APPLICATION OF THE THEORY OF STOCHASTIC CONTROL TO FINANCIAL AND ECONOMIC SYSTEMS

by

Caro Lucas

Memorandum No. ERL-M597

15 August 1976

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Application of the Theory of Stochastic Control to Financial and Economic Systems

Caro Lucas

Ph.D.

Dept. of Electrical Engineering
and Computer Sciences

Abstract

It is proved that under some very general conditions the "true" present value of a share will follow a "random walk" rule. Given the dividend policy of a firm, the differential equation for the evolution of the "true price" is derived and it is shown that the "true price" function is the potential corresponding to the dividend policy function. Also for capital markets where several different stocks are traded an intertemporal pricing model is discussed and conditions under which "mean-variance" efficient stocks are optimal for all the participants in the market are derived. The vertical market models that result from these conditions are examined and differential equations for "rational" prices for options are obtained therefrom.

Eugene Wong

Chairman of Committee

Table of Contents

Acknowledgement

I wish to acknowledge the great amount of help and direction given to me by my research advisor Professor Eugene Wong. I have benefited from his special attention, and encouragement throughout the period of research and preparation of the dissertation.

My gratitude is also extended to Professors P. P. Varaiya, K. R. A. Artle, C. A Desoer, as well as other faculty members at the University of California, Berkeley, and my colleages here, without whose help this research would not have been possible.

Finally, I would also like to acknowledge the financial support provided by the Joint Services Electronics Program Contract F44620-71-C-0087 and the National Science Foundation Grant ENG74-06651-A01.

Introduction

The purpose of this dissertation is to apply some recent developments in the Theory of Stochastic Processes to dynamical-system models arising in the theory of "intrinsic" or "rational" prices, and in the theory of Capital Markets. In particular, we shall make heavy use of the martingale calculus, whose origin was the Ito calculus but whose recent development due to Kunita-Watanabe, Meyer and others have made it a natural tool in the study of stochastic systems. Application of the theory of martingales to economic problems is particularly appropriate, since the very nature of martingales gives a prominent role to the evolution of the information pattern, and such forms are necessary in many stochastic systems in economics.

Much of the recent literature on finance has emphasized the dynamical nature of the underlying models. That is not surprising since the main concern of the capital theory is to find the best allocation for a given amount of resource at any given time. There is also little doubt that because of the large number of uncertain factors involved, risk analysis plays an important role for a successful investment policy. The dissertation consists of five chapters and a Preliminary chapter.

In the preliminary part , we shall present a brief exposition of the mathematical material which is used in the remainder of the dissertation. Special attention will be given to martingales and Markov processes. The first of these processes plays a natural role whenever one's information is evolving in time. The martingale Calculus provides a powerful tool for the development of a dynamical theory for some important aspects of stochastic processes. The Markov process is important because it is used to develop a state-space model analogons to the deterministic case. In chapter 1, an economic interpretation for stochastically dominant stocks and their relative volatilities will be given for the dynamic case. In chapter 2, the Fundamentalist's true price model is examined. Under, very general conditions it is demonstrated that the "true prices" will have "random walk" fluctuations. Also for the normal growth model the optimal dividend policy to maximize the "true price" is obtained.

The case where the information is generated by a Markovian state variable is discussed in chapter 3. A differential equation for the evalation of the "true price" is derived. It is shown that the price is the r-potential corresponding to the dividend policy function where r is the discount rate prevailing in the market. The potential kernel is the Laplace transform of the probability transition function for the Markov state variable. Also the "true price" operator is shown to be the Laplace transform of the differential generator corresponding to the state process. This operator, also called the resolvant operator satisfies the resolvant equation, which gives the changes in the true prices due to a change in the discount rate. When one considers the fact that prices cannot become negative, this shows that "true prices" are monotonously decreasing functions of the discount rate. It also explicitely gives the amount of changes in the "true price" if the rate of discounting is changed.

The results are further generalized for the case of futures prices as well as the case of stochastic discount rates. Finally, the effect of various transformations of the Markov state process is examined and the equivalence of a termination of the process with discounting is shown, suggesting an interpretation for discounting the future incomes.

In the next chapter, a vertical market model which excludes the possibility of arbitrage is derived and applied to obtain a set of differential equations for the "rational prices" for options and warrants.

In the final chapter we examine market models where there are alternative opportunities for investment. The conventional mean-variance portfolio selection model is generalized for the dynamical case. It is shown, however, that when there are uncertainties concerning the future changes of market opportunities the mean-variance model is no longer valid and that investors would be willing to choose their portfolio so as to smooth the future variation in their consumption process. Therefore, the behaviour of the investors in a random and dynamic environment cannot be characterized by adding the effects of uncertainty introduced in the static problem to the consumption smoothing that gives the best allocation of consumption in the dynamic-deterministic problem. The best portfolio is seen to be a combination of a minimum-risk portfolio, a portfolio that will give the best expected returns, and m other portfolies that will give higher returns when future opportunities for investment become less favorable. Thus investors will always choose a portion of their portfolio so as to minimize the variation in their consumption due to uncertainties concerning future market conditions. These "insurance" portfolies are shown to be independent of the investors' information and/or probability beliefs. Also it is shown that as a consequence of the desire of the investors to minimize the investment risk the rates of return will satisfy a market equation that will exclude the possibility of arbitrage in the sense described in Chapter 4.

Chapter 0
Mathematical Preliminaries

0.1. Definitions and Interpretations

Uncertainty prevails in capital markets because of the complexity of determining how various stock and commodity prices respond to the different events that occur daily. The introduction of uncertainty into the analysis of capital markets was first done in the pioneering work of Louis Bachelier [3] which anticipated later developments in the theory of Probability and Stochastic Processes. The subject has now become of much interest to statisticians and probability theorists because of the greater availability of data concerning price fluctuations and other economic variables, and the fact that an ability to predict future price changes has an obvious speculative interest.

We begin with a discussion of the mathematical preliminaries. Firstly, one should model the information pattern for the participants in the market. Thus ones information is usually represented by a family $\mathcal{A}$ of events; i.e. subsets of a universal (certain) event Ω in the sense that the occurance of an event A is known if and only if $A \in \mathcal{A}$. Since the information concerning a countable number of events enables one to also know whether their conjunction or disjunction is true or not we can axiomize that a σ-algebra (information pattern) $\mathcal{A}$ is a family of subsets of Ω, including Ω itself, that is closed under countable Boolean operations. $(\Omega, \mathcal{A})$ will then be called a measurable space, and one could assign prior probability measures concerning the likelihood of occurrence of the events constituting the σ-algebra. Such measures should assign probability one to the certain event Ω and should be additive for mutually exclusive events i.e. we should have

$$P(\Omega) = 1 \tag{0.1}$$

$$P(\bigcup_n A_n) = \sum_n P(A_n) \quad \text{if} \quad A_i \cap A_j = 0 \quad \text{whenever} \quad i \neq j \tag{0.2}$$

The triple $(\Omega, \mathcal{A}, P)$ will be called a probability space if the corresponding axioms are satisfied.

If $(\Omega, \mathcal{A})$ and $(E, \mathcal{E})$ are two measurable spaces and f is a function mapping into E then f is said to be a measurable function or a random variable if and only if for any $A \in \mathcal{E}$ we have

$$f^{-1}(A) \in \mathcal{A}$$

Now if there is a probability measure P defined on $(\Omega, \mathcal{A})$ then a random variable will induce another measure P' on $(E, \mathcal{E})$ so that

$$P'(A) = P(f^{-1}(A)) \tag{0.3}$$

Alternatively a random variable can be viewed as mapping the measure on $(\Omega, \mathcal{A})$ into a measure on $(E, \mathcal{E})$ by Eq. (1.3) hereafter, sometimes abbreviated as

$$P' = Pf^{-1} \tag{0.4}$$

Now a collection of events (i.e. subsets of Ω) will generate an information-σ-algebra if we augment all the other events that can be obtained from them by a countable number of Boolean operations. Thus the σ-algebra generated by a family $\mathcal{A}'$ of subsets of Ω is the smallest σ-algebra containing $\mathcal{A}'$ and will be shown as $\sigma(\mathcal{A}')$. Similarly observing a function with a measurable range will generate a σ-algebra, namely the σ-algebra generated by all the inverse images [27].

The expected value of a random variable given its probability law is the integral

$$E[f] = \int_{\Omega} f(\omega) \ P(d\omega) = \int_{E} \varepsilon \, Pf^{-1}(d\varepsilon) \tag{0.5}$$

i.e. to calculate the expected value of a random variable one can either integrate over the domain the function with respect to the corresponding measure or one can integrate the values of the function over the range of the function with respect to the induced measure. The probability interpretation of the mathematical expectation is the ensamble average of the values that the random variable can assume and in that sense it is what one can expect "in the average" for the value of the random variable.

To incorporate the dynamics let us define an increasing family of σ-algebras to be the indexed family $\mathcal{F}_t$ of events such that we have $\mathcal{F}_t \supset \mathcal{F}_s$ whenever $t \geqq s$. Clearly, this corresponds to "not forgeting" the information at any time. ($\mathcal{F}_t$ represents ones information at time t).

Also a stochastic process is an indexed family of random variable X_t such that for each t X_t is measurable. Again, one can imagine a bigger σ-algebra generated by the cylinder sets $S_t(A) = \{X_t \in A\}$ for some t and some $A \in \mathcal{A}$ and thus view the process as a measurable mapping into this bigger space [49].

In practice the set of all time functions is usually too big and moreover, by a combination of observation measurement and intuitive judgement, one starts from a compatible family of finite dimensional probability distributions

$$P[X_{t_1} \leqq x_1, X_{t_2} \leqq x_2, \ldots, X_{t_n} \leqq x_n]$$

and is free to construct the probability space. Details are given in [49]. Two different stochastic processes that have the same finite dimensional distributions are said to be equivalent.

It can be proved that any stochastic process has a separable equivalent [27]. If it is also continuous in probability i.e. for $\varepsilon > 0$ $\lim_{s\to t} (|X_s - X_t| \geqq \varepsilon) = 0$ and satisfies the Kolmogorov condition

$$E|X_{t+h} - X_t|^{\alpha} \leqq Ch^{1+\beta} \tag{0.6}$$

for strictly positive constants α, β, c, then it is measurable and every sample function (of t) is uniformly continuous [49]. These conditions involve only the finite dimensional distributions.

We shall be mainly interested in two class of processes: namely martingales and Markov processes, which will be described in the subsequent sections. For now, let us proceed to define the conditional expectation, which will be required to study the evolution of factors determined by changing information.

If M_1 and M_2 are two different measures on $(\Omega, \mathcal{A})$ then there exists [49] a random variable Λ and a measure μ such that

$$M_2(A) = \int_A \Lambda M_1(d\omega) + \mu(A) \tag{0.7}$$

for $A \in \mathcal{A}$.

If $\mu = 0$ then M_1 and M_2 are said to be absolutely continuous and Λ

will then be called the Radon-Nikodym derivative of M_2 with respect to M_1.

If $\mathcal{A}'$ is a sub-σ-algebra of $\mathcal{A}$ and I_A denotes the indicator function that assumes the value one for the points belonging to $A \in \mathcal{A}'$ and zero otherwise then the Radon-Nikodym derivative of the indefinite integrals $E[I_A X]$ with respect to the restriction of the probability measure on $\mathcal{A}'$ is (almost surely) a random variable called the conditional expectation of X given $\mathcal{A}'$ and denoted by $\Lambda = E|X|\mathcal{A}']$. It is characterized by the properties of being measurable with respect to $\mathcal{A}'$ and having the smoothing property

$$\int_A X(\omega) \quad P(d\omega) = \int_A \Lambda(\omega) \quad P^*(d\omega) \tag{0.8}$$

for all $A \in \mathcal{A}'$ where P^* is the restriction of P on $\mathcal{A}'$.

From what was discussed above it is clear that if one's information is given by $\mathcal{A}'$ then Λ is a "smoother" version of X in the sense that it is adapted to the information $\mathcal{A}'$ and over any "known" event A it has the same average as X. Of course $\mathcal{A}'$ can be generated by a collection of sets or functions. For the latter case, the conditional expectation can also be denoted as $E[X|f]$.

0.2. Markov Processes

A stochastic process X_t is said to be Markovian if its future is conditionally independent of its past given its present. That is, the probability law for the future evolution of the process given the past and present values of the process depends only on the present.

If ones information is generated by observing a Markov process, then the knowledge of the present value suffices in determining the probability measure concerning the likelihood of the future events and thus is analogous with the state variable for the deterministic case.

Definition 2.1. Let $\mathcal{F}_t$ be an increasing family of σ-algebras. A nonnegative random variable T is called a stopping time if the events $\{T \leq t\}$ are always adapted to (measurable with respect to) $\mathcal{F}_t$. Thus at time t the information contained in $\mathcal{F}_t$ would enable one to determine whether or not T has "stopped," i.e. $\{T \leq t\}$.

Now if $(\mathcal{F}_t, T)$ is a stopping time and μ is a measure defined on the state space (range of X_t) then one can define a Markov process X_t on the interval (0,T) such that μ is the induced measure at time zero.

If $\mathcal{X}_t^+$ is the σ-algebra generated by X_s, $s \geq t$ then the Markov condition is stated as follows:

$$E[Z|\mathcal{F}_t] = E[Z|X_t] \tag{0.9}$$

where Z is any $\mathcal{X}_t^+$-measurable random variable. The general theory for Markov processes has been of interest for a very long time. There are many references e.g. [12.14]. One main result is its close connection with the potential theory which we shall see when we analyse the fundamentalist's "true prices." Rigourous analysis may be found in [14,8].

One special Markov process is the Brownian Motion, a natural phenomenon the existance of which was known in 1827. The theory was later developed by Bachelier, Wiener, and others. A Markov process defined on nonnegative times is a Brownian Motion if it is Gaussian, i.e. every linear combination of the form $\alpha_i X_{t_i}$ is a Gaussian random variable with density $f(n) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}}$ and if we have

$$EX_t X_s = \min(t,s) \tag{0.10}$$

Usually, a separable version is chosen, and it will then be sample continuous [12].

Also one important class of Markov processes are the (Ito) diffusion processes: let X_t be a sample continuous Markov process with

$$m(X_t,t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} E[X_{t+\Delta}-X_t | X_t] \tag{0.11}$$

$$\sigma^2(X_t,t) = \lim_{\Delta \downarrow 0} \frac{1}{\Delta} E[(X_{t+\Delta}-X_t)^2 | X_t] \tag{0.12}$$

then under some general conditions [49] X_t can be represented as the unique solution to the stochastic differential equation

$$dX_t = m(X_t,t)\,dt + \sigma(X_t,t)dW_t \tag{0.13}$$

where W_t is a Brownian Motion Process. Stochastic integrals and differential equations will be discussed in the next section. The functions m and σ can be applied to obtain the differential generator [50] as well as the equations for transition probability functions [12]. The main application of Eq. (0.13) however is to generalize the deterministic state equations of the form

$$\frac{d}{dt} x_t = f(x_t, t) \tag{0.14}$$

Using the celebrated Ito differentiation rule, one can use Eq. (0.13) for formulating solutions to problems concerning stochastic control of a system [10,13,25,47].

0.3. Martingales

It can be easily demonstrated that for a Brownian Motion Process the following is always true.

$$\varepsilon\{W_t|\mathcal{F}_{WS}\} = W_S \quad \text{for} \quad t \geqq s \tag{0.15}$$

where $\mathcal{F}_{WS}$ is the σ-algebra generated by observing W up to the time s.

