$$s.t. \sum_{t=1}^{n} \frac{\rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}} c_t = y. \tag{17.7}$$

Observe that the implicit state price density satisfy condition $E\widetilde{\pi} = 1$. Moreover, we have

$$v(z) = (\sum_{\tau=1}^{n} \beta^{\tau-1})\overline{v}(\frac{z + w\sum_{t=2}^{n} \rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}}) \tag{17.8}$$

Combining Proposition 47 with the assumption that P is larger than $2A$ yields

$$v'(z) = \frac{\sum_{\tau=1}^{n} \beta^{\tau-1}}{\sum_{\tau=1}^{n} \rho^{1-\tau}} \overline{v}'(\frac{z + w\sum_{t=2}^{n} \rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}}) \leq \frac{\sum_{\tau=1}^{n} (1/\beta)^{1-\tau}}{\sum_{\tau=1}^{n} \rho^{1-\tau}} u'(\frac{z + w\sum_{t=2}^{n} \rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}}).$$

If $1/\beta \leq \rho$, it implies that

$$v'(z) \leq u'(\frac{z + w\sum_{t=2}^{n} \rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}}).$$

The end of the proof is as for the case $\beta = \rho = 1$: we just have to check that $u'(\frac{z+w\sum_{t=2}^{n} \rho^{1-t}}{\sum_{\tau=1}^{n} \rho^{1-\tau}}) \leq u'(z) = v_c'(z)$ when z is less than w. This is immediate, since $z \leq w$ is equivalent to

$$\frac{z + w \sum_{t=2}^{n} \rho^{1-t}}{\sum_{t=1}^{n} \rho^{1-t}} \geq z.$$

The concavity of u concludes the proof.■

Chapter 18

The saving-portfolio problem

The concept of prudence has been introduced to treat the effect of an exogenous future risk on current savings. But many risks in life are endogenous. For example, people don't want to insure their house in full, and they purchase risky assets. The level of saving clearly affects the optimal behavior to these endogenous risks. Symmetrically, the opportunity to take risk should influence the saving behavior. Thus, it is important to examine the joint consumption and portfolio decisions.

Samuelson (1969) and Merton (1969,1971) were the first to provide a complete analysis of this problem in the case of HARA utility functions. Their motivation was not so much about the optimal saving behavior. Rather, they wanted to determine the optimal portfolio decision in a dynamic framework with intermediary consumption. Drèze and Modigliani (1972) focussed instead on the effect of the portfolio strategy on the optimal saving strategy. We will review the two branches of this literature in this Chapter.

18.1 Precautionary saving with an endogenous risk

We start with the analysis of the impact of an endogenous risk on the optimal saving. We consider two cases. In the first case, the agent lives in an Arrow-Debreu world with the possibility to take any risky position on the aggregate risk in the economy. In the second case, the agent lives in a world with one risk free asset and one risky asset. No other contingent contract can be exchanged. This refers to the standard portfolio problem.

18.1.1 The case of complete markets

Suppose that the state price density in the Arrow-Debreu economy be $\widehat{\pi}(.)$. The portfolio-saving problem is written as follows in this case:

$$\max\ u_0(w_0 - s) + Eu_1(c(\widetilde{x}))$$

$$s.t.\ E\widehat{\pi}(\widetilde{x})c(\widetilde{x}) = s + E\widehat{\pi}(\widetilde{x})w_1.$$

At date 0, the agent saves s from his income and sells his sure income w_1 on financial markets, yielding $w_1 E\widehat{\pi}(\widetilde{x})$. With this capital, he purchases a portfolio $c(.)$ of Arrow-Debreu securities. His decision problem is to select s and $c(.)$ in order to maximize his lifetime expected utility. Define $\rho = [E\widehat{\pi}(\widetilde{x})]^{-1}$ and $\pi(x) = \widehat{\pi}(x)\rho$. The budget constraint takes the following form:

$$E\pi(\widetilde{x})c(\widetilde{x}) = w_1 + \rho s$$

with the property that $E\pi(\widetilde{x}) = 1$. Notice that we can then rewrite this problem by using the value function v that has been defined by program (12.2) for price kernel $\pi(.)$ and utility function $u \equiv u_1$. It yields

$$\widehat{v}(w_0) = \max\ H(s) = u_0(w_0 - s) + v(w_1 + \rho s). \tag{18.1}$$

In parallel to what we have done for the exogenous future risk, we address the question of how does the presence of this endogenous risk affect the optimal saving. An application of this problem is for the effect of an economic policy related to pension funds. If one forces savings to be invested in bonds alone, will it induce more or less savings than if one allows pension funds to be invested in bonds and stocks? Intuitively, two contradictory effects are at work here if investing in stocks is allowed. The global effect is split into the effect of a pure risk and the effect of an increase in the expected future income. First, because it is optimal to purchase some shares since $E\widetilde{x} > 0$, the presence of the portfolio risk has a positive precautionary saving effect under prudence. The intensity of this effect is increasing in P. Second, this option to take risk has a positive wealth effect: it increases the expected future income of the agent. We know that this

increase in future income has a negative impact on the optimal saving if u_1 is concave (income smoothing). The intensity of this opposite effect is increasing in the degree of concavity of u_1, i.e., it is increasing in $A = -u_1''/u_1'$. Thus, the global effect of the option to invest in risky assets is ambiguous under prudence. But we may conjecture that it will have a positive impact on the aggregate saving if absolute prudence is large enough compared to absolute risk aversion. This is what we now show.

We compare the solution of the above problem with respect to the one in which the agent is forced to put all his wealth in a risk free portfolio. This second problem is written as:

$$\max\ u_0(w_0 - s) + u_1(w_1 + \rho s)$$

where $\rho = [E\widehat{\pi}(\widetilde{x})]^{-1}$ is the return of the safe portfolio. The first order condition is written as

$$u_0'(w_0 - s^*) = \rho u_1'(w_1 + \rho s^*).$$

The optimal saving when the agent is allowed to take a risky position is larger than s^* if $H'(s^*)$ is positive, i.e., if

$$v'(w_1 + \rho s^*) \geq u_1'(w_1 + \rho s^*).$$

The following Proposition is then a direct application of Proposition 47.

Proposition 63 *In an economy with a complete set of Arrow-Debreu securities, allowing investors to take a risky position increases (resp. reduces) their optimal saving if and only if absolute prudence is larger (resp. smaller) than twice absolute risk aversion.*

18.1.2 The case of the standard portfolio problem

Let us suppose alternatively that the only possible risky positions that are allowed be linear in the aggregate risk $\widetilde{x}$. Formally, the problem is to choose s and α in order to solve the following problem:

$$\max_{\alpha,s} \; u_0(w_0 - s) + Eu_1(w_1 + \rho s + \alpha\widetilde{x}).$$

This can be seen as a saving-portfolio problem in which s is the aggregate saving, α is the amount that is invested in a risky asset whose excess return is $\widetilde{x}$, whereas $s - \alpha$ is invested in a risk free asset whose gross return is ρ. Without loss of generality, we assume that $E\widetilde{x} > 0$. Because the choice of α has no impact on the first period utility, this problem can be seen as a dynamic problem in which s is selected in period 0 and α is selected in period 1. The value function can be written as:

$$v(z) = \max_{\alpha} \; Eu_1(z + \alpha\widetilde{x}) \tag{18.2}$$

and the first period problem can be rewritten as:

$$\widehat{v}(w_0) = \max_{s} \; H(s) = u_0(w_0 - s) + v(w_1 + \rho s). \tag{18.3}$$

We compare this solution to the one in which the agent cannot invest in the risky asset, in which case s maximizes $u_0(w_0 - s) + u_1(w_1 + \rho s)$. Let s^* be the solution of this problem. For exactly the same argument as in the case of complete markets, the opportunity to take a risky position increases saving if $Eu_1'(w_0+\rho s^*+\alpha^*\widetilde{x}) \geq u_1'(w_0 + \rho s^*)$, where α^* is the optimal demand for the risky asset of the agent with wealth $w_0 + \rho s^*$.

Normalizing α^* to unity and using condition (15.2), this condition is equivalent to

$$\forall z, \widetilde{x} : \; E\widetilde{x}u_1'(z + \widetilde{x}) = 0 \implies Eu_1'(z + \widetilde{x}) \geq u_1'(z) \tag{18.4}$$

where $z = w_1 + \rho s^*$. Because w_0, w_1 and ρ are arbitrary, we want this condition to hold for any z. We already know from Proposition 28 that condition (18.4) holds if and only if $P(z) \geq 2A(z)$ for all z.[1] This concludes the proof of the following Proposition, which is due to Gollier and Kimball (1997).[2]

[1] By property (13.11), this condition is equivalent to $T'(z) \geq 1$, where $T = 1/A$ is absolute risk tolerance.

[2] This result can also be proven by using Proposition 4 with $g(x) = x$, $h(x, 1) = u'(z + x)$ and $h(x, 2) = x^{-1}\left[u_1'(z) - u_1'(z + x)\right]$.

Proposition 64 *Consider the standard portfolio problem. The opportunity to invest in a risky asset raises (resp. reduces) the aggregate saving if and only if absolute prudence is larger (resp. smaller) than twice the absolute risk aversion.*

18.1.3 Discussion of the results

For the two structures of financial markets, we obtain that $P \geq 2A$ is necessary and sufficient for the property that allowing risk-taking enhances the willingness to save. As expected, prudence is not sufficient to guarantee that allowing households to invest their savings in risky asset increases aggregate saving. Notice that the condition is also stronger than decreasing absolute risk aversion which means $P \geq A$. Because credible estimations of absolute or relative prudence are still missing in the literature, let us look at the special case of HARA utility functions

$$u(z) = \zeta(\eta + \frac{z}{\gamma})^{1-\gamma}$$

for which we have:

$$A(z) \quad = \quad (\eta + \frac{z}{\gamma})^{-1}$$

and

$$P(z) \quad = \quad \frac{\gamma+1}{\gamma}(\eta + \frac{z}{\gamma})^{-1}.$$

Thus, for HARA utility functions, absolute prudence equals $\frac{\gamma+1}{\gamma}$ times the absolute risk aversion. Thus, absolute prudence is larger than twice absolute risk aversion if $\frac{\gamma+1}{\gamma} \geq 2$, or $\gamma \leq 1$. In the specific case of CRRA functions where γ represents relative risk aversion, this would mean that relative risk aversion be less than 1. As explained in section 3.6, this condition is not true for most households in the real economy. Thus, if we believe in HARA functions and in relative risk aversion being larger than unity, we should expect that the option to invest in risky assets would rather reduce aggregate saving. The (expected) consumption smoothing effect is here stronger than the precautionary saving effect.

More generally, observe that condition $P(z) \geq kA(z)$ is equivalent to

$$\frac{R(w)}{w}\left[(1-k)R(z)+1\right] \geq R'(w)$$

where $R(z)$ is the relative risk aversion measured at z. It implies that a set of sufficient conditions for $P \geq 2A$ is nonincreasing relative risk aversion with a coefficient of relative risk aversion bounded below unity. On the contrary, nondecreasing relative risk aversion combined with a relative risk aversion larger than 1 is enough for $P \leq 2A$, i.e., for options to take risk to reduce saving.

Another important remark is that our assumption that there are only two periods is without loss of generality, at least in the complete markets case. Indeed, we know that property that P is uniformly larger or uniformly smaller than $2A$ is inherited by the value function in this latter case. Thus, the effect of removing the opportunity to invest in complete markets this year is unambiguous. But only the property that $P > 2A$ is preserved in the case of the portfolio problem. In consequence, removing the opportunity to invest in stocks reduces saving if $-u'''/u''$ is larger than $-2u''/u'$. But it may be possible that it is also the case for some utility functions that violate this condition.

18.2 Optimal portfolio strategy with consumption

In the previous section, we examined how does the possibility to purchase stocks influence consumption. In this section, we examine how does the possibility to consume over time affects the optimal dynamic portfolio strategy. This is done basically by combining results from Chapters 10 and 14. The optimal dynamic portfolio strategy when there is no intermediary consumption was the focus of Chapter 10. In particular, we showed that a longer time horizon increases or reduces the demand for stocks depending upon whether absolute risk tolerance is convex or concave, at least in an Arrow-Debreu economy. When investors consume their wealth over several periods, there is also a time diversification effect that is at play. WSe assume in this section that $u_0 \equiv u$ and $u_1 \equiv \beta u$.

We add a risk-taking decision at date 0 to the problem that we examined in the previous section. The attitude towards risk at date $t = 0$ is given by the measure of concavity of the value function $\widehat{v}$ that has been defined in equations (18.1) or (18.3) depending upon whether the second period problem is standard or Arrow-Debreu.The only difference of this problem with respect to problem (10.1) for

example is that value function v is replaced by $\widehat{v}$ to express the time diversification effect. The effect of the existence of the second period on the optimal portfolio structure at date 0 is influenced by several factors:

- a background risk effect generated by the option to invest in risk in the second period;
- a wealth effect generated by this option;
- a time diversification effect generated by the possibility to allocate gains and losses at date 0 over two periods;
- an income effect generated by the flow of future incomes.

Combining results on whether u is more concave than v and whether v is more concave than v yields conditions for a positive effect of time horizon on risk-taking. An illustration is provided by the following Proposition.

Proposition 65 *Suppose that markets are complete and that investors have no income at old age ($w_1 = 0$). The young investor is always less risk-averse than the old investor if and only if the absolute risk tolerance of the felicity function u is convex and subhomogeneous.*

Proof: A direct consequence of Propositions 48 and 1.■

Another illustration is when $w_0 = w_1$, $\beta = \rho = 1$ and CRRA. In that case, the agent with two periods to go should invest exactly twice as much as the agent who has only one period to go. Only the time diversification effect is at play. The case of CRRA is examined in more details in the next section.

18.3 The Merton-Samuelson model

The discrete version of the Samuelson-Merton consumption-portfolio problem is written as follows:

$$v_t(y) = \max_{c_t, \alpha_{t+1}} \qquad u(c_t) + \beta E v_{t+1}(w_{t+1} + \rho(y - c_t) + \alpha_{t+1}\widetilde{x}_{t+1}), \tag{18.5}$$

with $v_n \equiv u$. As in the previous section, we assume that there is no other uncertainty than the portfolio risk. More specifically, we assume that the consumer

receives w_t at date $t = 1, ..., n$ from his human capital. Variable y is interpreted as the cash-on-hand at date t. At that date with cash y, the agent consumes c_t and carry the remaining by investing in assets to date $t + 1$. He invests α_{t+1} in stocks whose excess return from date t to date $t + 1$ is $\widetilde{x}_{t+1}$. The remaining is invested in the risk free asset whose return is ρ. We assume here that there is no serial correlation in the returns of stocks.

The first-order conditions for this problem are written as

$$u'(c_t) = \beta\rho E v'_{t+1}(\rho(y - c_t) + w_{t+1} + \alpha_{t+1}\widetilde{x}_{t+1}), \tag{18.6}$$

and

$$E\widetilde{x}_{t+1} v'_{t+1}(\rho(y - c_t) + w_{t+1} + \alpha_{t+1}\widetilde{x}_{t+1}) = 0. \tag{18.7}$$

Using the Envelope Theorem together with condition (18.6) for date $t + 1$ rather than t, we obtain that $v'_t(y) = u'(c_t(y))$. It implies that we can rewrite the above condition as follows:

$$u'(c_t) = \beta\rho E u'(\widetilde{c}_{t+1}) \tag{18.8}$$

$$E\widetilde{x}_{t+1} u'(\widetilde{c}_{t+1}) = 0. \tag{18.9}$$

Random variable $\widetilde{c}_{t+1}$ is the optimal consumption at date $t + 1$ given the cash-on-hand z at t. We recognize the Euler equation, which requires that the discounted marginal utility of 1 dollar must be equalized over time.

It happens that this model can be solved analytically when the utility function is HARA. To simplify the notation, let us limit the analysis to the case where $u(c) = \dfrac{c^{1-\gamma}}{1-\gamma}$. When facing a dynamic problem like this, the only practical way to find a solution is to guess it, then to check whether the trial solution verifies the Bellman equation 18.5. Remember that we showed in Section 10.2.2 that the value function is HARA when u is HARA in the pure investment problem. We could try to check whether this is also the case when intermediary consumption is introduced. Our trial solution is thus $v_t(y) = K_t \dfrac{(y + k_t)^{1-\gamma}}{1-\gamma}$, for some constants

$K_1, ..., K_{n-1}$ and $k_1, ..., k_{n-1}$to be specified. In particular, we will try with $k_t = \sum_{\tau=t+1}^{n} \rho^{t-\tau} w_\tau$. Let us define constants a_t, b_t and d_t as follows:

$$E\widetilde{x}_t(1 + a_t\widetilde{x}_t)^{-\gamma} = 0, \tag{18.10}$$

$$b_t = E(1 + a_t\widetilde{x}_t)^{-\gamma}, \tag{18.11}$$

$$d_t = E(1 + a_t\widetilde{x}_t)^{1-\gamma}. \tag{18.12}$$

One can then check that conditions (18.6) and (18.7) hold with

$$c_t(y) = \frac{[\beta\rho K_{t+1}b_{t+1}]^{-1/\gamma}\,\rho}{1 + [\beta\rho K_{t+1}b_{t+1}]^{-1/\gamma}\,\rho}\left[y + \sum_{\tau=t+1}^{n} \rho^{t-\tau} w_\tau\right] \tag{18.13}$$

and

$$\alpha_{t+1}(y) = \frac{a_{t+1}\rho}{1 + [\beta\rho K_{t+1}b_{t+1}]^{-1/\gamma}\,\rho}\left[y + \sum_{\tau=t+1}^{n} \rho^{t-\tau} w_\tau\right]. \tag{18.14}$$

To check that this trial solution was a good guess, we need to check Bellman equation (18.5). After simplifying by z, it can be written as

$$K_t = \left[\frac{\rho}{1 + [\beta\rho K_{t+1}b_{t+1}]^{-1/\gamma}\,\rho}\right]^{1-\gamma}\left[[\beta\rho K_{t+1}b_{t+1}]^{\dfrac{\gamma-1}{\gamma}} + \beta d_{t+1}K_{t+1}\right]. \tag{18.15}$$

Since this define a well-defined recursive equation for $K_1, K_2, ..., K_{n-1}, K_n = 1$, we obtain the confirmation that the trial solution at hand is indeed the optimal solution of our problem. It is generated by the optimal strategy characterized by conditions (18.13) and (18.14). Observe that $y + \sum_{\tau=t+1}^{n} \rho^{t-\tau} w_\tau$ is the aggregate wealth at date t, which is the sum of cash-on-hand at that time and the discounted

value of future incomes. The optimal portfolio strategy is to fix for each date the share of aggregate wealth that is invested in stocks. The optimal consumption strategy consists in fixing for each date a share of aggregate wealth that will be consumed at that date. These rates can vary over time, but they are independent of the actual aggregate wealth. In particular, it implies that the marginal propensity to consume out of wealth at each date is constant.

In the following numerical example, we assume that the stable distribution of excess returns is normal, with a mean of 0.06 and a variance of 0.027. As reported by Kocherlakota (1996), this is close to the distribution of the excess return of Standard and Poor 500 over the last century. We also assume that $\rho = 1.01$, which is also its historical value over the same period. Finally, we consider an agent with no impatience and a relative risk aversion equaling $\gamma = 2$. We obtain that $a = 1.12$. It is then easy to calculate the $\{K_t\}_{t=1,\ldots,n}$ and the corresponding optimal marginal propensity to consume together with the optimal share of aggregate wealth invested in stocks. In Figure 18.1, we reported the optimal MPC as a function of the time horizon. Without surprise, it decreases with n basically as $1/n$. The important point here is to compare the optimal MPC with the one that would have been optimal without the opportunity to invest in stocks. This optimal MPC when the risk free asset is the only asset than can be carried forward is represented in Figure 18.1 by the dotted curve. Because $P < 2A$ in this simulation, the opportunity to invest in stocks reduces the willingness to save. It increases consumption at every date.

It appears that the opportunity to invest in stocks has almost no effect the optimal MPC for short horizons. But the effect is quite important for longer horizon. In our simulation, with 20 years to go, the MPC when there is no opportunity to invest in stock equals 0.0500. But people who are allowed to invest in stock during their lifetime would rather select a MPC equaling 0.0595. This represents a 20% increase in the MPC.

A final remark on the optimal investment strategy that can be rewritten as

$$\alpha_{t+1}(y) = a\left[\rho(y - c_t) + \sum_{\tau=t+1}^{n} \rho^{t+1-\tau} w_\tau\right]. \tag{18.16}$$

The bracketed term is the above equality is the aggregate wealth available at the beginning of period $t + 1$. We conclude that the optimal investment in stocks

Figure 18.1: Optimal consumption expressed as a proportion of aggregate wealth with (plain) and without (dotted) the opportunity to invest in stocks.

is a constant share of the aggregate wealth. This is a consequence of the very specific properties of CRRA that yields an absolute risk tolerance that is linear and homogeneous.

18.4 Concluding remark

The level of wealth accumulated is obviously an important element to determine the optimal portfolio structure. Reciprocally, the opportunity to take risk in the future is expected to affect the optimal strategy of wealth accumulation. In this Chapter, we examined the joint determination of the portfolio strategy and the consumption strategy. The opportunity to take risk has a mixed effect on saving. Under prudence, it tends to increase the precautionary saving. But this option to take risk also generates a wealth effect. It makes the consumer implicitly wealthier in the future, and that tends to reduce her willingness to save today. We showed that the global effect of financial markets on saving is positive if prudence is larger than twice the absolute risk aversion.

Researchers dealing with this problem use to think about it by using the Merton-Samuelson model with a CRRA utility function. This has the big advantage that the complete solution to the problem can be obtained analytically. We presented a simple simulation of the Merton-Samuelson model by using the historical data of stock returns.

Chapter 19

Disentangling risk and time

There is some logic for decreasing marginal utility of consumption to generate at the same time an aversion to risk in each period and an aversion to non-random fluctuations of consumption over time . If the marginal gain of one more unit of consumption is less than the marginal loss due to one unit reduction of consumption, agents will reject the opportunity to gamble on a fifty-fifty risk to gain or lose one unit of consumption. For the same reason, if their current consumption plan is smooth, patient consumers will reject the opportunity to exchange one unit of consumption today against one unit of consumption tomorrow. Thus, the degree of concavity of the utility function on consumption measures the willingness to insure consumption across states and the willingness to smooth consumption across time. The argument for having a unique function to characterize these two features of agents' preferences is based on the assumption that their objective over their lifetime is the sum of their expected utility at each periods. This makes the objective function additive across states and across time. If people live for two periods, the canonical model has the following functional:

$$U(c_0, \widetilde{c}_1) = u_0(c_0) + Eu_1(\widetilde{c}_1). \tag{19.1}$$

Kreps and Porteus (1978) and Selden (1978) proposed an alternative model to disentangle the attitudes toward consumption smoothing over time and across states. Their models are equivalent. Because the presentation by Kreps and Porteus (1978) is more intuitive, we hereafter follow their definitions.

19.1 The model of Kreps and Porteus

The model is a direct extension of the additive model (19.1). It is written as

$$U(c_0, \widetilde{c}_1) = u_0(c_0) + u_1(v^{-1}(Ev(\widetilde{c}_1))), \tag{19.2}$$

where u_0, u_1 and v are three increasing functions. It is easier to interpret this preference functional by using the certainty equivalent functional $m(\widetilde{c}_1)$ which is defined as $v(m) = Ev(\widetilde{c}_1)$. U can then be rewritten as

$$U(c_0, \widetilde{c}_1) = u_0(c_0) + u_1(m(\widetilde{c}_1)). \tag{19.3}$$

We see that the lifetime welfare is computed by performing two different operations. First, one computes the certainty equivalent m of the future uncertain consumption $\widetilde{c}_1$ by using utility function v. This is done in a atemporal context. Thus, the concavity of v measures the degree of risk aversion alone. Second, one evaluates the lifetime welfare by summing the utility of the current consumption and the utility of the future certainty equivalent consumption, using functions u_0 and u_1. Because all uncertainty has been removed in this second operation, the concavity of these two functions are related to preferences for consumption smoothing over time.

Some particular cases of Kreps-Porteus preferences are worthy to mention. For example, if v and u_1 are identical, we are back to the additive model (19.1). Thus, the standard additive model is a particular case of Kreps-Porteus preferences. But this model is much richer because of its ability to disentangle preferences with respect to risk and time. Suppose for example that v is the identity function, but u_0 and $u_1 \equiv u_0$ are concave. In that case, the agent is willing to smooth the expected consumption over time, despite he is risk neutral. At the opposite side of the spectrum, we can imagine a risk-averse agent who is indifferent toward consumption smoothing. This would be the case if v is concave, but u_0 and u_1 are linear.

This additional degree of freedom in the way we shape individual preferences does not simplify the analysis of optimal behaviors and their comparative statics. Without surprise, many researchers as Epstein and Zin (1991) considered a particular specification of the model by using power functions. They supposed that

$$v(z) = \frac{z^{1-\gamma}}{1-\gamma} \quad \text{and} \quad u_1(z) = \beta u_0(z) = \frac{z^{1-\alpha}}{1-\alpha}, \tag{19.4}$$

with γ and α being two nonnegative scalars. Observe that γ is the degree of relative risk aversion whereas α is the relative resistance to intertemporal substitution. When $\gamma = \alpha$, the model is additive. As we have seen before, there are some empirical evidence in favor of the hypothesis that agents are more risk-averse than they are resistant to intertemporal substitution.

19.2 Preferences for an early resolution of uncertainty

In the classical case with $u_1 \equiv v$, the timing of the resolution of the uncertainty does not matter for the consumer. Suppose that the consumption plan under consideration be $(c_0, \widetilde{c}_1)$, where $\widetilde{c}_1$ is random. If the realization of $\widetilde{c}_1$ is not expected to be known before $t = 1$, the lifetime expected utility would be measured by $u_0(c_0) + Eu_1(\widetilde{c}_1)$. Suppose alternatively that the realization of $\widetilde{c}_1$ is expected to be known at $t = 0$. Conditional to $\widetilde{c}_1 = c_1$, the lifetime utility is $u_0(c_0) + u_1(c_1)$. Ex ante, before knowing $\widetilde{c}_1$, the expected lifetime utility is measured by $E\left[u_0(c_0) + u_1(\widetilde{c}_1)\right] = u_0(c_0) + Eu_1(\widetilde{c}_1)$. Thus, in the classical case, agents are indifferent about the timing of the resolution of the uncertainty affecting their consumption plan.

This is not the case anymore with Kreps-Porteus preferences. Indeed, if $\widetilde{c}_1$ is not expected to known before $t = 1$, the lifetime welfare of the agent is measured according to equation (19.3), with $v(m(\widetilde{c}_1)) = Ev(\widetilde{c}_1)$. On the contrary, suppose that the realization of $\widetilde{c}_1 = c_1$ is observed at $t = 0$. Then, because obviously $m(c_1) = c_1$ under certainty, the lifetime utility conditional to c_1 is $u_0(c_0) + u_1(c_1)$, as in the classical case. Ex ante, the lifetime welfare of the agent equals $E\left[u_0(c_0) + u_1(\widetilde{c}_1)\right]$. This is generally not equal to $u_0(c_0) + u_1(m(\widetilde{c}_1))$. We conclude that an agent with Kreps-Porteus preferences is in general not indifferent to the timing of the resolution of uncertainty.

We say that an agent has preferences for an early resolution of uncertainty (PERU) if he prefers to observe $\widetilde{c}_1$ at date $t = 0$ than at $t = 1$, whatever the distribution of $\widetilde{c}_1$. This is the case when

$$E\left[u_0(c_0) + u_1(\widetilde{c}_1)\right] \geq u_0(c_0) + u_1(m(\widetilde{c}_1)), \tag{19.5}$$

or, equivalently, when

$$u_1^{-1}(Eu_1(\widetilde{c}_1)) \geq m(\widetilde{c}_1) = v^{-1}(Ev(\widetilde{c}_1)). \tag{19.6}$$

In words, PERU requires that the certainty equivalent is always larger when using function u_1 than when using function v. Using Proposition (8), this is true if and only if u_1 is less concave than v.

Proposition 66 *Suppose that consumers have Kreps-Porteus preferences described by equation (19.3) with $m(\widetilde{c}_1) = v^{-1}(Ev(\widetilde{c}_1))$. Then, they prefer an early resolution of uncertainty if and only if u_1 is less concave than v in the sense of Arrow-Pratt.*

As mentioned before, there are some reasons to believe that agents are more risk-averse than they are fluctuation-averse, i.e., that v is more concave than u_1. This observation is compatible with preferences for an early resolution of uncertainty.

In Part VII of this book, we will examine in details another reason for why people prefer an early resolution of uncertainty. Early information may have a positive value because informed agents will take better decisions. This explanation is completely orthogonal to the one presented in this section, since we assumed here that agents don't take any decision.

19.3 Prudence with Kreps-Porteus Preferences

In this section, we reexamine the simplest decision problem in which time and risk are intricate, namely precautionary saving. We reconsider the model presented in section **??** by introducing Kreps-Porteus preferences. The income flow of the consumer who lives for two periods is $(w_0, w_1 + \widetilde{x})$, with $E\widetilde{x} = 0$. The gross return on savings between $t = 0$ and 1 is ρ. With Kreps-Porteus preferences, the optimal level of savings is given by the solution of the following maximization problem:

$$\max_s \quad u_0(w_0 - s) + u_1(M(s)) \tag{19.7}$$

where $M(s)$ is the certainty equivalent future consumption as a function of savings. It is implicitly defined by

$$v(M(s)) = Ev(w_1 + \widetilde{x} + \rho s). \tag{19.8}$$

Notice first that this problem is not necessarily well behaved, even if u_0, u_1 and v are concave. Indeed, the concavity of the objective function (19.7) requires that function M be concave in s. This is true in particular if we assume that the consumer in fluctuation-neutral, i.e., that $u_0 = u_1$ are linear. We already discussed whether the certainty equivalent is concave in a change in wealth. In Proposition **??** , we showed that M is concave if the absolute risk tolerance of v is concave. In particular, it is concave in the HARA case. If $-v'/v''$ is not concave, it may be possible that the solution to the first-order condition of problem (19.7) be a local minimum. We hereafter assume that this is not the case.

As in section **??**, we want to examine the impact of the presence of uncertainty on the optimal level of savings. Adding zero-mean risk $\widetilde{x}$ to future consumption increases saving if it raises its marginal value. Without $\widetilde{x}$, the marginal value of saving is measured by $\rho u_1'(w_1 + \rho s)$, whereas it equals $\rho M'(s) u_1'(M(s))$ when $\widetilde{x}$ is added to the future consumption. The agent is prudent if

$$M'(s) u_1'(M(s)) \geq u_1'(w_1 + \rho s) \tag{19.9}$$

where

$$v(M(s)) = v(w_1 + \widetilde{x} + \rho s) \tag{19.10}$$

and

$$M'(s) = \frac{Ev'(w_1 + \widetilde{x} + \rho s)}{v'(M(s))}. \tag{19.11}$$

Condition (19.9) does not hold in general. We must restrict the set of increasing and concave functions u_1 and v to obtain prudence. One such condition is that u_1 and v to be identical, with v' being convex. We can generalize this result when u_1 and v are not identical if we assume that u_1 is *more* concave than v. Remember that it implies that u_1 is centrally more risk-averse than v at $w_1 + \rho s$. Since $M(s)$ is less than $w_0 + \rho s$, this means that

$$\frac{u_1'(M(s))}{u_1'(w_1 + \rho s)} \geq \frac{v'(M(s))}{v'(w_1 + \rho s)}. \tag{19.12}$$

If we combine this with the assumption that v' is convex, which implies that $v'(w_1 + \rho s) \leq Ev'(w_1 + \widetilde{x} + \rho s)$, we obtain that

$$\frac{u_1'(M(s))}{u_1'(w_1 + \rho s)} \geq \frac{v'(M(s))}{Ev'(w_1 + \widetilde{x} + \rho s)}. \tag{19.13}$$

This is equivalent to condition (19.9). The problem with this sufficient condition is that we must assume that fluctuation aversion is larger than risk aversion, which seems to be unrealistic.

Another sufficient condition for prudence is that v be DARA, as shown by Kimball and Weil (1992). The proof of this statement comes from the fact that $M'(s)$ is larger than unity under DARA. Remember that DARA means that

$$Ev(w_1 + \widetilde{x} + \rho s) = v(M) \quad \Longrightarrow \quad Ev'(w_1 + \widetilde{x} + \rho s) \geq v'(M). \tag{19.14}$$

Observe also that $M(s)$ is smaller than $w_1 + \rho s$ under the concavity of v. The concavity of u_1 implies in turn that $u_1'(M(s))$ is larger than $u_1'(w_1 + \rho s)$. We conclude that v DARA is sufficient for prudence in the framework of Kreps-Porteus preferences. Notice that v DARA is also necessary in the limit case where u_1 is linear, where condition (19.9) simplifies to $M'(s) \geq 1$. To sum up, without any information on the felicity function u_1 than its concavity, the necessary and sufficient condition for prudence is that v be DARA. This is more restrictive than in the classical case in which we obtained the weaker condition $v''' \geq 0$. This is the cost to be paid to allow for any concave felicity function.

These findings are summarized in the following Proposition.

Proposition 67 *Suppose that consumers have Kreps-Porteus preferences described by equation (19.3) with $m(\widetilde{c}_1) = v^{-1}(Ev(\widetilde{c}_1))$. Then, the agent is prudent*

1. *if v' is convex and u_1 is more concave than v;*
2. *independent of the concave felicity function u_1 if and only if v is DARA.*

19.4 Conclusion

We have seen that the classical additive model can be generalized to disentangle consumption fluctuation aversion from risk aversion. The Kreps-Porteus criterion presented in this chapter has the merit to do that without eliminating the basic additive nature of the measurement of welfare under uncertainty, or of lifetime welfare under certainty. This is only when risk and time are considered together that this criterion differs from what we have seen before. Our standard techniques can still be used in this generalized framework. Because it is more general, additional conditions will be required for the comparative statics analysis of risk taking and consumption. In particular we have seen that consumers will accumulate precautionary saving in the face of uncertainty if the utility function is DARA, which

is stronger than prudence in the classical framework. Several other extensions of classical results remain to be performed with Kreps-Porteus preferences.

Part VI

Equilibrium prices of risk and time

Chapter 20

Efficient risk sharing

This chapter is devoted to the characterization of efficient allocations of risks in the economy. We start with a simple exchange economy.

20.1 The case of an exchange economy

Agents have different characteristics that are represented by θ, where θ belongs to some characteristics set Θ. The distribution function over θ is denoted H. We normalize the population to unity. The characteristics θ of an agent influence his preferences that are represented by a utility function $u(c, \theta)$, where c denotes the quantity that is consumed by this agent. We assume that u is increasing and concave in c.

Agents face risk. The simplest way to represent uncertainty in such an economy is to rely on the notion of a state of the world. The state of the world that can prevail in the future is denoted $x \in X$, where X is the set of possible states of the world. There is an agreed-upon cumulative distribution function F on X in the economy. $\widetilde{x}$ denotes the random variable whose cumulative distribution function is F. There is a single consumption good in the economy. Each agent θ gets an endowment of the consumption good that is contingent to the state of the world that will prevail. Let $\omega(x, \theta)$ denote the endowment of agent θ in state x. Let also $z(x)$ denote the quantity of the consumption good that is available per head in state x, i.e.,

$$z(x) = \int_{\Theta} \omega(x, \theta) dH(\theta) = E\omega(x, \widetilde{\theta}).$$

We say that there is an aggregate uncertainty if z is sensitive to variations of x.

An allocation in this economy is characterized by a function $c(x, \theta)$ which is the quantity consumed in state x by agents having characteristics θ. This function is in fact a risk sharing rule. This risk allocation is said to be feasible if the quantity of the good consumed per head in any state equals the quantity available per head in that state:

$$Ec(x, \widetilde{\theta}) = z(x) \qquad \forall x \in X. \tag{20.1}$$

The expected utility that is enjoyed by agent θ with this allocation is

$$Eu(c(\widetilde{x}, \theta), \theta) = \int_X u(c(x, \theta), \theta) dF(x). \tag{20.2}$$

Following the standard way by which economists use to define efficiency, we say that a feasible allocation $c(.,.)$ is Pareto-efficient if there is no other feasible allocation that increases the expected utility of at least one category of agents without reducing the expected utility of the other categories. Consider now any positive function $\lambda : \Theta \to R^+$. It is easy to check that any solution to the following program is a Pareto efficient allocation of risk:

$$\max_{c(.,.)} \quad \int_\Theta \lambda(\theta) Eu(c(\widetilde{x}, \theta), \theta) dH(\theta) \tag{20.3}$$

subject to the feasibility condition (20.1). This is done by contradiction: suppose that the solution to program (20.3) is not Pareto-efficient. Then, by definition, there would exist an alternative feasible risk-sharing rule $\widehat{c}(.,.)$ such that $Eu(\widehat{c}(\widetilde{x}, \theta), \theta)$ is larger than $Eu(c(\widetilde{x}, \theta), \theta)$ for all $\theta \in \Theta$, with a strict inequality for at least one θ of positive measure. It implies that

$$\int_\Theta \lambda(\theta) Eu(\widehat{c}(\widetilde{x}, \theta), \theta) dH(\theta) > \int_\Theta \lambda(\theta) Eu(c(\widetilde{x}, \theta), \theta) dH(\theta), \tag{20.4}$$

which contradicts the assumption that $c(.,.)$ is a solution to program (20.3). We conclude that for any positive function $\lambda(.)$, there is an associated Pareto-efficient risk-sharing rule $c(.,.)$ which is the solution to program (20.3). It can also be

shown that, under risk-aversion, we get all Pareto-efficient risk-sharing rules by using this procedure.

We now characterize the Pareto-efficient risk-sharing allocations. These allocations have two basic properties. The first one is governed by the so-called mutuality principle.

20.2 The mutuality principle

Proposition 68 *In a Pareto-efficient risk-sharing, the consumption of each member of the pool depends upon the state of the world only through the aggregate wealth available in that state.*

The proof of this result is as follows. Consider two different states of nature, y and y', that are such that the wealth available per head is the same in these two states:

$$z(y) = z(y'). \tag{20.5}$$

The mutuality principle states that, for the allocation to be efficient, all agents must consume the same amount in the two states: $c(y, \theta) = c(y', \theta)$ for all $\theta \in \Theta$. The intuition behind this result can be found in the way we prove it. By contradiction, suppose that there exists an agent θ_0 such that $c(x, \theta_0)$ is larger than $c(x', \theta_0)$. Then, consider an alternative risk-sharing rule $\widehat{c}(.,.)$ which is exactly the same as $c(.,.)$, except in states y and y'. More specifically, we assume that

$$\widehat{c}(x, \theta) = \begin{cases} c(x, \theta) & \text{if} \quad x \notin \{y, y'\} \\ \dfrac{pc(y, \theta) + p'c(y', \theta)}{p + p'} & \text{if} \quad x \in \{y, y'\} \end{cases} \tag{20.6}$$

where p and p' are the probability associated respectively to state y and y'. In words, agents replace their random consumption conditional to the realization of y or y' by its expected value. It is obvious that this alternative allocation is feasible. It is also obvious that all agents favor this alternative allocation of risk, since it yields a mean-preserving reduction in risk. This is seen by observing that, for each $\theta \in \Theta$,

$$c(\widetilde{x}, \theta) = \widehat{c}(\widetilde{x}, \theta) + \widetilde{\varepsilon}(\theta) \tag{20.7}$$

where $\widetilde{\varepsilon}(\theta)$ is a random variable that is zero if $\widetilde{x}$ is not y or y', and that is otherwise distributed as $(\frac{p'}{p+p'}(c(y,\theta)-c(y',\theta)), \frac{p}{p+p'}; \frac{p}{p+p'}(c(y',\theta)-c(y,\theta)), \frac{p'}{p+p'})$. Observe that $E\left[\widetilde{\varepsilon}(\theta) \mid \widetilde{x}\right] = 0$. Thus, $\widehat{c}(\widetilde{x},\theta)$ is less risky than $c(\widetilde{x},\theta)$ in the sense of Rothschild-Stiglitz, for all θ. We conclude that $c(.,.)$ may not be Pareto-efficient, since $\widehat{c}$ is unanimously preferred to it.