An important class of processes resultsfrom a generalization of this property of Brownian Motions. Let $\mathcal{F}_t$ be an increasing family of σ-algebras and let X_t be a process adapted to $\mathcal{F}_t$. Hereafter unless otherwise stated we shall be mainly concerned with sample continuous processes for which we have a better developed analytic theory. Let us define dX_t as the forward difference $X_{t+dt}-X_t$. (Thus we shall so following the stochastic calculus convention due to Ito). Now X_t is said to be an $\mathcal{F}_t$-martingale if

$$\varepsilon\{dX_t|\mathcal{F}_t\} = 0 \tag{0.16}$$

which is equivalent to

$$\varepsilon\{X_t|\mathcal{F}_s\} = X_s \quad \text{for} \quad t \geqq s \tag{0.17}$$

Martingales occur naturally whenever one considers conditional expectations with respect to evolving information (increasing family of σ-algebras). In particular, a process Y_t of the form

$$Y_t = \varepsilon\{R|\mathcal{F}_t\} \tag{0.18}$$

where R is a random variable and $\mathcal{F}_t$ is an increasing family of σ-algebras can easily shown to be a martinglae [45]. Perhaps the importance of martingale theory stems from the fact that it provides us with the powerful analytic tool of the martingale calculus. Unless otherwise stated, difficulties increasing family of information $_t$ is always assumed to be continuous in the sense that

$$\mathcal{F}'_t = \bigcap_{s>t} \mathcal{F}_s \tag{0.19}$$

$$\mathcal{F}_t = \sigma\left(\bigcup_{s<t} \mathcal{F}_s\right) \tag{0.20}$$

Now let X_t be a general stochastic process and define B_t and $\hat{X}_t$ by:

$$dB_t = \varepsilon\{dX_t | \mathcal{F}_t\} \tag{0.21}$$

$$\hat{X}_t = \varepsilon\{X_t | \mathcal{F}_t\} \tag{0.22}$$

Then $M_t = \hat{X}_t - B_t$ is a Martingale adapted to $\mathcal{F}_t$. When B_t is of bounded variation $\hat{X}_t$ is the sum of a martingale and a process of bounded variation adapted to $\mathcal{F}_t$. We shall call processes that admit such decompositions semi-martingales. In particular if $\mathcal{F}_t$ is the σ-algebra generated by $\hat{X}_t$ then we have $\hat{X}_t = X_t$ and then martingale term M_t given by

$$dM_t = dX_t - \varepsilon\{dX_t | \mathcal{F}_{Xt}\} \tag{0.23}$$

represents the new information contained in dX_t and is innovations process for X_t.

Also if one considers the decomposition of the process $X_t = M_t^2$ where M_t is a second order $\mathcal{F}_t$ martingale the predictable part is nondecreasing and nonnegative and we have

$$\varepsilon\{B_t\} = \varepsilon\{X_t - X_0\} = \varepsilon\{(M_t - M_0)^2\} \tag{0.24}$$

$$dB_t = \varepsilon\{dX_t|\mathcal{F}_t\} = \varepsilon\{(dM_t)^2|\mathcal{F}_t\} \tag{0.25}$$

Definition 3.1. This process denoted by $B_t = \langle M,M\rangle_t$ is called the quadratic variation of the martingale process M_t.

An important result is that the quadratic variation does not depend on the family of σ-algebras $\mathcal{F}_t$ and can be constructed by computing the following limit (described in [47,48].

$$\langle M,M\rangle_t = \lim_{n\to\infty}\sum_{\nu}[M(t_{\nu+1}^{(n)})-M(t_{\nu}^{(n)})]^2 \tag{0.26}$$

where the sequence of partitions (t_ν^n) refines to zero as n goes to infinity.

If N_t and M_t are two $\mathcal{F}_t$ martingales then the co-variation process is defined as

$$\langle M,N\rangle_t = \frac{1}{4}\,[\langle M+N,M+N\rangle_t - \langle M-N,M-N\rangle_t] \tag{0.27}$$

One important result of (0.26) is that if we consider two different decompositions of a semi-martingale X_t relative to two families of σ-algebras, then the martingale terms will have the same quadratic variations. Also it turns out that the quadratic variation of a semi-martingale is equal to that of its martingale term in a decomposition of the process, i.e. if $X_t = B_t+M_t$ where dB_t is adapted to $\mathcal{F}_t$, B_t is of bounded variation, and M_t is an $\mathcal{F}_t$ martingale, then

$$\langle X,X\rangle_t = \lim_{n\to\infty}\sum_{\nu}[X(t_{\nu+1}^{(n)})-X(t_{\nu}^{(n)})]^2 = \langle M,M\rangle_t \tag{0.28}$$

Now we are equipped to study the stochastic calculus. If M_t is a sample continuous $\mathcal{F}_t$ martingale and Φ_t is a process adapted to $\mathcal{F}_t$ such that

$$\int_0^T \Phi_t^2 \, d\langle M,M\rangle_t < \infty \text{ with probability 1} \tag{0.29}$$

then the stochastic integral $\int_0^T \Phi_t dM_t$ is defined by the following limit [48].

$$\int_0^T \Phi_t \, dM_t = \lim_{n\to\infty} \sum_\nu \Phi_{t_\nu^{(n)}} [M_{t_{\nu+1}^{(n)}} - M_{t_\nu^{(n)}}] \tag{0.30}$$

However, in light of the recent theory of stochastic integral developed by Kunita-Watanabe [24], stochastic integrals are considered as transformations of local martingales. If we define a sequence of stopping time

$$\tau_n = \min\{t: |M_t| \geqq n\} \tag{0.31}$$

then M_t is said to be an $\mathcal{F}_t$ local martingale if the stopped process $M_{\min(t,\tau_n)}$ is an $\mathcal{F}_t$ martingale.

Definition 3.2. Given a local martingale and a process Φ_t satisfying (2.29) $Z_t = \int_0^t \Phi_s dM_s$ is another $\mathcal{F}_t$ local martingale defined uniquely by

$$\langle Z,Y\rangle_t = \int_0^t \Phi_s d\langle Y,M\rangle_s$$

for every local $\mathcal{F}_t$ martingale Y_t.

Similarly, an integral of the type $\int_0^t \Phi_s dX_s$ is defined for local semi-martingales to be another local semi-martingale composed of a Stieltjes integral $\int_0^t \Phi_s dB_s$ (the predictable term) and a stochastic integral $\int_0^t \Phi_s dM_s$ defined as above. Now we can interpret a stochastic differential equation to be a differential equality in the sense that when integrated the two sides will be equal.

An important result is the stochastic differentiation rule. If X_t is a local semi-martingale then the differential of any function of X_t is given by

$$df(X_t,t) = f_t(X_t,t)\,dt + \sum_i f_i(X_t,t)\,dX_{it} + \sum_i \sum_j f_{ij}(X_t,t)\,d\langle X_i,X_j\rangle_t \tag{0.33}$$

where

$$f_t(x,t) = \frac{\partial f}{\partial t} \qquad f_{ij}(x,t) = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

$$f_i(x,t) = \frac{\partial f}{\partial x_i}$$

This equation shows clearly how a function of a stochastic process changes not only because of the changes of that variable, but also because of the uncertainty concerning its future fluctuation.

An active area of research in the theory of Martingales is the conditions under which a process can be represented by the stochastic

An important representation result on martingales is that if $\mathcal{F}_{Xt}$ is an increasing family of σ-algebras generated by a (local) semi-martingale X_t, and if the martingale term in the $\mathcal{F}_{Xt}$ decomposition is a Brownian Motion (for this it suffices that $\langle X,X\rangle_t = t$) [48] then every $\mathcal{F}_{Xt}$ local semi-martingale can be represented by a predictable term and an integral involving X_t i.e. if we have

$$X_t = B_t + M_t \tag{0.34}$$

where B_t is of bounded variation, dB_t and M_t are $\mathcal{F}_{Xt}$ measurable, and $\langle M,M\rangle_t = t$, then every $\mathcal{F}_{Xt}$ local semi-martingale y_t has the following representation

$$y_t - y_o = A_t + \int_0^t \Phi_s dM_s \tag{0.35}$$

where A_t is again the predictable part given by

$$dA_t = \varepsilon\{dY_t|\mathcal{F}_{Xt}\}$$

Finally, an important application of the martingale theory is in studying absolutely continuous changes of measure. Consider a σ-algebra $\mathcal{F}$ and mutually absolutely continuous probability measures P and P_0. Let Λ be the Random-Nikodym derivative of P with respect to P_0. If $\mathcal{F}$ is generated by X_t then Λ is a function of past values of X_t. Let $\mathcal{F}_t$ be an increasing information pattern and let us define

$$\Lambda_t = \varepsilon_0\{\Lambda|\mathcal{F}_t\} \tag{0.37}$$

Clearly, Λ_t is a martingale under the probability measure P_0, and so is the process M_t defined by

$$dM_t = \frac{d\Lambda_t}{\Lambda_t} \tag{0.38}$$

or in its integrated form

$$\Lambda_t = \exp(M_t - \frac{1}{2}\langle M,M\rangle_t) \tag{0.39}$$

Now if Z_t is a local martingale under P then it can be proved that $Z_t\Lambda_t$ and thus $X_t = Z_t + \langle Z,M\rangle_t$ are local martingale under P_0 [48]. Therefore, under that measure we would have

$$\varepsilon_0\{dZ_t|\mathcal{F}_t\} = -d\langle Z,M\rangle_t \tag{0.40}$$

If the family $\mathcal{F}_t$ is generated by a Brownian Motion process, or a local semi-martingale that can be decomposed into a Brownian Motion process, and a process of bounded variation, then M_t can be represented as a stochastic integral of that process and so can be the predictable part of any local martingale under one measure [48].

Many of the results stated above can be generalized to the case of martingale processes that are not sample continuous. However, the details are more complex and we will not consider them here.

Chapter 1

Market Analysis Under Uncertainty

1.1. Volatility of Investment Returns

An uncertain investment for a fixed period is usually characterized by a random variable R called the investment return, which shows the amount of money one gets at the end of the period for each dollar of investment. In the dynamical case, the investment opportunity can similarly be characterized by its return process. If the amount W_t is invested at time zero, the value one can get at each time t is the random variable W_t that depends on the information known at time t.

Let us define the return R_t to be a stochastic process adapted to the increasing information pattern $\mathcal{F}_t$ and given by the following equation

$$dR_t = \frac{dW_t}{W_t} \tag{1.1}$$

The following theorem provides us with a definition of absolute volatility of the return process at any time:

Proposition (1.1): If $\{X_t, \mathcal{F}_t\}$ is a sample continuous second order martingale and if there exists a nonnegative measurable process ψ_t adapted to $\mathcal{F}_t$ such that for t > s.

$$\varepsilon\{(X_t - X_s)^2 | \mathcal{F}_s\} = \int_s^t \varepsilon\{\psi_t | \mathcal{F}_s\}\, dt \tag{1.2}$$

and if the set $\{\psi = 0\}$ has zero d P dt measure (the product of the probability and the Legesgue measure) then there exists a Brownian Motion M_t adapted to $\mathcal{F}_t$ such that we have

$$dX_t = \psi_t^{1/2}\, dM_t \tag{1.3}$$

with probability one. With the hypothesis that ψ vanishes almost nowhere (1.3) is still valid with the adjunction of a Brownian motion to the probability space [47].

Clearly the derivative of the quadratic variation with respect to Lebesgue measure satisfies (1.2) [47]. The condition that $\langle X,X\rangle_t$ be almost surely continuous with respect to the Lebesgue measure is stringent and also hard to verify. One can extend the theorem for local martingale for right continuous completed family $\mathcal{F}_t$ with replacing some other F measure instead of the Lebesgue measure to get [47].

$$dX_t = \phi_t \, dM_{F(t)} \tag{1.4}$$

For the return process R_t this theorem can be applied to the martingale term in its decomposition with respect to $\mathcal{F}_t$. We will always assume that the return process R_t is continuous and a (local) semi-martingale relative to $\mathcal{F}_t$. This is true if $\mathcal{F}_t$ is generated by the past values of R_t. The implication of this last assumption is that ones information is such that the predictable term of future returns is always of bounded variation.

Since we know that the quadratic variation of a process does not depend on one's information $\mathcal{F}_t$, this representation theorem suggests that we should take the following process to be a measure of volatility of the investment at any time t.

$$\phi_t = [\frac{d}{dt} \langle R,R\rangle_t]^{1/2} \tag{1.5}$$

where the derivative is taken with respect to the Legesgue measure. Because one could write

$$d\,R_r = d\,\hat{R}_t + \phi_t dM_t \tag{1.6}$$

where

$$d\hat{R}_t = \varepsilon\{dR_t|\mathcal{F}_t\} \tag{1.7}$$

ϕ_t can be interpreted at the change in the return process in addition to its expected change due to a unit change in the sample outcome of the (unpredictable) Brownian Motion process. (The derivative of the Brownian Motion process is sometimes called white noise because its spectral density function is flat. If the process $\langle R,R\rangle_t$ is not absolutely continuous with respect to Lebesgue measure one can consider its derivative with respect to some other F measure. It is perhaps more convenient to consider the square of (1.5) as a generalized function to avoid technical difficulties.

It is clear that the more volatile an investment opportunity is the more changes one could expect in an infinitesimal time, because we have

$$d\langle R,R\rangle_t = \varepsilon\{(dR_t-d\hat{R}_t)^2|\mathcal{F}_t\} \tag{1.8}$$

Thus the volatility is roughly equivalent to the (conditional) standard deviation of the infinitesimal return dR_t given the information $\mathcal{F}_t$. We shall shortly see how this variable can be a measure of risk for the investors, however, a relative or "systematic" volatility is widely considered to be more insightful than the absolute measure thus defined. Consider a return process R_t and an index X_t both of which are (local) semi-martingales with respect to the information $\mathcal{F}_t$. The index could be any economic process such as the rate of return on another investment opportunity. Both R_t and X_t can be decomposed into their predictable and risky terms with respect to the information $\mathcal{F}_t$.

$$R_t = \hat{R}_t + M_t \tag{1.9}$$

$$X_t = X_t + \eta_t \tag{1.10}$$

Suppose we want to approximate M_t with a stochastic integral of η_t given by

$$N_t = \int_0^t \beta_s d\eta_s + M_o \tag{1.11}$$

such that the error process $e_t = M_t - N_t$ is "orthogonal" to η_t in the following sense.

$$d\langle e,\eta\rangle_t = d\langle M,\eta\rangle_t - \beta_t d\langle \eta,\eta\rangle_t = 0 \tag{1.12}$$

The coefficient β_t will be called the volatility of the return process R_t with respect to the index X_t. It can be determined from Eq. (1.12) to be

$$\beta_t = \frac{d\langle M,\eta\rangle_t}{d\langle \eta,\eta\rangle_t} = \frac{d\langle R,X\rangle_t}{d\langle X,X\rangle_t} \tag{1.13}$$

which shows that it only depends on the processes R_t, X_t and does not depend on the information pattern $\mathcal{F}_t$.

1.2. Stochastic Dominance

The purpose of this section is to extend the stochastic dominance concept as defined in [21,35,36] to the dynamical case. Let U(W) denote the utility function for a wealth level W. We assume that all investors are maximizers of expected utility in the sense of Von Neumann-Morgenstern [11,17] and that their utility functions are concave and increasing. This means that the investors are risk-averse in the sense of Arrow-Pratt [1,43], and they prefer more wealth to less wealth.