This proof shows that the mutuality principle states that all risks that can be diversified away by ex-ante contracting must be washed away. This is a necessary condition for Pareto-efficiency. Thus, there is no loss of generality to characterize a Pareto-efficient allocation of risk by a function $w(z,\theta)$ where $c(x,\theta) = w(z(x),\theta)$. In short, $w(z,\theta)$ is the consumption of agents θ in all states whose wealth per head is z. Agents θ are fully insured against the social risk $\widetilde{z}$ if $w(z,\theta)$ is independent of z. More generally, the sensitivity of w to changes of z measures the risk borne by category θ.

20.3 The sharing of the social risk

Let G denote the cumulative distribution function of the wealth per head $\widetilde{z}$. Using the mutuality principle, we may rewrite problem (20.3) as follows:

$$\max_{w(.,.)} \quad E\lambda(\widetilde{\theta})u(w(\widetilde{z},\widetilde{\theta}),\widetilde{\theta}) = \int\int \lambda(\theta)u(w(z,\theta),\theta)dG(z)dH(\theta) \tag{20.8}$$

$$s.t. \quad \int w(z,\theta)dH(\theta) = z \quad \forall z. \tag{20.9}$$

There is a simple intuitive way to solve to program. Let us define function v as

$$v(z) = \max_{w(z,.)} \quad \int \lambda(\theta)u(w(z,\theta),\theta)dH(\theta) \tag{20.10}$$

$$s.t. \quad \int w(z,\theta)dH(\theta) = z. \tag{20.11}$$

Obviously, problem (20.8) is just to find the social risk-sharing rule w that maximizes $Ev(\widetilde{z}) = \int v(z)dG(z)$. This is obtained by a pointwise optimization $\max$ $v(z)$, for all z, i.e., by solving the set of problems (20.10).

Here is the intuition for using this procedure. Without loss of generality, assume that $E\lambda(\widetilde{\theta}) = 1$. We can then interpret $\lambda(\theta)dH(\theta)$ as a probability. When the group has to determine ex ante the rule that they will use to share the cake $\widetilde{z}$, obvious conflicts of interest take place in the debate. Each category θ wants to increase its share $w(.,\theta)$. Suppose alternatively that agents have to determine the sharing rule $w(.,.)$ *before* observing their characteristics. To be more precise, consider the following structure: all agents are assigned the same probability $\lambda(\theta)dH(\theta)$ to belong to category θ. Before learning their type, all members of the group vote about the sharing rule to be used. Agents have to select the sharing rule under the veil of ignorance about their characteristics. At the second stage, nature chooses in a random way the characteristics of the agents according to the distribution function H for $\widetilde{\theta}$. At the last stage, $\widetilde{z}$ is realized, and the allocation of the cake is made according to the sharing rule that has been selected earlier.[1]

We see that we added an additional uncertainty to the problem. The notion of the veil of ignorance transformed an interpersonal, distributional problem into a well-known problem of decision of uncertainty. Indeed, at the first stage of the game, agents are all alike. They unanimously vote for the feasible allocation that maximizes the expected utility given the uncertainty about their characteristics. In fact, the uncertainty about the size of the cake becomes secondary here: all agents will agree to maximize there expected utility in each state z. That is, the allocation of any size z of the cake that will unanimously be selected is the one that solves problem (20.10).

Now, observe that problem (20.10) is technically equivalent to the selection of a portfolio of contingent claims, as described by program (13.1). Formally, this is seen by performing the following change of variable: define random variable $\widetilde{\pi}$ as the one having cumulative distribution function K, with $dK(\theta) = \lambda(\theta)dH(\theta)/E\lambda(\widetilde{\theta})$. Using this change of variable in program (20.10) makes it equivalent to program (13.1). We can thus use all properties that have been developed in section 13 to characterize the properties of an efficient risk-sharing rule. The first-order condition of problem (20.10) is written as

$$\lambda(\theta)u'(w(z,\theta),\theta) = v'(z). \tag{20.12}$$

[1]Notice that the only probability distribution that is compatible with the actual composition of the population is to assign $\lambda(\theta) = 1$ for all θ. This observation induced some economists to favor the Pareto-efficient solution that solves (20.3) with this neutral weighting function. This is the utilitarian approach to fairness.

This condition is formally equivalent to condition (13.3), yielding $w(z,\theta) = C(1/\lambda(\theta), z)$. This condition implies that

$$\frac{u'(w(z_1,\theta),\theta)}{u'(w(z_2,\theta),\theta)} = \frac{v'(z_1)}{v'(z_2)} \quad \forall\theta \in \Theta. \tag{20.13}$$

A Pareto-efficient allocation is such that the marginal rate of substitution between any two states of the world is the same for all agents. Using property (**??**), the full differentiation of the first-order condition directly implies that

$$\frac{\partial w}{\partial z}(z,\theta) = \frac{T(w(z,\theta),\theta)}{T_v(z)}, \tag{20.14}$$

where $T(w,\theta) = -u'(w,\theta)/u''(w,\theta)$ and $T_v(z) = -v'(z)/v''(z)$ are the indexes of absolute risk tolerance of respectively $u(.,\theta)$ and $v(.)$. This property of Pareto-efficient risk-sharing rules has first been obtained by Wilson (1968). It states that the share of the social risk that must be allocated to category θ is proportional to its degree of absolute risk tolerance. The larger the risk tolerance, the larger the share of the social risk.

The feasibility constraint implies that

$$E\frac{\partial w}{\partial z}(z,\widetilde{\theta}) = 1. \tag{20.15}$$

Any reduction of wealth per head must be absorbed by an equivalent average reduction of consumption per head. Combining conditions (20.14) and (20.15) implies that

$$T_v(z) = ET(w(z,\widetilde{\theta}),\widetilde{\theta}). \tag{20.16}$$

Proposition 69 *In a Pareto-efficient risk-sharing, the sensitivity of each member's consumption to changes in the wealth per head equals the ratio of the corresponding agent's absolute risk tolerance to the average absolute risk tolerance in the group.*

We now have a complete picture of how a group should share a risk. The first rule is to put all wealths in common, and to decide about a way to share the aggregate wealth. This is the mutuality principle, which states that the final

allocation of the consumption good should not depend on who brought what in the common pool. If the aggregate wealth is uncertain, those who are less risk-averse should bear a larger share of the risk. If no member is risk neutral, T_v is finite, and all members should bear a share of the social risk. This is another application of second-order risk aversion, which implies that nobody should be excluded from the sharing of the social risk, even if one's risk aversion is very large. It is always optimal to allocate some of the risk to each member, because the effect of risk on welfare is a second-order effect.

It is interesting to determine the conditions under which the Pareto-efficient risk-sharing rules are linear in z. A linear risk-sharing rule is easy to implement, since it is done by organizing lump-sum, ex ante, transfers combined with assigning a fixed percentage of the size of the cake to each member. In a market economy, this is done by offering to each agent a portfolio of a risk free asset and a share of the stock market. It allows for a Two-Fund Separation Theorem. The simplest case is when T is independent of z, i.e., when absolute risk aversion is constant. Another case is directly derived from Proposition 45. Namely, the Pareto-efficient risk sharing rules are linear if all members of the pool have the same HARA utility function. In the other cases, the Two-Fund Separation Theorem does not apply as efficient risk-sharing rules are locally concave or convex. As noticed by Leland (1980), a convex sharing rule can be obtained by a portfolio of call options on the social risk.

20.4 Group's attitude towards risk

Suppose that the group's members agreed to share their ex post wealth according to the Pareto-efficient rule associated to $\lambda(.)$. What should be the group's attitude towards risk? The answer to this question can be found in the characteristics of function v that has been defined by equation (20.10).

It may useful to start by providing an intuition for why the sharing rule (20.14) is Pareto-efficient. It will help us understanding how should the group behave in the face of uncertainty. Let us suppose that the wealth per head is initially certain and equal to z. It yields an allocation $w(z, \theta)$ of the sure wealth in the group. Suppose now that the group is considering taking a small risk $\widetilde{\varepsilon}$ per head,with mean $\mu > 0$ and variance σ^2. Should the group accept this risk? It depends upon how this risk would be shared in the group.

A possible procedure that is alternative to the one used above is to select the sharing rule that maximizes the mean of all certainty equivalents of the risk borne

by the different categories of members. Let $a(\theta)$ represent the sensitivity of category's θ consumption to change in $\widetilde{\varepsilon}$. This means that category's θ final wealth is $w(z,\theta) + a(\theta)\widetilde{\varepsilon}$. Feasibility imposes that $Ea(\widetilde{\theta}) = 1$. By the Arrow-Pratt approximation, the certainty equivalent of the risk borne by category θ is measured by

$$CE(\theta) = a(\theta)\mu + 0.5\frac{(a(\theta))^2\sigma^2}{T(w(z,\theta),\theta)}. \tag{20.17}$$

The problem is thus to select the function $a(.)$ that solves

$$CE = \max_{a(.)|Ea(\widetilde{\theta})=1} \quad E\left[CE(\widetilde{\theta})\right] \tag{20.18}$$

where CE is the mean of the optimized members' certainty equivalent of the risk borne by the group. The first-order condition to this problem is written as

$$\mu + \frac{a(\theta)\sigma^2}{T(w(z,\theta),\theta)} = \xi \tag{20.19}$$

for all $\theta \in \Theta$. Using the feasibility constraint, we can eliminate the Lagrangian multiplier ξ to obtain

$$a(\theta) = \frac{T(w(z,\theta),\theta)}{T_v(z)}. \tag{20.20}$$

This is the same risk-sharing rule than we obtained in the previous section. Thus, at least for small risks, Pareto-efficient risk-sharing rules are those which maximize the sum of the certainty equivalents of the risk borne by the group's members. There is other lessons that we can learn from this exercise. Observe that, using the optimal sharing rule (20.20), the certainty equivalent of the risk borne by category θ is

$$CE(\theta) = \frac{T(w(z,\theta),\theta)}{T_v(z)}\left[\mu + 0.5\frac{\sigma^2}{T_v(z)}\right] \tag{20.21}$$

Also, the optimized certainty equivalent per head of risk $\widetilde{\varepsilon}$ up equals

$$CE = E\left[CE(\widetilde{\theta})\right] = \mu + 0.5\frac{\sigma^2}{T_v(z)}. \tag{20.22}$$

Remember now the initial question, which was to determine whether the group should accept risk $\widetilde{\varepsilon}$ or not. Equation (20.21) provides a clear answer to this question: as soon as $CE = \mu + 0.5\frac{\sigma^2}{T_v(z)}$ is positive, all members of the pool are willing to vote in favor of accepting the risk. The way this social risk would be shared forces unanimity in the group. This is an important result, which is easily extended to larger risks. Indeed, the first-order condition (20.12) tells us that all marginal utilities are proportional to each other with a Pareto-efficient risk-sharing rule. It implies that the sensitivity of marginal utility to changes in z, which is the members' aversion to the social risk, is the same for all members of the group. We conclude that, once the sharing rule stipulated by function $\lambda(.)$ is determined, there is no conflict of interest among the group's members to select a strategy of decision under uncertainty.

The second lesson is that $T_v(z)$ is the relevant measure of absolute risk tolerance that must be taken into account to determine whether a risk should be undertaken. This is apparent for example in the way the group's certainty equivalent is measured in equation (20.22). Function $T_v(.)$ is the group's absolute tolerance to risk per head. In fact, the attitude towards risk of a group $\left\{u(w,\theta) \mid \widetilde{\theta} \sim H\right\}$ implementing the rule associated with weight function $\lambda(.)$ is the same as the attitude towards risk of a group of identical agents v implementing the symmetric, fair allocation. They all have utility function $v(.)$ that is defined by (20.10). An agent with such a utility function is called a "representative agent". Of course, such an egalitarian group of identical agent has an attitude towards a social risk $\widetilde{\varepsilon}$ per head that is identical to its member's attitude towards such a risk. This is because the group will transfer it unchanged to each of its members.

When the group is heterogenous, or when it does not implement the egalitarian solution, risk-sharing arrangements within the group is likely to affect the group's attitude towards undiversifiable risks. To illustrate, let us consider an economy of identical agents that implements an unequal risk-sharing arrangement. This is done by using a non-constant weight function $\lambda(.)$. Parallel to Proposition 48, we directly derive from Jensen's inequality and condition (20.14) that an unequal risk-sharing arrangement reduces the willingness of the group to take risk if its members' absolute risk tolerance is concave with wealth. To obtain this result, we

don't need to make any assumption on whether risk aversion must be increasing or decreasing. But under increasing absolute risk tolerance, the intuition is as follows: introducing wealth inequality in the economy alters the group's attitude towards risk in two ways. First, poorer agents are more risk averse. They will take a smaller share of the social risk. This reduces the group's tolerance to risk. Second, wealthier people are less risk averse. They increase their share of the bearing of the social risk. That raises the group's willingness to take risk. Under linear risk tolerance, it happens that the two effects exactly compensate each other, so that the group's willingness to take risk is unaffected. Some intuition can be obtained by observing that, for a small social risk, the certainty equivalent of the risk borne by each member is proportional to his absolute risk tolerance (equation (20.21)). Combining this with the fact the group's certainty equivalent of social risk per head is the mean of the members' certainty equivalent implies the result.

Two final remarks on the group's attitude towards risk. First, observe that, by the mutuality principle, only the undiversifiable risk matters for the analysis of the effect of an additional risk to the group's wealth. The second remark can be seen as an illustration of the first. Namely, if the group is large enough, his behavior towards risk should be almost risk neutral. This point is made for example by Arrow and Lind (1970), but it is systematically used in the theory of finance. To explain this, consider an economy with n identical agents. The fair allocation of the consumption good is implemented. These assumptions are made for simplicity and they can be generalized easily by using the trick of the representative agent. Each agent is allocated an initially sure amount w of the consumption good. The group is considering accepting an aggregate risk $\widetilde{y}$. Since this risk would be shared in a fair way, each agent would get $1/n$ of it. If the size n of the population is large enough, this is a small risk. Under the assumption of second order risk aversion, the members of the group will unanimously accept the risk as soon as $E\widetilde{y}$ is positive. That is, the group is neutral to the aggregate risk. This can also be seen by using condition (20.16) together with relaxing the assumption that the size of the population be one. Alternatively, take $\int dH(\theta) = n$. Condition (20.16) is then written as

$$T_v(y) = nT(w + \frac{y}{n}) \tag{20.23}$$

where y is the realization of the aggregate risk $\widetilde{y}$ and T is the absolute risk tolerance of the identical agents in the economy. We see that the group's degree of absolute risk tolerance tends to infinity when the size of the group tends to infinity.

Stated differently, the group's willingness to accept risk increases proportionally with the size of the group. This is true for a group of identical agents that implement the fair allocation of the consumption good. When agents are not identical, or when the distribution of wealth is unequal, this is true up to the second-order effect of the size of the group on the degree of concavity of the representative agent's utility function.

20.5 Introducing time and investments

Taking risk is usually done by investing money today in the anticipation of an uncertain payoff in the future. Thus, taking risk had to do not only with the attitude towards risk, but also with the attitude towards time. Introducing time in the model is easy if the members' utility function is time-additive. Consider a model with two dates indexed by $t = 0$ or $t = 1$. There is a single good that is perishable. At $t = 0$, there is uncertainty about the state of the world that will prevails at $t = 1$. The uncertainty is given by the state-contingent endowment $\omega(\widetilde{x}, \theta)$, where $\widetilde{x}$ is the random variable characterizing the state of the world that will prevail at date $t = 1$. Let $\omega_0(\theta)$ denote the endowment of agents θ that is available to them at date $t = 0$. Let also z_0 denote the wealth per head that is available at that time, i.e., $z_0 = E\omega_0(\widetilde{\theta})$. An allocation is characterized by a pair $(c_0(\theta), c(x, \theta))$ where $c_0(\theta)$ is the consumption of category θ at date $t = 0$. It is feasible if condition (20.1) is satisfied together with the date 0 feasibility constraint $Ec_0(\widetilde{\theta}) = z_0$. The lifetime utility attained by category θ is given by

$$u(c_0(\theta), \theta) + \beta E u(c(\widetilde{x}, \theta), \theta) \tag{20.24}$$

where β is the discount factor.

Using the same argument as in the static model, a feasible allocation is Pareto-efficient if it is a solution of the following class of problems:

$$\max_{c_0(.);c(.,.)} \quad E\lambda(\widetilde{\theta}) \left[u(c_0(\widetilde{\theta}), \widetilde{\theta}) + \beta u(c(\widetilde{x}, \widetilde{\theta}), \widetilde{\theta}) \right] \tag{20.25}$$

under the above mentioned feasibility constraints. Let us define a new state of the world, $x_0 \notin X$, and let compound random variable $\widetilde{y}$ takes value x_0 with probability $1/\beta$, otherwise it is $\widetilde{x}$. Finally, let us define $c(x_0, \theta) = c_0(\theta)$ and $z(x_0) = z_0$. Then, program (20.25) can be rewritten as

$$\max_{c_0(.);c(.,.)} \quad E\lambda(\widetilde{\theta})u(c(\widetilde{y},\widetilde{\theta}),\widetilde{\theta}) \tag{20.26}$$

$$s.t. \quad Ec(y,\widetilde{\theta}) = z(y) \quad \forall y \in X \cup x_0. \tag{20.27}$$

This is formally equivalent to program (20.3), with an additional state of the world that represents date $t = 0$ allocation of the perishable good. Since program (20.3) is solved pointwise anyway, that does not affect at all the properties of the solution. This is another illustration of the symmetric role that are played by time and states of nature in an additive world.

The mutuality principle applied in this intertemporal framework tells us that if there is a state of the world that yields the same wealth per head than at date 0, the consumption patterns in that state should be the same that at date 0. Turning now to the group's attitude towards time, we can use the same trick than for the analysis of the group's attitude towards risk. Let us define the representative agent's utility function v as we did in equation (20.10). Consequently, the representative agent's objective function can be written as

$$v(z_0) + \beta E v(\widetilde{z}) \tag{20.28}$$

It implies that the concavity of v does not only represent the degree of risk aversion of the group. It also represents the group's resistance to the intertemporal substitution of consumption.

Chapter 21

The equilibrium price of risk and time

In this chapter, we examine the Pareto-efficient allocation of risk that is generated by competitive markets for contingent claims. It will allow us to examine the driving forces that govern the pricing of risk and time in such an economy. In fact, the model that we present in this Chapter is a two-period version of the Continuous-time Capital Asset Pricing Model (CCAPM).

21.1 An Arrow-Debreu economy

The description of our economy is exactly the same as in section 20.1. Time will be added to this picture later in this chapter. We now assume that at the beginning of the period, viz before the realization of $\widetilde{x}$, agents are in a position to trade their state contingent endowment. This is done by opening a complete set of markets for contingent claims. Agents are allowed to buy or sell any contingent claim that promises the delivery of one unit of the consumption good if and only if state x is realized. Following the notation that we used in chapter 6, let $\pi(x)$ be the price of such a contract per unit of probability. Markets for contingent claims are assumed to be competitive. This means that consumers take prices as given, and that prices clear markets. The problem of agent θ is written as

$$\max_{c(.,\theta)} \quad Eu(c(\widetilde{x},\theta),\theta) \tag{21.1}$$

$$s.t. \quad E\pi(\widetilde{x})c(\widetilde{x},\theta) = E\pi(\widetilde{x})\omega(\widetilde{x},\theta) \tag{21.2}$$

where $c(x,\theta)$ is the consumption of agent θ in state x. In other terms, $c(x,\theta) - \omega(x,\theta)$ is his demand for the contingent claim associated to state x. This problem is equivalent to problem (12.2), where $E\pi(\widetilde{x})\omega(\widetilde{x},\theta)$ is the wealth of the agent evaluated at the market prices.

An allocation in this Arrow-Debreu economy is characterized by a couple $(c(.,.),\pi(.))$. Such an allocation is a competitive equilibrium if

1. $c(.,\theta)$ is the solution to program (21.1) given price kernel $\pi(.)$, for all $\theta \in \Theta$;
2. all markets clear, i.e.,

$$E\left[c(x,\widetilde{\theta}) - \omega(x,\widetilde{\theta})\right] = 0 \quad \forall x \in X. \tag{21.3}$$

We hereafter assume that such an equilibrium exists and that it is unique. The remaining of this chapter is devoted to the analysis of the properties of this equilibrium.

21.2 The first theorem of welfare economics

One of the first achievements of mathematical economics has been to prove that the competitive equilibrium is Pareto-efficient. This is seen by writing the first-order condition associated to agent θ's portfolio problem (21.1). It yields

$$u'(c(x,\theta),\theta) = \pi(x)\xi(\theta) \quad \forall x \in X,\ \forall \theta \in \Theta. \tag{21.4}$$

It implies that

$$\frac{u'(c(x_1,\theta),\theta)}{u'(c(x_2,\theta),\theta)} = \frac{\pi(x_1)}{\pi(x_2)} \tag{21.5}$$

This means that the competitive equilibrium is such that the marginal rate of substitution between any two states of the world is the same for all agents, and is equal to the ratio of the corresponding state prices. But, as stated by equation (20.13), this property characterizes a Pareto-efficient allocation of risk. Thus, the competitive equilibrium is Pareto-efficient.

All properties of Pareto-efficient risk-sharing arrangements that we derived in the previous chapter hold for the competitive equilibrium. It means for example that if there are two states yielding the same wealth per head $E\omega(x_1,\widetilde{\theta}) =$

$E\omega(x_2, \widetilde{\theta})$, all agents will consume the same quantity of the consumption good in these two states: $c(x_1, \theta) = c(x_2, \theta)$. By condition (21.5), it implies that the equilibrium prices will be the same in these two states: $\pi(x_1) = \pi(x_2)$. Thus, there is no loss of generality to characterize a state of the world by the wealth per head that is available in that state (-as we did in the previous chapter), or by the equilibrium price in that state (as we did in chapter 6).

21.3 Pricing Arrow-Debreu assets

Because the competitive equilibrium is Pareto efficient, we know that there exists a representative agent that represents this economy. More precisely, consider an economy in which all agents are alike, with the same utility function v and the same state contingent endowment $z(\widetilde{x})$. Obviously, the homogeneity of the population eliminates any source of mutually advantageous transaction. It implies that the autharcic allocation in which agents just consume their endowment in every state is the competitive equilibrium of this economy. The competitive price kernel which sustains such an allocation is such that

$$\frac{v'(z(x_1))}{v'(z(x_2))} = \frac{\pi(x_1)}{\pi(x_2)} \qquad \forall z_1, z_2. \tag{21.6}$$

Because competitive prices are defined up to multiplicative factor, we can take $v'(z(x)) = \pi(x)$ for all x. The price kernel is just the marginal utility of the representative agent. Thus, obtaining the price kernel $\pi(.)$ of the heterogenous economy directly yields the characterization of the corresponding representative agent of this economy. In fact, the modern theory of finance often use the argument the other direction. Using the fact that a representative agent exists for any Arrow-Debreu economy, pricing models postulate an homogenous economy of identical agents. This simplification is thus without loss of generality, since any heterogenous economy has its representative agent representation. The difficulty of obtaining function v is put aside with this approach. We will follow this line in the remaining of this chapter.

We presented in section 20.4 the technique to build the utility function of the representative agent from the utility functions of the different categories of agents in the heterogenous economy. We showed there that the risk aversion of actual agents is transferred to the representative agent's attitude toward risk. The absolute risk tolerance of the representative agent is the mean absolute risk tolerance

of actual agents in each state:

$$T_v(z(x)) = ET(c(x, \widetilde{\theta}), \widetilde{\theta}) \tag{21.7}$$

where $c(x, \theta)$ is the consumption of agent θ in state x, in the competitive equilibrium. If one agent is risk-neutral, so is the representative agent.

To sum up the approach, we assume that the economy is composed of identical agents with the same utility function v and the same state contingent endowment $z(\widetilde{x})$. There is no trade at the competitive equilibrium in this Arrow-Debreu economy. The price kernel that sustains such a competitive equilibrium is just the marginal utility of consumption:

$$\pi(x) = v'(z(x)). \tag{21.8}$$

Under the risk aversion of the representative agent, the equilibrium price of a state contingent asset is decreasing with the wealth per head available in that state. This induces the agent to reduce his consumption in states of relative scarcity of the consumption good. Let $\phi(z) = \pi(z^{-1}(z))$ be the equilibrium state price a function of the wealth per head. The sensitivity of the price to a change of the wealth per head equals the absolute risk aversion of the representative agent:

$$\frac{-\phi'(z)}{\phi(z)} = A_v(z) = \frac{-v''(z)}{v'(z)}. \tag{21.9}$$

The larger the absolute risk aversion, the larger the price sensitivity. When the risk aversion is large, the agent is ready to pay much to smooth his consumption across states. Large price discrepancies in state contingent prices are then necessary to force him to bear the social risk.

21.4 Pricing by arbitrage

In the previous section, we provided a simple way to price Arrow-Debreu assets. But most assets in the real world are not Arrow-Debreu assets. In this section, we explain how to price any asset that is not an Arrow-Debreu asset. This is done by using the technique of arbitrage.

An asset in this static economy is entirely described by the state contingent payoff that it will generate at the end of the period. Let $q_j(x)$ be the payoff of asset j in state x. Knowing the competitive price of the Arrow-Debreu assets, we can

derive the competitive price P_j of asset j in the following manner. Build a portfolio of Arrow-Debreu assets with, for any state x, $q_j(x)$ units of the Arrow-Debreu asset associated to state x. The market value of this portfolio is $E\pi(\widetilde{x})q_j(\widetilde{x})$. This portfolio has exactly the same risk profile as asset j itself. Thus, we now have two strategies of taking risk $q_j(\widetilde{x})$. By an argument of arbitrage, the cost of using these two strategies must be the same. Suppose by contradiction that the equilibrium price of asset j is larger than the market value of the equivalent portfolio of Arrow-Debreu assets: $P_j > E\pi(\widetilde{x})q_j(\widetilde{x})$. A risk-free arbitrage that has a positive payoff is obtained by selling short asset j, and buying the portfolio of Arrow-Debreu assets that duplicates it. This strategy has no effect on the contingent consumption plan of the investor, since the promise to pay $q_j(x)$ due to the short-selling of asset j is just compensated by the payoff generated by the portfolio of contingent claims. Thus, this strategy is risk free, with a zero net payoff with probability 1 at the end of the period. But it allows the investor to relax his budget constraint, since he extract a sure profit $P_j - E\pi(\widetilde{x})q_j(\widetilde{x}) > 0$ from this strategy. This cannot be an equilibrium, as all agents will duplicate this strategy like crazy. It would raise the supply of asset j up to infinity. The symmetric argument by contradiction can be made if the price of the asset is less than the market value of the portfolio which duplicates it. We conclude that the price of any asset $q_j(\widetilde{x})$ must equal the market value of the portfolio of Arrow-Debreu assets which has the same risk profile than it, i.e.,

$$P_j = E\pi(\widetilde{x})q_j(\widetilde{x}) \tag{21.10}$$

There are two branches in the pricing theory in finance. The first branch is based on the optimal portfolio management of investors. Demand curves and supply curves for the different assets are build from solving the investors' expected utility maximization problem. A typical example of this branch is given in the previous section. Another example is the capital asset pricing theory. The second branch is pricing by arbitrage, in which the price of a subset of assets is taken as given. The equilibrium price of all other assets whose payoffs can be duplicated by this subset are derived by arbitrage. Option pricing formulas are typically obtained by following this approach. For example, the equilibrium price of a call option on asset j with strike price q_0 equals $E\pi(\widetilde{x})\max(0, q_j(\widetilde{x}) - q_0)$. No other information on preferences than those already contained in the prices of the subset is necessary to derive such prices.

21.5 The competitive price of risk

What are the determinants of risky asset pricing? To answer this question, let us normalize the price of the asset promising an unconditional delivery of one unit of the good to unity:

$$E\pi(\widetilde{x}) = 1. \tag{21.11}$$

In this section, this asset is the numeraire. It implies that the equilibrium price of the asset promising a payoff profile $q_j(\widetilde{x})$ equals

$$P_j = Eq_j(\widetilde{x}) + cov(\pi(\widetilde{x}), q_j(\widetilde{x})) = Eq_j(\widetilde{x}) + cov(v'(z(\widetilde{x})), q_j(\widetilde{x})). \tag{21.12}$$

The equilibrium price of any asset equals its expected payoff plus a risk premium. This risk premium is measured by the covariance of the payoff of the risky asset with the price kernel. This means that only the undiversifiable risk in $q_j(\widetilde{x})$ gets a risk premium. This result is a classical one in the modern theory of finance. The intuition is simple. Consider an asset whose payoff is independent of the social risk $z(\widetilde{x})$. This means that adding a marginal amount of this asset in everyone's portfolio increases everyone's expected utility in proportion of the expected payoff of the asset:

$$\left.\frac{\partial Ev(z(\widetilde{x}) + kq_j(\widetilde{x}))}{\partial k}\right|_{k=0} = Eq_j(\widetilde{x})v'(z(\widetilde{x})) = Eq_j(\widetilde{x}) \tag{21.13}$$

if $q_j(\widetilde{x})$ and $z(\widetilde{x})$ are independent. This is another consequence of the fact that expected utility exhibits second order risk aversion. Because a marginal increase in the holding of this asset does not affect the portfolio's risk, this asset's risk will not get a risk premium. On the contrary, if a risk premium would be included, all investors would increase their demand for it.

The risk premium equals the covariance of the state contingent price with the payoff function of the asset. Since the sensitivity of the price kernel to changes in the wealth per head equals the absolute risk aversion of the representative agent, we understand that the risk premium will be an increasing function of the degree of risk aversion and of the covariance of the wealth per head and the payoff of the asset. This can be seen by approximating $v'(z(x))$ in equation (21.12) in a linear way around z_0, where z_0 is defined by $v'(\widetilde{z}) = 1 = Ev'(z(\widetilde{x}))$. It yields

$$\begin{aligned} P_j &= Eq_j(\widetilde{x}) + E\left[\{v'(z(\widetilde{x})) - v'(z_0)\}\{q_j(\widetilde{x}) - Eq_j(\widetilde{x})\}\right] \\ &\simeq Eq_j(\widetilde{x}) + v''(z_0)E\left[\{z(\widetilde{x}) - z_0\}\{q_j(\widetilde{x}) - Eq_j(\widetilde{x})\}\right] \\ &= Eq_j(\widetilde{x}) - A_v\, cov(z(\widetilde{x}), q_j(\widetilde{x})), \end{aligned} \tag{21.14}$$

where $A_v = A_v(\widehat{z})$ is the market's degree of risk aversion, and $z(\widetilde{x})$ is the market risk. This is the standard Capital Asset Pricing (CAPM) formula, which states that the equilibrium price of the risk associated to an asset equals the product of the market's degree of risk aversion by the covariance of the asset risk with the market risk.

CAPM formula (21.14) uses to be written in a slightly different way. Consider the price P_M of an asset that exactly duplicates the risk profile of the market. Its price equals

$$P_M = Ev'(z(\widetilde{x}))z(\widetilde{x}). \tag{21.15}$$

Using the same linear approximation as above, we obtain

$$P_M \simeq \mu_M - A_v\sigma_M^2 \tag{21.16}$$

where $\mu_M = Ez(\widetilde{x})$ and $\sigma_M^2 = Var(z(\widetilde{x}))$ are respectively the mean and the variance of the market payoff. Eliminating A_v from (21.14) and (21.16) yields

$$P_j = Eq_j(\widetilde{x}) - \Pi_M\ \beta_j \tag{21.17}$$

where $\Pi_M = Ez(\widetilde{x}) - P_M$ is the so-called market price of risk, and $\beta_j = cov(z(\widetilde{x}), q_j(\widetilde{x}))/\sigma_M^2$ is the so-called beta of asset j. The market price of risk, Π_M, is the expected return of equity. The mean value of the beta is one in the economy. It has the same sign as the covariance of the asset risk with the market risk. Both Π_M and the betas can quite easily be estimated by using market data.

21.6 The competitive price of time

Up to now in this chapter, we considered a static model. This did not allow us to examine the determinants of the risk free rate of interest, i.e., the price of time. To do this, we reexamine the two-date model that we used in section 20.5. The (perishable) wealth per head is z_0 at date $t = 0$, and $z(x)$ in state x at date $t = 1$. As explained in that section, adding time to the initial picture does not change the representative agent result: to any heterogenous or unequal economy, there is an homogenous and fair economy that duplicates its behavior toward risk and time. Let v be the utility of the representative agent. His objective is to maximize his lifetime utility $v(c_0) + \beta Ev(c(\widetilde{x}))$, where β represents here the discount factor. The equilibrium conditions are $c_0 = z_0$ and $c(x) = z(x)$, for saying that the consumption is not transferable across time or states.

The difference with what we did above is the possibility to consume at date $t = 0$. The Arrow-Debreu security associated to state x guarantees the delivery of one unit of the consumption good in state x, date 1, against the paiement of price $\pi(x)$ at date 0. Obviously, we cannot normalize the price in such a way that $E\pi(\widetilde{x}) = 1$, because the numeraire is now the consumption good itself: at date 0, promises of delivery at date 1 are exchanged against some units of the good today. It implies that $\pi(.)$ contains information on the attitude towards risk and time at the same time. The equilibrium price of time can be measured from the number of units of the consumption good that can be obtained at date $t = 1$ for certain when giving up one unit of the consumption good at date $t = 0$ on the credit market. Let $1+r$ be this number. It is an equilibrium if, at the margin, the representative agent is indifferent upon signing up a credit contract, from the autharcic allocation. This is the case if

$$-v'(z_0) + \beta(1+r)Ev'(z(\widetilde{x})) = 0, \tag{21.18}$$

or, equivalently,

$$r = \frac{v'(z_0)}{\beta Ev'(z(\widetilde{x}))} - 1. \tag{21.19}$$

We now have a clear picture of the determinants of the price of time in an Arrow-Debreu economy, as they have been stated by Bohm-Bawerk. First, the larger the impatience of agents (β small), the larger is the equilibrium interest rate to induce them to select a feasible consumption pattern over time. This is a determinant generated by a pure preference for the present. There is a second effect due to the growth of GNP per head over time. To isolate this second effect, suppose that $\beta = 1$ (no impatience) and $z(\widetilde{x}) = z_1 > z_0$ with probability one. Then, using a first-order approximation, we get that

$$1 + r = \frac{v'(z_0)}{v'(z_1)} = \frac{v'(z_0)}{v'(z_0) + (z_1 - z_0)v''(z_0)} \simeq 1 + A_v(z_0)(z_1 - z_0). \tag{21.20}$$

It implies that

$$r \simeq z_0 A_v(z_0)\frac{(z_1 - z_0)}{z_0}, \tag{21.21}$$

i.e., the risk-free rate is approximately equal to the product of the relative risk aversion of the representative agent by the growth rate of the economy. Here, the

relative risk aversion plays the role of measuring the resistance to intertemporal substitution, as explained in section 14.3. When the expected growth rate is large, households are willing to transfer more of the future purchasing power into today consumption, for an income smoothing reason. To induce them not to do that, the interest rate must be increased.

A third determinant of the equilibrium risk free rate is uncertainty about the growth rate of incomes. As we have seen in chapter 14, the fact that future incomes are uncertain makes prudent agents more willing to accumulate precautionary savings to forearm themselves against potential losses of their incomes. In order to induce them not to do that, the risk free rate must be reduced. This is seen from equation (21.19) by observing that if v' is convex,

$$\frac{v'(z_0)}{\beta E v'(z(\widetilde{x}))} - 1 \leq \frac{v'(z_0)}{\beta v'(E z(\widetilde{x}))} - 1. \tag{21.22}$$

Under prudence, an increase in the uncertainty affecting the future growth of the economy reduces the risk free rate.

To sum up, we can say that expectations about the growth of the economy are the driving forces behind the equilibrium price of time. Assuming risk aversion and prudence, a large expected growth increases the equilibrium interest rate, whereas the uncertainty affecting this growth rate has an adverse effect on it.

21.7 Disentangling the prices of risk and time

As said above, the introduction of time makes the price kernel $\pi(.)$ a little bit more complex to interpret, because time and risk are intertwined in it. Different prices can be inferred from it. In addition to the risk free rate that we examined above, we obtain that the price of equity in terms of today's consumption good is P_e that must satisfy the following equilibrium condition:

$$0 = \frac{\partial}{\partial k}\left[v(z_0 - kP_e) + \beta E v((1+k)z(\widetilde{x}))\right]\big|_{k=0} = -P_e v'(z_0) + \beta E z(\widetilde{x}) v'(z(\widetilde{x})), \tag{21.23}$$

which implies that

$$P_e = \frac{\beta E z(\widetilde{x}) v'(z(\widetilde{x}))}{v'(z_0)}. \tag{21.24}$$

The pure preference for the present and the expectations about the growth of the economy will affect the price of equity.

Because equity yields an uncertain payoff tomorrow against a paiement *today*, the price of equity, P_e, is a measure that combines the price of risk and the price of time. The equilibrium price of risk can be obtained by looking at the price of a financial asset called a "future". A "future" contract on equity is exactly the same as the equity itself, except that the paiement of the price will be called at date $t = 1$, not at $t = 0$. Thus, a future contract organizes the exchange of a sure amount of the good tomorrow against an uncertain amount of the same good tomorrow. Its price is thus the price of risk in the economy. The price of this future is denoted P_f. It must satisfy the following equilibrium condition:

$$0 = \frac{\partial}{\partial k}\left[v(z_0) + \beta E v((1+k)z(\widetilde{x}) - kP_f)\right]\big|_{k=0} = \beta\left[Ez(\widetilde{x})v'(z(\widetilde{x})) - P_f E v'(z(\widetilde{x}))\right], \tag{21.25}$$

which implies that

$$P_f = \frac{Ez(\widetilde{x})v'(z(\widetilde{x}))}{Ev'(z(\widetilde{x}))}. \tag{21.26}$$

Because P_f is the equilibrium price of risk in this economy, we may inquire about whether an increase in risk aversion of the representative agent reduces this price. Let v_1 be more risk-averse than v_2. More risk aversion reduces the equilibrium price of equity if

$$\frac{Ez(\widetilde{x})v_2'(z(\widetilde{x}))}{Ev_2'(z(\widetilde{x}))} = P_{f2} \Longrightarrow \frac{Ez(\widetilde{x})v_1'(z(\widetilde{x}))}{Ev_1'(z(\widetilde{x}))} \leq P_{f2}. \tag{21.27}$$

Let $\widetilde{y}$ denote random variable $z(\widetilde{x}) - P_{f2}$. The above property may then be rewritten as

$$E\widetilde{y}v_2'(P_{f2} + \widetilde{y}) = 0 \Longrightarrow E\widetilde{y}v_1'(P_{f2} + \widetilde{y}) \leq 0. \tag{21.28}$$

This condition is formally equivalent to condition (5.7). Using Proposition 14, we conclude that an increase in risk aversion reduces the equilibrium price of equity in the economy. Notice that, following the developments made in sections 5.5 and 8.2.2, the effect of an increase of risk in the payoffs of equity is ambiguous, even under risk aversion and prudence.