Here we are addressing the question as to whether an investment can be regarded as being superior to another one for all such investors. The question as to whether to invest and how much will be considered in later chapters. First, let W_{1t} and W_{2t} denote respectively the wealth processes associated with the first and second investment opportunities, and let R_{1t} and R_{2t} be the corresponding return processes. From the increasing property of the utility function we conclude that the first investment opportunity is superior to the second one if the following is true for all t

$$dW_{1t} \geq dW_{2t} \quad \text{a.s.} \tag{1.14}$$

which is equivalent to (considering (1.1.1)

$$dR_{1t} \geq dR_{2t} \quad \text{a.s.} \tag{1.15}$$

However, these are only necessary conditions based on the increasing property of u. If now we consider the charge in the expected utility of wealth at each time given the information $\mathcal{F}_t$ at that time

$$\varepsilon\{du(W_t)|\mathcal{F}_t\} = u'(w_t)\ \varepsilon\{dw_t \setminus \mathcal{F}_t\} + \frac{1}{2} u''(w_t)\ d\langle w,w\rangle_t \tag{1.16}$$

where u' and u" stand for the first and second derivatives of the utility function and we have

$$u'(w_t) \geq 0 \tag{1.17}$$

$$u''(w_t) \leq 0 \tag{1.18}$$

we can see from (1.16) that among the investments with the same expected change in one's wealth the one with the least quadratic variation dominates the others. Using (1.1) once more and taking notice of

$$d\langle R,R\rangle_t = \frac{1}{w_t^2} d\langle w,w\rangle_t \tag{1.19}$$

we can state the equivalent statement that among the investment oppportunities with the same expected rate of return (given the information at that time) $\varepsilon\{dR_t|\mathcal{F}_t\}$, the one with the least quadratic variation term $d\langle R,R\rangle_t$ dominates the others.

From the above discussion it is apparent that in investing one seeks to minimize the volatility of his investment returns at any time. Any risk averse investor would demand higher expected rate of return for a stock of higher volatility, the rate of compensation, which is to be considered as the price of risk is equal to

$$-\frac{1}{2}\frac{u''(w_t)}{u'(w_t)} \quad \text{for wealth process and} \tag{1.20}$$

$$-\frac{1}{2}\frac{u''(w_t)}{w_t u'(w_t)} \quad \text{for the rate of return} \tag{1.21}$$

process. These terms are the respective measures of absolute and relative risk aversions of the investor [1,43] i.e. the more risk-averse investors (sometimes called conservative investors) would demand higher expected rate of return for the volatility of their stock.

Chapter 2

A Random Walk Theory of Fundamentalists' "Intrinsic Price"

2.1. Preliminaries

The professional stock market analysts are widely [9] catagorized according to which one of the following two main schools they belong to.

On the one hand there are the "fundamentalists," who believe that a share must have a "true" value, which depends on the performance of the corporation and thus strive to find the relationships between prices and such external factors that lie behind price changes as earnings and dividends paid by various companies; so that by getting the necessary information concerning the prospects for profits in different industries and individual firms, and using this information in the "fundamental" equations they could detect the trends in future price changes and make a profit from that knowledge.

The "technicians," on the other hand, argue that since the stock markets are nothing but speculative exchange markets the price of a share is what someone is willing to pay for it. Accordingly they study the past history of the price fluctuations in order to predict future prices.

However, empirical studies showed that [9,37] the price fluctuations obeyed a "Random Walk" rule and could hardly be told from independent random numbers, which seemed to negate any economic law governing these price changes. Although in its extreme position the random walk theory suggests that past history of prices has no predictive value for future changes, the possibility that observing some external economic processes might help one infer something about future prices is in no way precluded. Some of the recent works [19,26,38,39] have shown that the fundamentalist's pricing mechanisms and the random walk theories are not only not irreconcilable but in fact, imply each other. A starting point is Fisher's present discounted value rule, [17] which defines the "true" value of a stock as the discounted sum of its future dividend payments. For a constant discount rate r we have

$$V_t = \int_t^{\infty} e^{-r(s-t)} \, dD_s \qquad (2.1)$$

where D_t is the cumulative dividends paid up to time t.

Even for the deterministic case the "true value" will change as a function of time due to changes in future dividend payment.

Simple differentiation of (2.1) yields

$$dV_t + dD_t = rV_t \, dt \qquad (2.2)$$

or if we define the pseudo-return by

$$dR_t = (dV_t + dD_t)/V_t \qquad (2.3)$$

then (2.2) implies that this must be equal to the interest (discount) rate.

$$dR_t - rdt = 0 \qquad (2.4)$$

Therefore, one cannot change the pseudo-return by changing the dividend payments, because the "true value" will also change accordingly. Similar results can be obtained for the futures prices, i.e. prices for contracts for transactions that should be carried out at future date. If we have

$$V_t^T = \exp[-r(T-t)] \; C_T \qquad (2.5)$$

where C_T is the cash to be received as a result of the transaction at time T and V_t^T, the price of the contract, is the present value of C_T. We then have

$$dV_t^T/V_t^T - rdt = 0 \tag{2.6}$$

Thus we conclude that in the deterministic case the rate of return for the "true" present value is equal to the interest rate and does not depend on the present value of other economic factors. This is to be expected for otherwise an arbitrage possibility [17] tending to readjust prices to their "true" value levels would exist. Next, we want to find a stochastic generalization to (2.4) and (2.6) for the case where future cash flows are not known with certainty.

2.2. The Random Walk Theorem

It is widely assumed [39] that the market capitalizes a stock at the mathematical expectation of its present value. That is, if we model the information at each time as an increasing family of σ-algebras $\{\mathcal{F}_t\}$ then the "true" or "intrinsic" price of a stock is the conditional expectation of the sum of its discounted future dividend payments given by

$$P_t = \varepsilon\{\int_t^\infty e^{-r(s-t)} \, dD_s | \mathcal{F}_t\} \tag{2.7}$$

where the discount rate r is the sum of the (risk-free) interest rate and a positive term due to security risk, and the dividend payments D_t are assumed to be $\mathcal{F}_t$ (local) semi=martinglaes so that the integral (2.7) can be defined.

In a similar manner the "true" futures price for a share or a commodity is given by:

$$P_t^T = \varepsilon\{e^{-r(T-t)} P_T | \mathcal{F}_t\} \tag{2.8}$$

where P_T is the price at time T.

For security prices the increasing property of the $\{\mathcal{F}_t\}$ implies

$$P_t^T = \varepsilon\{\int_T^\infty e^{-r(s-t)} dD_s | \mathcal{F}_t\} \tag{2.9}$$

We shall shortly develop a model for the evolution of the information σ-algebras $\{\mathcal{F}_t\}$. However, even without a dynamical model for $\mathcal{F}_t$ we can derive a version of random walk model [20] for the "true" prices by simply differentiating (2.7), (2.8), and (2.9), and using the increasing property of the $\{\mathcal{F}_t\}$ and the identity

$$\varepsilon\{dX_t | \mathcal{A}_t\} = \varepsilon\{d\hat{X}_t | \mathcal{A}_t\} \qquad (2.10)$$

where $\hat{X}_t = \varepsilon\{X_t | \mathcal{A}_t\}$ and $\mathcal{A}_t$ is the appropriate family of σ-algebras. In so doing, we obtain

$$\varepsilon\{d_t P_t + d_t D_t | \mathcal{F}_t\} = rP_t dt \qquad (2.11)$$

$$\varepsilon\{d_t P_t^T | \mathcal{F}_t\} = rP_t^T dt \qquad (2.12)$$

$$\varepsilon\{d_T P_t^T + e^{-r(T-t)} dD_T | \mathcal{F}_t\} = 0 \qquad (2.13)$$

where d_h denotes differentiation with respect to h. Therefore, we conclude that the true prices suitably discounted, normalized, and corrected for the dividend payments (rates of "excess" returns) are uncorrelated with the past information in the sense of martingales. That is, we have

$$(dP_t + dD_t)/P_t - rdt = dM_t \qquad (2.14)$$

$$dP_t^T/P_t^T - rdt = dN_t \qquad (2.15)$$

where M and N are $\mathcal{F}_t$ martingales.

Thus the fundamentalist's interpretation of the random walk hypothesis is that all the information is being discounted in the present "true price" and has no value in estimating the excess returns due to future price changes, for we have

$$\varepsilon\{dM_t | \mathcal{F}_t\} = 0 \qquad (2.16)$$

$$\varepsilon\{dN_t | \mathcal{F}_t\} = 0 \qquad (2.17)$$

Neither can one profit by buying futures contract with the right delivery date, for (2.13) shows that the change in the price of the futures due to a change in the closing data is equal to the discounted present value of the dividends that are expected to be paid at that time.

The following dilemma has been posed by Granger in [19]. Suppose that everyone believes in the fundamentalist's pricing mechanism and acts accordingly. But the actual market price P_{mt} is only an estimate of the "true value." If P_{mt} satisfies the same dynamics as the fundamentalist's intrinsic value, then for positive discount rate $r > 0$ the error θ_t in (2.11) and (2.12) will satisfy the equation [26]

$$d\theta_t = r\theta_t \, dt + d\xi_t \tag{2.18}$$

where ξ_t is an $\mathcal{F}_t$ martingale with $\varepsilon\{d\xi_t | \mathcal{F}_t\} = 0$. Equation (2.18) is an unstable equation and any initial error causes $|\theta_t| \underset{t\to\infty}{\to} \infty$.

A possible resolution of the dilemma is as follows. Since there is a nonzero error the market will attempt to not only predict the future dividend payments, but also estimate the probability law at each time. The question of achieving an asymptotically stable error when the system has to be continuously identified has been studied [2], even though the answers are only partially known.

2.3. An Optimal Dividend Policy for Normal Growth Model

It can be argued with some cogency that the investors of a given company share the common goal of maximizing the "true present value" or the "intrinsic price" of the shares [17]. Therefore, the dividend policy should be so chosen as to maximize P_t the current "true price" of the share. Here, we examine such a policy for the normal growth model.

Let X_t denote the book value at time t. Then

$$dX_t + dD_t = (\text{earnings in } dt) \tag{2.19}$$

and let the model for earnings be

$$(\text{earnings in } dt) = X_t[adt+bd\eta_t] \tag{2.20}$$

where η_t is a standard Wiener process. Denoting the dividend rate as dD_t and assuming that it is measurable with respect to $\mathcal{F}_t$, we get:

$$dX_t + dD_t = X_t[adt+bd\eta_t] \tag{2.21}$$

Clearly, a is the rate of return on invested capital while b is the level of risk.

Suppose the firm is scheduled to be liquidated at T. Then the present value might reasonably be taken to be the expected cumulative dividend payments plus the liquidation value both discounted to the present i.e.

$$P_t = \varepsilon[\int_t^T e^{-r(s-t)}dD_s + e^{-(T-t)}X(T)|\mathcal{F}_t]. \tag{2.22}$$

Using (2.21) and (2.22) we get

$$P_t = X_t + (a-r)\ \varepsilon[\int_0^{T-t} e^{-rs}X_{t+s}ds|\mathcal{F}_t] \tag{2.23}$$

Observe:

(a) If $a = r$ then $P_t = X_t$ and the dividend policy is a matter of indifference.

(b) If $a < r$ then $P_t \lneqq X_t$. The optimal policy, if it is a feasible one, is to liquidate at once.

(c) If $a > r$ write (2.21) using the Ito lemma as:

$$d\,\ell n\,X(t) = \frac{1}{X(t)}\,dX(t) - \frac{1}{2}\,\frac{b^2X^2(t)}{X^2(t)}\,dt$$

$$= a\,dt + bd\eta(t) - \frac{1}{2}\,b^2dt - \frac{dD_t}{x(t)} \tag{2.24}$$

so that

$$\ell n\,X(t+s) - \ell n\,X(t) = as + b[\eta(t+s)-\eta(t)] - \frac{1}{2}\,b^2s$$

$$- \int_t^{t+s} \frac{dD_\tau}{X_\tau} \tag{2.25}$$

or

$$X(t+s) \leq X(t)e^{as+b[\eta(t+s)-\eta(t)]-\frac{1}{2}b^2s} \tag{2.26}$$

It follows that

$$E[X(t+s)|\mathcal{F}_t] \leq X(t)e^{as} \tag{2.27}$$

and from (2.23) one could write for a constant T

$$P(t) \leq X(t)e^{(a-r)(T-t)} \tag{2.28}$$

with equality attained by setting $dD(s) = 0$, $t \leq s \leq T$, i.e., the optimal policy is to pay no dividend.

In short, the optimal policy for the normal growth model is a rather obvious one, viz., for $a > r$ pay no dividend, for $a < r$ liquidate at once, for $a = r$ it matters not what dividend policy is pursued.

That the dividend policy is a matter of indifference with respect to the "true" price in all cases when there exists a risk-free opportunity for lending or borrowing with an interesting rate equal to r has already been discussed by Miller and Modigliani [17], because if the optimal dividend policy is D_t^* one can always pay that amount by combining any other dividend policy D_t with the difference $I_t = D_t^* - D_t$ borrowed at an interest equal to the discount rate. The present value of the stream I_t is equal to the difference of the "true" prices corresponding to D_t^* and D_t respectively.

$$\varepsilon\{\int_t^\infty e^{-r(s-t)} dI_S | \mathcal{F}_t\} = \varepsilon\{\int_t^\infty e^{-r(s-t)} dD_s^* | \mathcal{F}_t\} - \varepsilon\{\int_t^\infty e^{-r(s-t)} dD_s | \mathcal{F}_t\}$$

$$= P_t^* - P_t \tag{2.29}$$

Even when one cannot borrow or lend money in the market, companies can achieve an equivalent financing by issuing new shares and selling them at their "true" present price.

Chapter 3

A Quantitative Theory of the Fundamentalist's Price

3.1. Price as a Resolvent Function

Let us now assume that the information $\mathcal{F}_t$ is generated by a vector valued Markov process X_t which might be interpreted as the state of the company at time t. Let the expected dividend payments at time interval (t,t+dt) be given by $f(x_t)$. Let us also assume for simplicity that the process X_t is time-homogenuous i.e. if we define the operator $\mathcal{H}_t^s$ by

$$\mathcal{H}_t^s f(x) = \varepsilon\{f(x_S)|X_t = x\} \tag{3.1}$$

for bounded and continuous f then $\mathcal{H}_t^s$ will only depend on s-t (and we shall hereafter denote it is $\mathcal{H}_{s-t}$). Extension to arbitrary Markov processes, can easily be accomplished by argumenting the state variable with the time variable. Now since we have [49]

$$\|\mathcal{H}_t f\| \leqq \|f\| \tag{3.2}$$

$\mathcal{H}_t$ is a contraction mapping and is therefore continuous.

Also because of the increasing property of $\{\mathcal{J}_t\}$ we have the semigroup property:

$$\mathcal{H}_{t+s} = \mathcal{H}_t \mathcal{H}_s \tag{3.3}$$

And, from definition

$$\lim_{t \downarrow 0} \mathcal{H}_t f = f \tag{3.4}$$

(3.3) and (3.4) imply that $\mathcal{H}_t$ must be of the form

$$\mathcal{H}_t = \exp(t\mathcal{L}) \tag{3.5}$$

where the differential generator $\mathcal{L}$ is given by

$$\mathcal{L}f = \lim_{t \downarrow 0} \frac{1}{t} (\mathcal{H}_t f - f) \tag{3.6}$$

wherever such limit exists and $\mathcal{L}$ is bounded.

Since $\mathcal{H}_t$ is the operator for the expected future value of a function the interpretation for the operator $\mathcal{L}$ is the expected rate of change of value of that function.