21.8 Corporate finance in an Arrow-Debreu economy

It is easy to determine the rules that should determine investment strategies of investors and entrepreneurs in an Arrow-Debreu economy. For the sake of simplicity, let us go back to the static version of the model, with consumption taking place only at the end of the period. Consider an agent θ who has an investment project that would generate a state contingent payoff $q(\widetilde{x}, \alpha)$, where α is a decision variable that affects the risk profile of the investment. Should she invest in this project and, if yes, how much should invest? Without any financial markets or any other risk-sharing arrangement, she would select the α that solves

$$\max_{\alpha} \; Eu(\omega(\widetilde{x}, \theta) + q(\widetilde{x}, \alpha), \theta). \tag{21.29}$$

Thus, the degree of risk aversion of agent θ and the riskiness of the project will be the determinants of her investment in the project.

Obviously, the existence of financial markets will affect entrepreneurs select their investment strategy. The problem of corporate finance is not only to determine where and how much to invest, but also how to finance the selected projects. An important message of modern finance is that these two aspects of the problem may not be disentangled.

In the previous chapter, we have seen that it is socially efficient to aggregate all risks in the economy in order to optimize diversification. We have also seen that it will be in the private interest of all agents to do so if financial markets are competitive, frictionless and complete. This can be easily seen by observing that the true problem of our investor θ is to select at the same time her selling strategy on contingent claim markets and the characteristics of her investment. This problem is written as

$$\max_{c(.,\theta)\alpha} \quad Eu(c(\widetilde{x}, \theta), \theta) \tag{21.30}$$

$$s.t. \quad E\pi(\widetilde{x})c(\widetilde{x}, \theta) = E\pi(\widetilde{x}) \left[\omega(\widetilde{x}, \theta) + q(\widetilde{x}, \alpha)\right] \tag{21.31}$$

It is clear that this problem can be decomposed into two steps:

1. Select the α that maximizes the market value of the project $E\pi(\widetilde{x})q(\widetilde{x}, \alpha)$,
2. Rebalance the investor's risk profile by an active strategy on financial markets.

This is because of the possibility to compensate the risk taken by investing in the project by trading on financial markets that it is optimal to select the investment strategy which maximizes its market value. The effect of the riskiness of the project on the income of the investor does not matter. We see that the only thing which is important for determining whether a project should be implemented is the statistical relationship that exists between the payoff of this project and the market risk. Observe that the notion of a firm does not exist in an Arrow-Debreu economy, since the value of a bundle of different investment projects that a firm could have is just the sum of the value of these projects. Mergers and acquisitions are just irrelevant for improving the market value of a firm.

There are several way for an entrepreneur to rebalance the risk profile of her income. She can do it by purchasing insurance contracts for specific contingencies, which are nothing else that contingent claim contracts. Or she can purchase a portfolio of risky assets whose returns are negatively correlated with the return of her investment project. She can also sell shares of her project on financial markets, creating a stockholders company. Finally, she can issue bonds or borrow from banks. The point is that these strategies have no effect on the market value of the project. Again, the value of the firm is just the sum of the market value of its component. Because the market did already diversify risks that can be washed out in the economy, purchasing a contingent claim contract on the market to reduce the riskiness of the firm will not affect its value. The increase in the market value of the firm due to this risk reduction is just compensated by the market price that the firm has to pay for this contingent claim contract. The shareholders of the firm are just indifferent about the financial engineering used by the firm to finance its investment. Any financial arrangement that could be done by the firm can be undone by shareholders purchasing the symmetric arrangement on financial markets. This is the essence of the famous Modigliani-Miller Theorem.

The only thing that matters for shareholders is to make sure that managers select the investment project — the α — that maximizes the market value of the firm. That may be a problem if managers have private informations about the risk profiles of the investment projects that are available to the firm. Most of the developments in corporate finance for the last twenty years are contained in this sentence.

Chapter 22

Searching for the representative agent

The computation of the equilibrium price of financial assets is simple when the preferences of the representative agent are known. The difficulty lies in fact in characterizing these preferences. Most often, these characteristics are not the mean of the corresponding characteristics in the population. For example, the degree of absolute risk aversion of the representative agent is not the mean value of the actual absolute risk aversion in the population. The same remark holds for marginal utility, prudence, and so on. In this chapter, we reexamine this aggregation problem.

We consider exactly the same problem as in the previous chapter: each agent θ is characterized by her utility function $u(.,\theta)$ and by a endowment $\omega(z,\theta)$ that is a function of the GDP per capita z that will prevail. Observe that we simplified the notation by assimilating the state of world x to the GDP per capita z that prevails in that state. By the mutuality principle, we know that this is without loss of generality. For any distribution of $(\widetilde{z},\widetilde{\theta})$, there is an equilibrium distribution of consumption that is characterized by function $c(z,\theta)$. The preferences of the representative agent is characterized by the utility function v. It is obtained by solving the following system of equations that we derived in the previous chapter:

$$\left\{\begin{array}{lcll} u'(c(z,\theta),\theta) & = & \xi(\theta)v'(z) & \text{for all } \theta, z; \\ E\left[c(\widetilde{z},\theta)-\omega(\widetilde{z},\theta)\right]v'(\widetilde{z}) & = & 0 & \text{for all } \theta; \\ Ec(z,\widetilde{\theta}) & = & z & \text{for all } z. \end{array}\right. \tag{22.1}$$

As a reminder, the first equation is the first-order condition of the decision problem of agent θ associated to the Arrow-Debreu security z, whereas the second equation

is her budget constraint. The last one is the market-clearing condition for the Arrow-Debreu security associated to state z. The unknowns are the c, v' and ξ functions. Once this system is solved, we can easily derive the equilibrium prices of the Arrow-Debreu securities, since we have $\pi(z) = v'(z)$.

22.1 Analytical solution to the aggregation problem

In general, there is no simple way to characterize the attitude towards risk of the representative agent, except in four cases. The first case has already been encountered in this book: when there is no source of heterogeneity in the economy, i.e. if u and ω are independent of θ, then, trivially, the above system is solved with

$$c(z,\theta) = z, \quad v'(z) = u'(z), \quad \xi(\theta) = 1, \quad \text{for all } z, \theta. \tag{22.2}$$

In this trivial case, v coincides with u. The second case is when there exists a risk-neutral agent in the economy. System (22.1) is solved with $v'(z) = 1$ for all z in such a case.

The third case is when the utility functions are CARA: $u'(c,\theta) = \exp(-\theta c)$. Then, the above system of equation is solved by

$$c(z,\theta) = c_0(\theta) + \frac{\theta_*}{\theta} z, \quad v'(z) = \exp(-\theta_* z), \quad \xi(\theta) = \exp(-\theta c_0(\theta)), \tag{22.3}$$

where $c_0(\theta)$ solves the budget constraint of agent θ, and θ_* is defined as follows:

$$\frac{1}{\theta_*} = E\left[\frac{1}{\tilde{\theta}}\right]. \tag{22.4}$$

When all agents have CARA preferences, the representative agent will also be CARA. Notice that we could have proven this result directly by using property (21.7) of the representative agent's absolute risk tolerance. Moreover, her constant degree of absolute risk aversion is the harmonic mean of the different coefficients of absolute risk aversion in the population. Because the harmonic mean is smaller than the harmonic mean, we conclude, with CARA preferences, the representative agent has a degree of absolute risk aversion that is smaller than the mean absolute risk aversion in the population. The intuition is simple, because those who are less risk-averse will accept to purchase a larger share of the aggregate risk. It

implies that their degree of risk aversion be overrepresented in the aggregation. That pushes the risk aversion of the representative agent down. It is noteworthy that this rule cannot directly be extended for preferences that are not CARA. This is because, in equation (21.7), the degrees of absolute risk tolerance are averaged when measured at different levels of consumption.

In the last case, the only source of heterogeneity that is allowed comes from wealth inequality. All agents have the same utility function that is assumed to be HARA:

$$u'(c,\theta) = u'(c) = (\eta + \frac{c}{\gamma})^{-\gamma}, \tag{22.5}$$

for some constants η and $\gamma \geq 0$. The first-order condition yields

$$c(z,\theta) = -\eta\gamma + \gamma(v'(z))^{-\frac{1}{\gamma}}(\xi(\theta))^{-\frac{1}{\gamma}}.$$

Taking the expectation of both sides of this equality over $\widetilde{\theta}$ and using the market-clearing condition yields

$$z = -\eta\gamma + \gamma(v'(z))^{-\frac{1}{\gamma}}E(\xi(\widetilde{\theta}))^{-\frac{1}{\gamma}}$$

or,

$$\begin{aligned} v'(z) &= \left[E(\xi(\widetilde{\theta}))^{-\frac{1}{\gamma}}\right]^{\gamma}(\eta + \tfrac{z}{\gamma})^{-\gamma} \\ &= Ku'(z) \end{aligned} \tag{22.6}$$

Thus, when all agents have the same HARA utility function, the representative agent has the same attitude towards risk — and the same elasticity to intertemporal substitution — as the original agent in the economy, even if the distribution of wealth is unequal. In particular, wealth inequality will have no effect on asset pricing in this economy.

This last result also holds in an economy with just one risk free asset and one risky asset, in the absence of any independent background risk. Indeed, we know that the the demand for the risky asset is linear in the investor's initial wealth under HARA. Thus, a mean-preserving spread on the initial wealth levels, i.e. the introduction of wealth inequality, has no effect on the demand for the risky asset per capita. It implies that equilibrium prices are not affected.

22.2 Wealth inequality, risk aversion and the equity premium

From now on, we assume that all agents have the same utility function $u(.)$. Our objective will be to determine the effect of wealth inequality on the degree of risk aversion of the representative agent. We know that the distribution of wealth has no effect if u is HARA. More generally, we have that

$$T_v(z) = ET(c(z, \widetilde{\theta})). \tag{22.7}$$

Together with the market-clearing condition $Ec(z, \widetilde{\theta}) = z$, it implies that $T_v(z)$ is larger than $T(z)$ if T is convex, whereas $T_v(z)$ is smaller than $T(z)$ if T is concave. We conclude that wealth inequality increases (resp. reduces) the degree of risk aversion of the representative agent if the actual agents' absolute risk tolerance is concave (resp. convex).

An application of this result is found in the analysis of the effect of wealth inequality on the equity premium. The equity premium with wealth inequality equals

$$\phi = \frac{E\widetilde{z}Ev'(\widetilde{z})}{E\widetilde{z}v'(\widetilde{z})}, \tag{22.8}$$

where v is the solution of system (22.1). This should be compared to the equilibrium price in the absence of wealth inequality which equals $E\widetilde{z}u'(\widetilde{z})/Eu'(\widetilde{z})$. As stated in chapter 6, the equity premium is an increasing function of the degree of risk aversion. Thus in order to sign the effect of wealth inequality on the equity premium, we just need to know whether v is more or less concave than u. We thus end up with the following Proposition:

Proposition 70 *Assume that all agents have the same utility function u. If absolute risk tolerance of u is concave (resp. convex), then the equilibrium price of equity in an unequal economy is smaller (resp. larger) than the equilibrium price in the egalitarian one, with the same aggregate uncertainty.*

This proposition generalizes what we already know when absolute risk tolerance is linear, which corresponds to HARA utility functions for which wealth inequality has no effect on asset pricing.

To obtain this result, we don't need to make any assumption on whether risk aversion must be increasing or decreasing. But under increasing absolute risk tolerance, the intuition is as follows: introducing wealth inequality in the economy alters the aggregate demand for equity in two ways. First, poorer agents are more risk averse. This reduces their demand for equity. Second, wealthier people have a larger demand for it, since they are less risk averse. Under linear risk tolerance, we know that the demand for equity is linear in wealth. It implies that the two effects exactly compensate each other, so that the aggregate demand at the original equilibrium price is unaffected. Some intuition can be obtained by observing that,at least for a small aggregate risk, the demand for equity is linear with respect to absolute risk tolerance (see equation (5.5)). Combining this with the fact the absolute risk tolerance of the representative agent is the mean of the actual agents' absolute risk tolerance implies the result. Although this technique cannot be extended to larger risks, Proposition 70 indicates that the same result holds in general.

A consequence of this analysis is that the equity premium is increased due to wealth inequality if T is concave, whereas it is decreased if T is convex. Following Weil (1992), suppose that an outside observer is willing to confront the equity premium obtained from the model above with the observed equity premium data. Suppose that the observer believes that an egalitarian allocation is implemented in the economy. This will lead to an underprediction (resp. overprediction) of the equity premium if T is concave (resp. convex). In order to help solving the equity premium puzzle, we thus need T to be concave.

22.3 Wealth inequality and the risk-free rate

We now turn to the analysis of the impact of wealth inequality on the risk-free rate. The equilibrium rate of return of a zero-coupon bond, i.e. the risk free interest rate, equals

$$r = \frac{v'(z_0)}{\beta E v'(\tilde{z})} - 1, \tag{22.9}$$

where z_0 is the current GDP per capita and $\tilde{z}$ is the GDP per capita at the maturity of the zero-coupon bond. The level of the risk-free rate depends upon three different characteristics of the representative agent's preference. First, the risk-free rate is affected by the pure preference for the present characterized by $1/\beta$. Assets

yielding payoffs in the distant future are less valued to compensate investors for their patience. Second, the risk-free rate depends upon the growth of the representative agent's consumption. If v is concave, smoothing consumption over time is valuable. In the absence of uncertainty about the positive growth of consumption, increasing the risk-free rate above $1/\beta$ is necessary in order to compensate investors for not smoothing their consumption over time. By how much the risk-free rate must be increased depends upon the elasticity of intertemporal substitution of the representative agent. Third, if the representative agent is prudent and if future consumption is uncertain, a reduction in the risk-free rate is required in order to convince agents not to accumulate wealth for a precautionary saving motive. The size of this effect depends upon the degree of absolute prudence of the representative agent.

In order to address the question of the effect of wealth inequality on the risk-free rate, we must determine its impact on each of these three motives to save. We do this by starting with a simple case that will be extended in three subsequent steps. This benchmark case is when there is no uncertainty on future GDP per head (to exclude the precautionary saving motive) and no growth in aggregate consumption (to exclude the income smoothing effect). In that case, credit markets will be at work at date 0 to smooth all consumption plans at equilibrium in the unequal economy. This equilibrium is sustained by a gross risk-free rate equaling $1/\beta$. Wealth inequality has no effect on it. We now introduce growth.

22.3.1 The consumption smoothing effect

In this paragraph, we suppose that $\widetilde{z}$ takes value z_1 with probability 1.Moreover, we assume that $z_1 > z_0$. There is no uncertainty about the positive growth of the GDP per head. In this economy, insurance markets will be at work to eliminate all individual risks. Because growth is certain and positive, the concavity of v in equation (22.9) implies that R is larger than $1/\beta$. This is the pure income smoothing effect. Because the level of wealth of agents affect their willingness to smooth consumption over time, we may expect that wealth inequality affects the equilibrium risk-free rate in this economy. Proposition 57 tells us that we just have to determine the impact of the wealth inequality on the degree of concavity of v. We already know the condition under which this is the case. The following result is an immediate Corollary of Proposition 70.

Corollary 6 *Consider an economy with homogenous preferences. Suppose that*

there is no uncertainty about the positive growth of GNP per head. If the absolute risk tolerance is concave (resp. convex), then the equilibrium interest rate in an unequal economy is larger (resp. smaller) than the equilibrium interest rate in the egalitarian one with the same growth. The opposite results hold if growth is negative.

Another way of understanding this result is to observe that the marginal propensity to consume out of wealth is decreasing in wealth under certainty if and only if absolute risk tolerance is concave (Proposition 50). In other words, the concavity of T is equivalent to the concavity of saving with respect to wealth. The Jensen's inequality yields that wealth inequality reduces the average/aggregate saving under this condition. This must be compensated by an increase in the interest rate in order to sustain the equilibrium.

22.3.2 The precautionary effect

We now turn to the general case with an aggregate uncertainty at date 1. It is useful to decompose the global effect of wealth inequality on the risk-free rate into a consumption smoothing effect and a precautionary effect. We have seen in Corollary 6 that there is no consumption smoothing effect under certainty if R equals $1/\beta$. If we maintain the assumption $R = 1/\beta$ in the uncertain but egalitarian world, we obtain a pure precautionary effect of wealth inequality. We know from Proposition 58 that wealth inequality reduces the equilibrium risk free rate if v' is more convex than u'.

Thus, to determine whether wealth inequality reduces the risk-free rate, we have to examine how it affects the degree of prudence of the representative agent. This is in parallel with what we have done in the previous section where we had to look at the degree of risk aversion of the representative agent. But contrary to this first result, this new problem is much more complex. Indeed, from condition $v'(z) = u'(c(z,\theta))/\xi(\theta)$ and after some tedious manipulations, one can verify that

$$v''(z) = \xi(\theta)^{-1}u''(c(z,\theta))\frac{\partial c}{\partial z}(z,\theta) = -\frac{u'(c(z,\theta))}{\xi(\theta)ET(c(z,\widetilde{\theta}))},$$

and

$$v'''(z) = \frac{u'(c(z,\theta))E[T(c(z,\widetilde{\theta}))]^2 P(c(z,\widetilde{\theta}))}{\xi(\theta)[ET(c(z,\widetilde{\theta}))]^3}.$$

This yields

$$-\frac{v'''(z)}{v''(z)} = \frac{E[T(c(z,\widetilde{\theta}))]^2 P(c(z,\widetilde{\theta}))}{[ET(c(z,\widetilde{\theta}))]^2}. \tag{22.10}$$

On the left-hand side of this equation, we have the degree of absolute prudence of the representative agent that we want to be larger than $P(z)$ to get the result. We learn from this equality that the degree of prudence of the representative agent is a rather complex function of the degree of prudence of the original agent. In particular, contrary to what we got for the degree of risk tolerance, the degree of prudence (or its inverse) of the representative agent is *not* a weighted sum of the degree of prudence (or its inverse) of the original agent measured at different levels of consumption.

To solve this problem, let us define the precautionary equivalent consumption C of consumption plan $c(\widetilde{z})$ at date 1 as

$$C(c(\widetilde{z})) = u'^{-1}(Eu'(c(\widetilde{z})). \tag{22.11}$$

We would be done if this functional would be concave.[1] This is best understood by assuming that there is a finite number of states, with $\widetilde{z}$ being distributed as $(z_1, p_1; ...; z_m, p_m)$. Assuming that $\beta\rho = 1$ in the egalitarian economy means that $z_0 = C(z_1, ..., z_m)$. At the given interest rate, wealth inequality has the effect to transform the consumption plan for agent θ from $(z_1, ..., z_m)$ to $(c(z_1, \theta), ..., c(z_m, \theta))$, with $Ec(z_i, \widetilde{\theta}) = z_i$, for all i. The willingness to consume at date 0 of agent θ is measured by $C(c(z_1, \theta), ..., c(z_m, \theta))$. If C is a concave function of the vector of contingent consumptions, then the mean-preserving spread of the distribution of these vectors in the population will reduce the mean value of C. This reduction in the precautionary equivalent future consumption will induce an increase in the saving rate. The market will react to this shock by reducing the equilibrium risk free rate. Lemma **??** provides a sufficient condition

[1] For a more formal proof of this statement, see Gollier (1998).

for C to be concave. Indeed, using this lemma with $g(t) = -u'(t)$ yields that the precautionary equivalent consumption is concave in the vector describing the consumption plan if the inverse of absolute prudence is concave. Proposition 71 is a consequence of this Lemma.

Proposition 71 *Suppose that agents are prudent ($u''' \geq 0$) and that $R = 1/\beta$ in the egalitarian economy. Then, wealth inequality reduces the equilibrium risk-free rate, if the inverse of absolute prudence ($P^{-1}(z) = -u''(z)/u'''(z)$) is concave.*

Proposition 71 states that, if the inverse of absolute prudence is positive and concave, then wealth inequality has a negative precautionary effect on the risk-free rate. Thus, the concavity of the inverse of absolute prudence may explain why the representative agent à la Lucas overpredicts the rate of return on safe bonds if wealth inequality is not taken into account.

We are now in a situation to combine the consumption smoothing effect with the precautionary effect to obtain a general picture of the effect of wealth inequality on the equilibrium risk-free rate. Our findings are summarized in the following Proposition.

Proposition 72 *Suppose that the risk free rate is less (resp. larger) than $1/\beta$ in the egalitarian economy. Then, wealth inequality reduces the risk-free rate if the two following conditions are satisfied:*

1. *the inverse of absolute prudence is concave;*

2. *the absolute tolerance to risk is concave (resp. convex).*

Proof: See the Appendix.

Observe that the inverse of absolute risk aversion and the inverse of absolute prudence are both linear in wealth when the utility function is HARA. We know that the distribution of wealth does not affect asset prices in that case. Again, HARA is a limit case.

22.4 Conclusion

The attractiveness of many asset pricing models available in the finance literature comes from the assumption that there is a representative agent in the economy. There are two potential deficiencies to this approach however. First, when the economy is heterogenous, the existence of a representative agent who maximizes the expected value of his utility should not be taken as granted. But as seen in the previous chapter, this assumption is automatically satisfied if markets are complete.

The second potential deficiency comes from the difficulty to characterize the preferences of the representative agent when it exists. A helpful result to solve this problem is that the absolute risk tolerance of the representative agent equals the expected absolute risk tolerance in the population, in each state of the world. But, for example, no similar formula can be obtained for the measure of prudence.

The most part of the chapter has been devoted to the search of the representative agent, and of the corresponding equilibrium prices, when the only source of heterogeneity in the population is wealth inequality. If we consider an egalitarian economy whose risk-free rate is small, we concluded that under the concavity of both the inverse of the absolute risk aversion and the inverse of absolute prudence, the equity premium will be underpredicted and the risk-free rate will be overpredicted by a calibrator who does not take into account of wealth inequality.

Part VII

Risk and information

Chapter 23

The value of information

An important aspect of dynamic risk management is the existence of a flow of information linked to the structure of future risks. Young workers can observe signals about their long-term productivity in the early stages of their career. Entrepreneurs obtain information linked to the profitability of their investment project. Policy-makers may want to wait for better scientific knowledge about the risk of global warming before deciding how much effort should be spent to reduce emissions of greenhouse gases. Technological innovations provide information about the prospect of future growth of the economy. The common feature of these examples is that some signals are expected to be observed before taking a final decision on the exposure to the risk.

This part of the book is devoted to the analysis of the interaction between risk and information. We begin with the examination of how people should value future risks in the presence of a flow of information related to the distribution of the risk. The notion of the value of information is introduced. We also examine here which information structure should be preferred by agents. We finish with the examination of the link between the value of information and risk aversion.

23.1 The general model of risk and information

23.1.1 Structure of information

We consider a risky situation that is characterized by X possible states of the world indexed by $x = 1, ..., X$. Vector $\Phi = (\phi_1, ..., \phi_X)$ measures the probability of occurrence of these states. Prior to the realization of the state, the decision

maker is in a position to observe a signal. Let M be the number of potential signals indexed by $m = 1, ..., M$. Vector $Q = (q_1, ..., q_M)$ denotes the vector of unconditional probabilities of the different signals. Signals are potentially statistically related to the states. Thus, observing the signal allows the decision maker to revise the probability distribution of the states, using Bayes rule. This statistical relationship is characterized by the matrix

$$P = [p_{mx}] \begin{array}{l} m = 1, ..., M \\ x = 1, ..., X \end{array} \tag{23.1}$$

of conditional/posterior probabilities. Namely, p_{mx} is the probability that state x be realized if signal m is observed. Matrix P has M rows and X columns. Vector $p_{m.} = (p_{m1}, ..., p_{mX})$ denote the posterior probabilities conditional to signal m. It corresponds to row m of matrix P. Basic probability calculus yields that the prior probability of state x, ϕ_x, equals $\sum_m q_m p_{mx}$. Using the matrix notation, this is equivalent to

$$\Phi = \sum_{m=1}^{M} q_m p_{m.} = QP. \tag{23.2}$$

Because we will not use the matrix of joint probabilities, its notation is not introduced here. In the remaining, we use the term ”prior” and ”posterior” to refer to situations respectively before and after the observation of the signal.

23.1.2 The decision problem

Posterior to the observation of the signal, but prior to the observation of the state of the world, the agent must take a decision in the face of the remaining uncertainty. In its most general formulation, the agent's final utility $v(\alpha, x)$ depends upon his decision α and the state of the world x. The agent is limited in his choice of α, which must belong to a non-empty set $B \subset R^n$, where n is the dimensionality of the decision variable α. Except when explicitly stipulated, we will hereafter assume that $n = 1$.

Suppose that the agent's preferences under uncertainty satisfy the axioms of von Neumann and Morgenstern. His decision problem is then written as:

$$U(P, Q) = \sum_{m=1}^{M} q_m \quad \left[\max_{\alpha \in B} \quad \sum_{x=1}^{X} p_{mx} v(\alpha, x) \right] \tag{23.3}$$

The bracketed term is the maximal expected utility that the decision maker can obtain when he observes signal m, which generates a vector of posterior probabilities $p_{m.} = (p_{m1}, ..., p_{mX})$. Prior to the observation of the signal, the maximum expected utility $U(P, Q)$ is the sum of these terms, weighted by the probability of the different possible signals that he can receive. Let us define function g from the simplex $\mathcal{L}$ in R^X to R as follows:

$$g(p_{m.}) = \max_{\alpha \in B} \quad \sum_{x=1}^{X} p_{mx} v(\alpha, x) \tag{23.4}$$

It represents the maximal expected utility given a vector of posterior probabilities $p_{m.}$ of the states. Associated to it is the optimal solution $\alpha(p_{m.})$ given this vector of posterior probabilities. We can then rewrite equation (23.3) as

$$U(P, Q) = \sum_{m=1}^{M} q_m g(p_{m.}) \tag{23.5}$$

It is important to examine the main properties of the g function.

23.1.3 The posterior maximum expected utility is convex in the vector of posterior probabilities

It is important to observe that function g is linear in the vector of posterior probabilities $p_{m.}$ when the optimal decision α is independent of this vector. This is in fact the main feature of expected utility. But in general, the optimal decision will be affected by a change in the probabilities. By definition (23.4), function g is the maximum of a set of functions $f(\alpha, p_{m.}) = \sum_x p_{mx} v(\alpha, x)$, $\alpha \in B$, that are linear in $p_{m.}$. Therefore, it is a *convex* function of vector $p_{m.}$. Remember that a convex function g from $\mathcal{L} \subset R^X$ to R is defined by the property that for any couple $(a, b) \in \mathcal{L}^2$ and any scalar $q \in [0, 1]$, we have

$$qg(a) + (1 - q)g(b) \geq g(qa + (1 - q)b). \tag{23.6}$$

Because this result is of primary importance for the remaining of this book, we illustrate this property by an example. Let us consider an uncertain environment with three states of the world. In this case, a specific posterior probability distribution is characterized by a point (p_{m1}, p_{m3}) in the Machina triangle, with $p_{m2} = 1 - p_{m1} - p_{m3}$. We consider the standard portfolio decision problem in

Figure 23.1: The expected utility as a function of the posterior probabilities with three possible states and two choices.

which $v(\alpha, x) = v(w_0 + \alpha y_x)$ where α is the number of euros invested in a risky project, and y_x is the conditional net payoff of the project per euro invested. Suppose first that $B = \{0, 1\}$, i.e., that the agent can invest either zero or one euro in the project. Then, we have that

$$g(p_{m.}) = \max \left[v(w_0), \sum_{x=1}^{3} p_{mx} v(w_0 + y_x) \right] \tag{23.7}$$

This is indeed the maximum of two linear functions. In Figure 23.1, we depicted these two linear functions by using a representation in the Machina triangle, when $y_1 = -0.5$, $y_2 = 0.5$, and $y_3 = 1$, $w_0 = 1$ and $v(z) = -z^{-1}$. The two plans describe the expected utility respectively for $\alpha = 0$ and $\alpha = 1$ as a function of the posterior probabilities. The upper envelope of the two plans represents the maximum expected utility. This is clearly a convex function. When p_{m1} is large, it is better not to invest in the project and $g(p_{m.}) = v(w_0)$. On the contrary, when it is small enough, it is optimal to invest in the project and $g(p_{m.}) = \sum_{x=1}^{3} p_{mx} v(w_0 + y_x)$.

In Figure 23.2, we describe function g when $B \equiv R$, i.e., when there is no

Figure 23.2: The maximal posterior expected utility in the portfolio problem with three possible states.

restriction on the amount that can be invested in the risky project. To each scalar α, it corresponds a plan above the Machina triangle that represents the expected utility if this decision α is taken. For each potential vector of posterior probabilities, the decision maker selects the α which maximizes the conditional expected utility. Because the optimal α is a smooth function of this vector, so is the optimal expected utility. In Figure 23.2, we just represented the optimal expected utility as a function of the posterior probabilities, which is the upper envelope of an infinite number of plans that have not been represented.

Thus, the g function is convex whatever the decision problem is. Reciprocally, one can prove that any convex function from $\mathcal{L} \subset R^X$ to R can be obtained by (23.4) from a specific decision problem (v, B). This is due to the well-known separating hyperplane theorem which implies that, to any convex surface, there exists a set of hyperplanes for which this surface is the upper envelope. We summarize this in the following Lemma.

Lemma 6 *Consider the definition of function $g : \mathcal{L} \subset R^X \to R$ as given by equation (23.4). The following two statements are true:*

- *Function g is convex*;

- *To any convex function g, there exists a decision problem (v, B) such that equation (23.4) holds for all $p_{m.} \in \mathcal{L}$.*

23.2 The value of information is positive

In this section, we examine how the attitude towards the uncertainty about the state of the world x is affected by the flow of information. To do this, we compare the optimal prior expected utility with the structure of information (P, Q) to the optimal expected utility without any information. In the absence of any source of information before having to take a decision, the problem of the decision maker is written as:

$$g(\Phi) = \max_{\alpha \in B} \quad \sum_{x=1}^{X} \phi_x v(\alpha, x) \tag{23.8}$$

We want to compare this with the maximal expected utility that we can attain when the decision can be taken after observing the signal coming from information structure (P, Q), which we denoted $U(P, Q)$ in equation (23.5). The information structure has a value if the optimal expected utility $U(P, Q)$ obtained with it is larger than the optimal expected utility $g(\Phi)$ obtained without it. In short, the information structure has a value if

$$\sum_{m=1}^{M} q_m g(p_{m.}) \geq g(\Phi) = g\left(\sum_{m=1}^{M} q_m p_{m.}\right) \tag{23.9}$$

since $\Phi = QP$.

Inequality (23.9) is a direct consequence of our earlier observation that g is convex in the vector of probabilities. We conclude that any information structure has a nonnegative value. This is true independent of matrixes (P, Q) and of the characteristics of the decision problem given by B and v. At worst, the information structure has no value. This is the case when g is linear. We know that this is the case only when observing the signal never affects the decision of the agent. As soon as there is at least one signal that generates a different action, g becomes convex and the inequality in (23.9) is strict: the information structure has a positive value. This means that the value of information is extracted from the fact that the observation of a signal allows the agent to better adapt his decision to the risky environment that he faces. Flexibility is valuable.

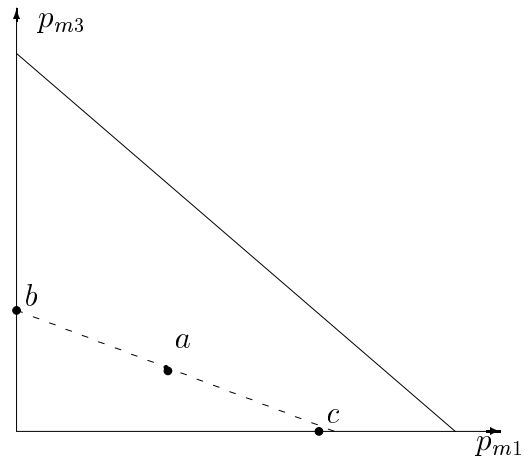

Figure 23.3: A simple information structure with two possible messages.

There is a simple intuition for why the value of information may not be negative. The action which is chosen without information is also feasible when information is available. Therefore, the informed agent can at least duplicate his rigid strategy with the information structure by choosing uniformly the same action, no matter what information is received. As explained by Savage (1954), "the person is free to ignore the observation. That obvious fact is the theory's expression of the commonplace that knowledge is not disadvantageous." You can always decide not to take your umbrella if rain is forecasted for today...

In Figure 23.3, we illustrate this in a simple example that we borrow from the previous section: there are 3 possible states of the world and $v(\alpha, x) = -(1 + \alpha y_x)^{-1}$ and $\alpha \in B \equiv \{0, 1\}$. The agent has the opportunity to purchase a lottery ticket whose possible net payoffs are $(y_1, y_2, y_3) = (-0.5, 0.5, 1)$. In the absence of any additional information, the distribution of probabilities is $\Phi = (0.35, 0.55, 0.1)$. This probability distribution is represented by point a in Figure 23.3. In spite of the positive unconditional expectation of the net payoff, the reader can easily check that it is not optimal for him to take this risk. His expected utility equals -1. It is represented by point A in Figure 23.4.

Suppose now that the agent has access to a source of information about his chance to win the lottery. There are two possible messages of equal probabilities ($q_1 = q_2 = 0.5$). The vector of posterior probabilities is represented respectively

Figure 23.4: The value of information in the Machina triangle.

by points b and c in the Machina triangle of Figure 23.3. If message b is received, the posterior distribution is $(0, 0.8, 0.2)$. The probability of a net loss is zero, the agent purchases the lottery in this case, and his expected utility equals -0.633. If message c is received, the posterior distribution is $(0.7, 0.3, 0)$. Because the expectation of the net payoff is negative, the agent does not purchase the lottery ticket, and his expected utility equals -1. Prior to receiving the message, the expected utility equals $0.5((-0.633) + (-1)) = -0.8166$. This is larger than -1, the expected utility of the uninformed agent. In Figure 23.4, the ex-ante expected utility of the agent is represented by the point at the center of the interval BC. It is strictly above point A, which means the information structure has a positive value.

23.3 Refining the information structure

23.3.1 Definition and basic characterization

In this section, we compare pairs of information structures like (P^1, Q^1) and (P^2, Q^2). We assume that they describe the same uncertainty ex ante, i.e., that

$$\Phi = Q^1 P^1 = Q^2 P^2 \tag{23.10}$$

More specifically, we try to determine whether it is possible to rank (P^1, Q^1) and (P^2, Q^2) in the following sense. We say that an information structure is finer than another if the first is always preferred to the second.

Definition 7 *(P^1, Q^1) is a better information structure than (P^2, Q^2) if and only if all agents enjoy a larger expected utility with (P^1, Q^1) than with (P^2, Q^2), no matter which decision problem (v, B) is considered. This is rewritten as follows:* $\forall B \subset R, \quad \forall v : R^2 \to R :$

$$U(P^1, Q^1) \geq U(P^2, Q^2), \tag{23.11}$$

or equivalently,

$$\sum_{m=1}^{M_1} q_m^1 \quad \left[\max_{\alpha \in B} \quad \sum_{x=1}^{X} p_{mx}^1 v(\alpha, x) \right] \geq \sum_{m=1}^{M_2} q_m^2 \quad \left[\max_{\alpha \in B} \quad \sum_{x=1}^{X} p_{mx}^2 v(\alpha, x) \right]. \tag{23.12}$$

This is a relevant question when the decision maker has the opportunity to choose between various information structures. Blackwell (1951, 1953) uses the terminology of an ”experiment” to characterize the process of information acquisition. The agent faces various possible experiments. He must select one from whose he will extract a signal before taking decision α. For each possible experiment i, he determines the probability q_m^i of occurrence of each signal $m = 1, ..., M_i$ and their corresponding vector $(p_{m1}^i, ..., p_{mX}^i)$ of posterior probabilities. Does there exist an experiment that is preferred by all decision makers? One faces this question when one has to evaluate medical diagnosis processes, experts in finance, but also to select a sample of the population, to improve weather forecasting, to select among various oil drilling strategies, etc...

We will hereafter indifferently use the terms of a ”better information structure”, a ”more informative structure”, a ”finer information structure” or an ”earlier resolution of uncertainty” to describe structure 1 with respect to structure 2 if condition (23.12) is satisfied. A direct consequence of Lemma 6 is given in the following Proposition, which is due to Bohnenblust, Shapley and Sherman (1949) and is for example in Marschak and Miyasawa (1968), Epstein (1980) and Jones and Ostroy (1984).

Proposition 73 *(P^1, Q^1) is a better information structure than (P^2, Q^2) if and only if*

$$\sum_{m=1}^{M} q_m^1 g(p_{m.}^1) \geq \sum_{m=1}^{M} q_m^2 g(p_{m.}^2) \tag{23.13}$$

for all convex function $g : \mathcal{L} \subset R^X \to R$.

23.3.2 Garbling messages and the theorem of Blackwell

Proposition 73 is not very helpful to provide an intuition. We hereafter assume that the two structures have the same set of signals $m = 1, ..., M$. This is without loss of generality, since we can always set some of the q_m^i to zero if signal m never occurs in structure i.

The most extreme forms of information structures are respectively the completely unimformative one and the completely conclusive one. (P, Q) is completely uninformative if the vector of posterior probabilities is the same for all signals. It is completely conclusive if there is only one possible signal for each state of the world. In that latter case, the vector of posterior probabilities is of the form $(0, ..., 0, 1, 0, ..., 0)$: the observation of the signal completely specifies the state. In that case, the uncertainty is completely resolved by observing the signal. If the two structures describe the same prior uncertainty, it is obvious that the completely conclusive one is better than the completely uninformative one. This is a direct consequence of our observation that the value of information is always nonnegative. More generally, any information structure is better than the corresponding completely uninformative structure. This is the meaning of property (23.9). Similarly, the completely informative experiment is better than any incomplete one with the same prior.