Now (3.7) becomes

$$F_r(x) = \int_0 e^{-rs} \mathcal{H}_s f(x) \, ds \tag{3.7}$$

where $P_t = F_r(x_t)$ is the "true" price corresponding to the discount rate r. Thus the "true price" operator is the well known resolvent of the semigroup $\mathcal{H}_.$ i.e. it is the Laplace transform of $\mathcal{H}_.$ operating on the expected incremental dividend payment policy f.

$$F_r(x) = \mathcal{R}_r f(x) = \left(\int_0^\infty e^{-rs} \mathcal{H}_s ds\right) \; f(x) \tag{3.8}$$

It can be proved [49] that $\mathcal{L} F_r$ exists and thus F_r is the unique solution to the following equation:

$$rF_r(x) - \mathcal{L}F_r(x) = f(x) \tag{3.9}$$

which is equivalent of the random walk Eq. (2.11). It leads to the interesting observation that the dividend payments should be equal to the interest minus the expected rate of change for the "true price."

Example 3.1.1.

As mentioned earlier the decision criteria on the dividend policy of a company often is to maximize the current "true price" of its shares. We assume that the choice of the dividend policy will affect the dynamics of the information process. Let us denote the differential generator corresponding to dividend policy $f(\cdot)$ by $\mathcal{L}_f$ and let the discount rate be $\xi(\cdot)$. The optimal dividend policy will then satisfy the following optimality condition.

$$\sup_f\{\mathcal{L}_f F - \xi F + f\} = 0. \tag{3.10}$$

for some function F, which will be the optimal "true price" function. The optimization is to be carried out for all $f(\cdot)$ satisfying the constraints of the problem.

Example 3.1.2.

Let X_t be a diffusion process satisfying the following stochastic differential equation

$$dX_t = m(X_t)dt + N(X_t)\, d\eta_t \tag{3.11}$$

where η_t is an $\mathcal{F}_t$ measurable Wiener-Bachelier process (vector of Brownian Motions). The operator $\mathcal{L}$ in this case will coincide with the differential operator given by

$$\mathcal{L}g(x) = m(x)\ \frac{d}{dx}\, g(x) + \frac{1}{2}\, tr[N'(x)\, \frac{d^2}{dx^2}\, g(x)\ N(x)] \tag{3.12}$$

which means the capital change in the "true price" is the sum of a term due to trends in the information process and a risk-related correction term implying that there is a cost to the uncertainties about the future states that should be compansated by either a dividend payments or an expected change in the future "true price."

Solving (3.9) we can derive the formula for the "true" price explicitely, as a function of the state, and although knowledge of the present state can not help to estimate future excess returns the quadratic variation will be given by

$$d\langle P,P\rangle_t = [\frac{d}{dx} F(x)]' N(x)N'(x)[\frac{d}{dx} F(x)]dt\Big|_{x=X_t} \tag{3.13}$$

so that unless this function is independent of x (which e.g. is the case for the lognormal model) the present state does help in estimating the risk level.

Equations similar to (3.9) can be derived for futures prices. Thus we have

$$\frac{\partial}{\partial h} F_r^h(x) = \mathcal{L}F_r^h(x) - rF_t^h(x) \tag{3.14}$$

$$F_r^0(x) = F(x) \tag{3.15}$$

where $F_r^h(x)$ is the futures price given by

$$F_r^h(x) = \varepsilon\{e^{-rh}F(X_h)|X_0 = x\} \tag{3.16}$$

and the interpretation is as one approaches the closing time the futures prices will change because of the changes in the information process and because there is less interest to be discounted and at the end it should be equal to the "true" spot price. In operator rotation we have

$$F_r^h(x) = e^{-(rI-\mathcal{L})h}F(x) = e^{-(rI-\mathcal{L})h}(rI-\mathcal{L})^{-1}f(x) \tag{3.17}$$

where f is the dividend policy, and I is the identity operator.

3.2. Derivation of a Kernel Function

Now suppose that the information process has a transition function given by

$$P_t(x,A) = Pr[X_t \in A | X_0 = x] \tag{3.18}$$

Since we have

$$\mathcal{H}_t f(x) = \int P_t(x,dy)\ f(y) \tag{3.19}$$

we can rewrite (3.7) as

$$F_r(x) = \int f(y)\ K_r(x,dy) \tag{3.20}$$

where

$$K_r(x,dy) = \int_0^\infty e^{-rt} P_t(x,dy) \tag{3.21}$$

i.e. $K_r(x,dy)$ is the price per each dollar divident payment policy for the neighbourhood of the state y when the economy is at state x. Therefore, the "true price" is the potential corresponding to the dividend policy distribution (in the state space) $f(\cdot)$ with the Kernel $K_r(\cdot,\cdot)$ being the Laplace transform of the transition function which has obvious economic interpretation. Interesting problems arise when the process X_t is defined only in a region G with smooth boundary . Different sets of boundary conditions for (3.9) correspond to different behavior of the process on the boundary. In general the only possible types of behavior are stopping, disappearance, reflection, diffusion along the boundary, and their combination which means linear combination of the corresponding boundary conditions. There is no general theory for the case of random domains. (See e.g. [42] for a discussion of boundary conditions. Also

[14] gives the solutions for the one-dimensional diffusion processes).

Example 3.2.1. Consider the n-dimensional Wiener process which transition probability given by

$$P_t(x,\Gamma) = \int_\Gamma \frac{1}{(\sqrt{2\pi t})^n} \exp(- \frac{1}{2t} \|x-y\|^2) \, dy \tag{3.22}$$

The differential generator for this process is such to be the Laplace operator [14] given by

$$\mathcal{L}f(x) = \frac{1}{2} \nabla^2 f(x) = \frac{1}{2} \sum_{i=1}^{n} \frac{\partial^2 f}{\partial x_i^2} \tag{3.23}$$

for n > 2 the corresponding Green's function is given by

$$K_o(x,dy) = \frac{1}{(2\pi)^{n/2}} \Gamma(\frac{n}{2} - 1) \; \|y-x\|^{2-n} \, dy \tag{3.24}$$

<u>Example 3.2.2.</u>

Now let X_t be a one-dimensional Brownian motion in $[0,\infty)$. The corresponding Kernel will be of the form

$$K_r(x,dy) = \{A(y) \exp(-\sqrt{2r}\, x) \, 1(x-y) + [B(y) \exp(-\sqrt{2r}\, x) + c(y) \exp(\sqrt{2r}\, x)]$$

$$\cdot 1(y-x) + D(y) \exp(-\sqrt{2r}\, x) \; \delta(y)\} \, dy \tag{3.25}$$

where $1(\cdot)$ and $\delta(\cdot)$ are the unit step function and Dirac δ-function respectively A, B, C, D depend on the behavior of the process at the left boundary.

Finally, for the future prices given by (3.16) we have

$$F_r^h(x) = \int e^{-rh} P_h(x,dy) \; F(y) \tag{3.26}$$

or if $F(\cdot)$ is the "true" price given by (3.7)

$$F_r^h(x) = \int K_r^h(x,dy)\ f(y) \tag{3.27}$$

where K_r^h is given by:

$$K_r^h(x,\Gamma) = \int_0^\infty e^{-r(s+h)} P_{s+h}(x,\Gamma) \tag{3.28}$$

yielding the same potential interpretation as stated before.

3.3. Price as a Function of the Discount Rate

The "true price" given by (2.7) and derived in (3.7) for the Markovian state proves case is a function of the capitalization rate r. To see the dependence we notice that the prices for different discount rates satisfy the following resolvent equation [9,14]:

$$\mathcal{R}_\alpha f - \mathcal{R}_\beta f + (\alpha-\beta)\,\mathcal{R}_\alpha \mathcal{R}_\beta f = 0 \tag{3.29}$$

which when coupled with the fact that because of limited liability prices are nonnegative tells us that prices are decreasing for increasing discount rates. This is consistent with the observation that during inflation times prices go down. Moreover, (3.29) explicitely gives us the change in prices due to a change in the discount rate. In operator notation we have

$$\frac{d}{dr}\mathcal{R}_r = -\mathcal{R}_r^2 \tag{3.30}$$

or using (3.21) we have

$$\frac{d}{dr} F_r(x) = -\int K_r(x,dy)\; F_r(y) \tag{3.31}$$

This observation gives us an often useful method for evoluating a project with its expected future cash flow given by a function $f(X_t)$. The rule of thumb is to look for this so called internal rate of return, which is the rate for which we have

$$\mathcal{R}_r f - I = 0 \tag{3.32}$$

where I is the required investment.

The investment should be chosen if the internal rate of return is higher than the interest rate prevailing in the market. However, for the case where f is not always nonnegative (e.g. future investments are required under certain conditions) there may not be a unique internal rate of return, or if one wants to compare two different investment opportunities the one with the highest internal rate of return is not necessarily the best opportunity. A modified rule is to use the following algorithm in order to choose from among different investment opportunities [17]:

(i) Find the internal rate of return for each candidate. Reject those with internal rates of return less than the market rate r.

(ii) Choose the alternative with the highest internal rate of return provisionally as the defender.

(iii) Take the alternative with the second (if any) highest internal rate of return as challenger. Compute the rate of return of the challenger over the defender; i.e. the rate of return on the difference of the two. If this is greater than r accept the challenger, otherwise, accept the defender. Repeat this step for all the alternatives.

This algorithm will work if the internal rate of returns computed at each stop is unique.

3.4. Stochastic Rate for Discounting

The constant discounting model given by Eqs. (2.7) and (2.8) can be criticized on grounds that the observable interest rate varies with time. Furthermore, it can be argued that future rates for discounting are random because they cannot be known given the current information. A more general model therefore, is to assume an instantanuous rate for discounting to be a stochastic process r_t measurable with respect to $\mathcal{F}_t$. The exponential form of discounting is because of the requirement that it be consistent. That is, if the "present value" of a cash flow P_t at time XT is

$$D_t^T = \varepsilon\{k_t^T.P_T|\mathcal{F}_t\} \tag{3.33}$$

then one should be able to arrive at this value by first considering the "true" value of P_T discounted to some time s (t<s<T) and then discounting it again to the present time t.

$$p_t^T = \varepsilon\{k_t^s\ \varepsilon\{k_s^T P_T|\mathcal{F}_s\}|\mathcal{F}_t\} \tag{3.34}$$

Using the increasing property of $\mathcal{F}_t$ and comparing (3.33) with (3.34) one can conclude that in order to be consistent the discounting functional k should satisfy the semigroup property given by

$$k_t^T = k_t^s k_s^T \tag{3.35}$$

Therefore, under some regularity conditions the infinitesimal discounting functional k_t^{t+dt} will be given by

$$k_t^{t+dt} = \exp(r_t dt) \tag{3.36}$$

for some process r_t, which will be called the rate of discounting at time t.

Equations (2.7), (2.8), and (2.9) thus become

$$P_t = \varepsilon\{\int_t^\infty \exp(-\int_t^s r_u du).\ dD_s \mid \mathcal{F}_t\} \tag{3.37}$$

$$P_t^T = \varepsilon\{\exp(-\int_t^T r_s ds)\ P_T \mid \mathcal{F}_t\} \tag{3.38}$$

$$P_t^T = \varepsilon\{\int_T^\infty \exp(-\int_t^s r_u du).dD_s \mid \mathcal{F}_t\} \tag{3.39}$$

differentiating (3.37) and (3.38) we get

$$(d_t P_t + d\ D_t)/P_t - r_t\ dt = dM_t \tag{3.40}$$

$$d_t P_t^T / P_t^T - r_t dt = dN_t \tag{3.41}$$

where M_t and N_t are $\mathcal{F}_t$ martingales with

$$\varepsilon\{dM_t \mid \mathcal{F}_t\} = \varepsilon\{dN_t \mid \mathcal{F}_t\} = 0 \tag{3.42}$$

again indicating that the present information has no value in predicting the excess returns due to future changes in the "true" prices. For the Markovian state-space case we have

$$F(x) = \varepsilon\{\int_0^\infty \exp(-\int_0^t k(X_s)ds)\ f(X_t)dt \mid X_0 = x\} \tag{3.43}$$

where k(x) is the discount rate corresponding to the state x and f and F are the dividend and "true" price functions respectively. Then the celebrated Kac's theorem yields [14].

$$\mathcal{L}F(x) - k(x)\ F(x) = -f(x) \tag{3.44}$$

which has the same interpretation as (3.9) and may be solved explicitely to give the true price function F(x).

Finally, one can again, determine explicitely the change in the "true price" due to any change in the discounting method $k(\cdot)$. Let $\mathcal{R}_{k_i(.)}f$ be the "true" price corresponding to the discount function $k_i(.)$ and dividend function f. One can write

$$\mathcal{R}_{k_1(.)}f - \mathcal{R}_{k_2(.)}f + (k_1-k_2)\mathcal{R}_{k_1(.)}\mathcal{R}_{k_2(.)}f = 0 \tag{3.45}$$

i.e. for a change $\delta k(.)$ of the discount rate we have the price change F given by

$$\delta F(x) = - \delta k(x) \cdot \mathcal{R}_{k(.)}F(x) \tag{3.46}$$

3.5. Random Stopping

Suppose now that the dividend process is only defined in the [o,T] interval, i.e. the share will produce an (uncertain) stream of future dividend payments up to a (random) stopping time T after which a bankrupt, a default, a decision to sell the stock, or some other unaccounted for event would stop the cash flow and yield a final price P_T price, which is a random variable adapted to $\mathcal{F}_T$. Equation (3.37) now becomes

$$P_t = \varepsilon\{\int_t^T \exp(-\int_t^s r_u du) \cdot dD_s + P_T|\mathcal{F}_t\} \tag{3.47}$$

Letting

$$\Pr[t < T < t + dt|\mathcal{F}_t] = \xi_t dt \tag{3.48}$$

where ξ_t is the conditions probability of stopping and is measurable with respect to $\mathcal{F}_t$, we have

$$(dP_t+dD_t)/P_t - r_t dt = \xi_t dt + dM_t \tag{3.49}$$

where M_t is an $\mathcal{F}_t$ martingale.. i.e. the conditional expectation of the excess return is equal to the conditional probability of stopping. The quadratic variation of M_t will in general depend on the evolution of $\mathcal{F}_t$. For the case where $\mathcal{F}_t$ is generated by a Markovian state process X_t let us consider the subprocess obtained by terminating X_t at a random stopping time T. Let $\xi(x)$ be the conditional probability of termination corresponding to the state x i.e. the conditional probability of the event {T > t} is given by [14]

$$\Pr[T > t|\mathcal{F}_t] = \exp(-\int_0^t \xi(X_s)ds) \tag{3.50}$$

Then it can be shown that the differential generator $\tilde{\mathcal{L}}$ corresponding to the "killed" process is given by [14]

$$\tilde{\mathcal{L}}f = \mathcal{L}f - \xi \cdot f \qquad (3.51)$$

where $\mathcal{L}$ is the generator for the original Markov process X_t.

Combining (3.51) with (3.44) one concludes that the "true price" corresponding to the discounting function $k(\cdot)$ is equal to the true price of the same cash flow with no discounting, but with a random termination corresponding to the termination density [14] $k(.)$.

$$\varepsilon\{\int_0^\infty \exp(-\int_0^t k(X_s)ds)\ f(X_t)dt | X_o = x\} = \varepsilon\{\int_0^T f(X_s)ds | X_0 = x\} \qquad (3.52)$$

where T is given by (3.50) with $\xi(\cdot)$ replaced by $k(\cdot)$.

Thus we once more conclude that the conditional expectation of excess returns should compensate for the conditional probability of termination.

It is often the case that the process X_t is confined to a region $\mathcal{X}$ with a smooth boundary $\partial\mathcal{X}$. Then, in order to know the process we should know its differential generator and in particular, the way it behaves on the boundary. The possibilities are stopping, disappearance, reflection, or diffusion along the boundary, as well as their various linear combinations. To each type of behaviour, there corresponds a boundary condition associated with differential equations involving the generator such as (3.9) and (3.44) see [49] for a discussion of boundary conditions and the criterion that determines whether the region is closed or open or whether the boundary is of "regular" or "exit" type. Also it turns out that the potential kernel given by (3.20) and (3.21) has a singular term on the boundary, which is again the Laplace transform of the transition probability in the boundary, and is given by

$$F_r(y) = \int_{\mathcal{X}} f(y)\ K_r(x,dy) + \int_{\partial\mathcal{X}} f(y)\ G_r(x,dy) \tag{3.53}$$

$$K_r(x,\Gamma) = \int_0 e^{-rt} P_t(x,\Gamma) \tag{3.54}$$

$$G_r(x,\Delta) = \int_0^{\infty} e^{-rt} P_t(n,\Delta) \tag{3.55}$$

where the singular term arises because of stopping or diffusion along the boundary. In cases where the process terminates when it hits the boundary $\partial\mathcal{X}$ the differential generator is again given by (3.51) where $\xi(x)$ is the density associated with the hitting time T and is given by

$$\Pr[X_s \in \dot{\mathcal{X}},\ \ 0 \leqq s \leqq t|\mathcal{F}_t] = \exp(-\int_0^t \xi(x_s)ds) \tag{3.56}$$

where $\dot{\mathcal{X}}$ is the interior of $\mathcal{X}$.

A general theory of Markov processes with random domains has not yet been constructed [14]. However, the theory of additive functional provides us with the means of interpreting the above observations and find its link to the (generalized) Brownian Motion.

<u>Example 3.5.1</u>. Let X_t satisfy the linear model given by

$$dX_t = AX_t dt + Bd\eta_t \tag{3.57}$$

where A and B are nxn matrices and η_t is an n-dimensional Wiener process. Let the dividend payments be given by

$$dD_t = CX_t dt \tag{3.58}$$

where C is an n-dimensional row vector. Considering the limited liability of prices we have

$$V(x) = \varepsilon\{\int_0^T e^{-rs} dD_s|X_0 = x\} \tag{3.59}$$

where T is the first time X = 0 (i.e. $\mathcal{X} = \mathbb{R}^n_+$ and the boundary behavior is stopping) and [49]

$$V(0) = 0. \tag{3.60}$$

Solving the "true" price Eq. (3.9)

$$(Ax)' \frac{dV}{dx} + \frac{1}{2} \operatorname{tr} B' \frac{d^2V}{dx^2} B - rV = Cx \tag{3.61}$$

subject to the boundary condition (3.60), we get

$$V(x) = C(rI-A)^{-1}x \tag{3.62}$$

i.e. the "true" price will be given by

$$P_t = C(rI-A)^{-1}X_t \tag{3.63}$$

which is similar to what Granger [19] obtained for the discrete time model.

For future contracts given by (2.5) the dynamics of the process given by (2.6) will not be changed if the expiration T is a random stopping time. Thus for simplicity let us consider the case with Markov state space where the stopping time is the first occurance of an event and there is no discounting i.e.

$$\phi(x) = \varepsilon\{F(X_T)|X_0 = x\} \tag{3.64}$$

where T is the first exit time of G then we have [14]

$$\mathcal{L}\phi(x) = 0 \quad \text{for } x \in G \tag{3.65}$$

$$\phi(x) = F(x) \text{ for } x \in \partial G \tag{3.66}$$

which is called the Dirichlet Problem).

Yet another way to apply the theory of the terminated subprocess to the case of "true" futures prices is that of the contingent contracts i.e. contracts that should be transacted upon at a certain time conditional on the non-occur nce of an event, i.e.

$$F_h(x) = \varepsilon\{F(X_h)\,\mathrm{Ind}(T>h)\,|\,X_0=x\} \tag{3.67}$$

where Ind(·) is the indicator function and T is the first exit time of G. Again we have

$$\frac{\partial}{\partial h} F_h(x) = \mathcal{L} F_h(x) \qquad \text{for } x \in G \tag{3.68}$$

$$F_0(x) = F(x) \tag{3.69}$$

$$F_h(x) = 0 \qquad \text{for } x \in \partial G \tag{3.70}$$

Other boundary conditions will correspond to the various agreements (other than non-transacting) in the case when the boundary of G has been reached (the "taboo" event has happened) before the contract date.

3.6. The Effect of Time-Change and Changes of Probability Measure or Price

We have already seen that a random termination of the dividend process has the same effect as higher discount rate would have on the "true" prices, and for Markov processes, its effect can be summerized in the transformed differential generator given by (3.51). Another possible transformation is a random change in the time variable. It can be shown [44] that martingales are Brownian Motions whose time has been changed randomly with the quadratic variation of the processes. In other words let M_t be an integrable local martingale and $\langle M,M\rangle_t$ be its quadratic variation let B_t be an independent Brownian Motion and let us define a random time change given by the below equation.

$$T_t = \inf\{s \mid \langle M,M\rangle_s > t\} \text{ if this is finite} \tag{3.71}$$

$$T_t = \infty \quad \text{otherwise}$$

and let X_t be defined by

$$X_t = M_{T_t} \quad \text{if } T_t < \infty \tag{3.72}$$

$$X_t = M_\infty + B_{t-\langle M,M\rangle_\infty} \quad \text{otherwise}$$

then X_t is a Brownian Motion [44].

This observation e.g. might explain why when the price fluctuations almost seem to be a Brownian Motion and one changes the time dimension into a "move" dimension [9] (i.e. each new time unit is when the prices have changed % x) then the picture changes and the new process is no longer a Brownian Motion. When we have a Markovian State process a random time change characterized by

$$\int_0^{T_t} V(X_s)ds = t \tag{3.73}$$

where V(x) is a positive function, will transform the differential generator [14] according to

$$\tilde{\mathcal{L}}f(x) = \frac{1}{V(x)} \mathcal{L}f(x) \qquad 3.74)$$

A Markov process is called a generalized Brownian Motion [14] if it can be obtained from the Weiner process by a construction of a subprocess and a random time change. The construction of the subprocess, or random killing as we have seen is associated with an additive function ϕ_{su} such that if T is the termination time then conditional probability of stopping is given by

$$\Pr[T > t|\mathcal{F}_t'] = \exp(-\phi_{0t}) \tag{3.75}$$

And the random time change corresponds to another additive function ψ_{su} such that if τ_t is the new time we have [14]

$$\psi_{0\tau_t} = t \tag{3.76}$$

Since we know that each additive function corresponds to a measure such that

$$\phi_{st} = \int_s^t \frac{d\mu}{d\ell} (X_u)\ du \tag{3.77}$$

$$\psi_{st} = \int_s^t \frac{d\nu}{d\ell} (X_u)\ du \tag{3.78}$$

where $\frac{d}{d\ell}$ denotes density with respect to the Lebesgue measure. Each generalized Brownian Motion is thus characterized by two measures; the "killing measure" μ determines when the process is terminated and the "speed measure" ν determines the random time change. The differential generator of a generalized Brownian Motion is then given by

$$f = \frac{1}{2}\,(1/\,\frac{d\nu}{d\ell})\nabla^2 f - \frac{d\mu}{d\nu}\,f \tag{3.79}$$

where ∇^2 is the Laplacian operator. Finally, let us consider a change in state space. For simplicity let us consider the case where the differential generator is identical to the second order operator given by

$$\mathcal{L}f = \sum_{i,j} a_{ij}(x)\,\frac{\partial^2}{\partial x_i \partial x_j}\,f + \sum_i b_i(x)\,\frac{\partial}{\partial x_i}\,f - c(x)f \tag{3.80}$$

Now let us change the state x into another state $\tilde{x}$ where the transformation to the new coordinate system is twice continuously differentiable then the new differential generator is given by [14]

$$\mathcal{L}f = \sum_{i,j} \tilde{a}_{ij}(x)\,\frac{\partial^2}{\partial x_i \partial x_j}\,f + \sum_i \tilde{b}_i(x)\,\frac{\partial}{\partial x_i}\,f - \tilde{c}(x)\,f \tag{3.81}$$

where we have

$$\tilde{a}_{ij} = \sum_{\alpha,\beta} a_{\alpha\beta}\,\frac{\partial \tilde{x}_i}{\partial x_\alpha}\,\frac{\partial \tilde{x}_i}{\partial x_\beta} \tag{3.82}$$

$$\tilde{b}_i = \sum_\alpha b_\alpha\,\frac{\partial \tilde{x}_i}{\partial x_\alpha} + \frac{1}{2}\sum_{\alpha,\beta} a_{\alpha\beta}\,\frac{\partial^2 \tilde{x}_i}{\partial x_\alpha \partial x_\beta} \tag{3.83}$$

$$\tilde{c} = c \tag{3.84}$$

Let us consider the economic interpretation of the transformation mentioned above. Regarding (3.51) we observe that there is no difference in the "true price" of a terminated process with that of a non-terminated process that has been discounted by an amount equal to the conditional probability of termination, because if one considers (3.44) then one can conclude that they both satisfy the same differential equation. Therefore, we conclude that one discounts future streams of income in anticipation of a termination in the stream due to unexpected events that have not been incorporated. On the other hand, a random time change will change the quadratic variation of the process and unless one is indifferent towards risk, there should be a reverse relationship between a "speed" rate and the discount rate. But, (3.79) does not give an indication of what the relation should be. Our conjecture is that since under very general conditions [14] Markov processes can be obtained from the Wiener process by successively carrying out a random termination, a random time change, and a transformation of the coordinates, and since the term due to termination in the differential generator is invariant under the transformation of coordinates (Eq. (3.84)), the discount rate from (3.79) should be the density of the killing measure of the process with respect to the speed measure. In general the discount increases with increasing termination probability and with increasing quadratic variation, implying risk aversion on the part of participants in the market.

Chapter 4

A Markov Theory of Options

Let us consider markets where by a proper composition of his portfolio one can eliminate all the uncertainties concerning the immediate future return of his investment. Mathematically, this means that one can find such a combination of investment opportunities such that the total incremental return on his wealth is measurable with respect to the information that he has at the time. For example, let us consider the case where the information is generated by an N dimensional Markov process. Now consider N+1 different assets, none of which can be obtained from a combination of the other N. Let the return (adjusted for dividend payments if necessary) of each of these assets be given by

$$dR_i(X_t) = \left(\sum_{k=1}^{N} \frac{\partial R_i}{\partial x_k}\right)_t dX_{kt} + \sum_{k=1}^{N} \sum_{\ell=1}^{N} \left(\frac{\partial^2 R_i}{\partial x_k \partial x_\ell}\right)_t d\langle X_k, X_\ell\rangle_t \qquad (4.1)$$

where i stands for the asset i and X_{kt} is the k-th component of X_t, and

$$\left(\frac{\partial R_i}{\partial x_k}\right)_t = \left.\frac{\partial R_i}{\partial x_k}\right|_{x=X_t}$$

$$\left(\frac{\partial^2 R_i}{\partial x_k \partial x_\ell}\right)_t = \left.\frac{\partial^2 R_i}{\partial x_k \partial x_\ell}\right|_{x=X_t}$$

The second term in (4.1) is $\mathcal{F}_t$ measurable, and one can find a combination of the N assets with α_i being the preportion for the asset i such that the uncertainty is eliminated

$$\sum_{i=1}^{N+1} \alpha_i dR_i = \sum_{i=1}^{N} \sum_{k=1}^{N} \alpha_i \frac{\partial R_i}{\partial x_k} dx_k + \sum_{i=1}^{N+1} \alpha_i dC_i(x\ ) \qquad (4.2)$$

where $C_i(X_t)$ is the correction term and is given by (4.1). This can be done if one choses the α_i such that

$$\sum_{i=1}^{N\ 1} \alpha_i \frac{\partial R_i}{\partial n_k} = 0 \quad \text{for } k = 1,2,\ldots,N \qquad (4.3)$$

If now a new investment opportunity is introduced in the market, one can again combine this with N of the existing assets in such a way that the risk is eliminated. Unless the resulting risk-free rates of returns turn out to be equal, there is a clear arbitrage possibility and one could make infinite amount of profit by selling short the resulting portfolio with the lower rate of return and buying long the one with the higher return rate. The riskless opportunity for profit will force the rates to resettle into an equal risk-free rate. Therefore, the return on the new investment will be determined by the rates of returns for the N+1 assets.

Example 4.1.1.

Let the information process be a one-dimensional diffusion process given by

$$dX_t = m(X_t)\,dt + \sigma(X_t)dB_t \tag{4.4}$$

where B_t is a one-dimensional Wiener process. Let us assume that there is a market portfolio whose rate of return is given by

$$dR_{mt} = \sigma_m(X_t)dt + \beta_m(X_t)dB_t \tag{4.5}$$

Now consider a share with incremental dividend payments f(x) and a "true" price F(x) corresponding to the state x. The dividend-adjusted rate of return is equal to

$$dR_{st} = \frac{\mathcal{L}F(X_t)+f(X_t)}{F(X_t)}\,dt + \frac{F'(X_t)}{F(X_t)}\,\sigma(X_t)dB_t \tag{4.6}$$

Finally, let us assume that there is a risk-free rate of borrowing or lending equal to $r(X_t)$. Let us now consider a combined portfolio composed of $w_1(X_t)$ dollars of share, $w_2(X_t)$ dollars invested on the market portfolio, and w_1+w_2 dollars borrowed with the rate $r(X_t)$ to finance the investment. The total return will be given by

$$dR_t = [w_1(\frac{\mathcal{L}F+f}{F} - r) + w_2(\alpha_m - r)]\ dt + (w_1 \frac{F'\sigma}{F} + w_2\beta)dB_t \tag{4.7}$$

if one choses w_1 and w_2 such that

$$w_2 = \frac{-\sigma F'}{\beta F} w_1 \tag{4.8}$$

then he can get an arbitrarily high return with zero investment and zero risk, a clear arbitrage, unless we have

$$\frac{\mathcal{L}F(x)+f(x)}{F(x)} - r(x) = \frac{\sigma(x)F'(x)}{\beta(x)F(x)} (\alpha_m(x)-r(x)) \tag{4.9}$$

We thus have a rational discounting rate that excludes arbitrage where the rate of excess return on each portfolio is related to that of a market portfolio and the higher the risk level, the more return one should expect in compensation.

Example 4.1.2.