As an example, consider three possible messages $m = 1, 2, 3$ that are equally likely. Information structure $i = 1$ gives a posterior distribution a in Figure 23.5 for all three messages. This information structure is uninformative. Information structure $i = 2$ generates posterior distribution c if signal $m = 1$ is received, and posterior distribution b otherwise. This information structure is better than information structure 1. Observe that information structure 2 is a *mean-preserving spread* of information structure 1 *in the space of probability distributions*. Equivalently, a is the center of gravity of an object having one third of its mass in c

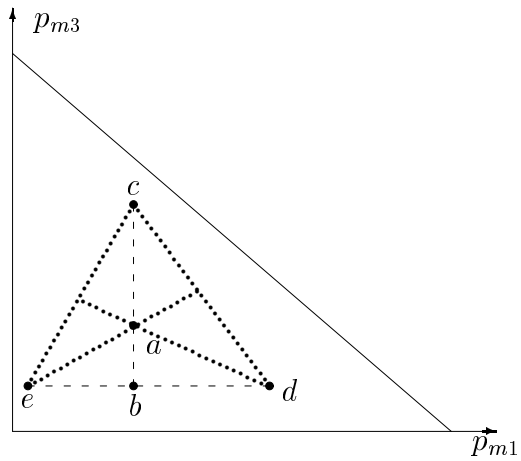

Figure 23.5: Information structures with three messages.

and two third of it in b. Because the maximal expected utility is a convex function of the various posterior probability distributions, it is clear that structure 2 is preferred to structure 1: all decision makers like mean-preserving spreads in the space of posterior distributions function. And this has nothing to do with a risk-loving behavior, which corresponds to mean-preserving spread in the space of wealth.

Now, consider information structure $i = 3$ which generates posterior distributions c, d and e respectively for signals $m = 1, 2$ and 3. Observe that b is the center of gravity of (d, e) where the mass is equally split in the two points. In other terms, (d, e) is a mean-preserving spread of b. Thus, information structure $i = 3$ is a mean-preserving spread of information structure $i = 2$. Again, because the maximal expected utility is convex in the posterior distribution, information structure $i = 3$ is preferred to information structure $i = 2$, independent of the decision problem under consideration. This can be seen as follows:

$$\begin{aligned} \frac{1}{3}g(p_d) + \frac{1}{3}g(p_e) + \frac{1}{3}g(p_c) &= \frac{2}{3}\left[\frac{1}{2}g(p_d) + \frac{1}{2}g(p_e)\right] + \frac{1}{3}g(p_c) \\ &\geq \frac{2}{3}g(p_b) + \frac{1}{3}g(p_c) \end{aligned} \tag{23.14}$$

because $p_b = 0.5(p_d + p_e)$ and g is convex in p. We conclude that any mean-

preserving spread in the distribution of the vectors of posterior probabilities improves the degree of informativeness of the structure.

An alternative way to explain this is by using the technique of garbling messages. One can duplicate information structure 2 from information structure 3 by using the following procedure: consider a garbling machine which receives the true signal from information structure 3 and send another message by using these rules:

- if message 1 is received, the machine sends message 1;
- if message 2 is received, the machine sends message 2 or 3 with equal probabilities;
- if message 3 is received, the machine sends message 2 or 3 with equal probabilities.

Suppose that the decision maker only observes the message coming out of this garbling machine. Then, if message 1 is observed, he immediately infers that the message received by the machine was message $m = 1$ from the original information structure 3: this message is not garbled by the machine. But if message 2 is received, the agent infers that the original message received by the machine was either $m = 2$ or $m = 3$ with equal probabilities. Thus the vector of probabilities on the states is either d or e with equal probabilities. In other words, the vector of posterior probabilities in this case is $b = 0.5(d + e)$. He concludes in the same way if the machine gives him message 3. We conclude that using information structure 3 coupled with the garbling machine generates an information structure equivalent to information structure 2. It gives him posterior c with probability $1/3$ and posterior b with probability $2/3$. More generally, any sequence of mean-preserving spread in the space of probability distribution can be obtained by this kind of garbling machine.

A garbling machine is characterized by a function $\mu : R^2 \to R$ and a random variable $\widetilde{\xi}$ independent of $\widetilde{m}^1$ such that $\widetilde{m}^2$ is distributed as $\mu(\widetilde{m}^1, \widetilde{\xi})$ for all x, where $\widetilde{m}^1$ and $\widetilde{m}^2$ are the messages coming respectively in and out of the machine:

$$\widetilde{m}^2 \text{ is distributed as } \mu(\widetilde{m}^1, \widetilde{\xi}) \tag{23.15}$$

This technique is easy to represent by using the matrix notation. Consider a machine which sends message j if it receives message i with probability h_{ij}:

$$h_{ij} = \Pr\left[\text{machine sent } j \mid \text{machine received } i\right]. \tag{23.16}$$

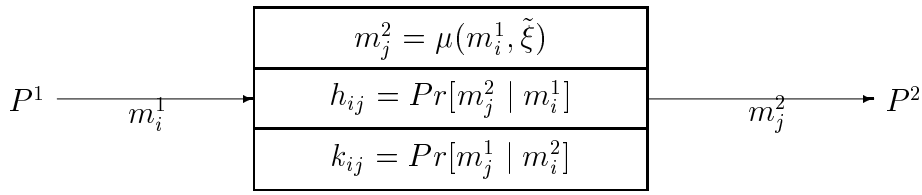

Figure 23.6: The garbling machine.

The knowledge of matrix $H = [h_{ij}]$ allows us to completely describe the distribution of the message coming out of the machine if we would know the message that has been received by it, independent of the state. Because the random noise introduced by the machine is not correlated to the state, observing the message coming out of it does not bring more information than observing the message coming in. In statistics, we say that the $\widetilde{m}^1$ is a sufficient statistic for $\widetilde{m}^2$. This garbling machine is represented in Figure 23.6.

¿From matrix H, we can determine k_{ij}, the probability that message j has been received by the garbling machine if it sent message i:

$$k_{ij} = \Pr\left[\text{machine received } j \mid \text{machine sent } i\right]. \tag{23.17}$$

Using Bayes rule, we have that

$$k_{ij} = \frac{q_j h_{ji}}{\sum_{k=1}^{M} q_k h_{ki}}. \tag{23.18}$$

It yields

$$p_{i.}^2 = \sum_{j=1}^{M} k_{ij} p_{j.}^1. \tag{23.19}$$

Of course, we have $k_{ij} \geq 0$ and $\sum_j k_{ij} = 1$ for all i. Thus, all posterior probabilities p_i^2 in structure 2 are a convex combination of the posterior probabilities p_j^1 of structure 1. In other words, all points representing posterior distributions of experiment 2 are within the smallest convex envelope of points representing posterior distributions of experiment 1. Two examples are given in Figure 23.7. This is another way to say that structure 2 is a mean-preserving contraction of structure

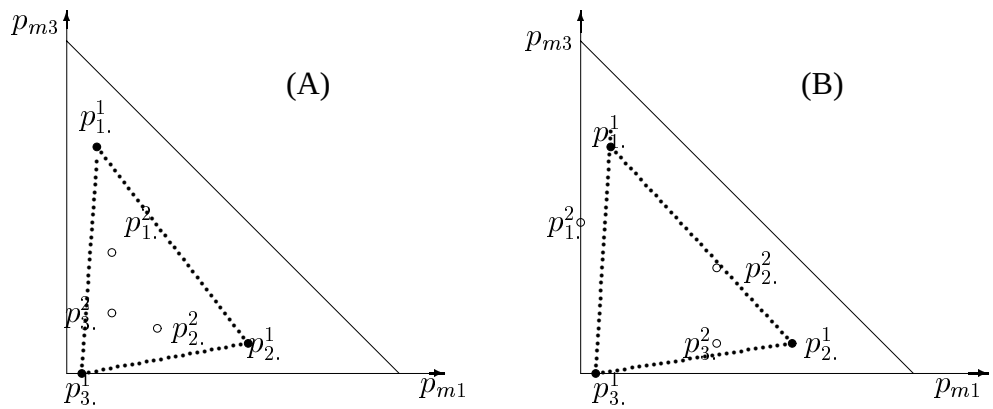

Figure 23.7: Experiment 1 is better than experiment 2 in case (A), but not in case (B).

1. In matrix notation, this means that $P^2 = KP^1$ where $K = [k_{ij}]$. In order to preserve the distribution of prior probabilities $\Phi = Q^1P^1 = Q^2P^2 = Q^2KP^1$, it must be true that $Q^1 = Q^2K$.

Property (23.19) says nothing else than structure 2 is obtained by garbling messages from structure 1. Or that structure 2 is obtained from structure 1 by a sequence of mean-preserving contractions in the space of probability distributions. We show in the following Proposition, which is due to Blackwell (1951,1953), that this property characterizes the set of finer information structures.[1]

Proposition 74 *Consider two information structures, (P^1, Q^1) and (P^2, Q^2), which have the same number M of possible messages, and the same vector of unconditional probabilities $\Phi = Q^1P^1 = Q^2P^2$. Information structure 1 is finer than information structure 2 if and only if there exists a MxM matrix $K = [k_{ij}]$ such that all $k_{ij} \geq 0$ and $\sum_{j=1}^{M} k_{ij} = 1$ for all i, and:*

$$P^2 = KP^1 \quad \textit{and} \quad Q^1 = Q^2K. \tag{23.20}$$

[1]Ponssard (1975) and Cremer (1982) provided two alternative proofs of this theorem.

Proof: We start with the proof of sufficiency. We have that

$$\begin{array}{rcl} U(P^2, Q^2) & = & \sum_{m=1}^{M} q_m^2 g(p_{m.}^2) \\ & = & \sum_{m=1}^{M} q_m^2 g(\sum_{j=1}^{M} k_{mj} p_{j.}^1) \\ & \leq & \sum_{j=1}^{M} \left[\sum_{m=1}^{M} q_m^2 k_{mj}\right] g(p_{j.}^1) \\ & = & \sum_{j=1}^{M} q_j^1 g(p_{j.}^1) = U(P^1, Q^1) \end{array} \tag{23.21}$$

We used the convexity of g to obtain the inequality on line 3 of this sequence of operations. This concludes the proof of sufficiency.

The proof of necessity follows from Lemma 6: let C be the subset of the simplex in R^X which corresponds to the smallest convex envelope of $\{p_1^1, ..., p_m^1, ..., p_M^1\}$. Suppose by contradiction that information structure 1 is not obtained by garbling messages from information structure 2, i.e., that condition (23.19) is not satisfied with $k_{ij} \geq 0$ for all i. Or, that some posterior probability distributions of structure 2 are outside C. By Lemma 6, we know that there exists a decision problem (v, B) such that the corresponding g function is linear in C and is strictly convex outside C. This implies that

$$g(p_{m.}^2) = \sum_{j=1}^{M} k_{mj} g(p_{j.}^1) \qquad \text{for all } m \text{ such that } p_{m.}^2 \in C \tag{23.22}$$

and

$$g(p_{m.}^2) > \sum_{j=1}^{M} k_{mj} g(p_{j.}^1) \qquad \text{for all } m \text{ such that } p_{m.}^2 \notin C \tag{23.23}$$

because g is convex outside C. We conclude that

$$\sum_{m=1}^{M} q_m^2 g(p_{m.}^2) > \sum_{m=1}^{M} q_m^2 \sum_{j=1}^{M} k_{mj} g(p_{j.}^1) = \sum_{j=1}^{M} q_j^1 g(p_{j.}^1). \tag{23.24}$$

It implies that experiment 1 may not be better than experiment 2. ■

If we allow for a continuum of messages and states, condition (23.19) is rewritten as

$$f^2(x \mid m^2) = \int k(m^1 \mid m^2) f^1(x \mid m^1) dm^1, \tag{23.25}$$

with $\int k(m^1 \mid m^2)dm^1 = 1$ for all m^2, where f^i is the conditional density of the states in structure i. Another way to express the notion of garbling is to refer to definition (23.16). It yields

$$g^2(m^2 \mid x) = \int h(m^2 \mid m^1) f^1(m^1 \mid x) dm^1, \tag{23.26}$$

for all x, where g^i is the conditional density of messages in structure i, and $\int h(m^2 \mid m^1)dm^2 = 1$ for all m^1,. Conditions (23.25) and (23.26) are two alternative ways to define the notion of sufficiency in statistics.

23.3.3 Location experiments

An important class of information structures is such that, for all x :

$$\widetilde{m}^i \mid x \text{ is distributed as } x + \widetilde{\varepsilon}^i \tag{23.27}$$

where $\widetilde{\varepsilon}^i$ is independent of x. An information structure belonging to such a set is called a "location experiment". If $E\left[\widetilde{\varepsilon}^i\right] = 0$, signal $\widetilde{m}^i$ is an unbiased estimator of the true state of the world. Random variable $\widetilde{\varepsilon}^i$ linked to experiment i can be interpreted as the error of observation of the unknown variable x. The intuition suggests that an experiment with a smaller dispersion of the error around its mean provides a better information. To illustrate, let us compare two location experiments. Experiment $i = 1$ has an error $\widetilde{\varepsilon}^1$ which is uniformly distributed over interval $[-1/2, 1/2]$, whereas experiment $i = 2$ has an error $\widetilde{\varepsilon}^2$ which is more dispersed since it is uniformly distributed over interval $[-1, 1]$:

$$\widetilde{\varepsilon}^1 \sim U\left[-1/2, 1/2\right]; \qquad \widetilde{\varepsilon}^2 \sim U\left[-1, 1\right] \tag{23.28}$$

It is easily shown that all agents prefer experiment 1 over experiment 2, i.e., that experiment 1 is better than experiment 2 in the sense of Blackwell. Indeed, remember that one way to prove this is to show that there exists a function $\mu(.,.)$ and a random variable $\widetilde{\xi}$ such that signal $\widetilde{m}^2$ is distributed as $\mu(\widetilde{m}^1, \widetilde{\xi})$ for all x. This is immediate in our case by taking $\mu(m, \xi) = m + \xi$ and $\xi \sim (-1/2, 0.5; 1/2, 0.5)$.

Another example is when $\widetilde{\varepsilon}^1$ and $\widetilde{\varepsilon}^2$ are both normally distributed:

$$\widetilde{\varepsilon}^1 \sim N(0, V^1); \qquad \widetilde{\varepsilon}^2 \sim N(0, V^2). \tag{23.29}$$

Suppose that $\widetilde{\varepsilon}^1$ is less dispersed than $\widetilde{\varepsilon}^2$, i.e., that V^1 is smaller than V^2. Is structure 1 more informative than structure 2? The answer is positive, since $\widetilde{\varepsilon}^2$

is distributed as $\widetilde{\varepsilon}^1$ plus a normally distributed variable $\widetilde{\xi}$ with mean zero and variance $V^2 - V^1$.

These two examples suggests the possibility that the comparison of two location experiments can be reduced to the comparison of the dispersion of the errors in the observation of x. We hereafter show that this suggestion is misleading. Following Boll (1955), cited by Lehmann (1988), we know that condition (23.15) can be true for two location experiments (23.27) only when function μ takes the form $\mu(m, \xi) = m + \xi$. This greatly simplifies the analysis, but it also shows how restrictive is Blackwell's notion of better information.

For example, suppose again that $\widetilde{\varepsilon}^2$ is normally distributed. It is well-known that if a random variable $\widetilde{m}^2$ is normal and if it is the sum of two independent random variables $\widetilde{m}^1$ and $\widetilde{\xi}$, then $\widetilde{m}^1$ and $\widetilde{\xi}$ *must* also be normal. In consequence, $\widetilde{m}^1$ cannot be more informative than a normally distributed $\widetilde{m}^2$ unless $\widetilde{m}^1$ (and $\widetilde{\varepsilon}^1$) is also normal. If $\widetilde{m}^1$ is not normal, it can never be more informative than $\widetilde{m}^2$, however concentrated or dispersed the distributions are.

Another example is in Lehmann (1988) and Persico (1998). Suppose that $\widetilde{\varepsilon}^1$ and $\widetilde{\varepsilon}^2$ are both uniformly distributed, with

$$\widetilde{\varepsilon}^1 \sim U\left[-\rho, \rho\right]; \qquad \widetilde{\varepsilon}^2 \sim U\left[-1, 1\right], \tag{23.30}$$

with $0 < \rho < 1$, which implies that the error is less dispersed in experiment 1 than in experiment 2. We showed above that experiment 1 is better than experiment 2 when $\rho = 1/2$. The same proof can be used to show that this result holds for all ρ such that $\rho = 1/k$ for some positive integer k. But Lehmann (1988) showed that this condition is not only sufficient, but is also necessary. Thus, taking ρ which is not the inverse of an integer, does not lead to a finer information structure, whatever small ρ!

These two examples confirm two claims: first, the Blackwell's order is partial. Second, it is very restrictive. This is due to the fact that Blackwell requires that all agents must prefer structure 1 to structure 2 *independent of the decision problem under scrutiny*. Some recent progresses have been made to enlarge the degree of comparability by restricting the set of decision problems. In particular, Lehmann (1988) and Persico (1998) focus on problems satisfying the single crossing property, i.e., problems where the derivative of $v(\alpha, x)$ with respect to α crosses zero at most once from below as a function of x. Athey and Levin (1998) examines decision problems where the optimal decision α is an increasing function of the signal.

23.4 The value of information and risk aversion

23.4.1 A definition of the value of information

In section 23.2, we have seen that the value of information is nonnegative. The ability to adapt the decision to the information makes the information structure always more valuable than no information at all. Notice that this result is obtained without making any assumption about the utility function of the decision maker, who can be risk-averse or risk-lover. We now examine the link between the value of information and risk aversion. Because we will evaluate information structures in monetary terms, we need to assume that the agent's utility v is a function of his consumption, which is itself a function $c(\alpha, x)$ of his decision α and of the state x:

$$v(\alpha, x) = u(c(\alpha, x)). \tag{23.31}$$

Under this additional assumption, one can define the value of information structure (P, Q) as the sure amount V of consumption that is necessary to compensate the agent for the loss of the access to the information. V is thus formally defined as

$$\max_{\alpha \in B} \quad \sum_{x=1}^{X} \phi_x u(c(\alpha, x) + V) = \sum_{m=1}^{M} q_m \quad \left[\max_{\alpha \in B} \quad \sum_{x=1}^{X} p_{mx} u(c(\alpha, x)) \right] \tag{23.32}$$

where ϕ_x is the unconditional probability of state x. The right-hand side of this equality is the expected utility of the agent who has access to the information. The left-hand side is the expected utility of the uninformed agent who got compensation V.

We can also interpret the value of information as the difference between two certainty equivalents. Let $C(\alpha, p_., z)$ denote the certainty equivalent associated to lottery $(c(\alpha, 1), p_1; ...; c(\alpha, X), p_X)$ when the agent has an exogenous initial wealth z. This means that $u(z + C(\alpha, p_.)) = \sum p_x u(z + c(\alpha, x))$. Let also $C^+(p_.)$ be the supremum of the $C(\alpha, p_., 0)$ over all α in B. $C^+(p_.)$ is the maximal posterior certainty equivalent that the agent can obtain if the posterior probability distribution is $p_.$. Before observing the message, the agent who will have access to the signal faces lottery $\widetilde{C}^+ = (C^+(p_{1.}), q_1; ...; C^+(p_{M.}), q_M)$ whose certainty equivalent is denoted IC. This certainty equivalent of the informed agent is such that $u(IC) = \sum q_m u(C^+(p_{m.}))$. It is the certainty equivalent of the maximum

posterior certainty equivalents. IC represents the value of the game against nature for the informed agent. Let UC be the maximum certainty equivalent of the game for the uniformed agent with wealth V. This means that UC is the supremum of the $C(\alpha, \Phi, V)$ over all α in B. This implies that condition (23.32) can be rewritten as

$$u(UC + V) = u(IC), \tag{23.33}$$

or, equivalently,

$$V = IC - UC. \tag{23.34}$$

The value of the information is the difference between the certainty equivalents of the game respectively for the informed agent and the uninformed agent.