One field of application of the random walk theory of the stock prices is the determination of the "true" prices of such derivative assets as puts, calls, warrants and convertibles. It is usually assumed [30,37] that the underlying security prices are Markov processes and thus the "true" prices of the otpions are functions of security prices. One problem is that since the options are more volatile the discount rate should be higher than that of the unlevered securities. Using the above example, however, one can obtain the discount rate that would exclude arbitrage possibilities when there exists a risk-free rate for lending or borrowing money in the market. Let X_t be the price of the security and let the differential generator $\mathcal{L}$ be given by

$$\mathcal{L} = \frac{1}{2} a^2(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx} - c(x) \tag{4.10}$$

Now let f(x,t) be the "true" price of the option when the security price is x and let r(x) be the risk-free rate corresponding to that price. From (4.9) we have

$$(\mathcal{L}f + \frac{\partial}{\partial t} f)/f - r = (f'/f)(b-r) \tag{4.11}$$

or after rearranging terms we get the following differential equation

$$\frac{1}{2} a^2(x) \frac{\partial^2 f}{dx^2} + r(x) \frac{\partial f}{\partial x} - [r(x)+c(x)]f + \frac{\partial f}{\partial t} = 0 \tag{4.12}$$

subject to the proper initial condition [6,30]. One important result of the above analysis is that the "true" prices for options are independent of the drift factor b(x). Thus two different investors with different beliefs on the rate of return of the security might agree on the "true" option price. Now let us go back to the general case where a portfolio return is given by (4.2) and we have

$$\sum_{i=1}^{N+1} \alpha_i(x) = 1 \tag{4.13}$$

each set of α_i, $i = 1,2,\ldots,N+1$ characterizes a portfolio and the one obtained by solving (4.3) together with (4.13) will be called the risk-free portfolio and its return rate will be denoted as dR_{0t}. By a process similar to the Gram-Schmidt [41] one can construct N different portfolios whose rates of returns, denoted by $d\tilde{R}_k$, $k = 1,2,\ldots,N$, will be uncorrelated in the sense that

$$d\langle \tilde{R}_k, \tilde{R}_\ell \rangle_t = 0 \quad \text{for } k \neq \ell \tag{4.14}$$

To do this let

$$d\tilde{R}_{1t} = dR_{1t}$$

$$d\tilde{R}_{2t} = -\frac{d\langle R_2, \tilde{R}_1 \rangle_t}{d\langle \tilde{R}_1, \tilde{R}_1 \rangle_t} (d\tilde{R}_{1t} - dR_{0t}) + dR_{2t} \tag{4.15}$$

and in general

$$d\tilde{R}_{kt} = dR_{kt} - \sum_{\ell=1}^{k-1} \frac{d\langle R_k, R\ \rangle_t}{d\langle \tilde{R}_\ell, \tilde{R}_\ell \rangle_t} (d\tilde{R}_{\ell t} - dR_{0t}) \tag{4.16}$$

then the obtained portfolios will satisfy (4.14). Now we can prove that any other investment opportunity should be linearly related to these "basis" portfolios so as to exclude an possibility of hedging. Suppose for an opportunity whose rate of return is given by dR_t one can find a portfolio combining the N "basis" and the risk-free portfolios such that one has

$$\sum_{i=0}^{N} \theta_i \frac{\partial R_i}{\partial x_k} = \frac{\partial R}{\partial x_k} \qquad k = 1, 2, \ldots, N \tag{4.17}$$

$$\sum_{i=0}^{N} \theta_i = 1 \tag{4.18}$$

where θ_i is the proportion invested in the i-th portfolio. Clearly, one can sell short the combined portfolio and invest on the new opportunity with the proceeds with a risk-free rate of profit equal to

$$W_0 (dR - \sum_{i=0}^{N} \theta_i d\tilde{R}_i) \tag{4.19}$$

unless one has

$$dR = \sum_{i=0}^{N} \theta_i d\tilde{R}_i \tag{4.20}$$

where the θ_i, $i = 0,\ldots,N$ are solutions to (4.17) and (4.18). To find the solutions, we observe that $dR - \sum_{i=0}^{N} \theta_i d\tilde{R}_i$ should be orthogonal to the dR_i thus:

$$\theta_i = \frac{d\langle R,\tilde{R}_i\rangle_t}{d\langle \tilde{R}_i,\tilde{R}_i\rangle_t} \quad \text{for} \quad i = 1,2,\ldots,N \tag{4.21}$$

and we have

$$dR = \sum_{i=1}^{N} \frac{d\langle R,\tilde{R}_i\rangle_t}{d\langle \tilde{R}_i,\tilde{R}_i\rangle_t} (d\tilde{R}_i - dR_o) + dR_o \tag{4.22}$$

The terms $\frac{d\langle \tilde{R},\tilde{R}_i\rangle_t}{d\langle \tilde{R}_i,\tilde{R}_i\rangle_t}$ in the above equations are the volatilities of the process R relative to the R_i.

Chapter 5

The Dynamic Capital Asset Pricing Model

5.1. Preliminaries

It has long been held that a satisfactory theory for analyzing stock price fluctuations must take into account the interdependence of the future price processes. When more than one investment opportunities are available, a possibility of diversification exists. That is, one can choose a portfolio of different stocks in such a way that when the price of one of the individual stocks is expected to go up the price of another will go down and thus hedge against the risks involved It has also been conjectured that since many people's investment goals are to maximize the expectation of a risk-averse concave increasing utility function in the sense of Von Newmann-Morgenstern, [34] they would prefer a portfolio with more expected rate of return and less risk to one with less expected rate of return and greater risk. To pursue this line of argument one needs a normative model for the degree of a portfolio. Among different such models [29,40] the so-called mean variance analysis, where the variance of the return of an asset is considered as a measure of the risk associated with that asset, has by far been the most popular because it deals with such empirically observable variable, as the first two moments of a random variable rather than Von Newmann-Morgenstern type utility functions and corresponding subjective probability distributions. The capital asset pricing model [23,42], which is based on this assumption has received considerable attention in the financial literature because it can be easily applied to such areas as portfolio selection and capital investment decisions, in addition to providing a theoretical foundation to diagonal market models and theorems on the separation of investment and financing decisions [46].

The mean-variance model, on the other hand, has been criticized on both theoretical and empirical grounds in recent years [7,18,22]. This model is true only if either the return variables are assumed to belong to a two parameter class (e.g., jointly normal) [17], or if the utility function is assumed quadratic. Both assumptions lead to gravely unrealistic implications [16].

Perhaps an even more important shortcoming of this model is its static nature, since the allocation of the resources in time is so important that it is said that the capital theory is the theory of time. There have been some recent attempts to extend the theory for the dynamic case [4.17,31,33]. Again, the known information at each time can be modelled as an increasing family of σ-algebras $\mathcal{F}_t$. Let us assume that there are N opportunities for investment the rate of return for each of which, duely adjusted for dividend payments if necessary, is a process R_{it} adapted to $\mathcal{F}_t$. Thus if at time t the amount m_{it} is invested in the i-th opportunity at time t+dt one should get the amount $m_{it}dR_{it}$ for that investment. Therefore, if θ_{it} is the proportion of the wealth at time t that is being invested in i-th opportunity and if c_t is the proportion of ones wealth that is being consumed at that time, then the wealth equation will be given by

$$dW_t/W_t = \sum_{i=1}^{N} \theta_{it}dR_{it} - c_t dt \tag{5.1}$$

where we have

$$\sum_{i=1}^{N} \theta_{it} = 1 \tag{5.2}$$

and the return processes are assumed to be (local) semi martingales with respect to the information $\mathcal{F}_t$ so that the integrals could be defined.

Given the initial wealth (5.1) could be integrated and we will have

$$W_t = W_o \exp[-\int_0^t c_s ds] \exp[\int_0^t (\sum_{i=1}^{N} \theta_{is} dR_{is} - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \theta_{is}\theta_{js} d\langle R_i, R_j \rangle_s$$

Therefore, the wealth at each time will be determined by the past consumption and the past selection of one's portfolio, and the effects of these can be separated. The consumption process is given by:

$$C_t = W_o \cdot c_t \exp[-\int_0^t c_s ds] \exp[\int_0^t (\sum_{i=1}^{N} \theta_{is} dR_{is} - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \theta_{is}\theta_{js} d\langle R_i, R_j \rangle_s)]$$

We shall assume that the individends in the market have "atomistic" influence in the sense that they do not believe that they can change the laws governing the return processes by changing their consumption and investment policies, in other words, the $\mathcal{T}_t$ measurable processes will not depend on the strategies θ_{it}, c_t.

5.2. Generalized Mean-Variance Models

Let us now consider only the investment decision. It is often the case that consumption is not the concern of the investor. Instead he acts so as to maximize the expected value of some increasing concave utility function for the wealth at each instance.

As we saw in section 1.2, the best strategy for such an investor is to choose among portfolios with the same rate of a-posteriori expected return the one with the least quadratic variation term.

Since the rate of overall return on ones portfolio is given by

$$dR_{pt} = \sum_{i=1}^{N} \theta_{it} dR_{it} \tag{5.5}$$

$$\sum_{i=1}^{N} \theta_{it} = 1 \tag{5.6}$$

one has to minimize the quadratic variation term

$$d\langle R_p, R_p\rangle_t = \sum_{i=1}^{N} \sum_{j=1}^{N} \theta_{it}\theta_{jt}\, d\langle R_i, R_j\rangle_t \tag{5.7}$$

subject to (5.6) and for a given expected rate of return

$$\varepsilon\{dR_{pt} | \mathcal{F}_t\} = \sum_{i=1}^{N} \theta_{it}\, \varepsilon\{dR_{it} | \mathcal{F}_t\} \tag{5.8}$$

Let us denote the portfolio with the minimum quadratic variation with the subscript 0 (i.e. dR_{0t} stands for the rate of return for this portfolio). The corresponding ratios $\{\theta_{01t}\}$ will be obtained by differentiating Eq. (5.7) subject to the condition (5.6). Using Lagrange multiplier technique, one gets the following condition for the rate of return for the minimum risk portfolio

$$d\langle R_0,R_i\rangle_t = d\langle R_0,R_0\rangle_t \tag{5.9}$$

Thus, it has the property that every other portfolio has a volatility equal to that of the minimum-risk portfolio, and furthermore, it does not depend on the individual's information or subjective probability beliefs. If the determinant whose element are the a-posteriori covariances $d\langle R_i,R_j\rangle_t$ has a rank lewer than N, then the minimum-risk portfolio at that time will have zero quadratic variation one can construct an instantaneously risk-free portfolio. However, in general R_0 is too conservative a portfolio, and depending on ones measure of risk aversion as shown in 1.2 one would choose a portfolio with greater risk in return for higher expected value of the conditional rate of return. We shall examine later implications of this model and the volatilty of the measure of risk involved in the investment. An important result mentioned in 1.1 is that the volatility factors do not depend on the information pattern $\mathcal{F}_t$ nor does it change when the subjective beliefs concerning the underlying probabilities are changed as long as the probability measures are absolutely continuous with respect to one another.

Before pursuing the discussion or the results that one can get from this model, let us consider the case where the consumption given by

$$C_t = c_t W_t \tag{5.10}$$

is the main concern of the investor. If we differentiate (5.10) we get

$$dC_t = c_t dW_t + W_t dc_t + d\langle c,W\rangle_t \tag{5.11}$$

The last term is measurable with respect to $\mathcal{F}_t$ the information at time t. However, the second term is not necessarily adapted to $\mathcal{F}_t$, and the possible variations of the consumption, therefore, constitute a further source of risk besides the volatility of the investment returns. One would then expect the investors to diversify their portfolio not only to minimize the variations on their portfolio returns but also so as to hedge against changes in their consumption ratios. In the next section we shall examine a model due to Merton [28,31] concerning the investment and consumption strategies.

5.3. Stochastic Control Model

Let us now consider the model presented in 5.1 in more detail. We assume for simplicity that there are N+1 investment opportunities the last one of which is the minimum-risk asset denoted by the subscript 0. We also assume that the dynamics of the return processes i.e. $\varepsilon\{dR_{it}|\mathcal{F}_t\{$, and $d\langle R_i,R_j\rangle_t$ for i,j = 1,...,N are generated by a Markovian state X_t, the evolution of which does not depend on the investment or consumption strategies of individual investors acting in the market.

Consider an individual whose wealth is given by (5.3)
We assume that the augumented vector process $\binom{X_t}{W_t}$ is also a Markov process. If at time t the individual would invest θ_{it}, i = 1,...,N proportion of his wealth in the i-th opportunity and $\theta_0 = 1 - \sum_{i=1}^{N} \theta_{it}$ in the minimum-risk portfolio, then his wealth equation will be given by:

$$dW_t/W_t = \sum_{i=1}^{N} \theta_{it}(dR_{it}-dR_{0t}) + dR_{0t} - c_t dt \tag{5.12}$$

and the quadratic variation of the wealth process will be:

$$\frac{d\langle W,W\rangle_t}{W_t^2} = \sum_{i=1}^{N}\sum_{j=1}^{N} \theta_{it}\theta_{jt} \; (d\langle R_i,R_j\rangle_t - d\langle R_0,R_0\rangle_t) + d\langle R_0,R_0\rangle_t \tag{5.13}$$

Now if one wants to maximize his expected utility of consumption given by

$$U = \int_0^T A(C_s,s)\, ds + B(W_T,T) \tag{5.14}$$

then one can use the dynamic programming technique by defining

$$V(x,w,t) = \varepsilon\{\int_t^T A(C_s,s)ds + B(W_T,T)\,|X_t = x,\ W_t = w\} \tag{5.15}$$

and for the optimal policy one would have

$$\varepsilon\{dV_t|\mathcal{F}_t\} = \sup\{(\frac{\partial V}{\partial w})_t \ \varepsilon\{dW_t|\mathcal{F}_t\} + \sum_{k=1}^{M} (\frac{\partial V}{\partial x_k})_t \ \varepsilon\{dX_{kt}|\mathcal{F}_t\}$$

$$+ (\frac{\partial V}{\partial t})_t + \frac{1}{2} (\frac{\partial^2 V}{\partial w^2})_t \ d\langle W,W\rangle_t + \frac{1}{2} \sum_{k=1}^{M} (\frac{\partial^2 V}{\partial w \partial n_k})_t \ d\langle W,X_k\rangle_t$$

$$+ \frac{1}{2} \sum_{k=1}^{M} \sum_{\ell=1}^{M} (\frac{\partial^2 V}{\partial x_k \partial x_1})_t \ d\langle X_k,X_t\rangle_t - A(C_t,t)dt\} = 0. \qquad (5.16)$$

After replacing for $\varepsilon\{dW_i|\mathcal{F}_t\}$ and $d\langle w,w\rangle_t$ from Eqs. (5.12) and (5.13) and considering

$$d\langle X_k,W\rangle_t = W_t[\sum_{i=1}^{N} \theta_{it}(d\langle R_i,X_k\rangle_t - d\langle R_0,X_k\rangle_t) + d\langle R_0,X_k\rangle_t]$$

could differentiate (5.16) with respect to the θ, to get the regular solution for the investment strategies:

$$(w^2 \frac{\partial^2 V}{\partial w^2})_t \ [\sum_{j=1}^{N} \theta_{jt}(d\langle R_i,R_j\rangle_t - d\langle R_o,R_o\rangle_t)] + (w \frac{\partial V}{\partial w})_t \ [\varepsilon\{dR_{it}-dR_{0t}|X_t\}]$$

$$+ \sum_{k=1}^{m} (w \frac{\partial^2 V}{\partial x_k \partial x})_t \ [d\langle R_i,X_k\rangle_t - d\langle R_o,X_k\rangle_t] = 0. \qquad (5.18)$$

Or if we define the excess return processes as

$$dR'_{it} = dR_{it} - dR_{0t} \qquad i = 1,\ldots,N \qquad (5.19)$$

from which we obtain

$$d\langle R'_i,R'_j\rangle_t = d\langle R_i,R_j\rangle_t - d\langle R_0,R_0\rangle_t \qquad (5.20)$$

$$d\langle R'_i,X_k\rangle_t = d\langle R_i,X_k\rangle_t - d\langle R_0,X_k\rangle_t \qquad (5.21)$$

we will have the following equation for the optimal investment policy

$$\sum_{j=1}^{N} \theta_{jt} \; d\langle R_i', R_j' \rangle_t = - \frac{(\partial V/\partial w)_t}{(w\partial^2 V/\partial w^2)_t} \; \varepsilon\{dR_{it}' \mid X_t\}$$

$$- \sum_{k=1}^{m} \left(\frac{\partial^2 V/\partial X_k \partial w}{w\partial^2 V/\partial w^2}\right)_t \; d\langle R_i', X_k \rangle_t \qquad (5.22)$$

Now if m << N, (5.22) will give an efficient policy set for the selection of ones portfolio. That is, no matter what an individual's utility function is, his optimal portfolio will be a combination of m+2 "mutual fund" portfolios consisting of the following, minimum-risk portfolio corresponds to the solution for the homogeneous equation, the growth-optimal portfolio corresponds to the first right-hand side term and has the property that almost certainly it would result higher terminal wealth than any other decision rule [22,32], and finally m portfolios corresponding to the m right-hand side terms of (5.22) can be used to hedge against unfavorable changes in investment opportunities. Thus one not only must "smooth" the consumption in the sense of maintaining its level against the changes in one's income, but he also has to smooth the consumption by keeping its variability (risk) at a minimum through time. The function V(w,x) is the conditional expectation of one'sutility and is therefore the implied utility of ones wealth at time t. The factor $-\left(\frac{w\partial^2 V/\partial w^2}{\partial V/\partial w}\right)_t$ is the relative risk aversion factor for the implied utility and we can see from (5.22) that the more risk-averse individuals would choose less risky portfolios. Now if we differentiate (5.16) with respect to the other decision variable, namely the consumption we get another condition for the regular optima

$$\left(\frac{\partial V}{\partial w}\right)_t = A'(C_t, t) \qquad 5.23)$$

where

$$A' = \frac{d}{dC} A(C_t, t) \qquad (5.24)$$

Further differentiation of (2.23) and application of the implicit function theorem yields [28],

$$-\left(\frac{w\partial^2 V/\partial w^2}{\partial V/\partial w}\right)_t = \left(\frac{w\partial C/\partial w}{C}\right)_t \left(-\frac{CA''(C,t)}{A'(C,t)}\right)_t \qquad (5.25)$$

$$\left(\frac{\partial^2 V/\partial x_k \partial w}{\partial^2 V/\partial w^2}\right) = \left(\frac{\partial C/\partial x_k}{\partial C/\partial w}\right)_t \qquad (5.26)$$

It can be seen from (5.25) that whenever A is an increasing concave function representing non-satiable risk-averse utility for consumption, the left-hand side factor, i.e. the relative risk-aversion for the implied utility of wealth, will be positive [28]. One can also see from Eq. (5.23) that V is an increasing function of the wealth (since $\frac{\partial V}{\partial w}$ is always positive).

5.4. Separation Properties and Market Models

As mentioned in the last section (5.22) represents a generalized separation of decisions, which can be as follows: under the stated conditions there would exist m+2 "mutual funds" the proportion of each fund's portfolio invested in the individual N+1 assets are purely "technological" i.e. do not depend on investors' utilities such that no matter what one's consumption policy is he will be indifferent between choosing portfolios from among the original N+1 investment opportunities or a combination of the m+2 "mutual funds," and that the investors' demands on the funds as given by (5.25) and (5.26) depends on their utilities and consumption policies, but require no knowledge of the N+1 investment opportunities or the proportions held by the funds.

Now if the changes in the investment opportunities are "uncorrelated" with their rates of returns, i.e.,

$$d\langle X_k, R_i\rangle_t = 0 \quad i = 1,2,\ldots,N \quad k = 1,\ldots,m \tag{5.22}$$

then there would be no need for the m portfolios used to hedge against unfavorable changes in the investment opportunities and therefore the consumption and investment decisions can be completely separated and no matter what one's consumption policy is his investment portfolios will consist of two "mutual funds." Notice that (5.22) in this case results in the same optimal investment policy that was outlined in section (5.2). One chooses his portfolio so that it has the minimum quadratic variation among the possible portfolios having the same expected rate of return. One has thus to solve the following quadratic programming problem:

$$\text{minimize } \frac{1}{2}\left[\sum_{i=1}^{N}\sum_{j=1}^{N} \theta_{it}\theta_{jt}\, d\langle R_i', R_j'\rangle_t + d\langle R_0, R_0\rangle_t\right] = \frac{1}{2} s_t^2 dt \qquad (5.28)$$

$$\text{subject to: } \sum_{i=1}^{N} \theta_{it}\, \varepsilon\{dR_{it}' | \mathcal{F}_t\} + \varepsilon\{dR_{0t} | \mathcal{F}_t\} = M_t dt \qquad (5.29)$$

(If either the conditional mean or the quadratic variation is not absolutely continuous(with respect to the Lebesgue measure)then the differentiation can be done with respect to some other measure [47]) Using the Lagrange multiplier technique one could see that the optimal policy would be such that one's portfolio will consist of the two "mutual funds" given by (5.22), and the optimal mean and variance will have a hyparbolic trajectory in the M_t-S_t plane [29]. These portfolios are thus the proper generalization of the mean-variance efficient portfolios forthe dynamical case.

It can easily be seen that every combination of two mean-variance efficient portfolios is itself mean-variance efficient and as one's risk-aversion factor changes the efficient frontier is spanned. If we denote the rate of return for any one of the efficient portfolio as dR_{et} it will have the property that the expected rate of return for any of the other N investment opportunities will be linearly related to its "correlation with the efficient portfolio in the following sense

$$\frac{\varepsilon\{dR_{pt}' | \mathcal{F}_t\}}{d\langle R_p', R_e'\rangle_t} = \frac{\varepsilon\{dR_{it}' | \mathcal{F}_t\}}{d\langle R_i', R_e'\rangle_t} \qquad i = 1,2,\ldots,N \qquad (5.30)$$

where we have

$$dR_{pt}' = \sum_{i=1}^{N} \theta_{it}(dR_{it}') \qquad (5.31)$$

$$dR_{et}' = \sum_{i=1}^{N} \theta_{it}^{*} dR_{it}'$$

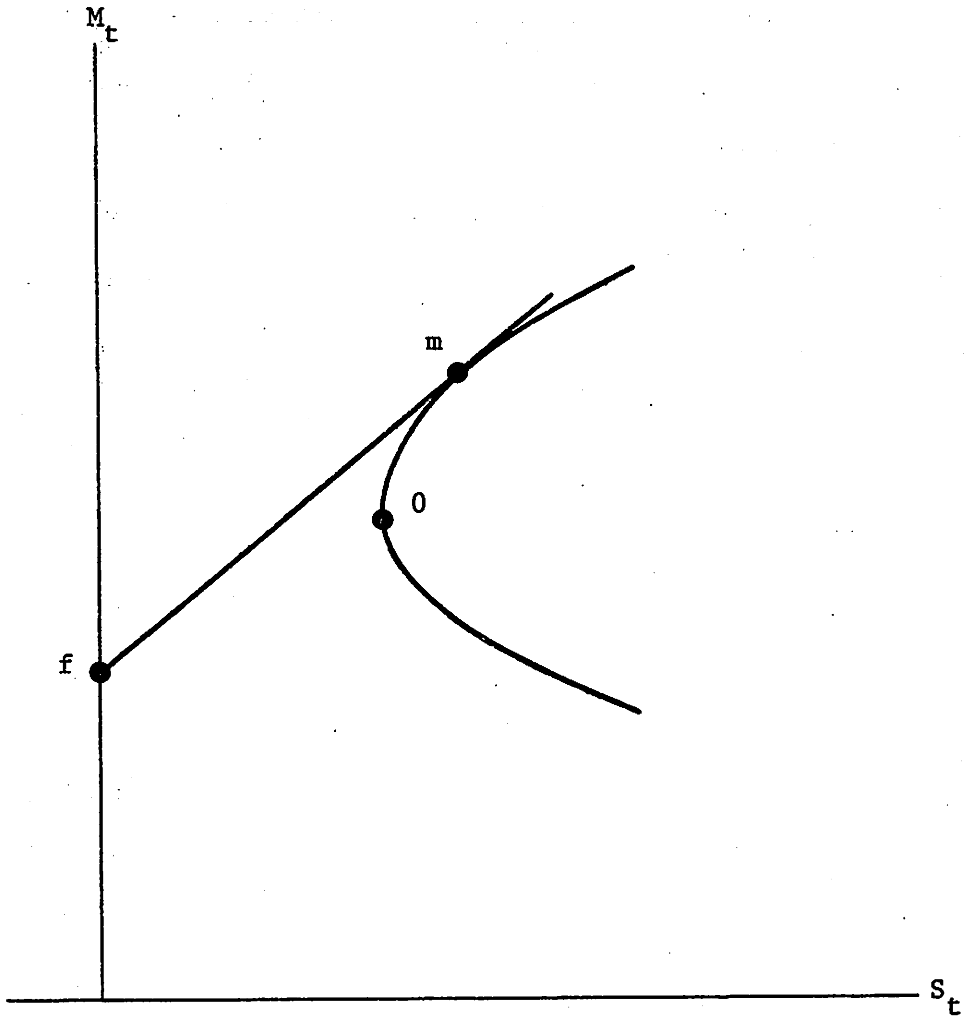

Fig. 5.1.

and * denotes the efficient portfolio. Furthermore, each of the efficient portfolios can be the parameter for a market model. To each one of these portfolios there corresponds another portfolio belonging to the efficient frontier which has an expected rate of return lower than that of the minimum-risk portfolio, and a zero correlation with the efficient portfolio. That is, if we denote the second portfolio by the subscript z we have

$$d\langle R_e, R_z\rangle_t = 0 \tag{5.33}$$

Now every other efficient portfolio is composed of the following combination of the two portfolios

$$dR_t = \beta_t^e \, dR_{et} + (1-\beta_t^e)\, dR_{zt} \tag{5.34}$$

where we have

$$\beta_t^e = \frac{d\langle R, R_e\rangle_t}{d\langle R_e, R_e\rangle_t} \tag{5.35}$$

For the non-efficient portfolios we will have the same model described by (5.34), but the zero-beta portfolio in this case is not one of the efficient frontier. However, for every portfolio in the market the following model will be true:

$$\varepsilon\{dR_{pt} | \mathcal{F}_t\} = \beta_{pt}^e \, \varepsilon\{dR_{et} | \mathcal{F}_t\} + (1-\beta_{pt}^e)\, \varepsilon\{dR_{zt} | \mathcal{F}_t\} \tag{5.36}$$

The β factors in the Eqs. (5.34) through (5.36) are the volatilities of the portfolios with the respect to the efficient portfolio.

Finally, let us consider the case where the minimum-risk portfolio has zero quadratic variation (conditionally risk-free portfolio). This will happen if one can lend or borrow money with no limitation at an interest rate that is adapted to the information at time t. Clearly, the risk-free asset will have zero correlation with any of the efficient portfolio. Thus the efficient frontier is a straight line generated by one point representing an efficient portfolio, duely levered by the risk-free asset. Equation (5.34), therefore will become

$$dR_t = \beta_t^m \, dR_{mt} + (1-\beta_t^m) \, dR_{ft} \tag{5.37}$$

where dR_{mt} is the rate of return for the (efficient) market portfolio. Also we could rewrite (5.36) as

$$\varepsilon\{dR_{pt}|\mathcal{F}_t\} - dR_{ft} = \frac{d\langle R_p,R_m\rangle_t}{d\langle R_m,R_n\rangle_t} [\varepsilon\{dR_{mt}|\mathcal{F}_t\} - dR_{ft}] \tag{5.38}$$

Equations (5.37) and (5.38) give the familiar market models described by the following equations

$$dR_{pt} = \beta_{pt}^e dR_{et} + (1-\beta_{pt}^e) \, dR_{zt} + d\mu_{pt} \tag{5.39}$$

$$dR_{it} = dR_{ft} = \frac{d\langle R_i,R_m\rangle_t}{d\langle R_m,R_m\rangle_t} [dR_{mt} - dR_{ft}] + d\nu_{it} \tag{5.40}$$

Where μ_{pt} and ν_{it} are $\mathcal{F}_t$ martingales "uncorrelated" with the efficient portfolio in the following sense:

$$d\langle \mu_p,R_e\rangle_t = 0 \tag{5.41}$$

$$d\langle \nu_i,R_m\rangle_t = 0 \tag{5.42}$$

The implied market model however, does not have the "independent noise" or "uncorrelated noise" property since for the optimal portfolio selection θ^* we have [15]:

$$\sum_{i=1}^{N} \theta_{it}^{*} \, d\nu_{it} = 0. \qquad (5.43)$$

Chapter 6

Summary of Conclusion and Discussion

6.1. Summary

The following have been concluded here

1. If the rate of return process dR_{it} is an $\mathcal{F}_t$ local semi-martingale where $\mathcal{F}_t$ is an increasing family of σ-algebras representing the information at time t, then its relative volatity with respect to another local-semi-martingale process X_t is given by

$$\beta_{it} = \frac{d\langle R_i,X\rangle_t}{d\langle X,X\rangle_t} \tag{6.1}$$

β_{it} gives the fluctuations of the rate of return process due to that of X_t in the sense that it minimizes the quadratic variation of the "error" process ε_t given by

$$d\varepsilon_t = dR_{it} - \beta_{it}dX_t. \tag{6.2}$$

2. If the "true price" of a futures contract is the conditional expectation of its discounted present value

$$P_t^T = \varepsilon\{e^{-r(T-t)}P_T|\mathcal{F}_t\} \tag{6.3}$$

then its rate of return in addition to the interest rate r is a (local) martingale process.

$$\frac{dP_t^T}{P_t^T} = rdt + d\xi_t \tag{6.4}$$

$$\varepsilon\{d\xi_t|\mathcal{F}_t\} = 0. \tag{6.5}$$

3. If the dividend payments for a share is a local $\mathcal{F}_t$ semi-martingale and if its "true price" is the conditional expectation of the sum of discounted future dividend payments given by

$$P_t = \varepsilon\{\int_t^\infty e^{-r(s-t)} dD_s | \mathcal{F}_t\} \qquad (6.6)$$

then P_t is another local semi-martingale with respect to $\mathcal{F}_t$ and we have

$$\frac{dP_t + dD_t}{P_t} - rdt = d\xi_t \qquad (6.7)$$

where ξ_t is again a local $\mathcal{F}_t$ martingale

$$\varepsilon\{d\xi_t | \mathcal{F}_t\} = 0. \qquad (6.8)$$

4. If $\mathcal{F}_t$ is generated by a time-homogeneous Markov process X_t with a given differential generator $\mathcal{L}$, and if the dividend policy is characterized by the function f(·) given by

$$f(X_t) = \varepsilon\{dD_t | X_t\} \qquad (6.9)$$

then the "true price" will be characterized by another function F(·), which is obtained by operating the resolvent operator of the Markov process $\mathcal{R}_r$ on the dividend policy function

$$P_t = F(X_t) = \mathcal{R}_r f(X_t) \qquad (6.10)$$

5. The true price function F is the solution to the following equation

$$\mathcal{L}F(x) - rF(x) = -f(x) \qquad (6.11)$$

6. Viewed in terms of the state space, the "true price" function is the potential corresponding to the dividend policy $f(\cdot)$

$$F(x) = \int G(x,dy).f(y) \tag{6.12}$$

The Green's kernel G is the Laplace transform of the probability transition function for the process X_t

$$G(x,\Gamma) = \int_0^\infty e^{-rt} P_t(x,\Gamma) \tag{6.13}$$

where

$$P_t(x,\Gamma) = P[X_t \in \Gamma | X_0 = x] \tag{6.14}$$

7. One can use the resolvent equation

$$\mathcal{R}_\alpha - \mathcal{R}_\beta + (\alpha-\beta)\mathcal{R}_\alpha \mathcal{R}_\beta = 0 \tag{6.15}$$

to explicitly get the rate of change of the "true price" due to a change in the discount rate r.

$$\begin{aligned} \frac{d}{dr} F_r(x) &= -\mathcal{R}_r F_r(x) \\ &= -\int F_r(y)\; G_r(x,dy) \end{aligned} \tag{6.16}$$

Considering the limited liability of stock prices, one can conclude that "true prices" are monotonously decreasing functions of the discount rate.