The value of information is a function of the structure of the information (P, Q), of the decision problem $(c(.,.), B)$, and of the decision maker's attitude towards risk characterized by u. Section 23.3 has been devoted to the analysis of the link between the information structure and the value of information. Indeed, we could have defined a finer information structure as any structure that yields a larger value of information, independent of the decision problem and of the attitude toward risk of the decision maker. As Freixas and Kihlstrom (1984) and Willinger (1991), we examine here and now the link between V and u. We immediately see from (23.34) that an increase in risk aversion has an ambiguous effect on the value of information, since it reduces IC and UC at the same time.

23.4.2 A simple illustration: the gambler's problem

The initial reaction to most people would be that more risk-averse agents would value information more. Because the flow of information will allow agents to better manage the risk, an increase in risk aversion would make the access to the information more valuable. This intuition is not true in general, as we now show. Let us illustrate this problem with the following example. A CRRA gambler is endowed with a wealth of 10. If he accepts to gamble, a coin is tossed and the payoff is $+10$ if Head, and -5 if Tail. Without any additional information, the gambler believes that there is an equal chance to have Head or Tail. We also consider the possibility to have access to some information. There are two possible messages, with $q_1 = q_2 = 0.5$. If signal $m = 1$ is received, the posterior probability to get Head becomes 0.75, whereas if signal $m = 2$ is received, this posterior probability is only 0.25. In this latter case, the ex post expected payoff of the lottery

is negative and the risk-averse decision-maker does not gamble. The value of the information V satisfies the following condition:

$$\max\left\{\frac{\frac{1}{2}(5+V)^{1-\gamma}+\frac{1}{2}(20+V)^{1-\gamma}}{1-\gamma},\frac{(10+V)^{1-\gamma}}{1-\gamma}\right\}=\frac{1}{2}\max\left\{\frac{\frac{1}{4}(5)^{1-\gamma}+\frac{3}{4}(20)^{1-\gamma}}{1-\gamma},\frac{(10)^{1-\gamma}}{1-\gamma}\right\} \tag{23.35}$$

The value of information is not a monotone function of risk aversion.

In Figure **??**, we represented the value of information as a function of the degree of relative risk aversion. We see that the value of information is first increasing, and then decreasing, in the degree of risk aversion. In fact, the kink in this curve corresponds to the critical level of risk aversion above which the uninformed agent does not gamble. We now explain why this curve is hump-shaped. Suppose first that the gambler's risk aversion is large enough to make him to decline the lottery when he is uninformed. In this case, the value of information comes from the decision to gamble when the signal is good ($m = 1$). The gambler gains the (negative) conditional certainty equivalent associated to this signal. In consequence, a more risk-averse gambler will value this information less. This is a case where UC is zero in equation (23.34), and $V = IC$.

To sum up, an increase in risk aversion has two competing effects on the value of information. First, it reduces the value of risks taken ex-post. Therefore, it has an adverse effect on the value of information. Second, it reduces the value

of the risk taken by the uninformed agent. This makes the information structure more valuable. But this second effect does not exist when the uninformed agent does not gamble at all. We formalize this idea in the following Proposition. The gambler's problem is defined as any decision problem such that $B = \{0, 1\}$ and $c(\alpha = 0, x) = c_0$ for all $x \in X$. In words, the gambler has only two options, one of which is risk free.

Proposition 75 *In the gambler's problem, the value of information is decreasing in the degree of risk aversion if the uninformed agent does gamble.*

Proof: We consider two agents $j = 1, 2$ with utility functions u_j. We assume that agent $j = 1$ is more risk-averse than agent $j = 2$. This implies that $C_1^+(p_{m.})$ is less than $C_2^+(p_{m.})$ for all m. By assumption, the value of information for agent j is obtained by solving the following equation:

$$u_j(c_0 + V_j) = \sum_{m=1}^{M} q_m u_j(C_j^+(p_{m.})) \tag{23.36}$$

We know that $\widetilde{C}_1^+$ is dominated by $\widetilde{C}_2^+$ in the sense of first-order stochastic dominance, which implies that

$$u_2(c_0 + V_2) = \sum_{m=1}^{M} q_m u_2(C_2^+(p_{m.})) \geq \sum_{m=1}^{M} q_m u_2(C_1^+(p_{m.})). \tag{23.37}$$

Because $j = 1$ is more risk-averse than $j = 2$, Proposition 7 together with the above inequality implies that

$$u_1(c_0 + V_2) \geq \sum_{m=1}^{M} q_m u_1(C_1^+(p_{m.})) = u_1(c_0 + V_1). \tag{23.38}$$

This implies in turn that V_1 is smaller than V_2. ■

This corresponds to relatively large degrees of risk aversion. But consider now the case where risk aversion is low enough so that the decision maker would gamble even when he has no access to the information. In that case, the value of the information comes from the decision to switch to the sure position if the bad signal $m = 2$ is received. The gambler saves the conditional expected loss and the risk premium associated to that message. This case is more difficult, since the value of information is now the difference between two functions, IC and UC, that are

decreasing with risk aversion. Because the risk premium ex post is increasing with risk aversion, this has a positive effect on the value of the information. However, the intuition suggests that *the informed agent faces a reduced risk with respect to the risk taken by the uninformed one*. It implies that an increase of risk aversion will reduce UC more than it will reduce IC. This would have a positive effect on the value of information. This intuition is confirmed in Figure **??**, but we will need some more restrictions to have a theoretical foundation for it. Indeed, we know from section **??** that an increase in risk aversion does not necessarily induce people to value more a given reduction in risk.

Proposition 76 *Consider a completely informative experiment in the gambler's problem. The value of information is increasing in the degree of risk aversion if the uninformed agent does not gamble.*

Proof: Suppose that u_1 is more concave than u_2. By assumption, the value of information is defined as follows:

$$Eu_j(\widetilde{c} + V_j) = Eu_j(\widetilde{C}_j^+), \tag{23.39}$$

where $\widetilde{c}$ is the random variable that is distributed has $(c(1,1), \phi_1, ..., c(1,X), \phi_X)$. Consider a completely informative experiment where $X = M$ and where the observation of signal m implies that state m is true. We have that $C_1^+(p_{m.}) = C_2^+(p_{m.}) = \max(0, c(1,x))$. This implies that $\widetilde{C}_1^+$ and $\widetilde{C}_2^+$ are identically distributed. Moreover, their cumulative distribution single crosses the cumulative distribution of $\widetilde{c} + V_1$ at 0, from below. Indeed, we have that

$$\Pr\left[\widetilde{C}_j^+ \leq t\right] = \Pr\left[\widetilde{c} \leq t\right] \geq \Pr\left[\widetilde{c} + V_1 \leq t\right] \tag{23.40}$$

for all $t > 0$, and

$$\Pr\left[\widetilde{C}_j^+ \leq t\right] = 0 \leq \Pr\left[\widetilde{c} + V_1 \leq t\right] \tag{23.41}$$

for all $t < 0$. This means that $\widetilde{c} + V_1$ is riskier than $\widetilde{C}_1^+$ in the sense of Jewitt (1987). Using Proposition **??**, this implies that

$$Eu_1(\widetilde{c} + V_1) = Eu_1(\widetilde{C}_1^+) \qquad \Longrightarrow \qquad Eu_2(\widetilde{c} + V_1) \geq Eu_2(\widetilde{C}_1^+). \tag{23.42}$$

Because $\widetilde{C}_1^+$ and $\widetilde{C}_2^+$ are the same, this implies in turn that V_2 is smaller than V_1. Notice that when the information is not complete, $\widetilde{C}_1^+$ and $\widetilde{C}_2^+$ do not coincide and we may not conclude. ■

23.4.3 The standard portfolio problem

Let us now examine the effect of more risk aversion on the value of information when $c(\alpha, x) = w_0 + \alpha y_x$ and $B = R$. As for the gambler's problem, this effect would unambiguously be negative if UC would be zero. This is the case here only if the unconditional expectation of $\widetilde{y}$ is zero. Because the uninformed agent does not take any risk, the information will induce the decision maker to increase his risk exposure ex post. Therefore, an increase of risk aversion reduces the value of information, which is the certainty equivalent of the posterior certainty equivalents of the optimal signal-dependent portfolios.

Persico (1998) shows that this result holds with CARA functions and normal returns, even when the unconditional expectation of $\widetilde{y}$ is not zero. Treich (1997) obtains the same result without assuming CRRA, but only for the case of small portfolio risks. The general intuition is again that the informed agent will take more risk than the uninformed one. Therefore, an increase in risk aversion has a larger impact on the certainty equivalent IC of the informed agent than on the certainty equivalent UC of the uninformed agent. It yields a reduction of the value of the information. This can be shown formally as follows. Let y_x be equal to x for all x. Suppose that the unconditional distribution of $\widetilde{x}$ is normal with mean $\mu_x \geq 0$ and variance σ_x^2. From conditions (**??**) and (**??**), we know that

$$\alpha(\Phi) = \frac{\mu_x \sigma_x^2}{\sigma_x^2 A} \quad \text{and} \quad IC = w_0 + \frac{1}{2}\frac{\mu_x^2}{\sigma_x^2 A}, \tag{23.43}$$

where A is the constant absolute risk aversion of u. Suppose also that $\widetilde{m} \mid x$ is distributed as $x + \widetilde{\varepsilon}$ for all x, with $\widetilde{\varepsilon}$ is normal with mean 0 and variance σ_ε^2 and is independent of $\widetilde{x}$. It implies that $\widetilde{x} \mid m$ is also normal with

$$E\left[\widetilde{x} \mid m\right] = m \quad \text{and} \quad Var\left[\widetilde{x} \mid m\right] = \frac{\sigma_x^2 \sigma_\varepsilon^2}{\sigma_x^2 + \sigma_\varepsilon^2}. \tag{23.44}$$

Using again conditions (**??**) and (**??**), we obtain that

$$\alpha(p_{m.}) = \frac{m\left[\sigma_x^2 + \sigma_\varepsilon^2\right]}{\sigma_x^2 \sigma_\varepsilon^2 A} \quad \text{and} \quad C^+(p_{m.}) = w_0 + \frac{1}{2}\frac{m^2\left[\sigma_x^2 + \sigma_\varepsilon^2\right]}{\sigma_x^2 \sigma_\varepsilon^2 A}. \tag{23.45}$$

Notice at this stage that our basic intuition that the information increases the optimal exposure to risk is true in this case, since

$$E\left|\alpha(p_{\widetilde{m}.})\right| \geq \left|E\alpha(p_{\widetilde{m}.})\right| = \frac{\mu_x\left[\sigma_x^2 + \sigma_\varepsilon^2\right]}{\sigma_x^2 \sigma_\varepsilon^2 A} = \alpha(\Phi) + \frac{\mu_x}{\sigma_\varepsilon^2 A} \geq \alpha(\Phi). \tag{23.46}$$

Combining (23.43) and (23.45), V is defined as

$$\exp\left[-\frac{1}{2}\frac{\mu_x^2}{\sigma_x^2}\right]\exp\left[-AV\right] = E\left[\exp\left[-\frac{1}{2}\frac{\widetilde{m}^2\left[\sigma_x^2+\sigma_\varepsilon^2\right]}{\sigma_x^2\sigma_\varepsilon^2}\right]\right]. \tag{23.47}$$

This immediately implies that V is inversely proportional to the degree of absolute risk aversion.

23.5 Conclusion

A Bayesian decision maker revises his belief about the distribution of a risk by observing signals that are correlated with it. If he cannot adapt his exposure to the risk to the posterior distribution, he is indifferent to any change in the information structure that preserves the prior distribution of the risk. This is because expected utility is linear in the probabilities of the various states of the world. The information has no value in that case. An experiment generating a signal that is correlated to the unknown state of the world has a value only if the agent can modify his decision to some signals. Flexibility is valuable. The intuition is that the informed agent can at worst duplicate the rigid strategy of the uninformed agent to secure the same level of expected utility. But in general, he can do better by using the information for a risk management purpose. We examined how the value of information is affected by a change in the structure of the information and by the degree of risk aversion of the decision maker. The first exercise lead us to the notion of a finer information structure and to the important Theorem of Blackwell.

Again, the independence axiom greatly simplified the analysis by making the expected utility linear with respect to probabilities. In non-expected utility models, only functional forms that are convex with respect to probabilities will generate nonnegative values of information. Notice that the value of information needs not to be nonnegative even in the expected utility if we take into account of indirect effects of an early resolution of uncertainty.[2] Suppose for example that the lifetime utility function is not time-additive, i.e., that the felicity function of the agent at any time is affected not only by his current level of consumption, but also by his future level of expected felicity. One can be happy this month even if one's current consumption is low, but just because one is offered a free vacation in dreamland next month. Suppose that the current felicity of an anxious person is affected by

[2] See also Drèze (197?) for some interesting counterexample.

his future certainty equivalent consumption in a concave way, as is the case for some Kreps-Porteus preferences. Such anxious persons can have negative values of information. To illustrate, consider the risk of a fatal illness and a medical test to determine who will be in good health and who will be ill. An early resolution of uncertainty may be undesirable because the positive effect of knowing oneself in good health is more than compensated by the risk of learning oneself close to death. If this is not compensated by the possibility to better adapt the medical treatment to the result of the test, its value is negative. This is an explanation for why some anxious people may prefer not to test themselves for AIDS and other predictable illnesses even when the test is free. This story has been discussed in section 19.2 for Kreps-Porteus preferences where agents may prefer or dislike an early resolution of uncertainty even without any flexibility in the decision process. In Chapter **??**, we will consider an alternative explanation that is based on the effects of information on equilibrium prices and on the opportunity to share risks efficiently.

Chapter 24

Bibliography

(Very preliminary)

Arrow, K.J., (1965), *Aspects of the Theory of Risk Bearing*, Yrjo Jahnsson Lectures, Helsinki. Reprinted in Essays in the Theory of Risk Bearing (1971), Chicago: Markham Publishing Co.

Arrow, K. J. (1971). *Essays in the Theory of Risk Bearing*. Chicago: Markham Publishing Co.Barsky, R.B., F.T. Athey, S., and J. Levin, (1998), The value of information in monotone decision problems, mimeo, MIT.

Blackorby, C., D. Primont and R.R. Russell, (1999), Separability: A survey, in *Handbook of Utility Theory, vol. 1: Principles*, , ed. S. Barbera, P.J. Hammond and C. Seidl, Kluwer Academic Press, Boston.

Boll, C.H., (1955), Comparison of experiments in the infinite case and the use of invariance in establishing sufficiency, Ph. D. thesis, Stanford Univ.

Cremer, J., (1982), A simple proof of Blackwell's "comparison of experiments" theorem, *Journal of Economic Theory*, 27, 439-443.

Juster, M.S. Kimball and M.D. Shapiro, (1997), Preferences parameters and behavioral heterogeneity: an experimental approach in the health and retirement study, *Quarterly Journal of Economics, 112*, 537-579.

Campbell, J. and N.G. Mankiw, (1990), Permanent income, current income, and consumption, *Journal of Business and Economic Statistics*, 8, 265-278.

Constantinides, G.M., (1990) Habit formation: A resolution of the equity premium puzzle, *Journal of Political Economy*, 98, 519-543.

Eeckhoudt, L. and C. Gollier, (1995), *Risk: Evaluation, Management and Sharing*, Hearvester Wheatsheaf, 347 pages.

Epstein, L., and S. Zin, (1991), Substitution, Risk aversion and the temporal behavior of consumption and asset returns: An empirical framework, *Journal of Political Economy*, 99, 263-286.

Gollier, C., (1995), The Comparative Statics of Changes in Risk Revisited, *Journal of Economic Theory* , 66, 522-536.

Gollier, C. and J.W. Pratt, (1996), Risk vulnerability and the tempering effect of background risk, *Econometrica*, 64, 1109-1124.

Gollier, C. and R.J. Zeckhauser, (1998), Horizon Length and Portfolio Risk, Discussion Paper, University of Toulouse.

Guiso, L., T. Jappelli and D. Terlizzese, (1996), Income risk, borrowing constraints, and portfolio choice, *American Economic Review*, 86, 158-172.

Hall, R.E., (1988), Intertemporal substitution of consumption, *Journal of Political Economy*, 96, 221-273.

Jagannathan, R. and N.R. Kocherlakota, (1996), Why should older people invest less in stocks than younger people?, *Federal Reserve Bank of Minneapolis Quarterly Review*, 20, 11-23.

Karlin, S., (1968), *Total Positivity*, Stanford University Press.

Kocherlakota, N.R., (1996), The Equity Premium: It's Still a Puzzle, *Journal of Economic Literature*, 34, 42-71.

Kreps, D.M., and E.L. Porteus, (1978), Temporal resolution of uncertainty and dynamic choice theory, *Econometrica, 46, 185-200.*

Lehman, E.L., (1988), Comparing location experiments, Annals of Statistics, 16, 521-533.

Leland, H.E., (1980), Who Should Buy Portfolio Insurance?, *The Journal of Finance*, 35, 581-596.

Lucas, D.J., (1994), Asset pricing with undiversifiable income risk and short sales constraints: Deepening the equity premium puzzle, *Journal of Monetary Economics*, 34, 325-341.

Mankiw, N.G., and S. Zeldes, (1991), The consumption of stockholders and non-stockholders, *Journal of Financial Economics*, 29, 97-112.

Mas-Colell, A., (1991), On the uniqueness of equilibrium once again, Chap 12 in *Equilibrium Theory and Applications*, edited by W. Barnett, B. Cornet, C. d'Aspremont, J. Gabszewicz, A. Mas-Colell, Cambridge University Press.

Merton, R.C., (1969), Lifetime portfolio selection under uncertainty: The continuous- time case, *Review of Economics and Statistics*, 51, pp. 247-257.

Milgrom, P., and R. Weber, (1982), "A theory of auctions and competitive bidding, Econometrica, 50, 1089-1122.

Mitiushin, L.G., and W.M. Polterovich, (1978), Criteria for monotonicity of demand functions (in russian), *Ekonomika i Mathematichskie Metody*, 14, 122-128.

Mossin, J., (1968), Optimal multiperiod portfolio policies, *Journal of Business*, 215-229.

Persico, N., (1998), Information acquisition in affiliated decision problems, mimeo, UCLA.

Ponssard, J.P., (1975), A note on information value theory for experiments defined in extensive form, *Management Sciences*, 22.

Pratt, J., (1964), Risk Aversion in the Small and in the Large, *Econometrica*, 32, 122-136.

Mehra, R. and E. Prescott, (1985), The Equity Premium: A Puzzle, *Journal of Monetary Economics*, 10, 335-339.

Salanié, B., (1997), *The Economics of contracts: A primer*, MIT Press, Boston.

Samuelson, P.A., (1963), Risk and uncertainty: the fallacy of the Law of Large Numbers, *Scientia*, 98, 108-113.

Samuelson, P.A., (1969), Lifetime portfolio selection by dynamic stochastic programming, *Review of Economics and Statistics*, 51, 239-246.

Samuelson, P.A., (1989), The judgement of economic science on rationale portfolio management: indexing, timing, and long-horizon effects, *Journal of Portfolio Management*, Fall 1989, 3-12.

Savage, L.J., (1954), *The foundations of statistics*, New York: Wiley.

Selden, L., (1979), An OCE analysis of the effect of uncertainty on saving under risk independence, *Review of Economic Studies*, 73-82.

Telmer, C.I., (1993), Asset-pricing puzzles and incomplete markets, *Journal of Finance*, 48, 1803-1832.

Townsend, R.M., (1984), Risk and insurance in village India, *Econometrica*, 62, 539-592.

Treich, N., (1997), Risk tolerance and value of information in the standard portfolio model, *Economics Letters*, 55, 361-363.

von Neumann, J. and O. Morgenstern, (1944), *Theory of Games and Economic Behavior*, Princeton: Princeton University Press, 1944.

Weil, P., (1990), Nonexpected utility in macroeconomics, *Quarterly Journal of Economics*,105, 29-42.

Wilson, R. (1968). "The theory of syndicates", *Econometrica* 36, 113-132.

Index