8. Similar results are obtained when the discount rate is a stochastic process adapted to $\mathcal{F}_t$. For example, if the instantaneous rate of discounting is a given function $r(\cdot)$ so that the "true price" is

$$F(x) = \varepsilon\{\int_0^\infty \exp(-\int_0^t r(X_s)ds)\; f(X_t)dt | X_0 = x\} \tag{6.17}$$

then using Kac's theorem the "true price" function will be the solution to

$$\mathcal{L}F(x) - r(x)F(x) = -f(x) \tag{6.18}$$

9. If instead of the Markov process X_t one considers a subprocess defined only on [0,T], where the conditional probability of termination is given by the function $\xi(\cdot)$ then the differential generator corresponding to the "killed" subprocess $\tilde{\mathcal{L}}$ will be given by

$$\tilde{\mathcal{L}}F(x) = \mathcal{L}F(x) - \xi(x)\, F(x) \tag{6.19}$$

Therefore the termination probability is not distinguished from the discount rate when one considers the "true price" equation (6.18).

10. When there are more than one investment opportunities, in order to exclude the possibility of arbitrage one should be able to express the rate of return on each investment in terms of N "orthogonal" processes S_{kt} and the risk-free rate of return as:

$$dR_t = \sum_{k=1}^{N} \beta_{kt} dS_{kt} + dR_{0t} \tag{6.20}$$

where

$$\beta_{kt} = \frac{d\langle R, S_k \rangle_t}{d\langle S_k, S_k \rangle_t} \tag{6.21}$$

N is the dimension of the information process X_t and dR_{0t}, the rate of return on the risk-free portfolio is adapted to $\mathcal{F}_t$ and therefore predictable at each time.

11. One application of the result stated in 10 is to find the "true price" of various options issued on a stock. If the stock price is assumed to be a terminated Markov diffusion process X_t whose

differential generator is given by

$$= \frac{1}{2} a(x) \frac{d^2}{dx^2} + b(x) \frac{d}{dx} - c(x) \quad (6.22)$$

then the "true price price" of the option F(·) that would exclude the possibility of arbitrage is a solution to the following differential equation

$$\frac{1}{2} a^2(x) \frac{\partial^2 F}{\partial x^2} + r(x) \frac{\partial F}{\partial x} - [r(x)+c(x)]F + \frac{\partial F}{\partial t} = 0.$$

where $r(x)$ is the risk-free rate of return.

12. If ones wealth is given in terms of ones investment return dR_t as

$$\frac{dW_t}{W_t} = dR_t \quad (6.24)$$

and if one acts to maximize a concave, increasing instantaneous utility of ones present wealth then at each time t cone should choose among all investments with the same expected rate of return $\varepsilon\{dR_t|\mathcal{F}_t\}$ the one with the least quadratic variation process $d\langle R,R\rangle_t$.

13. If one can choose a portfolio of N investment opportunities each characterized by its rate of return dR_{it} then at each time ones portfolio is given by the following quadratic programming problem [45]:

$$\text{minimize:} \quad \sum_{i=1}^{N} \sum_{j=1}^{N} \theta_{it}\theta_{jt} \, d\langle R_i,R_j\rangle_t \quad (6.25)$$

$$\text{subject to:} \quad \sum_{i=1}^{N} \theta_{it} = 1 \quad (6.26)$$

$$\sum_{i=1}^{N} \theta_{it}\varepsilon\{dR_{it}|\mathcal{F}_t\} = S_t \, dt \quad (6.27)$$

where θ_{it} is the ratio invested in the i-th opportunity.

This will give an efficient portfolio for every S_t.

14. If, in addition, one's ratio of consumption is denoted as c_t, and $\varepsilon\{dR_{it}|\mathcal{F}_t\}$, $d\langle R_i, R_j\rangle_t$ are generated by a Markov process X_t whose evolution is not dependent on the investment or consumption strategies, then the wealth equation will be given by

$$\frac{dW_t}{W_t} = \sum_{i=1}^{N} \theta_{it}\, dR_{it} - c_t\, dt \tag{6.28}$$

One can find a minimum-risk protfolio with rate of return dR_{0t} so that one's portfolio is a combination of the first N-1 opportunities for investment and the minimum-risk portfolio. Let θ_t be an N-1 dimensional vector representing this combination and suppose that one wants to maximize the expected value of a concave increasing utility for consumption. Then, the optimal portfolio is given by

$$\sum_{j=1}^{N-1} \theta_{it}\, d\langle R_i', R_j'\rangle_t = A_{rt}\varepsilon\{dR_{it}'|X_t\} + \sum_{k=1}^{m} A_{kt}\, d\langle R_i', X_k\rangle_t \tag{6.29}$$

where m is the dimension of X_t $dR_{ir}' = dR_{it} - dR_{0t}$ and $A_{rt} \geqq 0\ \forall t$.

6.2. Discussion

The volatility of the rate of return on an investment as defined in 1 is different from the conventional definition in that β is now time-dependent. Moreover, the future values of β are random. At each time however, the volatility is known given the information $_t$.

If the increasing family of σ-algebras $\mathcal{F}_t$ representing the information is changed or if there is an absolutely continuous change in the probability measure the value of β will remain unchanged. Therefore, the volatility is a property of the investment opportunity and not of one's information and/or subjective probability measures.

The results pertaining to the random walk theory of true prices is equivalent to the statement that if the market prices are the conditional expectation of the "true" values then the incoming information are being used in the estimation of the "true prices" and therefore, should have no value in predicting the rate of return in excess of the discount rate. Similar results are obtained in the discrete time framework by Granger [19], and Samuelson [38,39]. The model presented here is general and has no restriction on the dividend process or the information fields other than requiring D_t to be a local semi-martingale, so that the "true price" integral is well defined.

When $\mathcal{F}_t$ is generated by observing the present value of a state vector X_t Eq. (6.11) explicitely gives the "true price" function. The potential interpretation given by (6.12) reflect the fact that shares have value because if their potential dividend payments. The kernel is the Laplace transform of the probability transition function and is dependent on the discount rate but not the dividend policy.

The resolvent equation has the important implication that "true prices" are monotonic functions of the discount rate, and Eq. (6,15) yields explicitly the changes in "true prices" due to changes in the discount rate.

A very interesting result is the equivalence of a termination in the information process X_t with the rate of discounting, suggesting that one discounts the future for the possibility of a catastrophic termination of the system.

If the possibility for making arbitrage profit is to be excluded then the rate of return on each share should be given by the "market equation (6.20). The application on "rational option prices" has been known [37] for log-normal processes. Again a surprising result is that with the assumptions made in 4, option prices will not depend on the expected rate of return b(x) of the underlying security prices, which is hard to identity because it depends on $\mathcal{F}_t$ as well as on the probability measure. The initial conditions will depend on how the option has been written.

The results on the portfolio selection problem also have important implications. The generalized mean-variance analysis gives the proper dynamical generalization of the one-period case under the assumptions of normal returns or quadratic utility functions [16]. If all the participants in the market act accordingly, then the expected rate of return on each investment would be linearly related to its relative volatility with the market rate of return. The interpretation is that one is compensated for taking the risk and investing in the market to the extent that he is willing to undertake the market risk.

Since volatility of investments does not depend on the information or on absolutely continuous changes in probability, this gives a feedback rule that determines the expected rate of return on each portfolio in terms of their volatilities and the minimum-risk portfolio.

When one considers utilities for consumption, a more general result due to Merton [18] is outlined in 5.4. The optimal portfolio now is a combination of the generalized mean-variance efficient portfolio and m portfolio that are used to hedge against unfavorable changes in the future investment opportunities.

The surprising new result is that the m insurance portfolios are "objective" in the sense that they, again, do not depend on one's information or probability beliefs. Thus as long as everybody agrees on the "impossibility" of events they will also agree on the combination of the m "insurance" portfolios.

The result remains true if one drops the additivity assumption on the utility functions. Also the optimal portfolio does not change if one anticipates a random termination of the process X_t.

It would be interesting to extend the model when the sample-continuity assumption is dropped. Recent results [48] indicate that one can use the martingale calculus to generalize the above results to cases where the rate of return processes may have discontinuities. However, such a generalization has not been attempted here.

References

1. K. J. Arrow, Essays in the Theory of Risk Bearing, Chicago: Markham and London: North Holland, 1970.

2. K. J. Aström and B. Wittenmark, "On Self-tuning Regulators," Automatica, 9 (1973), 185-199.

3. L. Bachelier, "Theory of Speculation," in The Random Character of Stock Market, P. Cootner ed., The M.I.T. Press, 1964.

4. J. M. Bismut, "Growth and Optimal Intertemporal Allocation of Risks," Journal of Economic Theory, 10 (1975), 239-257.

5. F. Black, M. C. Jensen and M. Scholes, "The Capital Asset Pricing Model: Some Empircal Tests," in Studies in the Theory of Capital Markets, M. C. Jensen, ed., New York: Praeger Publishers, 1972.

6. F. Black and M. Scholes, "The Pricing of Options and Corporate Liabilities," Journal of Political Economy, (1973), 637-654.

7. F. Black, "Capital Market Equilibrium with No Riskless Borrowing or Lending," Journal of Business, 45 (1972), 444-456.

8. R. M. Blumental and R. G. Getoor, Markov Processes and Potential Theory, Academic Press, 1968.

9. P. Cootner, The Random Character of Stock Market, Cambridge, Massachusetts: The M.I.T. Press, 1964.

10. M. H. A. Davis and P. Varaiya, "Dynamic Programming Conditions for Partially Observable Stochastic Systems," SIAM Journal of Control, 11 (1970), 226-261.

11. G. Debreu, Theory of Value, New York: John Wiley and Sons, Inc., 1959.

12. J. L. Doob, Stochastic Processes, New York: John Wiley and Sons, 1953.

13. T. E. Dunkan and P. Varaiya, "On the Solutions of a Stochastic Control System," SIAM Journal of Control, 9 (1971), 354-371.

14. E. B. Dynkin, Theory of Markov Processes, D. E. Brown, Trans., London: Pergamon Press, and Englewood Cliffs, N. J.: Prentice Hall, 1961.

15. E. F. Fama, "Efficient Capital Markets: A Review of Theory and Empiracal Work," Journal of Finance, 25 (1970), 383-417.

16. E. F. Fama, "Risk Return and Equilibrium," Journal of Political Economy, 79 (1971), 30-55.

17. E. F. Fama and M. H. Miller, The Theory of Finance, New York: Holt Rinehart and Winston, 1972.

18. I. Friend and M. Blume, "Measurement of Portfolio Performance Under Uncertainty," The American Economic Review, 60 (1970), 561-575.

19. C. W. J. Granger, "Some Implications of the Fundamentalist's Valuation Model," Presented at the Berlin Workshop on International Capital Markets, September 1973.

20. C. W. J. Granger, "Empirical Studies of Capital Markets: A Survey," in Mathematical Methods in Investment and Finance, C. P. Szegö and K. Shell ed., North Holland, 1972.

21. J. Hadar, "Stochastic Dominance and Diversification," Journal of Economic Theory, 3 (1971), 288-305.

22. N. H. Hakansson, "Capital Growth and the Mean-Variance Approach to Portfolio Selection," Journal of Finance and Quantitative Analysis, 6 (1971), 517-557.

23. M. C. Jensen, "Capital Markets: Theory and Evidence," Bell Journal of Economics and Management Science, 3 (1972), 357-398.

24. H. Kunita and S. Watanabe, "On Square Integrable Martingales," Nagoya Math. J., 30 (1976), 209-245.

25. H. Kushner, Introduction to Stochastic Control, New York: Holt, Rinehart and Winston, 1971.

26 C. Lucas and E. Wong, "The Fundamentalist's View of Price Fluctuations -- An Application of Martingale Theory," Presented at the Stochastic Systems Symposium, University of Kentucky, Lexington, Kentucky, July 1975.

27. P. A. Mayer, Probability and Potentials, Blaisdell Publishing Co., 1966.

28. R. C. Merton, "An Intertemporal Capital Asset Pricing Model," Econometrica, 41 (1973), 868-887.

29. R. C. Merton, "An Analytic Derivation of the Efficient Portfolio Frontier," Journal of Financial and Quantitative Analysis, 7 (1972), 1851-1872.

30. R. C. Merton, "Theory of Rational Option Pricing," Bell Journal of Economics and Management Science, (1973), 141-183.

31. R. C. Merton, "Optimum Consumption and Portfolio Rules in a Continuous Time Model," Journal of Economic Theory, 3 (1971), 373-413.

32. R. C. Merton and P. A. Samuelson, "Fallacy of the Asymptatic Log-Normal Approximation in Portfolio Decision Making Over Many Periods," Working Paper 623-72 Sloan School of Management, Massachusetts Institute of Technology, 1972.

33. E. H. Neave, "Multiperiod Consumption-Investment Decision and Risk Preference," Journal of Economic Theory, 3 (1971), 40-53.

34. R. Radner, "Optimal Stationary Consumption with Stochastic Production and Resources," Journal of Economic Theory, 6 (1973), 68-90.

35. M. Rothschild and J. E. Stiglitz, "Increasing Risk: I. A Definition," Journal of Economic Theory, 9 (1970), 225-243.

36. M. Rothschild and J. E. Stiglitz, "Increasing Risk II: Its Economic Consequences," Journal of Economic Theory, 3 (1971), 66-84.

37. P. A. Samuelson, "Mathematics of Speculative Price," SIAM Review, 15 (1973), 1-42.

38. P. A. Samuelson, "Proof that Properly Anticipated Prices Fluctuate Randomly," Industrial Management Review, 6 (1965), 41-50.

39. P. A. Samuelson, "Proof that Properly Discounted Present Values of Assets Vibrate Randomly," Bell Journal of Economics and Management Science, (1974), 369-374.

40. W. F. Sharpe, Portfolio Theory and Capital Markets, New York: McGraw-Hill Book Co., 1970.

41. G. F. Simmons, Introduction to Topology and Modern Analysis, New York: McGraw-Hill Book Co., 1963.

42. G. P. Szegö and K. Shell, Mathematical Methods in Investment and Finance, North Holland, 1972.

43. J. Pratt, "Risk Aversion in the Small and in the Large," Econometrica, 32 (1964), 122-136.

44. J. H. Van Schuppen, "Estimation Theory for Continuous Time Processes, a Martingale Approach," Ph.D. Dissertation, Dept. E.E.C.S., University of California, Berkeley, September 1973.

45. P. P. Varaiya, Notes on Optimization, Van Nostrand Reinhold, 1972.

46. E. Wong, "Martinagle Theory and Applications to Stochastic Problems in Dynamical Systems," Publication 72119, Department of Computing and Control, Imperial College of Science and Technology, 1972.

47. E. Wong, "Representation of Martingales, Quadratic Variation and Applications," SIAM Journal of Control, 9 (1971), 621-633.

48. E. Wong, "Recent Progress in Stochastic Processes -- A Survey," IEEE Transactions on Information Theory, 19 (1973), 262-275.

49. E. Wong, Stochastic Processes in Information and Dynamical Systems, McGraw-Hill Inc., 1971.

50. W. M. Wonham, "Random Differential Equations in Control Theory," in Probabilistic Methods in Applied Mathematics, Vol. 2, A. T. Bharucha-Reid, ed., New York: Academic Press, 1970.